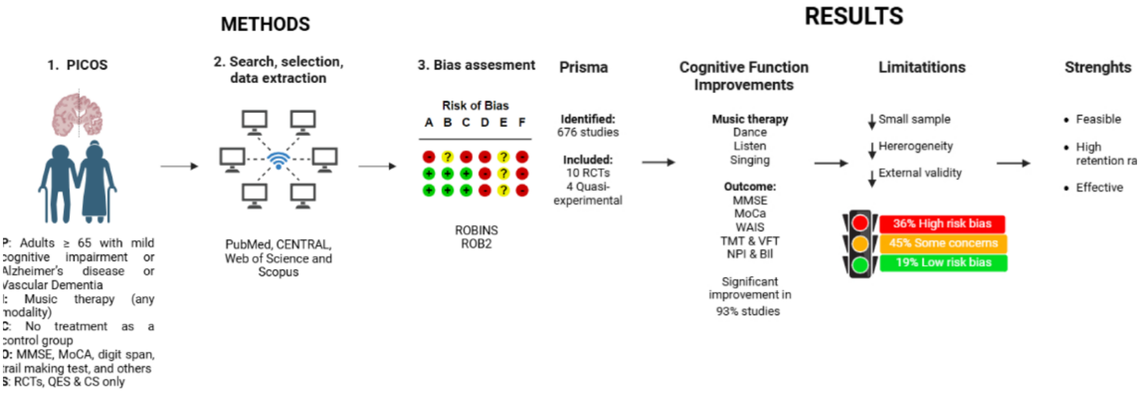
Peer-Review comments and authors responses

Reviewer #1:

Comments:

This study is complete. It describes the rationale for the review in the context of existing knowledge and the results of the search and selection process. The study uses flow diagrams and well-explained tables. The manuscript addresses the study's bias and limitations. Consider including a graphic abstract; it could help with understanding the study results.

Response: We greatly appreciate these positive comments. Following your suggestion, we have developed a comprehensive graphic abstract that synthesizes our key findings and methodological approach.



Reviewer #2:

METHODS: *Figure 2 - > The map would be better to add a gradient of articles instead of only categorizing > 2. (I think it is fine if the maximum number of articles for the same country is 2).*

Response: Thank you for this thoughtful suggestion regarding the visualization of geographical distribution. After careful consideration, we maintained the current categorization because the maximum number of articles per country was 2. Specifically:

- 2 articles each: China, United States, Japan, Spain
- 1 article each: Finland, Italy, France, Greece, Portugal, Brazil We believe that implementing a gradient scale for such a narrow range (1-2 articles) might not enhance the visual interpretation of the data.

Table 1 could be separated in two. It can be divided into Table 1 including intervention and control groups, then Table 2 outcomes, main findings, and dropout. For example, the main findings need more detail if they are calculating the pre- and post-mean differences or other measures using the scales. Reducing the amount of information will make it easier to read.

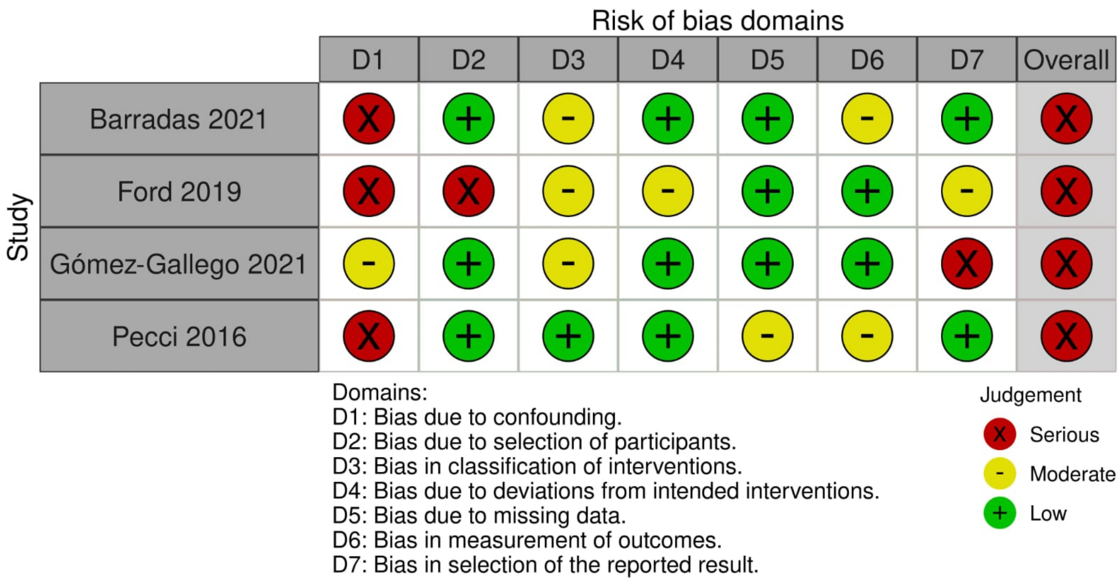
Response: We agree that dividing the information would enhance readability.

Is it possible to do a meta-analysis of the mean differences between the intervention and control arms? It seems for the measure of cognitive domains most of them use the MMSE scale (you can meta-analyze this). The same for the studies using Moca or Digit span score. Keep in mind it needs to be the same scale as a minimum requirement to do a meta-analysis, as well as the same population (AD, no early onset AD or other variations). In the sensitivity analysis, you can choose those with the more similar parameters for music therapy. If you do a mean difference meta-analysis you can do a subgroup analysis by intervention and control, which will give you a secondary p value of testing group differences. Maybe this can be a little more complicated, but the dropout rates can be presented as a meta-analysis of proportions. This meta-analysis of proportions could answer if the average proportion of dropouts is high-medium-low.

Response: Dear Reviewer, we appreciate your comment. During the design phase, we only planned the study to be a narrative review. We now discussed the possibility of performing a meta-analysis, however we chose not to do so because we included all types of therapy and all types of studies. It seems that there is too much heterogeneity among the studies to make valid interpretations from a meta-analysis.

There is no table/figure of risk of bias of the quasi-experimental studies using ROBINS-I. You can present this in a similar way to Rob2. In Figure 2, remember the consensus is to use + (green), ? (yellow) and - (red). There is a tool called Robvis that can help you to do this figure.

Response: We added "Figure 3".



In the first methods paragraph, please add that this systematic review was not previously registered in PROSPERO or other place.

Response: In the METHODS section, first paragraph, we added "This systematic review was not previously registered in PROSPERO".

The methods need a section about synthesis methods (which include narrative methods like using tables, or quantitative methods like meta-analysis) Also, a section about the effects measures (what are the primary and secondary outcomes of the trials is also important (for example the use of mean differences). Finally, add if you want to use a method to measure the certainty of evidence (the SOF table using the GRADE framework for interventions could be appropriate)

Response: Dear Reviewer, we added a new METHODS section: “Synthesis methods”.

Line 265 and 266 are confusing. Please rephrase that.

Response: We excluded the phrase “One study had more males and one showed differences between the two arms, in which males were more prevalent in the intervention group and females in the control.”

FORMAT: *Keep in mind the number of words per section (per intro, results, Discussion, etc). Remember some parts of the result can be merged together. You don’t need to repeat in writing what is stated in the tables and figures unless it is of high importance.*

Response: We excluded one phrase in the INTRODUCTION, and some paragraphs in the RESULTS section.

DISCUSSION: *I would suggest also exploring the heterogeneity of the studies in the discussion. Why do they use different scales and different types of patients? Analyzing the certainty of evidence with GRADE can also add to the discussion.*

Response: We moved two paragraphs from INTRODUCTION to DISCUSSION section, and added some statements about the heterogeneity of the studies.

Reviewer #3:

METHODS*: Table 1 lacks some information regarding the missing data techniques used for few studies. If you could, please address that and add that information.*

Response: Dear Reviewer, thank you so much for your comment. We added this information in Table 2, last column.

Reviewer #4:

Initially, I would like to inquire about the citation protocol you are using, as it seems to vary throughout the document.

Response: Dear Reviewer, thank you very much for your comment. We checked again our citation protocol, APA7, and added some corrections. In the RESULTS section, “Main Findings”, “MMSE changes” subsection, Doi et al., and Higuiti et al. In the “MoCA” section, we corrected for Lazarou et al. (2017). In the “Digit Span” subsection, we corrected for Lazarou et al. In “NPI” subsection, Higuiti et al., and Doi et al. In “TMT” subsection Doi et al. (2017), Pecci et al. (2016), Higuiti et al. (2021), Bisbe et al. (2020). In “Missing Data” subsection, (Xue et al., 2023), (Doi et al., 2017), (Ford et al., 2019), (Gomez et al., 2021), Gomez et al. (2021), Higuti et al.í (2021), Doi et al. (2017), and Wu-Chung et al. (2023). In “Feasibility of the interventions” subsection, Doi et al. (2017), (Ford et al., 2019), (Ford et al., 2019), (Särkämö et al., 2014; Wu-CGomez et al. (2021),hung et al., 2023). In “Assessment of risk of bias” subsection, Pecci et al., (2012), Ford et al., (2019), Barradas et al. (2021). In the DISCUSSION section, “Music intervention on specific domain of cognitive functions” subsection, Bisbe et al.and colleagues (2020). In “Strengths and Limitations” subsection, (Bisbe et al. 2020; Pecci et al., 2016; Pongan et al., 2017; Särkämö et al., 2014), (Särkämö et al., 2014), (Ho et al., 2020), (Ho et al., 2020), (Satoh et al., 2017). In “Implications for practice” subsection, (Bisbe et al., 2020; Ford et al., 2019; Gomez et al., 2021; Lazarou et al., 2017; Pecci et al., 2016; Pongan et al., 2017; Särkämö et al., 2014; Satoh et al., 2017; Xue et al., 2023), (Bisbe et al., 2020; Doi et al., 2017; Gomez et al., 2021; Lazarou et al., 2017; Pongan et al., 2017; Särkämö et al., 2014; Satoh et al., 2017; and Xue et al., 2023).

INTRODUCTION: *Your manuscript rightfully highlights the limited efficacy of pharmacological treatments in halting disease progression in neurodegenerative diseases, which rightly piques interest in alternative therapeutic approaches like music therapy (MT). However, your introduction asserts with undue positivism that “music therapy is one such method that may be beneficial for improving cognitive function...”. To prevent potential early red flags for the reader, it would be advisable to adopt a more conservative tone, perhaps stating that “some authors report potential benefits of music therapy...”.*

Response: In the INTRODUCTION section, we reformulated “Some authors report potential benefits of Music Therapy (MT) Music therapy (MT) is one such method that may be beneficial…”.

Furthermore, it would benefit the conventional reader to have a brief explanation of the neurological mechanisms by which music therapy purportedly exerts its effects on the cognitive variables under consideration.

Response: We added one paragraph about this topic to the DISCUSSION section.

When citing studies such as those by Dorris et al., 2021; Fusar et al., 2018; and others, you suggest a general approach to the topic by previous researchers. However, your manuscript claims to focus more narrowly on Alzheimer’s and Vascular dementia without adequately justifying this choice. It is crucial to clarify why this specific focus is necessary.

Response: We added one phrase about this topic to the INTRODUCTION section. We chose Alzheimer’s and Vascular dementia because they are the most frequent causes of dementia.

It is particularly striking that to address the "knowledge gaps," your study opts for a systematic review rather than generating new data through experimental research. Considering the diverse nature of the aspects you wish to explore, such as the effects of MT on different types of dementia, the specific types and durations of MT that are most *effective, and their impact across all cognitive domains, a review might seem insufficient. If these gaps are indeed valuable, why not consider an experimental approach to generate fresh data?*

Response: We reformulated the objectives paragraph in the INTRODUCTION section, in order to clarify the knowledge gaps.

Regarding your critique of pharmacological therapies for not halting disease progression, I question whether MT, your proposed alternative, halts the progression either. This argument might weaken the rationale for studying MT as it shares a common limitation with the pharmacological options.

Response: We agree with your comment. We excluded the phrase in the INTRODUCTION section that argues against pharmacological therapies.

METHODS: *I am also concerned about the 20% of the RCTs that did not specify their analysis type. When the analysis approach is unclear, it complicates the interpretation of the outcomes. How did you address this issue in your review? Would it be prudent to reassess and possibly limit your analysis to the 8 RCTs that at least report minimally robust methodologies?*

Response: The studies that did not specify the type of analysis were Särkämö et al. (2014) and Wu-Chung et al. (2023). After reviewing Särkämö et al. (2014), we found that it was a per-protocol analysis, and that Wu-Chung et al. (2023) performed an intention-to-treat analysis. We reformulated the RESULTS section.

CONCLUSION: *In conclusion, the use of the terms "potential" and "effective" alongside "despite these promising results" introduces a tone of bias toward positive outcomes, which may not be fully supported by the evidence presented. This enthusiasm seems misplaced, particularly in the conclusions section, which should be reserved for definitive statements supported by the results. Additionally, the conclusion section seems to blur with other sections by reintroducing methodological suggestions and discussing limitations that should be confined to their respective sections. In summary, while your review aims to synthesize existing evidence and offer clear recommendations on the usefulness of MT, it seems to overlook the need for generating new data to fill the identified gaps adequately, especially because those gaps are quite specific. This approach may not provide the robust foundation required for such recommendations.*

Response: We agree with your comment. We reformulated the INTRODUCTION and CONCLUSIONS sections, in order to change this unclear data.

Reviewer #5:

TITLE: *Try to change the ''Impact'' word on the title since it is a very big assumption to say that we are actually measuring the impact based on a Systematic Review. In addition patients with these disorders receive other therapies in parallel so assuming a direct causality impact based on music therapy might not be appropriate. I would suggest the word impact for ''effects'' or ''findings''.*

Response: Dear Reviewer. Thank you so much for your insightful comments. We agree, and we changed the title to “Effects of Music Therapy on Cognitive Function in Elderly Patients with Mild Cognitive Impairment and Early Dementia: A Systematic Review”.

INTRODUCTION: *Well structure for the introduction. However, in third paragraph other ‘’complementary therapies’’ should be introduced as well. Why researching directly for MT? What is MT offering that other are not and thus is worth researching more about it? These aspects are an essential component of significance of your study, so please develop more on this aspect.*

Good closing paragraph, however, why should a comprehensive review be enough to provide recommendations? This is a big and not justified assumption. Instead, the relevance of this study seems to be more focused on describing the evidence on MT to try to understand the effects reported so far.

Response: We agree. We reformulated our INTRODUCTION paragraph, specially in the section regarding the objectives of the study. We moved two paragraphs from INTRODUCTION to the DISCUSSION section, and adapted it.

METHODS: *Please justify why non-randomized, quasi-experimental intervention studies (QES) and cohort studies were included as part of the inclusion criteria. This significantly creates a heterogenous pool of studies in terms of quality evidence.*

Response: Thank you for your valuable comment. We chose to include non-randomized, quasi-experimental intervention studies and cohort studies in our inclusion criteria to capture a comprehensive range of evidence on the effects of music therapy on cognition. Given the variability in the study designs available in this field, we wanted to reflect real-world evidence by including a broader range of methodologies, even if this introduced some heterogeneity.

Including only studies where cognition was the primary outcome seems to be very restrictive and could have reduced the inclusion of potential studies with quality evidence. Even in the variables extracted from the studies, secondary outcomes were considered, so why no taking studies where cognition was reported as secondary outcome? Please justify.

Response: Regarding our decision to include only studies in which cognition was the primary outcome, we recognize that this may have limited the pool of studies. However, we aimed to prioritize studies that focused directly on cognitive outcomes to better balance the variables most likely to influence our main objective. While secondary outcomes were considered, we wanted to ensure that cognition was the primary target in each study to maintain a focused analysis of cognitive impacts.

Inclusion criteria based on only RCT’s, and cognition reported as outcome could have been a better systematic approach.

Response: We appreciate your suggestion to include only RCTs with cognition as an outcome for a more systematic approach. Although we agree that such a criterion would enhance the quality of evidence, we opted for a more inclusive review to reflect the diversity of studies and interventions in music therapy research.

RESULTS: *Please describe more in detail the essential characteristics of the type of interventions found. I would even suggest creating a flowchart or illustration grouping the essential categories of interventions such as the ''passive'' and ''active''. This would help the readers to understand better the type of Music Therapy Interventions available on Cognitive Function. I would suggest adding Table 1 and 2 as a supplementary material instead of the main text.*

Response: In response, we expanded the description of intervention types and added Table 3 to categorize music therapy interventions as "passive" or "active", improving clarity on the intervention types found.

Please state effect sizes if reported in the studies, since significant values might not represent significant clinical effects. I would like to see which are the therapies that are showing/reporting higher effect sizes after therapy.

Response: While we considered moving Tables 1 and 2 to supplementary material, we opted to keep them in the main text for ease of access, as they visually summarize essential study details. Regarding effect size, not all studies provided this information, limiting direct comparisons.

As noted in the Risk of Bias section, the Qualitative Evidence Syntheses (QES) were found to have numerous methodological deficiencies, including missing data. Therefore, one might question the inclusion of these studies in this systematic review. Including them seems to introduce more noise and confusion rather than contributing high- quality evidence. Please consider excluding these papers.

Response: Regarding the inclusion of Qualitative Evidence Syntheses (QES), we recognize the limitations in quality. We decided to keep these studies, as they offer valuable insights and have highlighted their limitations (as low quality) in the discussion, emphasizing the need for more rigorous future research.

DISCUSSION: *Very nice figure 3, however, why this was included in the discussion section. This would have been a better asset in the results section when describing the sites where MT is implemented.*

Before jumping directly to the effects of music intervention on specific domains of cognitive function, it is important to discuss the general effects found in cognition to give a general overview of this review compared with previous.

Response: We agree. We added two paragraphs to the DISCUSSION section. We also addressed Figure 3 (that now is Figure 4, after some corrections suggested by other reviewers) in the RESULTS section, as you suggested.

LIMITATIONS: *Some aspects would be important to add in this section. First, selecting only studies with cognition as a primary outcome might have limited the number of studies included and that could have been an asset for this review. Second, adding low-quality studies (QES) could have been introduced potential bias in the evaluation of Music*

Therapy on Cognitive Function. Finally, based on the heterogeneity of the results and type of studies, it is difficult to assume the actual effects of this type of therapy.

Response: We incorporated your suggestions in the Strengths and Limitations section: (1) restricted study pool by selecting cognition as the primary outcome, (2) potential bias from including low-quality studies, and (3) challenges in drawing broad conclusions due to study heterogeneity.

CONCLUSION: *Good structure of the conclusion section. However, based on the results, discussion, and methods used in this study, it might be inappropriate to state that music therapy shows potential as an effective non-pharmacological intervention to improve cognitive function. Please restate this section of your conclusions based on your findings. Also a systematic review alone, methodologically speaking is not able to bring conclusions regarding effectiveness.*

Response: We rewrote the first sentence of our conclusion: “Overall, music therapy seems to be a non-pharmacological intervention that benefits cognitive function in elderly patients with MCI and early dementia in isolation and in combination with physical exercise or dance.”