## PEER REVIEW HISTORY



## ARTICLE DETAILS

| | |
|---|---|
| **TITLE (PROVISIONAL)** | Mechanism evaluation of a lifestyle behavioural intervention for patients with musculoskeletal pain who are overweight or obese: protocol for a causal mediation analysis |
| **AUTHORS** | Lee, Hopin; Wiggers, John; Kamper, Steve; Williams, Amanda; O'Brien, Kate; Hodder, Rebecca; Wolfenden, Luke; Yoong, Sze Lin; Campbell, Elizabeth; Haskins, Robin; Robson, Emma; McAuley, James; Williams, Christopher |

## VERSION 1 - REVIEW

| | |
|---|---|
| **REVIEWER** | Cormac Ryan<br>Teeside University, UK |
| **REVIEW RETURNED** | 22-Oct-2016 |

| | |
|---|---|
| **GENERAL COMMENTS** | The authors are to be congratulated. This is a very interesting and important piece of work. It attempts to answer the question of the mediation of weight loss in relation to pain and function outcomes. I am a clinician scientist rather than a statistician. I can only comment on the clinical rationale for the DAGs presented, however in that respect I believe them to be solid. I would strongly encourage that a statistician provide comment on the statistical component of this work, as I am ill equiped to do so.<br><br>I would like to raise a number of minor issues:<br><br>1. In the last line of the introduction section of the abstract the word obese/overweight needs to be inserted before "patients". Same point for 1st bullet point of stenghts and limitations section.<br>2. In the methods and analysis section it needs to be made clearer which variables are outcomes and which are mediators (primary and alternative).<br>3. the hypothesis stated at the bottom of page 2 (or page 5 of 19) would be very useful in the abstract. (This is a writing structure issue I leave to authors discretion)<br>4. Randomisation subsection - Can the authors clarify how the patients and investigators are to be blinded?<br>5. Pain beliefs only appear to be being targetted in the back intervention - is this correct? If so one could argue that the pain beliefs mediation effects would only be appropriate in the back pain RCT analysis.<br>6. In the putative mediators section, can additional information be provided on the measurement properties (scale etc), validity and reliability of the physical activity measure and the food frequency measure.<br>7. In the limitation section, it would be important to recognise the limitation of measuring food intake and physical activity by self-report, see and cite Shephard (2003) in the BJSM with regards to |

| | |
|---|---|
| | physical activity in this respect. |

| | |
|---|---|
| **REVIEWER** | Elizabeth Dean<br>University of British Columbia<br>Faculty of Medicine<br>Department of Physical Therapy<br>Canada |
| **REVIEW RETURNED** | 04-Nov-2016 |

| | |
|---|---|
| **GENERAL COMMENTS** | Mechanism evaluation of a lifestyle behavioural intervention for patients with musculoskeletal pain who are overweight or obese. Protocol for a causal mediation analysis (bmjopen-2016-014652)<br><br>Overview<br>Given the prevalence and socioeconomic costs of low back pain (LBP) and knee osteoarthritis (OA) and that obesity has been implicated in both conditions, an established lifestyle intervention with an emphasis on weight loss (i.e., the New South Wales Get Healthy Information and Coaching Service, whose effectiveness has been documented (O'Hara et al, 2012) is being evaluated. The targets of the lifestyle behaviorial intervention are to increase physical activity, improve dietary choices, and address pain beliefs. Based on the findings of two, two-arm RCTs, the investigators propose using established causal mediation analysis (Imai et al, 2013; Imai et al, 2010) to better explicate the mechanisms that underlie the relationship between body weight and LBP and knee OA that may be mediated through weight loss and increased physical activity (specifically, 'why it worked or why it failed'). Inclusion criteria include those overweight individuals with LBP or knee OA awaiting orthopedic consultation.<br><br>Assessment<br>Justification for the study is well reasoned and supported; in fact, such a study is long overdue. Much of the methodology is also well reasoned, but I do have a few queries. Although the trial is underway, I believe their being addressed by the investigators will not negate their findings but could enhance their interpretation and potential usefulness.<br><br>1. Intervention for the patients with LBP and knee OA are not similar in every respect. Paragraph 4 on page 4 describes how participants with LBP are receiving an additional clinical consultation with the study physiotherapist before beginning the NSW Get Healthy Service program. The investigators' decision to include this additional consultation was based on the work of Williams et al, 2016 who reported on a RCT of a lifestyle behavioral intervention for patients specifically with LBP who were overweight/obese. A case could have been made to include this consult for both groups. Even though the study providing the supporting evidence for the decision was for patients with LBP, an argument could have been made (in the same way that the GHS program, designed for people who are overweight/obese in general, was extended to those who also have LBP and knee OA).<br><br>2. Further, the study by Williams et al (2016) that is used to support the additional intervention for the patients with LBP, included 'behavioural change techniques, informed by Self Determination |

| | |
|---|---|
| | Theory (refs 43 and 44)..'. These techniques '…aimed to develop autonomous motivation by increased perceived competence and self-regulation (ref 44)' are not a trivial addition to the LBP group. In my view, this is an important point, particularly in a study focusing on health behaviour change. Autonomous motivation and increased competence are related to self-efficacy, which is core to effective health behaviour change. I recommend including a self-efficacy measure and an autonomy measure (given effective self-management is predicated on perceived control and autonomy). Self-management is very much a 'western' value, less shared by individuals with more externally determined and fatalistic orientations and world views.<br><br>3. Related to point 2, Australia is a culturally and ethnically diverse country. In our work related to health behaviour change across countries and cultures (within and among countries), we attempt to capture these influences. Simply, recording 'language spoken at home' and to self-identify whether one is an 'Indigenous' person (as per the GHS program intake information) is likely insufficient to capture an individual's sociocultural context which is reflected in attributes such as self-efficacy (according to Bandura, 1977) perceived importance of a given health behavior change such as physical activity, and perceived confidence to effect such a change). Although there are 'general' self-efficacy scales, there is one strictly for exercise self-efficacy, which could be distinct to dietary change self-efficacy.<br><br>4. Lastly, with respect to studies of conditions with inflammatory components and dietary and exercise interventions, I believe it is important to distinguish diets for weight loss and diets for their anti-inflammatory qualities, e.g., the Mediterranean diet vs. the pro-inflammatory standard western diet. Much has been written on this topic and we have reported on the use of anti-inflammatory dietary regimen coupled with the anti-inflammatory qualities of reduced sitting and judicious physical activity, as first line interventions to address chronic systemic low-grade inflammation associated with these adverse lifestyle choices/attributes. Obesity itself is a pro-inflammatory state and whether this explains the lower pain threshold of people who are overweight/obese (at least men) is unclear. In terms of the nutritional regimen of the GHS or its underlying principles, little seems to be written.<br><br>5. Funding and trial registration details are shown. Although I did not see a quality control checklist for this trial appended, the trial as written appears to have all the core elements.<br><br>In sum, the investigators have given considerable thought to this study. I hope that my additional arm's length reflections are useful to them in augmenting their work and potentially the interpretation of the findings. I look forward to reading the final published results as they will most certainly add to our current understanding. The paradigm based on causal mediation analysis could certainly be extended to a host of other common conditions. Compared with drug and surgical interventions, we need to pay far more attention to the 'effect size' of healthy living practices on the common conditions of the day, particularly in this era of chronic non-communicable diseases and their shared risk factors. This study is a step in the right direction. |

| | |
|---|---|
| **REVIEWER** | Wei Wang<br>Division of Biostatistics, Center for Devices and Radiological Health, Food and Drug Administration<br>USA |
| **REVIEW RETURNED** | 25-Nov-2016 |

| | |
|---|---|
| **GENERAL COMMENTS** | In this manuscript, the authors presented an analysis plan for a mechanism evaluation of a lifestyle behavioral intervention for patients with knee OA and LBP who are overweight or obese. The proposed hypothesis is clear, the causal models in Figure 1 are reasonable and the presented analysis plan and causal analysis methods are appropriate. I do have some comments and requests that I think would improve and clarify the paper.<br>1) Since the outcome and some potential mediators are ordinal variables (e.g. the primary outcome average pain intensity) and not normally distributed variables, please clarify clearly in the data analysis section on Page 14, what are the regression models that will be used for the outcome and the mediators in the “mediation” R package.<br>2) Since the LBP and knee OA are two separate populations (reference 23 and 24), and the intervention is also not consistent for these two trials (the additional physiotherapy consultations exclusively delivered in the LBP trial), I strongly recommend performing the causal mediation analysis on two populations separately.<br>3) Please clearly clarify whether the treatment-mediator interaction term will be included in the outcome model and provide the justification why the interaction term is included or excluded. Of note, even when the treatment-mediator interaction term is not included in the outcome model, the ACME definition still has two different forms as follows, IE(x) = E(Y(x, M(1))) - E(Y(x, M(0))), where x = 0 or 1, and IE(0) and IE(1) may be slightly different, please clarify which version of ACME is calculated in the data analysis.<br>4) In the sample size calculation section on Page 13, please provide the justification why sample size calculation can assume mediation proportion 50%, and treatment-mediator effect size r = 0.5, and mediator outcome effect size r = 0.3, corresponding reference or preliminary data is needed.<br>Some minor comments:<br>1) Please clearly specify the intervention, primary mediator, alternative mediators and outcomes in Table 1 on Page 8. In addition, for the primary putative mediator (weight), it will be only measured 6 months after randomization, please correct Table 1 correspondingly (Table 1 includes 6 weeks weight measurement as well). |

| | |
|---|---|
| **REVIEWER** | Hanhua Liu<br>Dr Hanhua Liu<br>Research Fellow<br>Centre for Epidemiology (Occupational and Environmental Health)<br>School of Health Sciences<br>Faculty of Biology, Medicine and Health<br>The University of Manchester<br>Room C4.20, Ellen Wilkinson Building<br>Oxford Road<br>Manchester M13 9PL |
| **REVIEW RETURNED** | 02-Dec-2016 |

| | |
|---|---|
| **GENERAL COMMENTS** | Manuscript ID: bmjopen-2016-014652<br><br>Title: Mechanism evaluation of a lifestyle behavioural intervention for patients with musculoskeletal pain who are overweight or obese. Protocol for a Causal Mediation Analysis<br><br>1. Overall comments<br><br>It is an interesting manuscript describing a protocol of mechanism evaluation of a lifestyle behavioural intervention for patients with knee osteoarthritis and low back pain who are waiting for orthopaedic consultation.<br><br>This is a generally well-written manuscript with clear rationale for study. The study aims, methods and analysis were well described and the discussion was coherent.<br><br>As requested by the Journal my review and hence comments concentrate around the authors' use of contemporary methods for Causal Mediation Analysis with sensitivity analyses to evaluate the robustness of the estimated mediation effects to violation of sequential ignorability which is a critical assumption required for causal inference in mechanism evaluations.<br><br>2. Specific comments<br><br>Title<br><br>Appropriate.<br><br>BACKGROUND<br>Explaining underlying mechanisms<br>Clearly explained.<br><br>Mechanisms of a lifestyle behavioural intervention that aimed to address weight, diet, physical activity and pain beliefs<br>Clearly described; and the importance of examining the underling mechanisms of a lifestyle intervention for patients with LBP or OA who are obese or overweight is explained.<br><br>Objectives<br>Clearly defined.<br><br>METHOD<br>Design<br>Page 6 line 26: Something missing in "Combined Causal Mediation Analyses of two, two-arm RCTs.23,24"?<br><br>Page 6 line 35-45: The authors are recommended to expand their methods (especially with more technical description provided) in "Thus it is plausible that the two different clinical populations may respond differentially to their respective interventions. To accommodate this hypothesis, we will use moderated Causal Mediation Analysis to estimate trial-specific effects, and estimate averaged effects across both trials. If trial assignment is a significant moderator, we will interpret the mediation effects in separation; however, if trial assignment is not a significant moderator, we will |

| | |
|---|---|
| | interpret averaged mediation effects across both trials."<br><br>Participants and recruitment<br>Fine.<br><br>Randomisation<br>Intervention groups<br>Page 7 line 20-36: Different interviewers were involved, also different (10) individuals involved in the coaching calls. The authors should discuss (1) between interviewer effects (potential bias); (2) between (different callers) and within (calls made at different times) caller effects (potential bias), respectively; (3) how such potential biases were approached?<br><br>Page 7 line 38-40: Only participants with LBP received an additional clinical consultation. Linking this to my comment for Page 6 line 35-45, the authors need to describe in detail their method of dealing with this major treatment difference in their analysis of the combined cohort of the participants with LBP and OA.<br><br>Control groups<br>Fine.<br><br>Assessment time points<br>Table 1 indicates that the primary putative mediator will be measured 6 weeks and 6 months after randomisation, but your text says only measure at 6 months "The primary putative mediator (weight) will be measured 6 months after randomisation." So which is correct? The authors are recommended to measure 'weigh' at 6 weeks as well, this is because it will make the mediation analysis and its interpretation more precise.<br><br>Page 8 line 44-53: the authors need to clarify whether measure of the putative mediators will be done in both control and treatment participants or not. It appears measure in both participants (LBP and OA) is doable and so it should be measured in both participants.<br><br>Causal Mediation Analysis<br>Fine and clear.<br><br>Justification for primary and alternative mechanisms<br>The only concern is that the justification did not touch on the major difference with LBP treated participants receiving an additional clinical consultation. This should be incorporated into the discussion of justification here.<br><br>Figure 1: "Dotted lines" are not really dotted?<br><br>Table 2: their stepped mediation modelling methodology should be expanded to (1) briefly reflect on the Baron and Kenny approach (J Pers Soc Psychol 1986;51:1173–82); and (2) get some idea from a causal mediation methodology publication "Dunn G, Emsley R, Liu H, Landau S, Green J, White I, et al. Evaluation and validation of social and psychological markers in randomised trials of complex interventions in mental health: a methodological research programme. Health Technol Assess 2015;19(93)."<br><br>Sample size<br>Clearly described. |

| | |
|---|---|
| | Methodological considerations<br>Page 13 line 27-30: your discussion reflecting on this statement of "the treatment-mediator effect, and the mediator-outcome effect, are not confounded.25" is very important.<br><br>The authors are recommended to consider changing their 'post-treatment' to 'post-randomisation' as the latter is more accurate.<br><br>In addition, please link my comments/recommendations for Table 2 to expand their discussion on methodological considerations.<br><br>Data analysis<br>Please refer to my comments above under<br>- Justification for primary and alternative mechanisms<br>- Methodological considerations<br><br>Other than that, fine.<br><br>Conclusion<br>Fine. |

**VERSION 1 – AUTHOR RESPONSE**

Reviewer: 1
Reviewer Name: Cormac Ryan
Institution and Country: Teeside University, UK

I would like to raise a number of minor issues:

1. In the last line of the introduction section of the abstract the word obese/overweight needs to be inserted before "patients". Same point for 1st bullet point of strengths and limitations section.

Inserted: "obese or overweight"

1st strengths and limitations point deleted (as recommended by editor).

2. In the methods and analysis section it needs to be made clearer which variables are outcomes and which are mediators (primary and alternative).

Revised: "The primary mediator: weight, will be measured at 6 months' post-randomisation; alternative mediators including diet, physical activity, and pain beliefs will be measured 6 weeks' post-randomisation. All outcomes: pain, disability, and quality of life, will be measured 6 months' post-randomisation."

3. The hypothesis stated at the bottom of page 2 (or page 5 of 19) would be very useful in the abstract. (This is a writing structure issue I leave to authors discretion)

Due to word limit, we decided not to include the hypothesis in the abstract. Other features of the protocol had to be prioritised.

5. Randomisation subsection - Can the authors clarify how the patients and investigators are to be

blinded?

Added: “Patients were blind to group allocation by nature of the cohort multiple design. This design offers the intervention and control as part of a routine clinical service, where patients consent to routine data collection. Patients randomised to the intervention group were not aware of the control arm. Likewise, patients randomised to the control group were not aware of the intervention arm. Thus, the patients were not able to discriminate whether the intervention or control was being offered as part of a clinical trial. This reduces the risk of performance bias (how well the participants engage with the intervention). Service providers delivering the intervention were blind to treatment status as they were not aware that patients were being referred from a clinical trial. The outcome assessors did not have access to the randomisation schedule, thus were blind to group allocation. This reduces the risk of detection bias (differential outcome measurement between groups).”

6. Pain beliefs only appear to be being targeted in the back intervention - is this correct? If so one could argue that the pain beliefs mediation effects would only be appropriate in the back pain RCT analysis.

Inserted justification for pain beliefs to be tested in the OA trial: “Although patients with OA did not receive a clinical consultation that directly targeted pain beliefs, the Get Healthy Information and Coaching Service may have inadvertently changed pain beliefs through the promotion of physical activity. The physical activity component could enable the patients to realise that pain does not need to be a barrier to keeping a physically active lifestyle. This theory is informed by Albert Bandura’s techniques of verbal persuasion, modelling, and mastery.45”

7. In the putative mediators section, can additional information be provided on the measurement properties (scale etc), validity and reliability of the physical activity measure and the food frequency measure.

Added: “Physical activity will be measured using the Active Australia Survey,50 which has moderate reliability (Cohen’s Kappa = 0.52)51 and good face and criterion validity.52 Dietary intake will be measured using a Short Food Frequency Questionnaire,53 which has moderate reliability (Weighted Kappa range = 0.37 to 0.85)54,55 and criterion validity.55”

8. In the limitation section, it would be important to recognise the limitation of measuring food intake and physical activity by self-report, see and cite Shephard (2003) in the BJSM with regards to physical activity in this respect.

Added: “Putative mediators including diet and physical activity are measured using self-reported questionnaires.”

Added: “Putative mediators are measured using self-reported questionnaires with known limitations.57”

Reviewer: 2
Reviewer Name: Elizabeth Dean
Institution and Country: University of British Columbia, Faculty of Medicine, Department of Physical

Therapy, Canada

1. Intervention for the patients with LBP and knee OA are not similar in every respect. Paragraph 4 on page 4 describes how participants with LBP are receiving an additional clinical consultation with the study physiotherapist before beginning the NSW Get Healthy Service program. The investigators' decision to include this additional consultation was based on the work of Williams et al, 2016 who reported on a RCT of a lifestyle behavioral intervention for patients specifically with LBP who were overweight/obese. A case could have been made to include this consult for both groups. Even though the study providing the supporting evidence for the decision was for patients with LBP, an argument could have been made (in the same way that the GHS program, designed for people who are overweight/obese in general, was extended to those who also have LBP and knee OA).

Reviewers 3 and 4 raised similar concerns. Please see responses to their comments.

2. Further, the study by Williams et al (2016) that is used to support the additional intervention for the patients with LBP, included 'behavioural change techniques, informed by Self Determination Theory (refs 43 and 44)..'. These techniques '…aimed to develop autonomous motivation by increased perceived competence and self-regulation (ref 44)' are not a trivial addition to the LBP group. In my view, this is an important point, particularly in a study focusing on health behaviour change. Autonomous motivation and increased competence are related to self-efficacy, which is core to effective health behaviour change. I recommend including a self-efficacy measure and an autonomy measure (given effective self-management is predicated on perceived control and autonomy). Self-management is very much a 'western' value, less shared by individuals with more externally determined and fatalistic orientations and world views.

We agree that both self-efficacy and autonomy are important process variables that could provide insight into the success/failure of this behavioural intervention. Unfortunately, as the trial is underway, we cannot include these measures.

3. Related to point 2, Australia is a culturally and ethnically diverse country. In our work related to health behaviour change across countries and cultures (within and among countries), we attempt to capture these influences. Simply, recording 'language spoken at home' and to self-identify whether one is an 'Indigenous' person (as per the GHS program intake information) is likely insufficient to capture an individual's sociocultural context which is reflected in attributes such as self-efficacy (according to Bandura, 1977) perceived importance of a given health behavior change such as physical activity, and perceived confidence to effect such a change). Although there are 'general' self-efficacy scales, there is one strictly for exercise self-efficacy, which could be distinct to dietary change self-efficacy.

Please see response to previous comment.

4. Lastly, with respect to studies of conditions with inflammatory components and dietary and exercise interventions, I believe it is important to distinguish diets for weight loss and diets for their anti-inflammatory qualities, e.g., the Mediterranean diet vs. the pro-inflammatory standard western diet. Much has been written on this topic and we have reported on the use of anti-inflammatory dietary regimen coupled with the anti-inflammatory qualities of reduced sitting and judicious physical activity, as first line interventions to address chronic systemic low-grade inflammation associated with these adverse lifestyle choices/attributes. Obesity itself is a pro-inflammatory state and whether this

explains the lower pain threshold of people who are overweight/obese (at least men) is unclear. In terms of the nutritional regimen of the GHS or its underlying principles, little seems to be written.

5. Funding and trial registration details are shown. Although I did not see a quality control checklist for this trial appended, the trial as written appears to have all the core elements.

If the reviewer is referring to a type of reporting guideline, we did not include this because there are no accepted reporting guidelines for mechanism studies.

Reviewer: 3
Reviewer Name: Wei Wang
Institution and Country: Division of Biostatistics, Center for Devices and Radiological Health, Food and Drug Administration, USA

Since the outcome and some potential mediators are ordinal variables (e.g. the primary outcome average pain intensity) and not normally distributed variables, please clarify clearly in the data analysis section on Page 14, what are the regression models that will be used for the outcome and the mediators in the "mediation" R package.

Revised: "Continuous mediators and outcomes that are normally distributed will be modelled using linear models (lm); but if skewed, they will be modelled using generalised linear models (glm) with appropriate family and link functions.70 The ordinal mediator (diet) will be modelled using the proportional odds logistic model (polr).69"

Since the LBP and knee OA are two separate populations (reference 23 and 24), and the intervention is also not consistent for these two trials (the additional physiotherapy consultations exclusively delivered in the LBP trial), I strongly recommend performing the causal mediation analysis on two populations separately.

We agree with the hypothesis that mediation effects could differ across the two populations for the aforementioned reasons. Thus we proposed a moderated mediation analysis to test this hypothesis, and pre-specified the effects we would interpret – dependent on the outcome of the moderation effect:

"…we will use moderated Causal Mediation Analysis to estimate trial-specific effects, and averaged effects across both trials. If trial assignment (LBP trial vs OA trial) is a significant moderator, we will interpret trial-specific mediation effects in separation; however, if trial assignment is not a significant moderator, we will interpret the averaged mediation effects across both trials."

The technical details have been added to the analysis section:

"Trial assignment (OA trial vs LBP trial) could moderate indirect and direct effects. Therefore, we will test the moderating effect of trial assignment by using the test.modmed function. This function directly tests the difference in the ACME and ADE between two levels of the hypothesised moderator (OA trial vs LBP trial). If the ACME or ADE are statistically different, we will analyse the two trials separately to

estimate the ACME and ADE that are specific to each trial. However, if they are not different, we will estimate an averaged ACME and ADE across both trials."

Please clearly clarify whether the treatment-mediator interaction term will be included in the outcome model and provide the justification why the interaction term is included or excluded. Of note, even when the treatment-mediator interaction term is not included in the outcome model, the ACME definition still has two different forms as follows, IE(x) = E(Y(x, M(1))) - E(Y(x, M(0))), where x = 0 or 1, and IE(0) and IE(1) may be slightly different, please clarify which version of ACME is calculated in the data analysis.

Clarified: "Because it is plausible that the indirect and direct effect sizes might depend on treatment allocation, we will include a treatment-mediator interaction term in the outcome model. We will calculate two separate ACMEs that are conditional on treatment status (x=1 and x=0), and their marginal effects. We will interpret each conditional effect to generalise to their respective treatment group (treated and non-treated) and the marginal effect to generalise to the overall population.

In the sample size calculation section on Page 13, please provide the justification why sample size calculation can assume mediation proportion 50%, and treatment-mediator effect size r = 0.5, and mediator outcome effect size r = 0.3, corresponding reference or preliminary data is needed.

Justification added: "The sample sizes for both trials were primarily estimated to detect the main effect of the intervention on pain and weight. Therefore, this post-hoc power calculation provides indication that both trials would be powered to detect an indirect effect that consists of moderate treatment-mediator, and mediator-outcome effects. Moderate effects would be considered clinically meaningful effects based on previous work.67,68"

Some minor comments:
1) Please clearly specify the intervention, primary mediator, alternative mediators and outcomes in Table 1 on Page 8. In addition, for the primary putative mediator (weight), it will be only measured 6 months after randomization, please correct Table 1 correspondingly (Table 1 includes 6 weeks weight measurement as well).

Both tables have been revised.

Reviewer: 4
Reviewer Name: Hanhua Liu
Institution and Country: Dr Hanhua Liu, Research Fellow, Centre for Epidemiology (Occupational and Environmental Health), School of Health Sciences, Faculty of Biology, Medicine and Health, The University of Manchester, Room C4.20, Ellen Wilkinson Building, Oxford Road, Manchester M13 9PL

1. Overall comments

It is an interesting manuscript describing a protocol of mechanism evaluation of a lifestyle behavioural intervention for patients with knee osteoarthritis and low back pain who are waiting for orthopaedic consultation.

This is a generally well-written manuscript with clear rationale for study. The study aims, methods and

analysis were well described and the discussion was coherent.

As requested by the Journal my review and hence comments concentrate around the authors' use of contemporary methods for Causal Mediation Analysis with sensitivity analyses to evaluate the robustness of the estimated mediation effects to violation of sequential ignorability which is a critical assumption required for causal inference in mechanism evaluations.

2. Specific comments
METHOD
Design
Page 6 line 26: Something missing in "Combined Causal Mediation Analyses of two, two-arm RCTs.23,24"?

Revised: "We will conduct a Combined Causal Mediation Analyses of two, two-arm RCTs"

Page 6 line 35-45: The authors are recommended to expand their methods (especially with more technical description provided) in "Thus it is plausible that the two different clinical populations may respond differentially to their respective interventions. To accommodate this hypothesis, we will use moderated Causal Mediation Analysis to estimate trial-specific effects, and estimate averaged effects across both trials. If trial assignment is a significant moderator, we will interpret the mediation effects in separation; however, if trial assignment is not a significant moderator, we will interpret averaged mediation effects across both trials."

The technical details have been added to the analysis section (repeated response to Reviewer 3):

"Trial assignment (OA trial vs LBP trial) could moderate indirect and direct effects. Therefore, we will test the moderating effect of trial assignment by using the test.modmed function. This function directly tests the difference in the ACME and ADE between two levels of the hypothesised moderator (OA trial vs LBP trial). If the ACME or ADE are statistically different, we will analyse the two trials separately to estimate the ACME and ADE that are specific to each trial. However, if they are not different, we will estimate an averaged ACME and ADE across both trials."

Randomisation
Intervention groups
Page 7 line 20-36: Different interviewers were involved, also different (10) individuals involved in the coaching calls. The authors should discuss (1) between interviewer effects (potential bias); (2) between (different callers) and within (calls made at different times) caller effects (potential bias), respectively; (3) how such potential biases were approached?

Revised: "This service consists of 10 individually tailored coaching calls delivered by university qualified health coaches, including dieticians, exercise physiologists, and psychologists, over a 26-week period. All coaches undergo standardised training before delivering the GHS, thus reducing the potential for bias introduced through between coach effects. Coaching was provided on a tapered schedule. Six calls were made in the first 12 weeks to guide, monitor and improve uptake; and 4 calls were dispersed over the remaining 12 weeks to maintain adherence and avoid relapse. This tapered

schedule was kept consistent across all participants, reducing the potential for bias."

Page 7 line 38-40: Only participants with LBP received an additional clinical consultation. Linking this to my comment for Page 6 line 35-45, the authors need to describe in detail their method of dealing with this major treatment difference in their analysis of the combined cohort of the participants with LBP and OA.

Please see response to second comment (Page 6 line 35-45).

Assessment time points
Table 1 indicates that the primary putative mediator will be measured 6 weeks and 6 months after randomisation, but your text says only measure at 6 months "The primary putative mediator (weight) will be measured 6 months after randomisation." So which is correct? The authors are recommended to measure 'weigh' at 6 weeks as well, this is because it will make the mediation analysis and its interpretation more precise.

There was a typo in Table 1 – corrected in response to reviewer 1.

Weight was planned to be measured at 6 months because 6 weeks does not allow for sufficient time for clinically meaningful changes in weight. Although we agree that the temporal precedence of the mediation analysis would be more precise if weight was measured at 6 weeks, we chose to keep to the clinical relevance of this mechanism evaluation and measured weight at 6 months. In effect, this introduces a limitation in our analysis – which has been outlined in the limitations section:

"The primary mechanism (weight) and the outcomes will be captured at the same time-point. Thus, it will be challenging to test the possibility of reverse causation of the mediator-outcome effect."

Page 8 line 44-53: the authors need to clarify whether measure of the putative mediators will be done in both control and treatment participants or not. It appears measure in both participants (LBP and OA) is doable and so it should be measured in both participants.

Inserted: "All putative mediators are measured in both control and intervention groups in both trials."

Causal Mediation Analysis
Fine and clear.

Justification for primary and alternative mechanisms
The only concern is that the justification did not touch on the major difference with LBP treated participants receiving an additional clinical consultation. This should be incorporated into the discussion of justification here.

Added: "Finally, we hypothesise that the intervention may also exert its effect through changes in pain beliefs.39,63 This is because initial consultations in the LBP trial23 aimed to reassure patients and re-

frame erroneous beliefs about pain. Although patients with OA did not receive a clinical consultation that directly targeted pain beliefs, the Get Healthy Information and Coaching Service may have inadvertently changed pain beliefs through the promotion of physical activity. The physical activity component could enable the patients to realise that pain does not need to be a barrier to keeping a physically active lifestyle. This theory is informed by Albert Bandura's techniques of verbal persuasion, modelling, and mastery.64"

Figure 1: "Dotted lines" are not really dotted?

We have changed them to 'red' lines to improve clarity.

Table 2: their stepped mediation modelling methodology should be expanded to (1) briefly reflect on the Baron and Kenny approach (J Pers Soc Psychol 1986;51:1173–82); and (2) get some idea from a causal mediation methodology publication "Dunn G, Emsley R, Liu H, Landau S, Green J, White I, et al. Evaluation and validation of social and psychological markers in randomised trials of complex interventions in mental health: a methodological research programme. Health Technol Assess 2015;19(93)."

Our stepped approach does indeed reflect elements of the above papers. Thus, we have included these citations with the following sentence:

"These methods are an extension of the traditional methods (Baron and Kenny)69 and reflects contemporary advances in Causal Mediation Analysis.60"

Sample size
Clearly described.

Methodological considerations
Page 13 line 27-30: your discussion reflecting on this statement of "the treatment-mediator effect, and the mediator-outcome effect, are not confounded.25" is very important.

Revised: "It is critical that the treatment-mediator effect, and the mediator-outcome effect, are not confounded.25"

The authors are recommended to consider changing their 'post-treatment' to 'post-randomisation' as the latter is more accurate.

Revised: "post-randomisation"

In addition, please link my comments/recommendations for Table 2 to expand their discussion on methodological considerations.

Our stepped approach does indeed reflect elements of the above papers. Thus, we have included these citations with the following sentence:

“These methods are an extension of the traditional methods (Baron and Kenny)69 and reflects contemporary advances in Causal Mediation Analysis.60”

Data analysis
Please refer to my comments above under
- Justification for primary and alternative mechanisms
- Methodological considerations

All addressed under previous comments.

**VERSION 2 – REVIEW**

| | |
|---|---|
| **REVIEWER** | Cormac Ryan<br>Teesside University, UK |
| **REVIEW RETURNED** | 03-Jan-2017 |

| | |
|---|---|
| **GENERAL COMMENTS** | The authors have appropriately addressed the issues of I have raised primarily around blinding and the measurement properties of the outcomes being collected. I have not made any comments upon the statistical analysis as I feel I do not have the expertise upon which to comment and statistical issues are being addressed by a different reviewer. I wish the authors well with their study and look forward to seeing the results when the work is completed. |

| | |
|---|---|
| **REVIEWER** | Elizabeth Dean<br>University of British Columbia<br>Canada |
| **REVIEW RETURNED** | 30-Dec-2016 |

| | |
|---|---|
| **GENERAL COMMENTS** | Mechanism evaluation of a lifestyle behavioural intervention for patients with musculoskeletal pain who are overweight or obese: Protocol for a causal mediation analysis (bmjopen-2016-014652.R1)<br><br>The investigators have done a satisfactory job of not only addressing my comments but those of the other reviewers.<br><br>The investigators acknowledged that ‘self-efficacy and autonomy are important process variables through could provide insights into the success/failure of this behavioural intervention’. However, because the trial is underway, they acknowledge (appropriately) they could not include these variables that have been well documented in the literature to impact lifestyle behavior. However, it would be helpful to future investigators who wish to replicate and/or extend this trial, that this could be acknowledged in the study’s limitations. Refer to my earlier comments re the specific issues.<br><br>Further to the point that the trial is underway and this protocol has been prepared and edited at several points along the way, marked disagreement in the verb tenses throughout (future, present and past) has resulted. The verb tenses need to be congruent with the trial being on-going although the analysis, for example, would be future tense. |

| | |
|---|---|
| | Editorial<br>Re Causal Mediation Analysis. Does not need to be capitalized within the text.<br>I was unable to review the figures as their resolution was too poor. |

| | |
|---|---|
| **REVIEWER** | Wei Wang<br>Division of Biostatistics, Center for Devices and Radiological Health, Food and Drug Administration, USA |
| **REVIEW RETURNED** | 29-Dec-2016 |

| | |
|---|---|
| **GENERAL COMMENTS** | My concerns were adequately addressed in the revised manuscript. |

**VERSION 2 – AUTHOR RESPONSE**

Reviewers' Comments to Author:
Reviewer: 3
Reviewer Name: Wei Wang
Institution and Country: Division of Biostatistics, Center for Devices and Radiological Health, Food and Drug Administration, USA
Please state any competing interests or state 'None declared': None declared

My concerns were adequately addressed in the revised manuscript.

Reviewer: 2
Reviewer Name: Elizabeth Dean
Institution and Country: University of British Columbia, Canada
Please state any competing interests or state 'None declared': None declared.

The investigators have done a satisfactory job of not only addressing my comments but those of the other reviewers.

The investigators acknowledged that 'self-efficacy and autonomy are important process variables through could provide insights into the success/failure of this behavioural intervention'. However, because the trial is underway, they acknowledge (appropriately) they could not include these variables that have been well documented in the literature to impact lifestyle behavior. However, it would be helpful to future investigators who wish to replicate and/or extend this trial, that this could be acknowledged in the study's limitations. Refer to my earlier comments re the specific issues.

Response:
Although we share this view with the reviewer, we feel that the exclusion of 'self-efficacy' as a process variable would rather be a limitation to the results paper, not our current analysis protocol. The limitations stated in this paper refer directly to the analysis per se (e.g. self-reported outcomes). Broader limitations such as the one raised here will be included in our results paper.

Further to the point that the trial is underway and this protocol has been prepared and edited at several points along the way, marked disagreement in the verb tenses throughout (future, present and past) has resulted. The verb tenses need to be congruent with the trial being on-going although the analysis, for example, would be future tense.

Response:

Entire manuscript revised.

Editorial
Re Causal Mediation Analysis. Does not need to be capitalized within the text.
Revised.

Reviewer: 1
Reviewer Name: Cormac Ryan
Institution and Country: Teesside University, UK
Please state any competing interests or state ‘None declared’: None declared

The authors have appropriately addressed the issues of I have raised primarily around blinding and the measurement properties of the outcomes being collected. I have not made any comments upon the statistical analysis as I feel I do not have the expertise upon which to comment and statistical issues are being addressed by a different reviewer. I wish the authors well with their study and look forward to seeing the results when the work is completed.