

Review

Not peer-reviewed version

---

# The Effect of Music Therapy on Psychological Outcomes for Neurological Conditions: Systematic Review

---

[Sarah H. Gardener](#)\*, [Elizabeta B. Mukaetova-Ladinska](#)\*, [Nellinne Perera](#)

Posted Date: 13 August 2025

doi: 10.20944/preprints202508.0993.v1

Keywords: music therapy; neurological conditions; cognitive function; mood; emotion; behaviour; quality of life



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Review

# The Effect of Music Therapy on Psychological Outcomes for Neurological Conditions: Systematic Review

Sarah H. Gardener <sup>1,\*</sup>, Elizabeta B. Mukaetova-Ladinska <sup>1,2,\*</sup> and Nellinne Perera <sup>1</sup>

<sup>1</sup> School of Psychology and Visual Sciences, University of Leicester, University Road, Leicester LE1 7RH, UK

<sup>2</sup> The Evington Centre, Leicester General Hospital, Gwendolen Road, Leicester, Leicestershire, LE5 4QF

\* Correspondence: sg751@leicester.ac.uk (S.G.); eml12@le.ac.uk (E.B.M.-L.)

## Abstract

**Background and Objectives:** Music therapy has been used as a non-pharmacological treatment for neurological conditions, supporting cognitive, psychosocial, behavioural and motor functions. Although it is evidence-based, safe, and cost-effective, music therapy remains uncommon in neurorehabilitation services, largely due to a lack of quality research. This review aims to understand the effect of music therapy on psychological outcomes for neurological conditions using randomised clinical trials (RCTs). **Materials and Methods:** A systematic review of literature published in four databases (PsycINFO, CINAHL, PubMed, and Scopus which is inclusive of Medline and EMBASE) from 1 January 2015 to 31 January 2025 was performed. **Results:** Ten RCT's comprising of 469 participants were included in this review, with seven studies synthesised using a systematic review without meta-analysis approach. MT was found to improve cognitive function in individuals with Parkinson's disease (PD) and Traumatic Brain Injury (TBI), but not in those with stroke or Huntington's disease (HD). In HD improvements in cognitive function were noted in the comparator group. MT led to significant improvements in emotional well-being in stroke survivors, though not in mood. Behavioural outcomes did not significantly improve in HD, while quality of life improved following MT for individuals with multiple sclerosis (MS) and stroke. **Conclusions:** Overall, MT was perceived as a positive intervention with potential psychological benefits. However, findings were mixed, and the heterogeneity of population, intervention, comparator, outcomes and study design limited the ability to detect group differences and reduced confidence in the findings. Future research should focus on larger-scale RCTs that reduce bias, accommodate population diversity, and clearly define and distinguish MT from treatment-as-usual. Better standardisation of outcomes and improved reporting will be essential to evaluate the efficacy of music therapy in neurological conditions and strengthen the evidence base for its use in clinical practice.

**Keywords:** music therapy; neurological conditions; cognitive function; mood; emotion; behaviour; quality of life

---

## Introduction

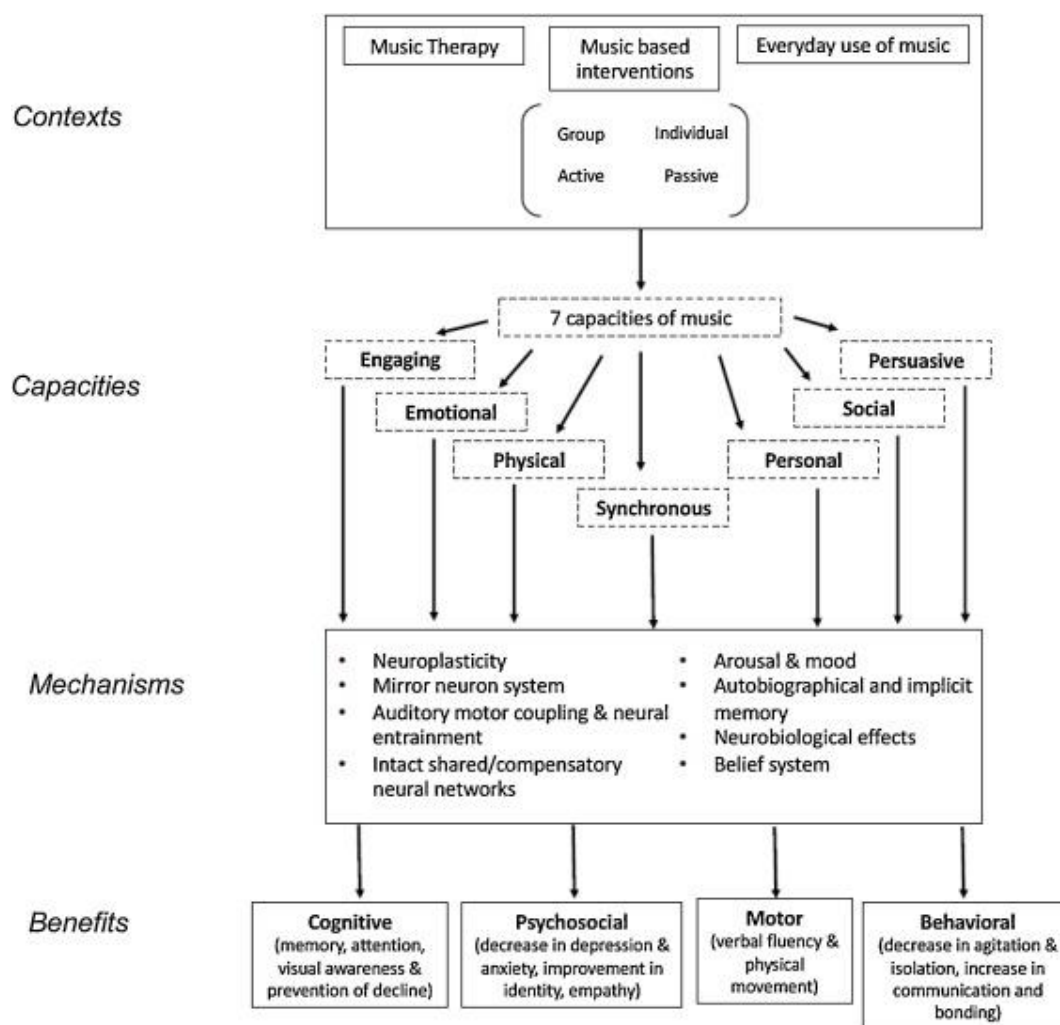
Music therapy (MT) is characterised by the use of music within a therapeutic relationship with the music therapist, encompassing multisensory stimulation [1–5]. MT, delivered by qualified, trained music therapists is distinguished from “music-based interventions” delivered by healthcare professionals that often comprises music listening, where intervention content varies depending on professional background [2,6].

MT supports psychological factors associated with neurological dysfunction, such as mood, emotion and distress management, and quality of life, while facilitating empowerment and enabling a supportive role towards others, which in turn support other areas of neurorehabilitation such as motor function [7–13]. Neurologic music therapy (NMT) is a specific evidence-based form of MT that

supports many presentations of a wide range of neurological conditions that affect the brain, spinal cord and nerves, such as traumatic brain injury (TBI), stroke, Parkinson's disease (PD), and multiple sclerosis (MS), among many others, by improving disorder of consciousness awake levels, to aiding sensorimotor, speech and language, and cognitive rehabilitation [11,14–19]. Therefore, MT is unique in its versatility and, NMT, can be used to support a range of needs during neurological rehabilitation [20].

### Music Therapy in Theory and Clinical Practice

The rational scientific mediating model (RSMM) was designed to translate music research into applied clinical music research. This led to the development of the transformational design model (TDM), used to translate clinical music research into functional MT [20–22]. This ensures therapeutic goals are both prospective rather than retroactively incorporated, a weakness of some music interventions, and better related to the functional therapeutic outcomes [20,22]. To further explore how MT benefits neurological conditions and establish a theory of change, the introduction of the Therapeutic Music Capacities Model (TMCM) was a step further and illustrates the neurological and psychological mechanisms that link the capacities impacted by MT to cognitive, psychosocial, motor, and behavioural benefits. In addition, this model provides a framework how MT can be personalised, adapted and applied to all neurological conditions (Figure 1) [14,23,24].



**Figure 1.** The Therapeutic Music Capacities Model (TMCM) [14]. This framework illustrated neurological and psychological mechanisms that link the capacities impacted by MT to cognitive, psychosocial, motor, and behavioural benefits. For more details, please see the main text.

MT is an evidence-based intervention for neurorehabilitation. However, limited resources are available for MT use in the United Kingdom (UK) National Health Services (NHS) [25]. MT research has a greater focus on dementia compared to other neurological conditions, where it is more widely used, suggestibly due to limited pharmacotherapy options for dementia unlike other neurological conditions that can be partially managed by pharmacological and multimodal non-pharmacological interventions (i.e. physiotherapy, speech and language therapy, electrotherapy etc.) [26–29] MT is not currently commissioned in specialised neurorehabilitation services for neurological conditions other than dementia, where in the latter there is still limited funding [30,31]. This raises ethical concerns about pharmacotherapy being heavily relied upon and promotes exploration of MT as an effective nonpharmacological treatment that is less invasive, side-effect-free and cost-reducing intervention in the treatment of neurological disorders [11,14].

### *Research in Music Therapy for Neurological Conditions*

Neurological conditions impede certain cognitive functions (i.e. executive functioning or speed of processing) impacting mental health, with them being more disabling than physical impacts in some instances [8,13]. MT has been suggested to improve mood and motivation, potentially increasing treatment adherence, implying psychological improvements from MT may benefit other domains, such as motor function [32]. Studies on MT for psychological benefits, such as cognitive rehabilitation, are limited. though cognition, mood, emotion, and quality of life for neurological condition have all been shown to benefit [11,16,33].

Research on MT is generally low in quality, potentially attributed to challenges with evaluating this procedure [6,33,34]. For example, the concept of MT has varying definitions across studies, and difficult to distinguish the MT effects from other treatment-as-usual interventions [16,36]. Even when other controlling for variation, music therapists bring their own training, musicality, style, and experience, and embrace creativity, which is an asset but difficult to quantify for research [37]. Furthermore, with few UK NHS services offering MT, research opportunities are limited [38,39]. National Institute for Health and Care Excellence (NICE) (2023) guidelines [40] recommend research on MT should include randomised-controlled trials (RCTs) to establish benefits for neurorehabilitation and patient-centred outcomes [41]. Therefore, this review will critique RCTs exclusively.

With literature on dementia establishing MT as a recommended non pharmacological intervention to support psychological outcomes [42,43] it is important to review other neurological conditions, and whether a shift from potentially unnecessary pharmacotherapy is possible [2,24]. The novelty of this review will establish the evidence and quality of research on MT for neurological conditions that meet criteria for UK specialised neurorehabilitation services, using the recommended RCT design [30,40]. Limitations and recommendations will be identified for future research and clinical implications. The research question is to review whether there is an effect of MT for psychological outcomes in adults with neurological conditions

## **METHOD**

### **Search Strategy**

To evaluate the effect of MT for psychological outcomes in adults with neurological conditions, the PICOS framework, well-established for developing systematic research, was employed to identify search terms and inclusion and exclusion criteria, consisting of the population (P) as adults with neurological conditions, intervention (I) as MT, comparison (C) as treatment-as-usual or specific alternative intervention, clinical outcome of interest (O) as psychological factors (see Table 1), and study design (S) as RCTs [9–12,14,18,44–47].

**Table 1.** Outcome variables and how they are defined by the included studies.

Outcome variable	Definition for included studies
Cognitive function	Executive functions, memory, visuospatial abilities, attention, communication [49, 50]
Mood	Depression, anxiety, anger, vigour, fatigue (11, 32, 51]
Emotion	Self-perceived emotional well-being, emotional awareness of self and others, sharing of emotions [32,49]
Self-efficacy	Sense of competence in managing new and challenging situations [52]
Behaviour	Communication and expressive skills, mental rigidity, aggression [53]
Affect	Valence, arousal, dominance [52]
Quality of life	An individual's perception of their physical and mental state, and social position [54]

The search strategy used search terms and synonyms of key terms including music therapy and types of neurological conditions (Appendix A) in representative databases frequently used in MT research: PsycINFO, CINAHL, PubMed, and Scopus (inclusive of Medline and EMBASE), covering psychology, healthcare, and peer-reviewed research [36,55]. Non-peer-reviewed research, including grey literature, was excluded due to risk of poor validity and unoriginality of research [56,57].

Five specific inclusion and exclusion criteria were used to determine eligible studies for inclusion in the review (Table 2). All neurological conditions were included in the search, except for dementia, spinal cord injuries, and neurodevelopmental conditions. Search results were filtered between 1st January 2015- 31st January 2025 to obtain the most recent research in the field worldwide. This captures time when MT is a registered profession requiring formal training, and before, during and after the covid-19 pandemic when alternative forms of delivering MT were required, such as distance, instrument-use limitations and online delivery, which remain incorporated clinically and therefore relevant to current practice [7,31,58–60]. Limiting the search to English language benefits

time constraints, reduces heterogeneity, and is the common language used in MT reviews, although may bias the range of studies despite opening the search to any location [61–63]. RCT-only studies were included since RCTs are recommended to improve the quality of findings establishing the effects of MT for neurological conditions (35, 36, 40). In the current review, RCT pilot studies were included with data relevant to MT effectiveness [64]. Furthermore, RCT pilot studies are increasingly comparable to main studies [65]. Five of the ten papers [11,32,50–52] included outcome measures not directly relevant to the research question but were included to ensure relevant data was captured and because different outcomes can be related since MT may be integrated into a multi-disciplinary approach [39]. However, the combined outcomes in these studies were accounted for. MT for dementia, including Alzheimer’s disease, Dementia with Lewy Body, Vascular dementia, Mixed dementia etc., and neurodevelopmental conditions (i.e. Autism Spectrum Disorder, Attention-Deficit/Hyperactivity Disorder (ADHD), Intellectual Disability, Specific Learning and Tic Disorders) are better established and follow differing treatment pathways to other neurological conditions, and were therefore excluded [28,30,66]. Spinal cord injuries have been excluded to focus on neurological conditions where there are direct deficits on psychological factors such as cognition [67]. Since RCTs require quantitative data for increased comparability and generalisability, qualitative studies were excluded [68]. Reviews were excluded due the research not being original, and protocols were excluded due to lack of data. Feasibility studies have been excluded with aims observing intervention feasibility rather than effectiveness. Studies researching non-psychological outcomes such as motor or physical were also excluded.

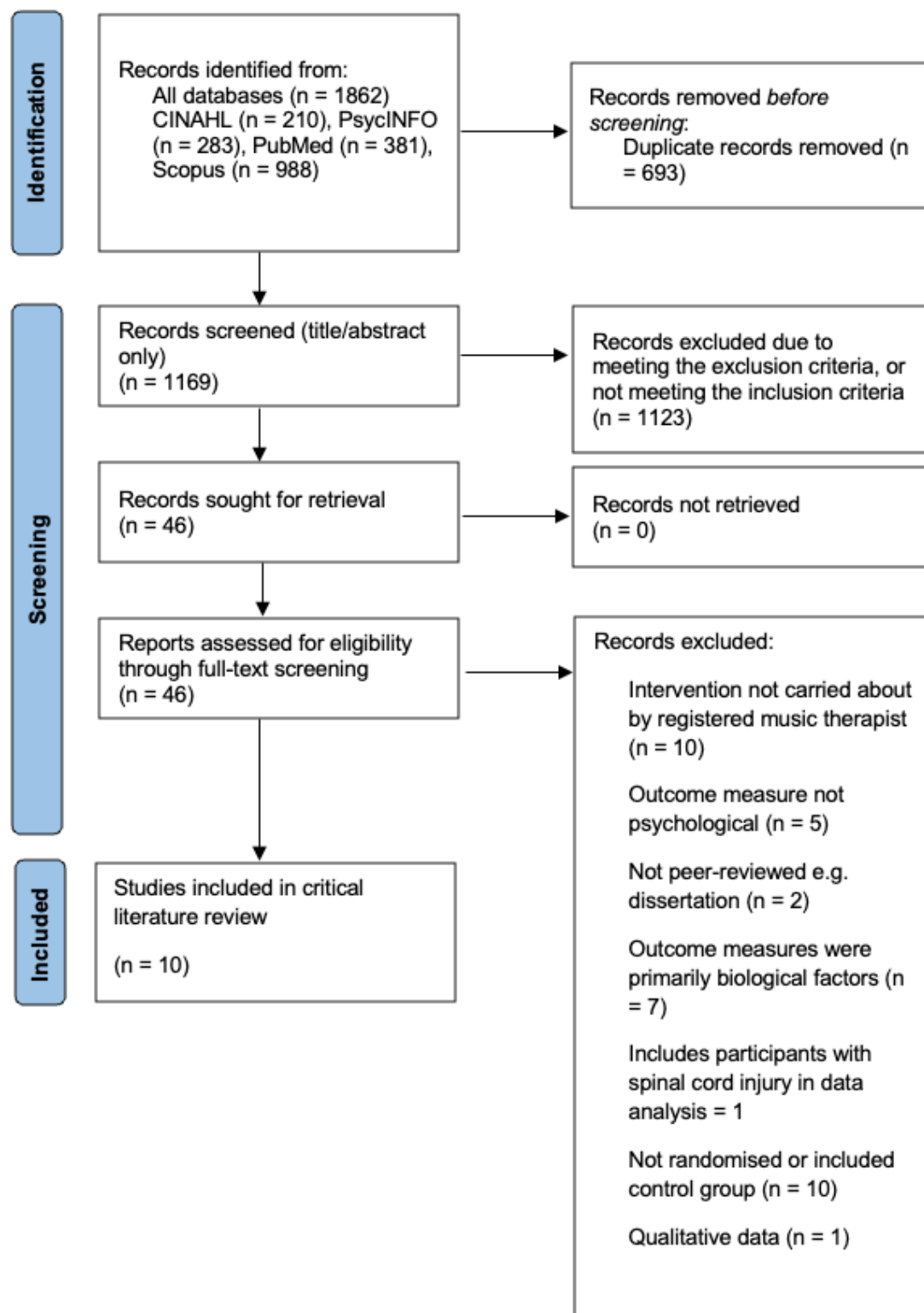
**Table 2.** Criteria used to determine eligible studies for inclusion in the review.

Inclusion Criteria	Exclusion Criteria
a) Peer-reviewed original empirical RCTs	a) Studies with dementia or spinal cord injury
b) Adults with neurological conditions	b) Studies with neurodevelopmental conditions
c) MT delivered by a board-certified music therapist or a therapist skilled in delivering specific evidence-based NMT	c) Study designs such as review, protocol, or feasibility studies
d) Published between 1st January 2015-31st January 2025,	d) Studies with qualitative or mixed methods data
e) Published in the English language	e) Studies containing non-psychological outcomes

### Search terms

To ensure interventions were conducted by music therapists, the search terms were limited to “music\* therap\*” rather than “music” or “music-based interventions” (Appendix A). This accounted for any relevant specific forms of MT, such as NMT [5]. Robb et al. (2018) explain the challenges in comparing music intervention research due to poor reporting of interventions [6], therefore important to ensure interventions are evidence-based. Consequently, only evidence-based MT interventions delivered by trained music therapists or professionals qualified to practice NMT were included.

Duplicates were removed manually using the Endnote function. Relevant data was extracted, and 10 papers included, illustrated in the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) flowchart (Figure 2) [69]. The small number of papers may reflect a lack of RCTs conducted on MT, hence the recommendations for research in the NICE guidelines [40]. Abstracts of identified articles and full texts of studies were independently screened and reviewed by two reviewers (SG, NP) with disagreements resolved by discussion.



**Figure 2.** PRISMA flowchart demonstrating the process of paper identification for review [69].

### Data Synthesis and Quality Appraisal

The Revised Cochrane Risk-of-Bias version 2 Tool for Randomised Trials, recommended and designed for assessing risk of bias in RCTs [70–72] was used for quality appraisal. The Synthesis Without Meta-analysis (SWiM) reporting guidelines were applied to the synthesis of data to answer the review aim, accounting for limited reporting of results [23]. The standardised metric was vote count for direction of effect, tabulated and determined by  $p$ -value significance and direction [23].

Combining  $p$ -values was not possible as a precise  $p$ -value for each outcome was not consistently available [23]. Average intervention effect sizes in a meta-analysis was not possible without sufficient data [73]. Data extracted included study characteristics (sample size, gender, age, location, outcome measures, type of MT) (Table 3), followed by data synthesis, where possible extracting the direction of effect,  $p$ -values, confidence intervals, and calculating Cohen's  $d$  effect size using mean scores, standard deviation, and number of participants [23,74] (Appendix B).

### SWiM Approach

The characteristics table, with informal heterogeneity investigation, and data synthesis tables helped grouped neurological conditions to show variation across conditions [45; 47] (Table 3, Appendices A and B). The effect direction table also grouped studies by psychological outcome (Appendix C). Studies were further excluded in the synthesis if between-group analyses were not undertaken since the intervention could not be reliably compared with the comparator [75,76].

To assess certainty of findings, the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) principles were applied narratively [73,77,78]. Limitations of the vote count direction of effect did not provide data for analysis of GRADE domains such as consistency, therefore a rank was not applied [73]. Furthermore, with a small number of studies and variety of psychological outcome measures, quantifying proportion of effects and indicating sample size visually has limited value and risks deception.

## RESULTS

### Characteristics of the Included Studies

Ten studies were included in the review and their characteristics are displayed in Table 3. The sample sizes ranged from 27-82 participants, with an averaging age 41-73 years. Each study recruited one neurological condition, including stroke, multiple sclerosis (MS), Parkinson's disease (PD), traumatic brain injury (TBI), and Huntington's disease (HD). The treatment consisted of specific types of MT or NMT, delivered over 1-24 sessions. The MT and NMT interventions were delivered by trained music therapists, except for studies by Impellizzeri et al. [49, 79] where interventions were delivered by NMT trained neuropsychologists. Comparators were treatment-as-usual (standard multi-disciplinary neurorehabilitation), except four studies comparing MT to an alternative intervention [32,51-53].

The psychological outcomes were assessed using different self-report and observational standardised measures, at baseline and after intervention in all studies, with three studies evaluating follow-up outcomes at 3-6-month time points [32,50,53] (Table 3).



**Table 3.** Characteristics of the included studies.

	Study design	Sample size	Age (mean years)	Sex (female)	Neurological condition	Location and research setting	Outcome measures	Music Therapy intervention	Comparator
Chou et al. (2024) [81]	RCT pilot study	82	58	28%	Stroke	Taiwan, inpatient setting	BDI-II, MMSE, MRS, BI Timepoints: Before and after intervention	Neurologic Music Therapy – Therapeutic Singing, Melodic Intonation Therapy, Rhythmic Speech Cueing, Therapeutic Instrument Music Playing, Music Cognitive Training (from neurologic music therapy) (in addition to treatment as usual) Frequency: Four hours over four weeks (extra to neurorehabilitation as normal)	Conventional therapy (treatment as usual)
Haire et al. (2021) [52]	RCT	30	55.9	47%	Stroke	Toronto, Canada, community setting	TMT-B, FDST, GSES, MAAC-R, SAM Timepoints: Two baseline assessments one-week apart. One post-intervention assessment.	Therapeutic Instrumental Music Performance (TIMP) Frequency: Three times a week for three weeks	The comparator groups consisted of TIMP plus cued motor imagery and TIMP plus motor imagery without external cues
Poćwierz-Marciniak &	RCT	61	64	78.7%	Stroke	Gdynia, Poland, inpatient neurological	SF-36, SA-SIP30, Cantril Ladder	Cognitive Music Therapy, Guided Imagery and Music, 1:1	Standard care (physiotherapy, ergotherapy,

Bidzan (2017) [54]						rehabilitation hospital		Timepoints: Before and after intervention	Frequency: Twice a week for five weeks	psychological diagnosis, maintenance psychotherapy)
Raglio et al. (2017) [11]	RCT pilot	38	72.7	58%	Stroke	Pavia, Italy, inpatient neurological rehabilitation hospital		HADS, MQOL-It Timepoints: Before and after intervention	Relational Active Music Therapy (RAMT) Frequency: Three sessions per week, 20 sessions total	Standard care (physiotherapy, occupational therapy)
Segura et al. (2024) [32]	RCT	58	63.2	24%	Stroke	Barcelona, Spain, ex-inpatient neuro-rehabilitation		BRIEF, SART, Figural Memory subtest from the WMS-R, AVLT, Verbal Fluency test in Spanish, BDI-II, self- and informant-version of AES, POMS, SIS, TSRQ, IMI, Strategies Used to Promote Health Timepoints: Before and after intervention, with 3-month follow-up	Enriched Music-supported Therapy Frequency: Once a week music therapy, plus three weekly individual self-training session, for 10 weeks	Graded Repetitive Arm Supplementary Program (GRASP) only
Impellizzeri et al. (2020) [79]	RCT pilot study	30	51	37%	Multiple Sclerosis	Messina, Italy, clinic setting		BRB-N, MSQOL-54, BDI, EAQ, MMF Timepoints: Before and after intervention	Neurologic Music Therapy – Associative mood and memory training, Music in psychosocial training and counselling (half of the	Treatment-as-usual (same number of sessions as the music therapy group)

										treatment-as-usual time replaced with music therapy intervention) Frequency: Three times per week for 8 weeks		
Impellizzeri et al. (2024) [49]	Pilot Quasi-RCT	40	62.45	30%	Parkinson's disease	Messina, Italy, clinic centre setting	MoCA, HRSD, FAB, Stroop test, Visual search test			Computer-Assisted Rehabilitation Environment (CAREN), Rhythmic Auditory Stimulation, Therapeutic Instrumental Music Performance	Standard treatment with CAREN selected scenarios three times per week	
Lee et al. (2024) [51]	RCT	27	73.3	52%	Parkinson's disease	Arizona, USA, Tremble Clefs therapeutic singing group	HY, GDS, VRQOL, VASM	Timepoints: Before and after intervention		Therapeutic Group Singing (TGS), Straw Phonation Combined with Therapeutic Singing (SP+TGS)	Speaking-only control group	
Siponkoski et al. (2020) [50]	Cross-over RCT	40	41.3	41%	Traumatic Brain Injury	Helsinki, Finland, brain injury clinic setting	FAB, Number-Letter Task, Auditory N-back Task, Simon Task, SART, Similarities, Block Design, and Digit Span subtests of the WAIS-IV, Words Lists I and II subtests of the WMS-III	Timepoints: Before and after intervention,		Rhythmical Training, Structured Cognitive-motor Training, Assisted music playing	Standard care (physiotherapy, occupational therapy, neuropsychological rehabilitation, speech therapy)	

							follow-up (3 and 6 months)			
Van Bruggen-Rufi et al. (2017) [53]	RCT	63	54.4	68.3%	Huntington's disease	Netherlands, set in four specialised Huntington's disease care facilities	BOSH – social-cognitive subscale and mental rigidity and aggression subscale, PBA	– social-functioning and the communication and self-expression"	Followed protocol "music therapy for Huntington's patients on improving and stimulating communication and self-expression"	Recreational therapy (with treatment guide offered in same circumstances as music therapy group e.g. reading the newspaper, cooking, arts and crafts, handwork, puzzles/games)

Outcome measures: Beck Depression Inventory-II (BDI-II), Mini-Mental State Examination (MMSE), Modified Rankin Scale (MRS), Barthel Index (BI), Trail-making test part B (TMT-B), Forward digit-span test (FDST), General self-efficacy scale (GSES), Multiple affect adjective check list revised (MAACL-R), Self-assessment manikin (SAM), Brief repeatable battery of neuropsychological test (BRB-N), Multiple sclerosis quality of life-54 (MSQOL-54), Beck depression inventory (BDI), Emotion awareness questionnaire (EAQ), McClelland motivational factors (MMF), Montreal Cognitive Assessment (MoCA), Hamilton Rating Scale for Depression (HRSD), Frontal Assessment Battery (FAB), Hoehn and Yar Stage Scale (HY), Geriatric Depression Scale (GDS), Voice-Related Quality of Life (VRQOL), Visual Analogue Scale for Mood (VASM), Medical outcome study 36-item short-form health survey (SF-36), Stroke-adapted 30-item version of the sickness impact profile (SA-SIP30), Hospital Anxiety and Depression Scale (HADS), Italian version of McGill Quality-of-Life Questionnaire (MQOL-It), Behaviour Rating Inventory of Executive Function (BRIEF), Sustained Attention Response Task (SART), Wechsler Memory Scale-Revised (WMS-R), Rey Auditory Verbal Learning Test (AVLT), Apathy Evaluation Scale (AES), Profile of Mood States (POMS), Stroke Impact Scale (SIS), Treatment Self-Regulation Questionnaire (TSRQ), Intrinsic Motivation Inventory (IMI), Sustained Attention to Response Task (SART), Wechsler Adult Intelligence Scale IV (WAIS-IV), Wechsler Memory Scale III (WMS-III), Behaviour Observation Scale Huntington (BOSH), Problem Behaviours Assessment-short version (PBA).

## Quality Appraisal and Risk of Bias

The Cochrane Risk of Bias (RoB) 2 tool [72] was used to evaluate the quality of the included studies (Appendix D).

### *Randomisation*

All participants were randomly allocated into intervention and comparator groups, using computerised randomisation methods to minimise bias. Despite this, two studies [11,49] did not detail the randomisation process or whether allocation was concealed, but deemed low risk of bias with no baseline group differences [72,80].

Three studies indicated a single difference between intervention and comparator groups at baseline (49, 50, 53). Thus, Van Bruggen-Rufi et al. (2017) [53] found one significant difference for Total Functional Capacity [this classification separates Huntington disease stage (as an inclusion criterion or endpoint in clinical trials accepted by the Food and Drug Administration), but no difference when adjusting means, suggesting it did not confound results. Siponkoski et al. (2020) [50] reported the cause of injury differed between groups but this item was not clinically important. These single baseline differences are therefore likely attributed to chance [49,50,72].

### *Deviations from Intended Interventions*

Concealing group allocation from participants and intervention deliverers was not possible due to the nature of MT. Inpatient settings are relatively controlled environments compared to community, and although it may be easier to monitor deviations from intended interventions, control group participants are more likely to be exposed to the intervention accidentally, increasing risk of bias [55] (Appendix D).

MT interventions were adapted to participant-centred needs, adjusting difficulty levels and intensity [32,50]. This would be expected in clinical practice and therefore was not deemed biased, and not adapting to needs could bias a negative effect. Adapting to participant needs without biasing intervention relies on therapist experience, hence the importance for trained music therapists to deliver MT.

In Chou et al. (2024) [81], one (therapeutic instrument music playing) of five (therapeutic singing, melodic intonation therapy, therapeutic instrument music playing, rhythmic speech cueing and music cognitive training according to the patient's abilities) MT interventions were not completed for an unknown number of participants due to time constraints, potentially impacting magnitude and direction of effect. Similarly, Haire et al. (2021) [52] carried out nine of the 12 planned intervention sessions, but this was changed prior to starting the intervention without deviation, unlike Chou et al. (2024) [81].

Three studies saw withdrawals, and intention-to-treat (ITT) analyses were used to minimise bias (32, 50, 53). Siponkoski et al. (2020) [50] used multiple imputation, considered a reliable method with missing data [82]. Segura et al. (2024) [32] used multiple imputation with a high missing data rate, reducing reliability, but found similar results between ITT and per-protocol analyses, strengthening reliability, while reporting the ITT analysis, deemed the least biased [32,41,83].

### *Missing Outcome Data*

All missing data was accounted for across studies. Siponkoski et al. (2020) [50] considered their data missing at random and used parallel datasets to minimise sampling variability [72] Despite this, in Van Bruggen-Rufi et al. (2017) study [53] some participants withdrew due to lack of motivation. However, the authors did not state the numbers in each group, therefore it is not clear if this reason relates to MT or lack of MT. However, linear mixed model analyses are less sensitive to missing data, therefore this may have impact on the results [53,84].

### *Measurement of the Outcome*

Consistent standardised measures were used to evaluate psychological outcomes before and after intervention for all participant groups across studies. However, given the range of studies used and locations included in this review, some measures have reduced validity and reliability following translation and cultural adaptations. For example, Segura et al. (2024) [32] translated the Profile of Mood States scale into Spanish, where meaning of psychological constructs may differ from the original measure in English, and therefore should be considered during interpretation [85]

Concealment of allocation was not possible for participants and therapists, raising some concerns for most studies. Bias reduced by concealing participant allocation from assessors across studies, with three exceptions. Lee et al. (2024) [51] assessors were aware of allocation and familiar with participants, increasing potential bias for participants seeking to please facilitators when assessing outcomes. Similarly, in van Bruggen-Rufi et al. study (2017) [53], some assessors were blinded to participant allocation but nursing staff assessing behaviour outcomes may have been aware and influenced by allocation [72]. Poćwierz-Marciniak & Bidzan (2017) [54] assessors also delivered the intervention; however, measures were self-report not observer-dependent, minimising this risk of bias. Nonetheless, allocation awareness may have influenced participant-reported outcomes, raising some concerns for all studies except Siponkoski et al. (2020) [50] where no self-report measures were conducted, and assessors were blinded (Appendix D).

### *Selection of the Reported Results*

Some studies bias towards reporting significant over non-significant findings [45,71]. For example, Impellizzeri et al. (2024) [49] do not discuss the non-significant Hamilton Depression Scale outcome, and only report statistically significant  $p$ -values without clear reporting of full statistical tests, suggesting optimism bias [86]. Similarly, several studies discuss significant outcomes where only select subtests for the outcome were significant, without balance for non-significant subtests [32,51–54,79]. This indicates bias towards rejection of the null hypothesis, increasing risk of type I error [87].

Finally, studies carrying out an ITT analysis demonstrated how planned analysis was altered to account for participant drop-outs [32,50]. Van Bruggen-Rufi et al. (2017) [53] followed the analytical procedure in the separately published protocol, but other studies did not detail planned analysis or refer to protocols [11,49,51,52,54,79,81,88]. As a result, there is potential bias towards reporting of significant findings, however, reported results were based on the planned outcomes measured, suggesting low risk of bias from changes to analysis [72].

### **Data Synthesis and Key Findings**

Seven studies included in the vote count of estimated direction of effect were used to determine whether MT had an effect for each psychological outcome. Data related to the research question was extracted, excluding data irrelevant to the review although considered when interpreting author's conclusions [11,32,50–52]. Effect size has been calculated for outcomes where data was provided, indicating the magnitude of effect for individual outcomes, but these were not synthesised due to lack of data across studies. If time constraints allowed authors of studies with missing data to estimate effect sizes interpreted using Cohen's guidelines of small ( $d>0.2$ ), medium ( $d>0.5$ ), and large ( $d>0.8$ ) effect [74] would have been contacted to inform this. Post-hoc power was not reported in the studies, and a priori power calculations were minimal due to the nature of pilot and small studies, increasing risk of type I and II errors [50,89]. This may account for the varying findings and cautious interpretations across studies.

### *Within-group Findings*

Within-group and between-group findings differed in two studies [49,81]. Impellizzeri et al. (2024) [49] found some subtests showed significant improvement for cognitive function following MT, whereas within-group MT results were consistently significant. Similarly, Chou et al. (2024) [81] found between-group cognitive function non-significant, whereas within-group findings were significant following MT. The increased significance for within-group MT findings indicates an effect of MT, but may not be more effective than the comparator, suggesting reduced power and increased risk of type II error [90]. Therefore, within-group significant effects cannot be used to solely conclude the success of MT [75,90].

### *Between-group Findings*

Between-group data comparing groups was used to indicate direction of effect (Appendix C) (75, 76). Haire et al. study (2021) [52] was not included in the synthesis for direction of effect due to all groups including MT, hence no relevant between-group data. Impellizzeri et al. (2020) [79] and Lee et al. (2024) [51] studies were also excluded from the data synthesis due to between-group comparisons being based on within-group data, rather than comparative statistical analysis.

**Cognitive Function.** Cognitive function was measured across all included neurological conditions, increasing heterogeneity for measures, population, and intervention, which may explain inconsistent findings (Table 3). Some aspects of cognitive function significantly improved following MT for PD and TBI, whereas no effect was found for stroke despite very small or no effect size ( $d = 0.00$ ) [49; 50, 81]. One study used virtual reality (VR) MT interventions, potentially confounding comparability to non-VR MT, but remained included in data synthesis because the comparison group also received VR, therefore controlled for [49].

In contrast, cognitive function significantly improved for the comparator group compared to the MT group in HD, suggesting MT is safe but recreational therapy (comparator) was more effective for improving cognitive function [53,55]. Alternatively, this may be attributed to the MT group having lower baseline cognitive function than the comparator, although not deemed a confounding factor, or the varying nature of advanced HD participants included [53]

**Mood.** Mood was measured using a variety of measures including the Hospital Anxiety and Depression Scale (HADS), Beck Depression Inventory-II, and Profile of Mood State (POMS) [11, 32, 81]. Following MT intervention there was no effect found for stroke, but small effect sizes ( $d = 0.02$ , 0.04 for depression and  $d = 0.24$  for anxiety) [11,32,81]. This reflects participants' variability, small sample sizes, and varying measures that have enabled an effect not to be found. Additionally, the measures have been validated in Taiwan and Italy, but should be interpreted with caution when measuring anxiety and depression separately with the HADS, and translating the POMS into Spanish, where meaning of constructs may differ, suggesting conclusions cannot be drawn reliably [11,32,85,91–93].

**Emotion.** Emotion was measured for one study researching stroke participants, finding significant improvements in emotional well-being following MT intervention with a medium effect size ( $d = 0.69$ ) post-intervention [32]. A non-significant result was found at three-months follow-up, with very small effect size however ( $d = 0.10$ ), suggesting the improvement may not have sustained but the difference is negligible [32]. These were estimated based on a subtest result, rather than overall test battery conclusions from the Stroke Impact Scale [32]. Single subtests can be interpreted effectively to measure change, but further analysis is required to establish effect and impact [94].

**Behaviour.** Change in behaviour was measured by one study, observing no significant effect of MT for HD [53]. The results may differ if compared to treatment-as-usual rather than an alternative therapy [95,96]. However, the behaviour measures may not be sensitive to changes for the advanced HD participants [53]. Therefore, concluding the effect of MT cannot be confirmed.

**Quality of Life.** Quality of life found significant improvement for MS and post-stroke [11,54]. However, one study based its findings on subtests rather than overall measures [54]. Therefore, effect size is not known, and the significant findings reported do not represent the breadth of the outcome

[54,97]. The estimated effect size for Raglio et al. (2017) [11] finding was calculated as small ( $d = 0.18$ ) indicating reduced applicability in practice.

### **Certainty of Evidence using GRADE**

#### ***Risk of Bias***

The risk of bias increases for outcomes using self-report measures, such as mood and quality of life [72]. Behaviour is more likely assessed using observer-reported judgement, increasing bias except when assessors are blinded to allocation, whereas cognitive function is more likely assessed objectively, minimising bias [72,98].

There is potential bias from type of MT and its impact on an outcome. For example, Chou et al. (2024) [81] used speech and language-related NMT, whereas Raglio et al. (2017) [11] used relational NMT that facilitates rapport between the therapist and participant [99]. Both studies measured mood outcomes, however, the interventions approached different intentions, suggesting this heterogeneity may cause the outcomes from different approaches improving mood with differing effects.

#### ***Inconsistency***

Statistical measures such as  $I^2$  or Cochran's  $p$ -value cannot be used to establish consistency due to heterogenic data [100]. However, inconsistencies in effect direction can be explained by the known variation across PICOS and the likely lack of power due to small sample sizes across studies reducing the likelihood of finding effect (45, 47, 78, 90).

#### ***Indirectness***

The MT interventions used across studies are evidence-based NMT techniques and protocols applicable to clinical practice, except Impellizzeri et al. (2024) [49], which incorporated VR into MT, less widely implemented clinically. Similarly, the quantity of intervention varied, such as Chou et al. (2024) [81] receiving MT extra to treatment-as-usual, which may confound findings by suggesting input intensity influences effect rather than MT. Alternatively, Impellizzeri et al. (2020) [79] replaced half of the treatment-as-usual therapy with MT, as recommended by NICE (2023) [40]. This increases applicability to settings where funding for MT replaces treatment-as-usual, rather than added [7].

The neurological conditions studied reflect service provision that may consider including MT into the multi-disciplinary team [30]. The treatment-as-usual comparators reflect clinical practice, and no study compared MT to no treatment, deemed unrealistic and unethical [95, 96, 101].

#### ***Imprecision***

Confidence intervals were not available to clearly determine the certainty of precision across studies. However, in Chou et al. (2024) [81], the confidence interval for cognitive function and mood between-groups is much wider than the mean difference, and span positive and negative values, indicating reduced certainty of effect. Sample heterogeneity and size reduces statistical power and therefore precision.

Precision cannot be determined from power due to lack of reporting. Four of the ten studies are pilots, less likely to meet statistical power than main studies, however main study sample sizes remained small, with a pilot study obtaining the largest sample included [81,89]. Power calculations were carried out a priori for three studies, where Chou et al. (2024) [81] had a sample size that met criterion, but Siponkoski et al. (2020) and van Bruggen-Rufi et al. (2017) criterion [50,53] were not sustained due to drop-out rates, reducing precision.

#### ***Publication Bias***



Magnitude of effect was not clear, risking inflated conclusions, particularly given effect sizes were small where available [11,52,78,81]. Furthermore, main studies did not differ in precision or sample size from pilots, risking exaggerated generalisability from reduced certainty.

### *Overall Certainty of Evidence and Importance of Outcome*

The overall certainty of evidence was low, which can be attributed to the limited reporting of effect size, small sample sizes, and heterogeneity within the elements of PICOS [45,47]. This outcome suggests findings should be interpreted with caution but does not disregard the potential effects of MT on psychological outcomes for neurological conditions, given the limitations in the consistency, precision, and bias within and between the studies [55].

### **Discussion**

This systematic review aimed to establish whether there is an effect of MT on psychological outcomes for neurological conditions. In summary, cognitive function was measured across all included neurological conditions in this review, showing MT improved cognitive function in PD and TBI, but not for stroke or HD, with HD showing improvement in the comparator group [32,49,50,53,81]. A significant improvement was found for emotional well-being, but not mood following MT for people who had stroke [32,81]. Behaviour was not found to significantly improve following MT in HD. However, quality of life improved in MS and stroke following MT [11,53,54].

All studies included randomised participants. However, participants and therapists could not be blinded from allocation, thus increasing risk of bias. Missing outcome data was accounted for across studies, and standardised measures were used across studies, but not always valid and reliable for the population studied, again increasing risk of bias. Studies often limited reporting to significant findings, and had limited transparency between planned and actual analyses, suggesting optimism bias [87]. One study [81] indicated high risk, and one low risk [50], with the remaining eight studies having some concerns of bias, (Appendix D).

The GRADE discussion suggests reduced certainty due to risk of bias increasing for self-reported outcomes but is improved through assessor blinding. Consistency of direction of effect is limited by heterogeneity across all PICOS elements, despite being relevant to clinical practice, improving directness [45,47]. Precision is reduced with wide confidence intervals when provided and no potential power differences between pilot and main studies. Publication bias may be impacted by inflated conclusions from lack of effect size reporting.

The findings in this review are similar to previous reviews. MT is perceived as a positive intervention that has potential impact on psychological outcomes, but PICOS heterogeneity reduce findings of group difference and certainty of findings, which are limitations across MT research [6,36,45,47,62]. Neurological conditions range across studies, with differing severities, from different countries. This heterogeneity means that consideration for cultural applicability of the MT intervention and the meaning and translation validity of outcome measures, including specific subtests, is required [16,102]. Similarly, the definition of "treatment-as-usual" for comparator groups is rarely clarified, suggesting there may be variations in the comparator group and the extent of similarity to the intervention, potentially reducing certainty of results [103,104].

The current review is limited in synthesising outcomes using vote counting for direction of effect [23]. It does not account for extent or magnitude of effect, so the risk of type I or II error is uncertain, and also limits the applicability of GRADE to assess certainty of outcomes [23]. Similarly, not all studies were included in the data synthesis due to lack of reporting of between-group results and homogeneity between treatment intervention and comparator intervention [51, 52; 79]. This reduces an already small number of studies included, reducing the value in vote counting direction of effect, since proportion of effects could not be determined reliably. Therefore, there is reduced generalisability of findings and certainty across studies.

The limited papers found for this review are an attempt to reduce heterogeneity by narrowing MT practice, rather than general music-based interventions, but still presents a range of interventions

[62]. However, narrowing the intervention criteria to specific NMT may not account for the diversity of methods and symptoms within disorders that MT can apply [36,45,47],

Despite recommendations for RCTs to more accurately determine the effectiveness of MT, the RCTs reviewed have limitations in quality and reporting [6,34,36,40]. Since the heterogenic nature of these studies cannot be controlled for without reducing clinical applicability, future research should focus on larger-scale RCTs to account for diversity, with comparison of MT to treatment-as-usual that is explicitly defined and distinguished from the intervention [28,36,105]. MT is well established for dementia compared to other neurological conditions, enabling increased access to participants [28]. Hence, funding is required to enable MT to be established for large-scale RCT research to be conducted.

Considering the limited quality of evidence and reporting of findings, it has been established that MT has potential for positive effects on psychological outcomes. However, the certainty of findings is low and, therefore, conclusions about MT effect on psychological outcomes for neurological conditions cannot be drawn.

## Conclusion

This review demonstrates the potential effect of MT to improve psychological outcomes including cognitive function, mood, emotion, behaviour, and quality of life, for people living with neurological conditions. RCTs are limited in certainty and bias due to the heterogenic nature of MT, neurological conditions, and measures of psychological outcomes, but often carried out on a small scale.

MT has the potential to be a cost-effective, safe, and valuable intervention that could minimise unnecessary pharmacotherapy [14,55]. Future research should endeavour to define and measure the interventions and corresponding effects on a large scale, to better inform intervention outcomes and prevent missed opportunity that MT could offer.

## Appendix A

*Search terms inputted into the four databases.*

---

Search terms for PubMed and Scopus:

[“music\* therap\*”] AND [neurorehab\* OR neuro-rehab\* OR “neuro\* rehab\*” OR neurologic\* OR “neurologic\* condition\*” OR “neurologic\* disorder\*” OR Parkinson\* OR “brain injur\*” OR TBI OR “traumatic brain injur\*” OR ABI OR “acquired brain injur\*” OR “brain tumour\*” OR “brain tumor\*” OR “multiple sclerosis” OR MS OR stroke\* OR encephal\* OR epileps\* OR “motor neurone disease\*” OR MND OR Huntington\* OR HD OR “disorder\* of conscious\*” OR DOC OR “minimal\* conscious\*” OR vegetative OR “unresponsive wakeful\* syndrome\*”]

---

Search terms for PsycInfo and CINAHL:

---

---

[music\* therap\*] AND [neurorehab\* OR neuro-rehab\* OR neuro\* rehab\* OR neurologic\* OR neurologic\* condition\* OR neurologic\* disorder\* OR Parkinson\* OR brain injur\* OR TBI OR traumatic brain injur\* OR ABI OR acquired brain injur\* OR brain tumour\* OR brain tumor\* OR multiple sclerosis OR MS OR stroke\* OR encephal\* OR epileps\* OR motor neurone disease\* OR MND OR Huntington\* OR HD OR disorder\* of conscious\* OR DOC OR minimal\* conscious\* OR vegetative OR unresponsive wakeful\* syndrome\*]

---

*Note.* Search fields were limited to title and abstract for all databases, with the addition of keywords in PsycINFO and Scopus.

Search terms were identified from previous literature and neurorehabilitation service provisions, alongside support from a specialist librarian at the University of Leicester

## Appendix B

Overview of synthesis and summary of findings from included studies.

	Neurological condition	Psychological outcome	Reported effect direction	Effect size (Cohen's <i>d</i> )	<i>p</i> -value	Mean difference [95% CI]
Chou et al. (2024) [81]	Stroke	Cognitive function	↑	Not reported.	<0.001 (within-group)	1.04 [0.51-1.57]
			→	Calculated as 0.40 (within-group intervention), 0.00 (between-group)	0.355 (between group)	0.02 [-2.88-2.83]
		Mood (depression)	→	Not reported.	0.129 (within-group)	-0.84 [-1.93-0.25]
			→	Calculated as 0.16 (within-group intervention), 0.02 (between-groups)	0.740 (between-group)	-0.05 [-2.80-2.90]
Haire et al. (2021) [52]	Stroke	Cognitive function	→	Not reported.	Trail-making test:	Not reported
				Trail-making test: Calculated as 0.12 (within-group MT only) 0.27 (within-group MT+cMI) 0.27 (within-group MT+MI) Digit span test: Calculated as 0.09 (within-group MT only) 0.00 (within-group MT+cMI) 0.10 (within-group MT+MI)	Trail-making test: 1.00 (MT only) 0.125 (MT+cMI) <0.05 (MT+MI) Digit span test: 0.459 (MT only) 0.865 (MT+cMI) 0.669 (MT+MI)	
		Affect	→	Not reported.	0.105 (MT only) <0.05 (MT+cMI)	Not reported

				Calculated as 0.34 (within-group MT only)	0.261 (MT+MI)	
				0.40 (within-group MT+cMI)		
				0.47 (within-group MT+MI)		
	Mood	→		Not reported. Calculated as 0.28 (within-group MT only)	0.147 (MT only)	Not reported
					<0.05 (MT+cMI)	
				0.90 (within-group MT+cMI)	0.492 (MT+MI)	
				0.05 (within-group MT+MI)		
	Self-efficacy	→		Not reported. Calculated as 0.21 (within-group MT only)	0.202 (MT only)	Not reported
					0.098 (MT+cMI)	
				0.35 (within-group MT+cMI)	1.00 (MT+MI)	
				0.02 (within-group MT+MI)		
Póćwierz-Marciniak & Bidzan (2017) [54]	Quality of life	↑*		Not reported. Not calculated due to multiple measures used for outcome.	<0.05 (between-groups) for some scales	Not reported
Raglio et al. (2017) [11]	Quality of life	↑		Not reported. Calculated as 0.18 (between-group)	0.19 (between-group)	Not reported
	Mood (depression)	↑		Not reported. Calculated as 0.04 (between-group)	<0.05 (between-group)	Not reported
	Mood (anxiety)	→		Not reported. Calculated as 0.24 (between-group)	0.25 (between-group)	Not reported
Segura et al. (2024)	Cognitive functions	→*		Not reported. Not calculated due to volume of measures for outcome.	>0.05 on subscales (between-group) at post-intervention or follow-up.	Refer to original paper, not reported here due to volume of measures used.
	Emotion	↑*		Not reported. Not calculated as 0.69 (between-group) post-intervention, 0.10 (between-group) at follow-up.	<0.05 subscale (between-group) at post-intervention. Not significant at follow-up.	Refer to original paper, not reported here due to

					volume of measures used.	
	Mood	↑*	Not reported. Not calculated due to volume of measures for outcome.	<0.05 on one subscale (between-group) at post-intervention. Not significant at follow-up.	Refer to original paper, not reported here due to volume of measures used.	
Impellizzeri et al. (2020) [79]	Multiple sclerosis	Cognitive function	↑*	Not reported. Not enough information provided (no standard deviation)	Overall battery outcome not reported. 4/9 subscales <0.05 within-group experimental) otherwise non-significant.	No overall outcome results provided.
	Mood (depression)	↑	Not reported. Not enough information provided (no standard deviation)	<0.05 (within-group experimental) 0.278 (within-group control)	5.60 [3.7-7.72] -0.66 [-1.93-0.60]	
	Emotion	↑*	Not reported. Not enough information provided (no standard deviation) No overall outcome results provided.	<0.05 for all subscales (within-group experimental)	No overall outcome results provided.	
	Quality of life	↑	Not reported. Not enough information provided (no standard deviation).	<0.05 for all subscales (within-group experimental)	No overall outcome results provided.	

				↑			
Impellizzeri et al. (2024) [49]	Parkinson's disease	Cognitive function		↑	Not reported. Reported the median, skewed data effect sizes to be discussed with statistician if time allowed.	<0.05 (within-group experimental)	Not reported for change scores.
				↑*		<0.05 within some sub-scales (between- group)	
Lee et al. (2024) [51]	Parkinson's disease	Mood		↑*	Not reported. Not calculated due to multiple subtests for outcome.	<0.05 sad, anxious, angry (within-group) 0.127 happy (within-group)	Not reported
Siponkoski et al. (2020) [50]	Traumatic brain injury	Cognitive function		↑*	Not reported. Executive function - $\eta^2p=0.093$ (medium-large) Set-shifting - $\eta^2p=0.112$ (medium-large) Note these are not directly comparable to Cohen's <i>d</i> .	<0.05 (between-group) for executive functions, >0.05 (between-group) set-shifting, >0.05 (between-group) for reasoning and verbal memory	
Van Bruggen-Rufi et al. (2017) [53]	Huntington's disease	Social-cognitive function		↓	Not reported. Not calculated due to limited data available.	<0.05 (between-groups)	2.88 [0.108-5.65]
		Behaviour		→	Not reported. Not calculated due to limited data available.	0.125 (between groups) on BOSH 0.630 (between groups on PBA-S)	4.60 [-1.32-10.52] -1.39 [-7.16-4.38]

↑ = music therapy significantly improved the outcome, over the control group; → = no significant effect of music therapy was found for the outcome, over the control group; ↓ = the control group significantly improved the outcome, over music therapy; \* = the direction of effect is not based on one standardised measure, but a number of subtests where a proportion showed significant results. MI=motor imagery, cMI=metronome-cued motor imagery, CI=confidence interval.

## Appendix C

Effect direction plot for between-group effects of psychological outcomes for neurological condition (Schünemann et al., 2019).

Studies (outcome(s) measured in study)	Neurological condition	Cognitive function (CF)	Mood (M)	Emotion (E)	Behaviour (B)	Quality of life (QoL)
Chou et al. (2024) (CF, M) [81], Poćwierz- Marciniak & Bidzan (2017) (QoL) [54], Raglio et al. (2017) (M, QoL) [11], Segura et al. (2024) (CF, E, M) [32]	Stroke	→*	→	↑*	-	↑*
Impellizzeri et al. (2024) (CF) [49]	Parkinson's disease	↑*	-	-	-	-
Siponkoski et al. (2020) (CF) [50]	Traumatic brain injury	↑*	-	-	-	-
van Bruggen-Rufi et al. (2017) (CF, B) [53]	Huntington's disease	↓	-	-	→	-

Note. Includes 7/10 studies, excluded Haire et al. (2021) [52] due to the lack of comparator group not containing music therapy, and Impellizzeri et al. (2020 [79] and Lee et al. (2024) [51] due to only reliably reporting within-group. ↑ = music therapy significantly improved the outcome, over the control group; → = no significant effect of music therapy was found for the outcome, over the control group; ↓ = the control group significantly improved the outcome, over music therapy; \*= the direction of effect is not based on one standardised measure, but a number of subtests where a proportion showed significant results.



## Appendix D

Cochrane Risk of Bias (RoB) 2 summary [72].

	Randomisation process	Deviations from intended interventions	Missing outcome data	Measurement of outcome	Selection of the reported result	Overall risk of bias
Chou et al. (2024) [81]	Low risk (randomised, no group differences at baseline)	Some concerns	Low risk	Some concerns (assessor blinded to participant allocation, self-report measures used)	Low risk	High risk
Haire et al. (2021) [52]	Low risk (randomised, no group differences at baseline)	Low risk	Low risk	Some concerns (assessor blinded to participant allocation, self-report measures used)	Low risk	Some concerns

Impellizzeri et al. (2020) [79]	Low risk (randomised, no group differences at baseline)	Low risk	Low risk	Some concerns (assessor blinded to participant allocation, used self-report measures)	Low risk	Some concerns
Impellizzeri et al. (2024) [49]	Low risk (randomised, differences in Stroop test at baseline, likely due to chance)	Low risk	Low risk	Some concerns (assessor blinded to participant allocation, uses one self-report measure)	Low risk	Some concerns
Lee et al. (2024) [51]	Low risk (randomised, no group differences at baseline)	Low risk	Low risk	Some concerns (assessors not blinded to allocation and familiar with participants, potential for bias)	Low risk	Some concerns

Poćwierz- Marciniak & Bidzan (2017) [54]	Low risk (randomised, no group differences at baseline)	Low risk	Low risk	Some concerns (assessor not blinded to allocation as they delivered the intervention, used self-report measures)	Low risk	Some concerns
Raglio et al. (2017) [11]	Low risk (randomised, no group differences at baseline)	Low risk	Low risk	Some concerns (assessors blinded to participant allocation, used self-report measures)	Low risk	Some concerns
Segura et al. (2024)	Low risk (randomised, no group differences at baseline)	Low risk (intention-to-treat analysis carried out)	Low risk	Some concerns (assessors blinded to participant allocation, used self-report measures)	Low risk	Some concerns

Siponkoski et al. (2020) [50]	Low risk (randomised, only group difference was deviation on cause of injury, but not considered clinically important and attributed to chance)	Low risk (intention-to-treat analysis carried out)	Low risk	Low risk (assessors blinded to participant allocation)	Low risk	Low risk
Van Bruggen-Rufi et al. (2017) [53]	Low risk (randomised, no group differences at baseline)	Low risk (intention-to-treat analysis considered)	Low risk (unclear which groups the participants were allocated to that withdrew due to lack of motivation)	Some concerns (some assessors blinded to participant allocation, not the nursing staff assessing behaviour, observer-reported assessment requiring judgement)	Low risk	Some concerns

*Note.* Low risk = the study is low risk of bias for all domains. Some concerns = the study raises some concerns in at least one domain, but not at high risk for any domain. High risk = the study is at high risk of bias in at least one domain, or the study has some concerns for multiple domains.

## References

1. Devlin, K., Alshaikh, J. T., & Pantelyat, A. (2019). Music therapy and music-based interventions for movement disorders. *Curr Neurol Neurosci Rep*, *19*(11), 83-95. <https://doi.org/10.1007/s11910-019-1005-0>
2. De Witte, M., Pinho, A. D. S., Stans, G. J., Moonen, X., Bos, A. E., & Van Hooren, S. (2022). Music therapy for stress reduction: a systematic review and meta-analysis. *Health Psychol Rev*, *16*(1), 134-159. <https://doi.org/10.1080/17437199.2020.1846580>
3. Hurkmans, J., de Bruijn, M., Boonstra, A. M., Jonkers, R., Bastiaanse, R., Arendzen, H., & Reinders-Messelink, H. A. (2012). Music in the treatment of neurological language and speech disorders: A systematic review. *Aphasiology*, *26*(1), 1-19. <https://doi.org/10.1080/02687038.2011.602514>
4. Koelsch, S. (2009). A neuroscientific perspective on music therapy. *Ann NY Acad Sci*, *1169*(1), 374-384. <https://doi.org/10.1111/j.1749-6632.2009.04592.x>
5. Leins, A. K., & Spintge, R. (2008). Music therapy in medical and neurological rehabilitation settings. In S. Hallam, I. Cross, & M. Thaut (Eds.), *The Oxford Handbook of Music Psychology* (pp. 526-535). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199298457.013.0049>
6. Robb, S. L., Hanson-Abromeit, D., May, L., Hernandez-Ruiz, E., Allison, M., Beloat, A., Daughtery, S., Kurtz, R., Ott, A., Oyedele, O. O., Polasik, S., Rager, A., Rifkin, J., & Wolf, E. (2018). Reporting quality of music intervention research in healthcare: A systematic review. *Complement Ther Med*, *38*, 24-41. <https://doi.org/10.1016/j.ctim.2018.02.008>
7. Grau-Sánchez, J., Jamey, K., Paraskevopoulos, E., Dalla Bella, S., Gold, C., Schlaug, G., Belleville, S., Rodríguez-Fornells, A., Hackney, M. E., & Särkämö, T. (2022). Putting music to trial: Consensus on key methodological challenges investigating music-based rehabilitation. *Ann NY Acad Sci*, *1518*(1), 12-24. <https://doi.org/10.1111/nyas.14892>
8. Howlett, J. R., Nelson, L. D., & Stein, M. B. (2022). Mental health consequences of traumatic brain injury. *Biol Psychiatry*, *91*(5), 413-420. <https://doi.org/10.1016/j.biopsych.2021.09.024>
9. McCaffrey, T. (2018). Evaluating music therapy in adult mental health services: Tuning into service user perspectives. *Nord J Music Ther*, *27*(1), 28-43. <https://doi.org/10.1080/08098131.2017.1372510>
10. Moore, K. S. (2013). A systematic review on the neural effects of music on emotion regulation: Implications for music therapy practice. *J Music Ther* *50*(3), 198-242. <https://doi.org/10.1093/jmt/50.3.198>
11. Raglio, A., Zaliani, A., Baiardi, P., Bossi, D., Sguazzin, C., Capodaglio, E., Imbriani, C., Gontero, G., & Imbriani, M. (2017). Active music therapy approach for stroke patients in the post-acute rehabilitation. *Neurol Sci*, *38*(5), 893-897. <https://doi.org/10.1007/s10072-017-2827-7>
12. Thompson, N., Iyemere, K., Underwood, B. R., & Odell-Miller, H. (2023). Investigating the impact of music therapy on two in-patient psychiatric wards for people living with dementia: retrospective observational study. *BJPsych Open*, *9*(2), 1-8. <https://doi.org/10.1192/bjo.2023.20>
13. Wilson, L., Horton, L., Kunzmann, K., Sahakian, B. J., Newcombe, V. F., Stamatakis, E. A., von Steinbuechel, N., Cunitz, K., Covic, A., Maas, A., Van Praag, D., & Menon, D. (2021). Understanding the relationship between cognitive performance and function in daily life after traumatic brain injury. *J Neurol Neurosurg Psychiatry*, *92*(4), 407-417. <https://doi.org/10.1136/jnnp-2020-324492>
14. Brancatisano, O., Baird, A., & Thompson, W. F. (2020). Why is music therapeutic for neurological disorders? The Therapeutic Music Capacities Model. *Neurosci Biobehav Riv*, *112*(1), 600-615. <https://doi.org/10.1016/j.neubiorev.2020.02.008>

15. Leins, A. K., & Spintge, R. (2008). Music therapy in medical and neurological rehabilitation settings. In S. Hallam, I. Cross, & M. Thaut (Eds.), *The Oxford Handbook of Music Psychology* (pp. 526-535). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199298457.013.0049>
16. Machado Sotomayor, M.J., Arufe-Giráldez, V., Ruíz-Rico, G., & Navarro-Patón, R. (2021). Music therapy and Parkinson's disease: A systematic review from 2015–2020. *Int J Environ Res Public Health*, *18*(21), 11618–11633. <https://doi.org/10.3390/ijerph182111618>
17. Magee, W. L. (2019). Why include music therapy in a neuro-rehabilitation team? *ACNR* *19*(2), 10-12. <https://doi.org/10.47795/STUI1319>
18. Mercier, L. J., Langelier, D. M., Buchanan, J., Robinson, S., & Plamondon, S. (2024). Development and integration of a music therapy program in the neurologic inpatient setting: a qualitative study. *Disabil Rehabil*, *47*(7), 1-10. <https://doi.org/10.1080/09638288.2024.2393439>
19. Murtaugh, B., Morrissey, A. M., Fager, S., Knight, H. E., Rushing, J., & Weaver, J. (2024). Music, occupational, physical, and speech therapy interventions for patients in disorders of consciousness: An umbrella review. *NeuroRehabilitation*, *54*(1), 109-127. <https://doi.org/10.3233/NRE-230149>
20. Thaut, M. H. (2014). Assessment and the transformational design model (TDM). In M. Thaut & V. Hoemberg (Eds.), *Handbook of neurologic music therapy* (pp. 60-68). Oxford University Press.
21. Breuer, E., Lee, L., De Silva, M., & Lund, C. (2015). Using theory of change to design and evaluate public health interventions: a systematic review. *Implement Sci*, *11*(63), 1-17. <https://doi.org/10.1186/s13012-016-0422-6>
22. de l'Etoile, S. K. (2014). Processes of music therapy: Clinical and scientific rationales and models. In S. Hallam, I. Cross, & M. H. Thaut (Eds.), *The Oxford handbook of music psychology* (2nd edition) (pp. 805-818). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780198722946.013.48>
23. Campbell, M., McKenzie, J. E., Sowden, A., Katikireddi, S. V., Brennan, S. E., Ellis, S., Hartmann-Boyce, J., Ryan, R., Shepperd, S., Thomas, J., Welch, V., & Thomson, H. (2020). Synthesis without meta-analysis (SWiM) in systematic reviews: reporting guideline. *BMJ*, *368*(6890), 1-6. <https://doi.org/10.1136/bmj.l6890>
24. Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Rodgers, M., Britten, N., Roen, K., & Duffy, S. (2006). *Guidance on the conduct of narrative synthesis in systematic reviews: A Product from the ESRC Methods Programme Version*. Lancaster University. <https://doi.org/10.13140/2.1.1018.4643>
25. Altenmüller, E., & Schlaug, G. (2013). Neurologic music therapy: The beneficial effects of music making on neurorehabilitation. *Acoust Sci Technol*, *34*(1), 5-12. <https://doi.org/10.1250/ast.34.5>
26. Lam, H.L., Li, W.T.V., Laher, I. and Wong, R.Y., 2020. Effects of music therapy on patients with dementia—A systematic review. *Geriatrics*, *5*(4), 62-75. <https://doi.org/10.3390/geriatrics5040062>
27. Lanb, L. C. L. S. H., Lanc, S. J., & Hsiehe, Y. P. (2024). Effectiveness of the Music Therapy in Dementia: A Systematic Review and Meta-Analysis of Randomized Controlled Trials. *Dement Geriatr Cogn Disord*, *54*(1), 1-20. <https://doi.org/10.1159/000542464>
28. Moreno-Morales, C., Calero, R., Moreno-Morales, P. and Pintado, C., 2020. Music therapy in the treatment of dementia: A systematic review and meta-analysis. *Front Med*, *7*(160), 1-11. <https://doi.org/10.3389/fmed.2020.00160>
29. National Institute for Health and Care Excellence (NICE) (2019). Dementia quality standard [NICE Guideline Quality Standard No. 184]. <https://www.nice.org.uk/guidance/qs184>
30. NHS England (2014). *NHS Standard Contract for Specialised Rehabilitation for Patients with Highly Complex Needs (All Ages)*. <https://www.england.nhs.uk/wp-content/uploads/2014/04/d02-rehab-pat-high-needs-0414.pdf>
31. Thompson, N., & Odell-Miller, H. (2024). An audit of music therapy in acute National Health Service (NHS) settings for people with dementia in the UK and adaptations made due to COVID-19. *Approaches Music Ther*, *16*(1), 1-16. <https://doi.org/10.56883/ajmt.2024.70>

32. Segura, E., Grau-Sánchez, J., Cerda-Company, X., Porto, M. F., De la Cruz-Puebla, M., Sanchez-Pinsach, D., Cerquides, J., Duarte, E., Palumbo, A., Turry, A., Raghavan, P., Särkämö, T., Münte, T. F., Arcos, J. L., & Rodríguez-Fornells, A. (2024). Enriched music-supported therapy for individuals with chronic stroke: a randomized controlled trial. *J Neurol* 271(10), 6606-6617. <https://doi.org/10.1007/s00415-024-12570-3>
33. Fusar-Poli, L., Bieleninik, L., Brondino, N., Chen, X. J., & Gold, C. (2018). The effect of music therapy on cognitive functions in patients with dementia: a systematic review and meta-analysis. *Aging Ment Health*, 22(9), 1103-1112. <https://doi.org/10.1080/13607863.2017.1348474>;
34. Freitas, C., Fernández-Company, J. F., Pita, M. F., & Garcia-Rodriguez, M. (2022). Music therapy for adolescents with psychiatric disorders: an overview. *Clin Child Psychol Psychiatry* 27(3), 895-910. <https://doi.org/10.1177/13591045221079161>
35. Tramontano, M., De Angelis, S., Mastrogiacomo, S., Princi, A. A., Ciancarelli, I., Frizziero, A., Iosa, M., Paolucci, S., & Morone, G. (2021). Music-based techniques and related devices in neurorehabilitation: a scoping review. *Expert Rev Med. Devices*, 18(8), 733-749. <https://doi.org/10.1080/17434440.2021.1947793>
36. Mishra, R., Florez-Perdomo, W. A., Shrivatava, A., Chouksey, P., Raj, S., Moscote-Salazar, L. R., Rahman, M. M., Sutar, R., & Agrawal, A. (2021). Role of music therapy in traumatic brain injury: a systematic review and meta-analysis. *World Neurosurg*, 146, 197-204. <https://doi.org/10.1016/j.wneu.2020.10.130>
37. Odell-Miller, H. (2016). The role, function and identity of music therapists in the 21st century, including new research and thinking from a UK perspective. *BJMT*, 30(1), 5-12. <https://doi.org/10.1177/1359457516639549>
38. Carr, C. E., Tsiris, G., & Swijghuisen Reigersberg, M. (2017). Understanding the present, re-visioning the future: An initial mapping of music therapists in the United Kingdom. *BJMT*, 31(2), 68-85. <https://doi.org/10.1177/1359457517728379>
39. Wood, J., Sandford, S., & Bailey, E. (2016). 'The whole is greater'. Developing music therapy services in the National Health Service: A case study revisited. *BJMT*, 30(1), 36-46. <https://doi.org/10.1177/1359457516637324>
40. National Institute for Health and Care Excellence (NICE) (2023). Stroke rehabilitation in adults [NICE Guideline No. 236]. <https://www.nice.org.uk/guidance/ng236/chapter/Recommendations>
41. Hariton, E., & Locascio, J. J. (2018). Randomised controlled trials—the gold standard for effectiveness research. *BJOG*, 125(13), 1716. <https://doi.org/10.1111/1471-0528.15199>.
42. **García-Navarro E.B., Buzón-Pérez A., Cabillas-Romero M. Effect of Music Therapy as a Non-Pharmacological Measure Applied to Alzheimer's Disease Patients: A Systematic Review.** *Nurs Rep*, 2022;12(4):775–790. <https://doi.org/10.3390/nursrep12040076>,
43. Bleibel, M., El Cheikh, A., Sadier, N.S. et al. The effect of music therapy on cognitive functions in patients with Alzheimer's disease: a systematic review of randomized controlled trials. *Alz Res Therapy* 15, 65 (2023). <https://doi.org/10.1186/s13195-023-01214-9>
44. Falzon, L., Davidson, K. W., & Bruns, D. (2010). Evidence searching for evidence-based psychology practice. *Prof Psychol Res Pr*, 41(6), 550-557. <https://doi.org/10.1037/a0021352>
45. Higgins, J. P. T., & Green, S. (Eds.) (2011). *Cochrane handbook for systematic reviews of interventions*. The Cochrane Collaboration. <https://doi.org/10.1002/9780470712184>
46. Chiu, E. C., Chen, Y. J., Wu, W. C., Chou, C. X., & Yu, M. Y. (2022). Psychometric comparisons of three depression measures for patients with stroke. *AJOT*, 76(4), 1-7.
47. Nishikawa-Pacher, A. (2022). Research questions with PICO: a universal mnemonic. *Publications*, 10(3), 21-30. <https://doi.org/10.3390/publications10030021>
48. Richardson, W. S., Wilson, M. C., Nishikawa, J., & Hayward, R. S. (1995). The well-built clinical question: a key to evidence-based decisions. *ACP J Club*, 123(3), 12-13. <https://doi.org/10.7326/ACPJC-1995-123-3-A12>

49. Impellizzeri, F., Maggio, M. G., De Pasquale, P., Bonanno, M., Bonanno, L., De Luca, R., Paladina, G., Alibrandi, A., Milardi, D., Thaut, M., Hurt, C., Quartarone, A., & Calabrò, R. S. (2024). Coupling neurologic music therapy with immersive virtual reality to improve executive functions in individuals with Parkinson's disease: A Quasi-Randomized Clinical Trial. *Clin Park Relat Disord*, *11*, 1-8. <https://doi.org/10.1016/j.prdoa.2024.100277>
50. Siponkoski, S. T., Martínez-Molina, N., Kuusela, L., Laitinen, S., Holma, M., Ahlfors, M., Jordan-Kilkki, P., Ala-Kauhaluoma, K., Melkas, S., Pekkola, J., Rodriguez-Fornells, A., Laine, M., Ylinen, A., Rantanen, P., Koskinen, S., Lipsanen, J., & Särkämö, T. (2020). Music therapy enhances executive functions and prefrontal structural neuroplasticity after traumatic brain injury: evidence from a randomized controlled trial. *J Neurotrauma*, *37*(4), 618-634.
51. Lee, S. J., Dvorak, A. L., & Manternach, J. N. (2024). Therapeutic Singing and Semi-Occluded Vocal Tract Exercises for Individuals with Parkinson's Disease: A Randomized Controlled Trial of a Single Session Intervention. *J Music Ther*, *61*(2), 132-167. <https://doi.org/10.1093/jmt/thae004>
52. Haire, C. M., Vuong, V., Tremblay, L., Patterson, K. K., Chen, J. L., & Thaut, M. H. (2021). Effects of therapeutic instrumental music performance and motor imagery on chronic post-stroke cognition and affect: A randomized controlled trial. *NeuroRehabilitation*, *48*(2), 195-208. <https://doi.org/10.3233/NRE-208014>
53. van Bruggen-Rufi, M. C., Vink, A. C., Wolterbeek, R., Achterberg, W. P., & Roos, R. A. (2017). The effect of music therapy in patients with Huntington's disease: a randomized controlled trial. *J Huntington's Dis*, *6*(1), 63-72. <https://doi.org/10.3233/JHD-160229>
54. Poćwierz-Marciniak, I., & Bidzan, M. (2017). The influence of music therapy on quality of life after a stroke. *Health Psychol Rep*, *5*(2), 173-185. <https://doi.org/10.5114/hpr.2017.63936>
55. Blackburn, R., & Bradshaw, T. (2014). Music therapy for service users with dementia: A critical review of the literature. *J Psychiatr Men. Health Nurs*, *21*(10), 879-888. <https://doi.org/10.1111/jpm.12165>
56. Hoffecker, L. (2020). Grey Literature Searching for Systematic Reviews in the Health Sciences. *Ser Lib* *79*(3-4), 252-260. <https://doi.org/10.1080/0361526X.2020.1847745>
57. Kelly, J., Sadeghieh, T., & Adeli, K. (2014). Peer review in scientific publications: benefits, critiques, & a survival guide. *J Int Fed Clin Chem Lab Med*, *25*(3), 227-243.
58. Barrington, A. (2015). Perspectives on the development of the music therapy profession in the UK. *Approaches*, *7*(1), 118-122. <https://doi.org/10.56883/ajmt.2015.385>
59. British Association for Music Therapy. (2020). *Guidelines on professional titles for music therapists*. British Association for Music Therapy. <https://www.bamt.org/music-therapy/what-is-a-music-therapist/guide-to-professional-practice>
60. Chandler, G., & Maclean, E. (2024). "There has probably never been a more important time to be a music therapist": Exploring how three music therapy practitioners working in adult mental health settings in the UK experienced the first year of the COVID-19 pandemic. *Approaches*. <https://doi.org/10.56883/ajmt.2024.63>
61. Helbach, J., Pieper, D., Mathes, T., Rombey, T., Zeeb, H., Allers, K., & Hoffmann, F. (2022). Restrictions and their reporting in systematic reviews of effectiveness: an observational study. *BMC Med Res Methodol*, *22*(230), 1-10. <https://doi.org/10.1186/s12874-022-01710-w>
62. Kamioka, H., Tsutani, K., Yamada, M., Park, H., Okuizumi, H., Tsuruoka, K., Honda, T., Okada, S., Park, S., Kitayuguchi, J., Abe, T., Handa, S., Oshio, T., & Mutoh, Y. (2014). Effectiveness of music therapy: a summary of systematic reviews based on randomized controlled trials of music interventions. *Patient Prefer Adherence*, *8*, 727-754. <https://doi.org/10.2147/PPA.S61340>
63. Pieper, D., & Puljak, L. (2021). Language restrictions in systematic reviews should not be imposed in the search strategy but in the eligibility criteria if necessary. *J Clin Epidemiol* *132*, 146-147. <https://doi.org/10.1016/j.jclinepi.2020.12.027>



64. Bond, C., Lancaster, G. A., Campbell, M., Chan, C., Eddy, S., Hopewell, S., Mellor, K., Thabane, L., & Eldridge, S. (2023). Pilot and feasibility studies: extending the conceptual framework. *PFS*, 9(1), 24-33.
65. Arain, M., Campbell, M. J., Cooper, C. L., & Lancaster, G. A. (2010). What is a pilot or feasibility study? A review of current practice and editorial policy. *BMC Med Res Methodol*, 10(67), 1-7. <https://doi.org/10.1186/1471-2288-10-67>
66. Mayer-Benarous, H., Benarous, X., Vontron, F., & Cohen, D. (2021). Music therapy for children with autistic spectrum disorder and/or other neurodevelopmental disorders: a systematic review. *Front Psychiatry*, 12(643234), 1-21. <https://doi.org/10.3389/fpsy.2021.643234>
67. Ahuja, C. S., Wilson, J. R., Nori, S., Kotter, M., Druschel, C., Curt, A., & Fehlings, M. G. (2017). Traumatic spinal cord injury. *Nat Rev Dis Primers*, 3(1), 1-21. <https://doi.org/10.1038/nrdp.2017.18>
68. Guyatt, G. H., Oxman, A. D., Vist, G. E., Kunz, R., Falck-Ytter, Y., Alonso-Coello, P., & Schünemann, H. J. (2008). GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*, 336(7650), 924-926. <https://doi.org/10.1136/bmj.39489.470347.AD>
69. Matthew J. Page, Joanne E. McKenzie, Patrick M. Bossuyt, Isabelle Boutron, Tammy C. Hoffmann, Cynthia D. Mulrow, Larissa Shamseer, Jennifer M. Tetzlaff, Elie A. Akl, Sue E. Brennan, Roger Chou, Julie Glanville, Jeremy M. Grimshaw, Asbjørn Hróbjartsson, Manoj M. Lalu, Tianjing Li, Elizabeth W. Loder, Evan Mayo-Wilson, Steve McDonald, Luke A. McGuinness, Lesley A. Stewart, James Thomas, Andrea C. Tricco, Vivian A. Welch, Penny Whiting, David Moher. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021;372:n71. <https://doi.org/10.1136/bmj.n71>
70. Higgins, J. P., Savović, J., Page, M. J., Elbers, R. G., & Sterne, J. A. (2019). Assessing risk of bias in a randomized trial. In J. P. T. Higgins, J. Thomas, J. Chandler, M. Cumpston, T. Li, M. J. Page, V. A. Welch (Eds.), *Cochrane Handbook for Systematic Reviews of Interventions* (pp. 205-228). The Cochrane Collaboration. <https://doi.org/10.1002/9781119536604.ch8>
71. Nejadghaderi, S. A., Balibegloo, M., & Rezaei, N. (2024). The Cochrane risk of bias assessment tool 2 (RoB 2) versus the original RoB: A perspective on the pros and cons. *Health Sci Rep*, 7(6), 1-6. <https://doi.org/10.1002/hsr.2.2165>;
72. Sterne J. A. C., Savović J., Page M. J., Elbers R. G., Blencowe N. S., Boutron I., Cates C. J., Cheng H-Y., Corbett M. S., Eldridge S. M., Herrán M. A., Hopewell S., Hróbjartsson A., Junqueira D. R., Jüni P., Kirkham J. J., Lasserson T., Li T., McAleenan A., ... Higgins J. P. T. (2019). RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ*, 366(14898), 1-8. <https://doi.org/10.1136/bmj.14898>
73. Thomson, H. J., & Thomas, S. (2013). The effect direction plot: visual display of non-standardised effects across multiple outcome domains. *Res Synth Methods*, 4(1), 95-101. <https://doi.org/10.1002/jrsm.1060>
74. Cohen, J. (1992). Quantitative methods in psychology: A power primer. *Psychol Bull*, 112(1), 1155-1159. <https://doi.org/10.1037//0033-2909.112.1.155>
75. Bauchner, H., Golub, R. M., & Fontanarosa, P. B. (2019). Reporting and interpretation of randomized clinical trials. *JAMA*, 322(8), 732-735. <https://doi.org/10.1001/jama.2019.12056>
76. Bhide, A., Shah, P. S., & Acharya, G. (2018). A simplified guide to randomized controlled trials. *AOGS*, 97(4), 380-387. <https://doi.org/10.1111/aogs.13309>
77. Murad, M. H., Mustafa, R. A., Schünemann, H. J., Sultan, S., & Santesso, N. (2017). Rating the certainty in evidence in the absence of a single estimate of effect. *BMJ EBM*, 22(3), 1-3. <https://doi.org/10.1136/ebmed-2017-110668>
78. Schünemann, H. J., Higgins, J. P., Vist, G. E., Glasziou, P., Akl, E. A., Skoetz, N., Guyatt, G. H. (2019). Completing 'summary of findings' tables and grading the certainty of the evidence. In J. P. T. Higgins, J. Thomas, J. Chandler,

- M. Cumpston, T. Li, M. J. Page, V. A. Welch (Eds.), *Cochrane Handbook for Systematic Reviews of Interventions* (pp. 375-402). The Cochrane Collaboration. <https://doi.org/10.1002/9781119536604.ch14>
79. Impellizzeri, F., Maggio, M. G., De Pasquale, P., Bonanno, M., Bonanno, L., De Luca, R., Paladina, G., Alibrandi, A., Milardi, D., Thaut, M., Hurt, C., Quartarone, A., & Calabrò, R. S. (2024). Coupling neurologic music therapy with immersive virtual reality to improve executive functions in individuals with Parkinson's disease: A Quasi-Randomized Clinical Trial. *Clin Park Relat Disord*, *11*, 1-8. <https://doi.org/10.1016/j.prdoa.2024.100277>
  80. Estellat, C., Torgerson, D. J., & Ravaud, P. (2009). How to perform a critical analysis of a randomised controlled trial. *Best Pract Res Clin Rheumatol*, *23*(2), 291-303. <https://doi.org/10.1016/j.berh.2009.03.003>
  81. Chou, C. H., Chen, P. C., Huang, Y. C., Yang, T. H., Wang, L. Y., Chen, I. H., Lee, H. J., & Lee, Y. Y. (2024). Neurological music therapy for poststroke depression, activity of daily living and cognitive function: A pilot randomized controlled study. *Nord. J Music Ther*, *33*(3), 226-237. <https://doi.org/10.1080/08098131.2023.2280968>
  82. Van Ginkel, J. R., Linting, M., Rippe, R. C., & Van Der Voort, A. (2020). Rebutting existing misconceptions about multiple imputation as a method for handling missing data. *J Pers Assess*, *102*(3), 297-308. <https://doi.org/10.1080/00223891.2018.1530680>
  83. Van Buuren, S., & Groothuis-Oudshoorn, K. (2011). MICE: Multivariate imputation by chained equations in R. *J Stat Softw.*, *45*(3), 1-67. <https://doi.org/10.18637/jss.v045.i03>
  84. Austin, P. C., White, I. R., Lee, D. S., & van Buuren, S. (2021). Missing data in clinical research: a tutorial on multiple imputation. *Can J Cardiol*, *37*(9), 1322-1331. <https://doi.org/10.1016/j.cjca.2020.11.010>
  85. Andrade, E., Arce, C., Torrado, J., Garrido, J., De Francisco, C., & Arce, I. (2010). Factor structure and invariance of the POMS mood state questionnaire in Spanish. *Span J Psychol*, *13*(1), 444-452. <https://doi.org/10.1017/S1138741600003991>
  86. Boutron, I., Dutton, S., Ravaud, P., & Altman, D. G. (2010). Reporting and interpretation of randomized controlled trials with statistically nonsignificant results for primary outcomes. *JAMA*, *303*(20), 2058-2064. <https://doi.org/10.1001/jama.2010.651>
  87. Carter, E. C., Schönbrodt, F. D., Gervais, W. M., & Hilgard, J. (2019). Correcting for bias in psychology: A comparison of meta-analytic methods. *Adv Meth Pract Psychol Sci*, *2*(2), 115-144. <https://doi.org/10.1177/2515245919847196>
  88. van Bruggen-Rufi, M., Vink, A., Achterberg, W., & Roos, R. (2016). Music therapy in Huntington's disease: a protocol for a multi-center randomized controlled trial. *BMC Psychol*, *4*(38), 1-9. <https://doi.org/10.1186/s40359-016-0146-z>
  89. Kraemer, H. C., Mintz, J., Noda, A., Tinklenberg, J., & Yesavage, J. A. (2006). Caution regarding the use of pilot studies to guide power calculations for study proposals. *Arch Gen Psychiatry*, *63*(5), 484-489. <https://doi.org/10.1001/archpsyc.63.5.484>
  90. Beato, M. (2022). Recommendations for the design of randomized controlled trials in strength and conditioning. Common design and data interpretation. *Front Sports Act Living*, *4*(981836), 1-7. <https://doi.org/10.3389/fspor.2022.981836>
  91. Annunziata, M. A., Muzzatti, B., Bidoli, E., Flaiban, C., Bomben, F., Piccinin, M., Gipponi, K. M., Mariutti, G., Busato, S., & Mella, S. (2020). Hospital Anxiety and Depression Scale (HADS) accuracy in cancer patients. *Support Care Cancer*, *28*(8), 3921-3926. <https://doi.org/10.1007/s00520-019-05244-8>
  92. Chiu, E. C., Chen, Y. J., Wu, W. C., Chou, C. X., & Yu, M. Y. (2022). Psychometric comparisons of three depression measures for patients with stroke. *AJOT* *76*(4), 1-7.
  93. Costantini, M., Musso, M., Viterbori, P., Bonci, F., Del Mastro, L., Garrone, O., M Venturini M., & Morasso, G. (1999) Detecting psychological distress in cancer patients: validity of the Italian version of the Hospital Anxiety and Depression Scale. *Support Care Cancer* *7*, 121-127 (1999). <https://doi.org/10.1007/s005200050241>;

94. Annell, S., Sjöberg, A., & Sverke, M. (2014). Use and interpretation of test scores from limited cognitive test batteries: how g+ Gc can equal g. *Scand J Psychol*, *55*(5), 399-408. <https://doi.org/10.1111/sjop.12140>
95. Essers, B., Veerbeek, J. M., Luft, A. R., & Verheyden, G. (2024). The feasibility of the adapted H-GRASP program for perceived and actual daily-life upper limb activity in the chronic phase post-stroke. *Disabil Rehabil*, *46*(24), 5815-5828. <https://doi.org/10.1080/09638288.2024.2313121>
96. Simpson, L. A., Eng, J. J., & Chan, M. (2017). H-GRASP: the feasibility of an upper limb home exercise program monitored by phone for individuals post stroke. *Disabil Rehabil*, *39*(9), 874-882. <https://doi.org/10.3109/09638288.2016.1162853>
97. Russell, E. W., Russell, S. L., & Hill, B. D. (2005). The fundamental psychometric status of neuropsychological batteries. *Arch Clin Neuropsychol*, *20*(6), 785-794. <https://doi.org/10.1016/j.acn.2005.05.001>
98. Casaletto, K. B., & Heaton, R. K. (2017). Neuropsychological assessment: Past and future. *JNS*, *23*(9-10), 778-790. <https://doi.org/10.1017/S1355617717001060>
99. Thaut, M., & Hoemberg, V. (Eds.). (2014). *Handbook of neurologic music therapy*. Oxford University Press.
100. von Hippel, P. T. (2015). The heterogeneity statistic I<sup>2</sup> can be biased in small meta-analyses. *BMC Med Research Methodol*, *15*(35), 1-8. <https://doi.org/10.1186/s12874-015-0024-z>
101. Königs, M., Beurskens, E. A., Snoep, L., Scherder, E. J., & Oosterlaan, J. (2018). Effects of timing and intensity of neurorehabilitation on functional outcome after traumatic brain injury: a systematic review and meta-analysis. *Arch Phys Med Rehabil*, *99*(6), 1149-1159. <https://doi.org/10.1016/j.apmr.2018.01.013>
102. Kern, P., & Tague, D. B. (2017). Music therapy practice status and trends worldwide: An international survey study. *J Music Ther*, *54*(3), 255-286. <https://doi.org/10.1093/jmt/thx011>
103. Dobran, S. A., & Gherman, A. (2024). Neurorehabilitation across continents: the WFNR -EFNR regional meeting in conjunction with the 19th congress of the society for the study of neuroprotection and neuroplasticity and the 19th international summer school of neurology in Baku, Azerbaijan. *J Med Life*, *17*(9), 825-829. <https://doi.org/10.25122/jml-2024-1014>
104. Nasios, G., Messinis, L., Dardiotis, E., & Sgantzios, M. (2023). Neurorehabilitation: Looking Back and Moving Forward. *Healthcare*, *11*(10), 1-6. <https://doi.org/10.3390/healthcare11101452>
105. Fernainy, P., Cohen, A. A., Murray, E., Losina, E., Lamontagne, F., & Sourial, N. (2024). Rethinking the pros and cons of randomized controlled trials and observational studies in the era of big data and advanced methods: a panel discussion. *BMC Proc*, *18*(2), 1-7. <https://doi.org/10.1186/s12919-023-00285-8>

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.