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)	The prevalence of chronic pain in the United Kingdom: a systematic review and meta-analysis of population studies
AUTHORS	Fayaz, Alan; Croft, Peter; Langford, Richard; Donaldson, Liam; Jones, Gareth

VERSION 1 - REVIEW

REVIEWER	Kathryn Mansfield London School of Hygiene and Tropical Medicine United Kingdom I have written a systematic review on a similar topic (CWP prevalence in the general population). I have previously worked with one of the authors (Peter Croft).
REVIEW RETURNED	18-Nov-2015

GENERAL COMMENTS	This paper reports the results of a systematic review and meta-analysis of the prevalence of chronic pain in the UK general population. Results are stratified using the following pain phenotypes: i) chronic pain; ii) fibromyalgia; iii) chronic widespread pain (CWP); and iv) chronic neuropathic pain. Search results allowed meta-analysis of prevalence CWP and the more generic chronic pain phenotype. The authors considered variability in pain prevalence by: age; gender; pain severity; and year of publication (depending on availability of data for each pain phenotype). I feel there are a few revisions that would benefit your manuscript: 1. Exclusion of studies at high risk of bias I understand your decision to exclude studies at high risk of bias from your meta-analyses. However, given that risk of bias assessment is a subjective exercise I would argue that it is inappropriate to exclude these studies entirely from your review. I feel you should at least discuss the high risk studies narratively and the reasons why you felt it necessary to drop them from the meta-analysis rather than simply excluding them entirely (I note that the excluded studies are named in the table in Appendix D, however full references are not given for these studies making it difficult for the reader to identify the studies and confirm whether they agree with your decision to exclude them). A more robust approach would be to present an overall prevalence estimate (including the high risk studies) with follow-up sensitivity analyses first excluding high risk and (then if numbers allow) moderate risk studies, and possible follow-up with a meta-regression to explore the effect of risk of bias. The emphasis in the text would clearly be on the results from the low and moderate risk of bias studies in order not to confuse the reader.

	2. Limitations Any presentation of study limitations in the paper (Article summary and Discussion sections) fails to acknowledge that the search strategy may have missed important articles. Having recently completed a similar study I am aware of at least two studies of chronic widespread pain (CWP) prevalence in the UK that you do not appear to have considered for inclusion in your review [1,2]. Other than a reference in your methods section regarding approaching 'field experts' to identify additional grey literature studies, I have not been able to find any reference to the possibility of publication bias in your paper. 3. Chronic widespread pain meta-analysis In your methods section you state estimates that "were not restricted to age-specific or gender-specific cohorts were incorporated into a meta-analysis". However, Figure 2 includes a CWP prevalence estimate from the Croft et al. 2003 study; since this study is restricted to women only I wonder if it should have been included in the meta-analysis. 4. Consideration of possible reasons for differences in study estimates It would have been helpful to see some attempt to tie up the prevalence estimates with each study's methods (risk of bias assessment). For example, might some aspects of study design account for any particularly low or high estimates? You have considered the year the study was undertaken as one reason, but what about the age/sex distribution of the study populations, or the way the data was collected? I accept this might be difficult to do given the limited number of studies you have, but simply stating the range of prevalence estimates found does not do justice to the amount of work you have done to extract all the relevant information about each study. 5. High I-squared statistic Could you please comment on the accuracy of the pooled prevalence estimates for the various pain phenotypes in light of the high level of heterogeneity demonstrated by the I-squared statistic (98.9% for chronic pain, and 95.2% for CWP). Both in your discussion and in the article summary you state that the studies were 'homogenous'. I think you can state that prevalence estimates were consistent across studies and perhaps pass some comment regarding similar/dissimilar methodological approaches taken by the studies included; however, I think stating that the studies were 'homogenous' in light of the very high I-squared statistic is over-stating your findings. 6. Forest plots To avoid having to cross-reference between the forest plots and relevant text/tables, I would prefer to see all prevalence estimates and their 95% CIs (particularly those for pooled estimates) presented in text alongside the relevant graphical elements of the forest plot. I would also like to see I-squared statistics presented in the figure legends. 7. Nested case-control studies Can you confirm that the prevalence estimates included from the three nested case-control studies were derived from figures from the cohort in which they were nested, as prevalence figures can only be

	derived from case-control studies in very specific circumstances. If this is the case, could you note this in some way in Table 1, as I think many readers will question whether it is appropriate to derive prevalence estimates from case-control studies. 8. Figure 3: Chronic pain prevalence by age I found the x-axis on this figure somewhat misleading. I understand that it is difficult to synthesize this type of data as different papers use different age bands. I think you have taken the mid-point of the age bands presented in each paper for the x-axis figures. I do not think that presenting categorical data as continuous data in this way is really appropriate. I am also not comfortable with your choice x-axis title (median age) as this does not seem to really capture the figures that you are using. Having recently grappled with this problem myself, one approach might be to create one figure containing five panels – each panel would be a graph representing the data from a paper, this would allow you to use the age bands from each paper appropriately on the x-axis. 9. Tables 2i and 2ii Table 2i presents studies of chronic pain prevalence ordered by study size and Table 2ii the chronic pain prevalence studies included in the meta-analysis stratified by publication date. It was difficult to get a sense of these studies with the results repeated and distributed across the two tables. Perhaps, rather than ordering by sample size, you could consider creating one table stratified by the following categories of study: i) studies excluded from the meta-analysis; ii) studies published between 1990 and 2005; iii) studies published between 2006 and 2009; and iv) studies published between 2010 and 2015. In the sixth column of Table 2i, the Mallen et al. 2005 and Croft et al. 1993 papers are missing the number of pain patients (% figures are given). 10. Table 3 Perhaps you could consider stratifying this table by whether the study was included in the meta-analysis, e.g. i) Studies restricted by age; and ii) Studies included in meta-analysis. 11. Chronic widespread pain results - page 8 At the end of the first paragraph on this page you write: “Prevalence estimates were again higher in female (9.0% to 10.4%) than in male participants (12.3% to 17.9%)”. I think you must have transposed the figures for women with those for men. 12. Fibromyalgia results - page 8 On the third paragraph on this page you write: “Due to the practical restrictions of formally diagnosing a patient with fibromyalgia (requiring a history and examination in order to exclude alternative causes for widespread pain 11), only one study [REFERENCE] was able to provide comprehensive data from populations representative of the general population.” Please could you include a reference to the study that you are discussing in the appropriate place. 13. Mean prevalence You make reference to ‘mean prevalence’ in your abstract. Could you please stick to ‘pooled prevalence’ for clarity.

REVIEWER	Steven Kamper The George Institute, University of Sydney, Australia
REVIEW RETURNED	04-Jan-2016

GENERAL COMMENTS	General comment • The authors report on a systematic review of the prevalence of chronic pain in the UK. The authors can be congratulated on a well-conducted study and a well-written manuscript, the content is appropriate for the journal readership. I have no major concerns with the study but offer a few suggestions and have some small queries that I would appreciate being addressed. Specific comments Introduction 1. The Introduction is very brief, which I understand conforms to the style of the journal, but I think comes at the expense of a solid justification for the study. The section would benefit from addition of a few sentences outlining why this study is 'much needed', providing some context in terms of the need from a research, clinical and/or policy perspective. Methods 2. Pg 5, 2nd paragraph; on my first reading, inclusion criterion (i) and the last sentence of the paragraph appeared to contradict one another. After some effort I realise that they don't but perhaps rewording of the last sentence would improve clarity. 3. I miss specification of a prevalence period. In the Results it seems point, and 1-month prevalence are combined, I have no issue with this but it is worth setting out whether this decision made a-priori, and whether other periods were excluded. 4. How were studies assessed as 'High' risk of bias using the tool? 5. Were primary study authors contacted in the cases of unreported data? Results 6. Some studies used GP practice data, it might be worth a sentence explaining how these data can be used to generate prevalence estimates. 7. Pg 7; by presenting Table 2ii, the authors hint at a time effect on prevalence but do not mention the results in the Results text. Given that this is mentioned in the Discussion, a line in the Results is appropriate. Some indication of the precision of the estimates is necessary, as is comment on whether this analysis was pre-planned or post-hoc. 8. Pg 9, 1st paragraph; it would be worth confirming here that all data reported in this section refers to the population prevalence of moderate-severe pain, as opposed to the proportion of people with pain who report moderate-severe symptoms. 9. Pg 9; some indication of the precision of the prevalence estimates in the age strata are necessary, here and/or in the Figure. I realise that that no pooled estimate is calculated but confidence intervals around the individual studies would make the data more interpretable. 10. Table 1; what does '(corrected)' with respect to the Response rate mean? 11. Table 2i; Given that Smith et al 2004 reported data only from women, I question the entry in the 'Prevalence total' column, I think this cell would be better left blank. 12. Table 3; Similar to the comment immediately previous, I question the inclusion of data from Croft 203 in the 'Prevalence total' column, and more importantly in the meta-analysis (Figure 2).

	Discussion 13. Pg 15, 1st paragraph; regarding comparison of the pooled prevalence estimate to the Pain in Europe survey data. Is it not the case that the 13% reported in the Pain in Europe survey was for moderate-severe pain? If so, this seems to align quite well with the estimates for more intense pain reported in Table 4. 14. Pg 15, 1st paragraph; as mentioned previously, CIs around the time-stratified estimates are necessary to interpret these data. 15. The Discussion would benefit from an extra paragraph placing the findings in the context of other studies, for example pain prevalence estimates globally and from other parts of the world. 16. Pg 15, 2nd paragraph; while the issue of heterogeneity in the meta-analyses is mentioned, the limitation associated with pooling heterogeneous estimates is not addressed directly. Given the level of heterogeneity reported, a reader may ask whether meta-analysis is warranted at all. 17. Pg 15, 2nd paragraph; further to the issue of heterogeneity, the authors offer two potential reasons both of which could be investigated quite easily ie. time, and geographical area. 18. Pg 15, 1st paragraph; the authors mention the criticism of prevalence estimates in the absence of information about the impact/severity of the condition. This is followed by a statement that the prevalence is high in younger people, the criticism is not addressed. 19. Pg 15, 1st paragraph; the last sentence of this paragraph is somewhat of a diversion, especially given reference to the comment in the following paragraph regarding the need for care. I recommend deletion. 20. Similar to the Introduction that lacks strong justification, the Discussion lacks a paragraph that speaks to the research/clinical/policy implications of this study. Such a paragraph at the end of the Discussion would place the study into context and leave the reader with a strong sense of its value.

REVIEWER	J. Berkhof VUMC, Netherlands
REVIEW RETURNED	13-Jan-2016

GENERAL COMMENTS	This systematic review is in general well described. One remark is that the authors can be more explicit about the criterion used to exclude a study on the basis of high risk of bias. Furthermore, there are a few issues with regard to the interpretation. • This study suggests that UK pain prevalence figures are higher than figures from European telephone surveys (Discussion p.15 line 1-5). However, the European telephone survey figures are similar to those reported in Table 4. • The selected studies are very heterogeneous and pooled meta-analytical estimates are of limited value. It is much more informative to gain insight into the factors that contribute to the heterogeneity. I like figure 3, but gender, year of publication, severity, and mode of data collection play an important role as well. It would be useful to display multiple factors in one graph and/or to carry out a multiple regression analysis to assess the role of individual factors. Please check errors, like e.g. p.8.l.15. “Prevalence estimates were again higher in female (9.0% to 10.4%) than in male participants (12.3% to 17.9%)”.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Reviewer Name
Kathryn Mansfield

Institution and Country
London School of Hygiene and Tropical Medicine
United Kingdom

Please state any competing interests or state 'None declared':
I have written a systematic review on a similar topic (CWP prevalence in the general population).
I have previously worked with one of the authors (Peter Croft).

Please leave your comments for the authors below
This paper reports the results of a systematic review and meta-analysis of the prevalence of chronic pain in the UK general population. Results are stratified using the following pain phenotypes: i) chronic pain; ii) fibromyalgia; iii) chronic widespread pain (CWP); and iv) chronic neuropathic pain. Search results allowed meta-analysis of prevalence CWP and the more generic chronic pain phenotype. The authors considered variability in pain prevalence by: age; gender; pain severity; and year of publication (depending on availability of data for each pain phenotype).

I feel there are a few revisions that would benefit your manuscript:

1. Exclusion of studies at high risk of bias
I understand your decision to exclude studies at high risk of bias from your meta-analyses. However, given that risk of bias assessment is a subjective exercise I would argue that it is inappropriate to exclude these studies entirely from your review. I feel you should at least discuss the high risk studies narratively and the reasons why you felt it necessary to drop them from the meta-analysis rather than simply excluding them entirely (I note that the excluded studies are named in the table in Appendix D, however full references are not given for these studies making it difficult for the reader to identify the studies and confirm whether they agree with your decision to exclude them)…

We appreciate the point being made and the reviewer's recent paper illustrates well the value of the approach being suggested, but this is about a design decision as to how the quality assessment is going to be used. We accept this is a matter for debate and have added this acceptance to our discussion. However, for the purposes of our review (to attempt to identify a reasonable prevalence estimate of chronic pain in the UK), we chose a priori that low quality studies would be excluded. We would argue that this is one defensible way of deciding which prevalence estimate to actually use in practice and would prefer to stick with this a priori decision.

However we can understand the confusion if papers are 'included' up to the end of quality assessment (part of the review) and then are 'excluded'. So we are following the helpful advice of the reviewer to:

(a) Clarify that the 6 papers were excluded from the review – but obviously at a later stage than other exclusions
(b) Provide more information on the papers excluded by including full references

… A more robust approach would be to present an overall prevalence estimate (including the high risk

studies) with follow-up sensitivity analyses first excluding high risk and (then if numbers allow) moderate risk studies, and possible follow-up with a meta-regression to explore the effect of risk of bias. The emphasis in the text would clearly be on the results from the low and moderate risk of bias studies in order not to confuse the reader.

Following on our response to the previous point, we have conducted a sensitivity analysis with the high-risk studies included and generated a new pooled estimate of 39.1% (95% CI 31.4 - 46.8), at the expense of an increase in heterogeneity (I2 99.8%). We have however stuck with our a priori plan (see first point) to exclude studies deemed to be high-risk-of-bias in order to base our summary estimates on better quality studies.

2. Limitations
Any presentation of study limitations in the paper (Article summary and Discussion sections) fails to acknowledge that the search strategy may have missed important articles. Having recently completed a similar study I am aware of at least two studies of chronic widespread pain (CWP) prevalence in the UK that you do not appear to have considered for inclusion in your review [1,2].

We have added a sentence that acknowledges this point, and used the opportunity to reference the referee's own recent meta-analysis of chronic widespread pain prevalence globally (not available when we first submitted our own review) as evidence from another source that papers were not missed.

(We are assuming that the studies referred to above were part of that meta-analysis and we have gone through the reference list in that paper. We cannot identify any new articles that would have matched our inclusion criteria: Papageorgiou et al. (retrieved in our searches but excluded as data duplicated another analysis already included in review), 32 Croft et al.13 (included), Carnes et al.10 (not in our searches but focus on MSK pain therefore did not match our search criteria), Choudhury et al.11 (retrieved in our searches but excluded as not deemed 'general population'), Hunt et al.19,26 (retrieved in our searches but excluded as not ACR definition of FM), Macfarlane et al.25 (retrieved in our searches but excluded as not deemed 'general population') Aggarwal et al.2 (included).

Other than a reference in your methods section regarding approaching 'field experts' to identify additional grey literature studies, I have not been able to find any reference to the possibility of publication bias in your paper.

We are not altogether sure how publication bias would be identified in studies reporting prevalence estimates. We have therefore spent time preferentially focusing on methodological and reporting bias in the review.

3. Chronic widespread pain meta-analysis
In your methods section you state estimates that "were not restricted to age-specific or gender-specific cohorts were incorporated into a meta-analysis". However, Figure 2 includes a CWP prevalence estimate from the Croft et al. 2003 study; since this study is restricted to women only I wonder if it should have been included in the meta-analysis.

Thank you, this has been removed from the meta-analysis.

4. Consideration of possible reasons for differences in study estimates
It would have been helpful to see some attempt to tie up the prevalence estimates with each study's methods (risk of bias assessment). For example, might some aspects of study design account for any particularly low or high estimates? You have considered the year the study was undertaken as one reason, but what about the age/sex distribution of the study populations, or the way the data was

collected? I accept this might be difficult to do given the limited number of studies you have, but simply stating the range of prevalence estimates found does not do justice to the amount of work you have done to extract all the relevant information about each study.

Thank you. Unfortunately there were not enough tangible variables amongst the studies remaining in the meta-analysis to allow for regression; only one study was deemed to be low-risk of bias; the gender and age distributions did not really vary enough amongst the included studies to justify different categories; nor was there significant variability in survey methodology. This may reflect our fairly strict inclusion criteria.

However, in response to this and similar points raised by the other two referees, we have now realigned the descriptions of year and location of studies as stratified explorations of between-study variation, and presented both figuratively as well as narratively.

5. High I-squared statistic
Could you please comment on the accuracy of the pooled prevalence estimates for the various pain phenotypes in light of the high level of heterogeneity demonstrated by the I-squared statistic (98.9% for chronic pain, and 95.2% for CWP). Both in your discussion and in the article summary you state that the studies were 'homogenous'. I think you can state that prevalence estimates were consistent across studies and perhaps pass some comment regarding similar/dissimilar methodological approaches taken by the studies included; however, I think stating that the studies were 'homogenous' in light of the very high I-squared statistic is over-stating your findings.

Thank you – the use of homogenous in the article summary was an error and this has in fact been changed to heterogeneous! We agree that further use of this term in the discussion is misleading and we have therefore amended the text in line with your recommendations. Confidence intervals for pooled estimates have been added.

6. Forest plots
To avoid having to cross-reference between the forest plots and relevant text/tables, I would prefer to see all prevalence estimates and their 95% CIs (particularly those for pooled estimates) presented in text alongside the relevant graphical elements of the forest plot. I would also like to see I-squared statistics presented in the figure legends.

Forest Plots using STATA have now been presented in order to achieve a more conventional format as requested

7. Nested case-control studies
Can you confirm that the prevalence estimates included from the three nested case-control studies were derived from figures from the cohort in which they were nested, as prevalence figures can only be derived from case-control studies in very specific circumstances. If this is the case, could you note this in some way in Table 1, as I think many readers will question whether it is appropriate to derive prevalence estimates from case-control studies.

Thank you – this issue has been addressed in the tables and in the methods section.

8. Figure 3: Chronic pain prevalence by age
I found the x-axis on this figure somewhat misleading. I understand that it is difficult to synthesize this type of data as different papers use different age bands. I think you have taken the mid-point of the age bands presented in each paper for the x-axis figures. I do not think that presenting categorical data as continuous data in this way is really appropriate. I am also not comfortable with your choice x-axis title (median age) as this does not seem to really capture the figures that you are using. Having

recently grappled with this problem myself, one approach might be to create one figure containing five panels – each panel would be a graph representing the data from a paper, this would allow you to use the age bands from each paper appropriately on the x-axis.

Thank you we have re-presented the data as suggested.

9. Tables 2i and 2ii
Table 2i presents studies of chronic pain prevalence ordered by study size and Table 2ii the chronic pain prevalence studies included in the meta-analysis stratified by publication date. It was difficult to get a sense of these studies with the results repeated and distributed across the two tables. Perhaps, rather than ordering by sample size, you could consider creating one table stratified by the following categories of study: i) studies excluded from the meta-analysis; ii) studies published between 1990 and 2005; iii) studies published between 2006 and 2009; and iv) studies published between 2010 and 2015.

As suggested. Table 2ii has removed with the data and CI presented in the results text.

In the sixth column of Table 2i, the Mallen et al. 2005 and Croft et al. 1993 papers are missing the number of pain patients (% figures are given).

'n' for Mallen has been added – the 'n' was not provided in the paper by Croft et al 1993 and has therefore purposely been omitted from the paper

10. Table 3
Perhaps you could consider stratifying this table by whether the study was included in the meta-analysis, e.g. i) Studies restricted by age; and ii) Studies included in meta-analysis.

As suggested

11. Chronic widespread pain results - page 8
At the end of the first paragraph on this page you write: "Prevalence estimates were again higher in female (9.0% to 10.4%) than in male participants (12.3% to 17.9%)". I think you must have transposed the figures for women with those for men.

Yes, thank you – this has been amended.

12. Fibromyalgia results - page 8
On the third paragraph on this page you write: "Due to the practical restrictions of formally diagnosing a patient with fibromyalgia (requiring a history and examination in order to exclude alternative causes for widespread pain 11), only one study [REFERENCE] was able to provide comprehensive data from populations representative of the general population."
Please could you include a reference to the study that you are discussing in the appropriate place.

Thank you, the reference had been placed after presentation of the results further on in the paragraph but we have included it where suggested.

13. Mean prevalence
You make reference to 'mean prevalence' in your abstract. Could you please stick to 'pooled prevalence' for clarity.

This has been corrected. Thank you.

Reviewer: 2

Reviewer Name
Steven Kamper

Institution and Country
The George Institute, University of Sydney, Australia

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

Please leave your comments for the authors below
General comment
• The authors report on a systematic review of the prevalence of chronic pain in the UK. The authors can be congratulated on a well-conducted study and a well-written manuscript, the content is appropriate for the journal readership. I have no major concerns with the study but offer a few suggestions and have some small queries that I would appreciate being addressed.

Specific comments
Introduction
1. The Introduction is very brief, which I understand conforms to the style of the journal, but I think comes at the expense of a solid justification for the study. The section would benefit from addition of a few sentences outlining why this study is 'much needed', providing some context in terms of the need from a research, clinical and/or policy perspective.

Agreed, we have edited the introduction in order to address this

Methods
2. Pg 5, 2nd paragraph; on my first reading, inclusion criterion (i) and the last sentence of the paragraph appeared to contradict one another. After some effort I realise that they don't but perhaps rewording of the last sentence would improve clarity.

Yes – we appreciate the issue. I hope this has been resolved by adding the words 'only' following the site-specific location.

3. I miss specification of a prevalence period. In the Results it seems point, and 1-month prevalence are combined, I have no issue with this but it is worth setting out whether this decision made a-priori, and whether other periods were excluded.

Thank you – we have amended the methods section to highlight that point and period estimates were allowed – there was no restriction on time period for estimates.

4. How were studies assessed as 'High' risk of bias using the tool?

This is described in the text "The tool consists of ten items addressing the external validity (risk of selection and nonresponse bias), as well as the internal validity (risk of measurement bias, and bias related to the data analysis) of observational studies in order to generate an overall risk of bias assessment". The breakdowns are also included under the Appendices. We have now also included references for articles excluded on basis of quality assessment. .

5. Were primary study authors contacted in the cases of unreported data?

No, we did not contact authors directly – this has now been iterated in the methods section

Results
6. Some studies used GP practice data, it might be worth a sentence explaining how these data can be used to generate prevalence estimates.

The GP consultation data were not used as such. Only population registers from GP practices, regardless of any consultation, were used to identify potential survey populations. A sentence has been added in the methods to clarify.

7. Pg 7; by presenting Table 2ii, the authors hint at a time effect on prevalence but do not mention the results in the Results text. Given that this is mentioned in the Discussion, a line in the Results is appropriate. Some indication of the precision of the estimates is necessary, as is comment on whether this analysis was pre-planned or post-hoc.

This has now been included in the results section.

8. Pg 9, 1st paragraph; it would be worth confirming here that all data reported in this section refers to the population prevalence of moderate-severe pain, as opposed to the proportion of people with pain who report moderate-severe symptoms.

The text has been amended to highlight that these estimates are from the total population. Thank you

9. Pg 9; some indication of the precision of the prevalence estimates in the age strata are necessary, here and/or in the Figure. I realise that that no pooled estimate is calculated but confidence intervals around the individual studies would make the data more interpretable.

Agreed, but unfortunately CI were only available for a few of the studies reporting age-specific data. We felt that only including CI for a subset of the studies might be confusing to the reader, but we can add these to the graphs where available if the editing team would prefer this.

10. Table 1; what does '(corrected)' with respect to the Response rate mean?

Corrected after excluding individuals on the survey list who did not actually receive the survey (due to death or change in address) or were unable to be completed due to illness or learning disability. A detail has been added to the methods section describing this.

11. Table 2i; Given that Smith et al 2004 reported data only from women, I question the entry in the 'Prevalence total' column, I think this cell would be better left blank.

The table has been amended.

12. Table 3; Similar to the comment immediately previous, I question the inclusion of data from Croft 203 in the 'Prevalence total' column, and more importantly in the meta-analysis (Figure 2).

This study has been removed from the meta-analysis and the tables amended. As have the figures presented in the abstract/results/discussion.

Discussion
13. Pg 15, 1st paragraph; regarding comparison of the pooled prevalence estimate to the Pain in Europe survey data. Is it not the case that the 13% reported in the Pain in Europe survey was for

moderate-severe pain? If so, this seems to align quite well with the estimates for more intense pain reported in Table 4.

Agreed and we also agree that this point was unclear in our discussion. We have therefore clarified this further in the discussion and alluded to the similarity between summary estimates from severe pain from our study and those drawn from the pain in Europe study.

14. Pg 15, 1st paragraph; as mentioned previously, CIs around the time-stratified estimates are necessary to interpret these data.

These have now been included

15. The Discussion would benefit from an extra paragraph placing the findings in the context of other studies, for example pain prevalence estimates globally and from other parts of the world.

Thank you for raising this. We have linked this to your points 19 and 20 below and to referee 3 on the issue of the wider objectives of this study and added material on this to both the introduction and discussion.

16. Pg 15, 2nd paragraph; while the issue of heterogeneity in the meta-analyses is mentioned, the limitation associated with pooling heterogeneous estimates is not addressed directly. Given the level of heterogeneity reported, a reader may ask whether meta-analysis is warranted at all.

This we accept is a difficult issue, but the reality is that we either reject any idea of 'getting' an average and then have no way of summarising what might be a plausible prevalence, or we go ahead on the basis that any summary of studies that are of reasonable quality is going to be more useful than a range. Your point 15 and referee 1's recent CWP review are helpful here in allowing us to highlight the plausibility of the summary estimates.

17. Pg 15, 2nd paragraph; further to the issue of heterogeneity, the authors offer two potential reasons both of which could be investigated quite easily ie. time, and geographical area.

Please also refer to response to Comment 4, Reviewer 1.
We have added an analysis by dates, having pooled results over three time periods, as well as geographical location.

18. Pg 15, 1st paragraph; the authors mention the criticism of prevalence estimates in the absence of information about the impact/severity of the condition. This is followed by a statement that the prevalence is high in younger people, the criticism is not addressed.

Thank you – the point we had hoped to make was that chronic non-disabling conditions alone could not explain the high prevalence estimates (as evidenced by the fairly high figures from younger population groups). We have edited the text to better reflect our original intention.

19. Pg 15, 1st paragraph; the last sentence of this paragraph is somewhat of a diversion, especially given reference to the comment in the following paragraph regarding the need for care. I recommend deletion.

This has now been deleted but a more general point about the value of prevalence estimates for policy decisions made in response to referee 3 and to your earlier point 15 and your next point 20.

20. Similar to the Introduction that lacks strong justification, the Discussion lacks a paragraph that

speaks to the research/clinical/policy implications of this study. Such a paragraph at the end of the Discussion would place the study into context and leave the reader with a strong sense of its value.

Agreed and we have attempted to address this.

Reviewer: 3

Reviewer Name
J. Berkhof

Institution and Country
VUMC, Netherlands

This systematic review is in general well described. One remark is that the authors can be more explicit about the criterion used to exclude a study on the basis of high risk of bias. Furthermore, there are a few issues with regard to the interpretation.

• This study suggests that UK pain prevalence figures are higher than figures from European telephone surveys (Discussion p.15 line 1-5). However, the European telephone survey figures are similar to those reported in Table 4.

Yes in fact the telephone survey looks specifically at moderate to severe pain and so would be expected to mirror these figures – however these estimates have also been used to describe the burden of chronic pain in the UK (not just severe pain) which was why the parallel was drawn in the first paragraph of the discussion. We have clarified and elaborated further on this in the discussion.

• The selected studies are very heterogeneous and pooled meta-analytical estimates are of limited value. It is much more informative to gain insight into the factors that contribute to the heterogeneity. I like figure 3, but gender, year of publication, severity, and mode of data collection play an important role as well. It would be useful to display multiple factors in one graph and/or to carry out a multiple regression analysis to assess the role of individual factors.

Please also refer to response to Comment 4, Reviewer 1.
We have added an analysis by dates, having pooled results over three time periods, as well as geographical location.

Please check errors, like e.g. p.8.l.15. “Prevalence estimates were again higher in female (9.0% to 10.4%) than in male participants (12.3% to 17.9%)”.

Amended – thank you

VERSION 2 – REVIEW

REVIEWER	Kathryn Mansfield London School of Hygiene and Tropical Medicine, UK I have written a systematic review on a similar topic (CWP prevalence in the general population). I have previously worked with one of the authors (Peter Croft).
REVIEW RETURNED	24-Feb-2016

GENERAL COMMENTS	The authors have addressed all reviewers’ comments on the

	first submission. This is a well-conducted study and a clearly written manuscript. However, I still have a couple of very minor comments. 1. Chronic pain - sources of variation In the discussion (paragraph 4) the authors state that their findings suggest calendar year and geographic region as potential sources of systematic variation between studies. However, review of the forest plots in Figures 2 and 3 show that this is based on comparison of pooled estimates with high heterogeneity and wide overlapping confidence intervals (with only 2 to 3 studies in each category). I appreciate that the authors are restricted by the small number of eligible studies, however, I think that conclusions based on these results need to be moderated a little; they should perhaps be presented with a word of caution highlighting the limited number of studies and the broad and overlapping CIs. 2. Figures Figures 1 and 2 present much the same information. Figure 1 is a forest plot showing prevalence estimates from studies reporting chronic pain prevalence. Figure 2 presents the same data stratified by date of publication. Perhaps the authors might consider deleting Figure 1 as it seems somewhat redundant. All forest plots need a label on the x-axis. Figure 3 needs a label on the y-axis. 3. Table 1 In his comments (10) Reviewer 2 asked, “Table 1; what does ‘(corrected)’ with respect to Response rate mean?’. I note that the authors have addressed this with additional text in the Methods section of the paper, however, I feel the Table would also benefit from a footnote explaining it.

REVIEWER	Steven Kamper The George Institute, University of Sydney
REVIEW RETURNED	18-Feb-2016

GENERAL COMMENTS	I thank the authors for their efforts and additional work in addressing my comments. I am satisfied with the manuscript as it stands.

REVIEWER	J.Berkhof VUMC, The Netherlands
REVIEW RETURNED	02-Mar-2016

GENERAL COMMENTS	The authors might consider adding the response to reviewer 1 Q4 to the manuscript: "the gender and age distributions did not really vary enough amongst the included studies to justify different categories; nor was there significant variability in survey methodology".

VERSION 2 – AUTHOR RESPONSE

Reviewer: 1

Reviewer Name
Kathryn Mansfield

Institution and Country
London School of Hygiene and Tropical Medicine, UK

Please state any competing interests or state 'None declared':
I have written a systematic review on a similar topic (CWP prevalence in the general population).
I have previously worked with one of the authors (Peter Croft).

Please leave your comments for the authors below
The authors have addressed all reviewers' comments on the first submission. This is a well-conducted study and a clearly written manuscript. However, I still have a couple of very minor comments.

1. Chronic pain - sources of variation
In the discussion (paragraph 4) the authors state that their findings suggest calendar year and geographic region as potential sources of systematic variation between studies. However, review of the forest plots in Figures 2 and 3 show that this is based on comparison of pooled estimates with high heterogeneity and wide overlapping confidence intervals (with only 2 to 3 studies in each category). I appreciate that the authors are restricted by the small number of eligible studies, however, I think that conclusions based on these results need to be moderated a little; they should perhaps be presented with a word of caution highlighting the limited number of studies and the broad and overlapping CIs.

Thank you, we have moderated our conclusions in keeping with your suggestion. Discussion: Page16/17

2. Figures
Figures 1 and 2 present much the same information. Figure 1 is a forest plot showing prevalence estimates from studies reporting chronic pain prevalence. Figure 2 presents the same data stratified by date of publication. Perhaps the authors might consider deleting Figure 1 as it seems somewhat redundant.

Removed

All forest plots need a label on the x-axis. Figure 3 needs a label on the y-axis.

Done

3. Table 1
In his comments (10) Reviewer 2 asked, "Table 1; what does '(corrected)' with respect to Response rate mean?'. I note that the authors have addressed this with additional text in the Methods section of the paper, however, I feel the Table would also benefit from a footnote explaining it.

Done

Reviewer: 2

Reviewer Name
Steven Kamper

Institution and Country
The George Institute, University of Sydney

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

Please leave your comments for the authors below
I thank the authors for their efforts and additional work in addressing my comments. I am satisfied with the manuscript as it stands.

Many thanks

Reviewer: 3

Reviewer Name
J.Berkhof

Institution and Country
VUMC, The Netherlands

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

Please leave your comments for the authors below
The authors might consider adding the response to reviewer 1 Q4 to the manuscript:

"the gender and age distributions did not really vary
enough amongst the included studies to justify different categories; nor was there
significant variability in survey methodology".

Thank you, this has been added to the Results section (Page 7) and is re-iterated in the Discussion (Page 17)