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)	Association between physical activity and musculoskeletal pain: an analysis of international data from the ASAP survey
AUTHORS	Rhim, Hye Chang; Tenforde, Adam; Mohr, Lisa; Hollander, Karsten; Vogt, Lutz; Groneberg, David; Wilke, Jan

VERSION 1 – REVIEW

REVIEWER	Smith, Benjamin Derby Teaching Hospitals NHS Foundation Trust, Physiotherapy Outpatients
REVIEW RETURNED	11-Jan-2022

GENERAL COMMENTS	Thank you for asking me to review this paper about physical activity and musculoskeletal pain. This is an important topic and a worthwhile project. Overall, this is an important study, and worthwhile conducting and publishing. My main concern is the language used in the manuscript. The title and aims of the study make it clear that you're looking for an association. And study design doesn't allow for causation to be established – you mention this in the strengths and limitations section. However, I did find the wording used in many parts of the manuscripts edging towards implying causation. Despite this, the underlying premise and the constant thread of your manuscript reads to me clearly about causation. I have documented these below, with a few other comments/questions. Abstract: I find with the word 'optimal' in your conclusion edges towards inferring causation. It may be people with MSK pain are less able to keep physically active. Rather than low physical activity increasing the risk of MSK pain. Background aim. Again, I find your hypothesis is written with a hint of causation in mind. It may be that PA reduces overall MSK pain. But we can't rule out, MSK pain being the driver that reduces PA. ?? Method: Where did the body locations come from? Looks to be some overlap to me, e.g. shoulder and upper arm; glutes and hip. Patient and Public Involvement: How many patients were involved?

	Discussion: Page 10 line 3. “Simple guideline compliance (150-300 min per week) was weakly associated with MSK pain, showing lower odds of developing pain only in thoracic spine but higher odds in foot/toes.”. This sentence, in the way it is written, implies causation to me. This mainly due to the word “developing”. Developing implies the MSK pain came after the lower PA levels. Perhaps a better word would be “having”. You used the word “developing” a few times in the discussion and I would advise changing it. “Beneficial effects of PA” page 11 line 19. Again – this implies causation to me. We simply do not know if the lower MSK pain is because of the effect of PA. An RCT is the only method of knowing for sure. The limitation section is well written and covers the main limitations well. Conclusion: “Prevent” and “beneficial” both imply causation to me. Why would someone with degenerative changes need to be cautious? Seems an odd throwaway comment to make that is not in line with the aims of the paper. This is the first mention of degenerative changes in this paper – I suggest removing it.

REVIEWER	Niederstrasser, Nils University of Portsmouth, Psychology
REVIEW RETURNED	03-Mar-2022

GENERAL COMMENTS	Thank you for the opportunity to review this interesting paper by Rhim, Tenforde, and colleagues examining the relationship between physical activity and pain in various anatomical locations. The authors collected data from more than 14,000 participants in 14 countries, which is impressive. Nevertheless, the following issues need to be addressed prior to publication: Abstract: Please mention the total number of participants. There seems to be a double mention of “questionnaire” here: PA volumes were assessed with an adapted version of the Nordic physical activity questionnaire-short questionnaire. Please rephrase the conclusion to avoid suggesting any directionality of the effects considering this is a cross-sectional investigation. Intro How do the authors explain the vast difference in MSK prevalence between the studies referenced in the second and third sentence of the introduction? The two seem contradictory of each other but just show the vast range of estimates. This should be made clearer. I think it should also be mentioned, in line with arguments made later in the paragraph, that the prevalence of pain changes with age. Could the authors please elaborate on why MSK impairments

	contribute to functional limitations particularly in developed countries? At the end of the first paragraph, do the authors suggest that the increase in DALYs is due to an increase in MSK pain? If so, this should be made clear, and reasons should be provided. I disagree with the notion that there is a substantial literature gap when it comes to studies investigating the relationship between pain and physical activity in general. The rationale for the distinction between lower and upper extremity is not convincingly made. Similarly, the hypotheses are only partially supported by the introduction. Also, the research does not actually examine differences between upper and lower extremities but rather looks at the association of varying doses of PA and pain in a variety of anatomical locations. The distinction feels rather arbitrary. Methods: In the participants section, please mention the number of participants and a breakdown of participants per country with relevant demographic information and physical activity levels. Please could you provide more detail on the NPAQ-Short? How is it scored and what are the response options? I am concerned about the questionnaire used to measure the prevalence of MSK. It appears that no distinction is made as to which side of the body pain occurs. Therefore, pain in the knee is scored the same, regardless of whether pain is experienced in one or both knees. Could the authors also provide details on how this checklist was adapted for the study? In the data processing section when you describe the formula, should it be: moderate PA – vigorous PA) + vigorous PA *2? Also, please ensure it is clear you are referring to minutes of physical activity. I have several concerns about the adequacy and level of detail provided for the analysis. First, what is the reasoning behind categorising what are potentially continuous data into compliers, double compliers, etc.? If PA is reported in minutes and by intensity, which is not clear from the paper at this point, would it not make more sense to keep data continuous? Second, since for each body region a separate analysis was run, it would seem necessary to correct for multiple testing. Third, how were the confounding variables recorded? Finally, please can you clarify if all confounding variables were included in all multivariate logistic regressions and report the results of these in full? Could you comment on how all materials were translated into the various languages and how equivalency of translations was ensured? Results: Please, could you break down results by country?

	For participants, please could you report mean age with SD? What were the maximum and minimum ages? Please could you report full results from all analyses, i.e., beta, standard error, and ORs (with CI) for PA and confounding variables and p-values? Discussion: Please could you discuss more on potential pathways specifically between PA and upper and lower extremity pain? Please can you comment on the sample's mean age? Overall, the discussion is too descriptive. Other than being mentioned on page 9, what is the relevance of Table 2? How were the findings derived? Why are they not discussed?

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Mr. Benjamin Smith, Derby Teaching Hospitals NHS Foundation Trust, University of Nottingham School of Medicine
Comments to the Author:
Thank you for asking me to review this paper about physical activity and musculoskeletal pain. This is an important topic and a worthwhile project.

Overall, this is an important study, and worthwhile conducting and publishing.

My main concern is the language used in the manuscript. The title and aims of the study make it clear that you're looking for an association. And study design doesn't allow for causation to be established – you mention this in the strengths and limitations section. However, I did find the wording used in many parts of the manuscripts edging towards implying causation. Despite this, the underlying premise and the constant thread of your manuscript reads to me clearly about causation. I have documented these below, with a few other comments/questions.

Authors' Response: As you recommended, we tried to avoid the words that might suggest causation throughout our manuscript.

Abstract:
I find with the word 'optimal' in your conclusion edges towards inferring causation. It may be people with MSK pain are less able to keep physically active. Rather than low physical activity increasing the risk of MSK pain.

Authors' Response: In order to avoid language that might suggest causation as you pointed out, we revised our conclusion in the abstract: "A dose of 300-450 min WHO-equivalent PA/week was associated with reduced MSK-pain. On the other hand, excessive doses of PA were associated with increased pain in certain body locations."

Background aim.
Again, I find your hypothesis is written with a hint of causation in mind. It may be that PA reduces overall MSK pain. But we can't rule out, MSK pain being the driver that reduces PA??

Authors' Response: In order to avoid a hint of causation, we revised our hypothesis to "We hypothesize that greater time spent in PA than WHO recommendation would be associated with reduction of MSK-pain, but excess time performing PA might be associated with higher MSK-pain".

Method:
Where did the body locations come from? Looks to be some overlap to me, e.g. shoulder and upper arm; glutes and hip.

Authors' Response: The body locations derived from the consensus statement (Bahr R, Clarsen B, Derman W, et al. International Olympic Committee consensus statement: methods for recording and reporting of epidemiological data on injury and illness in sports 2020 (including the STROBE extension for sports injury and illness surveillance (STROBE-SIIS)). The number of categories was slightly reduced when compared against the cited reference.

Patient and Public Involvement:
How many patients were involved?

Authors' Response: The questionnaire was face validated for each language with five non-academic individuals. Feedback on comprehension and clarity of the wording was used. This has now been described in the Methods section.

Discussion:
Page 10 line 3. "Simple guideline compliance (150-300 min per week) was weakly associated with MSK pain, showing lower odds of developing pain only in thoracic spine but higher odds in foot/toes.". This sentence, in the way it is written, implies causation to me. This mainly due to the word "developing". Developing implies the MSK pain came after the lower PA levels. Perhaps a better word would be "having". You used the word "developing" a few times in the discussion and I would advise changing it.

Authors' Response: Following your advice, we changed the word "developing" to "having" throughout discussion.

"Beneficial effects of PA" page 11 line 19. Again – this implies causation to me. We simply do not know if the lower MSK pain is because of the effect of PA. An RCT is the only method of knowing for sure.

Authors' Response: In order to avoid implying causation, we revised the phrase into "PA in the range of 300-600 min was also associated with lower odds of having pain in several locations…"

The limitation section is well written and covers the main limitations well.

Conclusion:
"Prevent" and "beneficial" both imply causation to me.

Authors' Response: We deleted "prevent" and "beneficial" and rephrased the conclusion.

Why would someone with degenerative changes need to be cautious? Seems an odd throwaway comment to make that is not in line with the aims of the paper. This is the first mention of degenerative changes in this paper – I suggest removing it.

Authors' Response: As you suggested, we removed this part from our conclusion.

Reviewer: 2
Dr. Nils Niederstrasser, University of Portsmouth
Comments to the Author:

Thank you for the opportunity to review this interesting paper by Rhim, Tenforde, and colleagues examining the relationship between physical activity and pain in various anatomical locations. The authors collected data from more than 14,000 participants in 14 countries, which is impressive. Nevertheless, the following issues need to be addressed prior to publication:

Authors' Response: We really appreciate your time and effort to provide us with meaningful comments. We tried to address the issues you were concerned with. Please see below for our response to your comments.

Abstract:
Please mention the total number of participants.

Authors' Response: A total of 13,741 participants completed the survey, and this was added to the abstract.

There seems to be a double mention of "questionnaire" here: PA volumes were assessed with an adapted version of the Nordic physical activity questionnaire-short questionnaire.

Authors' Response: We changed it to "Nordic Physical Activity Questionnaire-short"

Please rephrase the conclusion to avoid suggesting any directionality of the effects considering this is a cross-sectional investigation.

Authors' Response: We rephrased our conclusion in the abstract to avoid any suggestion of causality: "A dose of 300-450 min WHO-equivalent PA/week was associated with reduced MSK-pain. On the other hand, excessive doses of PA were associated with increased pain in certain body locations."

Intro

How do the authors explain the vast difference in MSK prevalence between the studies referenced in the second and third sentence of the intoduction? The two seem contradictory of each other but just show the vast range of estimates. This should be made clearer. I think it should also be mentioned, in line with arguments made later in the paragraph, that the prevalence of pain changes with age.

Authors' Response: The difference in prevalence of MSK-pain may derive from the diverse research methods and populations across the studies. As you pointed out, providing multiple numbers may be contradictory to each other and confusing to the readers, so we provided ranges instead and mentioned that the prevalence of pain increases with age.

Could the authors please elaborate on why MSK impairments contribute to functional limitations particularly in developed countries?

Authors' Response: The article we cited (Cimmino MA, Ferrone C, Cutolo M. Epidemiology of chronic musculoskeletal pain. *Best Pract Res Clin Rheumatol* 2011;25(2):173-83.) summarized epidemiological studies on chronic musculoskeletal pain, and most of these studies were conducted in the developed countries. Therefore, the authors of that article used the phrase, "in most developed countries" to refer to countries where the studies have taken place rather than trying to make distinction between developed and developing countries. This sentence can be confusing and may not even be relevant to our paper, so we deleted the following sentence: "Musculoskeletal impairments may contribute to functional limitations particularly in developed countries."

At the end of the first paragraph, do the authors suggest that the increase in DALYs is due to an increase in MSK pain? If so, this should be made clear, and reasons should be provided.

Authors' Response: In order to be clearer, we added the following sentence, "Most of the increased burden has derived from disability due to increased aging population affected by MSK conditions"

I disagree with the notion that there is a substantial literature gap when it comes to studies investigating the relationship between pain and physical activity in general.

Authors' Response: We agree that there are studies investigating the relationship between PA and pain in general. We meant to say, compared to studies investigating the association of PA with non-communicable disease, there are less. We revised the original sentence to "However, compared to the number of studies investigating the association of PA with non-communicable disease, there seems to be a literature gap regarding MSK-pain"

The rationale for the distinction between lower and upper extremity is not convincingly made. Similarly, the hypotheses are only partially supported by the introduction. Also, the research does not actually examine differences between upper and lower extremities but rather looks at the association of varying doses of PA and pain in a variety of anatomical locations. The distinction feels rather arbitrary.

Authors' Response: In order to address your concerns, we changed the final paragraph of the Introduction: "The purpose of this study was to investigate the association of total PA with MSK-pain in a variety of anatomical locations including both upper and lower extremities. We hypothesize that greater time spent in PA than WHO recommendation would be associated with reduction of MSK-pain, but excess time performing PA might be associated with higher MSK-pain." As we mentioned in the earlier paragraph that it is unclear whether the amounts of PA recommended by WHO would be sufficient to elicit benefits in terms of addressing MSK-pain, our interest was to explore if there was certain time that would be associated reduction of MSK-pain, but excess amount would rather be associated with more MSK-pain.

Methods:

In the participants section, please mention the number of participants and a breakdown of participants per country with relevant demographic information and physical activity levels.

Authors' Response: Thank you for the suggestion. This information is now provided as Table 1 in the Results section.

Please could you provide more detail on the NPAQ-Short? How is it scored and what are the response options?

Authors' Response: We provided details on the NPAQ-Short in our Methods section: "The instrument retrospectively assessed the amounts of moderate and combined moderate and vigorous activities (min/week) during leisure and occupational time. Moderate activities were defined as those that increase heart rate or breathing, and vigorous activities were defined as those that make heart racing, sweating, and shortness of breath. The questionnaire asked how much time participants spent in total on both moderate and vigorous PA on a typical week, and the time spent in all activities with a minimal duration of 10 minutes was asked to be added and entered in the form."

I am concerned about the questionnaire used to measure the prevalence of MSK. It appears that no distinction is made as to which side of the body pain occurs. Therefore, pain in the knee is scored the same, regardless of whether pain is experienced in one or both knees. Could the authors also provide details on how this checklist was adapted for the study?

Authors' Response: The number of categories was slightly reduced when compared against the cited reference (Bahr R, Clarsen B, Derman W, et al. International Olympic Committee consensus statement: methods for recording and reporting of epidemiological data on injury and illness in sports 2020 (including the STROBE extension for sports injury and illness surveillance (STROBE-SIIS)). *Orthop J Sports Med* 2020;8(2):2325967120902908). This was done based on validation with members of the target population which confirmed that the checklist was easy-to-understand and precise. With regard to the potential presence of pain in both body sides, we agree this may have an impact and added this aspect to the limitations section.

In the data processing section when you describe the formula, should it be: moderate PA – vigorous PA) + vigorous PA *2? Also, please ensure it is clear you are referring to minutes of physical activity.

Authors' Response: No, the formula is correctly described. MVPA is the sum of moderate and vigorous PA. Subtracting vigorous PA from MVPA yields the amount of moderate PA. Moderate and vigorous PA (*2) in sum yield the value for determining guideline compliance. As you recommended, we added "minutes" in front of moderate-to-vigorous PA or vigorous PA.

I have several concerns about the adequacy and level of detail provided for the analysis. First, what is the reasoning behind categorizing what are potentially continuous data into compliers, double compliers, etc.? If PA is reported in minutes and by intensity, which is not clear from the paper at this point, would it not make more sense to keep data continuous? Second, since for each body region a separate analysis was run, it would seem necessary to correct for multiple testing. Third, how were the confounding variables recorded? Finally, please can you clarify if all confounding variables were included in all multivariate logistic regressions and report the results of these in full?

Authors Response: (1) We used categorizations to be congruent with international recommendations of physical activity. (2) With regard to the high number of comparisons, we totally agree with the reviewer that this may lead to alpha error inflation. We now specify that our analysis was explorative in nature in the objective and discussion section and point towards this limitation with regard to the interpretation of the results. (3) As per methods section, all potential cofounders were tested with univariate regression, and subsequently those significantly associated were included into the multivariate model.

Could you comment on how all materials were translated into the various languages and how equivalency of translations was ensured?

Authors' Response: The questionnaire was available in 7 languages (Dutch, English, German, French, Italian, Brazilian-Portuguese, Spanish), and clarity and comprehensibility were validated by native speakers through forward and backward translation. This was added in the Methods section.

Results:

Please, could you break down results by country?

Authors' Response: We fully understand that breaking down the results by country would be of interest; however, considering that we had 14 countries along with other countries that were not specified, this would massively hamper comprehensibility and readability.

For participants, please could you report mean age with SD? What were the maximum and minimum ages?

Authors' Response: This will be provided in the Results section, and the breakdown of mean age with SD by countries is provided in the new table (Table 1). Minimum age was 18, and maximum age was 100. This was added in the Results section.

Please could you report full results from all analyses, i.e., beta, standard error, and ORs (with CI) for PA and confounding variables and p-values?

Authors' Response: Thank you very much for this suggestion. Due to the very high number of analyses, we decided to keep reporting concise. The provided OR and 95% CI represent the standard of reporting and, e.g. the p value, does not provide additional information (particularly, in view of the high number of tests, the 95% CI is more informative).

Discussion:

Please could you discuss more on potential pathways specifically between PA and upper and lower extremity pain?

Authors' Response: Unfortunately, there is no explicit pathway that can specifically explain the relationship between PA and upper and lower extremity pain. Therefore, we rather provided broader pathways that are implicated in relationship between PA and MSK-pain:

"The underlying mechanisms of how PA modulates pain are not completely understood, but several pathways have been proposed. Animal study findings suggest regular PA may act on the central nervous system (CNS) and alter rate of pain hypersensitivity, dysregulation of pain modulation, and development of chronic pain. In humans, it has been proposed that PA may intervene excitability and inhibition in the CNS, and anti-inflammatory and antioxidant effects of regular PA might diminish the processes contributing to central sensitization. Other proposed mechanisms in humans include the activation of opioid and serotonin pathways or involvement of endocannabinoid system induced from regular PA which could exert analgesic effects."

Please can you comment on the sample's mean age?

Authors' Response: Because the prevalence of MSK-pain increases with age and our univariate analysis showed that age was associated with MSK-pain, our multivariable model adjusted for age. Therefore, the mean age of the sample would not have significant meaning in the context of discussion. However, in order to make this part clear, we added the following sentence at the beginning of our discussion: "Our large-scale multinational study is novel in that it identified the associations between different degrees of compliance to PA recommended by WHO and multiple body locations in the general population after adjusting for multiple cofounding factors including age, which is known to be positively associated with MSK-pain prevalence."

Overall, the discussion is too descriptive.

Authors' Response: We understand your concerns but given the explorative and cross-sectional nature of the study, discussion had to revolve around the association that we found and inevitably became descriptive.

Other than being mentioned on page 9, what is the relevance of Table 2? How were the findings derived? Why are they not discussed?

Authors' Response: The number of pain locations Table 2 (now Table 3) was added for each individual, and this was categorized into having pain in minimum 1 location, 3 locations, 5 locations, and 10 locations. Then, the association between PA and number of MSK-pain locations was analyzed. This analysis was intended to overview how the volume of PA was asociated with the numbers of pain locations rather than specific pain locations. The finding was somewhat consistent with the theme that 300-450 min (PA above the WHO recommendation was associated with lower odds of multiple pain locations while excessive amount was associated with higher odds of having pain), and this theme has been discussed by anatomical locations, specifically. Addressing your concerns, we added these findings in the clinical implication part of the Discussion. "Also, this range was associated with lower odds of having pain in multiple number of locations… Furthermore, PA above 750 minutes was associated with having at least one pain location."

VERSION 2 – REVIEW

REVIEWER	Smith, Benjamin Derby Teaching Hospitals NHS Foundation Trust, Physiotherapy Outpatients
REVIEW RETURNED	23-May-2022

GENERAL COMMENTS	Thank you for address my comments. I have no further concerns and I am happy to recommend publication.

REVIEWER	Niederstrasser, Nils University of Portsmouth, Psychology
REVIEW RETURNED	27-May-2022

GENERAL COMMENTS	Thank you for the opportunity to review the revised version of this manuscript. The authors have done an excellent job at improving the manuscript. Nevertheless, the following issues need to be addressed prior to publication: Abstract: I think the language should be adapted to reflect the exploratory nature of this investigation. For example, the objective section in the abstract should convey to the reader that this was an exploratory study. Further, in the abstract's conclusion I would be hesitant to say "reduced" or "increased" pain, as you did not assess intensity but rather presence and absence of pain and number of pain sites. This should be made clear throughout the manuscript. Introduction: I am still unsure about your suggestion of there being a literature gap between pain and physical activity, as implied by the revised sentence: "However, compared to the number of studies investigating the association of PA with non-communicable disease, there seems to be a literature gap regarding MSK-pain." I think it is acceptable to say there are possibly fewer studies specifically on the topic of pain and physical activity compared to between physical activity and non-communicable diseases, but I would not go so far as to suggest that this disparity necessarily constitutes a gap in the literature, especially in the light of evidence you have provided up to that point. You may want to consider rewording this to say that this particular association has perhaps received less attention up to now or more closely delineate the literature gap you have identified. In your final paragraph, I would replace "investigate" with "explore" to establish consistency in the manuscript. It might be worth clarifying in this paragraph as well that your objective, presumably, was to explore the association of total PA with presence of MSK-pain in various body locations. Finally, as in the abstract, please clarify that you are looking for the presence or absence of pain in body locations and not intensity. So, PA would be associated with the presence of pain in fewer or more body regions rather than increased or reduced pain. Method: Please describe the data analysis used to examine the association between PA and number of body locations. What type of logistic regression was used to assess this? What is the reasoning behind categorising the number of pain locations? I am wondering if a linear regression model would have been more appropriate here. Could you explain how you assessed the confounding variables used in your analyses? For example, what were the response options for employment status and how was depression risk assessed? I think this is mentioned in the footnote for Table 1, but a description of the WHO-5 at least should be in the main text. I am concerned that you refer to the WHO-5 as depression risk, rather than wellbeing. Please could you clarify this? I am still concerned regarding the high number of logistic regressions (i.e., 20). It might be worth applying a multiple-testing correction, such as Bonferroni. Results: In the interest of consistency, the language describing the main

	findings should be changed to adequately reflect the analyses. Instead of saying “reduced pain in one body location” it should be made clear that the finding represents a lower likelihood of suffering from pain in a single body location. Discussion: Please adjust the first sentence to reflect the exploratory nature of the study. Please could you clarify whether the distinctions you report on page 9 regarding upper and lower extremities were actually assessed through analyses or whether this is based on a count of associations. Regarding the distinction in general, I am not sure regarding the merit of making this distinction as it does not seem to feature in the analyses and is not justified in the introduction. You say “However, our study found that participating in PA between 450-900+ min was associated with lower odds of having pain in neck/cervical spine”. Table 2, however, suggests that that those engaging in 750-900 minutes of PA do in fact not show this association, when looking at the adjusted OR. Please can you clarify this?

VERSION 2 – AUTHOR RESPONSE

Reviewer: 1
Mr. Benjamin Smith, Derby Teaching Hospitals NHS Foundation Trust, University of Nottingham School of Medicine
Comments to the Author:

Thank you for address my comments. I have no further concerns and I am happy to recommend publication.

Authors’ Response: We thank you again for your time and effort to review our manuscript.

Reviewer: 2
Dr. Nils Niederstrasser, University of Portsmouth
Comments to the Author:
Thank you for the opportunity to review the revised version of this manuscript. The authors have done an excellent job at improving the manuscript. Nevertheless, the following issues need to be addressed prior to publication:

Authors’ Response: We thank you again for taking time and effort to provide thoughtful comments. We tried our best to address your questions and concerns and hope that our revised manuscript meets your expectation.

Abstract:
I think the language should be adapted to reflect the exploratory nature of this investigation. For example, the objective section in the abstract should convey to the reader that this was an exploratory study. Further, in the abstract’s conclusion I would be hesitant to say “reduced” or “increased” pain, as you did not assess intensity but rather presence and absence of pain and number of pain sites. This should be made clear throughout the manuscript.

Authors’ Response: Following your recommendation, we changed “investigate” to “explore” in the objective of our abstract. Furthermore, we changed our conclusion to “A dose of 300-450 min WHO-equivalent PA/week was associated with lower odds of MSK-pain in six locations. On the other hand, excessive doses of PA were associated with higher odds of pain in certain body locations,”

recognizing that “reduced” or “increased” may sound like referring to intensity, which we did not assess.

Introduction:
I am still unsure about your suggestion of there being a literature gap between pain and physical activity, as implied by the revised sentence: “However, compared to the number of studies investigating the association of PA with non-communicable disease, there seems to be a literature gap regarding MSK-pain.” I think it is acceptable to say there are possibly fewer studies specifically on the topic of pain and physical activity compared to between physical activity and non-communicable diseases, but I would not go so far as to suggest that this disparity necessarily constitutes a gap in the literature, especially in the light of evidence you have provided up to that point. You may want to consider rewording this to say that this particular association has perhaps received less attention up to now or more closely delineate the literature gap you have identified.

Authors’ Response: We agree that we might have gone to far as to suggest that there is a gap in the literature. As you recommended, we revised our sentence to “However, compared to the number of studies investigating the association of PA with non-communicable disease, there seems to be fewer number of studies on the topic of PA and MSK-pain.”

In your final paragraph, I would replace “investigate” with “explore” to establish consistency in the manuscript. It might be worth clarifying in this paragraph as well that your objective, presumably, was to explore the association of total PA with presence of MSK-pain in various body locations. Finally, as in the abstract, please clarify that you are looking for the presence or absence of pain in body locations and not intensity. So, PA would be associated with the presence of pain in fewer or more body regions rather than increased or reduced pain.

Authors’ Response: Following your recommendation, we modified the final paragraph of the introduction: “The purpose of this study was to explore the association of total PA with presence of MSK-pain in a variety of anatomical locations including both upper and lower extremities. We hypothesize that greater time spent in PA than WHO recommendation would be associated with the absence of MSK-pain in more body regions, but excess time performing PA might be associated with the presence of MSK-pain in more body regions.”

Method:
Please describe the data analysis used to examine the association between PA and number of body locations. What type of logistic regression was used to assess this? What is the reasoning behind categorising the number of pain locations? I am wondering if a linear regression model would have been more appropriate here.

Authors’ Response: Thank you for raising this point. Generally, we have now better clarified that univariate binary logistic regression (identifying univariate associations between PA or potential confounders and pain) and multivariate binary logistic regression (identifying corrected associations between PA and pain) was performed.

Regarding the number of pain locations, we indeed realized that this had not been mentioned in the analysis section and we have amended this part accordingly: “As participants may have a strongly varying number of pain locations and as the impact of pain on the individual may vary with the number of affected body regions, additional analyses, using the same procedures as described above (binary logistic regression corrected for confounders), were performed to obtain adjusted OR for pain in only one, at least 3, 5, or 10 body locations.”

We decided to use logistic regression here too because obtaining OR was congruent with the prior analysis steps, more easily interpretable for the reader (e.g., if yielding a coefficient with decimals), and most importantly because assumptions for linear regression were unmet.

Could you explain how you assessed the confounding variables used in your analyses? For example, what were the response options for employment status and how was depression risk assessed? I think this is mentioned in the footnote for Table 1, but a description of the WHO-5 at least should be in the main text. I am concerned that you refer to the WHO-5 as depression risk, rather than wellbeing.

Please could you clarify this?

Authors' Response: As stated under "Data Processing and Statistical Analysis" of our Methods section, univariate logistic regression was run for the potential confounding variables we could obtain from the questionnaire. Variables listed in the manuscript (sex, age, employment status, and depression risk) were found statistically significant in univariate logistic regression and therefore used in our multivariate logistic regression.

In terms of employment status, participants were asked where they worked in multiple choices which included no employment option, and this question was used to categorize participants into employment or no employment for our analysis. Furthermore, although the name of WHO Well Being Index (WHO-5) does not explicitly seem to name depression risk, numerous studies have validated the WHO-5 as a measure for depression. We will be citing a study by Topp et al. who performed a systematic review on WHO-5 and showed its wide use in depression research. The following was added to address your concerns: "In addition to the assessment of PA, participants were asked where they worked in multiple choices which also included a 'no employment' option, and the answers to this question were used to categorize participants into being employed or not employed for our analysis. Also, the WHO-Well-Being Index (WHO-5) was used to capture depression risk as validated by previous research."

I am still concerned regarding the high number of logistic regressions (i.e., 20). It might be worth applying a multiple-testing correction, such as Bonferroni.

Authors' Response: We agree with the reviewer. Kindly note this is acknowledged in the limitations section ("Furthermore, because a separate analysis was run for each body region, there is a risk of multiple testing problem."). However, as, revising the manuscript, we marked the explorative character of the analysis, not performing an alpha error correction is consistent as exploratory studies generating hypotheses do not necessarily need to account for alpha error inflation (which may thwart the above objective).

Results:
In the interest of consistency, the language describing the main findings should be changed to adequately reflect the analyses. Instead of saying "reduced pain in one body location" it should be made clear that the finding represents a lower likelihood of suffering from pain in a single body location.

Authors' Response: Following your recommendation, we revised the language describing the main findings throughout the Results section.

Discussion:
Please adjust the first sentence to reflect the exploratory nature of the study.

Authors' Response: Following your suggestion, we adjusted the first sentence to reflect the exploratory nature of the study: "The purpose of the present study was to explore the relation between PA and MSK-pain."

Please could you clarify whether the distinctions you report on page 9 regarding upper and lower extremities were actually assessed through analyses or whether this is based on a count of associations. Regarding the distinction in general, I am not sure regarding the merit of making this distinction as it does not seem to feature in the analyses and is not justified in the introduction.

Authors' Response: Our distinction of upper and lower extremities was based on the observation of our results, and the purpose of distinction was to better organize the manuscript rather than sporadically discussing different body parts throughout the manuscript. Because our goal was not to specifically look at the difference between upper and lower extremities, justification may not be necessary in the introduction. We hope you understand that this distinction was rather a way of presenting and organizing our results.

You say "However, our study found that participating in PA between 450-900+ min was associated with lower odds of having pain in neck/cervical spine". Table 2, however, suggests that that those

engaging in 750-900 minutes of PA do in fact not show this association, when looking at the adjusted OR. Please can you clarify this?

Authors' Response: We apologize for the confusion. There are some confidence intervals that either start or end with 1.00, and these results rather derive from rounding the decimals than not being significant. We added the following in the footnote of Table 2 for clarification: "The numbers in bold denote significant results, and the confidence interval that starts or ends with 1.0 derives from rounding the decimals."

VERSION 3 – REVIEW

REVIEWER	Niederstrasser, Nils University of Portsmouth, Psychology
REVIEW RETURNED	25-Aug-2022

GENERAL COMMENTS	The authors have done an excellent job at revising the manuscript. All of my concerns have been addressed satisfactorily. While I recommend this manuscript for publication, I ask the authors to please check Table 2. The adjusted OR for "ribs" and "450-600 minutes" is printed in bold indicating a significant relationship (i.e., 0.74), while the 95% CI would suggest that this association is in fact not significant (95% CI = 0.46-1.17).