

## PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

### ARTICLE DETAILS

#### Title (Provisional)

Secondary causal mediation analysis of a pragmatic clinical trial to evaluate the effect of chiropractic care for US active-duty military on biopsychosocial outcomes occurring through effects on low back pain interference and intensity

#### Authors

Shannon, Zacariah; Long, Cynthia R.; Chrischilles, Elizabeth; Goertz, Christine; Wallace, Robert; Casteel, Carri; Carnahan, Ryan M.

---

### VERSION 1 - REVIEW

---

<b>Reviewer</b>	<b>1</b>
<b>Name</b>	<b>Lalji , Rahim</b>
<b>Affiliation</b>	<b>Balgrist University Hospital, Department of Chiropractic Medicine</b>
<b>Date</b>	<b>27-Feb-2024</b>
<b>COI</b>	<b>I am currently a post-doctoral researcher within the advanced methods department of Novartis in Basel, Switzerland. I have no competing interests with regards to the peer-review of this manuscript.</b>

---

Thank you for the opportunity to review this manuscript on the effect of chiropractic care on whole health factors. I found the article interesting, clinically applicable, and well written. Below I provide some comments mainly on the need to elaborate further on certain areas of the manuscript.

#### Minor Corrections

1. Line 56: "These interventions target patient beliefs, psychosocial function, and physical pain"

Line 57: "Mechanistic research of manual therapy has previously demonstrated effects on physical processes such as joint stiffness, [6,7] temporal summation, [8,9] and muscle activity, [10–12] which are associated with changes in pain."

Further rationale is needed to describe the role of pain being an appropriate mediator between chiropractic care and whole health factors. For example, it can similarly be argued

that improvements in whole health factors (e.g., psychosocial factors) can be the mediating factor for pain reduction following chiropractic care, particularly when applied to a chronic low back pain population. Rationale as to the use of PROMIS-29 subsections to adequately measure pain and whole health factors should also be included.

2. Line 75. "Do pain interference and pain intensity mediate the effect of chiropractic care on physical function, fatigue, sleep disturbance, anxiety, depression, and satisfaction with social roles?"

Reviewer comment: Given that this secondary analysis was performed on active-duty members with low back pain in the United States, I believe this patient population and setting should be added into the proposed research question. As a result, the research question should be more focused than it is currently stated.

3. Line 85. "A de-identified dataset was generated and a limited dataset was obtained and used for these analyses"

Reviewer comment: I would like to gain further understanding on how this limited dataset was obtained/generated. At the moment it is unclear as to why a full dataset was not used for this secondary analysis. It would be also useful to understand the proportion of patients within the full dataset which were subsequently used for this secondary analysis and provide rationale as to why patients were subsequently excluded for this specific study.

4. Line 120: We followed the intention-to-treat principle by defining our exposure as assignment to the intervention arm (usual medical care plus chiropractic care vs. usual medical care alone).

Reviewer comment: This is unclear as the intention-to-treat approach is based on the principle that all participants who are randomized are analyzed rather than a means for defining exposure assignment. I believe there may be 2 points within this one sentence and can be spaced out.

5. Line 163 "We used multiple imputation by chained equations with 25 imputations and 5 iterations to impute missing mediator and outcome values and 1000 bootstrap samples to generate bias-corrected standard errors and 95% confidence intervals"

Reviewer comment: I appreciate the robustness of this missing data approach taken, however what I believe is missing is a short description of the underlying missing data mechanism found (Missing at random, missing completely at random, missing not at random) and the rationale taken for the approach described. For example, multiple imputation alone may not be a satisfactory approach in all missing data situations.

6. Figure 1. Arrow from C to A

Reviewer comment: Given that A is treatment group assignment which was determined through the process of randomization, a confounder (C) could not be causally associated with A. Given that the process of randomization should mitigate confounding in treatment

selection. This is my interpretation of the directed acyclic graph within the context of this mediator analysis and I would ask to check and confirm.

7. References: In few instances I noticed inconsistencies in the referencing format. For example, reference 8 journal name is provided in full, while the remainder are provided in ISO4 format.

---

<b>Reviewer</b>	<b>2</b>
<b>Name</b>	<b>Kazal, Louis</b>
<b>Affiliation</b>	<b>Dartmouth College Geisel School of Medicine, Community and Family Medicine</b>
<b>Date</b>	<b>04-Mar-2024</b>
<b>COI</b>	<b>None.</b>

---

#### **Review of BMJopen-2023-083509**

##### General Comments:

Overall, the effort in organization and clarity of the study as presented is appreciated. Explaining a mediated analysis had the potential to become overly wordy, but did not. The subject is an important one as evidence-based guidelines make spinal manipulation one of the first-line options in treatment of back pain. Continuing to add to the evidence base and refine it is very important going forward to help the clinical management of low back pain become more effective and less costly, as currently it is a huge financial burden to the U.S. health care system and a leading cause of disability in the U.S.

##### Specific Comments:

Line 83: Reference links to "Assessment of Chiropractic Treatment for Low Back Pain and Smoking Cessation in Military Active-Duty Personnel." Is this because it was a secondary analysis of that data?

Line 85-6: Please address why it wasn't published or registered.

Line 90: It would be helpful to explain why you choose care to be "as usual and not prescribed by trial procedures"?

Line 91: Did you consider measuring cointerventions? If not, why not?

Line 92-93: Move to next paragraph where it would logically follow in sequence after the start date.

Line 97: Please address why it took 4 years to achieve the target goal. The study period began in 2012 and concluded in 2016. Some acknowledgement would be helpful about this being relatively old data at this point. Since treatment was not prescribed by a study protocol, did this 4-year study period create an opportunity for unforeseen external factors to influence care in the study; things like newer techniques and different approaches to care that develop over time?

Lines 102-3: Why exclude this? It would be helpful to know as it's not immediately obvious.

Line 107: Could you address the limitations of a study not designed for the statistics ultimately used to analyze it?

Lines 115-117: Would put an "s" on outcome to make it clear you're referring to the aforementioned outcomes. When read it, the wording may trip up readers; same with the direct effect sentence.

Lines 121-124: How do you know "single visit only" for medical care over a 6-week period constitutes usual medical care? How many participants fall into this category vs. more than one visit over the 6 weeks? Were these patients seen by a physician or associate clinician once yet referred to PT and thus may have received evidence-based care? How representative is a single visit to compare against?

Line 149-151: "It is also assumed that mediator-outcome confounding is accounted for by adjusting for the baseline values of the PROMIS-29 measures considered in these analyses." How was this done?

Line 182: One-half of the participants had acute back pain and the other half chronic back pain. Is this a limitation of the study not accounted for? Does this compromise the statistics and conclusion given acute and chronic back pain have different natural histories? Studying this group as whole is akin to

lumping apples and oranges together but treating it as the same fruit. Additionally, anxiety and other psychosocial issues would be expected to be less in those with acute back pain compared to chronic back pain. Should these two groups be analyzed separately and compared? Would the "n" for both categories be large enough once divided in half?

Lines 189-90: May want to address why you don't report p-values in Table 1? (These could be helpful when glancing through a table to compare the two group's similarities.)

Lines 194-6: Would be helpful to explain the general population; is it normed group at T50 like mentioned before? This isn't clear and also unsure where or how you got the population data from?

Lines 199-200: Please clarify to readers if the distribution created the floor effect, or is it instead a consequence of the data? e.g., "This resulted in a floor effect and a heavily right-skewed distribution in anxiety and depression outcomes. This constrained variability makes it challenging to detect meaningful changes, assess the impact of interventions, and establish mediation effects, particularly in the context of mental health outcomes."

Lines 203-4:

- (A) What limitations might there be if you did not include exposure-mediator interactions in the final models due to having little effect on direct and indirect effect estimates?
- (B) Can you provide a quantitative value for the effect it had, e.g. direct (#) and indirect (#)?

Lines 216-9: If the mediated proportion is a proportion, i.e., bounded by 0 and 1, how can the CI go beyond 1? Should there be an explanation, or is this just the reviewer's lack of understanding the analysis?

Line 244: Was a rationale provide for why the effects are considered moderate or cutoff points for

what is considered moderate?

Line 248: Would note that this work is unpublished or submitted for publication.

Lines 289-90: Making the association between fatigue and mental health as currently presented appears like a stretch. The number of conditions that can cause fatigue is broad and calling it out in this population as likely to be mental health related makes it seem like there is a reach to tie it together with the conclusion sentence in the paragraph. It could be a fair sentence if more support or information for the association can be provided to give the reader context.

---

<b>Reviewer</b>	<b>3</b>
<b>Name</b>	<b>X.W. Liew, Bernard</b>
<b>Affiliation</b>	<b>University of Essex, School of Sport, Rehabilitation and Exercise Sciences</b>
<b>Date</b>	<b>16-May-2024</b>
<b>COI</b>	<b>None</b>

---

I have been requested to assess the statistical methods of this paper. Overall, the authors performed the analysis correctly and the reporting is largely concise. They adopted some of the latest approaches in mediation analysis, alongside multiple imputation. My only issue is in the reporting of the results.

I would like the authors to be clear what kind of indirect and direct effects are they reporting. Is it Pure or Total natural indirect effect? Is it controlled or other type of indirect effect? Also for imputation, why was 5 iterations used? was it sufficient?

---

## **VERSION 1 - AUTHOR RESPONSE**

Reviewer: 1

Dr. Rahim Lalji, Balgrist University Hospital, University of Zurich

Comments to the Author:

Thank you for the opportunity to review this manuscript on the effect of chiropractic care on whole health factors. I found the article interesting, clinically applicable, and well written. Below I provide some comments mainly on the need to elaborate further on certain areas of the manuscript.

Minor Corrections

1. Line 56: "These interventions target patient beliefs, psychosocial function, and physical pain"

Line 57: "Mechanistic research of manual therapy has previously demonstrated effects on physical processes such as joint stiffness, [6,7] temporal summation, [8,9] and muscle activity, [10-12] which are associated with changes in pain."

Further rationale is needed to describe the role of pain being an appropriate mediator between chiropractic care and whole health factors. For example, it can similarly be argued that improvements in whole health factors (e.g., psychosocial factors) can be the mediating factor for pain reduction following chiropractic care, particularly when applied to a chronic low back pain population. Rationale as to the use of PROMIS-29 subsections to adequately measure pain and whole health factors should also be included.

This is a good point. We agree that chiropractic care could affect pain through biopsychosocial factors and vice versa and that the biopsychosocial model of pain does not currently suggest a causal direction to suggest it occurs “one way”. We have added rationale to justify the analysis of this paper while recognizing this limitation of the biopsychosocial model.

2. Line 75. “Do pain interference and pain intensity mediate the effect of chiropractic care on physical function, fatigue, sleep disturbance, anxiety, depression, and satisfaction with social roles?”

Reviewer comment: Given that this secondary analysis was performed on active-duty members with low back pain in the United States, I believe this patient population and setting should be added into the proposed research question. As a result, the research question should be more focused than it is currently stated.

This is a great point. We have incorporated this suggestion by adding to the research question.

3. Line 85. “A de-identified dataset was generated and a limited dataset was obtained and used for these analyses”

Reviewer comment: I would like to gain further understanding on how this limited dataset was obtained/generated. At the moment it is unclear as to why a full dataset was not used for this secondary analysis. It would be also useful to understand the proportion of patients within the full dataset which were subsequently used for this secondary analysis and provide rationale as to why patients were subsequently excluded for this specific study.

We have updated the language to be clearer about the data used. We had outcome data from all participants in the trial. The dataset was “limited” to outcome variables and to information that could not be used to re-identify the participants, such as PHI or, for instance, substituting participant ID for name or other individual identifiers. A limited dataset was used for the purposes of maintaining participant confidentiality and in accordance with the data principle of least privilege.

4. Line 120: We followed the intention-to-treat principle by defining our exposure as assignment to the intervention arm (usual medical care plus chiropractic care vs. usual medical care alone).

Reviewer comment: This is unclear as the intention-to-treat approach is based on the principle that all participants who are randomized are analyzed rather than a means for defining exposure assignment. I believe there may be 2 points within this one sentence and can be spaced out.

We have tried to make this clearer by editing the wording of the sentence. By stating that we are treating exposure in our models as assignment to intervention we are trying to communicate that we are analyzing according to intention to treat rather than a per protocol analysis which

would instead define exposure as those that received chiropractic care vs. those that received chiropractic care plus usual medical care and omit those that did not receive the interventions as intended.

5. Line 163 “We used multiple imputation by chained equations with 25 imputations and 5 iterations to impute missing mediator and outcome values and 1000 bootstrap samples to generate bias-corrected standard errors and 95% confidence intervals”

Reviewer comment: I appreciate the robustness of this missing data approach taken, however what I believe is missing is a short description of the underlying missing data mechanism found (Missing at random, missing completely at random, missing not at random) and the rationale taken for the approach described. For example, multiple imputation alone may not be a satisfactory approach in all missing data situations.

We have added additional details regarding our missing data approach along with an additional reference that explains the approach more thoroughly. (Vansteelandt, Stijn; Bekaert, Maarten; and Lange, Theis (2012) "Imputation Strategies for the Estimation of Natural Direct and Indirect Effects," Epidemiologic Methods: Vol. 1: Iss. 1, Article 7.)

6. Figure 1. Arrow from C to A

Reviewer comment: Given that A is treatment group assignment which was determined through the process of randomization, a confounder (C) could not be causally associated with A. Given that the process of randomization should mitigate confounding in treatment selection. This is my interpretation of the directed acyclic graph within the context of this mediator analysis and I would ask to check and confirm.

The parent trial used adaptive allocation to balance group assignment on sex, age, LBP duration, and pain intensity. Therefore, these are the only factors that could be considered causal for intervention group assignment and are why arrows were from the group of confounding factors to the exposure. We have added a statement in the outcomes, mediators, and confounding section of the manuscript to clarify this point.

7. References: In few instances I noticed inconsistencies in the referencing format. For example, reference 8 journal name is provided in full, while the remainder are provided in ISO4 format.

We have edited the references as suggested.

Reviewer: 2 [PLEASE SEE ATTACHED FILE FOR COMMENTS FROM REVIEWER 2]

Dr. Louis Kazal, Dartmouth College Geisel School of Medicine

Comments to the Author:

See attached.

Please note, Dr. Logan Benjamin (Department of Community and Family Medicine, Geisel School of Medicine; Chiropractic Research Fellow at Dartmouth) assisted me in this review.

General Comments:

Overall, the effort in organization and clarity of the study as presented is appreciated. Explaining a mediated analysis had the potential to become overly wordy, but did not. The subject is an important one as evidence-based guidelines make spinal manipulation one of the first-line options in treatment of back pain. Continuing to add to the evidence base and refine it is very important going forward to help the clinical management of low back pain become more effective and less costly, as currently it is a huge financial burden to the U.S. health care system and a leading cause of disability in the U.S.

Specific Comments:

Line 83: Reference links to "Assessment of Chiropractic Treatment for Low Back Pain and Smoking Cessation in Military Active-Duty Personnel." Is this because it was a secondary analysis of that data? "Assessment of Chiropractic Treatment for Low Back Pain and Smoking Cessation in Military Active-Duty Personnel." was the title of the parent trial when the protocol was published.

Line 85-6: Please address why it wasn't published or registered.

The analysis plan was not published because this was an exploratory analysis. We have added language throughout the manuscript to state that this is an exploratory analysis. We have included the wording around publication of the analysis because we are following the suggested reporting standards of EQUATOR to increase transparency and quality in the reporting of this manuscript.

Line 90: It would be helpful to explain why you choose care to be "as usual and not prescribed by trial procedures"?

This was a detail of the parent trial of which we are analyzing the data and therefore was not specifically chosen by our study team for this manuscript. The parent trial was a pragmatic trial evaluating the addition of "chiropractic care" to usual care rather than subcomponents of chiropractic care. Therefore, the parent trial did not try to control which subcomponents of chiropractic care were delivered, but rather to allow it to be delivered as normal within the health system. This is an important design aspect of a pragmatic clinical trial which gives information about what outcomes to expect when chiropractic care is added to usual medical care for active-duty military. We have added some language to explain this is due to it being a pragmatic clinical trial.

Line 91: Did you consider measuring cointerventions? If not, why not?

We performed a secondary analysis of an existing data set. This information was not available to be included in our secondary analysis.

Line 92-93: Move to next paragraph where it would logically follow in sequence after the start date. We have rearranged this information as suggested.

Line 97: Please address why it took 4 years to achieve the target goal. The study period began in 2012 and concluded in 2016. Some acknowledgement would be helpful about this being relatively old data at this point. Since treatment was not prescribed by a study protocol, did this 4-year study period create an opportunity for unforeseen external factors to influence care in the study; things like newer techniques and different approaches to care that develop over time?

The primary outcomes paper referenced in our manuscript has more details on the timeline of the study. In brief explanation, it took from Sep 2012-Nov 2015 to recruit the 750 participants in the trial. Data collection (follow up) was completed in Feb 2016 and data analysis of the primary results in 2016, in about a month following data collection. The interventions delivered as part of chiropractic care are described in another publication which we have referenced in our paper. We have added a statement regarding the interventions delivered as part of chiropractic care being consistent with those described in the National Board of Chiropractic Examiners practice analysis survey. To our knowledge, there has not been major changes to the scope of chiropractic practice to suggest that interventions would be significantly different, there is very limited evidence comparing



manipulation techniques with most suggesting little difference on pain outcomes, and meta-analyses of exercise interventions also suggest no difference between different exercise interventions on pain outcomes. We do agree that the way in which chiropractic care is delivered has the potential to have an effect. We have addressed this by adding to the discussion regarding that these were the mediators of chiropractic care as delivered and that future work could evaluate subcomponents of care.

Lines 102-3: Why exclude this? It would be helpful to know as it's not immediately obvious.  
We have added text to clarify this.

Line 107: Could you address the limitations of a study not designed for the statistics ultimately used to analyze it?  
We have added text to clarify this.

Lines 115-117: Would put an "s" on outcome to make it clear you're referring to the aforementioned outcomes. When read it, the wording may trip up readers; same with the direct effect sentence.  
We have edited as suggested.

Lines 121-124: How do you know "single visit only" for medical care over a 6-week period constitutes usual medical care? How many participants fall into this category vs. more than one visit over the 6 weeks? Were these patients seen by a physician or associate clinician once yet referred to PT and thus may have received evidence-based care? How representative is a single visit to compare against?

This is a good point. First, we did not specify that the participants had only 1 visit with a medical doctor, but rather the % that had at least 1 visit. The data were not available to report the number of medical visits for each participant. The trial was a pragmatic testing of usual medical care vs. usual medical care plus chiropractic care. The participants were not restricted to only 1 medical visit and instead attempted to be scheduled to have an initial medical visit with a medical physician at the military treatment facility. Usual care was then delivered to participants by the medical physician as normal and the trial did not interfere. Because the care was not restricted, some participants were referred for PT or to other providers, which is described in the supplement of the primary results paper. It is true that participants in the usual medical care alone group may have received evidence-based care whether that was delivered by the medical physician or in combination with a PT. The primary trial sought to evaluate the effect of adding chiropractic care to usual medical care rather than to evaluate the effect of receiving evidence-based care or implementing a primary spine care model or similar.

We do agree that there are some unique aspects to the care delivery at military treatment facilities. To address this point we have added to the limitations to highlight that the availability of care could affect the outcomes and may not necessarily be transportable to other populations outside of military treatment facilities.

Line 149-151: "It is also assumed that mediator-outcome confounding is accounted for by adjusting for the baseline values of the PROMIS-29 measures considered in these analyses." How was this done?

We used the CMAR package to specify these values as confounders which adds coefficients to the models for each of the variables.

Line 182: One-half of the participants had acute back pain and the other half chronic back pain. Is this a limitation of the study not accounted for? Does this compromise the statistics and conclusion given acute and chronic back pain have different natural histories? Studying this group as whole is

akin to lumping apples and oranges together but treating it as the same fruit. Additionally, anxiety and other psychosocial issues would be expected to be less in those with acute back pain compared to chronic back pain. Should these two groups be analyzed separately and compared? Would the “n” for both categories be large enough once divided in half?

In our models, we considered baseline pain duration as a confounder and adjusted for it. The alternative approach of splitting the sample into distinct groups based on baseline pain duration would have several considerations. The first is that it addresses a different research question than the primary trial results which evaluated the effect of adding chiropractic care to usual medical care for pain of any duration. Our analysis provides additional insight/context to the findings of the parent trial. A second reason we did not opt to split the sample is that it would severely reduce our power. This was an unplanned secondary analysis and keeping the entire sample provided a greater opportunity to detect differences. We agree that there is value in disentangling the effects of chiropractic care on acute and chronic pain and that this could be done in future analyses. We have added to the discussion to address this.

Lines 189-90: May want to address why you don't report p-values in Table 1? (These could be helpful when glancing through a table to compare the two group's similarities.)

We are following the recommendations of the CONSORT statement and others that provide the rationale for not reporting p-values in table 1:

<https://journals.sagepub.com/doi/10.1177/1741826711421688>

Lines 194-6: Would be helpful to explain the general population; is it normed group at T50 like mentioned before? This isn't clear and also unsure where or how you got the population data from?

We have changed the wording here and up above to try to be more clear. The reviewer is correct in that a T-score of 50 would represent a general population average. Part of the development of the PROMIS questionnaires included taking representative samples of the US general population. The general population samples provides the central point of the scale to show if participants in the sample report more or less of the measure. This provides a frame of reference to which each study sample that uses the PROMIS questionnaires can be compared to. This is described in the PROMIS measurement development website: <https://www.healthmeasures.net/explore-measurement-systems/promis/intro-to-promis>

Lines 199-200: Please clarify to readers if the distribution created the floor effect, or is it instead a consequence of the data? e.g., "This resulted in a floor effect and a heavily right-skewed distribution in anxiety and depression outcomes. This constrained variability makes it challenging to detect meaningful changes, assess the impact of interventions, and establish mediation effects, particularly in the context of mental health outcomes."

This is a good point. Floor effect may not be an accurate description and has been removed.

Lines 203-4:

(A) What limitations might there be if you did not include exposure-mediator interactions in the final models due to having little effect on direct and indirect effect estimates? (B) Can you provide a quantitative value for the effect it had, e.g. direct (#) and indirect (#)?

We checked to see if the interaction terms affected our effect estimates. If the interaction term had changed the effect estimates we would have included it in the model to avoid biased effect estimates. Because the interaction terms did not affect the estimates much, we chose to use the most parsimonious model which is the one with fewer terms and similar effect estimates. We view this as a strength of our approach.

Lines 216-9: If the mediated proportion is a proportion, i.e., bounded by 0 and 1, how can the CI go beyond 1? Should there be an explanation, or is this just the reviewer's lack of understanding the analysis?

We have added a brief explanation of why this happens to the discussion.

In brief, in mediation modeling, when you include multiple mediators in one model this can cause some variables to have positive coefficients while some "take on" negative coefficients. This was the case in our models. This happens because once one of the mediators is in the model it impacts the estimates of the other variables. This is consistent with the "Table 2 fallacy" that makes individual coefficients in a multivariable model not be able to be interpreted individually and why we had to compare individual mediator models to the combined to try to parse out individual mediator contribution (Current recommended approach). When the model has variables with positive and negative coefficients simultaneously and this is used to calculate a proportion from multiple regressions, this can lead to confidence intervals that overlap 1. This is explained in Tyler Vanderweele's work, including "Explanation in Causal Inference", as a common occurrence when doing multiple mediator modeling, for which he suggests emphasizing the indirect effect, which does not have the problem of dividing the results from multiple regressions with wide CI's, rather than the proportion mediated in these cases.

Line 244: Was a rationale provide for why the effects are considered moderate or cutoff points for what is considered moderate?

No. We were using moderate as a qualitative adjective.

Line 248: Would note that this work is unpublished or submitted for publication.

Edited as suggested.

Lines 289-90: Making the association between fatigue and mental health as currently presented appears like a stretch. The number of conditions that can cause fatigue is broad and calling it out in this population as likely to be mental health related makes it seems like there is a reach to tie it together with the conclusion sentence in the paragraph. It could be a fair sentence if more support or information for the association can be provided to give the reader context.

This is a good point. Our original intent was to communicate that we suspect that the participants are likely underreporting mental health symptoms due to the population being sampled and that there is a form of reporting bias, rather than such a low prevalence of mental health symptoms. We recognize that the reviewer is correct that there are many causes of fatigue and that we aren't sure that they have mental health symptoms that are causing the fatigue. We have changed the wording to add to this point in an attempt to soften the language and remove the assertion that this is the case.

Reviewer: 3

Dr. Bernard X.W. Liew, University of Essex

Comments to the Author:

I have been requested to assess the statistical methods of this paper. Overall, the authors performed the analysis correctly and the reporting is largely concise. They adopted some of the latest approaches in mediation analysis, alongside multiple imputation. My only issue is in the reporting of the results.

I would like the authors to be clear what kind of indirect and direct effects are they reporting. Is it Pure or Total natural indirect effect? Is it controlled or other type of indirect effect? Also for imputation, why was 5 iteration used? was it sufficient?

We have added a more thorough description of our missing data approach with a relevant reference and included more specific wording on the type of effects throughout.

---

## **VERSION 2 - REVIEW**

---

**Reviewer**                    **2**  
**Name**                        **Kazal, Louis**  
**Affiliation**                **Dartmouth College Geisel School of Medicine, Community and Family Medicine**  
**Date**                        **21-Oct-2024**  
**COI**                         **None.**

---

Dr. Logan Benjamin and I are satisfied with the reviewer responses and corrections.

---

**Reviewer**                    **3**  
**Name**                        **X.W. Liew, Bernard**  
**Affiliation**                **University of Essex, School of Sport, Rehabilitation and Exercise Sciences**  
**Date**                        **18-Sep-2024**  
**COI**                         **None**

---

All my comments are addressed