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. Some articles will have been accepted based in part or entirely on reviews undertaken for other BMJ Group journals. These will be reproduced where possible.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Are job Strain and Sleep Disturbances Prognostic Factors for Neck/Shoulder/Arm pain? A Cohort Study of a General Population of Working Age in Sweden
AUTHORS	Rasmussen-Barr, Eva; Grooten, Wilhelmus; Hallqvist, Johan; Holm, Lena; Skillgate, Eva

VERSION 1 - REVIEW

REVIEWER	Victor CW Hoe Centre for Occupational and Environmental Health-UM c/o Department of Social and Preventive Medicine Faculty of Medicine, University of Malaya
REVIEW RETURNED	22-Apr-2014

GENERAL COMMENTS	Thank you for the opportunity to review your report on "Are job Strain and Sleep Disturbances on Prognostic Factors for Neck/Shoulder/Arm pain? A Cohort Study of a General Population of Working Age in Sweden". The report assessed the association between job strain using the Karasek Job-Content Model and sleep disturbance at baseline and the chances of developing neck/shoulder/arm pain (NSAP) after four years. 1. It is an interesting concept to look at sleep disturbance as an effect measure modifier for job strain as a prognostic factor for NSAP. However, I am confuse how sleep can be considered an effect measure modifier for job strain. Job strain according to the Karasek Job-Content Model does not measure stress as a consequence of work but classify different type of job into four category of work strain. Workers in the high job strain category is more likely to develop sleep disturbance as compared to the other three groups of workers. So, sleep disturbance is a consequence of job strain and cannot contribute to job strain. Further in your stratification analysis the results does not shows differences in the two strata; both the strata (no sleep disturbance and with sleep disturbance) showed an increase odds of NSAP, and there was not a major change in the odds ratios of job strain and active job pre- and post-stratification. I feel that sleep disturbance should be considered as a risk factors for NSAP and not really an effect modifier of job strain. 2. I am also curious the reason for choosing logistic regression as the choice for conducting your analysis. It has been pointed out in your report that reporting of odds is different from risk. In your cohort study I would expect the estimate measure to be reported as risk and not odds. It will be good if you provide an explanation on the choice of choosing your current analysis method.

3. Job strain and sleep disturbance as assessed in your study may be a transient effect and it may not be the same over the four years period. Is it possible to measure change in job strain and sleep disturbance instead of only measuring job strain and sleep disturbance only at baseline?

4. In the Methods section on potential confounders, you have described many variables that are not presented or are different from those in Table 1. For example the age as described in the methods included continuous and categorised into five categories, but in Table 1 only 2 categories was presented. It would be appropriate to only describe the variables included in your current report.

5. I feel that before presenting multivariate analysis, it would be appropriate to also present the univariate analysis. As many variables in Table 1 were not included in the multivariate analysis (Table 2). In table 3, please also indicate the variables that have been included in the multivariate analysis.

6. In your discussion there should be more in-depth discussion on the study. There are many other non-work and work factor for NSAP that was not studied in the current study which may contribute to the reporting of NSAP; e.g., somatization tendency and other psychosocial risk factor, and workplace physical risk factor.

REVIEWER	Alice Kongsted University of Southern Denmark and Nordic Institute of Chiropractic and Clinical Biomechanics Denmark
REVIEW RETURNED	22-Apr-2014

| GENERAL COMMENTS | The study is a population based cohort study investigating if job strain and sleep disturbances are associated with experiencing troublesome neck/shoulder/arm pain four years after baseline. An additional objective was to investigate if the effect of job strain was modified by sleep disturbances.

The study concerns an important health issue with large negative impact on global health, it is based on a large and comprehensive dataset, and the methods and the results are clearly presented.

My main concerns are: (1) that the outcome is described as reflecting "the development of troublesome pain" whereas the outcome measured is actually the presence of an episode of troublesome pain which was not necessarily different from the occasional pain reported at baseline, (2) that stratified analyses were used to test for effect modification, and (3) that stronger conclusions are made than I find supported by the results.

The language is absolutely readable but the manuscript would benefit from proof reading by a native English speaking person. |
| --- | --- |

Introduction

The background introduces the theme of the paper adequately. However, I would like to point out some issues that I find imprecise or somehow confusing:

1. After underlining the importance of studying modifiable factors you mention a study on economic stress and family income. I don't think of these as obviously modifiable and suggest that you consider if the flow of the background should be slightly altered.
2. It is stated that the above mentioned factors were associated with "a poor prognosis (i.e., development of troublesome neck pain)". I think it would be more correct to describe this as an association with *increased risk* of troublesome neck pain in order to distinguish clearly between prognosis and risk.
3. In some instances I believe you use the term psychosocial when actually referring to psychological factors. I would prefer distinguishing between social and psychological factors when possible. Please, have a look at that throughout the manuscript.
4. You describe the job-strain model as an "important" model but it is not clear to me why it is important. Is it widely used or has it been demonstrated somehow useful?
5. Is the job-strain model generally known as such or is it more often referred to as the "job demand-control" model? If these are the same I think it would be helpful stating that the model is also known as the "job demand-control" model.
6. After defining active versus passive job situation you describe the effects of job strain. Were these associations actually between passive job situations and prognosis/risk of NSAP?
7. I think you meant to write "affect" rather than "modify" in the sentence "Several factors most likely modify the trajectory from occasional to troublesome NSAP". Or "…modify the association between job strain and the trajectory …"
8. I think the expression "effect measure modifier" is hard to read. You could consider something like: "… explore whether an association between job strain and the prognosis of NSAP is modified by sleep disturbances"

METHODS
Study population
9. How were participants who received a questionnaire in 2006 selected?

Questionnaires
10. I think it would appear logical to name the job strain categories (i) and (iiii)[iv] "low strain" and "high strain"
11. Could all participants be allocated to the four categories? I wonder if no one answered yes to both of the questions a and b; or no to both? And no one answered yes to one of c and d and no to the other?

Outcome

12. I think there is a mismatch between the outcome measure and the intention of investigating factors associated with "the development of troublesome NSAP". People considered to have developed troublesome NSAP could have had just one short episode of NSAP within a 6 month period which they had recovered from when responding to the follow-up questionnaire and this may not be different from having had "a couple of days [with pain] in the last six months" at baseline. Please, reconsider if it is correct to describe these subjects as having developed a NSAP condition. I guess it is more correct to state that they have experienced an episode of troublesome NSAP within the 6 months preceding follow-up.

Statistical analyses

13. It appears that some of the confounders may be highly correlated to job strain. Did you test for collinearity?
14. Were analyses of the effect of jobs strain controlled by sleep disturbances? It appears to me that they were not but ought to be.
15. Why did you choose a stratified analysis rather than testing for interactions between job strain and sleep disturbances? Using a test for interactions would tell whether the stratas were significantly different which is not possible with a stratified analysis.
16. I think the information regarding ethical approval belongs more naturally to the beginning of the methods section rather than to the statistical analyses.

Results

17. Would it be possible to describe the source population in order to now if there were potential selection bias?

Discussion/Conclusion

18. I think it is too strong to state that the studied factors are important risk factors given that the effect sizes are small and these factors would not assist importantly in predicting individuals outcome (would be my best guess without results regarding predictive accuracy being available).
19. Again, I question whether you actually study factors related to the development of NSAP since it is not clear to me that the 'cases' had a worse condition at follow-up than they had at baseline.
20. I would also be hesitant to conclude that sleep disturbances modified the effect of job strain. Based on the confidence intervals I doubt that the change in OR was not statistically significant and even if so I think it was only potentially relevant in size in those with job strain.

	21. A general comment is that you are not consisting in describing your study as either a prognostic or a risk study.

VERSION 1 – AUTHOR RESPONSE

Reviewer Name Victor CW Hoe

1. It is an interesting concept to look at sleep disturbance as an effect measure modifier for job strain as a prognostic factor for NSAP. However, I am confuse how sleep can be considered an effect measure modifier for job strain. Job strain according to the Karasek Job-Content Model does not measure stress as a consequence of work but classify different type of job into four category of work strain. Workers in the high job strain category is more likely to develop sleep disturbance as compared to the other three groups of workers. So, sleep disturbance is a consequence of job strain and cannot contribute to job strain. Further in your stratification analysis the results does not shows differences in the two strata; both the strata (no sleep disturbance and with sleep disturbance) showed an increase odds of NSAP, and there was not a major change in the odds ratios of job strain and active job pre- and post-stratification. I feel that sleep disturbance should be considered as a risk factors for NSAP and not really an effect modifier of job strain.
Response
We will try to clarify the hypothesized mechanism in the associations investigated in this study. We don't consider that sleep might be an effect measure modifier for job strain. We consider instead that sleep might be an effect measure modifier for experience new episodes of troublesome neck/shoulder/arm pain (NSAP) in those with job strain or high strain at baseline.
We completely agree with you that it seems that sleep disturbances is a prognostic factor for developing episodes of troublesome NSAP and so is active jobs and high strain at base line. Since sleep disturbance probably is in the causal pathway between job strain and troublesome NSAP – we could not perform analyses of the biological (causal) interaction between job strain and sleep disturbance to understand if the effect of job strain on the risk of troublesome NSAP was modified by the persons level of sleep disturbance. That's why we stratified the analyses of job strain into two strata of no sleep disturbances or sleep disturbances. An aim with the study was: "to study whether sleep disturbances act as an effect measure modifier in the association between job strain and troublesome NSAP" – not if sleep acted as an effect modifier on job strain.
We agree that we might have drawn to strong conclusions regarding the effect measure modification since the OR in the stratified analyses in Table 3 do not differ much between persons with and without sleep disturbances. Still there is an effect of active job and high job strain on persons with sleep disturbance but not in persons without sleep disturbances. We have now reworded our results/conclusion to stating that our results indicate that high strain and sleep disturbances are prognostic factors and that we suggest that sleep disturbances may be an effect measure modifier in the association between high job strain and troublesome NSAP.

2. I am also curious the reason for choosing logistic regression as the choice for conducting your analysis. It has been pointed out in your report that reporting of odds is different from risk. In your cohort study I would expect the estimate measure to be reported as risk and not odds. It will be good if you provide an explanation on the choice of choosing your current analysis method.
Response
Thank you for your question. We have published several studies on risk and prognosis based on the Stockholm Public Health Study, using the logistic regression model for the analyses. Theoretically we could consider using the Cox proportional Hazard model to calculate the HRR. In that case all study participants gets the same person time since we have only one follow-up occasion. In earlier analyses

we found that the HRR get almost identical as the OR from the logistic regression. Another option, especially if the outcome is common, is to use the Log- binomial regression. We have also tested that in our earlier studies (Bohman et al BMC Public Health 2013) as well as in this study and found that the results were very similar as if analyzed with the logistic regression model. We don't consider choosing the Logistic regression model to be a serious limitation since the OR all are relatively low (< 2). We choose to use the OR since the results are more easily compared to these and other studies as Canivet et al (2008)(1) who also present results with calculated OR. Page 15, line13

3. Job strain and sleep disturbance as assessed in your study may be a transient effect and it may not be the same over the four years period. Is it possible to measure change in job strain and sleep disturbance instead of only measuring job strain and sleep disturbance only at baseline?
Response
Thank you for your wise remark. We absolutely get your point and that would be of great interested of course. This was not however possible to do in this study. We have discussed that as a potential limitation in the Discussion part of the manuscript (page 15, line 10)

4. In the Methods section on potential confounders, you have described many variables that are not presented or are different from those in Table 1. For example the age as described in the methods included continuous and categorized into five categories, but in Table 1 only 2 categories was presented. It would be appropriate to only describe the variables included in your current report.
Response
Thank you for your remark. All variables presented in Table 1 are also presented in the methodological section but might be in not the same wording and categorization. We have reviewed the table and the text in the methodological session to make sure that there is a consistency.

5. I feel that before presenting multivariate analysis, it would be appropriate to also present the univariate analysis. As many variables in Table 1 were not included in the multivariate analysis (Table 2). In table 3, please also indicate the variables that have been included in the multivariate analysis.
Response
We don't consider that presenting a univariate analysis of all potential confounding factors is relevant for the aim of this study. We tested all confounding factors one by one to the crude regression model to study if they changed the crude OR by 10% or more (2). For those that changed the OR we included those in the adjusted analyses. For table 3 we will include the variables of course. Thank you for pointing this out. If the confounding analysis would be of interested to the readers they may be presented as an appendix of course.

6. In your discussion there should be more in-depth discussion on the study. There are many other non-work and work factor for NSAP that was not studied in the current study which may contribute to the reporting of NSAP; e.g., somatization tendency and other psychosocial risk factor, and workplace physical risk factor.
Response
Thank you for pointing this out to us. Of course we are aware of that there are other factors that may contribute to troublesome episodes of NSAP. They are of importance for our analyses only if they are confounders and associated with the variables of interest, i.e. related to both exposure and outcomes. We have added some information about this to the Discussion section as unmeasured confounders. (Page 14, line 9)

Reviewer Name Alice Kongsted

The study is a population based cohort study investigating if job strain and sleep disturbances are associated with experiencing troublesome neck/shoulder/arm pain four years after baseline. An

additional objective was to investigate if the effect of job strain was modified by sleep disturbances. The study concerns an important health issue with large negative impact on global health, it is based on a large and comprehensive dataset, and the methods and the results are clearly presented.

My main concerns are: (1) that the outcome is described as reflecting "the development of troublesome pain" whereas the outcome measured is actually the presence of an episode of troublesome pain which was not necessarily different from the occasional pain reported at baseline, (2) that stratified analyses were used to test for effect modification, and (3) that stronger conclusions are made than I find supported by the results.

Response

1) We understand your concern, however the included subjects reported pain at baseline that was only reported as an occasional NSAP: "Occasional pain was indicated if participants responded to the question "During the previous six months, have you experienced pain in neck, shoulder and/or arms?" with either "Yes, a couple of days in the last six months" or "Yes, a couple of days each month."

For the follow-up the pain was reported such as: "During the last six months, have you felt pain in your neck or upper back and/or shoulder or arms? If so, have these restricted your work capacity or hindered you in daily activities to some degree or to a high degree?"

Nevertheless we cannot be sure that some of the participants with occasional pain at baseline might also have considered the pain to be troublesome. Therefor we have reworded the outcome to be experienced at least one episode of troublesome NSAP during the past six months as suggested further down and we have changed the manuscript accordingly.

2) Since sleep disturbance may be in the causal pathway between job strain and troublesome NSAP – we did not perform analyses of the biological interaction between job strain and sleep disturbances. To understand the effect of high job strain on the risk of troublesome NSAP we instead stratified the analyses of job strain into two strata of no sleep disturbances or sleep disturbances. In addition we entered an interaction term (job strain x sleep disturbances) to the un-stratified model and compared the models with and without this interaction term. The model with the interaction term differed statistically significant in fit from the model without the interaction term, which indicates a modifying effect.

3) We agree that we might have drawn to strong conclusions of our results and we have now reworded our results/conclusion to stating that our results indicate that high strain and sleep disturbances are prognostic factors and that we suggest that sleep disturbances may be an effect measure modifier in the association between high strain and troublesome NSAP.

The language is absolutely readable but the manuscript would benefit from proof reading by a native English speaking person.

Response

The manuscript has been revised by a certified language revision but obviously not good enough. Thank you for pointing this out! It has now been re-revised by an English native person.

Introduction

The background introduces the theme of the paper adequately. However, I would like to point out some issues that I find imprecise or somehow confusing:

1. After underlining the importance of studying modifiable factors you mention a study on economic stress and family income. I don't think of these as obviously modifiable and suggest that you consider if the flow of the background should be slightly altered.

Response

Thank you for this comment. We have changed the flow in the introductions part.

2. It is stated that the above mentioned factors were associated with "a poor prognosis (i.e., development of troublesome neck pain)". I think it would be more correct to describe this as an association with increased risk of troublesome neck pain in order to distinguish clearly between

prognosis and risk.

Response

We have changed that part following your recommendation. (Page 4, line 16)

3. In some instances I believe you use the term psychosocial when actually referring to psychological factors. I would prefer distinguishing between social and psychological factors when possible. Please, have a look at that throughout the manuscript.

Response

Thank you for pointing this out. We have checked the manuscript thoroughly. The job strain model might be seen as associated with both psychological and social work stress as there are interactions between both psychological factors a and social factors as the social surroundings at work or at home. A job situation may be considered a social phenomena with both psychological as biological consequences; leading to physiological processes as diseases or disorders .(3) The factors psychological demands and decision latitude may be classified as psychosocial factors.

4. You describe the job-strain model as an "important" model but it is not clear to me why it is important. Is it widely used or has it been demonstrated somehow useful?

Response

Thank you for this comment. We changed the passage (Page 4, line 19)

5. Is the job-strain model generally known as such or is it more often referred to as the "job demand-control" model? If these are the same I think it would be helpful stating that the model is also known as the "job demand-control" model.

Please see our response above

6. After defining active versus passive job situation you describe the effects of job strain. Were these associations actually between passive job situations and prognosis/risk of NSAP?

Response

Thank you for pointing this out. This has been changed to high strain throughout the paper including the tables,

7. I think you meant to write "affect" rather than "modify" in the sentence "Several factors most likely modify the trajectory from occasional to troublesome NSAP". Or "…modify the association between job strain and the trajectory …"

Response

We changed the wording and hope that the understanding is better (page 5 , line 9)

8. I think the expression "effect measure modifier" is hard to read. You could consider something like: "… explore whether an association between job strain and the prognosis of NSAP is modified by sleep disturbances"

Response

Thank you for your remark. We have changed the expression in several places in the revised manuscript. Among others as you suggested at page 5, line 21.

METHODS

Study population

9. How were participants who received a questionnaire in 2006 selected?

Response

The text is changed according to your remark (Page 6, line 10)

Questionnaires

10. I think it would appear logical to name the job strain categories (i) and (iiii)[iv] "low strain" and "high strain"

Response

Thank you. This is a wise comment and we have changed that throughout the manuscript,

11. Could all participants be allocated to the four categories? I wonder if no one answered yes to both of the questions a and b; or no to both? And no one answered yes to one of c and d and no to the other?
Response
No a participant could only be part of one group, thus they are mutual exclusive. .
Outcome
12. I think there is a mismatch between the outcome measure and the intention of investigating factors associated with "the development of troublesome NSAP". People considered to have developed troublesome NSAP could have had just one short episode of NSAP within a 6 month period which they had recovered from when responding to the follow-up questionnaire and this may not be different from having had "a couple of days [with pain] in the last six months" at baseline. Please, reconsider if it is correct to describe these subjects as having developed a NSAP condition. I guess it is more correct to state that they have experienced an episode of troublesome NSAP within the 6 months preceding follow-up.
Response
Thank you for this comment. We are aware that we don't know what happened during the years between the base line and the follow up -if the pain was recurrent or chronic. We defined the outcome comprising those who had experienced troublesome NSAIP that affected or hindered their work situation. This is to be considered troublesome pain compared to the occasional pain at baseline. To change this into "experienced at least one episode of troublesome NSAP during the past six months" is a wise remark. We have revised that throughout the manuscript.

Statistical analyses
13. It appears that some of the confounders may be highly correlated to job strain. Did you test for co linearity?
Response
Yes, we have tested for co-linearity and detected that there were none.

14. Were analyses of the effect of jobs strain controlled by sleep disturbances? It appears to me that they were not but ought to be'
Response
Since sleep disturbance may be in the causal pathway between job strain and troublesome NSAP we did not treat sleep disturbance as a potential confounder in the analyses of the associations between job strain and troublesome NSAP. But we tested if job strain was a confounder in the analyses of the association between sleep disturbances and troublesome NSAP, and it didn't change the OR with more than 10% and was accordingly not included in the adjusted model. Information about this has been added to the Method part of the revised manuscript.(Page 10, line 7)

15. Why did you choose a stratified analysis rather than testing for interactions between job strain and sleep disturbances? Using a test for interactions would tell whether the stratas were significantly different which is not possible with a stratified analysis.
Response
Since sleep disturbance may be in the causal pathway between job strain and troublesome NSAP – we did not perform analyses of the biological interaction between job strain and sleep disturbances. To understand the effect of job strain on the risk of troublesome NSAP we instead stratified the analyses of job strain into two strata of no sleep disturbances or sleep disturbances. In addition we entered an interaction term (job strain x sleep disturbances) to the un-stratified model and compared the models with and without this interaction term. The model with the interaction term differed statistically significant in fit from the model without the interaction term, which indicates a modifying effect.

16. I think the information regarding ethical approval belongs more naturally to the beginning of the methods section rather than to the statistical analyses.
Response
Thank you for observing this. It is now moved to study design. Page 6, line4

Results
17. Would it be possible to describe the source population in order to now if there were potential selection bias?
Response
The source population was stratified sample from the Stockholm County. The text is more thorough now. We have also discussed the selection bias in the discussion part page 15, line 20.

Discussion/Conclusion

18. I think it is too strong to state that the studied factors are important risk factors given that the effect sizes are small and these factors would not assist importantly in predicting individuals outcome (would be my best guess without results regarding predictive accuracy being available).
Response
We agree and have changed the conclusion to – that the results of our study indicate-
Our results indicate that high strain, active jobs and sleep disturbances are prognostic factors that should be taken into account when implementing preventive measures to minimize the risk of troublesome NSAP neck/shoulder/arm pain among people of working age. Further, we suggest that sleep disturbances may modify the association between job strain and troublesome neck/shoulder/arm pain.

19. Again, I question whether you actually study factors related to the development of NSAP since it is not clear to me that the 'cases' had a worse condition at follow-up than they had at baseline.
Response
We understand your concern and have changed this - please see our answer in the beginning of the text

20. I would also be hesitant to conclude that sleep disturbances modified the effect of job strain. Based on the confidence intervals I doubt that the change in OR was not statistically significant and even if so I think it was only potentially relevant in size in those with job strain.
Response
We agree that we might have drawn to strong conclusions regarding the effect measure modification since the OR in the stratified analyses in Table 3 do not differ much between persons with and without sleep disturbances. Still there is an effect of active jobs and high strain on persons with sleep disturbance but not in persons without sleep disturbance. We have now reworded our results/conclusion to stating that our results indicate that high strain and sleep disturbances are prognostic factors and that we suggest that sleep disturbances may be an effect measure modifier in the association between job strain and experience episodes of troublesome NSAP.
21. A general comment is that you are not consisting in describing your study as either a prognostic or a risk study.
Response
Thank you for your remark. We consider that we study the prognosis and we have looked through the manuscript to make it more consistent.

1. Canivet C, Ostergren PO, Choi B, Nilsson P, af Sillen U, Moghadassi M, et al. Sleeping problems as a risk factor for subsequent musculoskeletal pain and the role of job strain: results from a one-year follow-up of the Malmo Shoulder Neck Study Cohort. Int J Behav Med. 2008;15(4):254-62.

2. Rothman KJ, Greenland, S. Introduction to stratified analysis. 3rd edn ed. Philadelphia: Wolters Kluwer Health/Lippincott Williams&Williams; 2008.

3. Kivimaki M, Leino-Arjas P, Luukkonen R, Riihimaki H, Vahtera J, Kirjonen J. Work stress and risk of cardiovascular mortality: prospective cohort study of industrial employees. BMJ. 2002 Oct 19;325(7369):857.

VERSION 2 – REVIEW

REVIEWER	Victor CW HOE Centre for Occupational and Environmental Health-UM c/o Department of Social and Preventive Medicine Faculty of Medicine, University of Malaya,
REVIEW RETURNED	02-Jun-2014

GENERAL COMMENTS	The revised report is an improvement to the previous version, however there are still some issue that need to be address.
	1. Could you please explain how could sleep be considered as an effect measure modifier for job strain and troublesome NSAP? Job strain according to the Karasek Job-Content Model does not measure stress or distress as a consequence of work but classify different type of job into four category of work strain. Workers in the high job strain category is more likely to develop sleep disturbance as compared to the other three groups of workers. So, sleep disturbance is a consequence of job strain and cannot contribute to job strain.
	Further in your stratification analysis the results does not shows differences in the two strata; both the strata (no sleep disturbance and with sleep disturbance) showed an increase odds of NSAP, and there was not a major change in the odds ratios of active jobs and passive jobs category for the pre- and post-stratification. Only the high strain job category showed any real differences. (results extracted from Tables 2 and 3)
	Active jobs – not-stratified (total) OR 1.2 (1.0-1.4), no sleep disturbance OR 1.0 (0.8-1.3), sleep disturbance OR 1.2 (1.0-1.6)
	Passive jobs – not-stratified (total) OR 1.3 (1.1-1.6), no sleep disturbance OR 1.3 (1.0-1.7), sleep disturbance OR 1.3 (0.9-1.8)
	High strain job – not-stratified (total) OR 2.0 (1.3-3.0), no sleep disturbance OR 1.4 (0.8-2.6), sleep disturbance OR 2.6 (1.4-4.6)
	2. Job strain and sleep disturbance as assessed in your study are transient effect and it may not be the same over the four-year period. Is it possible to measure change in job strain and sleep disturbance instead of only measuring job strain and sleep disturbance only at baseline? If it was not possible then it should be explained as a limitation in the discussion section.
	3. In the revised manuscript the explanation of the various variables in the methods section has improved, however there are still some of the explanation that are not really clear. For example the age category is not complete as those aged 44 and 45 were not included, <44/>45 years does not include 44 and 45 years, and this

is different from the way you have presented in results and table 1, 18-44 and 45-61. You have described that age will also be presented as continuous in the methods section however there were not found in the results.

4. You should also be consistent in the way the variable categories are explained; for example for alcohol consumption, the categories were presented as – sometime during a period of 12 months/no, however some other variables you have only presented the positive response.

5. You should be consistent in the way a variable is presented or categorised. The sleep disturbance presented in table 1 have been categorised into none/mild and severe, where as in table 2 it was categorised as no sleep disturbance, mild sleep disturbance and severe sleep disturbance, and in table 3 the categories were no sleep disturbance and sleep disturbance.

6. The number of people with NSAP should also be presented in Table 1.

7. Why were the total numbers of participants in Table 1 differed from those in Table 2 and 3? For example those with low strain were 5,358 in Table 1, whereas it was only 4,023 in table 2 and 5,242 in table 3.

8. In any multivariate analysis all potential confounders should be included in the analysis and this is of particular importance for the main demographic variable like age and gender. As describe in table 2 and 3 both age and gender were not included in the multivariate analysis. Further the variables included in the multivariate analysis for job strain and sleep disturbance were different. In the job strain model socio-economic class, work load and support from superior were included, whereas for the sleep disturbance model only economic stress were included. It is difficult for the reader to follow your analysis as you have not presented univariate analysis for your confounder and outcome variable (NSAP).

9. The presentation of p-value in a logistic regression model is not important as the 95%CI will have provided the similar information. When presenting p-value of extremely small value it would be advisable to present as <0.01 or <0.001 (depending on the number of decimals spaces used in your table) and not as <0.00.

10. The number of decimal spaces used in your tables and report should be consistent, in Table 2, you have presented both 0.05 and 0.004.

| GENERAL COMMENTS | The authors have addressed my points and revised or explained their points of views adequately. I have only some very minor issues relating to the revised version:

Abstract
Looking at the abstract again I don't think it is clear from the abstract |

	that "job strain" is about psychological job demands or psychological job stress. I guess some readers would think the study is about physical work load. 1st word : The should be replaced with To Please explain abbreviations when first used (NSAP) Potential confounders Describing the grouped age variable as <44/>45 would imply that you did not include those aged 44 and 45. Would ≤44 / ≥45 be correct?

VERSION 2 – AUTHOR RESPONSE

Reviewer Name Alice Kongsted
Institution and Country University of Southern Denmark
Please state any competing interests or state 'None declared': none declared

The authors have addressed my points and revised or explained their points of views adequately. I have only some very minor issues relating to the revised version:

Abstract
Looking at the abstract again I don't think it is clear from the abstract that "job strain" is about psychological job demands or psychological job stress. I guess some readers would think the study is about physical work load.
1st word : The should be replaced with To
Please explain abbreviations when first used (NSAP)
Response
Thank you for pointing this out. This is now revised

Potential confounders
Describing the grouped age variable as <44/>45 would imply that you did not include those aged 44 and 45. Would ≤44 / ≥45 be correct?

Response
This is now revised to 18-44/45-61

Reviewer Name Victor CW HOE
Institution and Country Centre for Occupational and Environmental Health-UM
c/o Department of Social and Preventive Medicine
Faculty of Medicine,
University of Malaya,
50603 Kuala Lumpur,
MALAYSIA
Please state any competing interests or state 'None declared': None declared

The revised report is an improvement to the previous version, however there are still some issue that need to be address.

1. Could you please explain how could sleep be considered as an effect measure modifier for job strain and troublesome NSAP? Job strain according to the Karasek Job-Content Model does not

measure stress or distress as a consequence of work but classify different type of job into four category of work strain. Workers in the high job strain category is more likely to develop sleep disturbance as compared to the other three groups of workers. So, sleep disturbance is a consequence of job strain and cannot contribute to job strain.

Response

Thank You for discussing this further. However we find that we have already answered this. The term effect modification is applied to situations in which the magnitude of the effect of an exposure of interest (in this case job strain) on an outcome (in this case troublesome NSAP) differs depending on the level of a third variable (in this sleep quality) .The job strain model does not describe type of job per se but more combinations of job control with job decision latitude and is considered to be a model of both psychological as well as social consequences. We were interested to investigate if those with job strain and occasional NSAP at baseline developed episodes of troublesome NSAP at follow-up to a different extent depending on if sleep disturbance is present or not.

(Answer last letter) "We will try to clarify the hypothesized mechanism in the associations investigated in this study. We don't consider that sleep might be an effect measure modifier for job strain. We consider instead that sleep might be an effect measure modifier for developing troublesome neck/shoulder/arm pain (NSAP) in those with job strain or high strain at baseline. We will try to clarify the hypothesized mechanism in the associations investigated in this study. We don't consider that sleep might be an effect measure modifier for job strain. We consider instead that sleep might be an effect measure modifier for developing troublesome neck/shoulder/arm pain (NSAP) in those with job strain or high strain at baseline.

We completely agree with you that it seems that sleep disturbances are a risk factor for developing troublesome NSAP and so is active jobs and high strain at base line. Since sleep disturbance probably is in the causal pathway between job strain and troublesome NSAP – we could not perform analyses of the causal interaction between job strain and sleep disturbances to understand if the effect of job strain on the risk of experience troublesome NSAP was modified by the persons level of sleep disturbance. That's why we stratified the analysis of job strain into two strata of no sleep disturbances or sleep disturbances. The aim with the study was: "to study whether sleep disturbances act as an effect measure modifier in the association between job strain and troublesome NSAP" – not if sleep acted as an effect modifier on job strain.

Further in your stratification analysis the results does not shows differences in the two strata; both the strata (no sleep disturbance and with sleep disturbance) showed an increase odds of NSAP, and there was not a major change in the odds ratios of active jobs and passive jobs category for the pre- and post-stratification. Only the high strain job category showed any real differences. (results extracted from Tables 2 and 3)

Active jobs – not-stratified (total) OR 1.2 (1.0-1.4), no sleep disturbance OR 1.0 (0.8-1.3), sleep disturbance OR 1.2 (1.0-1.6)

Passive jobs – not-stratified (total) OR 1.3 (1.1-1.6), no sleep disturbance OR 1.3 (1.0-1.7), sleep disturbance OR 1.3 (0.9-1.8)

High strain job – not-stratified (total) OR 2.0 (1.3-3.0), no sleep disturbance OR 1.4 (0.8-2.6), sleep disturbance OR 2.6 (1.4-4.6)

Response

Thank you once again pointing this out, but we find that we answered your question in our last letter. This is our answer from the last letter; please read especially the last sentence that answer to your consideration.

(Answer last letter)We agree that we might have drawn to strong conclusions regarding the effect measure modification since the OR in the stratified analyses in Table 3 do not differ much between

persons with and without sleep disturbances. Still there is an effect of active jobs and high strain on persons with sleep disturbances but not in persons without sleep disturbances. We have now reworded our results/conclusion to stating that our results indicate that high strain and sleep disturbances are prognostic factors and that we suggest that sleep disturbances may be an effect measure modifier in the association between high job strain and troublesome NSAP."

2. Job strain and sleep disturbance as assessed in your study are transient effect and it may not be the same over the four-year period. Is it possible to measure change in job strain and sleep disturbance instead of only measuring job strain and sleep disturbance only at baseline? If it was not possible then it should be explained as a limitation in the discussion section.
Response
Thank You for pointing this out again. We already answered this in our last letter, but have now tried to expand the discussion further in the Discussion section, page 15 line 13-19.
"We lack information about the duration of the exposures prior to baseline or about the presence of the exposures during the four-year follow-up period. This may limit the interpretation of the results through a misclassification of exposure. Such a misclassification would most probably be non-differential. Some study participants classified as exposed at baseline might after a while be unexposed, and some study participants classified as unexposed at baseline may after a while be exposed, which might result in a dilution of a true association"

3. In the revised manuscript the explanation of the various variables in the methods section has improved, however there are still some of the explanation that are not really clear. For example the age category is not complete as those aged 44 and 45 were not included, <44/>45 years does not include 44 and 45 years, and this is different from the way you have presented in results and table 1, 18-44 and 45-61. You have described that age will also be presented as continuous in the methods section however there were not found in the results.
Response
Thank you for pointing this out. This has now been revised to 18-44/45-61.

4. You should also be consistent in the way the variable categories are explained; for example for alcohol consumption, the categories were presented as – sometime during a period of 12 months/no, however some other variables you have only presented the positive response.
Response
Thank you for pointing this out. We find that we are consistent in how we present the variables. We have tried to go through all again and made minor corrections.

5. You should be consistent in the way a variable is presented or categorised. The sleep disturbance presented in table 1 have been categorised into none/mild and severe, where as in table 2 it was categorised as no sleep disturbance, mild sleep disturbance and severe sleep disturbance, and in table 3 the categories were no sleep disturbance and sleep disturbance.
Response
Thank you for pointing this out. We have corrected this in table 1 and also clarified it in the text. Page 9, line 4.

6. The number of people with NSAP should also be presented in Table 1.
Response
The number with occasional NSAP at baseline is presented in Table 1 in the Table title - n = 6,979

7. Why were the total numbers of participants in Table 1 differed from those in Table 2 and 3? For example those with low strain were 5,358 in Table 1, whereas it was only 4,023 in table 2 and 5,242

in table 3.

Response
Thank you so much for pointing this out. We have corrected a mistake in table 2 and 3 where in the column – No of exp cases and (total) - we had made a mistake and instead of the total number put the non exposed cases in the brackets. We have no gone through and re-calculated everything again to see that the sum is correct.

8. In any multivariate analysis all potential confounders should be included in the analysis and this is of particular importance for the main demographic variable like age and gender. As describe in table 2 and 3 both age and gender were not included in the multivariate analysis. Further the variables included in the multivariate analysis for job strain and sleep disturbance were different. In the job strain model socio-economic class, work load and support from superior were included, whereas for the sleep disturbance model only economic stress were included. It is difficult for the reader to follow your analysis as you have not presented univariate analysis for your confounder and outcome variable (NSAP).

Response
Thank you for discussing this again, however, we already explained this in our last letter:
"We don't consider that presenting a univariate analysis of all potential confounding factors is relevant for the aim of this study. We tested all confounding factors one by one to the crude regression model to study if they changed the crude OR by 10% or more (2). For those that changed the OR we included those in the adjusted analyses. For table 3 we will include the variables of course. Thank you for pointing this out. If the confounding analysis would be of interested to the readers they may be presented as an appendix of course."

9. The presentation of p-value in a logistic regression model is not important as the 95%CI will have provided the similar information. When presenting p-value of extremely small value it would be advisable to present as <0.01 or <0.001 (depending on the number of decimals spaces used in your table) and not as <0.00.

Response
Thank you for pointing this out. We have now revised this.

10. The number of decimal spaces used in your tables and report should be consistent, in Table 2 you have presented both 0.05 and 0.00

Response
Thank you for pointing this out. We have revised this.