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)	Effect of motion control versus neutral walking footwear on pain associated with lateral tibiofemoral joint osteoarthritis: a comparative effectiveness randomized clinical trial.
AUTHORS	Paterson, Kade; Bennell, Kim; Metcalf, Ben; Campbell, P; McManus, Fiona; Lamb, Karen; Hinman, Rana

VERSION 1 – REVIEW

REVIEWER	Mike Frecklington Auckland University of Technology
REVIEW RETURNED	10-May-2022

GENERAL COMMENTS	Thank you for the opportunity to review this manuscript. This trial is well-designed and clearly reported. Despite not finding differences between the two treatment arms, reasons for this are discussed alongside considerations for future trials. The biomechanical measures outlined in the protocol document would also be of interest, although I can appreciate these may be presented in a subsequent manuscript.

REVIEWER	Cynthia Coffman Duke University Medical Center, Biostatistics and Bioinformatics
REVIEW RETURNED	31-May-2022

GENERAL COMMENTS	• This is a very well written paper • Given the small recruited sample size, the subgroup analysis/ moderator analysis is not warranted - even though it was planned is not really interpretable. • A few more details on the instrumental variable approach accounting for adherence is needed to understand the validity of this analysis with the smaller sample size and details on the instrumental variable used. Clarify if Table 4 in appendix is using instrumental variable approach - in discussion any limitations to the analysis used for this should be noted • In results, seems like Table 3 and 4 could be combined – not sure necessary to have all that data in two separate tables for the outcomes

VERSION 1 – AUTHOR RESPONSE

Reviewer 1:

1. Thank you for the opportunity to review this manuscript. This trial is well-designed and clearly reported. Despite not finding differences between the two treatment arms, reasons for this are discussed alongside considerations for future trials. The biomechanical measures outlined in the

protocol document would also be of interest, although I can appreciate these may be presented in a subsequent manuscript.

RESPONSE: Thank you for the positive feedback. That is correct, we did collect some data on immediate changes in biomechanics with the study footwear. However, these outcomes were not collected as measures of treatment efficacy (as per prospective trial registration), which is why they weren't included in the manuscript. As the reviewer suggests, we intended to report this data in a subsequent biomechanics manuscript.

ACTION: None

Reviewer 2:

2. Given the small recruited sample size, the subgroup analysis/ moderator analysis is not warranted - even though it was planned is not really interpretable.

RESPONSE: Thank you for this comment. Even if 110 participants had been recruited, the sample would only have been adequately (90%) powered for the primary efficacy analysis. The moderator analyses were always going to be exploratory rather than conclusive. The smaller than expected sample size did not change this interpretation; that is, the moderator analyses results remain exploratory. In keeping with our pre-specified statistical analysis plan and published trial protocol, we would prefer to report all planned analyses along with effect sizes, confidence intervals, and p values in order to let readers use their own judgment about the relative weight of the conclusions. This approach aligns with the approach favoured by the American Statistical Association [1]. We do not believe that inclusion of the moderator analyses detracts from the open and transparent reporting of our trial.

ACTION: None.

3. A few more details on the instrumental variable approach accounting for adherence is needed to understand the validity of this analysis with the smaller sample size and details on the instrumental variable used. Clarify if Table 4 in appendix is using instrumental variable approach - in discussion any limitations to the analysis used for this should be noted

RESPONSE: This secondary sensitivity analysis, using the instrumental variable approach, estimated the complier average causal effect (CACE). The CACE method indirectly compares participants in the motion control shoe group with a similar (sub)group of participants from the neutral walking shoe group - those who would have adhered to wearing the motion control shoe had they been randomized to the motion control shoe group - in order to obtain a valid estimate of the effect of full adherence to wearing the motion control shoe [2]. All participants, as randomized, were included in the analysis. Randomization was the instrumental variable used for adherence. Appendix Table 4 uses the instrumental variable approach. In addition to assuming participants were randomized to each group and did not interact with each other, this analysis assumed that not wearing motion control shoes for an average of $\geq$6 hours/day for 6 months was equivalent to wearing neutral shoes. So, to benefit from the motion control shoes, we assumed participants had to wear them for an average of $\geq$6 hours/day for 6 months. The analysis also assumed that being assigned to the motion control shoe group could only increase adherence (and could not decrease adherence) to wearing the motion control shoes.

ACTION: We have included further detail on the instrumental variable approach and clarified that the instrumental variable was randomization in the Statistical analysis section of the Methods:

"A sensitivity analysis, including all participants as randomized, estimated complier average causal effects, which are the treatment effects on the primary outcome assuming full adherence to shoe wear

(classified as average of ≥6 hours/day for 6 months, based on logbook data), using an instrumental variables approach (where randomization was the instrument for adherence) [3]. Two-stage least squares models were fit: first, a model for observed adherence, including terms for randomized group, the outcome at baseline and the stratifying variable (KL grade) and second, a model predicting the primary outcome, given observed adherence."

We have noted in the Discussion the limitations of this analysis:

"We did not identify any between-group differences on the primary outcome when assuming full adherence, however these results assumed that participants had to wear motion control shoes for an average of >6 hours/day for 6 months in order to benefit from them."

We have also clarified in Appendix Table 4, footnote B, that treatment effects on the primary outcome assuming full adherence used an instrumental variable approach:

"The treatment effect on the primary outcome assuming full adherence (where full adherence was defined as an average of ≥ 6 hours/day shoe wear over 6 months) was estimated using an instrumental variables approach (where randomization was the instrument for adherence)."

4. In results, seems like Table 3 and 4 could be combined – not sure necessary to have all that data in two separate tables for the outcomes

RESPONSE: We agree with this suggestion.

ACTION: We have combined Tables 3 and 4 as suggested, removed any reference to Table 4, and renumbered Table 5.