Peer-Review comments and authors responses

"Design Considerations for Clinical Trials Combining Music Therapy and Physical Activity in Alzheimer's Disease: Lessons from designing a Multicenter Randomized Protocol"

Dear Faculty and Editor,

The authors thank the coaching meetings professors for their commitment during the revision of the paper, for the time spent and the insightful comments and suggestions. Clearly the paper is much better after the modifications performed.

Below are the answers to the questions. The modifications performed in the main manuscript are highlighted in track changes in the main text.

Regards,

Reviewer 1

INTRODUCTION

1. **Comment:** Data is not fully overviewed.

Response: *We modified the introduction section. We adjusted the Background and Existing Knowledge sections. section:*

METHODS

2. **Comment:** Population: choose MCI OR early dementia - or change to subjective cognitive decline (MCI rarely reverts). You need to define your population well.

Response: *We changed the study population from MCI and early dementia to subjective cognitive decline. We agree that the significance and innovation are going to be greater by changing the population. We changed the Inclusion criteria section:*

"Individuals 65 years or older with a self reported subjective cognitive decline."

3. **Comment:** Change the scores: ADAS-Cog - MoCa (more sensitive for screening, used more in prior stages) - MMSE (use in later stages; problem as a repeated measure, patients can learn how to answer it - bias) - PANAS score (to track mood changes). Choose more efficient methods of scoring.

Response: *We changed the primary and secondary outcomes. Primary outcome is going to be the CANTAB score, a very sensitive score to track subjective cognitive decline. Secondary outcomes are going to be mood (PANAS), progression to cognitive impairment (MoCA), and progression to depression (GDS).*

4. **Comment:** Intervention: Music therapy was described well. However, do you choose music therapy to alleviate cognitive symptoms or to improve mood? How to differentiate between depression or early dementia? MT can improve mood, but we are not sure about improving cognition. You need to capture mood as a modifier (early dementia or older depression), and take it into consideration in the analysis of the data phase.

Response: *We opted to exclude patients currently experiencing depressive symptoms, and patients that have already been diagnosed with moderate to severe depression from the study; in order to address this issue.*

5. **Comment:** Control, You could add a group interaction after the PA in the control group.

Response: *We changed the control group. Control group is now going to have a 30 minute social leisure group interaction after the 30 minutes physical activity intervention. Before the control group there was only 30 minutes of physical activity.*

6. **Comment:** Time of intervention: 12 months of intervention is too long. You are going to have so many sessions that patients could potentially have a plateau effect (both improving or not at the same rate), and costs are extremely high. It is recommended to short the intervention to 3-4 weeks (first endpoint), and evaluate 6m and 12m after intervention. Suggested track in terms of cognitive symptoms instead of track progression. There is a study of 10 years follow up, but they did a 4-week intervention. (we need to find it). There are previous studies with shorter intervention data with a long term change. You can do a shorter intervention and a longer follow-up.

Response: *We changed the time of intervention from 12 months to 4 weeks. Follow-up assessments are going to be 4 weeks, 6 months and 12 months.*

7. **Comment:** You can use phase III or change to phase II, to understand the mechanisms.

Response: *We changed the phase of the study for a phase II, in order to assess efficacy MT intervention in subjective cognitive impairment.*

8. **Comment:** 3-week run-in period because you wanted to answer that the nursing homes out population can participate in the PA without a problem. Should you keep it?

Response: *Since we have shortened the intervention to 4 weeks, there is no need for a running period. No training on exercise or intervention needed that will require run-in.*

<u>SIGNICIFANCE, INNOVATION</u>

9. **Comment:** The trial is a story: we need to see the impact, what happened before the literature, gap, approach. Problem: people are getting older and older; life expectancy is increasing. Previous research with cognitive rehabilitation doesn't show a bigger result. That's why you should go to the earlier stage of subjective cognitive decline.

Response: *We agree, we changed the population for Subjective Cognitive Decline.*

10. **Comment:** Significance:

- Improving cognitive impairment in adults
- Preventive or delaying progression to AD
- Reducing health care costs
- Building evidence for non pharmacological intervention
- Assessing multi dimensional cognitive functioning
- Secondary impacts: people are less dependent; social benefits

Gap:

- Addressing limitations of previous studies and building a form from these previous studies.

Innovation:

- Combining two interventions while controlling for group interaction - novel
- Assessing subjective cognitive decline as the population (early stage of dementia)

Response: *We agree, we changed the Introduction section to address the significance, gap and innovation described above. Currently there is no information about the use of cognitive intervention like music therapy in order to delay subjective cognitive decline.*

METHODS

11. **Comment:** Primary outcome:

- CANTAB (computerized assessment - very good and standardized tool for cognitive assessment based on performance) - very sensitive - subset of tests with age related cognitive decline -> you want the clinical importance. Expect the variation of 3-4 points after training. Tracking reliable changes during time. After 65 years old there is a normal cognitive decline, you want to find the abnormal cognitive decline. You need to go with cognitive improvement rather than track progression.
- MoCa as a screening tool and follow-up

Response: *We agree. We changed the primary outcome for CANTAB.*

12. **Comment:** Secondary outcome

- Mood (PANES): music therapy can influence in mood and not cognitive decline
- GDS for controlling depression levels
- Quality of life is an indirect measure
- MoCa: after 65 there is the normal cognitive decline x abnormal cognitive decline (this is what you are trying to prevent). Search for cognitive improvement using the MoCa (improve in terms of cognition, that is worthy of publishing)

Response: *We agree. We changed the secondary outcomes for mood (PANAS), progression to cognitive impairment (MoCA), and progression to depression (GDS).*

13. **Comment:** Intervention: short intervention to 4 weeks, could do a follow at 12 months and 36 months after intervention - check about worsening of symptoms.

Response: *We have changed the duration of the intervention to 4 weeks, and modified the follow-up time to 4 weeks, 6 months and 12 months, as you suggested.*

14. **Comment:** Study design: change from phase 3 to phase 2.

Response: *We agree. We changed the phase of the study for a phase II.*

Reviewer 2

Outcomes, Exposure, Confounding

1. **Comment:** Confounders: two options to deal with. Minimize exclusion criteria in terms of enrollment. Stratify is the answer to everything:

- Stratify randomization based on what you are really concerned about (country - institution - education level - age 65-75, 75 or higher)
- Stratify analysis (exploratory) within each country or education level (adjust for educational level in the analysis)

Response: *We opted for cluster randomization, stratified according to site location, and age (strata of 65-79 years and 80+ were selected). Cluster randomization was chosen as the most suitable method to avoid contamination between participants in the same site.*

2. **Comment:** Intervention: suggested 4 weeks of intervention and 6 months and 12 months of follow up, to be more feasible.

Response: *We agree that 12 weeks of intervention would not be feasible. We have changed the duration of the intervention to 4 weeks, and modified the follow-up time to 4 weeks, 6 months and 12 months, as you suggested.*

3. **Comment:** Discuss the variable depression. Will you include patients with recent history of depression?

Response: *We opted to exclude patients currently experiencing depressive symptoms, and patients that have already been diagnosed with moderate to severe depression from the study; in order to address this issue.*

4. **Comment:** Discuss the variable stroke. What if one patient has a stroke during the trial? Discontinuing the intervention or excluding the patient? Post randomization is a red flag, you need to discuss this.

Response: *Only stable physical, mental and medical conditions will be included. Patients unable to perform the physical activity intervention (like after stroke sequelae) will be excluded. If the patients have a stroke during the intervention, they are immediately going to be discontinued. If they have a stroke during the follow-up period and are unable to respond to the questionnaires, the last observed score is going to be accounted for.*

5. **Comment:** Discuss the variable education. Lower education could have a more beneficial outcome with the intervention (potential confounder). Make sure to capture the educational status so we can compare it.

Response: *We are going to adjust the variable educational level during the sensitivity analysis phase.*

6. **Comment:** Changing the design to phase 2 is a good way to address everything.

Response: *We agree. We changed the phase of the study for a phase II.*

Reviewer 3

<u>METHODS</u>

1. **Comment:** Study Design: Contamination concern: What about the control group - are you going to tell the control group to not listen to music? Worry about how to keep the intervention clean. We can't control how much music each arm listens to. Collect data about listening to music externally? Similar to dietary studies, it is a lifestyle intervention.

Response: *The intervention is a really good way to avoid contamination (active music therapy, 30 minutes duration: global, country and personal preference) because it is difficult to replicate at home.*

2. **Comment:** Population: Maybe exclude musicians? Change inclusion criteria. Exclude depression? People should be stable. Exclude patients with recent episodes of depression or hospitalization in the last year.

Response: *We agree. We added musicians to exclusion criteria, once it is well known that musicians have a lower risk of cognitive decline when compared to the non-musician population. We opted to exclude patients currently experiencing depressive symptoms, and patients that have already been diagnosed with moderate to severe depression from the study; in order to address this issue.*

3. **Comment:** Run in period: What does the washing period do - you should screen for that in the period instead of a washing period.

Response: *Since we have shortened the intervention to 4 weeks, there is no need for a running period.*

4. **Comment:** What happens if the trial is successful?

Response: *We will keep following MoCa to see if they progress to MCI and depression with GDS. Next steps are to stimulate the performance of larger trials.*

5. **Comment:** Sample size: How to calculate? See studies from MCI to see standard deviation (probably larger, and it will be more conservative). Go for what is going to be clinical enough. Delta = 3-4 points. Justify the delta, the change between groups is going to be tricky - it is expected that both groups are going to improve. Outcome: what is the range from CANTAB? Is there a total? Or are you keeping them separate? Statistical analysis: t-test for difference between the means. Show what proportion in each group improves 3-4 points (stick to the mean changes). Stick in the score of the domain where we find more change.

Response: *Since there are few studies evaluating the effectiveness of musical therapy within the population described above, the effect size and variability necessary for the power calculations need to be estimated. We consulted a specialist (Prof. Jorge Leite) and extrapolated data from available literature. We acknowledge a limitation in accuracy of this sample size calculation, since the study uses a cluster stratification as well as stratification by two age groups, which will be considered after consulting a biostatistician.*

6. **Comment:** How to address generalizability?

Response*: We are going to adapt the music depending on the place where the study will take place. The MT intervention is going to be a 30 minute active MT, 10 minutes international songs, 10 minutes of national songs, 10 minutes of participants' chosen songs.*

7. **Comment:** How are you going to address selection bias?

Response: *We changed our intervention from moderate physical activity to light.*

Reviewer 4

METHODS

1. Comment: Blinding, Intervention**:** Explain who, how and why! You can call an open label, due to the nature of intervention. However, to address this situation, you are going to blind (explain who). You don't need to call it "open label"… you can remove it from the title. Patients and clinicians are most important to be blinded.

Response: *We changed the blinding session, as following:*

"In this study, participants and clinicians will not be blinded due to the nature of the intervention. However, to address this situation, outcome assessors and data analysts will be blinded to the group assignments."

Reviewer 5

METHODS

1. **Comment: Randomization, Study Population**

Sample size: As we age, the first loose processing (attention) first, and memory second. Working memory is the memory you use to do some tasks = if she had to pick one, she would choose. Maybe you can pick one domain (working memory) subset of primary outcomes and then take the other subsets as secondary outcomes.

***Response**: We consulted a specialist (Prof. Jorge Leite) and opted to keep all CANTAB subset scores as the primary outcome.*

2. **Comment:** Randomization: Strata for site and for age (two strata). Use a centralized randomization method + stratify by site, ensure allocation concealment. Balance or baseline cognitive performance (but if we keep 200 people we don't have to; if it is 100 we will have to balance).

***Response:** We agree. We opted for cluster randomization, stratified according to site location, and age (strata of 65-79 years and 80+ were selected). Cluster randomization was chosen as the most suitable method to avoid contamination between participants in the same site.*

3. **Comment:** What about using CANTAB once a month for 6 months?

***Response:** We opted to keep only at baseline and after 4-week intervention. CANTAB is too long and complex (it takes more than 30 minutes), it is more feasible to keep it twice.*

4. **Comment:** You should exclude severe / moderate depression, and unstable depression.

***Response:** We agree, we excluded patients currently experiencing depressive symptoms, and patients that have already been diagnosed with moderate to severe depression from the study; in order to address this issue.*

5. **Comment:** Subjective cognitive decline questionnaire (APA battery) could be used for screening.

***Response**: We agree. We added this questionnaire in the screening phase.*

Reviewer 6

<u>Sample Size</u>

1. **Comment:** Keep the style of the sample size paragraph, justify each choice, and use citations. Rely on clinical significance, play with the parameters of 0.2, 0.3, 0.4 Sample size does not necessarily need to be over 300. Phase II the point is to get information of the efficacy to go or not to a phase III. Do not game the number of sample sizes in order to get 300 individuals because of the randomization. Don't be afraid to do analogies, we can use cantab from other pathologies. Even if it's not the same outcome. Extrapolate the 10% on MMSE, then use 10% in the CANTAB. Is that effect size too big and not affordable? The worst sample size calculations are just one paragraph showing only the numbers. Reviewers check about plausibility, feasibility. Changes are reasonable, we need to rely more on clinical knowledge and experts. 300 hundreds (is the top limit), phase 2 is to obtain information, preliminary information. 200 hundreds, we could use stratified randomization.

***Response**: Since there are few studies evaluating the effectiveness of musical therapy within the population described above, the effect size and variability necessary for the power calculations need to be estimated. We consulted a specialist (Prof. Jorge Leite) and extrapolated data from available literature. We acknowledge a limitation in accuracy of this sample size calculation, since the study uses a cluster stratification as well as stratification by two age groups, which will be considered after consulting a biostatistician.*

Reviewer 7

Statistical Plan

1. **Comment:** Use the change from baseline as the primary outcome, continuous. No interim analysis. Analysis should be a linear mixed model comparing the two groups for a random intercept for clusters (deal with correlation - clusters).

Response*: We agree, and used CANTAB as continuous variable and the linear mixed model comparing the two groups for a random intercept for clusters (to deal with correlation).*

2. **Comment:** The first analysis should not be adjusted. You should do an ITT analysis => exposure variable is the group the patient was randomized to (does randomization to music cause an improvement of cognition?). The randomization can not be affected by confounders. Nothing confounds the exposure. Imbalance you can find is not a confounder, it is a random error. Sensitivity analysis can be performed after adjusting confounders (differences in baselines by chance can affect the outcomes). We can still adjust for variables: cluster level variables or patient level variables. In the per protocol analysis there can be confounders.

Response*: We agree. An intention-to-treat analysis will be performed for the primary outcome, and secondary sensitivity analysis will be performed in order to address potential imbalances between groups.*

3. **Comment:** You have to adjust the sample size calculation to clusters. You need the inter cluster correlation (much more complex) for calculating, if you do not find it, you need to do a pilot or do a sensitivity analysis (sensitivity sample size calculation). However, you need the help of a statistician to do this, and there is no time to do this right now.

Response: *Since there are few studies evaluating the effectiveness of musical therapy within the population described above, the effect size and variability necessary for the power calculations need to be estimated. We consulted a specialist (Prof. Jorge Leite) and extrapolated data from available literature. We acknowledge a limitation in accuracy of this sample size calculation, since the study uses a cluster stratification as well as stratification by two age groups, which will be considered after consulting a biostatistician.*

Reviewer 8

1. **Comment:** If they have depressive symptoms at the beginning,

Response: *We excluded patients currently experiencing depressive symptoms, and patients that have already been diagnosed with moderate to severe depression from the study; in order to address this issue.*

2. **Comment:** Exclude patients based on use of medications such as benzodiazepines, antihistamines and z-drugs.

Response*: We adjusted the exclusion criteria for only these 3 drugs.*

3. **Comment:** You can use all the outcome variables as continuous.

Response: *We adjusted all the outcome variables to continuous, changing MoCA from categorical to also continuous.*

4. **Comment:** During the presentation, highlight that your study is going to do a 6 and 12 month follow-up. However, if the investigators plan to do a longer follow-up, it is feasible due to the low cost.

Response: *We agree, we are going to mention this statement during the presentation.*

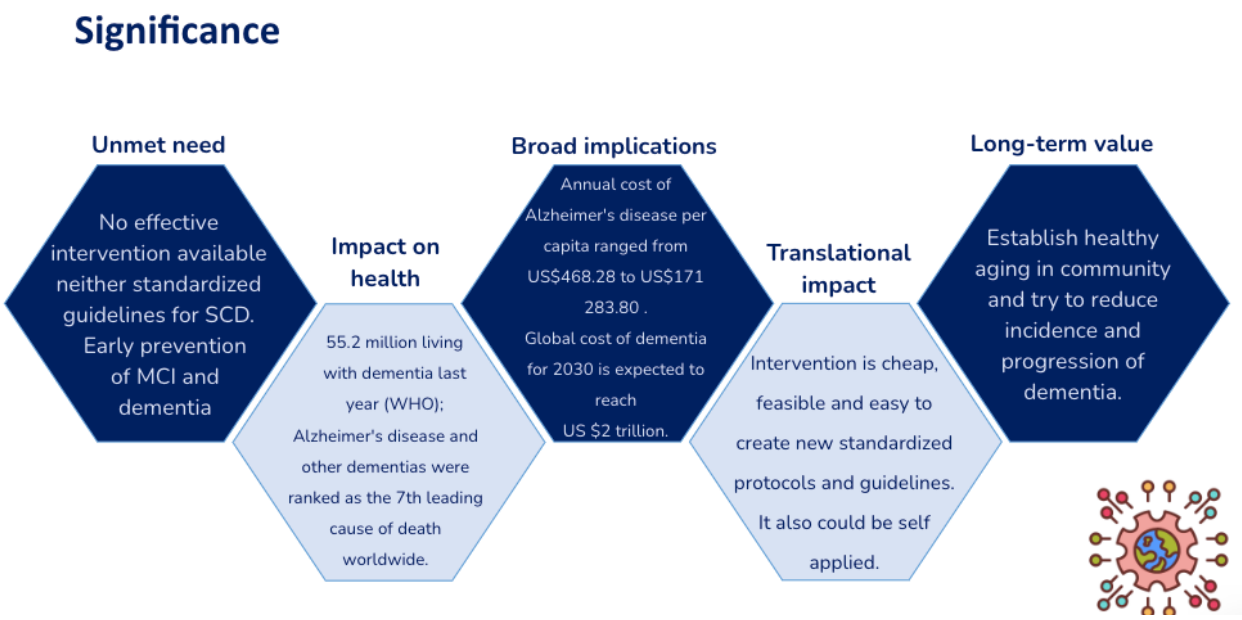
Reviewer 8

1. **Comment:** Significance strengths: real public health impact with a lot of potential; vast population, great economic impact; there's no non pharmacological guidelines for improving progression of dementia, it is the 7th leading cause of death (very significant). intervention is cheap.

Response: *Thank you for your comments.*

2. **Comment:** Significance weaknesses: not any identified; this is such an important disease but needs to be more explored in terms of knowledge about the pathology; vulnerable population.

Response: *We readjusted the slides in order to address these comments:*

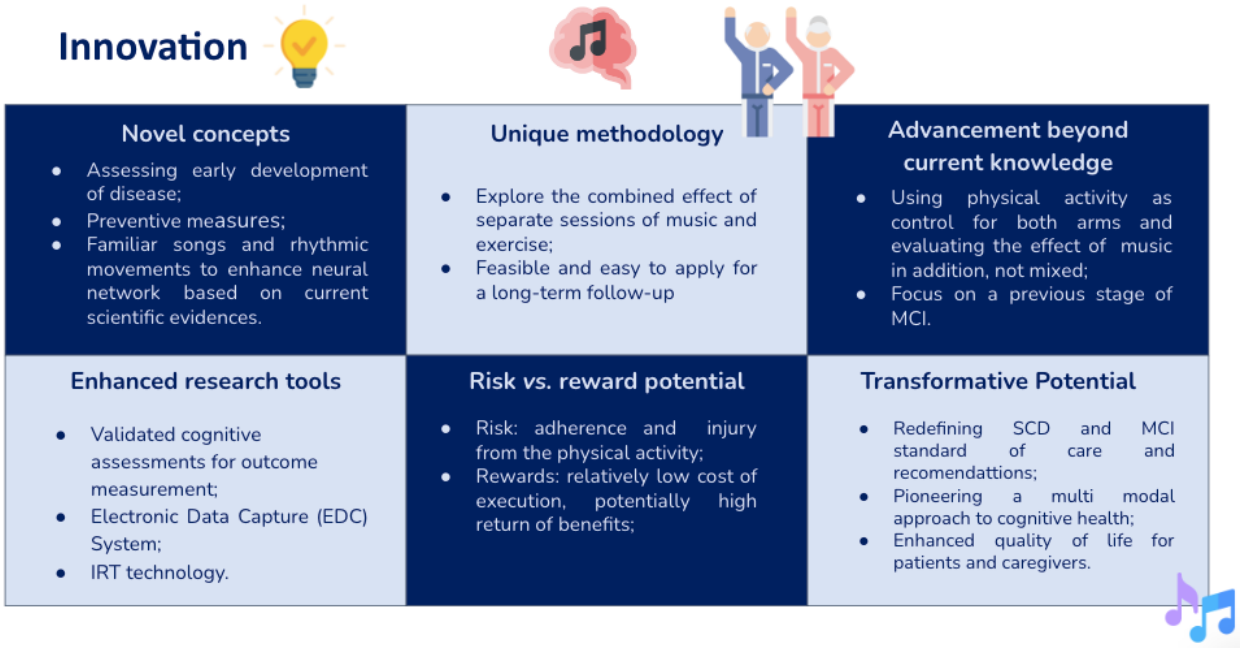


3. **Comment:** Innovation strengths: dual intervention, this can be mixed so it would be a better option to do it separately as they have addressed.; great idea to treat a giant ph problem; music as therapy for cognition, good assessments tools.

Response: *Thank you for your comments.*

4. **Comment:** Innovation weaknesses: not any identified; dementia can have different stages depending on the patients so how are they going to separate should be defined; vulnerable population.

Response: *We readjusted the slides in order to address these comments.*

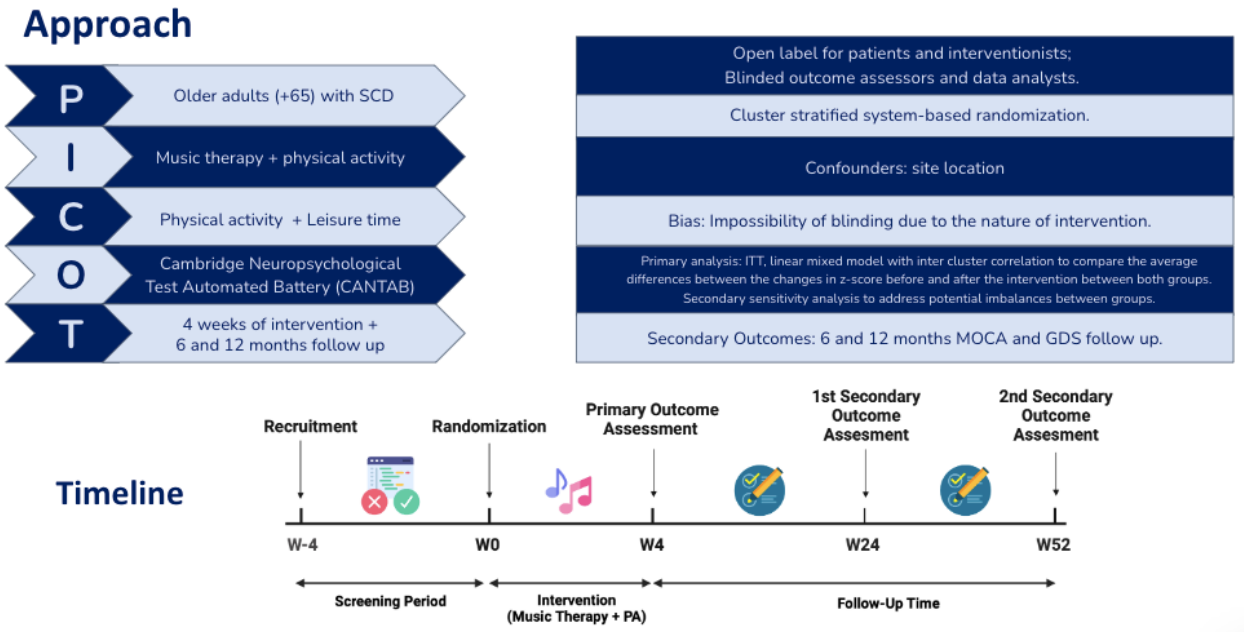


5. **Comment:** Approach strengths: block stratified randomization is always a good tool for these kinds of studies; easy application; great methods.

Response: *Thank you for your comments*.

6. **Comment:** Approach weaknesses: bias can appear due to the lack of blinding mentioned in their study.; attachment to the intervention; is 40 days enough to assess effects? What about long-term risks? Is it ethical? Is a sham + sham arm going to be accepted?

Response: *We readjusted the slides in order to address these comments*.



COMMENTS ABOUT THE INTRODUCTION PAGE DURING THE 5-DAY WORKSHOP 2024:

Significance, Innovation

1. **Comment:** Change the population to Subjective Cognitive Decline, as an earlier stage of Dementia:
   Response: *We changed from Mild Cognitive Impairment and Early Dementia to Subjective Cognitive Decline, in order to a more significant and innovative trial.*

2. **Comment:** Change the control group to physical exercise added to group leisure activity.

Response: *We changed from physical activity to physical activity followed by a 30 minute group interaction without music therapy, to control for social effect.*

3. **Comment:** Change the primary outcome to a more sensitive scale to SCD (CANTAB), and secondary outcomes to PANAS, GDS and MoCA:

Response: *We changed from MMSE to CANTAB (more sensitive to SCD); and from CDT, BADL, FAQ, QoL-AD to MoCA, PANAS, GDS (to track mood and cognitive impairment).*

4. **Comment:** Discuss about the study phase

Response: *We changed from phase III to phase II, to assess the efficacy of MT.*

5. **Comment:** Discuss time of intervention and follow-up:

Response: We changed from 12 weeks intervention to 4 weeks intervention; and 3, 6, 9, 12 month follow-up to 6, 12 month follow-up, to increase feasibility.

Outcomes, Exposure, Confounding - Heather Baer:

6. **Comment:** Discuss the variable depression as a potential confounder:

Response*: We excluded patients currently experiencing depressive symptoms.*

Study Design - Janis Breeze:

7. **Comment:** Contamination concern about the control group:

Response *We did not change the intervention, the specified type of active music therapy is difficult to replicate outside the study setting.*

Blinding, Intervention - Manuel Castillo Angeles:

8. **Comment:** You can call an open label, due to the nature of intervention. However, to address this situation, you are going to blind (explain who).

Response *We did not modify, open label, outcome assessors and data analysts will be blinded to the group assignments.*

Randomization, Study Populations - Sandra Carvalho:

9. **Comment:** Use a centralized randomization method and stratify by site and age.

Response *We changed from block randomization stratified by site, age, diagnosis to cluster randomization stratified by site and age.*

Sample Size - Jessica Paulus:

10. **Comment:** Rely on clinical significance, and extrapolate the data from other scales or outcomes if necessary.

Response: *We changed the sample size from 272 to 160 individuals. For calculation, a sensitivity analysis and adjustment for potential loss to follow-up were used.*

Statistical Plan - Armando Teixeira Pinto:

11. **Comment:** Analysis should be a linear mixed model comparing the two groups for a random intercept for clusters and perform a secondary analysis adjusting for confounders:

Response *We changed from unpaired t-test, Mann-Whitney and ANOVA tests to linear mixed model comparing the two groups for a random intercept for clusters.*

12. **Comment:** Adjust the sample size calculation to clusters:

Response*: We acknowledged a limitation in accuracy of our sample size calculation, once consulting a biostatistician would be required.*

COMMENTS BY PROFESSOR FELIPE FREGNI DURING THE REVISION PROCESS

1. **Comment:** Your manuscript has been carefully evaluated by the editorial team and reviewers, and we agree that it addresses a topic of significant scientific and clinical interest. We would therefore like to invite you to revise your paper in light of the editorial and reviewers' comments outlined below. Please ensure that all points raised are thoroughly addressed in your revision. As part of the editorial process, I have reorganized the structure of the manuscript and expanded several sections to transform it into a Brief Report focused on design lessons and methodological considerations for trials combining music therapy and physical activity in older adults. This approach highlights the conceptual and translational value of your work, even if the original trial has not yet been implemented. If you agree with these proposed changes, we kindly ask that all authors carefully review the revised text, verify accuracy throughout, and further expand selected sections with relevant recent references where appropriate. Once you return the updated version, I will ensure a prompt turnaround for the next review stage.

Response: *Dear Professor Fregni, thank you very much for your revision process. We carefully revised and agreed with all your suggestions. We accepted the tracked changes that you suggested and adjusted the Figure 1.*

2. **Comment:** You did not submit a response letter detailing how each editorial and reviewer comment was addressed. Please provide a point-by-point response explaining the changes made (or justifying why certain suggestions were not adopted).

Response*: Dear Professor Fregni. The only comments from the reviewers that we received were during the 5-Day Workshop in 2024. On that time, we developed a word document addressing each comment and each answer, and we sent it to PPCR Coordinator by e-mail. However, we did not submit the document to the PPCR Journal website. Here is the complete document, also responding the suggestions of November 2025 after the first review.*

3. **Comment:** In the Design Considerations section, parts of the manuscript are written as if the protocol is already being implemented. As this is a protocol paper, the language should consistently reflect recommendations rather than procedures already carried out. For example, instead of: "Participants are instructed to pedal at a self-selected, comfortable cadence, with resistance adjusted to maintain a target heart rate (THR) corresponding to approximately 50% of their age-predicted maximum heart rate (MHR), consistent with very light to light exercise intensity." Please revise to language such as: "We suggest that participants in a trial with exercise in cognitive decline should pedal at a self-selected, comfortable cadence, with resistance adjusted

to maintain a target heart rate (THR) corresponding to approximately 50% of their age-predicted maximum heart rate (MHR), consistent with very light to light exercise intensity." Please ensure that this distinction is applied consistently throughout the manuscript.

***Response**: Dear Professor, we agree. Thank you very much for revising and commenting this with us. We added tracked changes correcting this issue.*

4. **Comment:** You should expand the discussion of your design considerations by: Adding appropriate references to support key methodological decisions; Discussing alternative approaches and explaining why the chosen design was selected; Providing examples of how similar approaches have been used in prior trials. This will strengthen the methodological rigor and transparency of the protocol.

***Response**: Dear Professor, we agree. We expanded the discussion section. The changes are tracked in the manuscript.*