PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Face-to-face physiotherapy compared to a supported home exercise program for the management of musculoskeletal conditions: Protocol of a multicentre, randomised controlled trial - the REFORM trial
AUTHORS	Withers, Hannah; Glinsky, Joanne; Jennings, Matthew; Hayes, Alison; Starkey, Ian; Palmer, Blake; Szymanek, Lukas; Cruwys, Jackson; Wong, David; Duong, Kitty; Barnett, Anne; Tindall, Matthew; Lucas, Barbara; Lambert, Tara; Sherrington, Catherine; Maher, Christopher; Ferreira, Manuela; Taylor, Deborah; Chu, Jackie; Harvey, Lisa

VERSION 1 – REVIEW

REVIEWER	Are Hugo Pripp Oslo Centre of Biostatistics and Epidemiology Research Support Services Oslo University Hospital Norway Faculty of Health Sciences OsloMet - Oslo Metropolitan University Norway
REVIEW RETURNED	25-Aug-2020

GENERAL COMMENTS	The protocol is well described and makes it possible to replicate the study. It is a randomised controlled non-inferiority trial to compare a home exercise programme with face-to-face physiotherapy. The description of the sample size calculation and the non-inferiority analysis is not quite clear to me. Please specify the non-inferiority margin (often denoted delta). Is it 1.5 or 0.75 points? Do you plan to do a one-sided (often used in non-inferiority trials) or a two-sided test to assess non-inferiority. Is the significance level 5%? Do you plan to do a non-inferiority analysis for only the primary outcome at 6 weeks? What about the 26 weeks measurement and the secondary outcomes? There are many secondary outcomes and two follow-up measurements. Do you plan to do any adjustments for multiple comparisons? Due to COVID-19 the recruitment was stopped. If possible, please describe how you want to control for the likely delayed recruitment. Could it affect the analysis?

REVIEWER	Michelle Cottrell Royal Brisbane and Women's Hospital Queensland, Australia
REVIEW RETURNED	28-Aug-2020

GENERAL COMMENTS

Face-to-face physiotherapy compared to a supported home exercise program for the management of musculoskeletal conditions: Protocol of a multicentre, randomised controlled trial - the REFORM trial

Thank you for the opportunity to review this protocol paper that has been submitted to BMJ open for publication. This paper outlines the protocol for an Australian multi-centre, randomised (non-inferiority) controlled trial that commenced recruitment in March 2019. I have accepted this paper for publication with minor revisions and would appreciate if you could address the following comments/questions:

Section:	Comment / Question:
Abstract:	1. Line 8 (Methods) – please remove 'or better' as this protocol describes a non-inferiority trial.
Introduction:	2. Please update references 3 (there is a 2018 updated paper) and 4 (more recent papers put indirect costs around $45B per annum in Australia) 3. Lines 32-36 – for consistency please choose either telehealth or telerehabilitation. Since this paper uses technology to limit direct clinician-patient contact, it best fits the terminology of 'telemonitoring'. It might be worth defining and justifying your choice. 4. Line 44 – 'as effective or better' – again this implies superiority. Please change to 'as good as'.
Methods:	5. No hypothesis has been provided. 6. Page 3, Line12 – suggest changing 'mental illness' to cognitive/intellectual impairment. Please also acknowledge who decides as to whether the potential participant meets the outlined eligibility 7. Consider removing the 'public and patient involvement' section. Third sentence ('All participants for this trial …) can be moved to the below 'recruitment strategy and time frame' section. If deciding to keep this section, please elaborate on the second sentence with references – it reads as though the research team developed the Patient Specific Functional Scale (the primary outcome measure).

		8. 'Recruitment strategy and time frame' – please elaborate on how patients from the 4 different hospitals will be approached by the research team for trial participation. 9. Please acknowledge whether participants are/are not able to access other relevant concomitant care during their trial participation. 10. Outcome measures – Patient Satisfaction: Probably too late now as patient recruitment is underway but consider whether a 'patient experience' measure could be incorporated into this study. In general, patient satisfaction always rates highly (particularly in telehealth studies); patient experience measures are instead more objective and seem to where consumer engagement / feedback is heading at the moment (this might already be considered as part of the process evaluation!). 11. Sample size – please acknowledge the outcome measure (PSFS) that the non-inferiority sample size calculation is being measured against. 12. Statistical plan – will there be any attempts to analysed within-group changes (over 6 or 26 weeks) or group*time changes. Would be good to know that even if non-inferiority is achieved, that both interventions actually produced significant improvements in outcomes. 13. Statistical analysis – how will you account for the impact that the underlying MSK condition will have on PSFS (and the secondary outcomes) change scores? E.g. a patient with severe hip OA would be expected to have much smaller change scores when compared to a simple ankle sprain.
	General comments:	- consider changing 'face-to-face' to 'in-person' as this language in general has changed within the telehealth community over the past couple of years, as you are still technically 'face-to-face' in a VC consult (although this paper is talking about telephone support). - A process evaluation is briefly mentioned in the introduction but there is no elaboration of this within the methods section. Appreciate that several factors are being investigated within this protocol, but I would like to see how this process

		evaluation is going to be carried out. This is imperative as part of a wider implementation study design as I can imagine that even with positive clinical and economic outcomes, translating these findings into the public health sector will be challenging.

VERSION 1 – AUTHOR RESPONSE

Reviewer(s)' Comments to Author

Reviewer #1

Comment 1: The protocol is well described and makes it possible to replicate the study. It is a randomised controlled non-inferiority trial to compare a home exercise programme with face-to-face physiotherapy.

Response: Thank you for your positive comment.

Comment 2: The description of the sample size calculation and the non-inferiority analysis is not quite clear to me. Please specify the non-inferiority margin (often denoted delta). Is it 1.5 or 0.75 points?

Response: This suggestion has been adopted. We have rewritten this section to better explain the sample size calculation. The new text states:

> *"A sample size of 210 people is required to provide 80% power for a non-inferiority margin (delta) of -1.5 points (where a positive between-group difference favours the Experimental intervention) assuming a 15% loss to follow up, a standard deviation of 2 (19), a 15% treatment dropout rate, a correlation between baseline and final scores of 0.5 and a conservative estimate that the between-group difference favours the control group by 0.75 points."*

Comment 3: Do you plan to do a non-inferiority analysis for only the primary outcome at 6 weeks? What about the 26 weeks measurement and the secondary outcomes?

Response: We do not intend to set non-inferiority margins for anything other than the primary outcome at 6 weeks and we will not be dichotomising results as significant or not significant. Instead, the results of all other outcomes will be presented as point estimates and 95% (not p values). These results will only be used to help the interpretation of the primary analysis. We have added text to clarify this issue. The new text states:

> ***"Other analyses:*** *The results of all other analyses will be presented as point estimates (with 95% CI) and will not be interpreted with respect to non-inferiority margins (deltas) or statistical significance but instead used to aid the interpretation of the results of the non-inferiority analysis of the primary outcome at 6 weeks*."

Comment 4: There are many secondary outcomes and two follow-up measurements. Do you plan to do any adjustments for multiple comparisons?

Response: We will not do adjustments for multiple comparisons however we only intend to present the point estimates (and 95% CIs) and use these findings to aid the interpretation of the results of the non-inferiority analysis of the primary outcome at 6 weeks. Our interpretation will be cautious taking into account the number of outcomes and the two endpoints. We have added text to clarify this issue. The new text states:

> *"The results of all other analyses will be presented as point estimates (with 95% CI) and will not be interpreted with respect to non-inferiority margins (deltas) or statistical significance but instead used to aid the interpretation of the results of the non-inferiority analysis of the primary outcome at 6 weeks. We will not make any adjustments for multiple comparisons however we will interpret these findings cautiously taking into account the number of outcomes and the two endpoints."*

Comment 5: Due to COVID-19 the recruitment was stopped. If possible, please describe how you want to control for the likely delayed recruitment. Could it affect the analysis?

Response: It is not possible to "control" for the pause in recruitment and it is not known whether this interruption will affect the results. We will however discuss this possibility in our definitive paper but we don't feel it is appropriate to add text to the protocol regarding this issue. If the Editor feels strongly that we should then we will reconsider.

Reviewer #2

Comment 1: Line 8 (Methods) – please remove 'or better' as this protocol describes a non-inferiority trial.

Response: This suggestion has not been adopted because we are doing a non-inferiority trial, not an equivalence trial. So, the phrase "*or better*" is very important.

Comment 2: Please update references 3 (there is a 2018 updated paper) and 4 (more recent papers put indirect costs around $45B per annum in Australia)

Response: Thank you for this suggestion. We have updated these references and the text. The new text states:

> *"In 2008-9, costs attributed to musculoskeletal conditions were an estimated $5.7 billion"*

Comment 3. Lines 32-36 – for consistency please choose either telehealth or telerehabilitation. Since this paper uses technology to limit direct clinician-patient contact, it best fits the terminology of 'telemonitoring'. It might be worth defining and justifying your choice.

Response: This suggestion has been adopted. We have used the term 'telerehabilitation' throughout and provided a justification. We have not used the term "telemonitoring" because our intervention was more than just monitoring. The new text states:

> *"Telerehabilitation has enabled physiotherapists to continue to provide services to some of the many patients requiring physiotherapy thereby potentially preventing the escalation of symptoms and presentation to emergency departments at a time of burden for the health system."*

Comment 4. Line 44 – 'as effective or better' – again this implies superiority. Please change to 'as good as'

Response: This suggestion has not been adopted. Please see response above.

Comment 5: No hypothesis has been provided.

Response: This suggestion has not been adopted because according to the author guidelines of BMJ Open for protocols, a hypothesis is not required.

Comment 6. Page 3, Line12 – suggest changing 'mental illness' to cognitive/intellectual impairment. Please also acknowledge who decides as to whether the potential participant meets the outlined eligibility.

Response: This suggestion has been not been adopted because we feel that the term mental illness is not the same as an intellectual or cognitive impairment and we are unable to change the exclusion criteria mid-trial. We have changed the text to clearly acknowledge who decides as to whether the potential participant meets the outlined eligibility. The new text states:

> *"...has a mental illness which may affect adherence to the trial protocol. This will be determined in consultation with the treating physiotherapists and with review of past medical history."*

Comment 7. Consider removing the 'public and patient involvement' section. Third sentence ('All participants for this trial …) can be moved to the below 'recruitment strategy and time frame' section.

Response: This suggestion has not been adopted because it is a requirement of *BMJ Open* (as per the author guidelines). We will keep this section but will adopt the suggestion to elaborate on the second sentence.

Comment 8. If deciding to keep this section [recruitment strategy and time frame'], please elaborate on the second sentence with references – it reads as though the research team developed the Patient Specific Functional Scale (the primary outcome measure).

Response: This suggestion has been adopted. We have changed the text to ensure it does not imply that the team developed the PSFS. The next text states:

> *"The primary outcome measure was developed in 1995 (16) with input from patients."*

Comment 8. 'Recruitment strategy and time frame' – please elaborate on how patients from the 4 different hospitals will be approached by the research team for trial participation.
Response: This suggestion has been adopted. The new text states:

> *"Potential participants will be screened according to the inclusion/exclusion criteria from the waiting list of each outpatient physiotherapy department. This process will be completed by either the treating physiotherapists or admin staff of the department over the telephone. If appropriate, patients will be given an appointment to attend the outpatient department to complete the consent, baseline assessment and randomisation."*

Comment 9. Please acknowledge whether participants are/are not able to access other relevant concomitant care during their trial participation.
Response: This suggestion has been adopted. The new text states:

> *"Participants in both groups are permitted to continue with any concomitant treatments for any co morbidities. They are asked to not have any other physiotherapy for their musculoskeletal condition in addition to what is provided by the treating therapists of both groups for the 6-week intervention period."*

Comment 10. Outcome measures – Patient Satisfaction: Probably too late now as patient recruitment is underway but consider whether a 'patient experience' measure could be incorporated into this study. In general, patient satisfaction always rates highly (particularly in telehealth studies); patient experience measures are instead more objective and seem to where consumer engagement / feedback is heading at the moment (this might already be considered as part of the process evaluation!)
Response: Thank you for this suggestion. We are conducting an in-depth process evaluation to explore patients' experiences.

Comment 11. Sample size – please acknowledge the outcome measure (PSFS) that the non-inferiority sample size calculation is being measured against.
Response: Unfortunately, we are not sure what the reviewer means with this comment. We have adjusted the text for this section and hope that this has addressed this comment but if not please let us know what is meant and we will address if possible.

Comment 12. Statistical plan – will there be any attempts to analysed within-group changes (over 6 or 26 weeks) or group*time changes. Would be good to know that even if non-inferiority is achieved, that both interventions actually produced significant improvements in outcomes.
Response: We will not be doing within-group analyses because we do not believe that these analyses are meaningful for this trial. Our planned analyses using regression models to estimate between-group differences adjusting for baseline variables are essentially group*time analyses.

Comment 13. Statistical analysis – how will you account for the impact that the underlying MSK condition will have on PSFS (and the secondary outcomes) change scores? E.g. a patient with severe hip OA would be expected to have much smaller change scores when compared to a simple ankle sprain.
Response: It is correct that there will be a lot of variability in the change PSFS scores because of the differing MSK conditions. However, we have accounted for this in our power calculations as our anticipated SD reflects this possible variability. We anticipate that given the relatively large sample size, the randomisation will result in a balance of people with severe and less severe MSK conditions.

Comment 14: consider changing 'face-to-face' to 'in-person' as this language in general has changed within the telehealth community over the past couple of years, as you are still technically 'face-to-face' in a VC consult (although this paper is talking about telephone support).
Response: This suggestion has not been adopted because we have used the terminology "Face-to-face" throughout our protocol, CRF and source documents and we feel that at this stage it is too late to change. As acknowledged, the remote monitoring that takes place in this intervention is over the telephone only, so in this context face-to-face does only refer to in person physiotherapy within the outpatient department.

Comment 15: A process evaluation is briefly mentioned in the introduction but there is no elaboration of this within the methods section. Appreciate that several factors are being investigated within this

protocol, but I would like to see how this process evaluation is going to be carried out. This is imperative as part of a wider implementation study design as I can imagine that even with positive clinical and economic outcomes, translating these findings into the public health sector will be challenging.
Response: Thank you for your comment. We are preparing a protocol paper specifically for the process evaluation. This is a large piece of work and can't be adequately described in this paper.

VERSION 2 – REVIEW

REVIEWER	Pripp, Are Hugo Oslo universitetssykehus Ulleval, Oslo Centre for Biostatistics & Epidemiology
REVIEW RETURNED	24-Dec-2020

GENERAL COMMENTS	The statistical analysis is clearly described.

REVIEWER	Cottrell, Michelle Royal Brisbane and Women's Hospital
REVIEW RETURNED	17-Dec-2020

GENERAL COMMENTS

Thank you for the opportunity to review this protocol paper. I have attached a separate document with some comments and feedback. I hope that the trial is coming back on track with the lifting of social distancing restrictions - good luck for the remainder of the study. Cheers, Michelle

Abstract:	'is as good as or better than face-to-face …'. Remove the 'better than' as this not what a non-inferiority trial meaures. Please incorporate the final two sentences (where the results and conclusions headers have been removed) into the into the introduction /methods sections as relevant.
Introduction:	Some of the statements in the first paragraph need to be referenced. Final paragraph (aim) – the aim is to determine if the intervention is as good as the control. Please remove the 'better than' as per comment above.
Recruitment strategy and timeframe	Its probably not relevant to include the study timeline that is pre-COVID.
Interventions:	Is the amount of additional contact (via text, email, phone, in-person) being

		recorded for the intervention group? This may be of importance when interpreting your clinical outcomes. Previous studies that demonstrate telehealth as superior to standard care generally provide a lot more direct intervention which may be the reason for the results. An important consideration when trying to advocate for a reduced contact intervention.
	Sample Size:	Please be explicit that your sample size is based on the primary outcome measure of the PSFS.
	Data Analysis:	Non-inferiority analysis: Wording around how non-inferiority will be measured needs to be addressed. Please review and reference the CONSORT 2010 Statement for reporting non-inferiority and equivalence trials Between-group comparisons for each outcome are being measured – is this just at the 6 week mark? Are you considering to do within-group change, as well as group-by-time comparisons. I also think within-group statistics are interesting as non-inferiority may be met, but neither group achieved a statistical or clinically significant change. This might feed into your 'Other Analyse' section. It might read better to have the short introduction, then put in the 'non-inferiority analysis' section (this is the primary type of analysis you are doing) and then amalgamate the remaining first paragraph and the 'other analyse' section. I find have am having to re-read this section to piece it together.
	Data Collection:	How will the data collected as part of the economic evaluation be collected?
	Trial monitoring:	Could you give a brief definition / example of what would constitute a serious adverse event

	General:	The proposed process evaluation mentioned in the end of the introduction has not been explained in any detail. It is not particularly clear where the blinding will/will not occur. There is some general typograghical and grammatical errors throughout the paper that require review. Also there is a mixture of tense (past and future) being used that should be addressed. I have stated in the checklist that the supplementary reporting (the SPIRIT checklist) was not completed – this is just because I wasn't sure if you had to outline within the checklist where you met each of the criteria. Generally speaking, the majority of check points were met within the manuscript.

VERSION 2 – AUTHOR RESPONSE

Reviewer #1

Comment 1: Comments to the Author: The statistical analysis is clearly described.
Response: Thank you for your positive comment.

Reviewer #2

Thank you for the opportunity to review this protocol paper. I have attached a separate document with some comments and feedback. I hope that the trial is coming back on track with the lifting of social distancing restrictions - good luck for the remainder of the study. Cheers, Michelle

Comment 1: Abstract: 'is as good as or better than face-to-face …'. Remove the 'better than' as this not what a non-inferiority trial measures.

Response: We would respectfully argue that the reviewer is not correct on this point. A non-inferiority trial does measure whether an alternate treatment is as good or better than the gold standard. This is in contrast with an equivalence trial that only measures whether the alternate treatment is as good. For this reason, it is important that we retain the phrase - "*better than*".

This distinction is clarified by the European Agency for the Evaluation of Medicinal Products (1). They state:

"An equivalent trial is designed to confirm the absence of a meaningful difference between treatments……[A non-inferiority trial is used to determine] that a new treatment is no less effective..- it may be more effective or it may have a similar effect." Pg 310

Our terminology was used as recently as 2020 in an Editorial for Ophthalmology. The editor stated that*:*

"*If a new treatment is as good or better, then its ..[a] noninferiority trial."* (2) Pg. 711

Comment 2: Please incorporate the final two sentences (where the results and conclusions headers have been removed) into the introduction/methods sections as relevant.
Response: This suggestion has been adopted. New text states:

"***Ethics and Dissemination***

Ethical approval was obtained on the 17 March 2017 from the Northern Sydney Local Health District HREC, trial number HREC/16HAWKE/431-RESP/16/287. The results of this study will be submitted for publication to peer-reviewed journals and be presented at national and international conferences. Recruitment commenced in March 2019 and it is anticipated that the trial will be completed by September 2021. This trial will investigate two different models of physiotherapy care for people with musculoskeletal conditions."

Comment 3: Introduction: Some of the statements in the first paragraph need to be referenced.
Response: This suggestion has been adopted. Additional references have been added:

"In 2015 an estimated 30% of all people had at least one musculoskeletal condition in Australia.(1)"

Reference:
Australian Institute of Health and Welfare 2016. Australia's health 2016. Canberra: AIHW.

"This figure is reported to be as high as 72% for people aged over 75(3)"

Reference:
Australian Institute of Health and Welfare. Musculoskeletal conditions and comorbidity in Australia. Canberra: AIHW2019.

Comment 4: Final paragraph (aim) – the aim is to determine if the intervention is as good as the control. Please remove the 'better than' as per comment above.
Response: This suggestion has not been adopted. See response to comment1 above.

Comment 5: Recruitment strategy and timeframe. It's probably not relevant to include the study timeline that is pre-COVID.
Response: This suggestion has been adopted. The recruitment strategy and recruitment time frame have been updated. The new text states:

APPENDIX:

Table 1: Timeline for the study.

Phase	**Objective**	**Planned Completion Date**
Preparation	Finalise protocol Submit to ethics Finalise CRF	From October 2016

	Complete Database	
Recruitment	Commence Recruitment	April 2019
Dissemination	Publish Protocol	December 2020
Recruitment and data collection	Continue recruitment Collect data from 6 week and 26-week assessments Recruit 100% of participants	April 2019 to Dec 2020 Due to COVID -19 recruitment was stopped in March 2020, and will be resumed in January 2021. Currently n=113. Revised planned completion date: August 2021.
Analysis	Clean and lock data base Complete Analysis Submit papers for publication	From December 2021
Dissemination	Present results at seminars, conferences Disseminate results into policy and practice	From December 2021

Comment 6: Interventions: Is the amount of additional contact (via text, email, phone, in-person) being recorded for the intervention group? This may be of importance when interpreting your clinical outcomes. Previous studies that demonstrate telehealth as superior to standard care generally provide a lot more direct intervention which may be the reason for the results. An important consideration when trying to advocate for a reduced contact intervention.
Response: Thank you for your comment. We have added text to clarify that we are collecting details about all additional contact with the intervention group. Please note that there is no additional email contact. We have added the following text to clarify this issue. It states:

> *"Details about all additional text and telephone calls with the Intervention participants will be recorded including the number of text messages and telephone calls, and length of time spent on each telephone call. In addition, records are being kept on the number of failed attempts to contact participants by telephone."*

Comment 7: Sample Size: Please be explicit that your sample size is based on the primary outcome measure of the PSFS.
Response: This suggestion has been adopted. The new text states:

> *"A sample size of 210 people is required to provide 80% power for a non-inferiority margin (delta) of 1.5 points on the primary outcome (PSFS) where a positive between-group difference favours the Supported Home Exercise Group assuming a 15% loss to follow up, a standard deviation of 2(3), a 15% treatment dropout rate, a correlation between baseline and final scores of 0.5 and a conservative estimate that the between-group difference favours the Face-to-Face Group by 0.75 points."*

Comment 8: Between-group comparisons for each outcome are being measured – is this just at the 6-week mark?

Response: Between-group comparisons will be calculated for all outcomes at 6 weeks and 26 weeks. Text has been added to clarify this issue. The new text states:

> *"Between-group comparisons of each outcome will be conducted at 6- and 26-weeks using regression models in which the outcome will be a linear function of a dummy-coded variable representing group membership (Supported Home Exercise Group or Face-to-face Physiotherapy Group) and a dummy-coded variable for stratum, specifically site and duration since onset of injury (less than 12 weeks versus more than 12 weeks)."*

Comment 9: Non-inferiority analysis: Wording around how non-inferiority will be measured needs to be addressed. Please review and reference the CONSORT 2010 Statement for reporting non-inferiority and equivalence trials.

Response: We have reviewed the CONSORT statement. It states:

> "[clarify] the Statistical methods used to compare groups for primary outcome(s), specifying whether a 1- or 2-sided confidence interval approach was used. Methods for additional analyses, such as subgroup analyses and adjusted analyses."

We have addressed this item of the CONSORT statement with the following text (please note that the bolded text clarifies that we are using a 1-sided CI approach):

Current text: ***"****The Supported Home Exercise Group will be considered non-inferior to the Face-to-face Physiotherapy Group* ***if the upper limit of the 95% confidence interval*** *associated with the mean between group difference on the PSFS at 6 weeks indicates that Supported Home Exercise versus face-to-face physiotherapy is either better or no worse than 1.5 points out of 10. The non-inferiority cut-off point of 1.5 was decided by the investigators after taking into consideration the likely implications of this amount of difference on function and the cost of the intervention."*

Comment 10: Are you considering to do within-group change, as well as group-by-time comparisons. I also think within group statistics are interesting as non-inferiority may be met, but neither group achieved a statistical or clinically significant change. This might feed into your 'Other Analyses' section.

Response: This suggestion has not been adopted. We are not doing a within group analysis as stated in our initial rebuttal letter. We do not agree with the reviewers that within-group changes are helpful. We respectfully disagree with the assertion that these types of analyses can be used to indicate that "both interventions produced significant improvements

Comment 11: It might read better to have the short introduction, then put in the 'non-inferiority analysis' section (this is the primary type of analysis you are doing) and then amalgamate the remaining first paragraph and the 'other analyses' section. I find have am having to re-read this section to piece it together.

Response: This suggestion has been adopted. The text has not been altered but has been rearranged as suggested.

Comment 12: Data Collection: How will the data collected as part of the economic evaluation be collected?

Response: This suggestion has been adopted. The new text states:

> *"Data for the economic evaluation will be collected over the telephone by the trial physiotherapist after the 6-week blinded assessment has been completed"*

Comment 13: Trial monitoring: Could you give a brief definition / example of what would constitute a serious adverse event.

Response: We are using the standard definition of a serious adverse event, and the new text states:

> *"All SAE from the time of randomisation to the 26-week assessment will be recorded. These will include any events that result in death, disability, hospitalisation or prolongs existing hospitalisation. An SAE can occur at any time from randomisation until completion of the 26-week assessment."*

Comment 14: General: The proposed process evaluation mentioned in the end of the introduction has not been explained in any detail.
Response: We have not explained the process evaluation because we are preparing a protocol paper specifically for the process evaluation. This is a large piece of work and can't be adequately described in this paper. We have added text to clarify this. The new text states:

> *"A secondary process evaluation will also explore participants' opinions and experiences of the intervention and trial. A separate manuscript is being prepared to explain the protocol for the process evaluation."*

Comment 15: It is not particularly clear where the blinding will/will not occur.
Response: This suggestion has been adopted. The new text states:

> *"Regardless of the method used to collect the data, the assessor responsible for interacting with the participant and/or collecting the data over the telephone, will be blinded to the allocation of the participants."*

Comment 16: There is some general typographical and grammatical errors throughout the paper that require review. Also there is a mixture of tense (past and future) being used that should be addressed.
Response: Thank you for your comment. We have checked carefully and have made any necessary corrections to grammar and tense. We were aware that the tenses were a mix of past and present. This was necessary to reflect the status of the trial to date. For example, the randomisation schedule has already been created. However, to avoid confusion we have used future tenses throughout and just clarified the current status of the trial.

Comment 17: I have stated in the checklist that the supplementary reporting (the SPIRIT checklist) was not completed – this is just because I wasn't sure if you had to outline within the checklist where you met each of the criteria. Generally speaking, the majority of check points were met within the manuscript.
Response: This suggestion has been adopted. A SPIRIT checklist has been uploaded and the page number provided to indicate where the text that addresses each criterion can be found.

References

1. Products Committee for Proprietary Medicinal. Points to consider on switching between superiority and non-inferiority. *British journal of clinical pharmacology*. 2001;52:223-28.
2. Thomas M Lietman, Oldenburg Catherine E, O'Brien Kieran S. Noninferiority: It's All in the Margins. *Ophthalmology*. 2020;127:711-12.
3. K. Kowalchuk Horn, Jennings S, Richardson G, et al. The patient-specific functional scale: Psychometrics, clinimetrics, and application as a clinical outcome measure. *The Journal of Orthopaedic and Sports Physical Therapy*. 2012;42:30-42.