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)	Preoperative predictors of poor acute postoperative pain control: a systematic review and meta-analysis
AUTHORS	Yang, Michael M.H.; Hartley, Rebecca; Leung, Alexander; Ronksley, Paul; Jette, Nathalie; Casha, Steven; Riva-Cambrin, Jay

VERSION 1 - REVIEW

REVIEWER	Sukanya Mitra Government Medical College and Hospital, Chandigarh, India
REVIEW RETURNED	23-Aug-2018

GENERAL COMMENTS	The authors have done a commendable job in an important area, with pragmatic implications for clinical practice. Overall, they have followed sound methodology, appraised potential bias, and reported in a balanced manner. I have a few specific comments. 1. A central issue in this work is the definition of "poor acute postoperative pain control", because the entire work rests on this as the dependent variable. However, as evident from Table 1, this definition has varied very widely from study to study, in terms of time period of the studies, specific time when the pain measuring instrument was applied, whether at rest or on movement (mostly not mentioned), but most importantly, the wide range of scores implied to define poor pain control, for example, ranging from >3 up to >7 on NRS across different studies! There is another conceptual issue regarding defining and documenting adequate/poor postoperative pain control. Many studies utilize various patient-controlled analgesic devices these days, where the patient can control his/her pain beyond a certain point (say, >4 on NRS) by administering an additional dose of the analgesic (within the constraints of the lock-out period, of course). In these cases, pain control (or lack of it) cannot be judged by NRS/VRS/VAS pain scores but rather by postoperative consumption of rescue doses or additional doses of analgesics. All these issues make defining "poor acute postoperative pain control" a tough job. 2. The authors have skirted this issue by mentioning that poor postoperative pain control was "as defined by individual study authors" (page 6, line 22 of the PDF). Given the difficulties alluded to above, this solution, though not ideal, is acceptable to me, with THREE specific suggestions: (a) The title of the manuscript should reflect this, by adding "(as defined by individual study authors)" just after "postoperative pain control". (b) The Abstract should also reflect this, by adding the same phrase on page 2, line 21, within parentheses, just after "postoperative pain control".

	(c) The Limitations section of the Discussion must address this issue in some more details than the current cursory passing mention there. 3. Another issue that needs a more detailed discussion under the Limitations section is the wide heterogeneity observed between studies regarding quite a few study variables. To what extent does this heterogeneity weaken the conclusions about the predictors? 4. Table 1, line 5-6. The heading should include "definition of poor pain CONTROL" (the word control is currently missing),

REVIEWER	Benno Rehberg-Klug Hôpitaux Universitaires de Genève, Switzerland
REVIEW RETURNED	15-Sep-2018

GENERAL COMMENTS	This systematic review of risk factors for poor postoperative pain control is indeed an important study, since no such review has been performed recently. The methodology is adequate, although today systematic reviews should include trial sequential analysis to judge the adequacy of existing data. Major comments: It is difficult to judge the search strategy, but I do think that studies have been missed, potentially by the search term “pain measurement”. By a quick look in my reference library I identified several studies, which might have been included: 1. Holtzman S, Clarke HA, McCluskey SA, Turcotte K, Grant D, Katz J. Acute and chronic postsurgical pain after living liver donation: Incidence and predictors. Liver Transpl. 2014 Nov;20(11):1336–46. 2. Hetmann F, Schou-Bredal I, Sandvik L, Kongsgaard UE. Does chronic pre-operative pain predict severe post-operative pain after thoracotomy? A prospective longitudinal study. Acta Anaesthesiol Scand. 2013 Sep;57(8):1065–72. 3. Gerbershagen HJ, Dagtekin O, Rothe T, Heidenreich A, Gerbershagen K, Sabatowski R, et al. Risk factors for acute and chronic postoperative pain in patients with benign and malignant renal disease after nephrectomy. Eur J Pain. 2009 Sep;13(8):853–60. 4. Sommer M, Geurts JWJM, Stessel B, Kessels AGH, Peters ML, Patijn J, et al. Prevalence and Predictors of Postoperative Pain After Ear, Nose, and Throat Surgery. Arch Otolaryngol Head Neck Surg. 2009 Feb 16;135(2):124–30. 5. Strulov L, Zimmer EZ, Granot M, Tamir A, Jakobi P, Lowenstein L. Pain Catastrophizing, Response to Experimental Heat Stimuli, and Post–Cesarean Section Pain. The Journal of Pain. 2007 Mar 1;8(3):273–9. 6. Granot M, Ferber SG. The Roles of Pain Catastrophizing and Anxiety in the Prediction of Postoperative Pain Intensity: A Prospective Study. The Clinical Journal of Pain. 2005 Sep;21(5):439–45. 7. Pavlin DJ, Sullivan MJL, Freund PR, Roesen K. Catastrophizing: a risk factor for postsurgical pain. Clin J Pain. 2005 Feb;21(1):83–90. 8. Jacobsen PB, Butler RW. Relation of cognitive coping and catastrophizing to acute pain and analgesic use following breast cancer surgery. J Behav Med. 1996 Feb 1;19(1):17–29.

	When looking at the following non-systematic review: Khan RS, Ahmed K, Blakeway E, Skapinakis P, Nihoyannopoulos L, Macleod K, et al. Catastrophizing: a predictive factor for postoperative pain. Am J Surg. 2011 Jan;201(1):122–31. It may be that when including the references cited here catastrophizing could have been a significant factor. In addition, no ongoing or non-published studies were searched in registries like “clinicaltrials.gov”. From a methodological point in acute pain, planning subgroup analysis for surgical specialty is not ideal; ideally, there should be subgroups for individual types of surgery, or at least narrow groups. This should be mentioned in “limitations”. Minor comments: P7, line 35 and 40: reference 35 does not seem to adequate here P9, line 5: the phrase “studies not using a numeric scale were considered moderate pain” is not clear P10, line 21: the term “exposures” is not clear, better would be “predictive variables” or “predictors” P11 line 19ff: Figures 3 and 4 and Table 2 contain the same data, one of them can be deleted P12 line 28: it is mentioned that “surgical discipline” showed no differences in pooled estimates. I think the number of studies for some predictors is just not sufficient to do these stratifications. P13, line 3: it should be mentioned that “history of sleeping difficulties” was analysed in only 2 studies P13 line 47: you mention that that the association of age as a continuous variable was not significant. When looking at the Forest plot in Figure S1 b), it looks as if the study of Gerbershagen has not been included in the analysis for age. Why? P14, lines 42-54: please delete the discussion of BMI, there was only 1 study in which this predictor was significant! This study was on TKA patients, where BMI as a risk factor is obvious. The other non-significant study was breast surgery. The difference between these studies shows well that the type of surgery may indeed make a difference. P15, lines 10-22: delete the paragraph about pain catastrophizing, it does not contain any new ideas or facts. Table 1: - Please add a column of baseline risk, i.e. the incidence of “poor pain” for each study - Adding the reference numbers would be helpful - The study of Gerbershagen was performed in Germany Table 2: - Please add a column with the total number of patients for each pooled estimate - Please add also a column with the pooled baseline risk for each predictor Supplementary Figure S1: - Add a column with study size (and ideally also numbers of poor pain control in both groups for binary predictors) - add figures for the non-significant predictors

REVIEWER	Dr. Subas Neupane University of Tampere, Finland
REVIEW RETURNED	23-Nov-2018

GENERAL COMMENTS	This is an interesting study aimed to identify preoperative predictors of poor postoperative pain control in an adult population undergoing inpatient surgery. My review is focused on the statistical part only and my major concerns on the statistical part are presented below. It is not clear that whether the study focused on working or non-working population. Also, the preoperative predictors could have been specified for e.g. demographic or lifestyle or other medical conditions etc. in order to make the focus of the study in a particular aspect. It is interesting that the author included the study that has reported only the measure of association either as OR or RR. I wonder why it was not possible to include all studies and limit the meta-analysis only to that study that had reported OR or RR. Please provide for how many studies the author calculated ORs manually from the raw data? The authors said that they have used cohort or cross-sectional studies in their review. But in the meta-analysis, these studies are not separated. I suggest that these studies should be analyzed separately as the cohort study might have reported the incidence of pain while the cross-sectional studies report the prevalence, also the evidence based on cohort and cross-sectional studies can't be the same. It seems that there are only two cross-sectional studies included which can even be excluded from the study to make it more focused. Also, provide the follow-up duration for each of the cohort studies in Table 1. It is not very clear why statistically significant preoperative predictors are presented separately from statistically non-significant predictors? As it is shown in figure 3 and 4, the number of studies used for meta-analysis for each of the predictors are different and obviously the figure containing statistically non-significant pooled estimates included comparatively fewer studies. Figure 3 and 4 are the repetition of the information provided in Table 2, which can be deleted.

REVIEWER	Claire Cameron Biostatistics Unit, Dunedin School of Medicine, University of Otago, PO box 56, Dunedin, New Zealand
REVIEW RETURNED	28-Nov-2018

GENERAL COMMENTS	This is a very good manuscript and I appreciate having the opportunity to review it. It describes a systematic review and meta-analysis examining the association between pre-operative patient level predictors and poor postoperative pain control. It is more complex than a standard meta-analysis looking for a single pooled estimate. I enjoyed reading it and I am impressed at the clarity of the reporting of the results given it's complexity. The agree with the statistical decisions the authors have made. The analysis is very thorough and the checklist (MOOSE) itemises this. Note that I said 'N/A' to the question about research ethics as this study did not involve people at all. Also, I said 'N/A' to the question

	about the references because I was asked to provide a statistical review and the references are more to do with the content. I recommend this manuscript be accepted with minor revisions. I have just a few questions (and a couple of typos). Sadly, this list is longer than my positive comments – which does not reflect that fact that I think this manuscript is excellent. 1. You have included cohort and cross-sectional studies. Cross-sectional covers a multitude of approaches. Could you provide some information about the nature of those studies in terms of how the data was collected? It is just not clear to me how a cross sectional study would have data on preoperative exposures and post operative pain management. 2. Given all the studies are observational is seems surprising that a measure of quality in those studies is 'blinded outcome assessment'. I realise that is an important quality measure for a randomised trial but it is not so clear for an observational study. I also notice that many of the studies (the majority) do not clearly indicate their blinding (Figure 2). Could something be added about the importance of blinding in an observational study? 3. Feeling curious about the previous point I went to the citation from Table S1 and it took me to Higgins et al (2003) which is a paper about heterogeneity. I think it should go to Hayden et al (2006) – number 37 in your references. 4. Page 6, line 19, you say '…and studies that assessed for the association between postoperative patient-level predictors of poor postoperative pain control…'. Shouldn't that say '…and studies that assessed for the association between postoperative patient-level predictors and poor postoperative pain control…'. Sorry, I am slightly confused. You are looking at the association between the predictors and the postoperative pain, not the association between the predictors? 5. In the discussion, page 14, line 47, you say 'studies examining BMI as a dichotomous variable were inadequately powered to detect a statistical difference'. I am always uneasy about statements like this as they suggest that there is a significant difference there that could not detected – which may or may not be true. I assume that what is meant here is that the studies were small and so the inference from them was not convincing. I think it would be good to rephrase that sentence. 6. In table 2 (page 26, line 11), one of the preoperative predictors is 'Females sex' when I think it should read 'Female sex'.

REVIEWER	Prof Simon Skene University of Surrey, UK
REVIEW RETURNED	29-Nov-2018

GENERAL COMMENTS	This is an interesting and informative study which is comprehensively described. The authors have been clear about their methods, and unambiguous about the limitations. The biggest limitation is the great heterogeneity between studies, but the authors have attempted to investigate/explain this using stratified meta-analysis and meta-regression on appropriate factors. These analyses did not throw additional light on the heterogeneity, but the

	concordance of results across multiple methods gives some confidence. It would have been interesting perhaps to further expand on additional research into this topic which was alluded to at the end.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Reviewer Name: Sukanya Mitra
Institution and Country: Government Medical College and Hospital, Chandigarh, India
Please state any competing interests or state 'None declared': None declared

Please leave your comments for the authors below:

The authors have done a commendable job in an important area, with pragmatic implications for clinical practice. Overall, they have followed sound methodology, appraised potential bias, and reported in a balanced manner.

I have a few specific comments.

Comment #1:
A central issue in this work is the definition of "poor acute postoperative pain control", because the entire work rests on this as the dependent variable. However, as evident from Table 1, this definition has varied very widely from study to study, in terms of time period of the studies, specific time when the pain measuring instrument was applied, whether at rest or on movement (mostly not mentioned), but most importantly, the wide range of scores implied to define poor pain control, for example, ranging from >3 up to >7 on NRS across different studies! There is another conceptual issue regarding defining and documenting adequate/poor postoperative pain control. Many studies utilize various patient-controlled analgesic devices these days, where the patient can control his/her pain beyond a certain point (say, >4 on NRS) by administering an additional dose of the analgesic (within the constraints of the lock-out period, of course). In these cases, pain control (or lack of it) cannot be judged by NRS/VRS/VAS pain scores but rather by postoperative consumption of rescue doses or additional doses of analgesics. All these issues make defining "poor acute postoperative pain control" a tough job.

Response:
We would like to thank Dr. Sukanya Mitra for her insightful comments and suggestions. We agree with her assessment that this is a particularly difficult subject to perform a systematic review and meta-analysis given how heterogenous the studies are in terms of surgical specialty/procedure, timing of pain assessment, and definitions for poor pain control to name a few. We believe this systematic review serves to highlight your very point of the importance of standardizing pain studies in the postoperative period in the future (we have discussed and expanded this in the limitation section of our manuscript, see response to comment #3 below). Having common data elements in future pain studies will ensure we can reliably compare results between studies. However, this study represents the best summary of the current literature on this topic.

No changes were made to the manuscript as a result of this comment.

Comment #2:
The authors have skirted this issue by mentioning that poor postoperative pain control was "as defined by individual study authors" (page 6, line 22 of the PDF). Given the difficulties alluded to above, this solution, though not ideal, is acceptable to me, with THREE specific suggestions:
(a) The title of the manuscript should reflect this, by adding "(as defined by individual study authors)" just after "postoperative pain control".
(b) The Abstract should also reflect this, by adding the same phrase on page 2, line 21, within parentheses, just after "postoperative pain control".
(c) The Limitations section of the Discussion must address this issue in some more details than the current cursory passing mention there.

Response:
(a) and (b): We agree with the reviewer that we need to make clear for the readers that the definition of poor postoperative pain control was determined by the individual study authors.

As such, we added the following to the abstract section on page 2 (changes in bold): "Studies in any language were included if they evaluated postoperative pain using a validated instrument (e.g., visual-analogue-scale for pain) in adults (≥18 years) and reported a measure of association between poor postoperative pain control **(as defined by individual study authors)** and at least one preoperative predictor during the hospital stay".

However, we feel adding the statement "as defined by individual study authors" in the title of the article would make the title too long and distract prospective readers. Since the vast majority of readers will read the abstract, we believe having added the statement above in the methods section of the abstract will be sufficient. If the journal disagrees, we are happy to make this addition to the title.

(c): We will respond to this comment below with comment #3.

Comment #3
Another issue that needs a more detailed discussion under the Limitations section is the wide heterogeneity observed between studies regarding quite a few study variables. To what extent does this heterogeneity weaken the conclusions about the predictors?

Response:
We agree with the reviewer that this is a limitation of our study. Heterogeneity is a particular challenge that is common to all meta-analyses of prognostic studies. To our knowledge, there is no standardized way to report pain in the acute postoperative period. As such, it is often up to the individual study authors to determine their study design and method of reporting. We have attempted to limit this heterogeneity using our inclusion/exclusion criteria. For example, we limited studies that evaluated predictors of poor pain control during the inpatient period and only included studies that reported measures of association as an odds ratio, risk ratio, or presented raw data for its calculation. Future studies should attempt to develop common data elements so that studies can be reliably compared. This can be done through creation of research groups, seek consistency in cut-offs for continuous variables, adjustment factors, definition of poor pain control, and instrument used to measure pain. It is difficult to quantify how this heterogeneity weakens the conclusions of this study. Nevertheless, as indicated by reviewer #4, Dr. Claire Cameron, this study "was more complex than a standard meta-analysis looking [at] a single pooled estimate. I enjoyed reading it and I am impressed at the clarity of the reporting of the results given it's complexity." She also "agree[s] with the statistical decisions the authors have made."

In response to this reviewer's comments, we have made the following addition to the limitation section of the discussion on page 15:

"Further, we performed meta-analyses on studies that had appreciable heterogeneity as it pertains to definition of poor postoperative pain control (which was variably defined by individual study authors), surgical procedure/specialty, timing and instrument used for pain assessment, and threshold used to categorize continuous preoperative predictors between studies (e.g., young vs. old). Outcome heterogeneity may have been a potential source of bias if, for example, a particular predictor was associated with an increased risk of postoperative pain with one instrument (or cut-off) and a decreased risk of pain using a different instrument (or cut-off). In such cases, a pooled analysis might fail to detect either finding. Although we do not believe this issue biased our findings, future studies should attempt to standardize definitions (common data elements) to facilitate comparisons between studies."

We also updated the third bullet point under the "Article Summary" section on page 3 to the following:

"There were a variety of definitions for poor postoperative pain control, timing of pain assessment, and thresholds used to categorize continuous preoperative variables making the clinical and statistical interpretation of the meta-analysis more challenging.

Comment #4
Table 1, line 5-6. The heading should include "definition of poor pain CONTROL" (the word control is currently missing),

Response: We have modified Table 1 to include “Definition of Poor Pain **Control**”.

Reviewer: 2
Reviewer Name: Benno Rehberg-Klug
Institution and Country: Hôpitaux Universitaires de Genève, Switzerland
Please state any competing interests or state ‘None declared’: none declared

Please leave your comments for the authors below:

This systematic review of risk factors for poor postoperative pain control is indeed an important study, since no such review has been performed recently.
The methodology is adequate, although today systematic reviews should include trial sequential analysis to judge the adequacy of existing data.

Major Comments:

Comment #1
It is difficult to judge the search strategy, but I do think that studies have been missed, potentially by the search term “pain measurement”.

Response:
We would like to thank Dr. Benno Rehberg-Klug for his many helpful comments. We agree it is difficult to judge whether the current search strategy is inclusive enough; which is often the case for systematic reviews. For our study, we had two independent research librarians from the University of Calgary help develop our search strategy using the PRESS method. We had 4 themes in our search, of which 2 were “postoperative pain” and “pain measurement”. Both of these terms are MESH terms. We repeated the MEDLINE search using these two MESH terms and found there was considerable overlap between the two headings (11,103 articles). As such, we feel confident that our current search strategy supplemented with a manual check of reference lists of included study was comprehensive and unlikely to miss a number of key studies. Of the 9 articles this reviewer has listed, only 2 articles were not captured in our initial search (Holztman et al, and Strulov et al) of the 4 electronic databases (MEDLINE, EMBASE, CINAHL, Psychinfo) and the grey literature search. These 2 articles would have not met the entry criteria of our systematic review. The remaining 7 articles that were listed was captured in our initial database search and were eliminated either during the title and abstract or full-text review stage. No changes were made in the manuscript as a result of this comment.

Below we have outlined the rationale why each of the listed article did not meet criteria to be include in our systematic review:

1. Holtzman S, Clarke HA, McCluskey SA, Turcotte K, Grant D, Katz J. Acute and chronic postsurgical pain after living liver donation: Incidence and predictors. Liver Transpl. 2014 Nov;20(11):1336–46.
    - This article was not captured in our initial electronic database search. However, this article does not fit into the entry criteria of our systematic review as it does not report the association between preoperative predictors and poor postoperative pain control as an OR, RR, nor does it provide the raw data for its calculation.

2. Hetmann F, Schou-Bredal I, Sandvik L, Kongsgaard UE. Does chronic pre-operative pain predict severe post-operative pain after thoracotomy? A prospective longitudinal study. Acta Anaesthesiol Scand. 2013 Sep;57(8):1065–72.
    - This article was captured in our electronic database search but was eliminated during the full text review stage. The outcome of this article was not poor postoperative pain control, as such it does not fit into the entry criteria for our systematic review.

3. Gerbershagen HJ, Dagtekin O, Rothe T, Heidenreich A, Gerbershagen K, Sabatowski R, et al. Risk

factors for acute and chronic postoperative pain in patients with benign and malignant renal disease after nephrectomy. Eur J Pain. 2009 Sep;13(8):853–60.

- This article was captured in our electronic database search but was eliminated in the full text review stage because the study did not provide analysis on the predictors of poor postoperative pain control during the inpatient period.

4. Sommer M, Geurts JWJM, Stessel B, Kessels AGH, Peters ML, Patijn J, et al. Prevalence and Predictors of Postoperative Pain After Ear, Nose, and Throat Surgery. Arch Otolaryngol Head Neck Surg. 2009 Feb 16;135(2):124–30.

- This article was captured in our electronic database search but was eliminated because it was a secondary publication of a paper that was included in the systematic review: Sommer et al (2010): Predictors of acute postoperative pain after elective surgery.

5. Strulov L, Zimmer EZ, Granot M, Tamir A, Jakobi P, Lowenstein L. Pain Catastrophizing, Response to Experimental Heat Stimuli, and Post–Cesarean Section Pain. The Journal of Pain. 2007 Mar 1;8(3):273–9.

- This article was not captured in our initial electronic database search. However, this article does not meet the entry criteria of the systematic review because the outcome was not poor postoperative pain control after surgery. The authors performed a multiple linear regression analysis on postoperative pain with various preoperative variables. Further, the authors did not present a measure of variance/standard error (e.g., 95% confidence intervals) to allow for pooling of data required for a meta-analysis.

6. Granot M, Ferber SG. The Roles of Pain Catastrophizing and Anxiety in the Prediction of Postoperative Pain Intensity: A Prospective Study. The Clinical Journal of Pain. 2005 Sep;21(5):439–45.

- This article was captured in our electronic database search but was eliminated during our title and abstract screening stage. This article does not meet the entry criteria of the systematic review because the outcome was not poor postoperative pain control after surgery. The authors performed a hierarchal multiple linear regression analysis on postoperative pain (continuous) on various preoperative variables.

7. Pavlin DJ, Sullivan MJL, Freund PR, Roesen K. Catastrophizing: a risk factor for postsurgical pain. Clin J Pain. 2005 Feb;21(1):83–90.

- This article was captured in our electronic database search but was eliminated during the full text review stage. This article presented point estimates for measure of associations but did not provide a measure of variance to allow inclusion in the meta-analysis.

8. Jacobsen PB, Butler RW. Relation of cognitive coping and catastrophizing to acute pain and analgesic use following breast cancer surgery. J Behav Med. 1996 Feb 1;19(1):17–29.

- This article was captured in our electronic database search but was eliminated during the full text review stage. This article was excluded because it did not present measures of association in OR, RR or provide raw data for its calculation. This article used Pearson correlation coefficients to present their associations. Further, this article did not provide a measure of variance to allow pooling of results in a meta-analysis.

9. Khan RS, Ahmed K, Blakeway E, Skapinakis P, Nihoyannopoulos L, Macleod K, et al. Catastrophizing: a predictive factor for postoperative pain. Am J Surg. 2011 Jan;201(1):122–31.

- This article was captured in our electronic database search but was eliminated during the title and abstract stage because it was a review article.

Comment #2

In addition, no ongoing or non-published studies were searched in registries like "clinicaltrials.gov".

Response:

We agree this is potentially a limitation of our systematic review as we may miss important studies that could be included in our review. However, we determined a priori to not include studies that are ongoing or are not-published as indicated in our PROSPERO protocol. We did not search clinicaltrials.gov because this site mainly registers clinical trials, whereas the study design we were interested in including were observational studies.

No changes to the manuscript were made for this comment.

Comment #3
From a methodological point in acute pain, planning subgroup analysis for surgical specialty is not ideal; ideally, there should be subgroups for individual types of surgery, or at least narrow groups. This should be mentioned in "limitations".

Response:
We agree with this reviewer that planning subgroup analyses for surgical specialties is not ideal. Rather, subgroups should be analyzed for individual types of surgery to reduce the degree of heterogeneity. We did not perform subgroup analyses based on individual surgery types due to the limited number of studies (n=33) included in this systematic review and the fact that many of the included studies evaluated a variety of surgical procedures making subgroup analysis for specific surgery type impossible.

We have included the following in the limitation section of the manuscript which addresses this reviewer's comment (pages 15-16):

"Finally, there was significant statistical heterogeneity between studies, which could not be explained by stratified analysis or meta-regression based on a variety of clinical and study design factors (and the results should be interpreted with caution for surgical discipline as there were limited number of studies in each group). This heterogeneity was likely a product of important clinical differences as the included studies differed widely in surgery type and case-mix. Additional research may further define the influence of specific types of surgery on pain control."

Minor comments:

Comment #4: P7, line 35 and 40: reference 35 does not seem to adequate here

Response: We have replaced the reference between line 35-40 on page 7 with Hayden JA, Cote P, Bombardier C. Evaluation of the quality of prognosis studies in systematic reviews. Ann Intern Med 2006;144(6):427-37. [published Online First: 2006/03/22]. This article provides a guide to perform risk of bias assessments for systematic reviews on prognostic studies. This article was used to help develop our risk of bias/study quality assessment scheme.

Comment #5: P9, line 5: the phrase "studies not using a numeric scale were considered moderate pain" is not clear

Response: Auburn et al (2008) used morphine requirements > 15mg/kg in the post-anesthetic unit as their definition of poor postoperative pain control. Since this was not a numeric scale for pain (e.g., visual analogue scale for pain) this study was categorized as having "moderate pain" for the purpose of performing a stratified meta-analysis.

As such we have made the following clarifications on page 9 line 5 (changes in bold): "… definition of poor postoperative pain control (moderate vs. severe pain; moderate pain: 3-6, severe pain: >6 on an 11-point scale; studies not using a numeric scale **(e.g., morphine requirements as the definition for poor pain control)** were considered moderate pain)"

Comment #6: P10, line 21: the term "exposures" is not clear, better would be "predictive variables" or "predictors"

Response: We have updated the language from "exposures" to "predictors" throughout the manuscript, including the abstract, body of manuscript, Table 1, caption for Figure 2, and online supplementary Table S1.

Comment #7: P11 line 19ff: Figures 3 and 4 and Table 2 contain the same data, one of them can be deleted.

Response: Table 2 contains the definitions of preoperative predictors used by authors of the included studies (e.g., low ASA physical status is defined as ASA I compared to II or III). Hence, it is essential Table 2 is left in the manuscript. We agree with the reviewer that Figures 3 and 4 depict the same data as in Table 2. However, we feel displaying the results in a graphical form gives the reader a quick way to understand our results. In the interest of keeping the number of tables and figures within 5 (as indicated in the instructions for authors document), we have elected to remove Figure 4, which depicts the **non-significant predictors** of poor postoperative pain control. However, if the journal insists we remove both figures, we are happy to comply.

Comment #8: P12 line 28: it is mentioned that "surgical discipline" showed no differences in pooled estimates. I think the number of studies for some predictors is just not sufficient to do these stratifications.

Response: We agree with the reviewer's comments that we did not have sufficient studies to perform stratified analysis based on each surgical discipline. However, this was an a priori determined analysis and we feel compelled to include it in our manuscript. We have added a sentence in the limitation section on page 15 addressing this: "(the results should be interpreted with caution for surgical discipline as there were limited number of studies in each group)".

Comment #9: P13, line 3: it should be mentioned that "history of sleeping difficulties" was analysed in only 2 studies

Response: We have made the following changes in the first paragraph of the discussion on page 12 (changes in bold): "The strongest negative prognostic factors were a history of sleeping difficulties **(number of studies, n=2)** and depression **(n=8)**, which were independently associated with approximately 2-fold higher odds of poor postoperative pain control."

Comment #10: P13 line 47: you mention that that the association of age as a continuous variable was not significant. When looking at the Forest plot in Figure S1 b), it looks as if the study of Gerbershagen has not been included in the analysis for age. Why?

Response: The reviewer is correct in saying that the association for age as a **continuous variable** and poor postoperative pain control was not significant. Figure S1 b) is showing the Forest plot for age as a **dichotomous variable.** The study by Gerbershagen evaluated age as a continuous variable. We have included the Forest plot for age as a **continuous variable** below where the study by Gerbershagen was included for the reviewer's and editor's reference

Figure. Forest plot for age as a continuous variable.

Comment #11: P14, lines 42-54: please delete the discussion of BMI, there was only 1 study in which this predictor was significant! This study was on TKA patients, where BMI as a risk factor is obvious. The other non-significant study was breast surgery. The difference between these studies shows well that the type of surgery may indeed make a difference.

Response: We have deleted the following from the discussion on page 14 as per the reviewer's suggestion: "We found that every 5 kg/m^2 increase in BMI, was associated with a 10% higher odds of poor postoperative pain control (when BMI was examined as a continuous variable), though studies examining BMI as a dichotomous variable were inadequately powered to detect a statistical difference. The association between higher BMI levels and adverse pain outcomes may be a product of inadequate dosing of postoperative analgesia and/or greater tissue dissection in these patients leading to more postoperative pain.[48]"

Comment #12: P15, lines 10-22: delete the paragraph about pain catastrophizing, it does not contain any new ideas or facts.

Response: We have deleted the following from the discussion from page 15 as per the reviewer's suggestion: "Surprisingly, there was no detectable association between chronic pain or pain catastrophizing symptoms and poor postoperative pain control. Tasmuth and colleagues[70] described the memory of pain as determined by many factors such as current pain intensity, emotion, the expectation of pain and recent peak intensity of previous pain. Intuitively, chronic pain and the tendency to misinterpret or exaggerate threatening situations might be expected by many to increase the risk of poor postoperative pain outcomes. However, that relationship was not observed in our review."

Comment #13:
Table 1:
-Please add a column of baseline risk, i.e. the incidence of "poor pain" for each study
-Adding the reference numbers would be helpful
-The study of Gerbershagen was performed in Germany

Response: Table 1 has been updated to include an additional column "Incidence of Poor Post-operative Pain Control (%)". Out of the 33 included studies, 8 studies did not report the baseline incidence of poor postoperative pain. References for each study has also been added to Table 1. A correction was made under country of origin for the study by Gerbershagen et al in Table 1.

Comment #14:
Table 2:
-Please add a column with the total number of patients for each pooled estimate
-Please add also a column with the pooled baseline risk for each predictor

Response: We have added an additional column under Table 2 to indicate the total number of patients included in the pooled analysis for each predictor. We could not include a column containing the pooled baseline risk for each predictor because this data is unattainable in many of the included studies (and therefore cannot be presented in a meaningful manner). However, the odds ratios in Table 2 demonstrate the measures of associations for each predictor on poor postoperative pain control.

Comment #15:
Supplementary Figure S1
-Add a column with study size (and ideally also numbers of poor pain control in both groups for binary predictors) add figures for the non-significant predictors

Response: We have added new online supplemental Figures S4 to S8 which includes all the Forest plots for the non-significant predictors. Accordingly, we have added the following in the results section on page 11: "Detailed forest plots for each non-significant preoperative predictor are shown in online supplemental Figures S4 to S8."

Unfortunately, not all of the studies presented the raw number of patients in each of the group when calculating their measures of associations. To get around this, we used the "metan" command within the statistical software STATA Version 15, which just requires the point estimate (odds ratios) and the upper and lower limit of the 95% confidence interval for the odds ratio in order to calculate the pooled odds ratio. We agree having numbers in both groups would be very informative, however, we were not able to generate these figures given the limitations of the included studies. The sample size for each of the included studies and for each predictor are indicated in Table 1 and 2. We feel generating new forest plots for the online supplement with this information does not add to the quality of the manuscript.

Reviewer: 3
Reviewer Name: Dr. Subas Neupane
Institution and Country: University of Tampere, Finland
Please state any competing interests or state 'None declared': None declared

Please leave your comments for the authors below
This is an interesting study aimed to identify preoperative predictors of poor postoperative pain control in an adult population undergoing inpatient surgery. My review is focused on the statistical part only and my major concerns on the statistical part are presented below.

Comment #1
It is not clear that whether the study focused on working or non-working population. Also, the preoperative predictors could have been specified for e.g. demographic or lifestyle or other medical conditions etc. in order to make the focus of the study in a particular aspect.

Response:
We would like to thank Dr. Subas Neupane for his helpful comments. Of the 33 studies included in our systematic review, only the study by Lesin et al (2016) evaluated employment status as one of the predictors for poor postoperative pain control. Since this was the only study that reported this predictor, this predictor was not included in our meta-analysis. However, for the reviewer's interest, this study showed employment status had no impact on poor postoperative pain control after surgery (p=0.19).

Unfortunately, we were limited by the preoperative predictors evaluated by the included studies, and these studies did not evaluate the impact of lifestyle or other medical condition (except for smoking, alcohol use, and diabetes) on poor postoperative pain control.

No changes were made to the manuscript as a result of this comment.

Comment #2
It is interesting that the author included the study that has reported only the measure of association either as OR or RR. I wonder why it was not possible to include all studies and limit the meta-analysis only to that study that had reported OR or RR. Please provide for how many studies the author calculated ORs manually from the raw data?

Response:
While performing our title & abstract and full-text screening we found a wide variety methods of reporting measure of associations. For example, we came across authors using odds ratios, Pearson correlation coefficients, 2x2 tables, and comparing means and medians. One of the challenges in performing meta-analysis for prognostic studies is that there is no standardize way to report measures of associations. As such we decided to limit studies that presented their results in OR, RR, or its raw data for its calculation to facilitate the meta-analysis.

The main goal for this manuscript was to perform a meta-analysis on the highest quality papers on the subject. These papers represent those studies that report their measures of associations as ORs, RRs, or its raw data for its calculation (which allows the option for adjustment for potential confounders). Further, our systematic review serves as an update of a previous systematic review by Ip et al (2009) that included many of the studies that have been excluded in this review. Therefore, we did not feel repeating a systematic review would add significant value.

ORs were calculated when the authors presented their data as 2x2 tables, allowing the calculation of the unadjusted ORs. Unfortunately, we did not keep a detailed record of how many variables were hand calculated. However, all calculations were performed in STATA using the command "cci" which will give an OR point estimate and the 95% confidence interval.

No changes were made to the manuscript as a result of this comment.

Reviewer: 4
Reviewer Name: Claire Cameron
Institution and Country: Biostatistics Unit, Dunedin School of Medicine, University of Otago, PO box 56, Dunedin, New Zealand
Please state any competing interests or state 'None declared': None Declared

Please leave your comments for the authors below:

This is a very good manuscript and I appreciate having the opportunity to review it. It describes a systematic review and meta-analysis examining the association between pre-operative patient level predictors and poor postoperative pain control. It is more complex than a standard meta-analysis looking for a single pooled estimate. I enjoyed reading it and I am impressed at the clarity of the reporting of the results given it's complexity. The agree with the statistical decisions the authors have made. The analysis is very thorough and the checklist (MOOSE) itemises this.

Note that I said 'N/A' to the question about research ethics as this study did not involve people at all. Also, I said 'N/A' to the question about the references because I was asked to provide a statistical review and the references are more to do with the content.

I recommend this manuscript be accepted with minor revisions. I have just a few questions (and a couple of typos). Sadly, this list is longer than my positive comments – which does not reflect that fact that I think this manuscript is excellent.

Comment #1:
You have included cohort and cross-sectional studies. Cross-sectional covers a multitude of approaches. Could you provide some information about the nature of those studies in terms of how the data was collected? It is just not clear to me how a cross sectional study would have data on preoperative exposures and post operative pain management.

Response:
We would like to thank Dr. Claire Cameron for her commendatory comments towards our manuscript and for her helpful comments. Two of the included studies were identified by the study authors as cross-sectional studies. The first study by Liu et al (2012), was a multi-center, cross-sectional study evaluating the predictors for moderate to severe acute postoperative pain after total hip and knee replacement. Patients who met the inclusion criteria had their postoperative pain scored prospectively using a numeric rating scale for pain on postoperative day 1. All their candidate preoperative predictors were retrospectively collected from patient charts which include variables such as age, gender, home opioids use, alcohol use etc. The second study was by Storesund et al (2016) predictors of poor pain control after ankle surgery. Similar to the previous studies, pain data was prospectively collected while the preoperative predictors were obtained retrospectively though patient charts.

Despite these study authors labelling their studies as cross-sectional, after a detailed review of their methodology, we feel both of these studies are better categorized as retrospective cohort studies.

As such, we have made the following updates in the manuscript:

1) Both of these studies have been recategorized as "retrospective cohort studies" in Table 1. A footnote under Table 1 was created for both of the above studies to indicate that the study authors categorized their studies as cross-sectional: "Labelled as a cross-sectional study design by study authors, but methodology more represent a retrospective cohort study design."
2) We have updated the second paragraph of the result section to on page 9 to: "Twenty-six studies were prospective cohort studies (79%) and 7 were retrospective cohort studies (21%)."

Comment #2
Given all the studies are observational is seems surprising that a measure of quality in those studies is 'blinded outcome assessment'. I realise that is an important quality measure for a randomised trial but it is not so clear for an observational study. I also notice that many of the studies (the majority) do not clearly indicate their blinding (Figure 2). Could something be added about the importance of blinding in an observational study?

Response:
The definition of the quality indicator "blinded outcome assessment" in this study was: study reported that outcomes were assessed by persons without knowledge of prognostic factors or that the pain outcome was determined by personnel not aware of study objectives. This is important in observational studies because if the person involved in collecting postoperative pain data is aware of a particular patient's risk factors, and has a preconceived idea that a certain risk factor would increase the risk of poor pain control after surgery (e.g., if the person believed smoking is a risk factor for poor pain control, and he or she knows the patient is a smoker), this may influence the pain data results leading to misclassification bias. If the person collecting the outcome data is not aware of the patients' risk factor or is not involved in the study, the chances of misclassification bias are greatly reduced.

We have added the following to page 10 under the "assessment of study quality" in the results section (changes in bold):

"In 25 studies (76%), there was either no blinding or no reporting on whether there was blinding of predictors during outcome ascertainment. **The lack of blinding of predictors during outcome ascertainment in the majority of studies could lead to increased risk of misclassification bias.**"

Comment #3
Feeling curious about the previous point I went to the citation from Table S1 and it took me to Higgins et al (2003) which is a paper about heterogeneity. I think it should go to Hayden et al (2006) – number 37 in your references.

Response:
Thank you for catching this, we have now updated the Hayden et al reference to the appropriate location in Table S1 as well as in the "study quality assessment" section of the manuscript.

Comment #4
Page 6, line 19, you say '…and studies that assessed for the association between postoperative patient-level predictors of poor postoperative pain control…'. Shouldn't that say '…and studies that assessed for the association between postoperative patient-level predictors and poor postoperative pain control…'. Sorry, I am slightly confused. You are looking at the association between the predictors and the postoperative pain, not the association between the predictors?

Response:
The reviewer is correct in saying we are looking at the association between various preoperative predictors and poor postoperative pain control. We have made the following correction in the manuscript on page 6 line 19 (changes in bold):

"and studies that assessed for the association between preoperative patient-level predictors **and** poor postoperative pain control (as defined by individual study authors)."

Comment #5
In the discussion, page 14, line 47, you say 'studies examining BMI as a dichotomous variable were inadequately powered to detect a statistical difference'. I am always uneasy about statements like this as they suggest that there is a significant difference there that could not detected – which may or may not be true. I assume that what is meant here is that the studies were small and so the inference from them was not convincing. I think it would be good to rephrase that sentence.

Response:
We agree with the reviewer that statements such as this may potentially mislead readers in thinking if only we had a larger sample size, a significant difference will be observed. However, given comment #11 by reviewer #2, who recommended us deleting this sentence in question, this comment is no longer an issue.

Comment #6
In table 2 (page 26, line 11), one of the preoperative predictors is 'Females sex' when I think it should read 'Female sex'.

Response:
Thank you for catching this, we have made the appropriate correction in Table 2.

Reviewer: 5
Reviewer Name: Prof Simon Skene
Institution and Country: University of Surrey, UK
Please state any competing interests or state 'None declared': None declared

Please leave your comments for the authors below
This is an interesting and informative study which is comprehensively described.

The authors have been clear about their methods, and unambiguous about the limitations. The biggest limitation is the great heterogeneity between studies, but the authors have attempted to investigate/explain this using stratified meta-analysis and meta-regression on appropriate factors. These analyses did not throw additional light on the heterogeneity, but the concordance of results across multiple methods gives some confidence.

Comment #1
It would have been interesting perhaps to further expand on additional research into this topic which was alluded to at the end.

Response:
We would like to thank Professor Simon Skene for his commendatory comments and suggestions. We have expanded the conclusion on page 16 to include the following:

"Although acute postoperative pain is common, no standard criteria exist to classify outcomes. Future work is needed to develop consensus criteria for acute postoperative pain outcomes, ideally as an international, multicenter collaborative using the Delphi method."

VERSION 2 – REVIEW

REVIEWER	Benno Rehberg-Klug Hôpitaux Universitaires de Genève, Switzerland
REVIEW RETURNED	02-Jan-2019

GENERAL COMMENTS	The manuscript has much gained from the revision, and the authors have responded well to the remarks of the reviewer. I still have one minor point: whereas the authors have now well described in the article that ORs for some predictors are based on only few studies, this is not visible in the abstract. At least for "sleeping difficulties" and BMI this should be noted in the abstract. I know space is limited in the abstract, but many people tend to read only abstracts, and may erroneously conclude that evidence even for these predictors is based on 33 studies. Ideally, the number of studies should be given for each predictor.

REVIEWER	Subas Neupane, Tampere University, Finland
REVIEW RETURNED	08-Jan-2019

GENERAL COMMENTS	Thank you for the revised version of the manuscript and the response to my queries. Unfortunately only half of my concerned were answered in the response letter and other comments are not. Therefore I would like authors to look at my remaining comments and revise the paper accordingly. I have repeated my unanswered concerns here below: The authors said that they have used cohort or cross-sectional studies in their review. But in meta-analysis these studies are not separated. I suggest that these studies should be analyzed separately as the cohort study might have reported the incidence of pain while the cross-sectional studies reports the prevalence, also the evidence based on cohort and cross-sectional studies can't be the same. It seems that there are only two cross-sectional studies included which can even excluded from the study to make

	it more focused. Also, provide the follow-up duration for each of the cohort study in Table 1. It is not very clear why statistically significant preoperative predictors are presented separately from statistically non-significant predictors? As it is shown in figure 3 and 4, the number of studies used for meta-analysis for each of the predictors are different and obviously the figure containing statistically non-significant pooled estimates included comparatively fewer studies. Figure 3 and 4 are the repetition of the information provided in Table 2, which can be deleted.

REVIEWER	Claire Cameron Biostatistics Unit, Division of Health Sciences, University of Otago, PO Box 56, Dunedin New Zealand
REVIEW RETURNED	23-Jan-2019

GENERAL COMMENTS	I have just re-reviewed this manuscript in the light of the reviewers comments and their revision of the work. I was, originally, asked to provide a statistical review. I have studied all reviewers comments and subsequent revisions and feel that everything raised has been addressed adequately. With respect to my own original review I am satisfied that my concerns have been addressed. I recommend that this paper be accepted, however, I can't speak for the other reviewers.

VERSION 2 – AUTHOR RESPONSE

Reviewer: 2

Reviewer Name: Benno Rehberg-Klug

Institution and Country: Hôpitaux Universitaires de Genève, Switzerland

Please state any competing interests or state ‘None declared’: none declared

Comment #1

Please leave your comments for the authors below

The manuscript has much gained from the revision, and the authors have responded well to the remarks of the reviewer. I still have one minor point: whereas the authors have now well described in the article that ORs for some predictors are based on only few studies, this is not visible in the abstract. At least for "sleeping difficulties" and BMI this should be noted in the abstract.

I know space is limited in the abstract, but many people tend to read only abstracts, and may erroneously conclude that evidence even for these predictors is based on 33 studies. Ideally, the number of studies should be given for each predictor.

Response

We would like to thank Dr. Benno Rehberg-Klug for his additional helpful comments. We have added the number of studies next to each of the significant predictors in the abstract.

In response to this comment we have made the following changes in the abstract section of the manuscript (changes in bold):

"Thirty-three studies representing 53,362 patients were included in this review. Significant preoperative predictors of poor postoperative pain control included younger age (OR 1.18 [95%CI 1.05-1.32], number of studies, n=14), female sex (OR 1.29 [95%CI 1.17-1.43], n=20), smoking (OR 1.33 [95%CI 1.09-1.61], n=9), history of depressive symptoms (OR 1.71 [95%CI 1.32-2.22], n=8), history of anxiety symptoms (OR 1.22 [95%CI 1.09-1.36], n=10), sleep difficulties (OR 2.32 [95%CI 1.46-3.69], n=2), higher BMI (OR 1.02 [95%CI 1.01-1.03], n=2), presence of preoperative pain (OR 1.21 [95%CI 1.10-1.32], n=13), and use of preoperative analgesia (OR 1.54 [95%CI 1.18-2.03], n=6)…"

Reviewer: 3

Reviewer Name: Subas Neupane,

Institution and Country: Tampere University, Finland

Please state any competing interests or state 'None declared': None declared

Please leave your comments for the authors below

Thank you for the revised version of the manuscript and the response to my queries. Unfortunately only half of my concerned were answered in the response letter and other comments are not. Therefore I would like authors to look at my remaining comments and revise the paper accordingly. I have repeated my unanswered concerns here below:

Comment #1

The authors said that they have used cohort or cross-sectional studies in their review. But in meta-analysis these studies are not separated. I suggest that these studies should be analyzed separately as the cohort study might have reported the incidence of pain while the cross-sectional studies reports the prevalence, also the evidence based on cohort and cross-sectional studies can't be the same. It seems that there are only two cross-sectional studies included which can even excluded from the study to make it more focused. Also, provide the follow-up duration for each of the cohort study in Table 1.

Response:

We would like to thank Dr. Subas Neupane for his comments. We want to sincerely apologize for missing half of your comments in our previous submission. Below we will address your comments.

We were interested in evaluating the predictors for poor acute postoperative pain control while the patients were still in the hospital. Therefore, the "Time of Assessment" in Table 1 reflects the follow-up period. We did not include studies that only measured pain outcomes after the patients were discharged from hospital (e.g., 1-month after the operation). To illustrate an example, Liu et al (2012) followed their patients until postoperative day 1 and evaluated their pain outcome on postoperative day 1. As such, we recorded on Table 1 their "time of assessment" was 24 hours.

The reviewer is correct in saying that two of the included studies were identified as cross-sectional study by the individual authors. The first study by Liu et al (2012), was a multi-center, cross-sectional study evaluating the predictors for moderate to severe acute postoperative pain after total hip and knee replacement. Patients who met the inclusion criteria had their postoperative pain scored prospectively using a numeric rating scale for pain on postoperative day 1. All their candidate preoperative predictors were retrospectively collected from patient charts which include variables such as age, gender, home opioids use, alcohol use etc. The second study was by Storesund et al (2016) evaluating predictors of poor pain control after ankle surgery. Similar to the previous studies, pain data was prospectively collected while the preoperative predictors were obtained retrospectively though patient charts.

Despite these study authors labelling their studies as cross-sectional, after a detailed review of their methodology, we feel both of these studies are better categorized as retrospective cohort studies (as there was a temporal relationship between the preoperative predictors and the pain outcome). Therefore, we feel there was no need to repeat the meta-analysis and exclude these two “cross-sectional” studies.

As such, we have made the following updates in the manuscript:

1) Both of these studies have been recategorized as “retrospective cohort studies” in Table 1. A footnote under Table 1 was created for both of the above studies to indicate that the study authors categorized their studies as cross-sectional: “Labelled as a cross-sectional study design by study authors, but methodology more represent a retrospective cohort study design.”

2) We have updated the second paragraph of the result section on page 9 to: “Twenty-six studies were prospective cohort studies (79%) and 7 were retrospective cohort studies (21%).”

Comment #2

It is not very clear why statistically significant preoperative predictors are presented separately from statistically non-significant predictors? As it is shown in figure 3 and 4, the number of studies used for meta-analysis for each of the predictors are different and obviously the figure containing statistically non-significant pooled estimates included comparatively fewer studies. Figure 3 and 4 are the repetition of the information provided in Table 2, which can be deleted.

Response:

Our rationale for separating the statistically significant and non-significant predictors into two separate figures was simply to attempt to present the results in a more organized way. Table 2 also contains the definitions of preoperative predictors used by authors of the included studies (e.g., low ASA physical status is defined as ASA I compared to II or III). Hence, it is essential Table 2 is left in the manuscript. We agree with the reviewer that Figures 3 and 4 depict the same data as in Table 2. However, we feel displaying the results in a graphical form gives the reader a quick way to understand our results. In the interest of keeping the number of tables and figures within 5 (as indicated in the instructions for authors document), we have elected to remove Figure 4, which depicts the non-significant predictors of poor postoperative pain control.

Reviewer: 4

Reviewer Name: Claire Cameron

Institution and Country: Biostatistics Unit, Division of Health Sciences, University of Otago, PO Box 56, Dunedin New Zealand

Please state any competing interests or state ‘None declared’: None declared

Comment #1

Please leave your comments for the authors below

I have just re-reviewed this manuscript in the light of the reviewers comments and their revision of the work. I was, originally, asked to provide a statistical review. I have studied all reviewers comments and subsequent revisions and feel that everything raised has been addressed adequately. With respect to my own original review I am satisfied that my concerns have been addressed. I recommend that this paper be accepted, however, I can't speak for the other reviewers.

Response:

We would like to thank Dr. Claire Cameron for recommending the Journal to accept this manuscript.

VERSION 3 – REVIEW

REVIEWER	Benno Rehberg-Klug Hôpitaux Universitaires de Genève, Switzerland
REVIEW RETURNED	15-Feb-2019

GENERAL COMMENTS	My comments have been answered completely in this second revision, thank you!