![Cochrane Library logo]

Cochrane Database of Systematic Reviews

Music-based therapeutic interventions for people with dementia (Review)

van der Steen JT, van der Wouden JC, Methley AM, Smaling HJA, Vink AC, Bruinsma MS

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

[Intervention Review]

Music-based therapeutic interventions for people with dementia

Jenny T van der Steen[1,2,3], Johannes C van der Wouden[4], Abigail M Methley[5], Hanneke J A Smaling[1,6], Annemieke C Vink[7], Manon S Bruinsma[8,9]

[1]Department of Public Health and Primary Care, Leiden University Medical Center, Leiden, Netherlands. [2]Department of Primary and Community Care and Radboudumc Alzheimer Center, Radboud University Medical Center, Nijmegen, Netherlands. [3]Cicely Saunders Institute, King's College London, London, UK. [4]Department of General Practice and Department of Medicine for Older People, Amsterdam Public Health Research Institute, Amsterdam UMC, Vrije Universiteit, Amsterdam, Netherlands. [5]Innovative Clinical Psychology Solutions Ltd, London, UK. [6]University Network for the Care Sector Zuid-Holland, Leiden University Medical Center, Leiden, Netherlands. [7]Music Therapy Department, ArtEZ University of the Arts, Enschede, Netherlands. [8]Muzis, Praktijk voor Muziektherapie, Amersfoort, Netherlands. [9]Department of Creative Arts Therapies, HU University of Applied Sciences Utrecht, Utrecht, Netherlands

Contact: Jenny T van der Steen, jtvandersteen@lumc.nl.

A B S T R A C T

Background

Dementia is a clinical syndrome with a number of different causes. It is characterised by deterioration in cognitive, behavioural, social and emotional functioning. Pharmacological interventions are available but have limited effect on many of the syndrome's features. However, receptivity to music may remain until the late phases of dementia, and music-based therapeutic interventions (which include, but are not limited to, music therapy) are suitable for people with advanced dementia. As there is uncertainty about the effectiveness of music-based therapeutic interventions, trials are being conducted to evaluate this. This review updates one last published in 2018 and examines the current evidence for the effects of music-based interventions for people with dementia.

Objectives

To assess the effects of music-based therapeutic interventions for people with dementia on emotional well-being (including quality of life), mood disturbance or negative affect (i.e. depressive symptoms and anxiety), behavioural problems (i.e. overall behavioural problems or neuropsychiatric symptoms, and more specifically agitation or aggression), social behaviour and cognition, at the end of therapy and four or more weeks after the end of treatment, and to assess any adverse effects.

Search methods

We searched the Cochrane Dementia and Cognitive Improvement Group's Specialised Register, MEDLINE (Ovid SP), Embase (Ovid SP), PsycINFO (Ovid SP), CINAHL (EBSCOhost), Web of Science Core Collection (ISI Web of Science), LILACS (BIREME), ClinicalTrials.gov and the World Health Organisation's meta-register-the International Clinical Trials Registry Platform on 30 November 2023.

Selection criteria

We included randomised controlled trials of music-based therapeutic interventions (of at least five sessions) for people with dementia that measured any of our outcomes of interest. Control groups either received usual care or other activities with or without music.

Data collection and analysis

Two review authors worked independently to screen the retrieved studies against the inclusion criteria and then to extract data from included studies and assess their risk of bias. If necessary, we contacted trial authors to ask for additional data, such as relevant subscales. We pooled data using the random-effects model. We assessed the certainty of the evidence for our two comparisons and our main outcomes of interest using GRADE.

Main results

We included 30 studies with 1720 randomised participants that were conducted in 15 countries. Twenty-eight studies with 1366 participants contributed data to meta-analyses. Ten studies contributed data to long-term outcomes. Participants had dementia of varying degrees of severity and resided in institutions in most of the studies. Seven studies delivered an individual intervention; the other studies delivered the intervention to groups. Most interventions involved both active and receptive elements of musical experience. The studies were at high risk of performance bias and some were at high risk of detection or other bias.

For music-based therapeutic interventions compared to usual care, we found moderate-certainty evidence that, at the end of treatment, music-based therapeutic interventions probably improved depressive symptoms slightly (standardised mean difference (SMD) −0.23, 95% confidence interval (CI) −0.42 to −0.04; 9 studies, 441 participants), and we found low-certainty evidence that it may have improved overall behavioural problems (SMD −0.31, 95% CI −0.60 to −0.02; 10 studies, 385 participants). We found moderate-certainty evidence that music-based therapeutic interventions likely did not improve agitation or aggression (SMD −0.05, 95% CI −0.27 to 0.17; 11 studies, 503 participants). Low to very low certainty evidence showed that they did not improve emotional well-being (SMD 0.14, 95% CI -0.29 to 0.56; 4 studies, 154 participants), anxiety (SMD −0.15, 95% CI −0.39 to 0.09; 7 studies, 282 participants), social behaviour (SMD 0.22, 95% CI −0.14 to 0.57; 2 studies; 121 participants) or cognition (SMD 0.19, 95% CI −0.02 to 0.41; 7 studies, 353 participants). Low or very-low -certainty evidence showed that music-based therapeutic interventions may not have been more effective than usual care in the long term (four weeks or more after the end of treatment) for any of the outcomes.

For music-based therapeutic interventions compared to other interventions, we found low-certainty evidence that, at the end of treatment, music-based therapeutic interventions may have been more effective than the other activities for social behaviour (SMD 0.52, 95% CI 0.08 to 0.96; 4 studies, 84 participants). We found very low-certainty evidence of a positive effect on anxiety (SMD −0.75, 95% CI −1.27 to −0.24; 10 studies, 291 participants). For all other outcomes, low-certainty evidence showed no evidence of an effect: emotional well-being (SMD 0.20, 95% CI −0.09 to 0.49; 9 studies, 298 participants); depressive symptoms (SMD −0.14, 95% CI -0.36 to 0.08; 10 studies, 359 participants); agitation or aggression (SMD 0.01, 95% CI −0.31 to 0.32; 6 studies, 168 participants); overall behavioural problems (SMD −0.08, 95% CI −0.33 to 0.17; 8 studies, 292 participants) and cognition (SMD 0.12, 95% CI −0.21 to 0.45; 5 studies; 147 participants). We found low or very-low certainty evidence that music-based therapeutic interventions may not have been more effective than other interventions in the long term (four weeks or more after the end of treatment) for any of the outcomes.

Adverse effects were inconsistently measured or recorded, but no serious adverse events were reported.

Authors' conclusions

When compared to usual care, providing people with dementia with at least five sessions of a music-based therapeutic intervention probably improves depressive symptoms and may improve overall behavioural problems at the end of treatment.

When compared to other activities, music-based therapeutic interventions may improve social behaviour at the end of treatment. No conclusions can be reached about the outcome of anxiety as the certainty of the evidence is very low.

There may be no effects on other outcomes at the end of treatment. There was no evidence of long-term effects from music-based therapeutic interventions.

Adverse effects may be rare, but the studies were inconsistent in their reporting of adverse effects.

Future studies should examine the duration of effects in relation to the overall duration of treatment and the number of sessions.

PLAIN LANGUAGE SUMMARY

Does music-based therapy help people with dementia?

Key messages

· Compared to usual care (i.e. without specific activities), providing people with dementia with music-based therapy sessions probably improves depression, and may improve overall behavioural problems.

· Compared to other activities, music-based therapy may improve social behaviour, but we are unsure if it improves anxiety.

· These effects may not last beyond the end of treatment.

· There is no evidence that music-based therapy makes a difference to the emotional well-being, agitation or cognition (e.g. thinking and remembering) of people with dementia, although there is uncertainty about this.

· There is uncertainty in the evidence for long-term effects, but no long-term effects were observed in the studies.

· Adverse effects (unwanted side effects) may be rare, but the studies were inconsistent in their reporting, so we need more evidence before we can reach reliable conclusions.

Why offer music-based therapy to people with dementia?

People with dementia gradually develop increasing difficulty with thinking and daily activities. Dementia is often associated with emotional and behavioural problems and may decrease a person's quality of life. In the later stages of dementia, it may be difficult for people to communicate with words, but even when they can no longer speak, they may still be able to hum or play along with music. Therapy involving music may therefore be especially suitable for people with dementia to improve their lives.

Who provides music-based therapy?

Music therapists are certified to work with individuals or groups of people, using music to try to help meet their physical, psychological and social needs. Other professionals may also be trained to provide similar therapies.

What did we want to find out?

We wanted to find out if, for people with dementia, music-based therapy works better than usual care or other activities, such as painting. We were interested in whether the therapy changed these outcomes at the end of treatment:

· emotional well-being including quality of life;

· depression and anxiety;

· agitation or aggression and overall behavioural problems;

· social behaviour; and

· cognition (e.g. thinking and remembering).

We also wanted to find out if there were any long-term effects after therapy ended or any unwanted effects.

What did we do?

We searched for studies in which people with dementia were randomly allocated to receive music-based therapy (of at least five sessions) or to a comparison group with no activities or different activities. We combined the results of the studies to estimate the effects of music-based therapy as accurately as possible. We also rated our level of confidence in the findings.

What did we find?

We found 30 studies performed in 15 countries. The studies involved 1720 people with dementia of varying severity. In most of the studies, the participants lived in nursing homes. Seven studies delivered the music-based therapy to individuals; the other studies delivered the intervention to groups. We were able to use results from 28 studies involving 1366 people with dementia for one or more outcomes at the end of treatment. Ten studies contributed information about long-term results.

Main results

At the end of treatment

Music-based therapy probably improves depression and may improve overall behavioural problems, compared with providing usual care.

Music-based therapy may improve social behaviour compared to other activities. We were less confident about the effects of music-based therapy on anxiety compared to other activities.

The available evidence does not suggest any benefit of music-based therapy for emotional well-being (including quality of life), agitation and aggression, or cognition, but the evidence is limited and there is uncertainty about this.

In the longer term

Some studies measured outcomes four weeks or more after treatment ended. We did not find any lasting effects, but there were fewer results measured in the longer term, and we are uncertain about this evidence. Further studies are likely to have a significant impact on what we know about the effects of music-based therapy for people with dementia, so it is important that research continues.

What are the limitations of the evidence?

There was variation in the quality of the studies and how well they were reported. In all studies, participants and personnel might have known which treatment participants were getting, and in some, the assessors might have known this as well. This could have affected the results. Regarding effects at the end of music-based therapy, we are moderately confident in the evidence for the beneficial effect on depression compared to usual care. We have little confidence in the effects or lack of effects on any of the other outcomes. Adverse effects were rarely reported.

How up to date is this evidence?

This review updates our previous review. We added eight new studies and reached new conclusions. The evidence is current to 30 November 2023.

SUMMARY OF FINDINGS

Summary of findings 1. Music-based therapeutic interventions compared to usual care for people with dementia: end-of-treatment effects

Patient or population: people with dementia

Setting: long-term care facilities, specific hospital departments, or both; or unclear
Intervention: music-based therapeutic interventions
Comparison: usual care

Outcomes (end of treatment) measured with a variety of scales	Anticipated absolute effects, SMD* (95% CI)	Number of participants (studies)	Certainty of the evidence (GRADE)
	Score with music-based therapeutic interventions compared to usual care		
Emotional well-being including quality of life (at 6 weeks to 6 months)	The score in the intervention group was 0.14 SD higher (0.29 lower to 0.56 higher). (A positive result favours the intervention group.)	154[d] (4 RCTs)	⊕⊖⊖⊖ **Very low**[a]
Mood disturbance or negative affect: depression (at 6 weeks to 6 months)	The score in the intervention group was 0.23 SD lower (0.42 lower to 0.04 lower). (A negative result favours the intervention group.)	441[d] (9 RCTs)	⊕⊕⊕⊖ **Moderate**[b]
Mood disturbance or negative affect: anxiety (at 6 weeks to 6 months)	The score in the intervention group was 0.15 SD lower (0.39 lower to 0.09 higher). (A negative result favours the intervention group.)	282 (7 RCTs)	⊕⊕⊖⊖ **Low**[c]
Behavioural problems: agitation or aggression (at 2 weeks to 6 months)	The score in the intervention group was 0.05 SD lower (0.27 lower to 0.17 higher). (A negative result favours the intervention group.)	503 (11 RCTs)	⊕⊕⊕⊖ **Moderate**[b]
Behavioural problems: overall (at 6 weeks to 6 months)	The score in the intervention group was 0.31 SD lower (0.60 lower to 0.02 lower). (A negative result favours the intervention group.)	385[d] (10 RCTs)	⊕⊕⊖⊖ **Low**[c]
Social behaviour (at 12 to 24 weeks)	The score in the intervention group was 0.22 SD higher (0.14 lower to 0.57 higher). (A positive result favours the intervention group.)	121 (2 RCTs)	⊕⊕⊖⊖ **Low**[c]
Cognition (at 6 to 15 weeks)	The score in the intervention group was 0.19 SD higher (0.02 lower to 0.41 higher). (A positive result favours the intervention group.)	353[d] (7 RCTs)	⊕⊕⊖⊖ **Low**[c]

*Interpretation of SMD: a difference of < 0.20 SD can be regarded as little or no effect; 0.20 to 0.40 SD as a small effect; 0.40 to 0.70 as a moderate effect; and > 0.70 as a large effect (Higgins 2019, Chapter 15.5.3.1).

CI: confidence interval; **SMD:** standardised mean difference; **SD:** standard deviations; **RCT:** randomised controlled trial

GRADE Working Group grades of evidence (GradePro GDT)
High certainty: we are very confident that the true effect lies close to that of the estimate of the effect.
Moderate certainty: we are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low certainty: our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
Very low certainty: we have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

[a]Downgraded by three levels: one level for risk of bias: no blinding of therapists and participants (not possible), and often no or unclear blinding of outcome assessment; one level for imprecision: small number of participants (< 400); and one level for reporting bias

[b]Downgraded by one level for risk of bias: no blinding of therapists and participants (not possible), and sometimes no or unclear blinding of outcome assessment

[c]Downgraded by two levels: one level for risk of bias: no blinding of therapists and participants (not possible) and sometimes no or unclear blinding of outcome assessment; and one level for imprecision: small number of participants (< 400)

[d]Adjustments were made to account for clustering, so the number of participants shown does not correspond with the combined number of participants in the trials.

Summary of findings 2.   Music-based therapeutic interventions compared to usual care for people with dementia: long-term effects

Patient or population: people with dementia

Setting: long-term care facilities, specific hospital departments, or both
Intervention: music-based therapeutic interventions
Comparison: usual care

Outcomes (long-term effects) measured with a variety of scales	Anticipated absolute effects, SMD* (95% CI)	Number of partici-pants (studies)	Certainty of the evidence (GRADE)
	Score with music-based therapeutic interven-tions compared to usual care		
Emotional well-being in-cluding quality of life (2 to 6 months after end of treat-ment)	The score in the intervention group was 0.17 SD higher (0.80 lower to 1.14 higher). (A positive result favours the intervention group.)	86[c] (3 RCTs)	⊕⊝⊝⊝ **Very low**[a]
Mood disturbance or negative affect: depression (4 weeks to 6 months after end of treat-ment)	The score in the intervention group was 0.07 SD lower (0.31 lower to 0.18 higher). (A negative result favours the intervention group.)	276 (5 RCTs)	⊕⊕⊝⊝ **Low**[b]
Mood disturbance or nega-tive affect: anxiety (4 weeks to 2 months after end of treat-ment)	The score in the intervention group was 0.06 SD lower (0.48 lower to 0.37 higher). (A negative result favours the intervention group.)	141 (3 RCTs)	⊕⊕⊝⊝ **Low**[b]
Behavioural problems: agita-tion or aggression (4 weeks to 2 months after end of treat-ment)	The score in the intervention group was 0.17 SD lower (0.42 lower to 0.09 higher). (A negative result favours the intervention group.)	241 (4 RCTs)	⊕⊕⊝⊝ **Low**[b]
Behavioural problems: over-all (4 weeks to 6 months after end of treatment)	The score in the intervention group was 0.19 SD lower (0.52 lower to 0.14 higher). (A negative result favours the intervention group.)	245[c] (6 RCTs)	⊕⊕⊝⊝ **Low**[b]
Social behaviour	No data	0 (0 RCTs)	Not applicable
Cognition (1 to 3 months after end of treatment)	The score in the intervention group was 0.09 SD higher (0.24 lower to 0.41 higher). (A positive result favours the intervention group.)	146 (2 RCTs)	⊕⊕⊝⊝ **Low**[b]

*Interpretation of SMD: a difference of < 0.20 SD can be regarded as little or no effect; 0.20 to 0.40 SD as a small effect; 0.40 to 0.70 as a moderate effect; and > 0.70 as a large effect (Higgins 2019, Chapter 15.5.3.1).

CI: confidence interval; **SMD:** standardised mean difference; **SD:** standard deviations; **RCT:** randomised controlled trial

GRADE Working Group grades of evidence (GradePro GDT)
High certainty: we are very confident that the true effect lies close to that of the estimate of the effect.
Moderate certainty: we are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low certainty: our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
Very low certainty: we have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

[a]Downgraded by three levels: one level for risk of bias: no blinding of therapists and participants (not possible), and sometimes no or unclear blinding of outcome assessment; one level for imprecision: small number of participants (< 400) and broad CI includes both benefit and harm (1 level); and one level for inconsistency: effects in opposite directions
[b]Downgraded by two levels; one level for risk of bias: no blinding of therapists and participants (not possible), and sometimes no or unclear blinding of outcome assessment; and one level for imprecision: small number of participants (< 400)
[c]Adjustments were made to account for clustering, so the number of participants shown does not correspond with the combined number of participants in the trials.

Summary of findings 3. Music-based therapeutic interventions compared to other activities for people with dementia: end-of-treatment effects

Patient or population: people with dementia

Setting: long-term care facilities, specific hospital departments, or both
Intervention: music-based therapeutic interventions
Comparison: other activities

Outcomes (end of treatment) measured with a variety of scales	Anticipated absolute effects, SMD* (95% CI)	Number of participants (studies)	Certainty of the evidence (GRADE)
	Score with music-based therapeutic interventions compared to other activities		
Emotional well-being including quality of life (at 2 weeks to 6 months)	The score in the intervention group was 0.20 SD higher (0.09 lower to 0.49 higher). (A positive result favours the intervention group.)	298[c] (9 RCTs)	⊕⊕⊖⊖ **Low**[a]
Mood disturbance or negative affect: depression (at 4 weeks to 6 months)	The score in the intervention group was 0.14 SD lower (0.36 lower to 0.08 higher). (A negative result favours the intervention group.)	359[c] (10 RCTs)	⊕⊕⊖⊖ **Low**[a]
Mood disturbance or negative affect: anxiety (at 3 weeks to 4 months)	The score in the intervention group was 0.75 SD lower (1.27 lower to 0.24 lower). (A negative result favours the intervention group.)	291 (10 RCTs)	⊕⊖⊖⊖ **Very low**[b]
Behavioural problems: agitation or aggression (at 4 weeks to 4 months)	The score in the intervention group was 0.01 SD higher (0.31 lower to 0.32 higher). (A negative result favours the intervention group.)	168 (6 RCTs)	⊕⊕⊖⊖ **Low**[a]
Behavioural problems: overall (at 4 weeks to 6 months)	The score in the intervention group was 0.08 SD lower (0.33 lower to 0.17 higher). (A negative result favours the intervention group.)	292[c] (8 RCTs)	⊕⊕⊖⊖ **Low**[a]
Social behaviour (at 2 to 4 weeks)	The score in the intervention group was 0.52 SD higher (0.08 higher to 0.96 higher). (A positive result favours the intervention group.)	84[c] (4 RCTs)	⊕⊕⊖⊖ **Low**[a]

Cognition (at 4 to 16 weeks)	The score in the intervention group was 0.12 SD higher (0.21 lower to 0.45 higher). (A positive result favours the intervention group.)	147[c] (5 RCTs)	⊕⊕⊖⊖ **Low**[a]

*Interpretation of SMD: a difference of < 0.20 SD can be regarded as little or no effect; 0.20 to 0.40 SD as a small effect; 0.40 to 0.70 as a moderate effect; and > 0.70 as a large effect (Higgins 2019, Chapter 15.5.3.1).

CI: confidence interval; **SMD:** standardised mean difference; **SD:** standard deviations; **RCT:** randomised controlled trial

GRADE Working Group grades of evidence (GradePro GDT)
High certainty: we are very confident that the true effect lies close to that of the estimate of the effect.
Moderate certainty: we are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low certainty: our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
Very low certainty: we have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

[a]Downgraded by two levels: one level for risk of bias: no blinding of therapists and participants (not possible), and sometimes no or unclear blinding of outcome assessment; and one level for imprecision: small number of participants (< 400) and broad CI
[b]Downgraded by three levels: one level for risk of bias: no blinding of therapists and participants (not possible), and sometimes no or unclear blinding of outcome assessment; one level for imprecision: small number of participants (< 400) and broad CI; and one level for inconsistency: multiple non-overlapping CIs
[c]Adjustments were made to account for clustering, so the number of participants shown does not correspond with the combined number of participants in the trials.

Summary of findings 4. Music-based therapeutic interventions compared to other activities for people with dementia: long-term effects

Patient or population: people with dementia

Setting: long-term care facilities, specific hospital departments, or both
Intervention: music-based therapeutic interventions
Comparison: other activities

Outcomes (long-term effects) measured with a variety of scales	Anticipated absolute effects, SMD* (95% CI)	Number of participants (studies)	Certainty of the evidence (GRADE)
	Score with music-based therapeutic interventions compared to other activities		
Emotional well-being including quality of life (4 weeks to 6 months after end of treatment)	The score in the intervention group was 0.10 SD higher (0.29 lower to 0.49 higher). (A positive result favours the intervention group.)	130 (4 RCTs)	⊕⊕⊖⊖ **Low**[a]
Mood disturbance or negative affect: depression (4 weeks to 6 months after end of treatment)	The score in the intervention group was 0.07 SD lower (0.39 lower to 0.25 higher). (A negative result favours the intervention group.)	174 (4 RCTs)	⊕⊕⊖⊖ **Low**[b]
Mood disturbance or negative affect: anxiety (4 weeks to 2 months after end of treatment)	The score in the intervention group was 0.53 SD lower (1.31 lower to 0.25 higher). (A negative result favours the intervention group.)	124 (4 RCTs)	⊕⊕⊖⊖ **Low**[b]
Behavioural problems: agitation or aggression (4 weeks to	The score in the intervention group was 0.10 SD higher	89 (2 RCTs)	⊕⊖⊖⊖ **Very low**[c]

2 months after end of treatment)	(0.66 lower to 0.86 higher). (A negative result favours the intervention group.)		
Behavioural problems: overall (4 weeks to 6 months after end of treatment)	The score in the intervention group was 0.09 SD lower (0.39 lower to 0.22 higher). (A negative result favours the intervention group.)	197 (4 RCTs)	⊕⊕⊝⊝ **Low**[b]
Social behaviour (4 weeks after end of treatment)	The score in the intervention group was 0.53 SD higher (0.53 lower to 1.60 higher). (A positive result favours the intervention group.)	48 (2 RCTs)	⊕⊝⊝⊝ **Very low**[d]
Cognition (3 months after end of treatment)	The score in the intervention group was 0.04 SD higher (0.56 lower to 0.64 higher). (A positive result favours the intervention group.)	47 (1 RCT)	⊕⊝⊝⊝ **Very low**[d]

*Interpretation of SMD: a difference of < 0.20 SD can be regarded as little or no effect; 0.20 to 0.40 SD as a small effect; 0.40 to 0.70 as a moderate effect; and > 0.70 as a large effect (Higgins 2019, Chapter 15.5.3.1).

CI: confidence interval; **SMD:** standardised mean difference; **SD:** standard deviations; **RCT:** randomised controlled trial

GRADE Working Group grades of evidence (GradePro GDT**)**
High certainty: we are very confident that the true effect lies close to that of the estimate of the effect.
Moderate certainty: we are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low certainty: our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
Very low certainty: we have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

[a]Downgraded by two levels: one level for risk of bias: no blinding of therapists and participants (not possible), and sometimes no or unclear blinding of outcome assessment; and one level for imprecision: small number of participants (< 400) and broad CI includes both benefit and harm

[b]Downgraded by two levels: one level for risk of bias: no blinding of therapists and participants (not possible); and one level for imprecision: small number of participants (< 400) and broad CI includes both benefit and harm

[c]Downgraded by three levels: one level for risk of bias: no blinding of therapists and participants (not possible); one level for imprecision: small number of participants (< 400) and broad CI includes both benefit and harm; and one level for inconsistency: effects in opposite directions

[d]Downgraded by three levels: one level for risk of bias: no blinding of therapists and participants (not possible); and two levels for imprecision: small number of participants (< 400) and very broad CI includes both benefit and harm

BACKGROUND

Description of the condition

Alzheimer's Disease International's 2022 report estimated that there were 55 million people with dementia worldwide in 2019, and that this figure will increase to 139 million people in 2050 (Alzheimer's Disease International 2022). Dementia is a collective name for progressive degenerative brain syndromes that affect memory, thinking, behaviour and emotion (Alzheimer's Disease International 2015). Dementia of the Alzheimer's type is the most common form of dementia, followed by vascular dementia, Lewy body dementia and frontotemporal dementia (Alzheimer's Disease International 2015).

Symptoms may include:

- loss of memory;
- difficulty finding the right words or understanding what people are saying;
- difficulty performing previously routine tasks; and
- personality and mood changes.

It is important to help people with dementia and their caregivers to cope with the syndrome's social and psychological manifestations. As well as trying to slow cognitive deterioration, the aim of care should be to stimulate abilities, improve quality of life and reduce problematic behaviours associated with dementia.

Researchers are pursuing a variety of promising findings related to the aetiology and treatment of dementia. As dementia is due to damage to the brain, one approach is to limit the extent and rate of progression of the pathological processes producing this damage. Pharmacological interventions are available but have limited ability to treat many of the syndrome's features (Bosnjak Kuharic 2021; Lim 2024; McShane 2019; Mühlbauer 2021). However, there is ample research that shows that non-pharmacological treatment approaches can effectively improve relevant outcomes (Cho 2024; Sikkes 2021). The therapeutic use of music might be helpful, but there is uncertainty about the evidence for its effectiveness.

Description of the intervention

Many treatments for dementia symptoms depend on a person's ability to communicate verbally. When the ability to speak or understand language has been lost, music may still offer an alternative opportunity for communication. People who can no longer speak may still be able to hum or play along with music.

Music therapy is defined by the World Federation of Music Therapy (WFMT) as "the professional use of music and its elements as an intervention in medical, educational, and everyday environments with individuals, groups, families, or communities who seek to optimise their quality of life and improve their physical, social, communicative, emotional, intellectual, and spiritual health and wellbeing." Research, practice, education and clinical training in music therapy are based on professional standards according to cultural, social and political contexts (WFMT 2011). The American Music Therapy Association (AMTA) defines music therapy as "the clinical and evidence-based use of music interventions to accomplish individualised goals within a therapeutic relationship by a credentialed professional who has completed an approved music therapy program" (AMTA). It describes assessment of the person, interventions ("including creating, singing, moving to, and/ or listening to music"), benefits and research, and explains that music therapy is used "within a therapeutic relationship to address physical, emotional, cognitive, and social needs of individuals."

'Music as therapy' includes but is not limited to therapy provided by "a formally credentialed music major with a therapeutic emphasis" (Ing-Randolph 2015). Our review is not narrowly focused on music therapy but on 'music-based therapeutic interventions'. Therapeutic goals and the establishing of a therapeutic relationship are part of the music-based therapeutic interventions in this review even when the intervention is not provided by an accredited music therapist. Further, music-based interventions may also be used in ways that are less obviously therapy or therapeutic, for example, playing music during other activities such as meals or baths, or during physiotherapy or movement, or as part of an arts programme or another psychosocial intervention. In order to benefit people with dementia, those providing music-based interventions with a therapeutic goal may need to draw on the skills of both musicians and therapists to select and apply musical parameters tailored to a recipient's individual needs and goals. The training of the therapists and the requirements of training programmes and certification practice to deliver music-based therapeutic interventions vary in different countries, which implies that accredited music therapists are not the only people able to deliver music-based therapeutic interventions. Therefore, music-based therapeutic interventions comprise more broadly defined therapy than music therapy alone in terms of who delivers the interventions and how it is applied.

Two main types of music-based therapeutic interventions can be distinguished: receptive (or passive) and active, and these are often combined (Guetin 2013). Receptive therapeutic interventions consist of listening to music offered by the therapist who sings, plays or selects recorded music for the recipients. In active music therapy, recipients are actively involved in the music-making, for instance by playing on instruments. The participants may be encouraged to participate in musical improvisation with instruments or voice, with dance, movement or singing.

How the intervention might work

Music-based therapeutic interventions, including interventions with therapeutic goals provided by a certified music therapist and others such as musicians or experienced therapists with other backgrounds, mostly consist of singing, listening, improvising or playing along on musical instruments. The neural circuitry for singing is more diffusely located compared with speech, which is concentrated in the speech centres in the left hemisphere of the brain. Therefore, the neural circuitry needed for singing is less likely to be disrupted and more often preserved than neuronal connections needed for speech (Särkämo 2018). People with aphasia due to left-hemisphere lesions often show strikingly preserved vocal music capabilities (Riecker 2000). Singing can further help the development of articulation, rhythm and breath control. Singing in a group setting can improve social skills and foster a greater awareness of others. For people with dementia, singing may encourage reminiscence and discussions of the past, while reducing anxiety and fear. Music interventions can activate brain functions that play a role in maintaining mood stability and reducing anxiety and depression (Ting 2023). Indeed, singing interventions have been shown to enhance the effect of music

therapy in recent network analyses (Ting 2024). For people with compromised breathing, singing can improve oxygen saturation rates. For people who have difficulty speaking following a stroke, music may stimulate the language centres in the brain, promoting the ability to sing. In summary, singing may improve a range of physical and psychosocial parameters (Clift 2016). Live music may be particularly effective in persons with major neurocognitive disorder compared to those with no or a mild neurocognitive disorder (Hobeika 2022), while to reduce agitation, receptive music therapy may be equally or even more effective than active forms of therapy (Tsoi 2018). Playing instruments may improve gross and fine motor co-ordination in people with motor impairments or neurological trauma related to a stroke, head injury or a disease process (Magee 2017).

Whereas cognitive functions decline during disease progression, receptivity to music may remain until the late phases of dementia (Aldridge 1996; Baird 2009; Cowles 2003). Even in the latest stage of the disease, people may remain responsive to music where other stimuli may no longer evoke a reaction (Norberg 1986). This may be related to musical memory regions in the brain being relatively spared in Alzheimer's disease (Jacobsen 2015). Possibly, the fundamentals of language are musical, and precede lexical functions in language development (Aldridge 1996). Listening to music itself may decrease stress hormones such as cortisol, and help people to cope with, for instance, preoperative stress (Spintge 2000). Music can bring relaxation; potential mechanisms for this include decreasing of physiological arousal and altering of autonomic, endocrine and immunological pathways, but few studies have been conducted in persons with dementia, and studies that have been conducted have mixed findings (De Witte 2022; Sittler 2021). Musical interventions enable the recall of life experiences and the experience of emotions. Many important life events are accompanied by music; most of the time these 'musical memories' are stored for a longer time than the ones from the same period that were not accompanied by music (Baird 2009; Broersen 1995). When words are no longer recognised, familiar music may provide a sense of safety and well-being, which, in turn, may decrease anxiety. Musical rhythm may help people with Alzheimer's disease to organise time and space. People are able to experience group contact through musical communication with other participants, without having to speak. Owing to its non-verbal qualities, music-based interventions might help people with dementia to cope with the effects of their illness (Thompson 2024).

Why it is important to do this review

This is an update of a Cochrane review first published in 2002, and last updated in 2018. At that time, we concluded that music-based therapeutic interventions probably reduce depressive symptoms and improve overall behavioural problems at the end of treatment, and may also improve emotional well-being and quality of life and reduce anxiety. We found that they may have little or no effect on agitation or aggression or on cognition, while effects on social behaviour and long-term effects were uncertain. The review has been a well-used resource for those seeking high-quality evaluation of effects, but is now seven years old. Therefore, it lacks inclusion of the most recent studies. In this update, we examined current research literature to assess whether music-based therapeutic interventions, including music therapy, are an efficacious approach to the treatment of emotional, behavioural, social and cognitive problems in people with dementia. We also

investigated whether, in the presence or the absence of specific problems, these interventions have an effect on social behaviour and emotional well-being, including quality of life. Quality of life is often an important goal of care for people with dementia (Alzheimer's Disease International 2016), and it is important to assess evidence whether music-based therapeutic intervention can contribute to quality of life and related outcomes.

There is a lack of information about how often music-based therapeutic interventions are used for people with dementia. For music therapists, people with dementia form a major clientele; for example, in the UK, an estimated 250 of 900 music therapists work with people with dementia, and this is likely an underestimate as hundreds of therapists were not surveyed (Bowell 2018). Music-based therapeutic interventions, in particular group interventions, are relatively inexpensive; they are also suitable for people in more advanced stages of dementia for whom relatively few interventions are available, as playing or humming along is still possible up until the later stages of the disease. The use of music-based therapeutic interventions is gaining traction and hence it is important to keep updating the collation of the evidence in a systematic way, including assessing the level of certainty we can have in the available evidence.

OBJECTIVES

To assess the effects of music-based therapeutic interventions for people with dementia on emotional well-being (including quality of life), mood disturbance or negative affect (i.e. depressive symptoms and anxiety), behavioural problems (i.e. overall behavioural problems or neuropsychiatric symptoms, and more specifically agitation or aggression), social behaviour and cognition, at the end of therapy and four or more weeks after the end of treatment, and to assess any adverse effects.

METHODS

Criteria for considering studies for this review

Types of studies

We included parallel and cross-over randomised controlled trials (RCTs). Quasi-RCTs were not eligible for inclusion.

Types of participants

We included participants with dementia who were formally diagnosed as having any type of dementia according to the Diagnostic and Statistical Manual of Mental Disorders (DSM)-IV or DSM-5, International Classification of Diseases (ICD)-10 or other accepted diagnostic criteria. In order to be relevant to clinical practice, we also accepted physician diagnosis of dementia if no data on formal criteria, such as DSM-IV, DSM-5 or comparable instruments, were available. We included studies of people living in any setting (including in nursing homes and other long-term care facilities, the community and hospitals) and with all types of dementia, regardless of severity. We did not use age as a criterion for inclusion or exclusion. We included studies with identifiable subsets of participants with dementia; for studies with mixed samples, we used the results for the eligible subset, or we asked the authors for the specific findings for the subset of participants with dementia.

Types of interventions

We included any music-based intervention - active, receptive, or mixed - delivered to individuals or groups. We required a minimum of five sessions to ensure that a therapeutic intervention could have taken place. We defined therapeutic music-based interventions as: therapy provided by a qualified music therapist, or interventions based on a therapeutic relationship between the participant(s) and a health or social care professional and meeting at least two of the following criteria based on indicators of therapeutic goals and skill in engaging participants: 1) therapeutic objective for the individual participant, which may include communication, relationships, learning, expression, mobilisation, etc; 2) music matches individual preferences, not just group preferences; 3) active participation of the people with dementia using musical instruments or singing; 4) participants had a clinical indication for the intervention or were referred for the intervention by a clinician.

Simple participation in a choir would not meet our definition of a therapeutic intervention; neither would an individualised music listening intervention with preferred music if there was no communication or opportunity for an intervention provider to relate to the person with dementia during the session. Music had to be a main element of the intervention rather than, for example, merely moving with music or socialising. More than two-thirds of a session should comprise a music-based therapeutic intervention (therefore, less than one third of a session could comprise socialising, movement only or other non-music based activities). No family therapy should be provided at the same time; therefore, dyadic interventions for both persons with dementia and family were not eligible.

The music-based therapeutic interventions could be compared against no therapy or activity, or they could be compared against any other type of therapy or activity, including activities in which music was used. However, they could not receive any music-based therapeutic intervention that met the above criteria (even if it involved fewer sessions than the intervention group). We categorised the control groups as two comparators: usual care, i.e. no interventions, and other activities, i.e. any other activity that was not a music-based therapeutic intervention.

Types of outcome measures

We included studies that measured any of the following outcomes of interest.

- Emotional well-being including quality of life and positive affect. Facial expressions (in the absence of interaction with the observer) may also indicate emotional well-being.
- Mood disturbance or negative affect
  - Depression (depressive symptoms)
  - Anxiety
- Behavioural problems
  - Agitation or aggression (or both). We combined agitation and aggression outcomes consistent with the International Psychogeriatric Association consensus definition of agitation requiring presence of one of "excessive motor activity, verbal aggression, or physical aggression" (Cummings 2015).
  - Overall behavioural problems or neuropsychiatric symptoms.
- Social behaviour, such as (verbal) interaction (which could also be measured during therapy and control activities)
- Cognition
- In addition to the seven outcomes of interest above, we searched for any adverse effects.

For these outcomes, we accepted all assessment tools used in the primary studies. Irrespective of the duration and number of sessions of the music-based intervention, we used outcomes that had been assessed at the end of treatment (i.e. after a minimum of five sessions), in order to focus on therapeutic goals achieved by the intervention as a whole as opposed to an immediate effect after a session, which may not last. If there were multiple follow-up assessments with evidence of no different effect over time, then we could include outcomes assessed before five sessions had been administered. To assess long-term effects, we also looked at outcomes measured a minimum of four weeks after the treatment ended.

Primary outcomes

- Emotional well-being including quality of life
- Mood disturbance or negative affect:
  - depression; and
  - anxiety.
- Behavioural problems:
  - agitation or aggression; and
  - overall.

We selected these as the primary outcomes as outcomes related to emotions are of critical importance to the quality of life experienced by people with dementia (Banerjee 2009; Beerens 2014; Verkaik 2007). Moreover, depression and anxiety are prevalent and rather persistent during the course of dementia (Van der Linde 2016; Zhao 2016). We also prioritised behavioural problems because these affect relationships and caregiver burden (e.g. Van der Linde 2012), and some may also be indicators of distress.

Secondary outcomes

- Social behaviour
- Cognition
- Potential adverse effects

Social behaviour and cognition were important, but we considered them as secondary outcomes because the benefit for the participants themselves is not as obvious as for outcomes more closely related to their quality of life. Potential adverse effects were assessed qualitatively. We included any adverse effects that were reported in the studies regardless of whether adverse effects was an explicit outcome in the study.

Search methods for identification of studies

Electronic searches

For previous versions of the review and on 30 November 2023, we searched the Cochrane Dementia and Cognitive Improvement Group's Specialised Register (CDCIG SR) using the search terms music therapy, music, singing, sing and auditory stimulation. The CDCIG SR contains studies (including reports of trials from healthcare databases, the Cochrane Central Register of Controlled Trials (CENTRAL) and conference proceedings) in the areas of dementia prevention, dementia treatment and cognitive enhancement in healthy people. The Dementia and Cognitive

Improvement Group has now closed and the CDCIG SR is no longer being updated.

On 30 November 2023, we performed additional searches in each of the sources in the bullet point list below to cover the timeframe from the last searches performed for the CDCIG SR up to 30 November 2023. The search strategies are presented in Appendix 1.

- MEDLINE Ovid SP (1946 to 29 November 2023)
- Embase Ovid SP (1996 to 2023 week 47)
- PsycINFO Ovid SP
- CINAHL EBSCOhost
- Web of Science Core Collection ISI Web of Science
- LILACS (Latin American and Caribbean Health Science Information database)
- US National Institutes of Health Ongoing Trials Register ClinicalTrials.gov (clinicalTrials.gov)
- World Health Organisation International Clinical Trials Registry Platform (apps.who.int/trialsearch)

In addition, we searched Cairs, Carl Uncover/Ingenta, Geronlit/Dimdi, Musica and Research Index in January 2006 and June 2010, with the following search terms: music therapy, music, singing, dance, dementia, alzheimer. On the same dates, we also searched specific music therapy databases, as made available by the University of Witten-Herdecke on www.musictherapyworld.de, based in Germany.

Searching other resources

We checked the reference lists of all relevant articles, and a clinical librarian conducted a forward search from key articles using SciSearch. In addition, we handsearched conference proceedings of European and World Music Therapy conferences and European music therapy journals, such as the Nordic Journal of Music Therapy (archive), the British Journal of Music Therapy, the Musiktherapeutische Umschau and the Dutch Tijdschrift voor Vaktherapie, up to December 2023. Potentially eligible new studies (based on abstract review with two review authors working independently) published after our search date were included in the Characteristics of studies awaiting classification table.

Data collection and analysis

We used the standard methodological procedures expected by Cochrane.

Selection of studies

Two review authors (who were not authors on any of the studies they assessed) independently assessed publications for eligibility by checking the title, and abstract, if available, using MS Excel to record their decisions (Microsoft Excel). If there was any doubt about eligibility, they retrieved and assessed the full article. Disagreements were resolved through discussion, or in consultation with a third or fourth author if needed. The search and selection process was depicted in a PRISMA diagram (Page 2021), which showed the number of records identified in each version of the review. Our unit of interest in the review was the study: some studies were reported in multiple articles and some articles reported multiple studies.

Data extraction and management

Two review authors (who were not authors of the studies they assessed) independently extracted and cross-checked data to assess eligibility using a brief data collection form. If a study was eligible, we proceeded to an independent assessment using a longer data collection form to extract data describing the study and outcome data. The Word templates for the data collection forms were piloted by two authors and finalised after discussion. The two authors discussed any discrepancies or difficulties with a third review author. We used Microsoft Excel to keep track of the process of study inclusion and evaluation. We reviewed articles in English, French, German and Dutch, and we searched for Cochrane collaborators to assess articles in other languages. We emailed authors of potentially eligible studies for additional information when anything was unclear.

We first extracted data on the design (RCT), population (dementia diagnosis), criteria for music-based therapeutic interventions, outcomes and timing of outcome assessment, to evaluate the eligibility of the study. Of the eligible studies, we subsequently recorded the following characteristics.

- Data collection period
- Setting: nursing home, residential home, hospital, ambulatory care, other
- Participant characteristics: age, sex, severity and type of dementia
- Number of participants included, randomised and lost to follow-up
- Type, frequency and duration of active interventions and control interventions
- Description of activities in the control group if not usual care
- Outcomes: type of outcome measures for emotional well-being including quality of life or positive affect, mood disturbance or negative affect as indicating emotional problems, problematic or challenging behaviours (in general; and more specifically, agitation or aggression), social behaviour and cognition; whether outcomes were referred to as primary or secondary outcomes. We searched articles for any reporting of adverse or potential adverse effects.
- Timing of outcome measurement: after treatment ended; in the longer term
- Research hypotheses if specified, and a description of the results
- Any methodological problems and comments
- Funding sources and conflicts of interest
- Items for a risk of bias assessment (below)

For each study, we extracted relevant outcome data, that is, means, standard deviations and number of participants in each group for continuous data, and for dichotomous data, the number experiencing each outcome in each group. If necessary or helpful, we contacted study authors for clarification or data.

Assessment of risk of bias in included studies

Two review authors (who were not authors of the studies they assessed) independently assessed the risk of bias in the included studies according to the guidelines in the *Cochrane Handbook for Systematic Reviews of Interventions*, and using the risk of bias assessment tool RoB 1 (Higgins 2011). We looked at the following

domains: selection bias (random sequence generation, allocation concealment); performance bias (blinding of participants and personnel); detection bias (blinding of outcome assessment); attrition bias (incomplete outcome data); reporting bias (selective reporting) and other potential threats to validity (any potential source of bias not covered by the other domains, such as imbalances in the assessments or data, and unclear bias in studies in which the therapist was the sole author of the first report on the study). We assessed performance, detection and attrition bias for each outcome. For each domain, we judged the risk of bias as low, unclear or high. Disagreements between the independent assessments were resolved through discussion; if necessary, a third author adjudicated.

Measures of treatment effect

To summarise the effects, we used the mean difference (MD). If different instruments or scales were used, we used the standardised mean difference (SMD) for continuous variables and the risk ratio (RR) for any dichotomous outcome variables, with 95% confidence intervals (CI). We interpreted SMDs using generic effect size estimates (Higgins 2019, Chapter 15.5.3.1). SMD values below 0.20 were interpreted as little or no effect, values between 0.20 and 0.40 as a small effect, values between 0.40 and 0.70 as a moderate effect and values above 0.70 as a large effect.

Unit of analysis issues

Only participant-level outcomes were considered, and all were continuous measures. For cross-over trials, we extracted data for the first period only because of the likelihood of carry-over effects. For cluster-randomised trials, we assessed whether clustering had been taken into account in the analysis. If this was not the case, or was unclear, we used the intraclass correlation coefficients (ICC) of outcomes from another music-based therapeutic intervention study, and we adjusted the number of participants per group by dividing the original number by the design effect and rounding to whole numbers (Higgins 2019: Chapter 23).

Christian Gold, one of the senior investigators of the MIDDEL study, provided ICCs for various outcome measures, to be used for the other cluster-RCTs: MADRS: ICC = 0.09 (in Baker 2022). QOLAD: ICC = 0.010; NPI Severity, ICC 0.06. Mean ICC to be used for other outcomes: 0.09.

In studies with more than one control group contributing to both comparisons (versus usual care and versus other activities), in order to avoid double-counting the intervention group, we adjusted its size (i.e. reduced by 50% in case of two control groups) (Higgins 2019, Chapter 23.3.4).

Dealing with missing data

We considered, in reporting and risk-of-bias assessments, if there were missing outcome data, with reasons reported, for example, due to participants who were hospitalised, moved or died, and how these were dealt with (exclusion of cases for analyses or were dealt with otherwise). We did not contact authors about missing data or impute missing data. We considered availability of intention-to-treat data and sensitivity analyses employing different manners of managing missing outcome data when available.

Assessment of heterogeneity

We interpreted the I² statistic according to criteria in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2019: Chapter 10.10.2). It offers a rough guide, with no important heterogeneity for I² up to 40%, moderate heterogeneity between 30% and 60%, substantial heterogeneity between 50% and 90%, and considerable heterogeneity for I² of 75% and higher. Further, a low P value for the Chi² statistic indicated heterogeneity of intervention effects, which we evaluated against the combined 'usual care' and 'other activities' control groups. Because of small numbers of participants and studies for part of the outcomes, a non-significant P value was not decisive in the evaluation of consistency, and we also considered overlap of CIs in the forest plots. We considered possible clinical heterogeneity in type of intervention, setting and participants differing between studies, and possible methodological heterogeneity by variability in study design, outcomes and risk of bias.

Assessment of reporting biases

Selective outcome reporting is one of the elements of the risk of bias assessment, and for this we searched the articles on included studies and related articles for references to study protocols and trial registrations. If available, we compared with outcomes and prioritisation of outcomes in the article. If there was no research protocol available, we set risk of reporting bias to either unclear or high when appropriate. To detect possible publication bias, we examined funnel plots for outcomes with at least 10 studies available.

Data synthesis

We included studies about all eligible interventions in people in different stages of dementia, and we pooled the results of studies that examined effects on the same seven outcomes of interest. For the meta-analyses, we distinguished between effects at the end of treatment and long-term effects (a minimum of four weeks after treatment ended). We also distinguished between the comparators: control groups offered usual care versus control groups offered other activities. In the case of more control groups, we assigned proportionally less weight to the number of participants in the intervention group. We carried out meta-analyses, irrespective of the number of studies that were available. We used the random-effects (inverse-variance) model since we did not expect that the effect of music-based therapeutic interventions across studies is identical (Higgins 2019, Chapter 10.10.4). We used Review Manager software (RevMan 2024).

Subgroup analysis and investigation of heterogeneity

We examined effects of individual therapy versus effects of group therapies. We separately analysed if there were differences between the estimates of individual versus group therapy for outcomes for which at least five studies for each of the subgroups was available. This was the case for only one outcome. We used the test for subgroup differences available in RevMan. We reviewed if heterogeneity differed substantially, and if there were any patterns across outcomes.

Sensitivity analysis

Post hoc, we performed a series of sensitivity analyses because there are different possible criteria as to what constitutes music therapy or a therapeutic intervention considering the person

who delivers the intervention, and because funding related to music therapy potentially involves an intellectual conflict of interest. First, we reran all analyses on end-of-treatment effects with studies in which the intervention was definitely or possibly (when not mentioned explicitly) delivered by a professional music therapist (with or without support of other professionals). Second, we restricted the analyses to studies definitely delivered by a professional music therapist with no other professionals involved. Third, we restricted the analyses to studies definitely delivered by a professional music therapist with no other professionals involved and with no potential conflict of interest related to funding parties with a potential interest in promoting music-based therapeutic interventions or no reported funding source.

As masking participants and therapists was impossible, all studies were at high risk of bias for this item. Studies varied in the risk of bias on other items. We therefore also performed sensitivity analyses by removing studies at high risk of bias for any other item, to evaluate change and direction of change of the estimate.

Summary of findings and assessment of the certainty of the evidence

We used GRADE (GradePro GDT) to assess the certainty of evidence for each outcome listed below. Two review authors, neither of whom was an author on any of the studies they assessed, used GRADE methods to rate the certainty of evidence (high, moderate, low or very low) (Guyatt 2011). This rating refers to our level of confidence that the estimate reflects the true effect, taking account of risk of bias in the included studies, inconsistency between studies, imprecision in the effect estimate, indirectness in addressing our review question and the risk of publication bias. For each domain, the certainty of the evidence could be downgraded by one level (serious) or two levels (very serious). The certainty of the evidence was assessed with two authors independently initially, after which we discussed and resolved disagreements. Starting from initial high-certainty evidence, we downgraded for risk of bias for each domain if present. As all information was from studies at high risk of bias due to lack of blinding participants and therapists, we downgraded the certainty of the evidence to moderate. In case of additional risk of bias in other domains we downgraded further (to low or very low).

We produced summary of findings tables for end-of-treatment and long-term outcome comparisons to show the effect estimate and the quantity and certainty of the supporting evidence for the outcomes. The summary of findings tables were generated with data from RevMan (RevMan 2024) that we imported into the GRADEpro Guideline Development Tool (GradePro GDT); for the last two updates, the tables were revised manually. For this update, separate summary of findings tables were produced for music therapy versus usual care, and for music therapy versus other activities. For each comparison we report on the following outcomes, both at end of treatment and long-term effects:

- emotional well-being including quality of life;
- mood disturbance or negative affect: depression;
- mood disturbance or negative affect: anxiety;
- behavioural problems: agitation or aggression;
- behavioural problems: overall;
- social behaviour; and
- cognition.

RESULTS

Description of studies

Results of the search

As the PRISMA flow diagram shows (Figure 1), the total number of included studies for this update is 30, which is eight more than the previous update (Van der Steen 2018). These 30 studies were reported in 47 articles or communications. Twenty-eight studies contributed data to the meta-analyses, with no data available from Lord 1993 or Prieto Alvarez 2022.

Figure 1. PRISMA study flow diagram

Records identified through database searching:

2003: 354 records

2008: 107 records

2010: 188 records

2015: 874 records

2017: 286 records

2021: 2383 records

2023: 1438 records

0 additional records identified through other sources

Records after duplicates removed:

2003: 354 records

2008: 74 records

2010: 188 records

2015: 121 records

2017: 32 records

2021: 1387 records

2023: 986 records

Records screened:

2003: 354 records

2008: 74 records

2010: 188 records

2015: 121 records

2017: 32 records

2021: 1387 records

2023: 986 records

Records excluded:

2003: 328 records

2008: 56 records

2010: 172 records

2015: 96 records

2017: 26 records

2021: 1348 records

2023: 974 records

Full-text articles

Figure 1. (Continued)

Full-text articles assessed for eligibility: 142

2003: 26 full-text articles

2008: 18 full-text articles

2010: 16 full-text articles

2015: 21 full-text articles (count does not include 4 articles awaiting classification)

2017: 10 full-text articles (count includes 4 articles awaiting classification)

2021: 39 full-text articles on 32 studies (count includes 5 additional articles to the main article on one study, and 2 additional articles on an ongoing study)

2023: 12 full-text articles (count includes 1 article awaiting classification)

Full-text articles excluded: 108

2003: 21 + 3 full-text articles

2008: 15 + 2 full-text articles

2010: 14 full-text articles

2015: 10 full-text articles

2017: 5 full-text articles (0 articles awaiting classification were excluded)

2021: 27 full-text articles on 25 studies (includes 3 articles on an ongoing study)

2023: 11 full-text articles

Studies included: 30

2003: 5 studies (3 of which were excluded in 2015)

2008: 3 studies (3

Figure 1.   (Continued)

excluded in 2015)

2008: 3 studies (2 of which were excluded in 2015)

2010: 2 studies

2015: 12 studies (11 main articles; including an article with 2 studies)

2017: 5 studies

2021: 7 studies

2023: 1 study

studies included in quantitative synthesis (meta-analysis): 28 of 30

Since the last published version of the review (Van der Steen 2018), we have run the searches twice. Our search in 2021 resulted in 1387 unique records, and our search in 2023 found 986 unique records. We reviewed 39 full-text articles (reporting on 32 studies) from the 2021 search and included seven new studies (13 articles) (Baker 2022; Giovagnoli 2018; Gómez-Gallego 2021; Liu 2021; Reschke-Hernández 2019; Tang 2018; Werner 2017), and we reviewed 12 full-text articles from the 2023 search and included one (Prieto Alvarez 2022).

We have listed three ongoing studies in this update. Two compare individual interventions with an 'attention' control (Baroni Caramel 2024; NCT04666077). Baroni Caramel 2024 is conducted in the nursing home setting, while NCT04666077 is conducted at home. NCT04666077 also includes participants with mild cognitive impairment and a control group in which the family caregiver is supported to provide a singing intervention. The third ongoing study, which we have designated Gold 2018, is a multinational study, data from which we have already included in this review, i.e. the Australian data; see Baker 2022 for information about the study's intervention, setting and participants.

We also document 41 excluded studies and 19 studies await classification.

In Appendix 2, we describe the history of the searches undertaken and studies identified for inclusion in the review since it was first published.

Included studies

We emailed the authors of 26 of the 30 included studies for additional information about the type of control group or setting, and for additional data to enable us to include the study in meta-analyses (e.g. when estimates from graphical presentation were imprecise, standard deviations (SD) were lacking or when we needed item-level data from global tools used for relevant outcomes). We received responses from the authors of 24 of the 26 studies. We requested numerical data from 13 authors, and 12 of them provided unpublished data.

We present details of the included studies in the Characteristics of included studies table. One article (Narme and colleagues 2012: Narme 2012-study 1 and Narme 2012-study 2) reported on two studies with rather similar designs, indicated by study 1 and study 2 in the article. Six studies were reported in more than one article (Baker 2022; Cooke 2010; Lin 2011; Narme 2014; Raglio 2010a; Vink 2013).

Funding

For eight studies, funding sources were unknown (Clark 1998; Liesk 2015; Lin 2011; Lord 1993; Lyu 2014; Raglio 2010a; Raglio 2010b) or unclear (Liu 2021). Six studies did not receive external funding (Cho 2016; Gómez-Gallego 2021; Prieto Alvarez 2022; Raglio 2015; Tang 2018; Werner 2017). The other 16 studies reported a funding source, which was national, institutional or from a foundation or training programme. At least three studies were conducted in the context

of a Masters or PhD (Cho 2016; Reschke-Hernández 2019; Werner 2017).

Participants

The total number of participants randomised in the studies was 1720, 1366 of whom contributed to the meta-analyses. The total number of randomised participants varied between 14 (Narme 2012-study 2) and 239 of 318 assigned to a single condition (the others received two interventions; Baker 2022), with a median number of 47.5 participants across the studies. Twelve studies randomised fewer than 40 participants, and three had more than 100 participants. Twelve studies selectively included people with agitation, mood problems or behavioural problems (Baker 2022; Clark 1998; Cooke 2010; Guétin 2009; Hsu 2015; Liu 2021; Raglio 2010a; Raglio 2015; Ridder 2013; Sung 2012; Tang 2018; Vink 2013), while some studies excluded people with major psychiatric conditions such as psychosis or major depression, or people with other medical conditions such as hearing impairment or acute illness. Dementia severity varied.

Design

Twenty-five studies had a parallel-group design (Baker 2022; Ceccato 2012; Cho 2016; Giovagnoli 2018; Gómez-Gallego 2021; Guétin 2009; Hsu 2015; Liesk 2015; Lin 2011; Lord 1993; Liu 2021; Lyu 2014; Narme 2012-study 1; Narme 2012-study 2 ; Narme 2014; Raglio 2010a; Raglio 2010b; Raglio 2015; Sakamoto 2013; Sung 2012; Svansdottir 2006; Tang 2018; Thornley 2016; Vink 2013; Werner 2017). Five studies used a cross-over design (Clark 1998; Cooke 2010; Prieto Alvarez 2022; Reschke-Hernández 2019; Ridder 2013), with no data from the first period available for Prieto Alvarez 2022.

Five relatively recent studies employed a cluster-randomised design (Baker 2022; Gómez-Gallego 2021; Hsu 2015; Reschke-Hernández 2019; Werner 2017). Only one of these cluster-randomised studies took the cluster design into account in the analysis (Baker 2022).

Long-term effects were assessed between four weeks and six months after treatment ended, with a median of two months after the last session.

Settings

The 30 studies were performed in 15 countries. The two oldest studies were from the USA (Clark 1998; Lord 1993), as were three other studies (Cho 2016; Prieto Alvarez 2022; Reschke-Hernández 2019), and one study was conducted in Canada (Thornley 2016). Seventeen studies were conducted in Europe: Italy, Spain, France, Germany, the Netherlands, the UK and Iceland, with one study performed in two countries, Denmark and Norway (Ridder 2013). Six studies were conducted in Asia (Taiwan (Lin 2011; Liu 2021; Sung 2012), Japan (Sakamoto 2013) and China (Lyu 2014; Tang 2018)). Two studies were conducted in Australia (Baker 2022; Cooke 2010).

One study was conducted as a follow-up after hospitalisation (Giovagnoli 2018) (setting unclear), and some studies were conducted on specific inpatient units (psychiatric (Thornley 2016), older adults or geriatric (Lyu 2014; Narme 2012-study 1; Narme 2012-study 2), both unit for older adults and nursing home (Raglio 2015), both dementia ward and nursing home (Svansdottir 2006; Sakamoto 2013)). However, most studies were performed in institutional long-term care settings of nursing homes, residential homes or assisted living. We categorised the settings as long-term care facilities (i.e. nursing homes, residential homes, assisted living); hospital departments (i.e. geriatric, psychogeriatric, dementia or psychiatry unit); or unclear.

Interventions

The interventions were active (Prieto Alvarez 2022; Cho 2016; Cooke 2010; Giovagnoli 2018; Gómez-Gallego 2021; Hsu 2015; Liesk 2015; Liu 2021; Lyu 2014; Raglio 2010a; Raglio 2010b; Raglio 2015; Reschke-Hernández 2019; Sung 2012; Thornley 2016), receptive (listening interventions while there was communication with the therapist) (Clark 1998; Guétin 2009) or a mixture (Baker 2022; Ceccato 2012; Lin 2011; Lord 1993; Narme 2012-study 1; Narme 2012-study 2; Narme 2014; Ridder 2013; Sakamoto 2013; Svansdottir 2006; Tang 2018; Vink 2013; Werner 2017). Appendix 3 contains a description of the music-based therapeutic intervention and other activities in all studies.

Music included live music that met the preferences of the group or individual or recorded music that met the preferences of the individual. The active forms often combined instrument playing and singing activities, and some also combined these with movement such as clapping hands and dancing. In eight studies, the intervention was individual (Clark 1998; Guétin 2009; Hsu 2015; Raglio 2010a; Raglio 2015; Ridder 2013; Sakamoto 2013; Thornley 2016).

Session duration varied between 25 minutes and two hours. The total number of sessions ranged from six (Narme 2012-study 1) to 156 (Lord 1993), with a median total number of 14 sessions until the end-of-treatment assessment. The frequency ranged between one session per week (Gómez-Gallego 2021; Guétin 2009; Hsu 2015; Liu 2021; Sakamoto 2013) and seven sessions per week (daily, Lyu 2014), with a median and mode of two sessions per week (16 studies used two per week). These figures probably reflect the number of sessions offered, as the number of attended sessions may be lower. There were few reports about implementation fidelity, including adherence or dose received. However, Ridder 2013 reported that a minimum of 12 sessions were offered, but the participants received a mean of 10 sessions, and Thornley 2016, in their study in an acute inpatient psychiatric unit within an academic hospital, mentioned that the participants enroled in the study were generally hospitalised for two to three weeks, which limited the number of sessions attended. Baker 2022 and Werner 2017 reported on pragmatic trials, including process evaluations. Prieto Alvarez 2022 counted the total number of sessions attended in the intervention group and the two control groups and found the differences were small.

Intervention providers

In 19 of the studies, we could be sure from the report that the interventions had been delivered by an accredited music therapist (Baker 2022; Ceccato 2012; Cho 2016; Giovagnoli 2018; Gómez-Gallego 2021; Hsu 2015; Lin 2011; Lyu 2014; Prieto Alvarez 2022; Raglio 2010a; Raglio 2010b; Raglio 2015; Reschke-Hernández 2019; Ridder 2013; Svansdottir 2006; Tang 2018; Thornley 2016; Vink 2013; Werner 2017). In four studies, it was unclear whether a music therapist was involved: no profession was reported in the older studies (Clark 1998; Lord 1993); there was probable delivery by trained music therapists, but it was not explicitly stated in Guétin 2009; and in Cooke 2010, musicians trained in the delivery of

sessions and in working with older people with dementia delivered the intervention, but it was unclear if they were formally trained music therapists. In the other seven studies, the intervention was not delivered by a music therapist but by another professional (psychologist and other supervisor(s) with no training in music therapy: Narme 2012-study 1; Narme 2012-study 2; Narme 2014; trained research assistants: Sung 2012; music facilitator: Liu 2021 and Sakamoto 2013; music teacher specialised in teaching older people: Liesk 2015).

Control conditions

We distinguished between usual care and active control interventions; we labelled the latter 'other activities'. Usual care was a comparator in the following studies: Baker 2022; Ceccato 2012; Clark 1998; Giovagnoli 2018; Gómez-Gallego 2021; Hsu 2015; Lin 2011; Lyu 2014; Raglio 2010a; Raglio 2010b; Raglio 2015; Ridder 2013; Sakamoto 2013; Sung 2012; Svansdottir 2006; Tang 2018). In one of the studies, both groups received memantine 20 mg/day added to treatment as usual with cholinesterase inhibitors (Giovagnoli 2018).

About half of the studies compared the music-based therapeutic intervention with an active control intervention that had the same number of sessions and frequency as the music group, but it differed in more pragmatic trials such as Werner 2017. The two-armed studies compared a music-based therapeutic intervention with the following: reading (Cooke 2010; Guétin 2009; Liu 2021), a cognitive stimulation intervention (Liesk 2015), painting (Narme 2012-study 1), cooking (Narme 2012-study 2; Narme 2014), verbal discussion sessions (Reschke-Hernández 2019), recreational choir singing (Werner 2017), individual active engagement activities (Thornley 2016), or variable recreational activities, which included handwork, playing shuffleboard, and cooking and puzzle games (Vink 2013). Eight studies had three arms, with active control groups working on jigsaw puzzles (Lord 1993), reading familiar lyrics (Lyu 2014), watching television or nature video (Cho 2016; Gómez-Gallego 2021; Prieto Alvarez 2022), taking part in physical and cognitive exercises such as games (Prieto Alvarez 2022), or receiving a passive group music intervention that did not meet our inclusion criteria for a therapeutic music-based intervention (Baker 2022; Cho 2016; Gómez-Gallego 2021; Raglio 2015; Sakamoto 2013).

Outcomes and measures

Outcomes that were assessed often were emotional well-being, including quality of life; mood disturbance or negative affect (including as part of behavioural scales); behavioural problems (agitation or aggression, and behaviour overall); and cognition. Social behaviour was less commonly assessed (Giovagnoli 2018; Lord 1993; Narme 2012-study 1; Narme 2012-study 2; Narme 2014; Reschke-Hernández 2019; Tang 2018), and the meta-analyses of end-of-treatment scores included only six studies, three of which were from Narme and colleagues. In particular, the Cohen-Mansfield Agitation Inventory (CMAI, for agitation; Cohen-Mansfield 1986), Mini-Mental State Examination (MMSE, for cognition; Folstein 1975) and the Neuropsychiatric Inventory (NPI, for behaviour; Cummings 1994) were frequently used. Item-level NPI outcome data were reported in the articles or the authors provided data about depression, anxiety and agitation outcomes.

Excluded studies

We screened 3142 records and rejected 3000 (Figure 1). Of the remaining 142 records examined in full text, we excluded a total of 108 records. We also excluded five studies that had previously been included in the review. See Characteristics of excluded studies for examples of excluded studies.

Many studies were excluded because the participants did not have dementia or the trial was not an RCT. We excluded some studies because the intervention did not meet our criteria for a music-based therapeutic intervention, or it had fewer than five sessions. Further, some studies assessed outcomes (other than social behaviour) during the treatment sessions only, or combined with effects after multiple sessions during the cross-over wash out period (e.g. Gerdner 2000). Some studies were excluded because music was not the main or only therapeutic element, or was not provided with individual therapeutic intent. We excluded several studies on dyadic interventions that were fully or partially delivered by family caregivers at home after instruction, and studies with audio-recorded delivery of music but no interaction with a therapist.

Studies that await classification, 19 in total, included conference abstracts and articles about studies in Asia, which we could not retrieve or were unable to evaluate in time for the submission of this update (see Characteristics of studies awaiting classification table).

Risk of bias in included studies

We present the results of our assessment of the risk of bias in the studies in the Risk of bias in included studies tables and in figures (Figure 2; Figure 3).

Figure 2.   Risk of bias summary: review authors' judgements about each risk of bias domain for each included study

Figure 2.   (Continued)

Figure 3.   Risk of bias graph: review authors' judgements about each risk of bias domain presented as percentages across all included studies
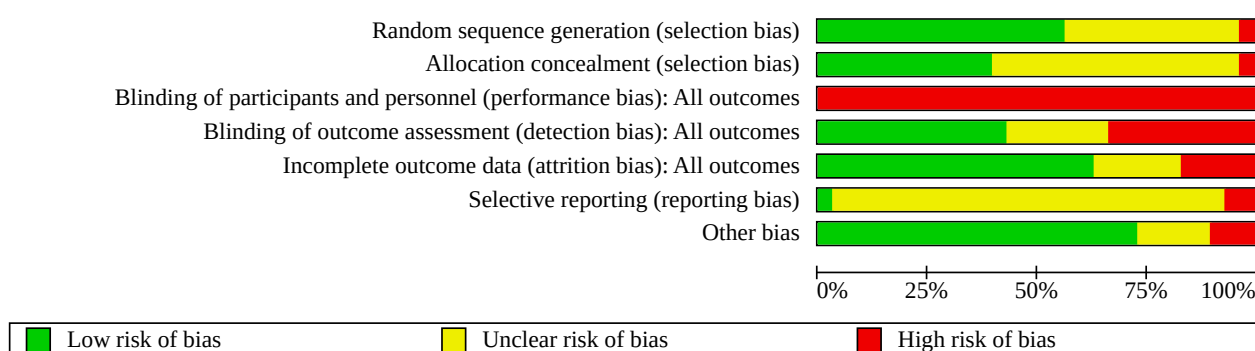
There were a number of possible biases, and often we could not assess the risk of bias due to poor reporting. The risk of performance bias was high for all studies because participants and staff could not be blinded to the intervention. Regarding the other items, in more recent studies, the risk of bias was lower. An exception was attrition bias. However, it is possible that this was reported more accurately in recent studies, i.e. the reporting in terms of interventions, rationale, chosen procedures, design and results was generally better in more recent studies. Still, we are unsure about the methodological quality of a number of studies because we rated several items as unclear.

Overall, all studies were at high risk of bias due to the fact that therapists and participants could not be blinded. Sixteen studies were also at high risk of bias for one or more of the other domains. In the section below, we provide details for each of the risk of bias domains.

The funnel plots we generated to consider possible publication bias are presented in Appendix 4.

Allocation

All the included studies were RCTs, but the randomisation procedure was not always described in detail (Figure 2). We assessed 17 studies at low risk of bias for the randomisation procedure and 12 were unclear. We considered allocation concealment to be at low risk of bias in 11 studies (all of which were published in 2010 or later) and unclear in 18 studies due to lack of information. We assessed one study to be at high risk of bias in both domains (Lord 1993); this study stated that participants were "non-systematically separated" into groups, without further detail.

Blinding

Blinding of therapists and participants to the intervention is not possible. Therefore, all studies were at high risk of performance bias even though therapists do not generally assess outcomes and participants may not be aware, have no specific expectations or be unable to self-report.

The outcome assessments were reported as blinded in 13 studies, and as unblinded, by the research team or unblinded nurses, for 10 studies; seven studies were unclear (Figure 2). For example, Narme and colleagues described two studies differing in detection bias (Narme 2012-study 1; Narme 2012-study 2). The first study involved a high risk of detection bias because anxiety (measured with the State-Trait Anxiety Inventory for adults, STAI-A) and, as assessed from the first two minutes of filmed interviews, emotions (from facial expressions) and social behaviour (discourse content), were assessed by nurses who were not blinded to the interventions (music intervention or painting) (Narme 2012-study 1). By contrast, in the second study, the risk of detection bias was low because five independent observers who were blinded to the type of intervention (music intervention or cooking) assessed the outcomes (Narme 2012-study 2). The risk of performance bias in several studies and the risk of detection bias for some outcomes resulted in downgrading of the certainty of the evidence for all end-of-treatment outcomes and for all long-term outcomes.

Incomplete outcome data

Self-reported outcomes were rarely used in the studies. Further, no outcome data were available when death, hospitalisation, acute illness or no interest in the therapy occurred across the different study arms. Cases with no outcome data were not included in the analyses.

We did not consider incomplete outcome data to be a problem in 19 studies (low risk of bias), but it was unclear in six studies. We considered incomplete outcome data to be problematic in five studies (Baker 2022; Cho 2016; Hsu 2015; Prieto Alvarez 2022; Thornley 2016). In Hsu 2015, three of nine participants in the intervention group died (and one of eight in the control group). Cho 2016 could include only eight of 17 participants in the television watching (control) group analysis of quality of life (one missed outcome data, four were lost to follow-up and four refused to come, while fewer were lost to follow-up in the music therapy and music listening groups) and suggested this was because individual preferences for television programmes were not taken into account. Thornley 2016 did not perform their study in a long-term care setting but in an inpatient psychiatric unit of a hospital, and some participants were discharged after having attended a few sessions. In the pragmatic trial of Baker 2022, there was loss to follow-up due to COVID-19 outbreaks and lockdown. The trial authors conducted a sensitivity analysis excluding data collected during the pandemic, when data were more often missing, and found little difference in outcome. Prieto Alvarez 2022 reported that two of three outcome measures were available for only 21/29 (72%) and 18/29 (62%) of participants.

Newer studies often presented cases lost to follow-up and missing outcome assessment in detail using flow diagrams. Several studies only reported the number of cases randomised (and analysed) and did not explicitly report reasons for missing outcome data by study arm, or how these were handled. Therefore, it is possible that attrition bias was problematic in more studies, but that the reporting of missing outcome data was better in newer studies.

Selective reporting

We considered only one study to be at low risk for selective reporting (Sakamoto 2013). We found some indications of inconsistent reporting of primary and secondary outcomes (Cooke 2010; Hsu 2015). These two studies were considered to be at high risk of bias. For 27 studies, selective reporting posed an unclear risk of bias. Most studies, including the newer studies, did not refer to initial plans, a study protocol or trial registration. Therefore, it was unclear to what extent bias due to selective outcome reporting was pertinent. One study referred to initial plans in the article reporting on the results (Ceccato 2012), and protocol articles were available for two studies (Baker 2022; Cooke 2010), in addition to trial registrations. Four other studies referred to a trial registration (Gómez-Gallego 2021; Hsu 2015; Reschke-Hernández 2019; Sakamoto 2013), three of which were registered in retrospect, after data collection started or ended (Cooke 2010; Gómez-Gallego 2021; Reschke-Hernández 2019). Two studies clearly referred to a change from their initial plan (Baker 2022; Ceccato 2012). Baker 2022 reported deviations from the protocol in the outcomes and timeline of outcome measurement (removal of a cognition measure and assessment at 24 months). Ceccato 2012 reported there was no follow-up because of lack of funding. We did not downgrade the certainty of the evidence because of the unclear risk of selective reporting.

For four outcomes, there were enough studies to construct funnel plots (Appendix 4).

Other potential sources of bias

We found some other potential sources of bias. In total, we considered three studies to be at high risk of other sources of bias, five studies to be unclear and 22 studies to be at low risk of other sources of bias.

One study noted that outcome assessment may be either imprecise or biased by the use of non-validated outcome measures with suboptimal distributions (such as skewed distributions, e.g. observed number of times yelling; Clark 1998). Further, we found problems with the reporting of outcomes in Lord 1993. (We also suspected errors in Hong 2011 and listed this study under Studies awaiting classification for this reason). Implementation fidelity, including non-adherence and problems with measurements, were infrequently described, but Liesk 2015, one of the few studies with null findings, reported on this in detail.

In some studies, the music therapists were the authors of the study publication: Baker 2022; Cho 2016; Giovagnoli 2018; Gómez-Gallego 2021; Hsu 2015; Prieto Alvarez 2022; Raglio 2015; Reschke-Hernández 2019; Svansdottir 2006; Tang 2018; Thornley 2016; Vink 2013; Werner 2017). We judged there to be an unclear risk of intellectual conflict of interest if the music therapist was the sole author of the first publication about the study, which was the case for three studies (Cho 2016; Prieto Alvarez 2022; Reschke-Hernández 2019). Other potential sources of bias that we judged to be unclear were the impact of COVID-19 in Baker 2022 and unclear baseline characteristics in Svansdottir 2006.

Effects of interventions

See: **Summary of findings 1** Music-based therapeutic interventions compared to usual care for people with dementia: end-of-treatment effects; **Summary of findings 2** Music-based therapeutic interventions compared to usual care for people with dementia: long-term effects; **Summary of findings 3** Music-based therapeutic interventions compared to other activities for people with dementia: end-of-treatment effects; **Summary of findings 4** Music-based therapeutic interventions compared to other activities for people with dementia: long-term effects

Music-based therapeutic interventions compared to usual care

Results at the end of treatment are summarised in Summary of findings 1 and longer-term effects in Summary of findings 2.

Sixteen studies addressed this comparison and contributed to the meta-analyses (Baker 2022; Ceccato 2012; Clark 1998; Giovagnoli 2018; Gómez-Gallego 2021; Hsu 2015; Lin 2011; Lyu 2014; Raglio 2010a; Raglio 2010b; Raglio 2015; Ridder 2013; Sakamoto 2013; Sung 2012; Svansdottir 2006; Tang 2018). Three of the studies used a cluster design (Baker 2022; Gómez-Gallego 2021; Hsu 2015). Only Baker 2022 took the cluster design into account in the data analysis. For the other two cluster-RCTs, we adapted the group sizes by the design effect using intraclass correlation coefficients from Baker 2022.

Results at end of treatment

Emotional well-being including quality of life

We included four studies with 154 participants in the analysis of end-of-treatment scores for the critically important outcome of emotional well-being and quality of life. The studies used the following instruments: the Quality of Life-Alzheimer's Disease (QOL-AD) (Baker 2022); the Cornell-Brown Scale for Quality of Life in Dementia (CBS-QoL) (Raglio 2015); a Dementia Care Mapping Wellbeing score (Hsu 2015); and a Danish translation of the Alzheimer's Disease-Related Quality of Life (ADRQL) (Ridder 2013).

We judged the certainty of the evidence for end-of-treatment effects of music-based interventions on emotional well-being and quality of life as very low. We found little or no clear evidence of a positive effect favouring the intervention at the end of treatment (SMD 0.14, 95% CI −0.29 to 0.56; Analysis 1.1), but the evidence is very uncertain. Heterogeneity was low ($I^2 = 29\%$). There was no blinding of outcome assessment in two of the four studies and whether blinding was successful was unclear in one study.

Mood disturbance or negative affect: depression

Nine studies contributed 441 participants to the analysis of end-of-treatment effects. Depression or depressive symptoms were measured with translated versions of the Geriatric Depression Scale (GDS), the Cornell Scale for Depression in Dementia, the Montgomery-Åsberg Depression Rating Scale, MADRS), the Alzheimer's Disease and Related Dementias Mood Scale (AD-RD), or with a subscale of the Behavioural Pathology in Alzheimer's Disease (BEHAVE-AD) or the NPI.

We judged the certainty of the evidence as moderate for depression at the end of treatment. We found that music-based therapeutic interventions probably result in a slight reduction in depressive symptoms at the end of treatment (SMD −0.23, 95% CI −0.42 to

0.04; Analysis 1.2; note that for this outcome, a negative SMD favours music-based therapeutic interventions). There was no heterogeneity ($I^2 = 0\%$ for both).

Mood disturbance or negative affect: anxiety

The other mood item we considered was anxiety. For this outcome, at the end of treatment, we included seven studies with 282 participants. A variety of translated outcome measures were used: Rating Anxiety in Dementia Scale (RAID), STAI-A, Hamilton Anxiety Scale and subscale scores of the BEHAVE-AD and NPI.

We judged the certainty of the evidence as low for anxiety at the end of treatment; therefore, we have little confidence in the results. Music-based therapeutic interventions may have little to no effect on anxiety (SMD −0.15, 95% CI −0.39 to 0.09; Analysis 1.3; a negative value favours music-based therapeutic interventions). There was no heterogeneity in end-of-treatment effects ($I^2 = 0\%$).

Behavioural problems: agitation or aggression

Eleven studies with 503 participants contributed to analysis of the end-of-treatment effects. Outcome measures used for agitation were translated versions of the CMAI and the agitation subscale of the NPI; and for aggression, the aggressiveness subscale of the BEHAVE-AD and counts of observed aggressive behaviour.

There was no clear evidence of publication bias regarding end-of-treatment effect (Appendix 4). We judged the certainty of the evidence as moderate for the end-of-treatment outcome. Music-based therapeutic interventions probably have little to no effect on agitation or aggression at the end of treatment (SMD −0.05, 95% CI −0.27 to 0.17; Analysis 1.4; a negative result favours music-based therapeutic interventions). Heterogeneity was moderate ($I^2 = 32\%$).

Behavioural problems: overall

Ten studies with 385 participants contributed to the end-of-treatment effect analysis. Outcome measures were (sometimes translated) versions of the BEHAVE-AD and NPI.

We judged the certainty of the evidence as low for end-of-treatment effects. The evidence suggests that music-based therapeutic interventions may result in a slight reduction in problematic behaviour overall at the end of treatment (SMD −0.31, 95% CI −0.60 to −0.02; Analysis 1.5; a negative result favours music-based therapeutic interventions). Heterogeneity was moderate ($I^2 = 42\%$). There was some evidence of publication bias regarding the end-of-treatment effect (Appendix 4).

Social behaviour

Only two studies (121 participants) contributed to the end-of-treatment effect analysis. No studies contributed to the analysis of long-term effects. Tang 2018 used the Holden Communication Scale, addressing various dimensions of communication. Giovagnoli 2018 used the Severe Impairment Battery (SIB), reporting on social interaction.

We judged the certainty of the evidence as low for end-of-treatment effects. Music-based interventions may have little to no effect on social behaviour at the end of treatment (SMD 0.22, 95% CI −0.14 to 0.57; Analysis 1.6; a positive result favours music-based therapeutic interventions). There was no heterogeneity ($I^2 = 0\%$).

Cognition

Seven studies (353 participants) contributed to the end-of-treatment effect analysis. Outcome measures used in the analyses were (translated or original) versions of the MMSE and the SIB. We used the MMSE data if these were available. Other measures used were cognition measures such as Prose Memory tests, the FAS-Test (Controlled-Oral-Word-Association Test) and the Alzheimer's Disease Assessment Scale Cognitive subscale (ADAS-cog).

We judged the certainty of the evidence as low. The analysis suggested that music-based interventions may have little or no effect on cognition at the end of treatment (SMD 0.19, 95% CI −0.02 to 0.41; Analysis 1.7; a positive result favours music-based therapeutic interventions). There was no heterogeneity (I^2 = 0%).

Long-term results

Emotional well-being including quality of life

Only three studies with 86 participants contributed to this outcome at long-term follow-up.

We judged the certainty of the evidence as very low. The SMD did not differ much from the SMD at the end of treatment, but results were inconsistent and imprecision was greater, so we are uncertain of the direction and magnitude of the effect (SMD 0.17, 95% CI −0.80 to 1.14). Heterogeneity was substantial (I^2 = 61%) (Analysis 2.1).

Mood disturbance or negative affect: depression

Five studies with 276 participants contributed to this outcome.

We judged the certainty of the evidence as low. There was no clear evidence of a reduction in the longer term, with a smaller estimate than at the end of treatment and a CI including no effect (SMD −0.07, 95% CI −0.31 to 0.18; Analysis 2.2). There was no heterogeneity (I^2 = 0%).

Mood disturbance or negative affect: anxiety

Three studies with 141 participants contributed to this outcome.

We judged the certainty of the evidence as low. Therefore, we have little confidence in the results. Music-based therapeutic interventions may have little to no effect on anxiety in the longer term (SMD −0.06, 95% CI −0.48 to 0.37; Analysis 2.3; a negative value favours music-based therapeutic interventions). Heterogeneity was low to moderate for longer-term effects (I^2 = 35%).

Behavioural problems: agitation or aggression

Four studies with 241 participants contributed to the long-term effect analysis.

We judged the certainty of the evidence as low. We found no clear evidence of an effect on agitation or aggression in the long term (SMD −0.17, 95% CI −0.42 to 0.09; Analysis 2.4; a negative result favours music-based therapeutic interventions). There was no heterogeneity (I^2 = 0%).

Behavioural problems: overall

Six studies with 245 participants contributed to the analysis of longer-term effects.

We judged the certainty of the evidence as low. There was no clear evidence of a long-term effect (SMD −0.19, 95% CI −0.52 to 0.14; Analysis 2.5). Heterogeneity was moderate (I^2 = 31%).

Social behaviour

We found no studies reporting on the long-term effect on this outcome.

Cognition

Two studies (146 participants) assessed long-term effects on cognition.

We judged the certainty of the evidence as low. The analysis suggested that music-based interventions may have little or no effect on cognition in the long term (SMD 0.09, 95% CI −0.24 to 0.41; Analysis 2.6; a positive result favours music-based therapeutic interventions). There was no heterogeneity (I^2 = 0%).

Music-based therapeutic interventions compared to other interventions

For this comparison, results at the end of treatment are summarised in Summary of findings 3 and longer-term effects in Summary of findings 4.

Seventeen studies addressed this comparison and contributed to the meta-analyses (Baker 2022; Cho 2016; Cooke 2010; Gómez-Gallego 2021; Guétin 2009; Liesk 2015; Liu 2021; Lyu 2014; Narme 2012-study 1; Narme 2012-study 2; Narme 2014; Raglio 2015; Reschke-Hernández 2019; Sakamoto 2013; Thornley 2016; Vink 2013; Werner 2017).

Results at end of treatment

Emotional well-being including quality of life

We included nine studies with 298 participants in the analysis of end-of-treatment scores for the critically important outcome of emotional well-being and quality of life. Most studies used a validated quality-of-life or well-being measure for more direct observation; the Dementia Quality of Life (DQOL) (Cooke 2010); a German translation of the Dementia Quality of Life Instrument (DEMQOL) (Liesk 2015); the Cornell-Brown Scale for Quality of Life in Dementia (CBS-QoL) – although it was unclear if this was a validated translated version (Raglio 2015); the Quality of Life-Alzheimer's Disease (QOL-AD) (Baker 2022; Cho 2016) and the QUALIDEM (Reschke-Hernández 2019). In the three studies conducted by Narme and colleagues, emotional well-being referred to counts of positive and negative facial expressions as assessed from the first minutes of filmed interviews (Narme 2012-study 1; Narme 2012-study 2; Narme 2014).

The certainty of evidence for the effects of music-based interventions on emotional well-being and quality of life at the end of treatment was low. We found little or no clear evidence of a positive effect favouring the intervention at the end of treatment (SMD 0.20, 95% CI −0.09 to 0.49; Analysis 3.1). Heterogeneity was low (I^2 = 26%). There was no blinding of outcome assessment in four studies. Narme 2012-study 2 is a clear outlier, with a large SMD compared to the others, but the study was very small.

Mood disturbance or negative affect: depression

Ten studies contributed 359 participants to the analysis on end-of-treatment effect. Depression or depressive symptoms were measured with (translated versions of) the Geriatric Depression Scale (GDS), the Cornell Scale for Depression in Dementia, the Montgomery-Åsberg Depression Rating Scale, MADRS), the Alzheimer's Disease and Related Dementias Mood Scale (AD-RD), or with a subscale of the Behavioural Pathology in Alzheimer's Disease (BEHAVE-AD) or the NPI.

We judged the certainty of the evidence as low. We found no clear evidence of an effect on depressive symptoms at the end of treatment (SMD −0.14, 95% CI −0.36 to 0.08; Analysis 3.2; note that for this outcome, a negative SMD favours the intervention group). There was no heterogeneity ($I^2 = 0\%$). We did not find clear evidence of publication bias (Appendix 4).

Mood disturbance or negative affect: anxiety

The other mood item we considered was anxiety. For this outcome, at the end of treatment, we included 10 studies with 291 participants. A variety of (translated) outcome measures were used; Rating Anxiety in Dementia Scale (RAID), STAI-A, Hamilton Anxiety Scale, and subscale scores of the BEHAVE-AD and NPI. Heterogeneity was substantial for end-of-treatment effects ($I^2 = 73\%$).

We judged the certainty of the evidence as very low for end-of-treatment effects. Therefore, we can have very little confidence in the results. Anxiety may be considerably lower in the music intervention group at the end of treatment (SMD −0.75, 95% CI −1.27 to −0.24; Analysis 3.3; a negative result favours the intervention group); however, the evidence is very uncertain. We did not find clear evidence of publication bias (Appendix 4).

Behavioural problems: agitation or aggression

Six studies with 168 participants contributed to the end-of-treatment effect analysis. Outcome measures used for agitation were (sometimes translated) versions of the CMAI and the agitation subscale of the NPI; and for aggression, the aggressiveness subscale of the BEHAVE-AD and counts of observed aggressive behaviour.

We judged the certainty of the evidence as low. We found no clear evidence of an effect on agitation or aggression at the end of treatment (SMD 0.01, 95% CI −0.31 to 0.32; Analysis 3.4; a negative result favours the intervention group). There was no heterogeneity ($I^2 = 0\%$).

Behavioural problems: overall

Eight studies with 292 participants contributed to the end-of-treatment effect analysis. Outcome measures were (sometimes translated) versions of the BEHAVE-AD and NPI.

We judged the certainty of the evidence as low. We found no clear evidence of an effect of music-based therapeutic interventions on problematic behaviour overall at the end of treatment (SMD −0.08, 95% CI −0.33 to 0.17); Analysis 3.5; a negative result favours the intervention group). There was no heterogeneity ($I^2 = 0\%$).

Social behaviour

Four studies with 84 participants contributed to the end-of-treatment effect analysis. For the three Narme studies, the outcome was the content of conversation (positive versus negative expressions when interviewed about current feelings and personal history). Reschke-Hernández 2019 reported social engagement observed with the MPES.

We judged the certainty of the evidence as low. At the end of treatment, music-based therapeutic interventions may be more effective than the other activities for social behaviour (SMD 0.52, 95% CI 0.08 to 0.96; Analysis 3.6). There was no heterogeneity ($I^2 = 0\%$).

Cognition

Five studies contributed 147 participants to the end-of-treatment effect analysis. Outcome measures used in the analyses were (sometimes translated) versions of the MMSE and the SIB. We used the MMSE data if these were available. Other measures used were cognition measures such as Prose Memory tests, the FAS-Test (Controlled-Oral-Word-Association Test) and the Alzheimer's Disease Assessment Scale Cognitive subscale (ADAS-cog).

We judged the certainty of the evidence as low. Results suggested that music-based interventions may have little or no effect on cognition at the end of treatment (SMD 0.12, 95% CI −0.21 to 0.45; Analysis 3.7; a positive result favours the intervention group). There was no heterogeneity ($I^2 = 0\%$).

Long-term results

Emotional well-being including quality of life

Four studies (130 participants) contributed to the analysis of long-term effects.

We judged the certainty of the evidence as low. The SMD did not differ much from the SMD at the end of treatment, but the imprecision was greater, so there is little evidence of any difference between groups. (SMD 0.10, 95% CI −0.29 to 0.49; Analysis 4.1). Heterogeneity was not important ($I^2 = 7\%$).

Mood disturbance or negative affect: depression

Four studies contributed 174 participants to the analysis of long-term effects.

We judged the certainty of the evidence as low. There was no evidence of a reduction in the longer term, with a smaller estimate than at the end of treatment and a CI including no effect (SMD −0.07, 95% CI −0.39 to 0.25; Analysis 4.2). There was no heterogeneity ($I^2 = 0\%$).

Mood disturbance or negative affect: anxiety

Four studies with 124 participants contributed to the analysis of long-term effects.

We judged the certainty of the evidence as low. Therefore, we can have little confidence in the results. Music-based therapeutic interventions may have little to no effect on anxiety (SMD −0.53, 95% CI −1.31 to 0.25; Analysis 4.3, a negative result favours the intervention group). Heterogeneity was substantial ($I^2 = 74\%$).

Behavioural problems: agitation or aggression

Two studies with 89 participants contributed to the analysis of long-term effects.

We judged the certainty of the evidence as very low. We are unsure about the effect of music-based therapeutic interventions on agitation or aggression in the long term (SMD 0.10, 95% CI −0.66 to 0.86; Analysis 4.4). Heterogeneity was substantial (I^2 = 67%).

Behavioural problems: overall

Four studies with 197 participants contributed to the analysis of long-term effects.

We judged the certainty of the evidence as low. We found no clear evidence of an effect of music-based therapeutic interventions on problematic behaviour overall (SMD −0.09, 95% CI −0.39 to 0.22; Analysis 4.5). There was no heterogeneity (I^2 = 0%).

Social behaviour

Two studies with 48 participants contributed to the analysis of long-term effects.

We judged the certainty of the evidence as very low. We are unsure about the effect of music-based therapeutic interventions on social behaviour (SMD 0.53, 95% CI −0.53 to 1.60; Analysis 4.6). There was moderate heterogeneity (I^2 = 54%).

Cognition

Only one study with 47 participants assessed long-term effects.

We judged the overall certainty of the evidence as very low. Results suggested that music-based interventions may have little or no effect on cognition in the long term (SMD 0.04, 95% CI −0.56 to 0.64; Analysis 4.7), but the evidence is very uncertain.

Potential adverse effects

Here we summarise the findings on adverse effects for both comparisons (i.e. the studies comparing music-based therapeutic interventions to usual care and those comparing the music-based therapeutic interventions to an active intervention). Some of the most recent trials explicitly reported adverse effects. In the Baker 2022 and Giovagnoli 2018 studies, documenting adverse events was planned in the methods, while results were also explicit in Werner 2017. Baker 2022 reported that two residents who left the group music therapy session became physically aggressive or demanded that the therapist stop, which the study authors called "related non-serious adverse events". Giovagnoli 2018 included assessment of somnolence, insomnia and depression and found these were the most frequently occurring adverse events in the intervention group (an adverse event in a total of 21.8% participants) and the control group (in 50.0% of participants). Werner 2017 found a worsening in depressive symptoms of more than 10 points on the MADRS in four participants of the recreational group singing control group; Werner 2017 found no other adverse events. Reschke-Hernández 2019 did not include adverse events explicitly but reported frequencies of participants trying to leave as part of observations of disengagement, which Reschke-Hernández 2019 defined as a person leaving or trying to leave an activity or behaving in a socially inappropriate manner. They found that participants tried to leave verbal sessions more often (10 people, 24 times) than music therapy sessions (3 people, 6 times). Liesk 2015 did not report adverse effects but reported dropout in the music-based therapeutic intervention group due to "problems in the group", though it is unclear if this was related to the intervention. Hsu 2015, Ridder 2013 and Thornley 2016 reported that no adverse effects were observed in the intervention group (Hsu 2015; Ridder 2013) or in the active control group (Thornley 2016), but adverse effects was not an outcome nor was it defined in the methods sections.

Subgroup and sensitivity analyses

We carried out a single subgroup analysis as the minimum number of five studies per subgroup was only met for the outcome of behaviour problems: agitation or aggression, in the comparison of music-based therapeutic interventions versus usual care at end of treatment. The test for subgroup differences showed that, compared to usual care, individual therapy and group therapy differed in effect (P = 0.02). Individual therapy showed evidence of a positive effect (SMD −0.31, 95% CI −0.60 to −0.02), but group therapy did not show evidence of an effect (SMD 0.13, 95% CI −0.12 to 0.37). The heterogeneity (I^2) in the total group was 32%; in the subgroup receiving individual therapy, it was 0%, and in the subgroup receiving group therapy, it was 11% (Analysis 5.1).

In the sensitivity analyses with analyses restricted to studies where the intervention was definitely or possibly delivered by a qualified music therapist, most effect estimates were similar. For some of the outcomes, results no longer showed evidence of an effect, probably due to the small number of studies. The presence of a potential conflict of interest did not affect the results (Appendix 5).

Masking therapists and participants was impossible; therefore, we judged all studies to be at high risk of bias for this domain in our risk of bias assessment. In the sensitivity analyses restricted to studies without a judgement of high risk of bias for any other domain, results were roughly similar. However, for the music-based therapeutic intervention versus usual care comparison at the end of treatment, the pooled effect on overall behavioural problems was reduced from −0.31 (95% CI −0.60 to −0.02; 10 studies) to −0.17 (95% CI −0.46 to 0.12; 6 studies) (Appendix 6; Analysis 1.5). For music-based therapeutic intervention versus other activities at the end of treatment, the effect on anxiety increased from −0.75 (95% CI −1.27 to −0.24; 10 studies) to −1.08 (95% CI −1.95 to −0.22; 5 studies) (Appendix 6; Analysis 3.3).

D I S C U S S I O N

Summary of main results

The aim of this review was to evaluate the effects of music-based therapeutic interventions on a range of outcomes relevant for people with dementia. The specific focus was to assess whether such interventions could improve emotional well-being including quality of life, mood disturbance or negative affect, behavioural problems, social behaviour and cognition, and if this could be done without causing adverse effects.

The review included 30 studies with 1720 randomised participants, and we performed meta-analyses with data from 28 studies (1366 participants) at the end of treatment and from 10 studies (560 participants) in the longer term (four weeks to six months after treatment ended).

For studies comparing music-based therapeutic interventions to usual care, we found moderate-certainty evidence that, at the end of treatment, music-based therapeutic interventions probably improve depressive symptoms but not agitation or aggression. We found low-certainty evidence that they may improve overall behavioural problems, while low or very low certainty evidence showed that they may not improve emotional well-being, anxiety, social behaviour or cognition.

Compared to other interventions, we found that, at the end of treatment, music-based therapeutic interventions may be more effective than the other activities for social behaviour and anxiety; however, these findings are based on low-certainty and very-low certainty evidence, respectively. For all other outcomes, the low-certainty evidence showed no effect.

For all outcomes measured in the long term (i.e. four weeks to six months after the end of treatment), we found low-certainty and very low-certainty evidence of no effect.

We could not quantify adverse effects as the reporting was inconsistent.

A subgroup analysis for individual versus group therapy suggested that for agitation or aggression, individual therapy may be effective at the end of treatment but group therapy may not.

In sensitivity analyses with analyses restricted to studies where the intervention was definitely or possibly delivered by a qualified music therapist, most effect estimates were similar. For some of the outcomes, confidence intervals widened, due to the small number of studies.

No patterns or differences were found in the other sensitivity analyses selecting studies based on funding source and risk of bias.

Overall completeness and applicability of evidence

We searched for studies reported in various languages, and we also included articles published in languages other than English.

We found no studies conducted in people's homes or a community setting only, although the participants in Giovagnoli 2018 were included after hospitalisation and may have resided in non-institutional settings. Institutional settings may differ internationally but were mostly residential long-term care settings or long-stay geriatric wards in or part of a hospital.

The evidence in this review relates to therapeutic effects of music-based therapeutic interventions after at least five sessions of the intervention. We excluded some group interventions that involved music but where music was not the main or only therapeutic element, where family caregivers provided the intervention or received the therapeutic intervention themselves, or where there was no interaction with individuals during the session. We therefore excluded listening interventions, and most interventions involved a mixture of active and passive elements, which reduces heterogeneity or variability in the range of possible music-based interventions.

Many studies used validated outcome measures for cognition (e.g. the MMSE (Folstein 1975)) and for behaviour (e.g. the NPI (Cummings 1994) or BEHAVE-AD (Reisberg 1987)), two widely used measures that are recommended because of favourable

psychometric properties (Cummings 2020; Jeon 2011)). We included subscales of the behavioural scales as outcome measures. However, there was less evidence for the validity of subscales compared to total scores (Lai 2014). Only six studies assessed social behaviour as an outcome.

Some included studies selected people with agitated behaviour before the intervention, or people who were more likely to be interested in music-based interventions. In contrast, there were studies in which people with musical knowledge were excluded (Raglio 2010b). Dropout was mostly for health reasons, including hospitalisation, illness or mortality. Dropout due to lack of interest was reported for particular control activities: cognitive stimulation programme (Liesk 2015); television watching (Cho 2016); and verbal discussion sessions (Reschke-Hernández 2019), and dropout due to "problems in the group" was reported in a music intervention group (Liesk 2015). However, few other studies reported any unfavourable effects of the music-based interventions, with only three newer studies explicitly reporting on the outcome of adverse effects for all groups. We do not know if there were any unreported adverse effects such as a sore throat after singing or cases of distress that were specifically related to the therapy. It is conceivable that distress or agitated behaviour occurs in a group setting, which would be particularly difficult to manage in the case of group therapy. This might explain that effects have been close to zero for the agitation or aggression outcome, with lower SMDs than for other outcomes consistently over the three updates of this review. We will not know until the occurrence or absence of such potentially related adverse effects are described in full. We also do not know whether participants have been included based on a subjective judgement of whether they would accept the intervention, and whether, without selective inclusion, some people with dementia might experience disadvantages of the intervention. It is possible that finding a beneficial effect in these studies depends on participants having significant problems at baseline (i.e. being selected as in need of treatment for specific problems) and hence having substantial room for improvement. Some specific subgroups might benefit from music-based therapeutic interventions more than others.

Certainty of the evidence

The certainty of the evidence (GRADE assessments) was moderate for depression and for agitation or aggression at the end of treatment in studies comparing music-based therapeutic interventions with usual care. For all other outcomes, at both time points and for both comparisons, it was low or very low. We downgraded all outcomes because of the risk of bias. Outcome assessment was often unblinded or blinding was unclear, and this may have inflated effects. Imprecision (wide confidence intervals) and inconsistency (effects in opposite directions) were other domains for which we downgraded.

The quality of reporting was sometimes poor, which resulted in uncertainty about the exact methodological quality of the included studies and the evidence for effects. The reporting of the intervention may be improved by using reporting guidelines for intervention description and replication.

Most of the studies had small sample sizes. Few studies reported on fidelity of the implementation of the music-based intervention and other activities, or on other elements of a process evaluation.

![Cochrane Library logo] Trusted evidence.
Informed decisions.
Better health.

Cochrane Database of Systematic Reviews

Implementation fidelity is often defined as the degree to which an intervention or programme is delivered as intended (Carroll 2007); and in music therapy trials specifically, treatment fidelity refers to "methodological strategies used to monitor the delivery of the music therapy intervention as described in the treatment manual" (Bradt 2012). Treatment fidelity includes adherence to an intervention, exposure or dose, quality of delivery, participant responsiveness and programme differentiation to identify essential components of the intervention (Carroll 2007), and therefore includes, but is not limited to, participant (or staff) adherence and responsiveness. Amano 2022 reviewed implementation strategies used in RCTs and other studies. Amano 2022 was unable to make a recommendation on the effectiveness of the implementation strategies, but suggested implementation studies should report on aspects such as acceptability, appropriateness, feasibility, adoption, sustainability and costs of music-based therapeutic interventions for people with dementia.

Notably, Hsu 2015 and Gómez-Gallego 2021 were outliers for one or two outcomes (Analysis 1.1; Analysis 1.5). Both studies showed larger effects than the other studies, contributing to heterogeneity in the analyses. Both were cluster-randomised trials at high risk of bias with small numbers of participants, which were further reduced for the analyses in order to correct for the cluster design.

There may be publication bias through selective outcome reporting in published study reports. Although more included studies had a registered protocol than previously, overall, there is still a lack of protocol registration, and we found inconsistencies in the reporting of outcome measures in two studies (inconsistency across multiple reports for Cooke 2010; inconsistency compared with trial registration in Hsu 2015). Moreover, although most of our meta-analyses found no evidence of an effect, 26 of 29 studies reported at least one significant effect (all except Liesk 2015, Raglio 2015 and Thornley 2016; while one study, Prieto Alvarez 2022, refrained from testing because the study was testing feasibility). For some studies, this included outcomes beyond the scope of this review, such as heart rate, but it could indicate selective reporting of significant findings or analytic methods that resulted in significant findings. One of the funnel plots we generated suggested publication bias; however, we did not downgrade the certainty of the evidence as the finding was based on one outlier.

There may be a financial conflict of interest if a study is funded by a source interested in the outcomes, or an intellectual conflict of interest when the intervention is performed by a music therapist who also is the author or a co-author of the study, but there were insufficient data to examine possible effects of conflicts of interest.

Potential biases in the review process

Although we did an extensive literature search in the most commonly used and relevant databases, and we thoroughly handsearched music therapy journals, it is still possible that we have missed one or more RCTs. We critically re-reviewed previously included studies as we had slightly modified the inclusion criteria. New studies using audio-recorded music often resemble receptive music-based therapeutic interventions, especially if they involve some degree of interaction with a facilitator, which complicates assessment of the eligibility of such studies. We recognise that "non-therapist led music-based interventions" might effectively improve outcomes in community-dwelling persons with dementia (Hofbauer 2022). In future updates, we may revise the inclusion

criteria again and consider limiting inclusion to active or mixed music-based interventions.

Further, although heterogeneity in effects on emotional well-being and quality of life was limited, we would suggest that studies include subgroup analyses for measurement of behaviour with direct observation (Hsu 2015; Narme 2012-study 1; Narme 2012-study 2; Narme 2014) versus measurement in retrospect. Direct effects during sessions were usually not included other than for social behaviour and behaviour during bathing.

We combined agitation and aggression in meta-analyses because this is consistent with the definition given by the International Psychogeriatric Association (Cummings 2015); and these items are also combined in the widely used CMAI (Cohen-Mansfield 1986). Some have raised concerns about conceptual issues such as overlap of a broad definition of agitation with resistance to care (Volicer 2007). Despite clinical heterogeneity, we pooled the data in meta-analyses. Nevertheless, we did not find excessive statistical heterogeneity, other than for the outcome of anxiety.

Agreements and disagreements with other studies or reviews

In the previous version of our review (Van der Steen 2018), we found an SMD of 0.27 for depression at the end of treatment. In this update, we found a similar effect of music-based therapeutic interventions compared to usual care (SMD 0.23) but little or no effect (SMD 0.14) compared to other activities. Amongst the five newly added studies with data on depression (Baker 2022; Giovagnoli 2018; Gómez-Gallego 2021; Liu 2021; Werner 2017), there were studies that found a result that favoured the active control group.

Inclusion criteria for music-based therapeutic interventions were more exclusive in this version of the review than in previous versions. We excluded studies on interventions termed music therapy when there was insufficient indication that the intervention had therapeutic goals and its delivery required skill, or when the intervention was combined with other types of interventions. In contrast, we included studies when the profession or training of the therapist was unclear if criteria for therapy and skill were met. The effects we found may be more modest than in many other reviews, but the sensitivity analyses indicated this is probably not explained by allowing inclusion of studies not or not clearly provided by a professional music therapist. A review by Lam 2020 included a wider range of studies (n = 82) and performed a narrative synthesis, vote counting statistically significant study results. The review reported "significant" effects of music therapy on mood, apathy and verbal fluency, while effects on agitation and quality of life were reported as "ambiguous" and on cognition as "insignificant".

One review and meta-analysis on effects of music therapy on behavioural and psychological symptoms of dementia found larger SMDs for behavioural problems overall (SMD –0.49, 95% CI –0.82 to –0.17) and for anxiety (SMD –0.64, 95% CI –1.05 to –0.24) compared with our findings (Ueda 2013), although we found a similarly large effect on anxiety when compared to other activities. However, that review included non-randomised trials, cohort studies and studies that we excluded because they did not meet our criteria for therapeutic interventions. They found an even larger effect for studies that lasted three months or longer (SMD –0.93, 95% CI

–1.72 to –0.13), a subgroup that we did not analyse separately. Reviews by Zhang 2023 and Ting 2023 also found large SMDs for reduction in anxiety (SMD in both reviews -0.67) and Zhang 2023 also found a large SMD for reduction in depression (-0.86). These reviews found larger effects for interventions with a longer duration but low frequency when compared to high-frequency short-duration interventions. They included pre-post test studies, which we excluded (e.g. Tamplin 2018).

The review by Chang 2015 included 10 studies, three of which we excluded from our review: Raglio 2008, which we excluded after inclusion in an earlier version of our review because after re-evaluation, we judged it to be quasi-randomised; Sung 2006, which after re-evaluation did not meet our criteria for a music-based therapeutic intervention (it was music with movement); and Janata 2012, which we excluded because streaming music also did not meet our criteria for a therapeutic intervention. Chang 2015 included studies that compared music therapy with usual care, excluding other activities other than reading sessions as the comparator. Like us, Chang 2015 found substantial heterogeneity in the analysis of anxiety. Effect sizes for cognition were smaller than for anxiety in both reviews, and effects on depression were not significant despite an effect size of -0.39 found in that review. Chang 2015 also found a significant effect on 'disruptive behaviours.' We did not find an effect on agitation or aggression, but we found an effect on overall behavioural problems. The scales used to assess behavioural problems, however, also included mood items.

Li 2019 reviewed RCTs for effect on depression and found an effect, also suggesting that "music therapy without a music therapist involved did not significantly reduce depression at any time." Because we received unpublished data from authors, we included more studies on mood outcomes than the seven included in Li 2019, two of which we excluded because of a quasi-experimental design or an intervention with streamed music with no contact with a facilitator.

One review by Zhang 2017 included non-randomised studies and studies that we excluded because of insufficient therapeutic-based goals and their methods and findings differed in a number of other ways. Their subgroup analyses for the effect on 'disruptive behaviour' (overall behavioural scales and agitation) suggested a higher SMD for non-randomised studies (–1.02 for non-randomised studies versus –0.65 (reported in the text) or –0.52 (reported in the table) for parallel RCTs). They found a larger SMD for disruptive behaviour (–0.42, 95% CI –0.74 to –0.11, compared to –0.28 for overall behavioural problems and –0.03 for agitation or aggression in our work). Compared to our review, they found a similar or somewhat larger SMD for cognition (SMD 0.20, 95% CI –0.09 to 0.49). Zhang 2017 performed different analyses, probably comparing scores before and after the intervention to calculate an SMD with a general check of whether there were baseline differences. This may explain the different SMDs they also reported for individual studies, and why the quality assessments of the same included studies rarely corresponded with ours. For example, Svansdottir 2006 was an outlier for effect on behaviour in Zhang 2017 (SMD –3.88) compared with an SMD of –0.06 for end-of-treatment score in our work. Also, in this study, Zhang 2017 assigned points for quality because of therapist blinding, whereas we rated all studies at high risk for performance bias (in view of standardised methods to allow for comparison of very different interventions and situations), and, moreover, Svansdottir 2006 also disclosed that the first author

"conducted the music therapy." Zhang 2017 judged all studies to be of acceptable quality, even those with a total score of 3 (reported in supplemental table) or higher than 4 (reported in text) on a 0 to 10 scale where one of the items was the random allocation. Finally, their secondary outcomes (depression, anxiety and quality of life) were prioritised in our review because of the importance for the person with dementia him/herself. In a more recent review on the effects on anxiety, Zhang 2023 did not downgrade the quality of the evidence of any of the included studies for lack of blinding of participants and personnel or for lack of blinding of outcome assessment. This was the same in a review by Ting 2024 and the grading of most studies in Ting 2023.

Multiple other reviews have summarised results and concluded, often without meta-analysis, that a music-based therapeutic intervention or music therapy can be beneficial. Some focused on specific outcomes, such as behavioural and psychological symptoms of dementia (e.g. Raglio 2012), or covered different types of outcomes, such as physiological outcomes (e.g. McDermott 2013, who also noted a lack of evidence on long-term effects). Petrovsky 2015 focused on the effects on anxiety and depression in people with mild dementia, but included studies with participants who had varying severity of dementia as long as it was not limited to severe dementia. They concluded, based on 10 studies, including some with a pre–post test design, that the evidence was inconclusive. We were able to include more RCTs in our review because trial authors provided data about agitation and mood items in overall behavioural scales. Ing-Randolph 2015 reviewed the effects of group music interventions, including music therapy, on anxiety; they found that music interventions reduced anxiety in seven of eight included studies.

The clinical importance of the effect of music-based interventions on depression is somewhat uncertain, with moderate certainty for an SMD of 0.23 when compared to usual care, and low certainty for an SMD of 0.14 when compared to other activities. A variety of scales was used, but there was no important heterogeneity in effects across the studies; in fact, for end-of-treatment outcomes, heterogeneity was substantial ($I^2 > 60\%$) only for anxiety in the comparison of music-based therapeutic intervention versus other activities. The SMDs for the effect on anxiety differed substantially between the comparison with usual care (SMD of 0.15) and the comparison with other activities (SMD 0.75). For depression, the difference was smaller (SMD 0.23 and 0.14, respectively). In terms of a comparison with medication for depression, a meta-analysis by Nelson 2011 found a pooled estimate of antidepressant medication effects in people with dementia of SMD 0.29, favouring antidepressant use (95% CI 0.02 to 0.60; six trials), while a meta-analysis by Sepehry 2012 of five trials of selective serotonin reuptake inhibitors (SSRIs) for people with Alzheimer's disease indicated a small to no effect (effect size estimates favouring medication of 0.06 to 0.10 depending on outcome scales). There may have been fewer adverse effects of music-based therapeutic interventions compared with medication, although this is uncertain because these are rarely reported for music-based therapeutic interventions.

AUTHORS' CONCLUSIONS

Implications for practice

Compared to usual care, music-based therapeutic interventions may be used for people with dementia as they probably improve

depression and may improve overall behavioural problems at the end of treatment. However, they probably do not improve agitation or aggression, and may not improve anxiety, social behaviour or cognition. Furthermore, the evidence is very uncertain about the effects of music-based therapeutic interventions on emotional well-being including quality of life.

Compared to other activities, music-based therapeutic interventions may improve social behaviour, but the effect on anxiety is very uncertain, and it may result in little to no difference to other outcomes.

It is not clear whether the effects observed persist beyond the intervention period. It is also not clear whether effects change when music-based therapeutic interventions are offered on a continuous basis.

It is also unclear whether the interventions cause any adverse effects, as these were inconsistently measured or recorded. However, no serious adverse events were reported.

Implications for research

Guidelines for the design and implementation of randomised controlled trials (RCTs) of music therapy are available (Bradt 2012). For dementia, more well-conducted studies are needed to establish more precisely the effects of music therapy and music-based therapeutic interventions in the treatment of people with dementia, including effects on positive outcomes such as emotional well-being, quality of life and social behaviour. Negative outcomes (adverse effects) should not be ignored, and their monitoring should be planned as part of the study methods. International consensus is needed on what the potential adverse effects are for music-based therapeutic interventions. How any effect is mediated may be an area for examination; for example, increased social behaviour may be mediated through decreased behavioural problems (Waters 2022). Outcomes could also include outcomes that are less obvious but compromise quality of life, such as apathy, which is highly prevalent and often highly persistent over the course of dementia (dementia or cognitive impairment, Van der Linde 2016; Alzheimer's disease, Zhao 2016). Arguably, apathy is a more important outcome than cognition, in particular for people with dementia in later stages of the disease for whom music-based therapeutic interventions are still suitable. Outcomes such as pain and discomfort have been used for testing the effects of music therapy at the end of life, mostly amongst people with cancer (McConnell 2016); these are also important outcomes for people with dementia. General behavioural scales (which include mood items, agitation and items on hallucinations, euphoria, etc.) might be rather broad for use as outcome scales for effects of music therapy.

Future studies should follow the CONSORT guidelines for reporting of randomised trials. In particular, they should use adequate methods of randomisation with adequate concealment of allocation of the participants to the parallel treatment groups, blind the outcome assessors to treatment allocation and report the method used for this, and be of sufficient duration to assess persistence of effects after the end of treatment. Blinding of participants is difficult but not impossible, especially with active control groups, if the participants are unaware of the hypothesis of the study and which intervention is considered the active Iintervention (Bradt 2012). We discourage the use of

cross-over designs because possible long-term effects of music-based interventions may carry over into the control phase. Study protocols should be registered and primary and secondary outcomes should be reported as prespecified. Reporting of effects should preferably include mean differences and standard deviations of differences between baseline and follow-up, or effect sizes, which, up till now, only a small proportion of studies have reported. Funding sources should be reported and any potential conflict of interest should be considered and disclosed, such as an interest in finding favourable effects of the therapy. This also includes cases where the therapist delivering the intervention is an author of the published report of the study.

More research is needed to differentiate between various therapeutic approaches using music; for example, to examine whether there is a difference between receptive and active approaches (De Witte 2024). Studies carried out in hospital wards should clarify the setting (i.e. long-stay versus acute care). We found that the interventions are usually described in reasonable detail, although a recent mapping by Lepping 2024 found poor reporting of some items. However, we note that full reproducibility as a scripted intervention with prescribed music is not desirable nor compatible with the delivery of a person-centred therapeutic intervention that is responsive to individual needs (e.g. through improvising), which characterises music-based therapeutic interventions. This is one aspect of intervention complexity (Higgins 2019), but other aspects do not involve complexity, such as a single discipline providing the intervention that takes place in the context of scheduled sessions only. With more studies becoming available in the future, we may be able to examine how response relates to duration of individual sessions (noting that any dose–response relationships may not be linear due to participants' difficulties sustaining concentration or the risk of overstimulation with longer sessions) and the number of sessions, taking into account that some outcomes are assessed directly after or during a therapy session and therefore include immediate effects. It is important to establish whether pre-existing problematic or challenging behaviour moderates the effects. More data are needed for subgroup analyses, but other types of reviews, such as realist reviews, may indicate how interventions may work and for whom. Further research is also required to compare music-based therapeutic interventions in which music is the main or only therapeutic element, to other group activities involving music. In the studies in this review, the professional background of the therapist was sometimes unclear, or there was no information about the training of the music therapists or their experience of delivering music-based therapeutic interventions specifically to people with dementia. It is important to provide details of who delivers the intervention in order to facilitate classification of interventions as music therapy delivered by a qualified, trained and experienced music therapist, other music-based therapeutic interventions and other interventions involving music. However, targeted studies may be more appropriate to evaluate the effects of training because subgroup analyses risk confounding if, for example, qualified therapists see people with more complex problems. Further studies may also include economic analyses, and focus on effects in special groups, such as people with early-onset dementia, or in different settings, including community settings with more people with dementia in its early stages.

ACKNOWLEDGEMENTS

Editorial and peer-reviewer contributions

REFERENCES

References to studies included in this review

Baker 2022 *{published data only (unpublished sought but not used)}* **10.1016/S2666-7568(22)00027-7.**

Baker FA, Stretton-Smith PA, Sousa TV, Clark I, Cotton A, Gold C, et al. Resource assessment in trials undertaken in residential care homes: experiences from the Australian MIDDEL cluster randomised controlled trial research team. *Contemporary Clinical Trials Communications* 2020;**20**:100675. [DOI: 10.1016/j.conctc.2020.100675]

ACTRN12618000156280. Music Interventions for Dementia and Depression in Elderly care: International Multicentre cluster randomised control trial. Australian New Zealand Clinical Trials Registry 2018. [TRIAL REGISTER: ACTRN12618000156280]

* Baker FA, Lee Y-EC, Sousa TV, Stretton-Smith PA, Tamplin J, Sveinsdottir V, et al. Clinical effectiveness of music interventions for dementia and depression in elderly care (MIDDEL): Australian cohort of an international pragmatic cluster-randomised controlled trial. *Lancet. Healthy Longevity* 2022;**3**:e153–65. [DOI: 10.1016/S2666-7568(22)00027-7]

Baker FA, Tamplin J, Clark IN, Lee YC, Geretsegger M, Gold C. Treatment fidelity in a music therapy multi-site cluster randomized controlled trial for people living with dementia: The MIDDEL project intervention fidelity protocol. *Journal of Music Therapy* 2019;**56**(2):125-48. [DOI: 10.1093/jmt/thy023]

Gold C, Eickholt J, Assmus J, Stige B, Wake JD, Baker FA, et al. Music Interventions for Dementia and Depression in ELderly care (MIDDEL): protocol and statistical analysis plan for a multinational cluster-randomised trial. *BMJ Open* 2019;**9**(3):e023436. [DOI: 10.1136/bmjopen-2018-023436]

Janus SIM, Vink AC, Ridder HM, Geretsegger M, Stige B, Gold C, et al. Developing consensus description of group music therapy characteristics for persons with dementia. *Nordic Journal of Music Therapy* 2021;**30**(1):24-40. [DOI: 10.1080/08098131.2020.1779790]

NCT03496675. Music Interventions for Dementia and Depression in Elderly Care (MIDDEL). https://clinicaltrials.gov/study/NCT03496675 (first posted 12 April 2018).

Rasing NL, Janus SIM, Kreutz G, Sveinsdottir V, Gold C, Nater UM, et al. The Impact of music on stress biomarkers: protocol of a substudy of the cluster-randomized controlled trial music interventions for dementia and depression in elderly care (MIDDEL). *Brain Sciences* 2022;**12**(4):485. [DOI: 10.3390/brainsci12040485] [PMID: 35448016]

Ceccato 2012 *{published data only}* **10.1177/1533317512452038**

Ceccato E, Vigato G, Bonetto C, Bevilacqua A, Pizziolo P, Crociani S, et al. STAM protocol in dementia: a multicenter, single-blind, randomized, and controlled trial. *American Journal of Alzheimer's Disease and Other Dementias* 2012;**27**(5):301-10.

Cho 2016 *{published data only}*

Cho HK. The effects of group singing on quality of life and affect of people with dementia: a randomized controlled trial [Dissertation]. Ann Arbor (MI): ProQuest Dissertations Publishing, 2016.

Clark 1998 *{published data only}* **10.3928/0098-9134-19980701-05**

Clark ME, Lipe AW, Bilbrey M. Use of music to decrease aggressive behaviors in people with dementia. *Journal of Gerontological Nursing* 1998;**24**(7):10-7.

Cooke 2010 *{published data only}*

ACTRN12612000616875. The effect of music on agitated behaviours in older people with dementia: a randomised cross-over study. Australian New Zealand Clinical Trials Registry 8 June 2012. [WEBPAGE: https://sandbox.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=362563]

* Cooke M, Moyle W, Shum D, Harrison S, Murfield J. A randomized controlled trial exploring the effect of music on quality of life and depression in older people with dementia. *Journal of Health Psychology* 2010;**15**(5):765-76. [DOI: 10.1177/1359105310368188]

Cooke ML, Moyle W, Shum DH, Harrison SD, Murfield JE. A randomized controlled trial exploring the effect of music on agitated behaviours and anxiety in older people with dementia. *Aging & Mental Health* 2010;**14**(8):905-16. [DOI: 10.1080/13607861003713190]

Harrison S, Cooke M, Moyle W, Shum D, Murfield J. Development of a music intervention protocol and its effect on participant engagement: experiences from a randomised controlled trial with older people with dementia. *Arts & Health* 2010;**2**(2):125-39.

Murfield J, Cooke M, Moyle W, Shum D, Harrison S. Conducting randomized controlled trials with older people with dementia in long-term care: challenges and lessons learnt. *International Journal of Nursing Practice* 2011;**17**(1):52-9. [DOI: 10.1111/j.1440-172X.2010.01906.x]

Giovagnoli 2018 *{published data only}* **10.1007/s10072-018-3316-3**

Giovagnoli AR, Manfredi V, Schifano L, Paterlini C, Parente A, Tagliavini F. Combining drug and music therapy in patients with moderate Alzheimer's disease: a randomized study. *Neurological Sciences* 2018;**39**(6):1021-8. [DOI: 10.1007/s10072-018-3316-3]

Gómez-Gallego 2021 *{published data only}* **10.3390/ijerph18158067**

* Gómez-Gallego M, Gómez-Gallego JC, Gallego-Mellado M, García-García J. Comparative efficacy of active group music intervention versus group music listening in Alzheimer's disease. *International Journal of Environmental Research and Public Health* 2021;**18**(15):8067. [DOI: 10.3390/ijerph18158067]

NCT04761497. Effects of group music therapy interventions on Alzheimer's Disease patients: a three-arm randomized controlled blind study. https://clinicaltrials.gov/

study/NCT04761497 (first posted 21 February 2021).
[CLINICALTRIALS.GOV IDENTIFIER: NCT04761497]

Guétin 2009 *{published data only}***10.1159/000229024**

Guétin S, Portet F, Picot MC, Pommie C, Messaoudi M,
Djabelkir L, et al. Effect of music therapy on anxiety and
depression in patients with Alzheimer's type dementia:
randomised, controlled study. *Dementia and Geriatric Cognitive
Disorders* 2009;**28**(1):36-46. [DOI: 10.1159/000229024]

Hsu 2015 *{published and unpublished data}***10.1186/
s12877-015-0082-4**

Hsu MH, Flowerdew R, Parker M, Fachner J, Odell-Miller H.
Individual music therapy for managing neuropsychiatric
symptoms for people with dementia and their carers: a
cluster randomised controlled feasibility study. *BMC Geriatrics*
2015;**15**:84. [DOI: 10.1186/s12877-015-0082-4]

NCT01744600. Music therapy in methodist homes: a study
investigating the impact of a music therapy programme. https://
clinicaltrials.gov/study/NCT01744600 (first posted 6 December
2012).

Liesk 2015 *{published data only}***10.1007/s00391-014-0661-6**

Liesk J, Hartogh T, Kalbe E. Cognitive stimulation and music
intervention for people with dementia in nursing homes: a
pilot study, problems and perspectives [Kognitive Stimulation
und Musikintervention bei stationär versorgten Menschen
mit Demenz: eine Pilotstudie, Probleme und Perspektiven].
Zeitschrift für Gerontologie und Geriatrie 2015;**48**(3):275-81.

Lin 2011 *{published data only}***10.1002/gps.2580**

Chu H, Yang CY, Lin Y, Ou KL, Lee TY, O'Brien AP, et al. The
impact of group music therapy on depression and cognition
in elderly persons with dementia: a randomized controlled
study. *Biological Research for Nursing* 2014;**16**(2):209-17. [DOI:
10.1177/1099800413485410]

* Lin Y, Chu H, Yang CY, Chen CH, Chen SG, Chang HJ, et al.
Effectiveness of group music intervention against agitated
behavior in elderly persons with dementia. *International
Journal of Geriatric Psychiatry* 2011;**26**(7):670-8. [DOI: 10.1002/
gps.2580]

Liu 2021 *{published data only}***10.12659/MSM.928714**

Liu MN, Liou YJ, Wang WC, Su KC, Yeh HL, Lau CI, et al. Group
music intervention using percussion instruments to reduce
anxiety among elderly male veterans with Alzheimer disease.
Medical Science Monitor 2021;**27**:e928714. [DOI: 10.12659/
MSM.928714]

Lord 1993 *{published data only}***10.2466/pms.1993.76.2.451**

Lord TR, Garner JE. Effects of music on Alzheimer patients.
Perceptual and Motor Skills 1993;**76**(2):451-5.

Lyu 2014 *{published data only}*

Lyu J, Gao T, Li M, Xie L, Li W, Jin W, et al. The effect of music
therapy on memory, language and psychological symptoms
of patients with mild Alzheimer's disease. *Chinese Journal of
Neurology* 2014;**47**(12):831-5.

Narme 2012-study 1 *{published and unpublished data}***10.1684/
pnv.2012.0343**

Narme P, Tonini A, Khatir F, Schiaratura L, Clément S,
Samson S. Non-pharmacological treatment for Alzheimer's
disease: comparison between musical and non-musical
interventions [Thérapies non médicamenteuses dans la
maladie d'Alzheimer: comparaison d'ateliers musicaux et
non musicaux]. *Gériatrie et Psychologie Neuropsychiatrie du
Vieillissement* 2012;**10**(2):215-24.

Narme 2012-study 2 *{published and unpublished data}***10.1684/
pnv.2012.0343**

Narme P, Tonini A, Khatir F, Schiaratura L, Clément S,
Samson S. Non pharmacological treatment for Alzheimer's
disease: comparison between musical and non-musical
interventions [Thérapies non médicamenteuses dans la
maladie d'Alzheimer: comparaison d'ateliers musicaux et
non musicaux]. *Gériatrie et Psychologie Neuropsychiatrie du
Vieillissement* 2012;**10**(2):215-24.

Narme 2014 *{published and unpublished data}***10.3233/
JAD-130893**

* Narme P, Clément S, Ehrlé N, Schiaratura L, Vachez S,
Courtaigne B, et al. Efficacy of musical interventions in
dementia: evidence from a randomized controlled trial.
Journal of Alzheimer's Disease 2014;**38**(2):359-69. [DOI: 10.3233/
JAD-130893]

Samson S, Clément S, Narme P, Schiaratura L, Ehrlé N.
Efficacy of musical interventions in dementia: methodological
requirements of nonpharmacological trials. *Annals of the New
York Academy of Sciences* 2015;**1337**:249-55. [DOI: 10.1111/
nyas.12621]

Prieto Alvarez 2022 *{published data only (unpublished sought but
not used)}***https://doi.org/10.1093/mtp/miab021**

Prieto Alvarez L. Neurologic music therapy with a habilitative
approach for older adults with dementia: a feasibility study.
Music Therapy Perspectives 2022;**40**:76-83.

Raglio 2010a *{published and unpublished
data}***10.1080/13607861003713158**

Raglio A, Bellelli G, Traficante D, Gianotti M, Ubezio MC,
Gentile S, et al. Addendum to 'Efficacy of music therapy
treatment based on cycles of sessions: a randomised
controlled trial' (Raglio et al., 2010). *Aging & Mental Health*
2012;**16**(2):265-7. [DOI: 10.1080/13607863.2011.630376]

* Raglio A, Bellelli G, Traficante D, Gianotti M, Ubezio MC,
Gentile S, et al. Efficacy of music therapy treatment based on
cycles of sessions: a randomised controlled trial. *Aging & Mental
Health* 2010;**14**(8):900-4. [DOI: 10.1080/13607861003713158]

Raglio 2010b *{published and unpublished
data}***10.2174/1874609811003030242**

Raglio A, Oasi O, Gianotti M, Manzoni V, Bolis S, Ubezio MC, et al.
Effects of music therapy on psychological symptoms and heart
rate variability in patients with dementia: a pilot study. *Current
Aging Science* 2010;**3**(3):242-6.

Cochrane
Library

Trusted evidence.
Informed decisions.
Better health.

Cochrane Database of Systematic Reviews

Raglio 2015 *{published and unpublished data}***10.1111/jgs.13558**

Raglio A, Bellandi D, Baiardi P, Gianotti M, Ubezio MC, Zanacchi E, et al. Effect of active music therapy and individualized listening to music on dementia: a multicenter randomized controlled trial. *Journal of the American Geriatrics Society* 2015;**63**(8):1534-9. [DOI: 10.1111/jgs.13558]

Reschke-Hernández 2019 *{published and unpublished data}*

NCT03643003. Music therapy for persons with dementia [Music therapy for persons with dementia: a randomized clinical crossover trial]. https://trialsearch.who.int/Trial2.aspx?TrialID=NCT03643003 (first posted 20 August 2018).

Reschke-Hernández AE, Gfeller K, Oleson J, Tranel D. Music therapy increases social and emotional well-being in persons with dementia: a randomized clinical crossover trial comparing singing to verbal discussion. *Journal of Music Therapy* 2023;**60**(3):314-42. [DOI: 10.1093/jmt/thad015]

Reschke-Hernández AE, Tranel D. Strategies to enhance treatment fidelity and music-based intervention reporting in dementia research. *Translational Behavioral Medicine* 2024;**14**(6):353-8. [DOI: 10.1093/tbm/ibae003]

* Reschke-Hernández AE. A Clinical Practice Model of Music Therapy to Address Psychosocial Functioning for Persons with Dementia: Model Development and Randomized Clinical Crossover Trial (thesis). Iowa (IA): University of Iowa, 2019. [PUBLISHER: ProQuest LLC, Ann Arbor, MI 48106 - 1346, USA] [THE UNIVERSITY OF IOWA PROQUEST DISSERTATIONS & THESES, 2019: 13806570]

Ridder 2013 *{published data only}***10.1080/13607863.2013.790926**

Ridder HM, Stige B, Qvale LG, Gold C. Individual music therapy for agitation in dementia: an exploratory randomized controlled trial. *Aging & Mental Health* 2013;**17**(6):667-78.

Sakamoto 2013 *{published data only}***10.1017/S1041610212002256**

JPRN-UMIN000008801. Examination of effective nursing intervention for music therapy for severe dementia elderly person. International Clinical Trials Registry Platform 2012. [TRIAL REGISTER: JPRN-UMIN000008801]

Sakamoto M, Ando H, Tsutou A. Comparing the effects of different individualized music interventions for elderly individuals with severe dementia. *International Psychogeriatrics* 2013;**25**(5):775-84.

Sung 2012 *{published data only}***10.1002/gps.2761**

Sung HC, Lee WL, Li TL, Watson R. A group music intervention using percussion instruments with familiar music to reduce anxiety and agitation of institutionalized older adults with dementia. *International Journal of Geriatric Psychiatry* 2012;**27**(6):621-7. [DOI: 10.1002/gps.2761]

Svansdottir 2006 *{published and unpublished data}***10.1017/S1041610206003206**

Svansdottir HB, Snaedal J. Music therapy in moderate and severe dementia of Alzheimer's type: a case-control study. *International Psychogeriatrics* 2006;**18**(4):613-21.

Tang 2018 *{published data only}***10.1016/j.gerinurse.2018.02.003**

Tang Q, Zhou Y, Yang S, Thomas WKS, Smith GD, Yang Z, et al. Effect of music intervention on apathy in nursing home residents with dementia. *Geriatric Nursing* 2018;**39**(4):471-6. [DOI: 10.1016/j.gerinurse.2018.02.003] [PMID: 29551546]

Thornley 2016 *{published and unpublished data}***10.1017/S1041610215001866**

Thornley J, Hirjee H, Vasudev A. Music therapy in patients with dementia and behavioral disturbance on an inpatient psychiatry unit: results from a pilot randomized controlled study. *International Psychogeriatrics* 2016;**28**(5):869-71. [DOI: 10.1017/S1041610215001866]

Vink 2013 *{published and unpublished data}***10.1002/gps.3924**

Vink AC, Zuidersma M, Boersma F, de Jonge P, Zuidema SU, Slaets JP. Effect of music therapy versus recreational activities on neuropsychiatric symptoms in elderly adults with dementia: an exploratory randomized controlled trial. *Journal of the American Geriatrics Society* 2014;**62**(2):392-3. [DOI: 10.1111/jgs.12682]

* Vink AC, Zuidersma M, Boersma F, de Jonge P, Zuidema SU, Slaets JP. The effect of music therapy compared with general recreational activities in reducing agitation in people with dementia: a randomised controlled trial. *International Journal of Geriatric Psychiatry* 2013;**28**(10):1031-8. [DOI: 10.1002/gps.3924]

Vink AC. Music Therapy for Dementia: the Effect of Music Therapy in Reducing Behavioural Problems in Elderly People with Dementia [thesis]. Groningen, The Netherlands: University of Groningen, 2013. [ISBN: 978-90-367-6191-8]

Werner 2017 *{published and unpublished data}***10.1080/13607863.2015.1093599**

Werner J, Wosch T, Gold C. Effectiveness of group music therapy versus recreational group singing for depressive symptoms of elderly nursing home residents: pragmatic trial. *Aging & Mental Health* 2017;**21**(2):147-55. [DOI: 10.1080/13607863.2015.1093599]

References to studies excluded from this review

Arroyo-Anlló 2013 *{published data only}*

Arroyo-Anlló EM, Díaz JP, Gil R. Familiar music as an enhancer of self-consciousness in patients with Alzheimer's disease. *BioMed Research International* 2013;**2013**:752965. [DOI: 10.1155/2013/752965]

Baker 2023 *{published data only}***10.1136/bmjopen-2019-031332**

Baker FA, Bloska J, Braat S, Bukowska A, Clark I, Hsu MH, et al. HOMESIDE: home-based family caregiver-delivered music and reading interventions for people living with dementia: protocol of a randomised controlled trial. *BMJ Open* 2019;**9**(11):e031332. [DOI: 10.1136/bmjopen-2019-031332]

* Baker FA, Pac Soo V, Bloska J, Blauth L, Bukowska AA, Flynn L, et al. Home-based family caregiver-delivered music and reading interventions for people living with dementia (HOMESIDE trial):

an international randomised controlled trial.. *EClinicalMedicine* 2023;**65**:102224. [DOI: 10.1016/j.eclinm.2023.102224]

Ballard 2009 *{published data only}***10.1097/JGP.0b013e3181b0f8c0**

Ballard C, Brown R, Fossey J, Douglas S, Bradley P, Hancock J, et al. Brief psychosocial therapy for the treatment of agitation in Alzheimer disease (the CALM-AD trial). *American Journal of Geriatric Psychiatry* 2009;**17**(9):726-33. [DOI: 10.1097/JGP.0b013e3181b0f8c0]

Brotons 2000 *{published data only}***10.1093/jmt/37.3.183.**

Brotons M, Koger SM. The impact of music therapy on language functioning in dementia. *Journal of Music Therapy* 2000;**XXXVII**(3):183-95. [DOI: 10.1093/jmt/37.3.183]

Bruer 2007 *{published data only}***10.1093/jmt/44.4.308**

Bruer RA, Spitznagel E, Cloninger CR. The temporal limits of cognitive change from music therapy in elderly persons with dementia or dementia-like cognitive impairment: a randomized controlled trial. *Journal of Music Therapy* 2007;**44**(4):308-28. [DOI: 10.1093/jmt/44.4.308]

Bugos 2005 *{published data only}*

Bugos JA. The effects of individualized piano instruction on executive functions in older adults (ages 60-85). Dissertation Abstracts International 2005;**66**(1-A):18. [UNIV OF FLORIDA: //ufdc.ufl.edu/AA00037129/00001]

Chae 2015 *{published data only}***10.7586/jkbns.2015.17.3.245**

Chae SK. Effects of laughing and music therapy on depression and activities of the autonomic nervous system in the elderly with dementia [요양시설에 거주하는 치매노인을 위한 웃음요법과 음악요법이 우울과 자율신경활성도에 미치는 영향] . *Journal of Korean Biological Nursing Science* 2015;**17**(3):245-52.

Clair 1996 *{published data only}*

Clair AA. The effect of singing on alert responses in persons with late stage dementia. *Journal of Music Therapy* 1996;**XXXIII**(4):234-47.

Cohen-Mansfield 2010 *{published data only}***10.1111/j.1532-5415.2010.02951.x**

Cohen-Mansfield J, Marx MS, Dakheel-Ali M, Regier NG, Thein K, Freedman L. Can agitated behavior of nursing home residents with dementia be prevented with the use of standardized stimuli? *Journal of the American Geriatrics Society* 2010;**58**(8):1459-64. [DOI: 10.1111/j.1532-5415.2010.02951.x]

Davidson 2011 *{published data only}***https://doi.org/10.1177/1029864911410954**

Davidson JW, Fedele J. Investigating group singing activity with people with dementia and their caregivers: problems and positive prospects. *Musicae Scientiae* 2011;**15**(3):402-22. [DOI: 10.1177/1029864911410954]

Dowson 2021 *{published data only}***10.1186/s40814-020-00759-y**

Dowson B, Schneider J, Guo B, Bath PM, McDermott O, Haywood LJ, et al. Pilot randomised evaluation of singing in dementia (PRESIDE): protocol for a two-arm, parallel-group randomised controlled feasibility study with waiting-list

control. *Pilot and Feasibility Studies* 2021;**7**(1):15. [DOI: 10.1186/s40814-020-00759-y]

Eun-Young 2020 *{published data only}***10.32451/KJOAPS.2020.16.2.273**

Eun-Young P, Eun-Young H. The effects of PERMA based music activities on depression, interpersonal relationship and the life qualities of the elderly mild dementia patients in long-term care facilities [PERMA기반 음악활동이 장기요양시설 경증치매 노인의 우울, 대인관계 및 삶의 질에 미치는 영향] . *Journal of Arts Psychotherapy* 2020;**16**(2):273-300.

Garland 2007 *{published data only}***10.1097/01.JGP.0000249388.37080.b4**

Garland K, Beer E, Eppingstall B, O'Connor DW. A comparison of two treatments of agitated behavior in nursing home residents with dementia: simulated family presence and preferred music. *American Journal of Geriatric Psychiatry* 2007;**15**(6):514-21. [DOI: 10.1097/01.JGP.0000249388.37080.b4]

Gerdner 2000 *{published data only}***10.1017/s1041610200006190**

Gerdner LA. Effects of individualized versus classical "relaxation" music on the frequency of agitation in elderly persons with Alzheimer's disease and related disorders. *International Psychogeriatrics* 2000;**12**(1):49-65. [DOI: 10.1017/s1041610200006190]

Groene 1993 *{published data only}*

Groene RW. Effectiveness of music therapy 1:1 intervention with individuals having senile dementia of the Alzheimer's type. *Journal of Music Therapy* 1993;**30**(3):138-57.

Hanser 1994 *{published data only}***10.1093/geronj/49.6.p265**

Hanser SB, Thompson LW. Effects of a music therapy strategy on depressed older adults. *Journal of Gerontology* 1994;**49**(6):P265-9. [DOI: 10.1093/geronj/49.6.p265]

Hicks-Moore 2008 *{published data only}***https://doi.org/10.1177/1471301207085369**

Hicks-Moore SL, Robinson BA. Favorite music and hand massage: two interventions to decrease agitation in residents with dementia. *Dementia* 2008;**7**(1):95-108. [DOI: https://doi.org/10.1177/1471301207085369]

Hokkanen 2008 *{published data only}***0.1111/j.1532-5415.2008.01611.x**

Hokkanen L, Rantala L, Remes AM, Härkönen B, Viramo P, Winblad I. Dance and movement therapeutic methods in management of dementia: a randomized, controlled study. *Journal of the American Geriatrics Society* 2008;**56**(4):771-2. [DOI: 0.1111/j.1532-5415.2008.01611.x]

Holmes 2006 *{published data only}***10.1017/S1041610206003887**

Holmes C, Knights A, Dean C, Hodkinson S, Hopkins V. Keep music live: music and the alleviation of apathy in dementia subjects. *International Psychogeriatrics* 2006;**18**(4):623-30. [DOI: 10.1017/S1041610206003887]

Janata 2012 *{published data only}*

Janata P. Effects of widespread and frequent personalized music programming on agitation and depression in assisted

living facility residents with Alzheimer-type dementia. *Music and Medicine* 2012;**4**(1):8-15.

Kwak 2016 *{published data only}***10.1093/geront/gnw162.806**

Kwak J, Valuch KC, Brondino M. Randomized crossover study of music & memory intervention with nursing home residents with dementia. *Gerontologist* 2016;**56**(Suppl 3):205.

Low 2016 *{published data only}***10.1016/j.ctim.2016.09.005**

Low LF, Carroll S, Merom D, Baker JR, Kochan N, Moran F, et al. We think you can dance! A pilot randomised controlled trial of dance for nursing home residents with moderate to severe dementia. *Complementary Therapies in Medicine* 2016;**29**:42-4. [DOI: 10.1016/j.ctim.2016.09.005]

Lyu 2018 *{published data only}***10.3233/JAD-180183**

Lyu J, Zhang J, Mu H, Li W, Champ M, Xiong Q, et al. The effects of music therapy on cognition, psychiatric symptoms, and activities of daily living in patients with Alzheimer's disease. *Journal of Alzheimer's Disease* 2018;**64**:1347-58. [DOI: 10.3233/JAD-180183]

Matziorinis 2023 *{published data only}*

Flo BK, Matziorinis AM, Skouras S, Sudmann TT, Gold C, Koelsch S. Study protocol for the Alzheimer and Music Therapy (ALMUTH) Study: a 12-month, randomised controlled trial to compare the efficacy of music therapy and physical activity on brain plasticity, depressive symptoms, and cognitive decline, in a population with and at risk for Alzheimer's disease. Research Square preprint 2022;**February 22**:1-27. [DOI: 10.21203/rs.3.rs-1380534/v1]

Flo BK, Matziorinis AM, Skouras S, Sudmann TT, Gold C, Koelsch S. Study protocol for the Alzheimer and music therapy study: an RCT to compare the efficacy of music therapy and physical activity on brain plasticity, depressive symptoms, and cognitive decline, in a population with and at risk for Alzheimer's disease. *PLoS One* 2022;**17**(6):e0270682. [CENTRAL: PMC9246122] [DOI: 10.1371/journal.pone.0270682] [PMID: 35771851]

* Matziorinis AM, Flo BK, Skouras S, Dahle K, Henriksen A, Hausmann F, et al. A 12-month randomised pilot trial of the Alzheimer's and music therapy study: a feasibility assessment of music therapy and physical activity in patients with mild-to-moderate Alzheimer's disease. *Pilot and Feasibility Studies* 2023;**9**(1):61. [DOI: 10.1186/s40814-023-01287-1] [PMID: 37076884]

Matziorinis AM, Flo BK, Skouras S, Dahle K, Sudmann TT, Gold C, et al. A 12-month randomised pilot trial of the Alzheimer's and music therapy study: a feasibility assessment of music therapy and physical activity in patients with mild-to-moderate Alzheimer's disease. Research Square preprint 7 January 2022;**v1**:n/a. [DOI: https://doi.org/10.21203/rs.3.rs-1238008/v1]

McCreedy 2022 *{published data only}***10.1016/j.jamda.2021.12.030**

McCreedy EM, Sisti A, Gutman R, Dionne L, Rudolph JL, Baier R, et al. Pragmatic trial of personalized music for agitation and antipsychotic use in nursing home residents with dementia. *Journal of the American Medical Directors Association* 2022;**23**(7):1171-7. [DOI: 10.1016/j.jamda.2021.12.030]

Noice 2009 *{published data only}***10.1080/13825580802233400**

Noice H, Noice T. An arts intervention for older adults living in subsidized retirement homes. *Neuropsychology, Development and Cognition. Section B. Aging, Neuropsychology and Cognition* 2009;**16**(1):56-79. [DOI: 10.1080/13825580802233400]

Otto 1999 *{published data only}***10.1093/jmt/36.3.182**

Otto D, Cochran VV, Johnson G, Clair AA. The influence of background music on task engagement in frail, older persons in residential care. *Journal of Music Therapy* 1999;**36**(3):182-95. [DOI: 10.1093/jmt/36.3.182]

Pomeroy 1993 *{published data only}*

Pomeroy VM. The effect of physiotherapy input on mobility skills of elderly people with severe dementing illness. *Clinical Rehabilitation* 1993;**7**(2):163-70.

Raglio 2008 *{published data only}*

Raglio A, Bellelli G, Traficante D, Gianotti M, Ubezio MC, Villani D, et al. Efficacy of music therapy in the treatment of behavioral and psychiatric symptoms of dementia. *Alzheimer Disease and Associated Disorders* 2008;**22**(2):158-62. [DOI: 10.1097/WAD.0b013e3181630b6f]

Riegler 1980 *{published data only}*

Riegler J. Comparison of a reality orientation program for geriatric patients with and without music. *Journal of Music Therapy* 1980;**17**(1):26-33. [DOI: 0.1093/jmt/17.1.26]

Sánchez 2016 *{published data only}*

Sánchez A, Maseda A, Marante-Moar MP, De Labra C, Lorenzo-López L, Millán-Calenti JC. Comparing the effects of multisensory stimulation and individualized music sessions on elderly people with severe dementia: a randomized controlled trial. *Journal of Alzheimer's Disease* 2016;**52**(1):303-15. [DOI: 10.3233/JAD-151150] [PMID: 27060958]

Särkämö 2014 *{published data only}*

Särkämö T, Tervaniemi M, Laitinen S, Numminen A, Kurki M, Johnson JK, et al. Cognitive, emotional, and social benefits of regular musical activities in early dementia: randomized controlled study. *Gerontologist* 2014;**54**(4):634-50. [DOI: 10.1093/geront/gnt100]

Satoh 2014 *{published data only}*

Satoh M, Ogawa J, Tokita T, Nakaguchi N, Nakao K, Kida H, et al. The effects of physical exercise with music on cognitive function of elderly people: Mihama-Kiho project. *PloS One* 2014;**9**(4):e95230. [DOI: 10.1371/journal.pone.0095230]

Sung 2006 *{published data only}***10.1016/j.ctim.2006.03.002**

Sung HC, Chang SM, Lee WL, Lee MS. The effects of group music with movement intervention on agitated behaviours of institutionalized elders with dementia in Taiwan. *Complementary Therapies in Medicine* 2006;**14**(2):113-9. [DOI: 10.1016/j.ctim.2006.03.002]

Tamplin 2018 *{published data only}*

Baker FA, Stretton-Smith P, Clark IN, Tamplin J, Lee YC. A group therapeutic songwriting intervention for family caregivers of people living with dementia: a feasibility study with thematic analysis. *Frontiers in Medicine* 2018;**5**:151. [DOI: 10.3389/fmed.2018.00151]

* Tamplin J, Clark IN, Lee YC, Baker FA. Remini-Sing: a feasibility study of therapeutic group singing to support relationship quality and wellbeing for community-dwelling people living with dementia and their family caregivers. *Frontiers in Medicine* 2018;**5**:245. [DOI: 10.3389/fmed.2018.00245]

Thompson 2005 *{published data only}*

Thompson RG, Moulin CJ, Hayre S, Jones RW. Music enhances category fluency in healthy older adults and Alzheimer's disease patients. *Experimental Aging Research* 2005;**31**(1):91-9. [DOI: 10.1080/03610730590882819]

Tomaino 2011 *{unpublished data only}*

Tomaino CM, Kim M. Enhancing community-based and home care with music therapy; final report. New York (NY): Institute for Music and Neurologic Function; Beth Abraham Family of Health Services 2011.

Tz-Han 2023 *{published data only}*

Tz-Han L, Wan-Ru W, I-Hui C, Hui-Chuan. Reminiscence music intervention on cognitive, depressive, and behavioral symptoms in older adults with dementia. *Geriatric Nursing* 2023;**49**:127-32. [DOI: 10.1016/j.gerinurse.2022.11.014] [PMID: 36495795]

Vanderark 1983 *{published data only}*

Vanderark S, Newman I, Bell S. The effects of music participation on quality of life of the elderly. *Music Therapy* 1983;**3**(1):71-81. [DOI: 10.1093/mt/3.1.71]

Van de Winckel 2004 *{published data only}* **10.1191/0269215504cr750oa**

Van de Winckel A, Feys H, De Weerdt W, Dom R. Cognitive and behavioural effects of music-based exercises in patients with dementia. *Clinical Rehabilitation* 2004;**18**(3):253-60. [DOI: 10.1191/0269215504cr750oa]

채경숙 2015 *{published data only}*

채경숙 . Effects of laughing and music therapy on depression and activities of the autonomic nervous system in the elderly with dementia. *Journal of Korean Biological Nursing Science* 2015;**3**:245-52.

References to studies awaiting assessment

Asmussen 1999 *{published and unpublished data}*

* Asmussen S, Shah V, Goldstein A, Tomaino C, Scheiby B, Ramsey D. The effect of a music therapy intervention on the levels of depression, anxiety/agitation, and quality of life experienced by individuals diagnosed with early and middle stage dementia. A controlled study (unpublished report). New York (NY): Beth Abraham Health Services, Center for Research and Education 1999.

Asmussen SA, Shah V, Goldstein A, Tomaino CM, Scheiby B, Ramsey D. The effects of a music therapy intervention on the levels of depression, anxiety/agitation, and quality of life experienced by individuals diagnosed with early and middle stage dementias: a controlled study (Final Report). New York (NY): Beth Abraham Health Services Center for Research and Education 1999:1-67.

Tomaino CM. The influence of music on memory in persons with dementia. New York (NY): Beth Abraham Center for Research and Education 1998.

Campbell 2022 *{published data only}*

* Campbell E, Wosch T. Music moves: audio and audio-tactile music therapy for dementia in German residential care homes. *British Journal of Music Therapy* 2022;**Online Special Edition**(Meeting abstracts):229. [CONFERENCE WEBSITE: https://www.qmu.ac.uk/conferences-and-events/emtc-2022/]

DRKS00023233. Music moves - Ettlinger Dementia Study [Musik beseelt - Ettlinger Demenzstudie]. https://drks.de/search/en/trial/DRKS00023233 (first posted 28 October 2020).

Curto Prieto 2015 *{published data only}*

Curto Prieto D, Sole Serrano C, De Castro M, Mercadal Brotons M, Asensio FM. Effects of music therapy in institutionalized elderly with dementia. *European Geriatric Medicine* 2015;**6**(Meeting abstracts):S48. [DOI: 10.1016/S1878-7649(15)30164-9]

Hong 2011 *{published data only}*

Hong IS, Choi MJ. Songwriting oriented activities improve the cognitive functions of the aged with dementia. *Arts in Psychotherapy* 2011;**38**:221-8. [DOI: https://doi.org/10.1016/j.aip.2011.07.002]

Hsiung 2013 *{published data only}*

Hsiung G-Y, Kirkland K, Hswen Y, Slack PJ, Summers S, Boyd L, et al. A pilot, randomized study of music therapy for people with Alzheimer's disease. *Alzheimer's & Dementia* 2013;**9**(4 (Meeting abstracts)):477.

Hsiung 2015 *{published data only}*

Hsiung G-YR, Kirkland K, Summers SG, Beattie BL, Jacova C. A randomized controlled trial of music therapy in managing behavioral symptoms in Alzheimer disease. *Alzheimer's & Dementia* 2015;**11**(S_Part_16 (Meeting abstracts)):749. [DOI: https://doi.org/10.1016/j.jalz.2015.06.1671]

ISRCTN11001662 *{unpublished data only}* **ISRCTN11001662**

ISRCTN11001662. Pilot trial of an interactive music programme for individuals living with dementia [Pilot trial of an interactive music intervention for individuals living with moderate to severe dementia]. https://trialsearch.who.int/Trial2.aspx?TrialID=ISRCTN11001662 (first posted 15 July 2016). [TRIAL REGISTER: ISRCTN11001662]

Kwak 2013 *{published data only}*

* Kwak K, Bae N, Jang WY. Music therapy with moderate Alzheimer's disease in a long-term care center. *Alzheimer's & Dementia* 2013;**9**(4 (Meeting abstracts)):880. [DOI: 10.1016/j.jalz.2013.08.226] [ABSTRACT NUMBER: P4-393]

Kwak K-P. Music therapy with moderate Alzheimer's disease in long-term care center. *16th International Congress of the International Psychogeriatric Association (IPA); 2013, Oct 1-4; Seoul; South Korea* 2013;**S1 (Meeting abstracts)**:S188.

Moreira 2023 *{published data only}*

Moreira SV, Justi FDR, Gomes CFA, Moreira M. Music therapy enhances episodic memory in Alzheimer's and mixed dementia: a double-blind randomized controlled trial. *Healthcare* 2023;**11**:2912. [DOI: 10.3390/healthcare11222912]

NCT00448318 *{unpublished data only}*

NCT00448318. Evaluating the effects of music interventions on hospitalised people with dementia. https://trialsearch.who.int/Trial2.aspx?TrialID=NCT00448318 (first posted 14 March 2007). [NCT00448318]

NCT01709188 *{published and unpublished data}*

NCT01709188. Musical dual task training to improve attention control for dementia [The effects of musical dual task training on attention control and associated gait stability of patients with mild to moderate dementia]. https://clinicaltrials.gov/study/NCT01709188 (first posted 18 October 2012). [NCT01709188]

NCT01958983 *{unpublished data only}*

NCT01958983. Music therapy to promote emotion and cognition for dementia [The effect of music therapy on emotion and cognition for individuals with mild to moderate dementia]. https://clinicaltrials.gov/study/NCT01958983 (first posted 9 October 2013). [NCT01958983]

NCT02020356 *{unpublished data only}*

NCT02020356. Music therapy in Alzheimer's disease (PRE-MUSICAL) [Impact of music therapy on prescription of neuroleptics in Alzheimer's Disease ou related symptoms associated with disruptive behaviours]. https://clinicaltrials.gov/study/NCT02020356 (first posted 24 December 2013). [NCT02020356]

NCT02147652 *{unpublished data only}*

Anor CJ, Moy S, Lake A, Shan R, Cockburn A, Krisman D, et al. Assessing the effects of a personalized music therapy with headphones on agitation in patients with dementia. *Canadian Geriatrics Journal* 2018;**21**(1 (Meeting abstracts)):120-1. [https://cgjonline.ca/index.php/cgj/article/view/321/377]

* NCT02147652. Personalized music therapy and agitation in dementia [Assessing the effects of a personalized music therapy with headphones on agitation in patients with dementia]. https://clinicaltrials.gov/study/NCT02147652 (first posted 28 May 2014). [NCT02147652]

NCT02670993 *{published data only}*

NCT02670993. Support by singing sessions on physical and moral pain: assessment of its effectiveness in Alzheimer's Disease (LACME) [Support by singing sessions on physical and moral pain: assessment of its effectiveness in Alzheimer's Disease. Multicenter study LACME]. https://clinicaltrials.gov/study/NCT02670993 (first posted 2 February 2016).

Prick 2024 *{published data only}***https://trialsearch.who.int/Trial2.aspx?TrialID=NL8861**

NL-OMON21011. Striking the right chord with mUsic in NurSing home resIdeNts with dEmentia [The effects of an individual listening music intervention and an individual music therapy intervention on behaviour and well-being in nursing home residents with dementia]. https://trialsearch.who.int/Trial2.aspx?TrialID=NL-OMON21011 (first posted 1 September 2020). [CCMO (DUTCH MEDICAL ETHICS REGISTRY): NL60766.096.17] [WHO INTERNATIONAL CLINICAL TRIALS REGISTRY PLATFORM: NL8861]

* Prick AJ, Zuidema SU, van Domburg P, Verboon P, Vink AC, Schols JM, et al. Effects of a music therapy and music listening intervention for nursing home residents with dementia: a randomized controlled trial. *Frontiers in Medicine* 2024;**11**:1304349. [DOI: 10.3389/fmed.2024.1304349]

권서령 2013 *{published data only}*

권서령, 강경선. The effect of music therapy using Korean folk song and rhythm on improving the cognitive function of elderly people with dementia. *Journal of Arts Psychotherapy* 2013;**2**:1-17.

김현정 2013 *{published data only}*

김현정, 정재원. A study on the effect of Korean traditional music therapy with on depression and quality of life of the elderly with mild dementia. *Journal of Arts Psychotherapy* 2013;**3**:19-38.

신보영, 황은영 2015 *{published data only}*

신보영, 황은영. A research on the effects of music activities adopting visual supportive strategies on cognitive rehabilitation of elderly dementia patients. *Journal of Arts Psychotherapy* 2015;**2**:41-62.

References to ongoing studies

Baroni Caramel 2024 *{unpublished data only}***https://trialsearch.who.int/Trial2.aspx?TrialID=NL-OMON47943**

* Baroni Caramel VM, Van der Steen JT, Vink AC, Janus SIM, Twisk JWR, Scherder EJA, et al. The effects of individual music therapy in nursing home residents with dementia to improve general well-being: study protocol of a randomized controlled trial. *BMC Geriatrics* 2024;**24**(1):290. [DOI: 10.1186/s12877-024-04863-z] [PMCID: PMC10976838] [PMID: 38539079]

NL-OMON47943. Individual music therapy and wellbeing in nursing home residents with dementia: a study protocol of a randomized controlled trial. https://trialsearch.who.int/Trial2.aspx?TrialID=NL-OMON47943 (first posted 27 June 2019).

Gold 2018 *{published and unpublished data}*

* Gold C, Eickholt J, Assmus J, Stige B, Wake JD, Baker FA, et al. Music Interventions for Dementia and Depression in ELderly care (MIDDEL): protocol and statistical analysis plan for a multinational cluster-randomised trial. *BMJ Open* 2019;**9**(3):e023436. [DOI: 10.1136/bmjopen-2018-023436]

NCT03496675. Music Interventions for Dementia and Depression in Elderly Care (MIDDEL) [Music Interventions for

Dementia and Depression in Elderly Care: international cluster-randomised controlled trial]. https://clinicaltrials.gov/study/NCT03496675 (first posted 12 April 2018).

NCT04666077 *{unpublished data only}*

NCT04666077. Music therapy experiences in patients with mild cognitive impairment (MCI) and Alzheimer's disease (AD) [A randomized, blinded, controlled trial of music therapy singing interventions for patients with MCI and AD]. https://clinicaltrials.gov/study/NCT04666077 (first posted 7 December 2020).

Additional references

Aldridge 1996

Aldridge D. Music Therapy Research and Practice in Medicine: From Out of the Silence. London, UK: Jessica Kingsley Publishers, 1996. [ISBN: 1853022969]

Alzheimer's Disease International 2015

Prince M, Wilmo A, Guerchet M, Ali G-C, Wu Y-T, Prina M. World Alzheimer Report 2015: the global impact of dementia: an analysis of prevalence, incidence, cost and trends; August 2015. Available at www.alz.co.uk/research/WorldAlzheimerReport2015.pdf (accessed prior to 10 July 2018).

Alzheimer's Disease International 2016

Prince M, Karagiannidou M, Guerchet M, Comas-Herrera A, Knapp M. World Alzheimer Report 2016: Improving healthcare for people living with dementia - coverage, quality and costs now and in the future; September 2016. Available at www.alz.co.uk/research/WorldAlzheimerReport2016.pdf (accessed prior to 10 July 2018).

Alzheimer's Disease International 2022

Benoist C, Arthurton L, Barbarino P, Dabas L, Lynch C, Paatalo T, et al. From plan to impact V. WHO Global action plan: the time to act is now; 25 May 2022. Available at www.alzint.org/resource/from-plan-to-impact-v/ (accessed 22 August 2022).

Amano 2022

Amano T, Hooley C, Strong J, Inoue M. Strategies for implementing music-based interventions for people with dementia in long-term care facilities: a systematic review. *International Journal of Geriatric Psychiatry* 2022;**37**(1):1-13. [DOI: 10.1002/gps.5641]

Baird 2009

Baird A, Samson S. Memory for music in Alzheimer's disease: unforgettable? *Neuropsychology Review* 2009;**19**(1):85-101. [DOI: 10.1007/s11065-009-9085-2]

Banerjee 2009

Banerjee S, Samsi K, Petrie CD, Alvir J, Treglia M, Schwam EM, et al. What do we know about quality of life in dementia? A review of the emerging evidence on the predictive and explanatory value of disease specific measures of health related quality of life in people with dementia. *International Journal of Geriatric Psychiatry* 2009;**24**(1):15-24. [DOI: 10.1002/gps.2090]

Beerens 2014

Beerens HC, Sutcliffe C, Renom-Guiteras A, Soto ME, Suhonen R, Zabalegui A, et al, RightTimePlaceCare Consortium. Quality of life and quality of care for people with dementia receiving long term institutional care or professional home care: the European RightTimePlaceCare study. *Journal of the American Medical Directors Association* 2014;**15**(1):54-61. [DOI: 10.1016/j.jamda.2013.09.010]

Bosnjak Kuharic 2021

Bosnjak Kuharic D, Markovic D, Brkovic T, Jeric Kegalj M, Rubic Z, Vuica Vukasovic A, et al. Cannabinoids for the treatment of dementia. *Cochrane Database of Systematic Reviews* 2021, Issue 9. Art. No: CD012820. [DOI: 10.1002/14651858.CD012820.pub2] [PMID: 34532852]

Bowell 2018

Bowell S, Bamford SM. What would life be - without a song or dance, what are we: a report from the Commission on Dementia and Music; 18 January 2018. Available at https://ilcuk.org.uk/what-would-life-be-without-a-song-or-dance-what-are-we/.

Bradt 2012

Bradt J. Randomized controlled trials in music therapy: guidelines for design and implementation. *Journal of Music Therapy* 2012;**49**(2):120-49. [DOI: 10.1093/jmt/49.2.120]

Broersen 1995

Broersen M, De Groot R, Jonker C. Music therapy for Alzheimer patients (In Dutch) [Muziektherapie bij Alzheimer patienten. Enkele richtlijnen op basis van de literatuur]. *Tijdschrift voor Kreatieve Therapie* 1995;**14**(1):9-14. [URL: hu.on.worldcat.org/oclc/72680878]

Carroll 2007

Carroll C, Patterson M, Wood S, Booth A, Rick J, Balain S. A conceptual framework for implementation fidelity. *Implementation Science* 2007;**30**(2):40. [DOI: 10.1186/1748-5908-2-40]

Chang 2015

Chang YS, Chu H, Yang CY, Tsai JC, Chung MH, Liao YM, et al. The efficacy of music therapy for people with dementia: a meta-analysis of randomised controlled trials. *Journal of Clinical Nursing* 2015;**24**(23-24):3425-40. [DOI: 10.1111/jocn.12976]

Cho 2024

Cho E, Lee JY, Yang M, Jang J, Cho J, Kim MJ. Symptom-specific non-pharmacological interventions for behavioral and psychological symptoms of dementia: an umbrella review. *International Journal of Nursing Studies* 2024;**159**:104866. [DOI: 10.1016/j.ijnurstu.2024.104866] [PMID: 39163681]

Clair 1990

Clair AA, Bernstein BA. Preliminary study of music therapy programming for severely regressed persons with Alzheimer's-type dementia. *Journal of Applied Gerontology* 1990;**9**(3):299–311. [DOI: 10.1177/073346489000900305]

Clift 2016

Clift C, Gilbert R, Vella-Burrows T. A choir in every care home. A review of research on the value of singing for older people. Working paper 6; June 2016. Available at https://achoirineverycarehome.wordpress.com/wp-content/uploads/2016/04/wp6-research-review-v2-1.pdf (accessed prior to 5 June 2018).

Cohen-Mansfield 1986

Cohen-Mansfield J, Billig N. Agitated behaviours in the elderly: I. A conceptual review. *Journal of the American Geriatrics Society* 1986;**34**(10):711-21. [DOI: 10.1111/j.1532-5415.1986.tb04302.x]

Cowles 2003

Cowles A, Beatty WW, Nixon SJ, Lutz LJ, Paulk J, Paulk K. Musical skill in dementia: a violinist presumed to have Alzheimer's disease learns to play a new song. *Neurocase* 2003;**9**(6):493-503. [DOI: 10.1076/neur.9.6.493.29378]

Cummings 1994

Cummings JL, Mega M, Gray K. The Neuropsychiatric Inventory: comprehensive assessment of psychopathology in dementia. *Neurology* 1994;**44**(12):2308-14. [DOI: 10.1212/wnl.44.12.2308]

Cummings 2015

Cummings J, Mintzer J, Brodaty H, Sano M, Banerjee S, Devanand DP, et al. Agitation in cognitive disorders: International Psychogeriatric Association provisional consensus clinical and research definition. *International Psychogeriatrics* 2015;**27**(1):7-17. [DOI: 10.1017/S1041610214001963]

Cummings 2020

Cummings J. The Neuropsychiatric Inventory: development and applications. *Journal of Geriatric Psychiatry and Neurology* 2020;**33**(2):73-84. [DOI: 10.1177/0891988719882102] [PMID: 32013737]

De Witte 2022

De Witte M, Pinho ADS, Stams GJ, Moonen X, Bos AER, Van Hooren. Music therapy for stress reduction: a systematic review and meta-analysis. *Health Psychology Review* 2022;**16**(1):134-59. [DOI: 10.1080/17437199.2020.1846580]

De Witte 2024

De Witte M, Nategh L, Antipas H, Westphal A, Lautenschlager NT, Baker FA, et al. The effects of music-based interventions and psychological symptoms of people living with dementia: a systematic review and network meta-analysis protocol. *Aging & Mental Health* 2024;**28**:1726-32. [DOI: 10.1080/13607863.2024.2373969] [PROSPERO: CRD 42023457452]

Engel 1980

Engel GL. The clinical application of the biopsychosocial model. *American Journal of Psychiatry* 1980;**137**(5):535-44. [DOI: 10.1176/ajp.137.5.535]

Folstein 1975

Folstein MF, Folstein SE, McHugh PR. 'Mini-mental state': a practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research* 1975;**12**(3):189-98. [DOI: 10.1016/0022-3956(75)90026-6]

Gerdner 2000

Gerdner LA. Effects of individualized versus classical 'relaxation' music on the frequency of agitation in elderly persons with Alzheimer's disease and related disorders. *International Psychogeriatrics* 2000;**12**(1):49-65. [DOI: 10.1017/s1041610200006190]

GradePro GDT [Computer program]

GradePro GDT. Version accessed 1 August 2016. Hamilton (ON): McMaster University (developed by Evidence Prime), 2023. Available at gradepro.org.

Guetin 2013

Guetin S, Charras K, Berard A, Arbus C, Berthelon P, Blanc F, et al. An overview of the use of music therapy in the context of Alzheimer's disease: a report of a French expert group. *Dementia* 2013;**12**(5):619-34. [DOI: 10.1177/1471301212438290]

Guyatt 2011

Guyatt GH, Oxman AD, Schünemann HJ, Tugwell P, Knottnerus A. GRADE guidelines: a new series of articles in the Journal of Clinical Epidemiology. *Journal of Clinical Epidemiology* 2011;**64**(4):380-2. [DOI: 10.1016/j.jclinepi.2010.09.011]

Hamberger 2011

Hamberger M. Musiktherapie im Alter: Demenz, Depression, Sterben und Tod [Music Therapy for Older People: Dementia, Depression, Dying and Death]. Freising: Laubsänger Verlag, 2011. [ISBN: 978-3-00-025482-6]

Hargreaves 2011

Hargreaves DJ, North AC. Experimental aesthetics and liking for music. In: Juslin PN, Sloboda JA, editors(s). Handbook of Music and Emotions: Theory, Research, Applications. New York: Oxford University Press, 2011:515-46. [ISBN: 978-0-19-923014-3]

Higgins 2011

Higgins JP, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from training.cochrane.org/handbook/archive/v5.1/.

Higgins 2017

Higgins JP, Altman DG, Sterne JA, editor(s). Chapter 8: Assessing risk of bias in included studies. In: Higgins JP, Churchill R, Chandler J, Cumpston MS, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 5.2.0 (updated June 2017). Cochrane, 2017. Available from www.training.cochrane.org/handbook.

Higgins 2019

Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page ML, et al, editor(s). Cochrane Handbook for Systematic Reviews of Interventions. 2nd edition. Chichester (UK): John Wiley & Sons, 2019. [ISBN: 9781119536628]

Hobeika 2022

Hobeika L, Ghilain M, Schiaratura L, Lesaffre M, Puisieux F, Huvent-Grelle D, et al. The effect of the severity of neurocognitive disorders on emotional and motor responses to music. *Annals of the New York Academy of Sciences* 2022;**1518**(1):231-8. [DOI: 10.1111/nyas.14923] [PMID: 36321882]

Hofbauer 2022

Hofbauer LM, Ross SD, Rodriguez FS. Music-based interventions for community-dwelling people with dementia: a systematic review. *Health & Social Care in the Community* 2022;**30**(6):2186-201. [DOI: 10.1111/hsc.13895]

Ing-Randolph 2015

Ing-Randolph AR, Phillips LR, Williams AB. Group music interventions for dementia-associated anxiety: a systematic review. *International Journal of Nursing Studies* 2015;**52**(11):1775-84. [DOI: 10.1016/j.ijnurstu.2015.06.014]

Jacobsen 2015

Jacobsen JH, Stelzer J, Fritz TH, Chételat G, La Joie R, Turner R. Why musical memory can be preserved in advanced Alzheimer's disease. *Brain* 2015;**138**(Pt 8):2438-50. [DOI: 10.1093/brain/awv135]

Jeon 2011

Jeon YH, Sansoni J, Low LF, Chenoweth L, Zapart S, Sansoni E, et al. Recommended measures for the assessment of behavioral disturbances associated with dementia. *American Journal of Geriatric Psychiatry* 2011;**19**(5):403-15. [DOI: 10.1097/JGP.0b013e3181ef7a0d]

Kitwood 1997

Kitwood T. Dementia Reconsidered. Bristol (PA): Open University Press, 1997. [ISBN: 9780335198559]

Lai 2014

Lai CK. The merits and problems of Neuropsychiatric Inventory as an assessment tool in people with dementia and other neurological disorders. *Clinical Interventions in Aging* 2014;**9**:1051-61. [DOI: 10.2147/CIA.S63504]

Lam 2020

Lam HL, Li WTV, Laher I, Wong RY. Effects of music therapy on patients with dementia - a systematic review. *Geriatrics* 2020;**5**(4):62. [DOI: 10.3390/geriatrics5040062]

Lepping 2024

Lepping RJ, Hess BJ, Taylor JM, Hanson-Abromeit D, Williams KN. Inconsistent music-based intervention reporting in dementia studies: a systematic mapping review. *Journal of Alzheimer's Disease* 2024;**100**(4):1145-59. [DOI: 10.3233/JAD-240255] [PMID: 38995790]

Li 2019

Li HC, Wang HH, Lu CY, Chen TB, Lin YH, Lee I. The effect of music therapy on reducing depression in people with dementia: a systematic review and meta-analysis. *Geriatric Nursing* 2019;**40**(5):510-6. [DOI: 10.1016/j.gerinurse.2019.03.017]

Lim 2024

Lim AWY, Schneider L, Loy C. Galantamine for dementia due to Alzheimer's disease and mild cognitive impairment. *Cochrane Database of Systematic Reviews* 2024, Issue 11. Art. No: CD001747. [DOI: 10.1002/14651858.CD001747.pub4]

Magee 2017

Magee WL, Clark I, Tamplin J, Bradt J. Music interventions for acquired brain injury. *Cochrane Database of Systematic Reviews* 2017, Issue 1. Art. No: CD006787. [DOI: 10.1002/14651858.CD006787.pub3]

McConnell 2016

McConnell T, Scott D, Porter S. Music therapy for end-of-life care: an updated systematic review. *Palliative Medicine* 2016;**30**(9):877-83. [DOI: 10.1177/0269216316635387]

McDermott 2013

McDermott O, Crellin N, Ridder HM, Orrell M. Music therapy in dementia: a narrative synthesis systematic review. *International Journal of Geriatric Psychiatry* 2013;**28**(8):781-94. [DOI: 10.1002/gps.3895]

McShane 2019

McShane R, Westby MJ, Roberts E, Minakaran N, Schneider L, Farrimond LE, et al. Memantine for dementia. *Cochrane Database of Systematic Reviews* 2019, Issue 3. Art. No: CD003154. [DOI: 10.1002/14651858.CD003154]

Microsoft Excel [Computer program]

Microsoft Excel. Version 2308. Microsoft Corporation, 2023. Available from https://office.microsoft.com/excel.

Muthesius 2010

Muthesius D, Sonntag J, Warme B, Falk M. Musik – Demenz - Begegnung: Musiktherapie für Menschen mit Demenz [Music - Dementia - Meeting: Music Therapy for People with Dementia]. Frankfurt am Main: Mabuse Verlag, 2010. [ISBN: 9783863214333]

Mühlbauer 2021

Mühlbauer V, Möhler R, Dichter MN, Zuidema SU, Köpke S, Luijendijk HJ. Antipsychotics for agitation and psychosis in people with Alzheimer's disease and vascular dementia. *Cochrane Database of Systematic Reviews* 2021, Issue 12. Art. No: CD013304. [DOI: 10.1002/14651858.CD013304.pub2]

Nelson 2011

Nelson JC, Devanand DP. A systematic review and meta-analysis of placebo-controlled antidepressant studies in people with depression and dementia. *Journal of the American Geriatrics Society* 2011;**59**(4):577-85. [DOI: 10.1111/j.1532-5415.2011.03355.x]

Norberg 1986

Norberg A, Melin E. Reactions to music, touch and object presentation in the final stage of dementia: an exploratory study. *International Journal of Nursing Studies* 1986;**40**(5):473-9. [DOI: 10.1016/S0020-7489(03)00062-2]

Page 2021

Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021;**372**:n71. [DOI: 10.1136/bmj.n71.]

Petrovsky 2015

Petrovsky D, Cacchione PZ, George M. Review of the effect of music interventions on symptoms of anxiety and depression in older adults with mild dementia. *International Psychogeriatrics* 2015;**27**(10):1661-70. [DOI: 10.1017/S1041610215000393]

Raglio 2012

Raglio A, Bellelli G, Mazzola P, Bellandi D, Giovagnoli AR, Farina E, et al. Music, music therapy and dementia: a review of literature and the recommendations of the Italian Psychogeriatric Association. *Maturitas* 2012;**72**(4):305-10. [DOI: 10.1016/maturitas.2012.05.016]

Reisberg 1987

Reisberg B, Borenstein J, Salob SP, Ferris SH, Franssen E, Georgotas A. Behavioral symptoms in Alzheimer's disease: phenomenology and treatment. *Journal of Clinical Psychiatry* 1987;**48 Suppl**:9-15. [PMID: 3553166]

Reisberg 1988

Reisberg B, Ferris SH, de Leon MJ, Crook T. Global Deterioration Scale (GDS). *Psychopharmacological Bulletin* 1988;**24**(4):661-3.

RevMan 2024 [Computer program]

Review Manager (RevMan). Version 7.12.0. The Cochrane Collaboration, 2024. Available at revman.cochrane.org.

Riecker 2000

Riecker A, Ackermann H, Wildgruber D, Dogil G, Grodd W. Opposite hemispheric lateralization effects during speaking and singing. *Neuroreport* 2000;**11**(9):1997-2000. [DOI: 10.1097/00001756-200006260-00038]

Sarafino 2008

Sarafino EP. Health Psychology: Biopsychosocial Interactions. 6th edition. Hoboken (NJ): Wiley, 2008. [ISBN: 9780470129166]

Schünemann 2023

Schünemann HJ, Higgins JP, Vist GE, Glasziou P, Akl EA, Skoetz N, et al. Chapter 14: Completing 'Summary of findings' tables and grading the certainty of the evidence. In: Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 6.4 (updated August 2023). Cochrane, 2023. Available from www.training.cochrane.org/handbook.

Sepehry 2012

Sepehry AA, Lee PE, Hsiung GY, Beattie BL, Jacova C. Effect of selective serotonin reuptake inhibitors in Alzheimer's disease with comorbid depression: a meta-analysis of depression and cognitive outcomes. *Drugs & Aging* 2012;**29**(10):793-806. [DOI: 10.1007/s40266-012-0012-5]

Sikkes 2021

Sikkes SAM, Tang Y, Jutten RJ, Wesselman LMP, Turkstra LS, Brodaty H, et al, ISTAART. Toward a theory-based specification of non-pharmacological treatments in aging and dementia: focused reviews and methodological recommendations. *Alzheimer's & Dementia* 2021;**17**(2):255-70. [DOI: 10.1002/alz.12188] [PMID: 33215876]

Sittler 2021

Sittler MC, Worschech F, Wilz G, Fellgiebel A, Wuttke-Linnemann A. Psychobiological mechanisms underlying the health-beneficial effects of music in people living with dementia: a systematic review of the literature. *Physiology & Behavior* 2021;**233**:113338. [DOI: 10.1016/j.physbeh.2021.113338]

Smith 2005

Smith SC, Lamping DL, Banerjee S, Harwood R, Foley B, Smith P, et al. Measurement of health-related quality of life for people with dementia: development of a new instrument (DEMQOL) and an evaluation of current methodology. *Health Technology Assessment* 2005;**9**(10):1-93, iii-iv. [DOI: 10.3310/hta9100]

Spintge 2000

Spintge R. Music and anesthesia in pain therapy [Musik in Anästhesie und Schmerztherapie]. *Anästhesiologie Intensivmedizin Notfallmedizin Schmerztherapie* 2000;**35**(4):254-61. [DOI: 10.1055/s-2000-10852-2]

Särkämo 2018

Särkämo T, Sihvonen AJ. Golden oldies and silver brains: deficits, preservation, learning, and rehabilitation effect of music in ageing-related neurological disorders. *Cortex* 2018;**109**:104-23. [DOI: 10.1016/j.cortex.2018.08.034]

Thompson 2024

Thompson N, Odell-Miller H, Underwood BR, Wolverson E, Hsu M-H. How and why music therapy reduces distress and improves well-being in advanced dementia care: a realist review. *Nature Mental Health* 2024;**2**(12):1532-1542. [DOI: 10.1038/s44220-024-00342-x]

Ting 2023

Ting B, Chen DT, Hsu WT, Liang CS, Malau IA, Li WC, et al. Does music intervention improve anxiety in dementia patients? A systematic review and meta-analysis of randomized controlled trials. *Journal of Clinical Medicine* 2023;**12**(17):5497. [DOI: 10.3390/jcm12175497] [PMID: 37685564]

Ting 2024

Ting B, Chen DT, Hsu WT, Tsai CL, Malau IA, Lee SL, et al. Multifaceted music therapy for depression in dementia: a network meta-analysis of randomized controlled trials. *European Journal of Investigation in Health, Psychology and Educacion* 2024;**14**(2):351-67. [DOI: 10.3390/ejihpe14020024] [PMID: 38391491]

Tsoi 2018

Tsoi KK, Chan JY, Ng YM, Lee MM, Kwok TC, Wong SY. Receptive music therapy is more effective than interactive music therapy to relieve behavioral and psychological symptoms of dementia:

a systematic review and meta-analysis. *Journal of the American Medical Directors Association* 2018;**19**(7):568-76. [DOI: 10.1016/j.jamda.2017.12.009]

Ueda 2013

Ueda T, Suzukamo Y, Sato M, Izumi S. Effects of music therapy on behavioral and psychological symptoms of dementia: a systematic review and meta-analysis. *Ageing Research Reviews* 2013;**12**(2):628-41. [DOI: 10.1016/j.arr.2013.02.003]

Van der Linde 2012

Van der Linde RM, Stephan BC, Savva GM, Dening T, Brayne C. Systematic reviews on behavioural and psychological symptoms in the older or demented population. *Alzheimer's Research & Therapy* 2012;**4**(4):28. [DOI: 10.1186/alzrt131]

Van der Linde 2016

Van der Linde RM, Dening T, Stephan BC, Prina AM, Evans E, Brayne C. Longitudinal course of behavioural and psychological symptoms of dementia: systematic review. *British Journal of Psychiatry* 2016;**209**(5):366-77. [DOI: 10.1192/bjp.bp.114.148403]

Verkaik 2007

Verkaik R, Nuyen J, Schellevis F, Francke A. The relationship between severity of Alzheimer's disease and prevalence of comorbid depressive symptoms and depression: a systematic review. *International Journal of Geriatric Psychiatry* 2007;**22**(11):1063-86. [DOI: 10.1002/gps.1809]

Volicer 2007

Volicer L, Bass EA, Luther SL. Agitation and resistiveness to care are two separate behavioral syndromes of dementia. *Journal of the American Medical Directors Association* 2007;**8**(8):527-32. [DOI: 10.1016/j.jamda.2007.05.005]

Waters 2022

Waters B, Sousa L, Orrell M, McDermott O. Analysing the use of music to facilitate social interaction in care home residents with dementia: narrative synthesis systematic review. *Dementia* 2022;**21**(6):2072-94. [DOI: 10.1177/14713012221100625]

Zhang 2017

Zhang Y, Cai J, An L, Hui F, Ren T, Ma H, et al. Does music therapy enhance behavioral and cognitive function in elderly dementia patients? A systematic review and meta-analysis. *Ageing Research Reviews* 2017;**35**:1-11. [DOI: 10.1016/j.arr.2016.12.003]

Zhang 2023

Zhang J, Yu Z, Zhang N, Zhao W, Wei B, He R, et al. Does music intervention relieve depression or anxiety in people living with dementia? A systematic review and meta-analysis. *Aging and Mental Health* 2023;**27**(10):1864-75. [DOI: 10.1080/13607863.2023.2214091]

Zhao 2016

Zhao QF, Tan L, Wang HF, Jiang T, Tan MS, Tan L, et al. The prevalence of neuropsychiatric symptoms in Alzheimer's disease: systematic review and meta-analysis. *Journal of Affective Disorders* 2016;**190**:264-71. [DOI: 10.1016/j.jad.2015.09.069]

References to other published versions of this review

Van der Steen 2017

Van der Steen JT, Van Soest-Poortvliet MC, Van der Wouden JC, Bruinsma MS, Scholten RJ, Vink AC. Music-based therapeutic interventions for people with dementia. *Cochrane Database of Systematic Reviews* 2017, Issue 5. Art. No: CD003477. [DOI: 10.1002/14651858.CD003477.pub3] [PMID: 28462986]

Van der Steen 2018

Van der Steen JT, Smaling HJ, Van der Wouden JC, Bruinsma MS, Scholten RJ, Vink AC. Music-based therapeutic interventions for people with dementia. *Cochrane Database of Systematic Reviews* 2018, Issue 7. Art. No: CD003477. [DOI: 10.1002/14651858.CD003477.pub4] [PMID: 30033623]

Vink 2002

Vink AC, Bruinsma MS, Scholten R. Music therapy in the care of people with dementia. *Cochrane Database of Systematic Reviews* 2002, Issue 1. Art. No: CD003477. [DOI: 10.1002/14651858.CD003477]

Vink 2003

Vink AC, Birks JS, Bruinsma MS, Scholten RJ. Music therapy for people with dementia. *Cochrane Database of Systematic Reviews* 2003, Issue 4. Art. No: CD003477. [DOI: 10.1002/14651858.CD003477.pub2]

* Indicates the major publication for the study

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Baker 2022

Study characteristics	
Methods	Cluster-RCT (parallel)
	"This large, pragmatic, multinational cluster-randomised controlled trial with a 2×2 factorial design" was to be run in Australia, Germany, the Netherlands, Norway, Turkey and the UK. This table relates to

Baker 2022 *(Continued)*

the Australian part of the study, which has been published. The ongoing study is also listed under Gold 2018 in Characteristics of ongoing studies.

Data were collected in Australia from 18 July 2018 until 17 March 2020. Due to COVID-19, the trial's conclusion was earlier than anticipated as recorded in the trial registration.

Participants	Country: Australia
	Nursing home residents were included in this study. Of 318 nursing home residents participating in one of four groups in a total of 20 units, 239 residing in 15 units participated in one of three included groups that represented single conditions (we did not include the fourth group that was provided two interventions).
	Eligible residents were diagnosed with dementia according to ICD-10 criteria, had depressive symptoms (8 or more on the Montgomery-Åsberg Depression Rating Scale) and CDR scores 0.5 to 2 (protocol), which was broadened later to 0.5 to 3, and MMSE 26 or less.
	In the music therapy group, 53/77 (69%) were women; in a choir singing group, 59/82 (72%) were women, and in a usual-care group, 55/80 (69%) were women. Mean ages were 86.0 (SD 7.5), 87.1 (SD 7.0) and 87.2 (6.5) years, respectively.
	Mean MMSE scores were 7.3 (SD 7.7), 8.2 (SD 7.6) and 9.2 (SD 8.1), respectively. About half (45%, 54% and 50%, respectively) had an unspecified type of dementia.
Interventions	Experimental group: 45-minute mixed (active and receptive) music therapy sessions were delivered by music therapists trained to deliver the sessions based on person-centred care principes. The intervention included singing, use of music instruments and music-evoked reminiscence. Sessions were delivered twice a week for 3 months, followed by a reduced frequency of once a week for 3 months.
	Control group 1: 45-minute recreational choir singing sessions based on person-centred principles delivered by trained community musicians. Sessions were delivered twice a week for 3 months, followed by a reduced frequency of once a week for 3 months.
	Control group 2: usual care
Outcomes	Primary outcome
	Depression measured with the Montgomery-Åsberg Depression Rating Scale at 6 months (end of treatment)
	Secondary outcomes included depression at 12 months (long-term outcome; up to 24 months according to the trial registration), overall behavioural problems measured with the NPI at 12 months (assessed also at 6 months), quality of life measured with the QoL-AD proxy and self-rated versions at 12 months (assessed also at 6 months) and also the generic EuroQol (EQ-5D-5L) rated by proxies. Cognition measured with the CDR at 12 months was proposed as a secondary measure despite the trial register reporting it was not measured at all sites; the protocol was amended on 26 September 2018 and the data were not reported. A number of other outcomes were measured, including serious and non-serious adverse events.
Notes	The author did not provide NPI-item level data upon our request. Christian Gold, one of the senior investigators of the MIDDEL STUDY, provided ICCs for various outcome measures to be used for the other cluster-RCTs. MADRS: ICC = 0.09 (in Baker 2022); QOLAD: ICC 0.010; NPI severity: ICC 0.06. Mean ICC to be used for other outcomes: 0.09.
	English was the first language for about three-quarters of residents (75%, 77% and 78%, respectively).
	A factorial design was employed, with a fourth group receiving both music therapy and recreational choir singing sessions. The trial was designed as a pragmatic trial; residents attended an average of 22.2 (SD 13.1) group music therapy sessions, or 20.0 (SD 13.3) recreational choir singing sessions. The analyses were conducted in an intention-to-treat sample.
	Patient and public involvement: in Australia, nursing home residents and family caregivers were involved in developing the research question, design and pilot testing of the study. In the UK and Den-

Baker 2022 *(Continued)*

mark, advocacy organisations considered the need for research, the interventions and outcomes. A trial steering committee involved service users or their relatives and representatives of stakeholder organisations such as Alzheimer Europe and Dementia Australia. Gold 2019 further reports planned continued involvement of user representatives.

Funding: the National Health and Medical Research Council, Australia funded this study. Protocol development was supported by the same source, and, additionally, by intramural support from Uni Research AS Norway and University Medical Center Groningen, the Netherlands

Conflict of interest: the authors declare that they have no conflict of interest (Gold 2019; Baker 2019; Baker 2020; Baker 2022).

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	A computer-generated randomisation list was used for block randomisation with a block size of four.
Allocation concealment (selection bias)	Low risk	"Four care home units were randomised at the same time (forming a wave), to ensure allocation concealment."
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible to blind the convener and participants. This was acknowledged: "Following assignment, interventionists, residents, and care staff could no longer be masked (study report, Baker 2022). "Blinding will be difficult to achieve. Intervention providers and study participants cannot be blinded to the intervention they receive or provide. However, participants may be unaware of the specific differences between GMT [group music therapy) and RCS [recreational choir singing]" (protocol, Gold 2019).
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	"All seven assessors were masked to treatment allocation until all follow-up assessments were completed and masking success was recorded" (Baker 2022). "Blinding of assessors (those evaluating outcomes) will be attempted by using assessors external to the care homes, but this may be incomplete because they will have to rely on information from proxy informants (care staff who know the participant ell) due to the inability of most residents to report on themselves. Assessors will remind informants not to reveal the unit's allocation to them. At the time of the last assessment, success of blinding will be verified by asking assessors whether they inadvertently discovered the unit's allocation" (Gold et al, 2019). "MADRS [Montgomery-Åsberg Depression Rating Scale] was scored by an external masked assessor but this scoring relied partly on information from care staff. Residents' secondary outcomes included neuropsychiatric symptoms and quality of life from unmasked proxy (care staff) or self-reports" (Baker 2022).
Incomplete outcome data (attrition bias) All outcomes	High risk	Substantial attrition, in part, COVID-19 related. Outcomes were also reported for the subsample with ≥ 50% attendance (Baker 2022).
Selective reporting (reporting bias)	Unclear risk	Trial registration: registration numbers NCT03496675 and AC-TRN12618000156280. The registry included depression up to 24 months as an outcome; reports in Baker 2022 were up to 12 months only. Cognition measured by the CDR=Clinical Dementia Rating (CDR) was omitted as a secondary outcome after a protocol amendment and described at baseline only.

Baker 2022 *(Continued)*

Other bias	Unclear risk	Unclear to what extent COVID-19 and measures taken affected the study

---

Ceccato 2012

Study characteristics	

Methods	RCT (parallel)
	No information on data collection period reported

Participants	Country: Italy
	The study was conducted in 5 "support centres". We assumed these were situated in nursing homes, given author affiliations and locations of support centres.
	Dementia diagnosis: formally diagnosed with the DSM-IV. Inclusion criterion was MMSE score from mild (MMSE 18 to 24) to moderate (MMSE 12 to 18). People with acute medical illness were excluded, and a number of additional inclusion criteria applied, including being "sensitive to sound/musical stimuli;" "the desire and capacity to remain in the setting;" "presence of sufficient (also residual) hearing and perceptive-communicative and relational skills."
	Type of dementia: not reported
	51 people with dementia and 50 of them were included in analyses (1 had only pretest data); experimental group: 28 participants (27 in analyses; 21 women); control group: 23 participants (19 women)
	Mean age: experimental group: 85.5 (SD 5.9) years; control group: 87.2 (SD 7.1) years

Interventions	Experimental group: Sound Training for Attention and Memory in Dementia ("STAM-Dem"). Mixed active-receptive group intervention with 24 sessions of 45 minutes in 12 weeks. STAM-Dem includes 4 phases: 1) stimulus-movement association, 2) reaction to acoustic stimuli, 3) shifting attention and 4) orderly and inverted repetition. The intervention combines listening to music, clapping hands, tapping the table and repeating sounds. The professional music therapists were trained to administer the STAM-Dem protocol. Supervision was provided throughout the course of the intervention by the protocol's author.
	Control group: normal "standard care" provided

Outcomes	Primary outcome
	• Cognitive functioning measured with MMSE, attentional matrices, forward and reverse digit-span exercise, MPI test and MPD test
	Secondary outcomes
	• Behaviour measured with the CMAI. Timeframe of CMAI was last 2 weeks
	• Mood measured with GDS
	• ADL was measured with the Index of Independence in Activities of Daily Living (ADL) by nurses, adequacy 6 functions.
	• Some other outcomes may have been measured only in the STAM-Dem group.
	• Follow-up was planned but not carried out. No follow-up was conducted after the intervention because of a lack of funding.

Notes	Randomisation was done separately for each centre (6 randomisations in total). This is also the reason why there were more people in the experimental group (28 participants) compared with the control group (23 participants).
	Patient and public involvement: no information

Ceccato 2012 *(Continued)*

Funding: F.S. Zerbato Centre at Tregnago (president, director and manager)

Conflict of interest: the authors declared that 7 of the 14 authors (who were named) are professional music therapists.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "They were divided up using an online randomization program by personnel not involved in the study, thereby ensuring totally ″blind″ conditions."
Allocation concealment (selection bias)	Low risk	An online randomisation programme was used by personnel not involved in the study.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible to blind the convenor and participants.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Quote: "Pre- and postintervention testing was also administered by professionals who had no other role in the project; blind conditions were thus obtained for assignment treatment."
Incomplete outcome data (attrition bias) All outcomes	Low risk	1 participant dropped out and 1 participant had no post-test data. Unclear if this was the same participant as the number allocated to the intervention group was incorrect in the figure.
Selective reporting (reporting bias)	Unclear risk	The investigators did not follow the plans here: no follow-up conducted after the intervention because of a lack of funding. We cannot verify if there were other factors that played a role, such as perceived limited benefit.
Other bias	Low risk	No other bias detected

Cho 2016

Study characteristics

Methods	RCT (parallel) with 3 groups Intervention provided in October 2015, for 4 weeks
Participants	Country: USA. Veterans Affairs skilled nursing home facility 52 people with dementia were randomised, and 35 or 36 (for different outcomes) were included in the analyses (experimental group: 14; control group 1: 14; control group 2: 7 for quality of life and 8 for affect outcomes). Residents were included when they had a diagnosis of dementia, were aged ≥ 65 years, had no significant hearing impairment and were able to sit in a chair or wheelchair for ≥ 1 hour. Residents with severe psychiatric conditions, or receptive or expressive language problems were excluded. Age, mean (SD), range: experimental group: 85.1 (SD 8.7), 67 to 99 years; control group 1: 87.9 (SD 5.9), 75 to 98 years; control group 2: 87.0 (SD 6.0), 74 to 97 years. There were only 3 women in each of the 3 groups of experimental group: 18; control group 1: 17; control group 2: 17.

Cho 2016 *(Continued)*

Mean BIMS scores (SD): experimental group: 10.2 (SD 4.4); control group 1: 10.2 (SD 4.0); control group 2: 9.9 (SD 3.6) (BIMS scores 8 to 12 refer to moderate impairment). Type of dementia was not reported. All participants were "Caucasian".

Interventions	Experimental group: music therapy-singing group: by "a music therapist with over 15 years of experience with dementia care."
	Control group 1: music listening group by nursing home activity assistants (for the purpose of our review, we regarded this as a control condition). The assistants "did not have same level of training as the music therapist, especially in facilitating a group process."
	Control group 2: TV watching group: control condition, watching a DVD
	All 3 groups ran 8 × 40-minute sessions in a period of 4 weeks (twice a week).
Outcomes	Outcome: quality of life (QoL-AD). Quality of life was assessed directly from the person with dementia. It was evaluated twice, once before the first intervention session and once after the last (8th) intervention session.
	An additional research question referred to differences in quality and affect over time between the 3 BIMS categories.
Notes	Specific population (more men than usual in nursing home populations)
	Randomisation was stratified by dementia severity (mild, moderate, severe, based on BIMS score).
	Other outcomes were general positive affect and negative affect measured with the PANAS.
	Patient and public involvement: no information
	Funding (author personal communication): institutional support with no external funding
	Conflict of interest: no information

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	For the random assignment, the list of participants was given to another nursing home activity assistant with specially assigned numbers in place of the participants' names.
Allocation concealment (selection bias)	Low risk	The participants' names were not revealed to the nursing home activity assistant who was responsible for the random assignment until the randomisation process was completed to ensure allocation concealment. The nursing home activity assistant randomly assigned participants to 1 of the 3 conditions within each stratum of the BIMS score using a random number table from a statistical text book.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible to blind the convener and participants
Blinding of outcome assessment (detection bias) All outcomes	High risk	Nursing home activity assistants who were involved in assessing the outcomes were not blinded.
Incomplete outcome data (attrition bias) All outcomes	High risk	Of the 17 participants who were assigned to the control (TV watching) group, only 8 (47%) completed the intervention. Dropout in this group was larger than

Cho 2016 *(Continued)*

for the other groups (with 83% in music therapy-singing group and 82% in music listening group completed).

Quote: "Furthermore, the participants' preferences for the TV group were not assessed, whereas music programs for singing and listening group were created based on their music preferences. This may have closely related to the inconsistent results regarding affect in the TV group, as well as the highest dropout rate of participants assigned to the TV group. Out of 17 participants who were assigned to the TV group, nine dropped out over the course of the study, and only eight completed the intervention."

Selective reporting (reporting bias)	Unclear risk	The study was not registered.
Other bias	Unclear risk	Music therapist sole author of first publication of the study (dissertation)

Clark 1998

Study characteristics	
Methods	RCT (cross-over 2 weeks + 2 weeks)
	No information on data collection period reported
Participants	Country: USA
	18 participants, (14 women, 4 men)
	Inclusion criteria: presence of dementia and a history of aggressive behaviour exhibited during care giving routines
	Exclusion criteria
	• Uncorrected hearing impairment
	• Absence of family member who could provide knowledge of a potential participant's music preferences
	Mean age: 82 (range 55 to 95) years, residents in a nursing home with Alzheimer-type dementia. Presence of dementia was assessed with the MMSE (mean 10, range 0 to 22). Type of dementia was not reported. Most residents had severe dementia.
Interventions	Experimental group: favourite music during bathing (receptive intervention)
	Control group: no music during bathing
	Following a 2-week (10 sessions) observation period, conditions were reversed. A total of 20 sessions (bathing episodes; 10 control, 10 experimental) were observed over a period of approximately 4 weeks. Probably the intervention was provided for all bathing episodes and all were observed.
Outcomes	Behaviour: frequency of aggressive behaviours (no specific measure was used, but counts and mean counts across specific behaviours)
Notes	The study also included younger people with dementia.
	Patient and public involvement: no information
	No information about funding available.
	Conflict of interest: no information

Clark 1998 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: "After being enrolled in the study, participants were randomly scheduled for observation during bath time under either a control (no music) condition or an experimental condition." No further information provided on randomisation
Allocation concealment (selection bias)	Unclear risk	No information provided
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible to blind the convener and participants
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	No information provided
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	No information provided
Selective reporting (reporting bias)	Unclear risk	Study protocol not available
Other bias	High risk	Questionable outcome measure and distribution. The authors reported in the article on the effects of the extreme intrasubject and intersubject variability characteristic of this population in this study. Quote: "For example, one subject was responsible for 408 and 84 occurrences of yelling behaviour in the no music and music conditions, respectively." Therefore, highly skewed distributions (the observation hardly occurred) causing imprecision.

Cooke 2010

Study characteristics

Methods	RCT (cross-over) Data collection from October 2008 to March 2009
Participants	Country: Australia 2 mixed-gender long-term care facilities, which provided low (assisted living) and high (nursing home) care 47 participants (33 women and 14 men) Dementia diagnosis: a confirmed diagnosis of early- to mid-stage dementia, OR probable dementia (i.e. a cognitive impairment level of 12 to 24 on MMSE) OR AD according to DSM-IV criteria. Participants had "a documented behavioural history of agitation/aggression on nursing/medical records within the last month."

Cooke 2010 *(Continued)*

Type of dementia was not reported.

Age: 3 people aged 65 to 74 years, 13 aged 75 to 84 years, 28 aged 85 to 94 years and 3 people ≥ 95 years. At baseline, the mean MMSE score was 16.51, representing middle-stage dementia (SD 6.737).

Interventions	Experimental group: active live group music programme (30 minutes per session) and listening to pre-recorded instrumental music (10 minutes per session) led by 2 musicians
	Control group: reading group chosen as the control group activity so as to provide a comparable activity. The facilitator of the 40-minute sessions was a trained research assistant.
	Both the active group music programme and the control activities ran 3 mornings a week (Monday, Wednesday and Friday) for 8 weeks, and the facilitators were trained in the delivery of the sessions and in working with older people with dementia.
Outcomes	**Primary outcome** • Agitation measured with the CMAI-SF and overall and subscale scores were reported for a modified 14-item short form. Timeframe: previous 2 weeks **Secondary outcomes** • Anxiety measured with the RAID. Timeframe: previous 2 weeks • Quality of life measured with DQOL using overall and subscale score • Depression measured with GDS • Outcomes measured at baseline, mid-point (after the first 8-week intervention arm) and postintervention (after the second 8-week intervention arm)
Notes	Patient and public involvement: no information Funding: funded by the National Health & Medical Research Council, Australia Conflict of interest: "none declared".

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "The randomisation process was conducted by the study's biostatistician, who was blinded to the identity of potential participants, using a computer-generated programme."
Allocation concealment (selection bias)	Low risk	Quote: "The randomisation process was conducted by the study's biostatistician, who was blinded to the identity of potential participants, using a computer-generated programme."
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible to blind the convener and participants
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Quote about CMAI-SF: "Aged care staff who provided most care to the participant, but blinded to treatment groups, were asked to rate the ..." Quote about RAID: "Research assistants (RAs) blinded to the treatment groups asked participants to rate, on a scale from '1 = absent' to '3 = severe,' how often he/she had experienced each symptom in the previous two weeks." Research assistant completed DQOL and GDS.

Cooke 2010 *(Continued)*

		Quote: "Both measures were conducted by trained RAs blinded to the treatment groups at a time most convenient for the participant (i.e. any day of the week from 9am–5pm). The RAs took the role as interviewer, taking the participants through the measures by asking them questions to elicit their response."
Incomplete outcome data (attrition bias) All outcomes	Low risk	Prior to all sessions, participants were asked if they wished to attend. This resulted in some refusals and differences in attendance levels among participants. Following a missing values analysis, which indicated data to be missing at random, an ITT analysis, in which all 47 randomised participants were included, was undertaken. Missing values in the outcome measures were imputed with multiple imputation methods.
Selective reporting (reporting bias)	High risk	Inconsistencies compared with the trial registration, which was retrospectively registered in 2012. Number of registration therefore not in article. Registration pointed to anxiety as a secondary outcome, not a primary outcome. Moreover, quality of life and depression were not reported as secondary outcomes.
Other bias	Low risk	No other bias detected

Giovagnoli 2018

Study characteristics	
Methods	RCT (parallel) Total study duration or begin and end dates are not reported
Participants	Country: the study was conducted in Italy as confirmed by the authors. The exact setting is unclear. After hospitalisation, participants with probable AD according to DSM-IV-TR and NINCDS-ADRDA revised criteria and with impaired language were recruited. Further, eligible patients were on stable treatment with acetylcholinesterase inhibitors for at least 4 weeks. Forty-five participants were randomly assigned to the experimental condition (mean age 74.3, SD 5.7; mean MMSE at baseline 16.59, SD 4.01; 16 of 23 women) or control condition (mean age 72.0, SD 7.3; mean MMSE at baseline 16.24, SD 4.10; 15 of 22 women).
Interventions	Experimental group: 40-minute active group music therapy sessions delivered by a music therapist, using a non-verbal approach. Memantine 20 mg/day was added to treatment as usual with cholinesterase inhibitors. The sessions were delivered twice a week for 24 weeks Control group: memantine 20 mg/day was added to treatment as usual with cholinesterase inhibitors
Outcomes	The secondary outcomes included cognition measured with the MMSE and behavioral problems overall measured with the NPI; NPI item scores of anxiety, depression and agitation were reported as well. SIB subscales ratings were reported including the social interaction subscale. Adverse events, somnolence, insomnia and depression were reported as well.
Notes	Language was the primary outcome. Agitation was reported in the article as "stirring." The authors explained it was a translation error and the reported values represent the NPI agitation item. Patient and public involvement: no information Funding: The study was supported by the Italian Ministry of Health and Lundbeck (Valby, Denmark).

Giovagnoli 2018 *(Continued)*

Conflict of interest: no information

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"The randomization was made using a computer-generated list of random numbers"
Allocation concealment (selection bias)	Unclear risk	Unclear who randomised although randomisation was after the baseline assessment ("Six to 8 months after the prescreening, consenting patients underwent the baseline assessment and were randomly assigned to a study group.")
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible to blind the convener and participants
Blinding of outcome assessment (detection bias) All outcomes	Low risk	"The patients were evaluated blindly by a neuropsychologist at baseline and at weeks 12 and 24."
Incomplete outcome data (attrition bias) All outcomes	Low risk	All 45 participants randomised were included in the analyses.
Selective reporting (reporting bias)	Unclear risk	No reference to a protocol or trial registration
Other bias	Low risk	No other bias detected

Gómez-Gallego 2021

Study characteristics

Methods	Cluster-RCT (parallel)
	The study was conducted between September and December 2018 (trial registration).
Participants	Country: Spain
	Nursing homes were randomised to three groups, and 90 or 92 nursing home residents with probable Alzheimer's dementia were included. Residents with aphasia or hearing impairment that "may affect participation in the activities" were excluded.
	The residents in the experimental group (71.5% women; 28 to 30 residents) had a mean age of 83.93 years (SD 8.01) years; residents in the music listening group (61.9% women; 21 residents) had a mean age of 78.67 (SD 5.73) year; and residents in the video watching group (54.5% women; 41 residents) had a mean age of 80.02 years (SD 5.78).
	Measured with the CDR scale, 70% of the total sample had mild dementia and 30% had moderate dementia; dementia severity was part of the inclusion criteria.
Interventions	Experimental group: active group music therapy sessions of 45 minutes delivered by a music therapist. Sessions included songs, music, movements and a music quiz. According to the article, it was provided "twice a week for 3 months (12 sessions in total)." Note an incongruency; the trial registry also mentions a duration of 3 months.

Gómez-Gallego 2021 *(Continued)*

Control group 1: 45-minute group music listening intervention based on an audio-recording of the same songs as used in the experimental group

Control group 2: in a group setting, the participants watched nature videos without music of nearly the same duration as the interventions.

Outcomes	Depression was measured with the GDS, overall behavioral problems measured with the NPI; cognition with the MMSE. All (including also ADL and functioning) were reported in the trial registry as primary outcome measures. Adverse events were not an outcome, but "When necessary, facilitators adapted their interventions to the needs of residents to avoid adverse reactions."
Notes	The frequency and total number of sessions do not correspond and remain unclear. Further, the total number of participants in the experimental group may be 30 (when adding up 6, 7, 8 and 9 residents in four therapy groups), 29 (one hospital admission) or 28 (in table). We departed from 28 as the smallest sample, assuming there were 28 cases in the analyses.
	Sample size adapted to adjust for clustering. ICCs applied from Baker 2022.
	Patient and public involvement: no information
	The article reports the study did not receive external funding.
	Conflict of interest: "The authors declare no conflict of interest."

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Random sequence generation not reported
Allocation concealment (selection bias)	Low risk	"Randomization was performed by an independent external researcher"
Blinding of participants and personnel (performance bias) All outcomes	High risk	The trial register reports double masking (participant and investigator). However, it is not possible to blind the convener and participants.
Blinding of outcome assessment (detection bias) All outcomes	High risk	"Information bias is possible since the administration of the instruments was not blinded."
Incomplete outcome data (attrition bias) All outcomes	Low risk	Only a few cases (0, 1 or 2; exact number unclear) could not be included in the analyses.
Selective reporting (reporting bias)	Unclear risk	The study has been registered (ID: NCT04761497). However, the study completion date 7 January 2019 reported in the trial register was before the Institutional Review Board approval date of 3 April 2019 reported in the main reference. Further, the trial was first registered only on 21 February 2021.
Other bias	Low risk	No other bias detected

Guétin 2009

Study characteristics

Guétin 2009 *(Continued)*

Methods	RCT (parallel); total duration 18 months, with a follow-up period of 6 months
	Participants resided in the nursing home between September 2007 and April 2008.

Participants	Country: France
	30 participants (22 women, 8 men), 1 centre
	Diagnosis of dementia: mild to moderate stage of AD
	Inclusion criteria
	• MMSE score 12 to 25 and Hamilton Anxiety Scale score ≥ 12
	• At baseline, MMSE mean score 19.8 (SD 4.4) for experimental group and 20.7 (SD 3.4) for control group
	Exclusion criteria
	• Major depressive disorder or other major psychiatric disorders
	• "patients considered highly likely not to comply with the protocol or to drop out of the study as well as those suffering from a life-threatening illness during the envisaged study period."
	Mean age: experimental group: 85.2 (range 75 to 93) years; control group: 86.9 (range 74 to 95) years

Interventions	Experimental group: individual receptive music therapy method, the 'U-sequence method,' which involved listening to music sequences, selected from preferred musical styles delivered through headphones, in the participant's room.
	Control group: reading sessions
	Weekly sessions for 16 weeks (total of 16 sessions)

Outcomes	Level of anxiety (HAM-A; total score 0 to 56)
	Level of depression (GDS; maximum score 30)
	MMSE score
	Outcomes assessed at day 0, and weeks 4, 8, 16 and 24 by an independent neuropsychologist assessor. Long-term outcomes were assessed 8 weeks after treatment ended.

Notes	Patient and public involvement: no information.
	Funding: this research could be carried out thanks to support from Centres Mémoire de Ressources et de Recherches, Les Violettes nursing home, Université René Descartes – Paris V, Institut Alzheimer, the Rotary Club and La Fondation Médéric Alzheimer.
	Conflict of interest: no information

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Probably yes, but no details provided
		Quote: "The study design corresponded to a randomised, controlled, comparative, single-centre study, with the results evaluated under blind conditions."
		Quote: "The patients were allocated to the different groups by randomisation at the end of the inclusion visit."
Allocation concealment (selection bias)	Unclear risk	No details provided

Guétin 2009 *(Continued)*

Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible to blind the convener and participants
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Participants and carers not blinded, outcome assessor blinded

Quote: "The results obtained at D0 [day], W4 [week], W8, W16 and W24 were collected by an independent neuropsychologist assessor (D.L.), not belonging to the care team and unaware of the type of intervention." |
| Incomplete outcome data (attrition bias) All outcomes | Unclear risk | Unclear whether dropouts caused bias

Quote: "Two patients were prematurely withdrawn from the study in the intervention group: 1 between W8 [week] and W16 owing to an intercurrent event not related to the study (life-threatening situation, hospitalisation), and the second died between W16 and W24. Four patients were withdrawn from the study in the control group: 1 between W4 and W8 due to dropping out, 1 between W4 and W8 owing to an intercurrent event not related to the study (hospitalisation), 1 patient died between W4 and W8, and the last patient dropped out between W16 and W24." |
| Selective reporting (reporting bias) | Unclear risk | No study protocol available |
| Other bias | Low risk | Baseline imbalances do not appear to have caused bias. |

Hsu 2015

Study characteristics	
Methods	Mixed quantitative-qualitative feasibility study, which included a parallel cluster-RCT (randomised at nursing home unit level)

Study took place from February to September 2013. |
| Participants | Country: UK

Nursing home residents with dementia (17 randomised; 13 contributed to the analyses) but also 10 staff from 2 nursing homes (see Notes)

Experimental group: 9 participants; control group: 8 participants

All diagnoses were made in accordance with the DSM-5. Other inclusion criteria, residents:

- presented with ≥ 2 neuropsychiatric symptoms of dementia
- aged ≥ 40 years
- no significant health problems

Women: experimental group: 89%; control group: 100%. Mean age: experimental group: 84.6 (SD 6.6) years; control group: 82.5 (SD 13.0) years. Overall range 56 to 98 years.

Almost half of the participants (41%) were diagnosed with dementia of AD type. The remaining residents had diagnoses of vascular, frontal lobe, Lewy Body and mixed type of dementia, while for 18% of the participants, the dementia diagnosis was unspecified. Mean Global Deterioration Scale: experimental group: 5.89 (SD 1.05); control group: 5.50 (SD 1.31) |

Hsu 2015 *(Continued)*

Interventions	Experimental group: individual active music therapy and training of care staff. Music therapists delivered the intervention consisting of individual active music therapy sessions in combination with training of care staff using video clips of the sessions. The sessions were delivered once a week for 5 months, in addition to standard care. Control group: "standard care." This consisted of medical and personal care, provision of basic needs, and activities carried out as usual within the home such as chaplaincy services, entertainment and leisure activities.
Outcomes	• Well-being: well-being score from DCM • Overall behavioural problems and its and disruptiveness, both measured with the NPI-NH (In addition, there were outcomes other than the 7 outcomes of interest for this review.) Long-term outcomes were assessed 2 months after treatment ended.
Notes	The author provided data upon our request. Sample size adapted to adjust for clustering. ICC applied from Baker 2022. Patient and public involvement: reported no involvement in study design ("The study design was developed in consultation with academic and clinical professionals".) Funding: Methodist Homes in Derby and Anglia Ruskin University Conflict of interest: "The authors declare that they have no competing interests".

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Randomisation by computer programme
Allocation concealment (selection bias)	Low risk	Cluster randomisation (between units) to reduce contamination across the control and intervention groups. After participants had been recruited by the researchers, randomisation was conducted by the study statistician independently of the researchers.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible to blind the convener and participants
Blinding of outcome assessment (detection bias) All outcomes	High risk	Outcome assessment was unblinded.
Incomplete outcome data (attrition bias) All outcomes	High risk	3/9 participants of the experimental group died versus 1/8 in the control group. They were excluded from all analyses.
Selective reporting (reporting bias)	High risk	Differences with trial registration (reported versus registration): secondary outcome was indicated as secondary only in the trial register. Moreover, there was no mention of disruptiveness as an outcome in the register. Clinicaltrials.gov number: NCT01744600
Other bias	Low risk	No other bias detected

Liesk 2015 *(Continued)*

Study characteristics

Methods	RCT (parallel)
	No information on data collection period reported

Participants	Country: Germany
	5 nursing homes
	Diagnosis of dementia: partly formally diagnosed with ICD-10 and partly not formally diagnosed. People with mild-to-moderate dementia were included. People with vision or hearing impairment or life-threatening illness were excluded.
	26 participants with dementia randomised. 2 had no complete baseline data, and 24 (12 in each group) were included in analyses. In each of the groups, there were 11 women and 1 male participant. Mean age: experimental group: 83.6 (SD 5.1; range 72 to 89) years; control group: 84.3 (SD 5.4; range 70 to 90) years. Dementia type was largely unknown in the setting

Interventions	Experimental group: active group music intervention 'Musikgeragogik' which included singing folk songs and canons and instrumental performance, 12 sessions of 90 minutes in 6 weeks.
	Control group: cognitive stimulation intervention: adapted cognitive training programme from NEUROvitalis, 12 sessions of 90 minutes in 6 weeks.

Outcomes	Cognition measured with the MMST, DemTect (and subscales), MTF/ROF, MAC-Q (German version, "Selbteinschatzung-Gedachtnis"), Trail Making Test A, FAS Test (Controlled-Oral-Word-Association Test), BTA.
	Quality of life measured with DEMQOL and DEMQOL-Proxy (no full name, developed by Smith and colleagues; Smith 2005).
	ADL measured with the Barthel Index, IADL and ADL (Aktivitaten des taglichen Lebens).
	Also the NOSGER was measured, but it is unclear for which outcome.
	Outcomes were measured at baseline (before randomisation) and 1 or 2 days after the last session.

Notes	No explanation about the instruments that were used. The instruments were only mentioned in the table with results. Unknown for which outcome the NOSGER observation scale was used.
	Low fidelity in music intervention group.
	Bottom effect cognitive measure and more problems described (also in Discussion section of the article) which was part of the goal of the article.
	Patient and public involvement: reported no involvement in the adaptation of the intervention as this done by author JL, gerontologist.
	No information about funding reported.
	Conflict of interest: the author reported that there is no conflict of interest.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "Die randomisierte Zuteilung der Programme auf die Einrichtungen fand computergestutzt statt." This means that randomised computer-assisted

Liesk 2015 *(Continued)*

allocation of the programmes (at the level of individuals with dementia) was performed at the facilities.

Allocation concealment (selection bias)	Unclear risk	No description about allocation concealment
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible to blind the convener and participants
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Unclear who administered the instruments and whether these people were blinded for the intervention type
Incomplete outcome data (attrition bias) All outcomes	Low risk	Few participants missed outcome data and this was clearly reported.
Selective reporting (reporting bias)	Unclear risk	No research protocol available
Other bias	High risk	Participants in the control group frequently developed an acute illness, which resulted in missing sessions. Quote: "Während keiner der 12 Teilnehmer des MP akut erkrankte, fielen 5 der 12 Teilnehmer des KS zwischen zwei und vier Sitzungen aus." (While none of the 12 participants in the music intervention group became acutely ill, 5 of the 12 participants in the cognitive stimulation group missed 2–4 sessions.) People who attended fewer than 8 of 12 sessions were excluded from the analyses, so these people still contributed to outcome data. Therefore, adherence or fidelity may be a problem even though they already preselected people who were probably interested in music therapy.

Lin 2011

Study characteristics

Methods	RCT (parallel) Data collection between August 2008 and January 2009
Participants	Country: Taiwan 3 nursing home facilities Diagnosis of dementia: participants had been diagnosed by a physician as having dementia, using the DSM-IV-TR. Of 104 included people with dementia (52 per group), 100 participants (experimental group: 49 participants; control group: 51) were included in analyses (53% women in total group; experimental group: 53.06% women; control group: 52.94%). Mean age: overall: 82 (range 65–97, SD 6.80) years; experimental group: 81.46 years; control group: 82.15 years. Type of dementia was not reported

Lin 2011 *(Continued)*

Interventions	Experimental group: mixed active-receptive music group intervention modified of the protocol developed by Clair and Bernstein (Clair 1990), 12 sessions of 30 minutes in 6 weeks; provided by a music therapist.
	Control group: continued to engage in their normal daily activities.
Outcomes	Physically non-aggressive behaviours, physically aggressive behaviours, verbally non-aggressive behaviours and verbally aggressive behaviours were measured with C-CMAI. The instrument rates a person's agitated behaviour and its frequency over the previous 2 weeks. The C-CMAI includes 29 items, each rated on a 7-point scale (1–7) ranging from never (1 point) to several times an hour (7 points), with a total score of 29 (minimum) to 203 (maximum). CMAI frequency referred to the previous 2 weeks.
	Depression measured with the C-CSDD.
	Cognition was measured with the C-MMSE.
	These outcomes were measured by another member of the research team in the experimental and control groups at baseline (1 week before start intervention), immediately after 6th and 12th sessions, and at 1 month after cessation of the intervention.
	Cortisol levels were used as a biomarker for depression and were measured at baseline, immediately after 6th and 12th sessions.
Notes	Patient and public involvement: no information.
	Funding: no information provided.
	Conflict of interest: "none declared."

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "subjects consisted of a total of 104 elderly persons who were randomly assigned to the experimental (n = 52) and control group (n = 52) by permuted block randomization." (p 671, Lin 2011) and "permuted block randomisation computer-based program" (p 672, Lin 2011).
		Quote: "Using permuted-block randomisation, a separate researcher randomized participants into the experimental or usual-care control group within each nursing home. We determined blocked randomization with a block size of 26 using the Research Randomizer computer program, which generates a list of random numbers to be used for allocating participants to the two groups. We generated the allocation sequence with the Research Randomizer program prior to the recruitment of participants and…" (Chu 2014, see under Lin 2011).
Allocation concealment (selection bias)	Low risk	Quote: "participants and …(continued) concealed the results in sequentially numbered and sealed opaque envelopes, which we opened when participant were ready for allocation. After four randomization series, we assigned the 104 participants to the experimental or control condition in a blinded manner" (Chu 2014, see under Lin 2011).
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible to blind the convener and participants
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Unclear who reported the C-CMAI
		However, Chu 2014 (see under Lin 2011) described that the C-CSDD and MSSE were reported by another member of the research team.

Lin 2011 *(Continued)*

		Quote: "Another member of the research team administered the study instruments 1 week before the start of the intervention (Time 1), immediately following the 6th (Time 2) and 12th (Time 3) sessions of the intervention, and 1 month after the final intervention session (Time 4) and collected salivary cortisol samples at Times 1–3. The same person administered the instruments each time" (Chu 2014, see under Lin 2011).
Incomplete outcome data (attrition bias) All outcomes	Low risk	Few cases lost to follow-up, and only 1 in the experimental group was not interested.
Selective reporting (reporting bias)	Unclear risk	No protocol available
Other bias	Low risk	No other bias detected

Liu 2021

Study characteristics	
Methods	RCT (parallel)
	No information on data collection period reported
Participants	Country: Taiwan
	The 50 participants in the study resided in veterans' homes in a long-term stay setting or dementia centre.
	Only male residents aged 75 years or older were included. All met the NINCDS-ADRDA diagnostic criteria for probable AD. Only those with a mild or moderate stage of dementia were included (CDR scale scores 1 and 2). There were also several inclusion criteria related to communication capacity. Further, only those with anxiety scores below 17 on the HAM-A were included and with stable psychotropic or anxiolytic treatment for 3 months.
	The average age was 86.6 (SD 4.5) years for the intervention group and 86.9 (SD 5.7) years for the control group.
	Type of dementia was unknown because not all residents had brain imaging diagnostics.
Interventions	Experimental group: the 25 residents randomised to the experimental condition received a 60-minute active group music session from a trained "music facilitator" once a week for 12 weeks. The main part of the session was using various percussion instruments to play familiar songs.
	Control group: the 25 residents randomised to the control condition received "rest and reading" sessions according to the same schedule and under the same conditions as the experimental group.
Outcomes	Outcomes were depression measured with the Geriatric Depression Scale (GDS) – short form and anxiety measured with the Hamilton anxiety rating scale (HAM-A), assessed at baseline, week 6 and end of treatment (week 12)
Notes	Specific male population residing in veterans' facilities. Randomised was matched by age and CDR score. Which of the outcomes was the primary outcome was not reported. Additional analyses compared effects stratified by musical aptitude but assessment and analyses was limited to the intervention group.
	Patient and public involvement: no information
	Some of the authors were responsible for "funds collection," but a funding source is not reported.

Liu 2021 *(Continued)*

		Conflict of interest: "none"

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"Patients with AD were randomized into an intervention or a control group with a 1: 1 ratio. The assessor, the music facilitator, and the participants were not able to know the group allocation in advance."
Allocation concealment (selection bias)	Low risk	"An external expert who had no direct contact with the patients generated the randomization list using a spreadsheet software program and kept each patient's allocation concealed."
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible to blind the convener and participants
Blinding of outcome assessment (detection bias) All outcomes	Low risk	"Only the outcome assessor was blinded to the group allocation throughout the trial. At the time of assessments, participants were reminded not to reveal their allocation to the assessor."
Incomplete outcome data (attrition bias) All outcomes	Low risk	All 50 who were randomised were in the analyses.
Selective reporting (reporting bias)	Unclear risk	No protocol and no preregistration mentioned
Other bias	Low risk	No other bias detected

Lord 1993

Study characteristics

Methods	RCT (parallel), total duration of 6 months
	No information provided about start and end dates of the study
Participants	Country: USA
	Residents in a privately funded home for older people were included. The 60 participants were "randomly selected from approximately 200 patients clinically diagnosed" as having AD.
	Of 60 residents, 42 were women and 18 were men. Age range: 72 to 103 years
Interventions	Experimental group: mixed active-receptive group intervention with music listening and playing along (30-minute sessions delivered 6 times per week for a period of 6 months)
	Control group 1: jigsaw puzzle activities (30-minute sessions 6 times per week for a period of 6 months)
	Control group 2: no special treatment, but involved in usual recreational activities of drawing, painting, and watching TV
Outcomes	Cognition, social skills (interaction) and emotional well-being as assessed with a self-made questionnaire: general impressions (assessed before and after intervention period) + participants' disposition and social coaction (assessed with a focused 30-seconds, observation on 1 participant for 3 periods

Lord 1993 *(Continued)*

during each activity session for the first 2 weeks and final 2 weeks of the study (resulting in 36 observations for each participant in the first 2 weeks and 36 observations in the last 2 weeks).

Notes	Randomisation stratified by gender.
	We included the study as it met the inclusion criteria but the data did not allow for a contribution to the meta-analysis.
	Patient and public involvement: no information
	No information reported about funding.
	Conflict of interest: no information

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	Quote: "The patients were non-systematically separated into three groups of equal size."
Allocation concealment (selection bias)	High risk	Quote: "To assure equal representation by gender, the random division was implemented first with the female and then with the male patients." No further information provided on method to conceal allocation sequence
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible to blind the convener and participants
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	No information provided on blinding of the outcome assessors
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	No information provided
Selective reporting (reporting bias)	Unclear risk	Not enough detail reported about the outcome measures. No study protocol available
Other bias	High risk	We were unable to abstract data or reproduce the results. No statistical tests were reported for the between-group comparisons, only for within-group. The article reported that the number of correct answers for each of the 3 groups was summed for baseline and post-treatment, and then a 1-way analysis of variance conducted. No information on how the data were analysed, whether the baseline was used as a covariate. Table 1 analysis of variance, although showing significant differences between the 3 therapies, did not seem valid. For example, the degrees of freedom within groups were not correct. To interpret this table, far more information is required. Even if the results in table 2 are accepted, all that can be deduced is that the treatments were different. They may be different in the level of participation in the therapies, but that does not explain whether the therapy itself brought any benefit.

Lyu 2014

Study characteristics

Methods	RCT (parallel)
	Recruitment took place between January 2012 and April 2014.
Participants	Country: China
	Participants were staying in a hospital for older adults. People with mild dementia (AD; CDR score 0.5 or 1.0) were enroled.
	A total of 93 participants were enroled. Experimental group: 32 participants; control group 1: 31 participants; control group 2: 30 participants .
	Women: experimental group: 69%; control group 1: 68%; control group 2: 70%. Mean age: experimental group: 68.8 (SD 7.0) years; control group 1: 70.4 (SD 8.4) years; control group 2: 69.9 (SD 7.84) years
Interventions	Experimental group: active music therapy group that included singing lyrics provided by a music therapist. Sessions were daily for 30 minutes for 3 months.
	Control group 1: "lyrics control group" where the same lyrics were read without music, supervised by the music therapist (daily 30 minutes for 3 months)
	Control group 2: "blank control group", which represented usual care
Outcomes	Cognition (overall cognitive functioning, verbal fluency, auditory verbal learning)
	• MMSE (primary outcome) • Verbal fluency: 1-minute animal naming test (secondary outcome) • Immediate recall and delayed recall: the World Health Organization-University of California Los Angeles Auditory Verbal Learning Test (secondary outcome)
	Overall behavioural problems
	• NPI, including the NPI Caregiver Distress Scale (secondary outcomes)
	Long-term outcomes were assessed 3 months after treatment ended.
Notes	Public and patient involvement: no information reported
	Funding: no information reported. There is no Acknowledgement section.
	Conflict of interest: this article does not report anything about conflict of interest.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Random sequence generated by random numbers table
Allocation concealment (selection bias)	Unclear risk	Not reported
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible to blind the convener and participants
Blinding of outcome assessment (detection bias)	Unclear risk	Not reported

All outcomes

Incomplete outcome data (attrition bias) All outcomes	Low risk	No missing data
Selective reporting (reporting bias)	Unclear risk	There was no protocol published in a peer-reviewed journal and it was not registered in any clinical trial registration platform.
Other bias	Low risk	No other bias detected

Narme 2012-study 1

Study characteristics

Methods	RCT (parallel) Lasted 6 weeks. Start and end dates not reported 1 article (Narme and colleagues 2012: Narme 2012-study 1 and Narme 2012-study 2) reported on 2 studies with similar designs indicated by study 1 and study 2 in the article
Participants	Country: France Enrolled 22 participants who resided on a unit for older adults, which was part of Valenciennes hospital. No diagnostic criteria for dementia were mentioned. Ten of 22 were women (experimental group: 6 of 11; control group: 4 of 11); age not described. The participants had Alzheimer's disease, MMSE 3–18
Interventions	Experimental group: mixed active-receptive group music therapy, 6 × 2-hour sessions, 2 per week (over 3 weeks) Control group: art therapy involving painting sessions with a variety of materials, 6 × 2-hour sessions, 2 per week Both interventions were delivered by 2 psychologists.
Outcomes	Outcomes were hypothesised to be more favourable for music therapy (experimental) compared with the other activity (control). • Emotional state; social behaviour from discourse content and EFEs (emotional facial expressions) as assessed from first 2 minutes of filmed interviews. • Further, emotional status was assessed as mood, with the STAI-A (timewindow not specified) For long-term outcomes, we used the assessment 4 weeks after treatment ended (there was also an assessment after 2 weeks)
Notes	The author provided data upon our request. Patient and public involvement: no information. Funding: l'Agence Nationale pour la Recherche du Ministère Français de l'Enseignement Supérieur et de la Recherche (ANR-09-BLAN-0310-02) et de l'Institut Universitaire de France à Séverine Samson. Conflict of interest: none.

Risk of bias

Narme 2012-study 1 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No explanation how random sequence was generated
Allocation concealment (selection bias)	Unclear risk	No information about allocation concealment
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible to blind the convener and participants
Blinding of outcome assessment (detection bias) All outcomes	High risk	High risk of bias because outcomes were assessed by nurses who were not blinded for the interventions.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Only a few participants were lost to follow-up.
Selective reporting (reporting bias)	Unclear risk	No study protocol available
Other bias	Low risk	No other bias detected

Narme 2012-study 2

Study characteristics	
Methods	RCT (parallel)
	Lasted 9 weeks. Start and end dates not reported
	1 article (Narme and colleagues 2012: Narme 2012-study 1 and Narme 2012-study 2 reported on 2 studies with similar designs. They were indicated by study 1 and study 2 in the article.
Participants	Country: France
	Participants resided on a unit for older adults, which was part of Valenciennes hospital.
	Participants had moderate-to-severe AD (MMSE < 12, no diagnostic criteria mentioned).
	Enrolled 14 participants, of whom 11 were included in the analyses. The participants had Alzheimer's disease. Sex/gender and age not described
Interventions	Experimental group: mixed active-receptive group music therapy, 8 × 2-hour sessions, 2 per week (over 4 weeks)
	Control group: cooking sessions, 8 × 2-hour sessions, 2 per week that included preparing a different recipe collectively, with roles distributed according to the participants' abilities. Participants were encouraged to taste ingredients, and verbalise remembrances.
	Both interventions delivered by 2 psychologists.
Outcomes	Outcomes for which stronger and more sustainable effects were hypothesised for music therapy (experimental) compared with the other activity (control) (measured 2 and 4 weeks after the last intervention).

Narme 2012-study 2 *(Continued)*

- Emotional state; social behaviour from discourse content and EFEs (emotional facial expressions) as assessed from first 2 minutes of filmed interviews
- Further, emotional status was assessed as mood, with the STAI-A (time window not specified).

Notes	The author provided data upon our request.
	Patient and public involvement: no information.
	Funding: l'Agence Nationale pour la Recherche du Ministère Français de l'Enseignement Supérieur et de la Recherche (ANR-09-BLAN-0310-02) et de l'Institut Universitaire de France à Séverine Samson
	Conflict of interest: none

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No explanation how random sequence was generated
Allocation concealment (selection bias)	Unclear risk	No information about allocation concealment
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible to blind the convener and participants
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Outcomes assessed by 5 independent and blinded observers
Incomplete outcome data (attrition bias) All outcomes	Low risk	Only a few participants were lost to follow-up.
Selective reporting (reporting bias)	Unclear risk	No study protocol available
Other bias	Low risk	No other bias detected

Narme 2014

Study characteristics

Methods	RCT (parallel)
	Lasted 10 weeks. Start and end dates not reported
Participants	Country: France
	Participants were living in a residential care home that was part of Reims University Hospital.
	Participants had AD or mixed dementia according to DSM-IV criteria. Inclusion criterion: MMSE ≤ 20. Exclusion criterion: high musical expertise as assessed by Musical Experience Questionnaire

Narme 2014 *(Continued)*

	Of 48 participants, at baseline, 37 were included in the analyses, of whom 32 were women (experimental group: 15 participants; control group: 17 participants). Mean age: experimental group: 86.7 (SD 6.4) years; control group: 87.5 (SD 6) years
	Mean MMSE: experimental group: 9.6 (SD 5.3); control group: 10.8 (SD 8.4)
Interventions	Experimental group: mixed active-receptive group music therapy, alternating listening and playing and singing along; 8 × 1-hour sessions, twice a week (during 4 weeks)
	Control group: cooking sessions as another pleasant activity in a group setting, which included preparing a different recipe during 8 sessions, twice a week, collectively, with roles distributed according to the participants' abilities
Outcomes	Main outcomes (outcomes for which improvement was hypothesised) were as follows.
	• Behaviour as assessed with the CMAI (total score up to 203; timeframe not reported but reference provided) and the NPI (total score up to 144; timeframe not reported but reference provided) • Emotional state; social behaviour from discourse content and EFEs (emotional facial expressions) as assessed from first 3 minutes of filmed interviews about current feelings and personal history. Emotional state was quantified through counting of numbers of negative and positive words, and positive and negative EFE. • Further, emotional status was assessed as mood, with the STAI-A (timewindow not reported, but reference provided).
	Another outcome (for which an effect "to a lesser extent" was hypothesised) was improved cognition measured with the SIB. Long-term outcomes were assessed 4 weeks after the last session.
Notes	The author provided data upon our request.
	Also, an effect "to a lesser extent" was hypothesised as improved professional carer's distress measured with an adapted version of the NPI, a distress scale.
	"Only native French speakers were recruited in order to ensure familiarity with the songs selected for music sessions."
	"All participants were receiving stable pharmacological treatments with constant doses during the clinical trials."
	Patient and public involvement: no information
	Funding: "Agence Nationale pour la Recherche" of the French Ministry of Research (contract number ANR-09-BLAN-0310-02).
	Conflict of interest: "nothing to disclose"

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No explanation as to how the participants were randomly assigned to groups
Allocation concealment (selection bias)	Unclear risk	No information about allocation concealment
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible to blind the convener and participants

Narme 2014 *(Continued)*

Blinding of outcome assessment (detection bias) All outcomes	Low risk	All observers were blind to the group to which the participant was allocated, although only one was blind to the pre- or post-test treatment phase. Further, only the first 3 minutes of interviews were analysed, which we feel decreased chances that raters could infer the group from the interviews. Regarding other outcomes, these were assessed by blinded carers and psychologist.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Probably about the same number was missing in each of the groups and health problems (6 participants) and death (2 participants) were unlikely related to the intervention. Refusal (3 participants) may have been more of a problem, but this was the case in only 3/48 randomised (although unknown in which group).
Selective reporting (reporting bias)	Unclear risk	No protocol available
Other bias	Low risk	No other bias detected

Prieto Alvarez 2022

Study characteristics	
Methods	RCT (cross-over)
	Study period: 12 weeks, early 2020 (before first COVID-19 lockdown)
Participants	Country: USA
	Inclusion criteria: clinical diagnosis of AD or related dementia with a stage 5 to 7 on the Global Deterioration Scale (GDS; Reisberg 1988), which indicates moderately severe to very severe cognitive decline; capacity to tolerate social interactions, musical stimuli and testing
	Exclusion criteria: very severe hearing impairment; delirium or psychosis
	29 participants recruited from an assisted living community in Massachusetts, USA (8 men, 21 women). Mean age 88.5 years (SD 6.7)
Interventions	Experimental group: applied "Neurologic Music Therapy" (NMT) techniques
	Control group 1: enrichment programme with physical and cognitive exercises
	Control group 2: tv watching, not actively facilitated
	Session frequency, timing, duration and themes addressed were the same for the three groups.
Outcomes	Overall behavioural problems (includes depression and anxiety items): NPI-NH
	Emotional well-being: self-reported DMPT (includes positive and negative emotions such as sad)
	Cogntion: MoCA
Notes	Published data not suitable for analysis: cross-over trial, we need first-period data. February 2024: asked for data. Data suitable for analysis not provided
	Patient and public involvement: no information
	Study not funded
	Conflict of interest: the author "is a neurologic music therapy fellow and therapeutic program manager". "The author declares no conflicts of interest".

Prieto Alvarez 2022 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Random permutation by independent researcher whose only role in the study was to randomise participants
Allocation concealment (selection bias)	Unclear risk	No information on this
Blinding of participants and personnel (performance bias) All outcomes	High risk	Impossible for these interventions
Blinding of outcome assessment (detection bias) All outcomes	High risk	All assessors were part of the research programme. Blinding not mentioned in paper
Incomplete outcome data (attrition bias) All outcomes	High risk	2 of 3 outcome measures available for only 21/29 (72%) and 18/29 (62%) of participants
Selective reporting (reporting bias)	Unclear risk	No trial or protocol registration referenced in the paper
Other bias	Unclear risk	Music therapist sole author of first publication on the study

---

Raglio 2010a

Study characteristics

Methods	RCT (parallel) March to November 2007 in 3 cycles of 12 sessions
Participants	Country: Italy Participants were residents from 5 nursing homes. Inclusion criteria • Diagnosis of dementia of the AD type, vascular dementia or mixed dementia (DSM-IV; MMSE (0 to 30) ≤ 18/30; CDR (1 to 5) ≥ 2/5). Mean MMSE: experimental group: 8.0 (SD 4.8); control group: 8.6 (SD 2.5). Mean CDR: experimental group: 2.8 (SD 0.4); control group: 2.9 (SD 0.6) • Presence of behavioural disturbances 60 participants (55 women, 5 men); mean age (age range): experimental group: 85.4 (74 to 99) years; control group: 84.6 (69 to 96) years
Interventions	All participants in the experimental and control groups received standard care (i.e. educational and entertainment activities such as reading a newspaper, performing physical activities, etc). Experimental group: received 3 cycles of 12 active music therapy sessions (total of 36 sessions) each, 3 times a week. Each session included a group of 3 people and lasted 30 minutes. Control group: standard care

Raglio 2010a *(Continued)*

	Each cycle of treatment was followed by 1 month of washout period (in the context of a parallel design) while the standard care activities continued over time. Total duration 6 months
Outcomes	NPI. Long-term outcomes were assessed 2 months after treatment ended (which included 1 month of washout)
Notes	The author provided data upon our request.
	Patient and public involvement: no information
	No information about funding reported; acknowledgement of contributions of persons and foundations are not specified.
	Conflict of interest: no information

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Probably yes, but no details provided
		Quote: "Sixty patients from 5 nursing homes [...] were eligible and were randomly assigned to experimental or control group."
Allocation concealment (selection bias)	Unclear risk	No details provided
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible to blind the convener and participants
Blinding of outcome assessment (detection bias) All outcomes	Low risk	The outcome assessor was blinded.
		Quote: "The assessments were made by NH [nursing home] healthcare assistants who were blinded to the aim of the study."
Incomplete outcome data (attrition bias) All outcomes	Low risk	Dropouts did not appear to cause bias.
		Quote: "During the study 7 patients dropped out, 3 in the experimental and 4 in the control group. The drops-out were due to death (n = 5), transfer to acute hospital because of hip fracture (n = 1) and transfer to another NH [nursing home] (n = 1)."
Selective reporting (reporting bias)	Unclear risk	Changes in Barthel Index scores and MMSE were not presented.
		Quote: "The patients' communicative and relational skills did not improve from baseline to the end of the treatment in the experimental group (data not shown)." No study protocol available
Other bias	Low risk	Baseline imbalances do not appear to have caused bias.

Raglio 2010b

Study characteristics

Methods	RCT (parallel)

Raglio 2010b *(Continued)*

	Study duration or start and end dates not reported
Participants	Country: Italy
	Participants resided in a nursing home.
	The participants had AD according to National Institute of Neurological and Communicative Diseases/Stroke and the Alzheimer's Disease and Related Disorders Association criteria or vascular dementia according to National Institute of Neurological Disorders and Stroke and Association criteria. "Patients with musical competence or knowledge about music therapy were excluded."
	Of 20 participants, 15 were women (experimental group: 8/10; control group: 7/10). Mean age: experimental group: 84 (SD 6) years; control group: 87 (SD 6) years
	CDR scale means: experimental group: 1.9 (SD 0.9); control group: 2.2 (SD 0.7). Mean MMSE scores at baseline: experimental group: 17 (SD 6); control group: 13 (SD 4)
Interventions	Experimental group: active, individual music therapy intervention in which free musical improvisation was used to build a relationship between participant and music therapist; 30 sessions of 30 minutes, twice a week (during 15 weeks)
	Control group: no music exposure but educational and occupational activities such as personal care, lunch, bath, cognitive stimulation, reading a newspaper, etc. Frequency or duration not reported, and these activities referred to as "standard care."
Outcomes	Main outcome (in line with study aims): behavioural and psychological symptoms of dementia measured with NPI (no timeframe reported but reference provided), including depression subscore
	Other outcomes were cognition, measured with MMSE and ADAS-cog, and depression measured with the NPI.
	Heart rate (variability) and (instrumental) ADL
Notes	The author provided data upon our request.
	Patient and public involvement: no information
	Funding source not reported
	Conflict of interest: no information

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Software mentioned: "patients were randomised to music therapy treatment or standard care by using the randomisation program QuickCalcs."
Allocation concealment (selection bias)	Low risk	Not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible to blind the convener and participants
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Not clear who assessed the outcomes

Raglio 2010b *(Continued)*

Incomplete outcome data (attrition bias) All outcomes	Low risk	No dropout
Selective reporting (reporting bias)	Unclear risk	No protocol of the pilot study available
Other bias	Low risk	No other bias detected.

Raglio 2015

Study characteristics

Methods	RCT (parallel)
	Recruitment from January 2013 to April 2014
Participants	Country: Italy
	People with moderate to severe dementia residing in 9 institutions (department for older adults, geriatric centre or nursing home)
	Inclusion criteria: aged ≥ 65 years; diagnosis of dementia according to DSM-IV Revised criteria; CDR score 17 of 1 to 4; MMSE score ≤ 18; NPI score ≤ 18; depression, anxiety, agitation or apathy NPI subitem scores > 6; residence in the nursing home > 2 months; and no significant variations in dosage of psychotropic medications during the previous month
	Exclusion criteria: severe cardiovascular, pulmonary or gastrointestinal disease; music therapy or 'listening to music' treatment in the previous year and refusal to participate
	Of a total of 120 participants: 40 each in experimental group, control group 1 and control group 2
	Women: 78.3% total: experimental group: 80%; control group 1: 72.5%; control group 2: 82.5%
	Age: experimental group: 81.7 (7.8) years; control group 1: 81.0 (7.6) years; control group 2: 82.4 (6.8) years
	There was no specification of dementia subtypes.
Interventions	Experimental group: individual active music therapy delivered by a music therapist in a separate room. Twice a week for 10 weeks, 30 minutes per session
	Control group 1: individualised listening that did "not involve any kind of direct relationship with a therapist" (30-minute sessions, twice a week for 10 weeks)
	Control group 2: usual care
Outcomes	Quality of life: CBS-QoL
	Overall behavioural problems: NPI
	Depression: CSDD
	Observed social behaviour in participants of the experimental group only
	Long-term outcomes were assessed 2 months after treatment ended.
Notes	The author provided data upon our request.
	Patient and public involvement: no information

Raglio 2015 *(Continued)*

Study not funded

Conflict of interest: "Sponsor's Role: None"

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Participants randomised to 1 of 3 treatments. Randomisation was centralised, and each participant was blindly associated to a sequential number.
Allocation concealment (selection bias)	Low risk	No potential problems detected
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible to blind the convener and participants.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Because participants were in the moderate to severe stages of dementia and were not able to provide adequate answers, the evaluators interviewed the formal carers on the participant's condition the previous week. All evaluators were blind to the type of treatment the participant was receiving.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Total loss to follow-up < 20%. 0/40 refused treatment in experimental group and 5/40 refused treatment in control group 1, which might be due to refusing to wear the headphone.
Selective reporting (reporting bias)	Unclear risk	No protocol available
Other bias	Low risk	No other bias detected

Reschke-Hernández 2019

Study characteristics

Methods	Cluster-RCT cross-over trial with a 2-week washout period between the conditions
	The study was conducted from June to September 2018. Checklists created by the author were used to prepare and evaluate sessions to ensure the treatment protocol was followed.
Participants	Country: USA
	Eligible residents had lived in the facility for at least 3 months, were 65 years or older and had English as their primary language. Residents already receiving music therapy, with severe sensory impairment, co-occuring Parkinson's or Huntington disease, Down syndrome or severe mental illness, and those bedridden or receiving hospice services were excluded from participation.
	32 nursing home residents (26 women) with mean age 82.13 years (SD 8.44), of whom 24 could self-report outcomes. More than half (56%; 18 participants) had an unspecified type of dementia.
Interventions	Experimental group: 25-minute active group singing sessions that included 15% conversations delivered by a music therapist who delivered the intervention according to a clinical practice model developed by the author based on literature, models such as person-centred care, and expert input. The sessions were delivered 3 times a week for 2 weeks.

Reschke-Hernández 2019 *(Continued)*

	Control group: 25-minute verbal discussion sessions without music were delivered by a music thera-pist. The sessions were delivered 3 times a week for 2 weeks.
Outcomes	Outcomes (none identified as primary or secondary) included quality of life measured with QUALIDEM and a social (non)engagement item of the OERS measured with the MPES scale of observed behaviour for which the author provided the exact data.
Notes	Other than social engagement, we did not use in the analyses any directly observed behaviour during the sessions. Cognition was measured with the CDR and the Montreal Cognitive Assessment (MoCA) to examine interaction with dementia severity and to evaluate change over time, but was not reported as an outcome measure. The Alzheimer's Disease and Related Dementias Mood Scale (AD-RD Mood Scale) was also an outcome measure, and it included an item on depression but the item-level data were not available.
	Sample size adapted to adjust for clustering. ICC applied from Baker 2022.
	Patient and public involvement: no information
	Funding: all through the University of Iowa; (1) departmental funds, (2) Caplan funds, (3) Graduate and Professional Student Government Research Grant, (4) Office of Outreach and Engagement Community Impact Grant, and (5) GoldRUSH crowdfunding campaign
	Conflict of interest: "None declared". (2023) "Alaine E. Reschke-Hernández and Daniel Tranel declare that they have no conflicts of interest". (2024)

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence genera-tion (selection bias)	Unclear risk	Cluster cross-over RCT (main study). For the main study, the facilities were sorted "into two groups of comparable participant enrollment size, and then a person not vested in study outcomes flipped a coin to determine order." Fur-ther, there were more groups per facility, and participants were assigned to groups through "stratified random assignment" to "avoid having all men in a single group" "(i.e., I randomly assigned men to groups first)." For this, a "ran-dom number generator (Random.org, 1998-2019)" was used.
Allocation concealment (selection bias)	Unclear risk	" a person not vested in study outcomes flipped a coin"
Blinding of participants and personnel (perfor-mance bias) All outcomes	High risk	Not possible to blind the convener and participants, which was acknowledged ("could not blind interventionists to the study condition they led")
Blinding of outcome as-sessment (detection bias) All outcomes	High risk	"Primary limitations to this research are: (1) some of the activities assistants who facilitated the comparison conditions were also non-blinded data takers"
		"Although I attempted to limit such bias through rater blinding, training, quali-ty assurance checks, and regular interrater reliability checks on observational data, bias was nonetheless present to some extent"
		"Note that aims and hypotheses were not revealed to data takers until study completion, and the importance of blinding was emphasized within this mod-ule. However, like interventionists, it is possible that data takers may have in-ferred the aims, or unintentionally held biases regarding the efficacy of music therapy as an intervention. It is possible that they unwittingly honed their ob-servations in on behaviors they wanted to see (i.e., self-fulfilling prophesy)"
Incomplete outcome data (attrition bias)	Low risk	No missing data for assessments other than self-report of other outcomes

Reschke-Hernández 2019 *(Continued)*

All outcomes

Selective reporting (reporting bias)	Unclear risk	Trial registration after enrolment started
Other bias	Unclear risk	Music therapist sole author of first publication on the study. Credentialed music therapist trained to administer both the intervention and the control condition

---

Ridder 2013

Study characteristics

Methods	RCT, cross-over with 2 periods of 6 weeks for the different conditions
	Quote: "Data were collected in three 15-week periods during fall 2010, spring 2011 and fall [autumn] 2011."
Participants	Countries: Denmark and Norway
	Participants were from 14 nursing homes (4 in Denmark and 10 in Norway); most were from Norway (76% of participants).
	The participants had a diagnosis of dementia ("stated in medical journal," no criteria mentioned). Eligible people had moderate-to-severe dementia. Included participants had symptoms of agitation.
	Of 42 participants, 69% were women. Mean age was 81 years (range 66 to 96 years) for the 26% of participants for whom this information was available.
	40% had AD; for 38%, the type was not specified; 22% had other types of dementia such as vascular, Lewy body, frontotemporal or mixed dementia. Mean baseline MMSE score: experimental group: 9.84 (SD 5.97); control group: 5.25 (SD 4.83). Global Deterioration Scale means: experimental group: 5.54 (SD 0.69); control group: 5.80 (SD 0.62)
Interventions	Experimental group: individual mixed active-receptive music therapy, a minimum of 12 sessions were offered, but the participants received a mean of 10 sessions (SD 2.82, range 0 to 13). Frequency: twice a week (over 6 weeks). Mean duration: 33.80 (SD 9.91) minutes
	Control group: received usual care, which for some participants meant participating in group sing-along sessions
Outcomes	Primary outcome: agitation measured with the CMAI. Time frame adapted from 2 to 1 week (previous week).
	In addition to the 7-point frequency scale, a later version of CMAI was used with a 5-point disruptiveness scale. The frequency scale, CMAI-fr, ranged from 1 (never) to 7 (several times per hour), and the disruptiveness scale, CMAI-di, from 1 (not at all) to 5 (extremely). The CMAI-fr 1- to 7-point scale was transformed to scores 0 to 6, leading to a maximum total score of 66, and the 1- to 5-point CMAI-di scale was transformed to scores 0 to 4, leading to a maximum total score of 44.
	Secondary outcome: quality of life measured with the ADRQL. Time frame adapted from 2 to 1 week (previous week).
Notes	Psychotropic medication use was measured and considered as an outcome.
	Patient and public involvement: no information
	Funding: GC Rieber Foundation in Bergen and Aalborg University

Ridder 2013 *(Continued)*

		Conflict of interest: no information

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Participants were randomly allocated to 1 of 2 groups (experimental or control first), but how this was done was not described.
Allocation concealment (selection bias)	Low risk	Quote: "[A] concealed sequence procedure" was used, witnessed and signed by someone who was not involved in the study.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible to blind the convener and participants
Blinding of outcome assessment (detection bias) All outcomes	High risk	Interviewers and proxy respondents were not blinded to the treatment allocation.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Only a few values were missing; and sensitivity analyses were performed with last observation carried forward.
Selective reporting (reporting bias)	Unclear risk	Quote: "The researchers designed the study protocol in collaboration with a group of clinicians from Denmark and Norway," but there is no reference to compare with.
Other bias	Low risk	No other bias detected

Sakamoto 2013

Study characteristics

Methods	RCT (parallel)
	Study duration, start and end dates not reported
Participants	Country: Japan
	Participants resided in 4 group homes or a special dementia hospital.
	Participants had AD according to DSM-IV criteria. Inclusion criterion: CDR scale 3 (severe dementia). Participants had no relevant hearing disorders and no experience of playing musical instruments.
	Of 39 participants, 32 were women.
	Experimental group: 11 women (mean age 81.2 years, SD 7.5) and 2 men (mean age 76 years, SD 7.1) Control group 1: 10 women (mean age 81.1 years, SD 11.0) and 3 men (78.7 years, SD 12.1) Control group 2: 11 women (mean age 81 years, SD 8.3) and 2 men (mean age 84.5 years, SD 4.95)
	Mean MMSE score at baseline: experimental group: 4.6 (SD 3.5); control group 1: 4.7 (SD 4.8); control group 2: 4.7 (SD 3.9)
Interventions	Experimental group: interactive mixed active-receptive music therapy intervention with 10 × 30-minute sessions once a week (over 10 weeks)

Cochrane
Library

Trusted evidence.
Informed decisions.
Better health.

Cochrane Database of Systematic Reviews

Sakamoto 2013 *(Continued)*

	Control group 1: passive individual music intervention (not therapy) with 10 × 30-minute sessions once a week
	Control group 2: "Each control group participant spent time with one caregiver in their own room as usual, without any music intervention (silent environment)."
Outcomes	Behavioural and psychological symptoms of dementia measured with the BEHAVE-AD rating scale
	Time frame: last 2 weeks, but any changes were by direct observation
	Another outcome was stress levels, which were also measured with the Faces Scale but only in the short term.
Notes	Patient and public involvement: no information
	Funding: MEXT KAKENHI grant numbers 19592567, 22592586 (2007–2009, 2010–2013)
	Conflict of interest: "None"

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence genera-tion (selection bias)	Unclear risk	Quote: "stratified randomization (gender and MMSE level)", but how exactly this was performed was not described.
Allocation concealment (selection bias)	Unclear risk	Quote: "Participants were randomly and blindly assigned to either control, passive, or interactive group," but there is no description of the blinding process.
Blinding of participants and personnel (perfor-mance bias) All outcomes	High risk	Not possible to blind the convener and participants
Blinding of outcome as-sessment (detection bias) All outcomes	Low risk	Quote: "The primary experimenters were not involved in the intervention or evaluation, and the evaluators did not act as music facilitators." Further, occu-pational therapists and nurses who did not work in the study institution com-pleted the BEHAVE-AD
		Quote: "The short- and long-term effects of intervention were evaluated by two trained occupational therapists and four trained nurses in a blinded fash-ion."
Incomplete outcome data (attrition bias) All outcomes	Low risk	No dropout
Selective reporting (re-porting bias)	Low risk	Study protocol available and all prespecified outcomes were reported in the article.
Other bias	Low risk	No other bias detected

Sung 2012

Study characteristics

Methods	RCT (parallel)

Sung 2012 *(Continued)*

	Total study duration or beginning and end dates are not reported.
Participants	Country: Taiwan
	Participants resided in a residential care facility.
	Diagnosis of dementia was not described. Inclusion criterion: "ability to engage in a simple activity and follow simple directions." The participants had the "ability to engage in a simple activity and follow simple directions, ability to understand Taiwanese or Chinese, no severe hearing impairment, presence of behavioural and psychological symptoms reported by nursing staff and no obvious symptoms of acute pain or infection."
	Of 60 recruited residents, 55 participated. 65.8% were women.
	Mean age: experimental group: 81.37 (SD 9.14) years; control group: 79.5 (SD 8.76) years
	The participants had mild-to-moderate cognitive impairment according to the Short Portable Mental Status Questionnaire (mean: experimental group: 6.56, SD 2.86; control group: 4.43, SD 3.17). Type of dementia was not described.
Interventions	Experimental group: active music intervention using percussion instruments, familiar music and movement. A nursing researcher and 2 trained research assistants delivered 12 sessions of 30 minutes, twice a week (over 6 weeks).
	Control group: usual care
Outcomes	Agitation assessed with a modified CMAI. Timeframe unclear with observations during music therapy session ("The behaviours of the participants during each music session were assessed by the observer assistants using modified CMAI"), and also "frequency of occurrence over 2 weeks." Unclear how the CMAI was modified.
	Anxiety assessed with RAID over previous 2 weeks
Notes	76.2% had not received any formal education.
	Included residents had behavioural and psychological symptoms as reported by nursing staff.
	Unclear if agitation effects included an immediate effect through observations during the music therapy sessions.
	Patient and public involvement: no information
	Funding: Taiwan National Science Council (NSC 96-2314-B-277-003-MY2)
	Conflict of interest: "None declared"

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "Participants were randomly assigned to either the experimental or the control group using simple random sampling method with a computer-generated list."
Allocation concealment (selection bias)	Unclear risk	Unclear who handled the allocation schedule
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible to blind the convener and participants

Cochrane
Library

Trusted evidence.
Informed decisions.
Better health.

Cochrane Database of Systematic Reviews

Sung 2012 *(Continued)*

Blinding of outcome assessment (detection bias) All outcomes	High risk	Detection bias (blinding of outcome assessment): observer assistants completed the CMAI and RAID over the last 2 weeks. Unclear if these were other people than the trained research assistants who gave the music therapy (probably, these were people who knew the person but they were also aware of the intervention because the assessment was during the intervention).
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Handling of missing data not reported; 60 were randomised and 55 were analysed.
Selective reporting (reporting bias)	Unclear risk	No published study protocol available
Other bias	Low risk	No other bias detected

Svansdottir 2006

Study characteristics		
Methods	RCT (parallel) 6-week intervention and 4-week follow-up No information reported about start and end dates of data collection	
Participants	Country: Iceland Participants resided in 2 nursing homes and 2 psychogeriatric wards. Diagnosis of dementia: all diagnosed with AD (ICD-10); Global Deterioration Scale score of 5 to 7 (moderate-to-severe dementia). Of 38 participants, sex/gender was not reported; age range: 71 to 87 years	
Interventions	Experimental group: group music therapy (3 or 4 participants per session), mixed active (playing instruments) and receptive (listening), 3 times a week for 6 weeks (total of 18 sessions), 30 minutes per session Control group: standard care as usual	
Outcomes	Behavioural and psychological symptoms of dementia assessed with the BEHAVE-AD scale. Long-term outcomes were assessed 4 weeks after the treatment ended.	
Notes	The author provided data upon our request. No clear baseline characteristics presented Patient and public involvement: no information Funded by the Research Fund for Alzheimer's Disease and Related Disorders, Landspitali University Hospital Conflict of interest: "None"	
Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**

Svansdottir 2006 *(Continued)*

Random sequence generation (selection bias)	Unclear risk	No information provided
Allocation concealment (selection bias)	Unclear risk	Quote: "...The 46 remaining patients were then randomised to a music therapy group or a control group, with 23 individuals in each group."
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible to blind the convener and participants
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Outcome assessors blinded Quote: "Two nurses were trained in using the BEHAVE-AD scale and they were blinded to the therapy used. The nurses were not part of the staff of the wards."
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	No information provided
Selective reporting (reporting bias)	Unclear risk	No data
Other bias	Unclear risk	No clear baseline characteristics presented. Also, the first author (HBS) provided the music therapy. ("Throughout the study the same qualified music therapist (H.B.S.) conducted the music therapy.")

Tang 2018

Study characteristics

Methods	RCT (parallel) "The baseline demographic data were collected during January 2016."
Participants	Country: China, Guangzhou "The participants in this study were recruited from a 1000-bed residential nursing facility, which specializes in the management of dementia-related symptoms". "The inclusion criteria for entry into this study was aged 60 years or older, mild to moderate dementia (defined by a MMSE score ranging from 10 to 27)." The participants also met the AES-C diagnostic criteria for apathy (the primary outcome). They had "no other serious co-morbidities and voluntary participation and the ability to communicate and cooperate with the research assistant to complete the questionnaires" 77 participants with dementia were randomised; 39 to the control group and 38 to the intervention group, one of whom was admitted to the hospital, resulting in 37 analysed in the intervention condition. Mean age of the 77 particpants at baseline was 75.88 years (SD 5.09; range: 65 to 90 years; 38 female); intervention group, 76.39 years SD 4.86, 17 female of 38 participants; control group 75.38 years, SD 4.94, 21 female of 39 participants. The overall mean MMSE score at baseline was 16.09 (SD 4.74; range 10 to 26). Type of dementia was not described.

Tang 2018 *(Continued)*

Interventions	Experimental group: 3 active and receptive music interventions per 50-minute session were offered in a group setting by a trained therapist assisted by a research assistant (different from the research assistant who conducted the measurements). The sessions were delivered 3 times a week for 12 weeks, totalling 36 sessions. Control group: usual care
Outcomes	Primary: apathy Secondary: communication measured with the Holden communication scale (social behaviour) and cognition measured with the MMSE
Notes	Patient and public involvement: no information No funding source was reported in the article, and the authors communicated there was no external funding. Conflict of interest: no information

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	A "computer-based random number allocation method" was used.
Allocation concealment (selection bias)	Unclear risk	Unclear when allocation was disclosed: "The participants in the intervention group were randomly divided into four independent subgroups."
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible to blind the convener and participants
Blinding of outcome assessment (detection bias) All outcomes	Low risk	"Raters in the study who were blinded to the delivery of the music intervention performed scoring"
Incomplete outcome data (attrition bias) All outcomes	Low risk	The methods section states that participants absent from the intervention 5 times or more would be removed from the analyses. However, of the 37 in the intervention group, "35 had full attendance at all the music intervention sessions, but two participants were absent for one time because of physical illness." Only one person lost to follow-up, due to hospitalisation.
Selective reporting (reporting bias)	Unclear risk	Advance study protocol or study registration not available
Other bias	Low risk	No other bias detected

Thornley 2016

Study characteristics

Methods	Pilot RCT (parallel) Data collection started September 2012 and ended September 2014.

Thornley 2016 *(Continued)*

Participants	Country: Canada
	People with dementia and moderate-to-severe cognitive impairment admitted to an inpatient psychiatric unit within a large academic hospital
	16 people (8 women and 8 men) randomised. Age: experimental group: 83.5 (SD 7.7) years; control group: 68.4 (SD 5.2) years (large difference; randomised before screening for eligibility may have caused imbalance)
	From the total sample, 11 (69%) had AD, 3 (19%) had vascular dementia and 2 (13%) had Lewy body dementia.
Interventions	Experimental group: individual, active music therapy provided by an accredited music therapist
	Control group: active engagement and attention intervention provided by a social worker
	Both groups had 60-minute sessions twice a week for 4 weeks with a maximum of 8 sessions.
Outcomes	Overall behavioural problems, and some individual item scores were reported as well from the NPI-Clinician version: frequency × severity and distress
	Agitation: CMAI
Notes	The author provided data upon our request.
	A number of the participants enroled in this study were hospitalised for 2 to 3 weeks, which limited the amount of data that could be collected. Moreover, end-of-treatment scores were reported for only some of the outcomes. Using data provided by the authors, and last observation after 5 sessions carried forward in case of missing assessments, we included 7 participants in the analyses of CMAI and NPI, and 8 participants for NPI depression and anxiety items.
	Patient and public involvement: no information
	Funding: Behavioral Supports Ontario program
	Conflict of interest: "none"

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Participants were randomised using an online randomisation programme.
Allocation concealment (selection bias)	Unclear risk	The sequence of allocation was concealed from the inpatient staff and clinical raters, but not reported for the therapists and the researchers.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible to blind the convener and participants
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Raters came from a pool of trained outpatient psychiatric nurses and social workers masked to treatment allocation.
Incomplete outcome data (attrition bias) All outcomes	High risk	Participants often did not stay long enough to attend sessions for more weeks (e.g. many did not have at least 5).
		7 participants (3 in experimental group, 4 in control group) received at least 5 therapy sessions (completed 3 weeks).

Thornley 2016 *(Continued)*

Selective reporting (reporting bias)	Unclear risk	No registration and no reference to a protocol
Other bias	Low risk	No other bias detected

Vink 2013

Study characteristics

Methods	RCT (parallel)
	Exact duration of total study or start and end dates were not reported, but therapy was provided over a period of 4 months.
Participants	Country: the Netherlands
	Residents of 6 nursing homes were included.
	Participants had any type of dementia according to DSM-IV criteria, CMAI score > 44.
	Of 77 participants, 10 did not specify type of dementia, 34 had AD, 16 had vascular dementia and 17 had another type of dementia.
	Of 94 included residents, 77 were included in the reported analyses. 54 (70%) were women.
	The mean age of all residents was 82.16 (SD 6.87) years.
Interventions	Experimental group: mixed active-receptive group music therapy, which involved listening to live music, interacting with the therapist and playing simple instruments. A maximum of 34 sessions of 40 minutes each were held, twice weekly, over 4 months.
	Control group: general recreational activities such as handwork, playing shuffleboard, cooking and puzzle games. Sessions lasted 40 minutes, twice weekly over 4 months.
Outcomes	Agitation assessed with the CMAI modified through dichotomising of items, resulting in a total score range of 0 to 29. Presence and absence of behaviour was presumably measured by direct observation or with very short time frames (because it was assessed 1 hour before the session, 1 hour after the session, 2 hours after the session and 4 hours after the session).
	Neuropsychiatric symptoms (behaviour overall, NPI)
Notes	The author provided data upon our request.
	Patient and public involvement: no information
	Funding: ZonMw (the Netherlands Organisation for Health Research and Development), the Dutch Alzheimer Foundation (Alzheimer Nederland) and the Triodos Foundation
	Conflict of interest: "None declared" (2013). "The authors declare that they have no conflicts of interest". "Sponsor's role: The funding sources are all nonprofit associations financed by public and private sectors and did not influence the study content in any way" (2014)

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "To ensure randomised allocation, sealed envelopes were used, with at least two persons present to ensure appropriate randomisation."

Vink 2013 *(Continued)*

Allocation concealment (selection bias)	Unclear risk	Only sealing was described; it remains unclear whether envelopes were sequentially numbered and opaque.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible to blind the convener and participants
Blinding of outcome assessment (detection bias) All outcomes	High risk	Quote: "Some of the nurse caregivers who rated the modified CMAI scores were at occasion responsible for taking the residents to either the activity or music therapy room. Complete blinding for some of the nurse caregivers could therefore not be guaranteed."
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	The explanation of missing data was unclear. There were 7 missing cases in the baseline data in the control group, and 4 of the 47 allocated participants died. It was unclear if baseline data were missing because participants died before the baseline assessment.
Selective reporting (reporting bias)	Unclear risk	Study protocol not available
Other bias	Low risk	No other bias detected

Werner 2017

Study characteristics	
Methods	Cluster-RCT (parallel) in 2 nursing homes
	The data collection period is not reported.
Participants	Country: Germany
	We obtained outcome data for a subset of participants (out of 117 randomised, 113 analysed), i.e. 67 residents who were diagnosed with dementia (41 in the experimental group and 26 in the control group). The residents selected for analysis either had a physician's diagnosis of dementia in their medical chart or received dementia medication. A gerontopsychiatric specialist made the diagnosis according to ICD-10 criteria.
	Broad inclusion criteria were employed; only residents who were bedridden or in short-term care (less than 4 weeks) were not eligible.
	Dementia type was not reported.
Interventions	Four clusters based on 14 nursing home wards were randomised.
	Experimental group: a music therapist with limited experience provided interactive, flexible, person-centred active and receptive music therapy sessions considering biographies and individual and group emotional states. A total of 20 sessions of 40 minutes were offered biweekly (quote: "the complete study period with 10 weeks of intervention.")
	Control group: 90-minute recreational singing sessions were delivered according to the same schedule as the music therapy but once a week during 10 weeks (weeks 1 to 11, therefore probably 10 in total).
Outcomes	Depression was the primary outcome, with depressive symptoms assessed with the MADRS at baseline and in week 6 and 12 (following the end of treatment in "week 11" or "12th week"). Change in individual MADRS scores was regarded as a secondary outcome.

Werner 2017 *(Continued)*

Notes	Upon our request, the authors provided data about the selection of residents diagnosed with dementia. Reported post hoc interaction analyses suggested greater effects of music therapy in those with dementia compared with those without dementia.
	Sample size adapted to adjust for clustering. ICC applied from Baker 2022.
	Patient and public involvement: no information
	Funding: no particular sources reported, but "The study was conducted as part of the master's degree programme at the University of Applied Sciences Würzburg Schweinfurt, Germany"
	Conflict of interest: "No potential conflict of interest was reported by the authors."

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Matching procedure unclear: "At first a geropsychiatric specialist excluded non-eligible participants by aforementioned criteria and prepared a list with the number of eligible participants and information about local conditions. After that we used this obtained list to combine several wards to clusters (two to six per cluster, depending on the number of participants and local similarities) and achieve clusters of similar size (20-42 participants per cluster) and clinical characteristics. These four clusters were then allocated randomly to music therapy or recreational singing (use of music in a non-therapeutic context)." Random sequence generation not reported.
Allocation concealment (selection bias)	Low risk	Allocation was not known to participants before decision to participate: "The participants were informed that the study aimed to examine the effect of music therapy and that music therapy was going to be offered in some of the wards. They were informed that their participation was voluntary and that confidentiality was ensured."
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible to blind the convener and participants
Blinding of outcome assessment (detection bias) All outcomes	High risk	"Previously completed forms were available for geropsychiatric specialists at the follow-up evaluations for comparison"
Incomplete outcome data (attrition bias) All outcomes	Low risk	Only a few participants from the total sample (with and without dementia) were lost to follow-up.
Selective reporting (reporting bias)	Unclear risk	No publised protocol or trial registration available
Other bias	Low risk	No other bias detected

AD: Alzheimer's disease; ADAS-cog: Alzheimer's Disease Assessment Scale Cognitive subscale; ADL: activities of daily living; ADRQL: Alzheimer's Disease-Related Quality of Life; AES-C: Apathy Evaluation Scale-Clinician; AR-RD: Alzheimer's Disease and Related Dementias Mood Scale; BEHAVE-AD: Behavioural Pathology in Alzheimer's Disease; BIMS: Brief Interview for Mental Status; BTA: Brief Test of Attention; C-CMAI: Chinese Version of the Cohen-Mansfield Agitation Inventory; C-CSDD: Chinese Version of the Cornell Scale for Depression in Dementia; C-MMSE: Chinese Version of the Mini-Mental State Examination; CBS-QoL: Cornell-Brown Scale for Quality of Life in Dementia; CDR: Clinical Dementia Rating; CMAI: Cohen-Mansfield Agitation Inventory; CMAI-SF: Cohen-Mansfield Agitation Inventory – Short Form; CSDD: Cornell Scale for Depression in Dementia; DCM: Dementia Care Mapping; DemTect: Demenz-Detektion; DMPT: Dementia Mood Picture Test; DQOL: Dementia Quality of Life; DSM-IV: Diagnostic and Statistical Manual of Mental Disorders, 4th edition;

DSM-IV-TR: Diagnostic and Statistical Manual of Mental Disorders, 4th edition Text Revision; DSM-5: Diagnostic and Statistical Manual of Mental Disorders, 5th edition; DVD: digital video disc; EFE: emotional facial expression; EQ-5D-5L: EuroQol-5 Dimensions 5 Levels questionnaire; FAS: Controlled-Oral-Word-Association-Test; GDS: Geriatric Depression Scale (note that Global Deterioration Scale is written in full only); HAM-A: Hamilton Anxiety (Rating) Scale; IADL: instrumental activities of daily living; ICC: intraclass correlation coefficient; ICD-10: International Classification of Diseases-10; ITT: intention to treat; MAC-Q: Subjective Memory Complaints Questionnaire; MADRS: Montgomery- Asberg Depression Rating Scale; MMSE: Mini-Mental State Examination; MMST: Mini Mental Status Test; MoCA: Montreal Cognitive Assessment; MPD: Deferred Prose Memory; MPES: Menorah Park Engagement Scale; MPI: Immediate Prose Memory; MTF/ROF: Modified Taylor Figure/Rey-Osseterrieth Figure; NINCDS-ADRDA: National Institute of Neurological and Communicative Diseases and Stroke/Alzheimer's Disease and Related Disorders Association; NOSGER: Nurses' Observation Scale for GERiatric patients; NPI: Neuropsychiatric Inventory; NPI-NH: Neuropsychiatric Inventory Nursing Home version; OERS: Observed Emotion Rating Scale; PANAS: Positive and Negative Affect Schedule; QUALIDEM: Quality of Life in Late-stage Dementia Scale; QoL-AD: Quality of Life-Alzheimer's Disease; RAID: Rating Anxiety in Dementia Scale; RCT: randomised controlled trial; SD: standard deviation; SIB: Severe Impairment Battery; STAI-A: State-Trait Anxiety Inventory for Adults; TV: television

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Arroyo-Anlló 2013	Not clear whether it was randomised; the outcome was self-consciousness.
Baker 2023	Family caregiver-delivered intervention
Ballard 2009	RCT, not a music-based therapeutic intervention. A small proportion of the study sample (35) followed individualised music as an intervention. There was a non-significant improvement on the total CMAI score.
Brotons 2000	Only 4 therapy sessions
Bruer 2007	Participants were involved in fewer than 5 sessions. Cross-over RCT of 8 weeks comparing group music therapy to video presentation on cognition (MMSE score)
Bugos 2005	RCT, people with dementia were excluded.
Chae 2015	Not randomised
Clair 1996	Not clear if participants were randomised, and they participated in fewer than 5 sessions.
Cohen-Mansfield 2010	No control group included
Davidson 2011	No control group included
Dowson 2021	Included dyadic therapy, not only for participants but also for family
Eun-Young 2020	Not randomised
Garland 2007	Fewer than 5 sessions, in which participants listened to preferred music
Gerdner 2000	Use of music in control group; tested individualised music compared to classical "relaxation" music
Groene 1993	Control group also received music therapy
Hanser 1994	RCT, participants had depression not dementia
Hicks-Moore 2008	RCT, comparison of favourite music and hand massage, fewer than 5 sessions
Hokkanen 2008	RCT, no music therapy, study involved dance and movement therapeutic methods

Study	Reason for exclusion
Holmes 2006	RCT, comparison of live interactive music, passive prerecorded music or silence for 30 minutes in a single session, fewer than 5 sessions
Janata 2012	The intervention did not meet our criteria for a therapeutic-based intervention in which contact with a therapist or facilitator is essential.
Kwak 2016	RCT, only music listening, no music therapist or interaction
Low 2016	Dance intervention; control group involved "music appreciation and socialisation" provided by the dance teacher.
Lyu 2018	Two intervention groups: a singing group and a lyric reading group
Matziorinis 2023	Unclear how many participants received at least 5 sessions. Further, the main part of the intervention was offering audio-recorded music and the family could provide the intervention at home.
McCreedy 2022	Music listening intervention
Noice 2009	RCT, no music therapy but "a theatrically based intervention"
Otto 1999	RCT, participants did not have dementia.
Pomeroy 1993	RCT, music was part of physiotherapy.
Raglio 2008	Not randomised, quasi-randomised study
Riegler 1980	RCT, not clear whether participants were diagnosed with dementia
Sánchez 2016	RCT, only music listening, no music therapist or interaction
Särkämö 2014	No music-based therapeutic intervention, but singing coaching for family carers and nurses, and listening to music
Satoh 2014	No music-based therapeutic intervention, but physical exercise combined with music
Sung 2006	No music-based therapeutic intervention, but music with movement intervention
Tamplin 2018	Not randomised, pre-post test study; the intervention included dyadic therapy for the family as well.
Thompson 2005	RCT, single test moment, music as cue to facilitate performance on a category fluency task. No therapeutic intervention
Tomaino 2011	Not randomised
Tz-Han 2023	More than one third of the 60-minute music therapy sessions was spent in reminiscence therapy involving virtual reality gaming and discussions; therefore, it did not meet our criteria for a music-based therapeutic intervention.
Vanderark 1983	RCT, not clear whether participants were diagnosed with dementia
Van de Winckel 2004	RCT, music-based exercises not music-based therapeutic intervention
채경숙 2015	Not randomised

CMAI: Cohen-Mansfield Agitation Inventory; MMSE: Mini-Mental State Examination; RCT: randomised controlled trial

Characteristics of studies awaiting classification *[ordered by study ID]*

Asmussen 1999

Methods	RCT
Participants	People with dementia (n = 121) living in two long-term care facilities in New York, USA
Interventions	Music therapy group sessions 3 times a week for 16 weeks
Outcomes	Depression, anxiety and agitation, quality of life, cognition, ADL, medication, memory, behaviour, expression, activities
Notes	In 2024, Dr Tomaino sent us two unpublished reports related to this study. We will decide on inclusion in a future update.

Campbell 2022

Methods	3-arm pragmatic RCT
Participants	Nursing home residents in Germany
Interventions	Active music therapy versus a control group receiving vibroacoustic therapy as a receptive approach versus a control group receiving usual care
Outcomes	Primary - behavioural problems overall Secondary - depression and quality of life
Notes	We have been unable to identify a full report on this study.

Curto Prieto 2015

Methods	Either RCT or quasi-experimental design
Participants	"Institutionalized" people with dementia (n = 24), "in phases 5 and 6" (moderate-to-advanced dementia)
Interventions	Experimental group: group music therapy Control group: reminiscence-recreation group
Outcomes	Mood and cognition, perhaps also (social) behaviour
Notes	Conference abstract. When a full report becomes available, the design needs careful evaluation (a "quasi-experimental study" with a "pre-post test design with a control group" wherein groups were "randomly assigned to a music therapy group or a reminiscence group").

Hong 2011

Methods	Parallel-group RCT

Hong 2011 *(Continued)*

Participants	30 nursing home residents in the Republic of Korea
Interventions	Experimental group: songwriting; music therapy programme employing songwriting activities. 3 stages: preparing songwriting, songwriting and reinforcing songwriting. A therapist administered the active individual intervention. Sessions of 60 minutes were given for 16 weeks (once per week).
	Control group: free time given
Outcomes	Cognition assessed with the MMSE-K
Notes	Presentation of results in the paper (Figures 2a and b) was incorrect. The intervention and control groups were reversed. There was little variability in MMSE-K scores with either no change or change in one direction only. The authors have not responded to remaining questions about whether outcome assessment was blinded, any review or approval of the protocol and the time between the repeated cognition tests for which only mean scores are presented.

Hsiung 2013

Methods	Pilot cross-over RCT
Participants	10 people with Alzheimer's disease, MMSE score range 6 to 28
Interventions	Experimental group: music therapy by a trained music therapist; no detail on type of intervention reported
	Control group: not reported
Outcomes	Overall behavioural problems was a primary outcome; secondary outcomes included quality of life, depression and cognition
	In addition, there were outcomes other than the 7 outcomes of interest for our Cochrane review.
Notes	Conference abstract. If a full report becomes available, the type of intervention will be reviewed against our criteria for music-based therapeutic interventions.

Hsiung 2015

Methods	Cross-over RCT
Participants	27 people with moderate Alzheimer's disease
Interventions	Experimental group: "music therapy by an accredited music therapist following a standardized structured protocol (Clair 1990)."
	Control group: "waiting" (probably usual care)
Outcomes	Overall behavioural problems was a primary outcome; secondary outcomes included quality of life, depression, agitation and cognition
	In addition, there were outcomes other than the 7 outcomes of interest for our Cochrane review.
Notes	Conference abstract. If a full report becomes available, the exact type of intervention should be reviewed against our criteria for music-based therapeutic interventions.

ISRCTN11001662

Methods	Feasibility RCT
Participants	Residents of a hospital ward with moderate to severe dementia
Interventions	Interactive group music workshops versus music listening workshop as the control condition
Outcomes	Primary outcomes are eligibility, recruitment and retention. Secondary outcome: agitation, observations and experiences
Notes	We were unable to identify a full report on this study.

Kwak 2013

Methods	"Case control study" but "...participants…were assigned randomly to a music therapy group and a control group."
Participants	People with moderate Alzheimer's disease residing in 1 of 4 participating long-term care centres (probably 120 were randomised and 82 participated)
Interventions	Experimental group: music therapy with active elements provided by music therapists Control group: "standard care"
Outcomes	Behavioural problems overall measured with the BEHAVE-AD; however, the study aims and results are about agitation disruptiveness. In addition, there were outcomes other than the 7 outcomes of interest for our Cochrane review.
Notes	Conference abstract. If a full report becomes available, we will carefully consider the study design to see if it qualifies as an RCT.

Moreira 2023

Methods	Parallel-group RCT
Participants	People with dementia, people with mild cognitive impairment, healthy older adults
Interventions	Music therapy with family education to communicate with the person with dementia, and to provide music interventions at home. The control group did not receive family education.
Outcomes	Depression, cognition
Notes	Unsure whether it fulfils inclusion criteria. If so, authors can provide data for the subgroup of participants with a clinical diagnosis of dementia.

NCT00448318

Methods	Single-blind parallel-group RCT

NCT00448318 *(Continued)*

Participants	"People with a clinical diagnosis of dementia who are currently in sub-acute aged care hospital wards"
Interventions	Music therapy
Outcomes	Primary outcomes: depression and cognition
Notes	We were unable to identify a full report on this study. The study's status in the trials register is "unknown".

NCT01709188

Methods	Parallel-group RCT
Participants	Estimated 30 people with mild-to-moderate dementia in Taiwan
Interventions	Experimental group: "Musical Dual Task Training protocol is structured with musical content and patients are required to do musical tasks including singing and playing instruments contingent on visual or auditory cues while walking" delivered by a "qualified music therapist."
	Control group: "walking and talking:" "read a newspaper article prior to a walk and have a conversation with the music therapist based on the content of the news while walking."
Outcomes	Cognition (primary outcome); agitation (secondary outcome
	In addition, there were outcomes other than the 7 outcomes of interest for our Cochrane review.
Notes	Estimated primary completion date October 2013. However, ClinicalTrial.gov reported (status 17 April 2017): "Study has passed its completion date and status has not been verified in more than two years."

NCT01958983

Methods	Parallel-group RCT
Participants	Estimated 30 people with "a mild dementia diagnosis" (or "mild to moderate" dementia) in Taiwan
Interventions	Experimental group: mixed active-receptive music therapy
	Control group: "no intervention" (usual care)
Outcomes	Quality of life, depression and agitation were secondary outcomes.
	In addition, there were outcomes other than the 7 outcomes of interest for our Cochrane review.
Notes	Estimated trial completion date: September 2014. However, ClinicalTrial.gov reported (status 17 April 2017): "Study has passed its completion date and status has not been verified in more than two years."

NCT02020356

Methods	Parallel-group RCT
Participants	35 people with Alzheimer's disease living in "an institution for the dependent elderly" in France, with MMSE score 5 to 20
Interventions	Experimental group: receptive intervention using "'U' sequence: the musical sequence lasts 20 minutes and is made up of several phases that progressively induce a relaxed state in the patient. The phase of maximum relaxation is followed by a stimulating phase."
	Control group: "interview with an occupational activity (such as discussion of personal pictures or news) with the caregiver in charge of music therapy sessions with the same period."
Outcomes	Quality of life, agitation and overall behavioural problems were secondary outcomes
	In addition, there were outcomes other than the 7 outcomes of interest for our Cochrane review.
Notes	ClinicalTrials.gov: the study was completed June 2015; the study has been terminated. No study results are posted (accessed 16 April 2017). If a report on possible results should become available, eligibility should be reviewed, in particular, whether the intervention meets our criteria for music-based therapeutic interventions.

NCT02147652

Methods	Unclear
Participants	Inclusion criteria
	• Diagnosis of dementia with possible or probable cause of Alzheimer's disease, vascular disease, mixed dementia
	• Moderate stage of dementia, MMSE score < 20
	• Age 60 to 90 years inclusive
	• Preserved hearing (hearing aids permissible)
	• Pittsburgh Agitation Scale score ≥ 3 on at least 3 occasions over 5 days
	Exclusion criteria
	• Auditory deficits requiring correction beyond hearing aids
	• No substitute decision maker available to indicate music preference and person unable to answer for themselves
	• Recent acute event, e.g. myocardial infarction, fractures, or major infection (not urinary tract infection)
	• Receiving standing orders of medication for personal care
Interventions	Listening to personalised and either non-personalised or no music during daily hygiene care (grooming)
Outcomes	Changes in agitation
Notes	Registered trial. Last update was in September 2019. Eligibility of the study has not yet been assessed.

NCT02670993

Methods	RCT (parallel)
Participants	59 people with mild Alzheimer's disease or mild cognitive impairment (but "Patient with a different etiology of cognitive disorder that of Alzheimer's disease" were excluded), in France
Interventions	Experimental group: singing sessions Control group: painting sessions
Outcomes	Primary outcome: "physical and moral pain" or "pain intensity" rated at "a simplified visual scale;" secondary outcome: other pain intensity scale (Brief Pain Inventory)
Notes	Study completed in June 2016. When study results become available, we will assess whether people with no dementia were included, whether we accept pain as an outcome for the review and whether analyses included outcomes assessed after < 5 sessions.

Prick 2024

Methods	3-arm parallel-group RCT
Participants	Nursing home residents with dementia in the Netherlands
Interventions	Individual music therapy versus another group listening to individualised music and a control group receiving usual care
Outcomes	Behavioural problems overall (primary outcome), quality of life (QUALIDEM as a primary outcome and Positive Response Scale as a secondary outcome measure for well-being) and communication behaviour (secondary outcome)
Notes	Eligibility of the study has not been assessed yet.

권서령 2013

Methods	"Pretest-posttest control group design" and "people were randomly assigned to the experimental and control groups"
Participants	34 people with dementia attending a daycare centre in South Korea
Interventions	Experimental group: music therapy Control group: usual care or other not reported in the abstract
Outcomes	Cognition
Notes	We could not retrieve the full text. First, we would like to evaluate if this was an RCT.

김현정 2013

Methods	Parallel-group RCT
Participants	20 people with mild dementia "who reside in G Welfare Foundation in D city" (Korea)

김현정 **2013** *(Continued)*

Interventions	Experimental group: group music therapy
	Control group: usual care or other not reported in the abstract
Outcomes	Quality of life and depression
Notes	We could not retrieve the full text. Type of analysis is not clear from the abstract. We would need to review if analyses were limited to effects after ≥ 5 sessions.

신보영, 황은영 **2015**

Methods	Unclear ("17 of them were assigned to experimental group and the other 17 people were assigned to control group. The musical activities with visual supportive strategies were carried out both experimental group and control group for 10 sessions")
Participants	34 people with dementia attending a daycare centre in South Korea
Interventions	Experimental group: musical activities with visual supportive strategies
	Control group: unclear
Outcomes	Cognition
Notes	Unclear if this was an RCT and how effectiveness could be derived if the control group received the same intervention ("According to this results, it was shown that the musical activities with visual supportive strategies were effective intervention for the cognitive rehabilitation of elderly people with dementia"). It is also unclear if this is music therapy or a combination of more types of therapy. We still need to retrieve the full text to evaluate eligibility.

ADL: activities of daily living; BEHAVE-AD: Behavioural Pathology in Alzheimer's Disease; MMSE: Mini-Mental State Examination; MMSE-K: Mini-Mental State Examination - Korean Version; n: number; RCT: randomised controlled trial

Characteristics of ongoing studies *[ordered by study ID]*

Baroni Caramel 2024

Study name	Individual music therapy in nursing home residents with dementia to improve general well-being
Methods	Parallel-group RCT
Participants	Nursing home residents in the Netherlands with behavioural problems
Interventions	Individual music therapy versus a control group receiving individual attention
Outcomes	Primary outcome: observed discomfort with DS-DAT
	Secondary outcomes: quality of life (measured with QUALID), depressive symptoms (CSDD), agitation (CMAI) and neuropsychiatric symptoms (NPI-NH)
Starting date	2019
Contact information	Dr. Baroni-Caramel, vanusa@baroni-caramel.com
Notes	

Gold 2018

Study name	Music Interventions for Dementia and Depression in ELderly care (MIDDEL)
Methods	Cluster-randomised trial
Participants	Nursing home residents recruited in Australia, Germany, the Netherlands, Norway, Turkey and the UK
Interventions	Care home units were cluster-randomised to one of four conditions: only group music therapy (GMT), only recreational choir singing (RCS), both GMT and RCS, or standard care (neither GMT nor RCS).
Outcomes	Primary outcome: depressive symptoms score at 6 months, measured using the MADRS
	Secondary outcomes included neuropsychiatric symptoms (NPI-Q) and quality of life from un-masked proxy (care staff) or self-reports (QoL-AD; self-rated), (EQ-5D-5L; proxy). Effect of the interventions on staff burden was assessed using the PCTB.
	See the study protocol for further outcomes.
Starting date	15 June 2018 (Australia)
Contact information	Dr Christian Gold, christian. gold@ norceresearch.no
Notes	The Australian part of the trial was published (Baker 2022) and included in this 2025 review update.

NCT04666077

Study name	Music therapy experiences in patients with mild cognitive impairment (MCI) and Alzheimer's disease (AD)
Methods	3-arm parallel-group RCT
Participants	Community-dwelling persons with AD or MCI in the USA
Interventions	Individual music therapy versus a control group receiving individual supported singing and a control group receiving individual attention
Outcomes	Primary outcomes: behavioural problems overall; cognition measured by MMSE
	Secondary outcome: cognition measured by ADAS-Cog
Starting date	2019
Contact information	Adeleine Dauray, adeleine.dauray@mssm.edu
Notes	Inclusion depends on whether data are available for the subgroup of persons with dementia. Status 2 March 2024: last refreshed on 22 May 2023 and still recruiting at that time

AD: Alzheimer's disease; ADAS-Cog: Alzheimer's Disease Assessment Scale-Cognitive Subscale; CMAI: Cohen-Mansfield Agitation Inventory; CSDD: Cornell Scale for Depression in Dementia; DS-DAT: Discomfort Scale - Dementia of Alzheimer Type; EQ-5D-5L: EuroQol-5 Dimensions 5 Levels questionnaire; MADRS: Montgomery- Asberg Depression Rating Scale; MCI: mild cognitive impairment; MIDDEL: Music Interventions for Dementia and Depression in ELderly; MMSE: Mini-Mental State Examination; NPI-NH: Neuropsychiatric Inventory Nursing Home version; NPI-Q: Neuropsychiatric Inventory Questionnaire; PAIC-15: Pain Assessment in Impaired Cognition; PCTB: Professional

Care Team Burden Scale; QoL-AD: Quality of Life in Alzheimer's Disease, QUALID: Quality of Life in Late-Stage Dementia; RCT: randomised controlled trial

DATA AND ANALYSES

Comparison 1. Music-based therapeutic interventions versus usual care: end of treatment

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1.1 Emotional well-being including quality of life	4	154	Std. Mean Difference (IV, Random, 95% CI)	0.14 [-0.29, 0.56]
1.2 Mood disturbance or negative affect: depression	9	441	Std. Mean Difference (IV, Random, 95% CI)	-0.23 [-0.42, -0.04]
1.3 Mood disturbance or negative affect: anxiety	7	282	Std. Mean Difference (IV, Random, 95% CI)	-0.15 [-0.39, 0.09]
1.4 Behaviour problems: agitation or aggression	11	503	Std. Mean Difference (IV, Random, 95% CI)	-0.05 [-0.27, 0.17]
1.5 Behaviour problems: overall	10	385	Std. Mean Difference (IV, Random, 95% CI)	-0.31 [-0.60, -0.02]
1.6 Social behaviour	2	121	Std. Mean Difference (IV, Random, 95% CI)	0.22 [-0.14, 0.57]
1.7 Cognition	7	353	Std. Mean Difference (IV, Random, 95% CI)	0.19 [-0.02, 0.41]

Analysis 1.1. Comparison 1: Music-based therapeutic interventions versus usual care: end of treatment, Outcome 1: Emotional well-being including quality of life

Study or Subgroup	Music-based therapeutic interventions Mean	SD	Total	Usual care Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI	Std. Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
Ridder 2013a	333.26	62.57	20	315.66	76.46	21	30.6%	0.25 [-0.37 , 0.86]		? + + + + ? +
Raglio 2015b	4.9	6.9	20	4.6	9.6	40	36.2%	0.03 [-0.50 , 0.57]		+ + + + + ? +
Hsu 2015c	1.8	0.59	3	0.61	0.49	5	4.4%	1.97 [-0.00 , 3.94]		+ + – – + – +
Baker 2022d	26.1	6.5	13	27	7	32	28.8%	-0.13 [-0.77 , 0.52]		+ + – ? + ? +
Total (Walde)			**56**			**98**	**100.0%**	**0.14 [-0.29 , 0.56]**		

Test for overall effect: Z = 0.63 (P = 0.53)
Test for subgroup differences: Not applicable
Heterogeneity: Tau² (DLf) = 0.05; Chi² = 4.21, df = 3 (P = 0.24); I² = 29%

-4 -2 0 2 4
Favours usual care   Favours music-based therapy

Footnotes
aHigher scores reflect higher quality of life.
bIntervention group size reduced by 50% because of two control groups. Higher scores reflect higher quality of life.
cHigher scores reflect higher well-being. Published data not corrected for cluster design, so group sizes adapted (ICC 0.10, design effect 1.5).
dIntervention group size reduced by 50% because of two control groups; higher scores: higher self-report quality of life.
eCI calculated by Wald-type method.
fTau² calculated by DerSimonian and Laird method.

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 1.2. Comparison 1: Music-based therapeutic interventions versus usual care: end of treatment, Outcome 2: Mood disturbance or negative affect: depression

Study or Subgroup	Music-based therapeutic interventions Mean	SD	Total	Usual care Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI	Std. Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
Svansdottir 2006a	6.1	4.3	16	6.4	4.8	10	6.0%	-0.06 [-0.85 , 0.73]		? ? – + + ? ?
Raglio 2010ab	1	1.819	27	1.5	2.735	24	12.3%	-0.21 [-0.77 , 0.34]		? ? – – + ? +
Raglio 2010bb	1	2.8	10	2	2.8	10	4.8%	-0.34 [-1.23 , 0.54]		+ + + ? + + +
Lin 2011	8.22	7.12	49	13.78	9.59	51	23.0%	-0.65 [-1.05 , -0.25]		+ + + – + + +
Ceccato 2012c	9.66	6.17	27	8.96	6.8	23	12.0%	0.11 [-0.45 , 0.66]		+ + – – + ? +
Raglio 2015d	7.7	4.421	20	8.83	6.042	40	12.9%	-0.20 [-0.74 , 0.34]		+ + + + + ? +
Giovagnoli 2018	1.86	2.82	23	1.9	3.48	22	10.9%	-0.01 [-0.60 , 0.57]		+ ? + + + + +
Gómez-Gallego 2021e	6.46	1.55	4	5.85	1.84	12	2.9%	0.32 [-0.82 , 1.46]		? + – – – ? +
Baker 2022f	12	9.7	23	13.7	7.8	50	15.2%	-0.20 [-0.69 , 0.30]		+ + – + – ? +
Total (Waldg)			**199**			**242**	**100.0%**	**-0.23 [-0.42 , -0.04]**		

Test for overall effect: Z = 2.35 (P = 0.02)
Test for subgroup differences: Not applicable
Heterogeneity: Tau² (DLh) = 0.00; Chi² = 7.31, df = 8 (P = 0.50); I² = 0%

-2 -1 0 1 2
Favours music-based therapy   Favours usual care

Footnotes
aDepression subscale of BEHAVE-AD data provided by the author
bDepression subscale of NPI data provided by the author
cWe calculated end-of-treatment scores from baseline and change scores, and we adopted the SD of the baseline scores.
dIntervention group size reduced by 50% because of two control groups. Means and SD of the Cornell scale were provided by the author
eTwo control groups so intervention group reduced by 50%; uncorrected for clusters, so group sizes adapted (ICC 0.09, design effect 3.34); assuming 2/30 dropped out
fIntervention group size reduced by 50% because of two control groups
gCI calculated by Wald-type method.
hTau² calculated by DerSimonian and Laird method.

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 1.3. Comparison 1: Music-based therapeutic interventions versus usual care: end of treatment, Outcome 3: Mood disturbance or negative affect: anxiety

Study or Subgroup	Music-based therapeutic interventions Mean	SD	Total	Usual care Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI	Std. Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
Svansdottir 2006[a]	0.7	1.3	20	0.4	1.1	18	13.9%	0.24 [-0.40 , 0.88]		? ? ● ⊕ ? ? ?
Raglio 2010a[b]	1	1.71	27	1.67	2.899	24	18.6%	-0.28 [-0.83 , 0.27]		? ? ● ⊕ ⊕ ? ⊕
Raglio 2010b[b]	3.1	3.9	10	3.1	2	10	7.4%	0.00 [-0.88 , 0.88]		⊕ ⊕ ● ? ⊕ ? ⊕
Sung 2012	3.89	4.02	27	5.36	4.34	28	20.0%	-0.35 [-0.88 , 0.19]		⊕ ? ● ● ? ? ⊕
Sakamoto 2013[c]	0.3	0.6	7	1.2	1.7	13	6.4%	-0.60 [-1.54 , 0.34]		? ? ● ⊕ ⊕ ⊕ ⊕
Raglio 2015[d]	2.64	2.769	18	3.69	3.225	35	17.3%	-0.34 [-0.91 , 0.24]		⊕ ⊕ ● ⊕ ⊕ ? ⊕
Giovagnoli 2018	2.41	3.58	23	1.71	2.49	22	16.5%	0.22 [-0.36 , 0.81]		⊕ ? ● ⊕ ⊕ ? ⊕
Total (Wald[e])			**132**			**150**	**100.0%**	**-0.15 [-0.39 , 0.09]**		

Test for overall effect: Z = 1.22 (P = 0.22)
Test for subgroup differences: Not applicable
Heterogeneity: Tau² (DL[f]) = 0.00; Chi² = 5.14, df = 6 (P = 0.53); I² = 0%

-2 -1 0 1 2
Favours music-based therapy   Favours usual care

Footnotes
[a]Anxieties and phobias subscale score of BEHAVE-AD, data provided by the author
[b]Anxiety subscale score of NPI, data about control group provided by the author
[c]Anxiety and phobias subscale of BEHAVE-AD. Intervention group size reduced by 50% because of two control groups
[d]Anxiety subscale score of NPI, data provided by the author. Intervention group size reduced by 50% because of two control groups
[e]CI calculated by Wald-type method.
[f]Tau² calculated by DerSimonian and Laird method.

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 1.4. Comparison 1: Music-based therapeutic interventions versus usual care: end of treatment, Outcome 4: Behaviour problems: agitation or aggression

Study or Subgroup	Music-based therapeutic interventions Mean	SD	Total	Usual care Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI	Std. Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
Clark 1998	65.56	58.02	18	121.56	119.23	18	7.9%	-0.58 [-1.25 , 0.08]		? ? ● ? ? ? ●
Svansdottir 2006a	1.2	1.7	20	1.3	1.6	18	8.5%	-0.06 [-0.70 , 0.58]		? ? ● ● ? ? ?
Raglio 2010ab	1.41	1.907	27	2.38	3.386	24	10.3%	-0.35 [-0.91 , 0.20]		? ? ● ● ● ? ●
Raglio 2010bb	2.5	4.2	10	1.6	2.1	10	5.2%	0.26 [-0.62 , 1.14]		● ● ● ? ● ? ●
Lin 2011c	36.37	10.64	49	38.55	10.27	51	15.2%	-0.21 [-0.60 , 0.19]		● ● ● ● ● ● ●
Ceccato 2012d	25.63	15.88	27	22.8	12.73	23	10.2%	0.19 [-0.37 , 0.75]		● ● ● ● ● ● ●
Sung 2012e	32.7	4.98	27	31	2.96	28	10.8%	0.41 [-0.12 , 0.95]		● ? ● ● ? ● ●
Ridder 2013f	26.09	13.54	17	28	18.15	18	8.0%	-0.12 [-0.78 , 0.55]		? ● ● ● ● ? ●
Sakamoto 2013g	0.7	1	7	3.2	3	13	4.4%	-0.95 [-1.93 , 0.02]		? ? ● ● ● ● ●
Raglio 2015h	3.78	3.053	18	3.77	3.011	35	10.0%	0.00 [-0.57 , 0.57]		● ● ● ● ● ? ●
Giovagnoli 2018i	1.41	2.82	23	0.38	0.97	22	9.4%	0.48 [-0.12 , 1.07]		● ? ● ● ● ? ●
Total (Waldj)			**243**			**260**	**100.0%**	**-0.05 [-0.27 , 0.17]**		

Test for overall effect: Z = 0.42 (P = 0.67)
Test for subgroup differences: Not applicable
Heterogeneity: Tau² (DLk) = 0.04; Chi² = 14.62, df = 10 (P = 0.15); I² = 32%

-2   -1   0   1   2
Favours music-based therapy   Favours usual care

Footnotes

aAggressiveness subscale score of BEHAVE-AD, data provided by coauthor
bAgitation subscale score of NPI, data about control group provided by the author
cNo SD of the mean difference of CMAI scores was reported; we applied the SD of the differences found by Ceccato 2012
dWe calculated end of treatment scores from baseline and change scores, and we adopted the SD of the baseline scores.
eOutcomes at 6 weeks, by direct observation in intervention group using some modified version of CMAI
fAdapted CMAI with different range; note that an effect size is reported but based on SD baseline
gAggressiveness subscale of the BEHAVE-AD. Intervention group size reduced by 50% because of two control groups
hAgitation subscale score of NPI, data provided by the author. Intervention group size reduced by 50% because of two control groups
iThe item in the article was called "stirring". The authors confirmed this was in fact the agitation item
jCI calculated by Wald-type method.
kTau² calculated by DerSimonian and Laird method.

Risk of bias legend

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 1.5. Comparison 1: Music-based therapeutic interventions versus usual care: end of treatment, Outcome 5: Behaviour problems: overall

Study or Subgroup	Music-based therapeutic interventions			Usual care			Weight	Std. Mean Difference IV, Random, 95% CI	Std. Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
	Mean	SD	Total	Mean	SD	Total				
Svansdottir 2006a	4.4	4.7	20	4.7	5.6	18	11.5%	-0.06 [-0.69 , 0.58]		? ? − + ? ? ?
Raglio 2010ab	8.86	7.317	27	19.04	21.666	26	13.3%	-0.63 [-1.18 , -0.07]		? ? − + ? + +
Raglio 2010b	14.8	17.3	10	13.9	8.6	10	7.7%	0.06 [-0.81 , 0.94]		+ + − ? + + +
Sakamoto 2013c	0.7	0.6	7	1.5	0.8	13	6.5%	-1.04 [-2.02 , -0.05]		? ? − + + + +
Lyu 2014d	13.52	11.63	16	15.14	11.58	30	12.1%	-0.14 [-0.74 , 0.47]		+ ? − ? + + +
Raglio 2015c	23.7	10.7	20	28.9	13.3	40	13.5%	-0.41 [-0.95 , 0.13]		+ + − + + + +
Hsu 2015e	12.33	11.2	4	26.57	7.14	5	3.0%	-1.39 [-2.96 , 0.18]		+ + − + − − +
Giovagnoli 2018	21.05	20.58	23	14.52	16.9	22	12.4%	0.34 [-0.25 , 0.93]		+ ? − + + ? +
Gómez-Gallego 2021f	11.36	4.01	5	27.9	13.14	16	5.5%	-1.34 [-2.44 , -0.24]		+ + − + + ? +
Baker 2022	6.6	6.4	23	7.2	5.2	50	14.7%	-0.11 [-0.60 , 0.39]		+ + − ? + ? +
Total (Waldg)			**155**			**230**	**100.0%**	**-0.31 [-0.60 , -0.02]**		

Test for overall effect: Z = 2.11 (P = 0.04)
Test for subgroup differences: Not applicable
Heterogeneity: Tau² (DLh) = 0.09; Chi² = 15.48, df = 9 (P = 0.08); I² = 42%

-2 -1 0 1 2
Favours music-based therapy Favours usual care

Footnotes

aSD provided by the author
bNPI end-of-treatment values and SD presented in Figure 1 in the main paper as provided by the author
cTotal scores, subscale scores included elsewhere. Intervention group size reduced by 50% because of two control groups
dIntervention group size reduced by 50% because of two control groups
ePublished data not corrected for cluster design, so group sizes adapted (ICC 0.06, design effect 1.36)
fTwo control groups so intervention group reduced by 50%; uncorrected for clusters, so group sizes adapted (ICC 0.06, design effect 2.56). We assumed 2/30 dropped out
gCI calculated by Wald-type method.
hTau² calculated by DerSimonian and Laird method.

Risk of bias legend

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 1.6. Comparison 1: Music-based therapeutic interventions versus usual care: end of treatment, Outcome 6: Social behaviour

Study or Subgroup	Music-based therapeutic interventions			Usual care			Weight	Std. Mean Difference IV, Random, 95% CI	Std. Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
	Mean	SD	Total	Mean	SD	Total				
Giovagnoli 2018	5.41	1.44	23	5.14	1.62	22	37.3%	0.17 [-0.41 , 0.76]		+ ? − + + ? +
Tang 2018a	28.32	8.97	37	26.18	8.54	39	62.7%	0.24 [-0.21 , 0.69]		+ ? − + + ? +
Total (Waldb)			**60**			**61**	**100.0%**	**0.22 [-0.14 , 0.57]**		

Test for overall effect: Z = 1.19 (P = 0.24)
Test for subgroup differences: Not applicable
Heterogeneity: Tau² (DLc) = 0.00; Chi² = 0.03, df = 1 (P = 0.86); I² = 0%

-2 -1 0 1 2
Favours usual care Favours music-based therapy

Footnotes

aMeans reverted by subtracting original means from 48 (scale maximum) in order to harmonize direction
bCI calculated by Wald-type method.
cTau² calculated by DerSimonian and Laird method.

Risk of bias legend

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 1.7. Comparison 1: Music-based therapeutic interventions versus usual care: end of treatment, Outcome 7: Cognition

Study or Subgroup	Music-based therapic interventions			Usual care			Weight	Std. Mean Difference IV, Random, 95% CI	Std. Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
	Mean	SD	Total	Mean	SD	Total				
Raglio 2010b	16	6	10	13	6	10	5.7%	0.48 [-0.41 , 1.37]		
Lin 2011	15.72	6.53	49	13.82	4.36	51	28.9%	0.34 [-0.05 , 0.74]		
Ceccato 2012a	16.26	3.66	27	16.39	3.9	23	14.6%	-0.03 [-0.59 , 0.52]		
Lyu 2014b	17.64	5.3	16	17.91	3.1	30	12.2%	-0.07 [-0.67 , 0.54]		
Tang 2018	16.71	5.03	37	15.01	4.52	39	21.9%	0.35 [-0.10 , 0.81]		
Giovagnoli 2018	15.82	8.04	23	15.43	7.18	22	13.2%	0.05 [-0.53 , 0.63]		
Gómez-Gallego 2021c	19.57	3.8	4	19.83	3.31	12	3.5%	-0.07 [-1.20 , 1.06]		
Total (Wald d)			**166**			**187**	**100.0%**	**0.19 [-0.02 , 0.41]**		

Test for overall effect: Z = 1.79 (P = 0.07)

Test for subgroup differences: Not applicable

Heterogeneity: Tau² (DLe) = 0.00; Chi² = 3.19, df = 6 (P = 0.78); I² = 0%

-2 -1 0 1 2
Favours usual care   Favours music-based therapy

Footnotes

a We calculated end of treatment scores from baseline and change scores and we adopted the SD of the baseline scores

b Intervention group size reduced by 50% because of two control groups

c Two control groups so intervention group reduced by 50%; uncorrected for clusters, so group sizes adapted (ICC 0.09, design effect 3.34). We assumed 2/30 residents dropped out.

d CI calculated by Wald-type method.

e Tau² calculated by DerSimonian and Laird method.

Risk of bias legend

(A) Random sequence generation (selection bias)

(B) Allocation concealment (selection bias)

(C) Blinding of participants and personnel (performance bias)

(D) Blinding of outcome assessment (detection bias)

(E) Incomplete outcome data (attrition bias)

(F) Selective reporting (reporting bias)

(G) Other bias

Comparison 2. Music-based therapeutic interventions versus usual care: long-term effects

Outcome or subgroup title	No. of studies	No. of partici-pants	Statistical method	Effect size
2.1 Emotional well-being including quality of life	3	86	Std. Mean Difference (IV, Random, 95% CI)	0.17 [-0.80, 1.14]
2.2 Mood disturbance or negative affect: depression	5	276	Std. Mean Difference (IV, Random, 95% CI)	-0.07 [-0.31, 0.18]
2.3 Mood disturbance or negative affect: anxiety	3	141	Std. Mean Difference (IV, Random, 95% CI)	-0.06 [-0.48, 0.37]
2.4 Behavioural problems: agitation or aggression	4	241	Std. Mean Difference (IV, Random, 95% CI)	-0.17 [-0.42, 0.09]
2.5 Behavioural problems: overall	6	245	Std. Mean Difference (IV, Random, 95% CI)	-0.19 [-0.52, 0.14]
2.6 Cognition	2	146	Std. Mean Difference (IV, Random, 95% CI)	0.09 [-0.24, 0.41]

Analysis 2.1. Comparison 2: Music-based therapeutic interventions versus usual care: long-term effects, Outcome 1: Emotional well-being including quality of life

Study or Subgroup	Music-based therapeutic interventions			Usual care			Weight	Std. Mean Difference IV, Random, 95% CI	Std. Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
	Mean	SD	Total	Mean	SD	Total				
Raglio 2015a	5.5	6.3	20	4.3	9.1	40	48.2%	0.14 [-0.39 , 0.68]		
Hsu 2015b	1.76	0.48	3	0.47	0.68	5	18.0%	1.81 [-0.09 , 3.70]		
Baker 2022c	24.1	7.2	5	28.2	5.4	13	33.8%	-0.66 [-1.72 , 0.40]		
Total (Waldd)			28			58	100.0%	0.17 [-0.80 , 1.14]		

Test for overall effect: Z = 0.34 (P = 0.73)
Test for subgroup differences: Not applicable
Heterogeneity: Tau² (DLe) = 0.44; Chi² = 5.12, df = 2 (P = 0.08); I² = 61%

Favours usual care   Favours music-based therapy

Footnotes

aHigher scores reflect better quality of life. Intervention group size reduced by 50% because of two control groups

bHigher scores reflect higher well-being, two months after end of treatment. Uncorrected for cluster design, so group sizes adapted (ICC 0.10, design effect 1.5)

cIntervention group size reduced by 50% because of two control groups; higher scores: higher self-report quality of life

dCI calculated by Wald-type method.

eTau² calculated by DerSimonian and Laird method.

Risk of bias legend

(A) Random sequence generation (selection bias)

(B) Allocation concealment (selection bias)

(C) Blinding of participants and personnel (performance bias)

(D) Blinding of outcome assessment (detection bias)

(E) Incomplete outcome data (attrition bias)

(F) Selective reporting (reporting bias)

(G) Other bias

Analysis 2.2. Comparison 2: Music-based therapeutic interventions versus usual care: long-term effects, Outcome 2: Mood disturbance or negative affect: depression

Study or Subgroup	Music-based therapeutic interventions			Usual care			Weight	Std. Mean Difference IV, Random, 95% CI	Std. Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
	Mean	SD	Total	Mean	SD	Total				
Svansdottir 2006a	6.9	6.6	12	7.2	4	10	8.5%	-0.05 [-0.89 , 0.79]		
Raglio 2010ab	1.41	3.238	27	1.33	2.792	24	19.7%	0.03 [-0.52 , 0.58]		
Lin 2011c	11.23	8.64	49	11.43	9.72	51	38.7%	-0.02 [-0.41 , 0.37]		
Raglio 2015d	8.27	5.449	20	8.5	6.437	40	20.7%	-0.04 [-0.57 , 0.50]		
Baker 2022e	9.6	8.3	11	13.3	9	32	12.5%	-0.41 [-1.10 , 0.28]		
Total (Waldf)			119			157	100.0%	-0.07 [-0.31 , 0.18]		

Test for overall effect: Z = 0.53 (P = 0.59)
Test for subgroup differences: Not applicable
Heterogeneity: Tau² (DLg) = 0.00; Chi² = 1.13, df = 4 (P = 0.89); I² = 0%

Favours music-based therapy   Favours usual care

Footnotes

aBEHAVE-AD depression sub scale data provided by the author and they represent the status four weeks after treatment ended

bNPI sub scale depression data provided by the author and represent the status one month after treatment ended (not used two months after treatment)

cThe data represent the status one month after treatment ended

dMeans and SD of the Cornell scale provided by the author. Intervention group size reduced by 50% because of two control groups

eIntervention group size reduced by 50% because of two control groups

fCI calculated by Wald-type method.

gTau² calculated by DerSimonian and Laird method.

Risk of bias legend

(A) Random sequence generation (selection bias)

(B) Allocation concealment (selection bias)

(C) Blinding of participants and personnel (performance bias)

(D) Blinding of outcome assessment (detection bias)

(E) Incomplete outcome data (attrition bias)

(F) Selective reporting (reporting bias)

(G) Other bias

Analysis 2.3. Comparison 2: Music-based therapeutic interventions versus usual care: long-term effects, Outcome 3: Mood disturbance or negative affect: anxiety

Study or Subgroup	Music-based therapeutic interventions			Usual care			Weight	Std. Mean Difference IV, Random, 95% CI	Std. Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
	Mean	SD	Total	Mean	SD	Total				
Svansdottir 2006[a]	0.8	1.5	19	0.3	0.6	18	29.0%	0.42 [-0.23 , 1.08]		? ? ⊖ ⊕ ? ? ?
Raglio 2010[a][b]	1.04	2.066	27	1.46	2.043	24	36.3%	-0.20 [-0.75 , 0.35]		? ? ⊖ ⊕ ⊕ ? ⊕
Raglio 2015[c]	2.14	2.428	18	3.03	3.101	35	34.7%	-0.30 [-0.87 , 0.27]		⊕ ⊕ ⊖ ⊕ ⊕ ? ⊕
Total (Wald[d])			**64**			**77**	**100.0%**	**-0.06 [-0.48 , 0.37]**		

Test for overall effect: Z = 0.26 (P = 0.80)
Test for subgroup differences: Not applicable
Heterogeneity: Tau² (DL[e]) = 0.05; Chi² = 3.05, df = 2 (P = 0.22); I² = 35%

-4 -2 0 2 4
Favours music-based therapy   Favours usual care

Footnotes
[a]BEHAVE-AD Anxieties and phobias sub scale SD provided by the author and the data represent the status four weeks after treatment ended
[b]NPI subscale anxiety data provided by the author and they represent the status one month after treatment ended (not used two months after treatment)
[c]Anxiety subscale score of NPI, data provided by the author. Intervention group size reduced by 50% because of two control groups
[d]CI calculated by Wald-type method.
[e]Tau² calculated by DerSimonian and Laird method.

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 2.4. Comparison 2: Music-based therapeutic interventions versus usual care: long-term effects, Outcome 4: Behavioural problems: agitation or aggression

Study or Subgroup	Music-based therapeutic interventions			Usual care			Weight	Std. Mean Difference IV, Random, 95% CI	Std. Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
	Mean	SD	Total	Mean	SD	Total				
Svansdottir 2006[a]	1.1	1.6	19	0.8	1.5	18	15.7%	0.19 [-0.46 , 0.84]		? ? ⊖ ⊕ ? ? ?
Raglio 2010[a][b]	1.63	2.115	27	2.54	3.464	24	21.5%	-0.32 [-0.87 , 0.24]		? ? ⊖ ⊕ ⊕ ? ⊕
Lin 2011[c]	35.69	9.99	49	37.75	9.7	51	42.5%	-0.21 [-0.60 , 0.19]		⊕ ⊕ ⊕ ? ⊕ ? ⊕
Raglio 2015[d]	3.11	2.964	18	3.8	3.833	35	20.3%	-0.19 [-0.76 , 0.38]		⊕ ⊕ ⊖ ⊕ ⊕ ? ⊕
Total (Wald[e])			**113**			**128**	**100.0%**	**-0.17 [-0.42 , 0.09]**		

Test for overall effect: Z = 1.26 (P = 0.21)
Test for subgroup differences: Not applicable
Heterogeneity: Tau² (DL[f]) = 0.00; Chi² = 1.49, df = 3 (P = 0.68); I² = 0%

-1 -0.5 0 0.5 1
Favours music-based therapy   Favours usual care

Footnotes
[a]BEHAVE-AD subscale aggressiveness. SD provided by the author. The data represent the status four weeks after treatment ended
[b]NPI subscale agitation data provided by the author and they represent the status one month after treatment ended (not used two months after treatment)
[c]The data represent the status one month after treatment ended
[d]Agitation subscale score of NPI, data provided by the author. Intervention group size reduced by 50% because of two control groups
[e]CI calculated by Wald-type method.
[f]Tau² calculated by DerSimonian and Laird method.
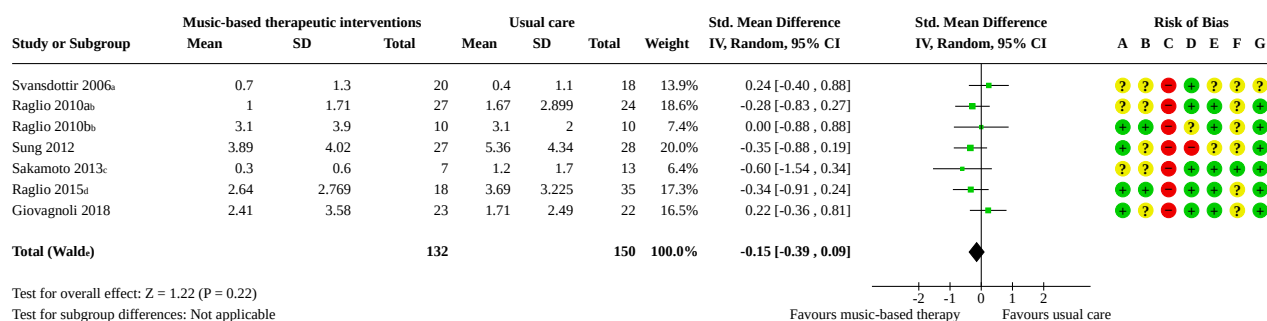
Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 2.5. Comparison 2: Music-based therapeutic interventions versus usual care: long-term effects, Outcome 5: Behavioural problems: overall

Study or Subgroup	Music-based therapeutic interventions			Usual care			Weight	Std. Mean Difference IV, Random, 95% CI	Std. Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
	Mean	SD	Total	Mean	SD	Total				
Svansdottir 2006[a]	5	4.9	19	3.5	3.3	18	17.7%	0.35 [-0.30 , 1.00]		? ? ⊖ ⊕ ? ? ?
Raglio 2010a[ab]	11.11	12	27	14.08	13.273	24	21.9%	-0.23 [-0.78 , 0.32]		? ? ⊖ ⊕ ⊕ ? ⊕
Lyu 2014[c]	13.01	11.72	16	15.42	9.72	30	19.4%	-0.23 [-0.84 , 0.38]		⊕ ? ⊖ ? ⊕ ? ⊕
Raglio 2015[c]	22.4	11.9	20	26.8	14.9	40	22.5%	-0.31 [-0.85 , 0.23]		⊕ ⊕ ⊖ ⊖ ⊕ ? ⊕
Hsu 2015[d]	8.67	9.54	4	34.43	7.37	4	2.1%	-2.63 [-4.91 , -0.35]		⊕ ⊕ ⊖ ⊖ ⊖ ⊖ ⊕
Baker 2022[c]	6.6	5.3	11	7.9	6.5	32	16.5%	-0.20 [-0.89 , 0.48]		⊕ ⊕ ⊖ ? ⊖ ? ⊕
Total (Wald[e])			**97**			**148**	**100.0%**	**-0.19 [-0.52 , 0.14]**		

Test for overall effect: Z = 1.12 (P = 0.26)
Test for subgroup differences: Not applicable
Heterogeneity: Tau² (DL[f]) = 0.05; Chi² = 7.27, df = 5 (P = 0.20); I² = 31%

Favours music-based therapy -4 -2 0 2 4 Favours usual care

Footnotes
[a]SD provided by the author and the data represent the status four weeks after treatment ended
[b]Data provided by the author and represent the status one month after treatment ended (not used two months after treatment)
[c]Intervention group size reduced by 50% because of two control groups
[d]Data two months after end of treatment, uncorrected for cluster design, so group sizes adapted (ICC 0.06, design effect 1.6)
[e]CI calculated by Wald-type method.
[f]Tau² calculated by DerSimonian and Laird method.

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 2.6. Comparison 2: Music-based therapeutic interventions versus usual care: long-term effects, Outcome 6: Cognition

Study or Subgroup	Music-based therapeutic interventions			Usual care			Weight	Std. Mean Difference IV, Random, 95% CI	Std. Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
	Mean	SD	Total	Mean	SD	Total				
Lin 2011[a]	14.24	6.39	49	13.5	4.6	51	70.5%	0.13 [-0.26 , 0.52]		⊕ ⊕ ⊖ ? ? ? ⊕
Lyu 2014[b]	17.81	4.7	16	17.94	4.7	30	29.5%	-0.03 [-0.63 , 0.58]		⊕ ? ⊖ ? ⊕ ? ⊕
Total (Wald[c])			**65**			**81**	**100.0%**	**0.09 [-0.24 , 0.41]**		

Test for overall effect: Z = 0.51 (P = 0.61)
Test for subgroup differences: Not applicable
Heterogeneity: Tau² (DL[d]) = 0.00; Chi² = 0.19, df = 1 (P = 0.67); I² = 0%

Favours usual care -2 -1 0 1 2 Favours music-based therapy

Footnotes
[a]The data represent the status one month after treatment ended
[b]Intervention group size reduced by 50% because of two control groups
[c]CI calculated by Wald-type method.
[d]Tau² calculated by DerSimonian and Laird method.

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
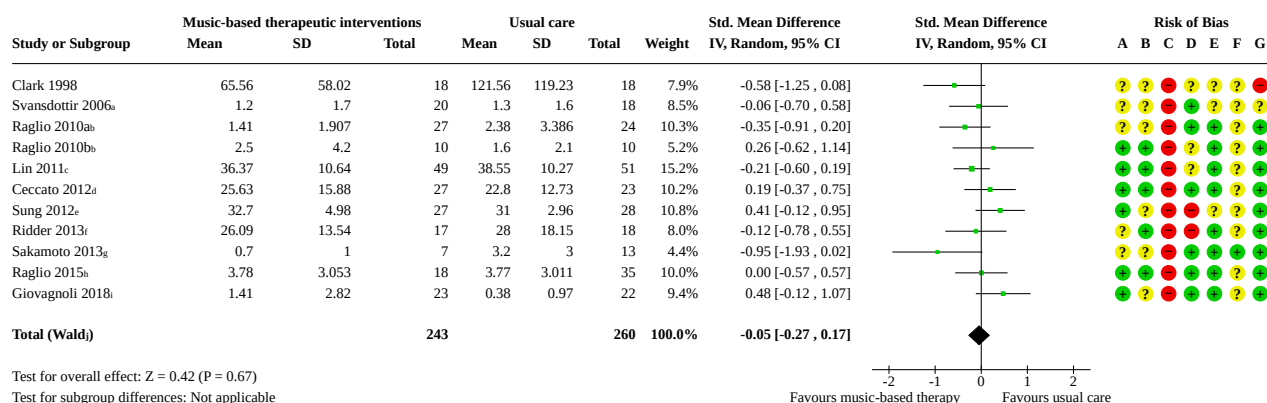(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
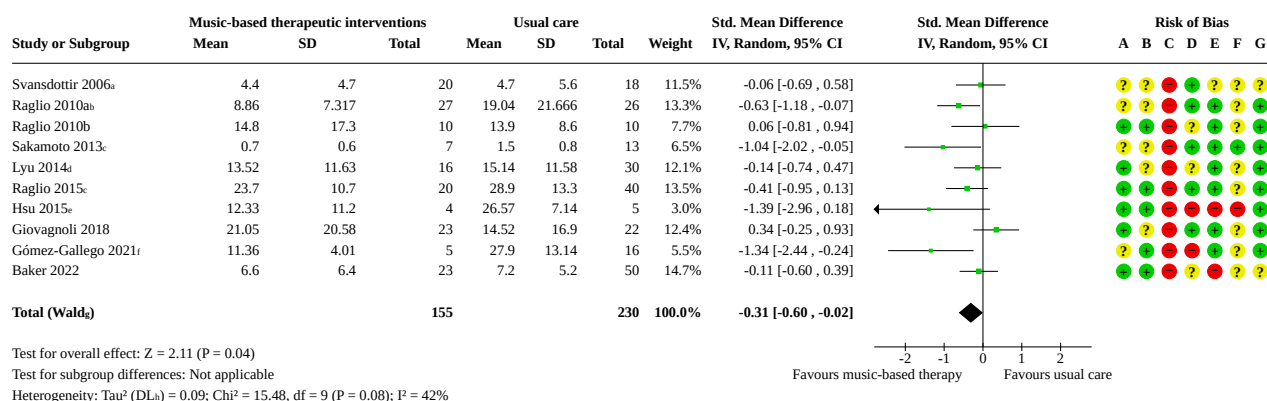(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 3. Music-based therapeutic interventions versus other activities: end of treatment

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
3.1 Emotional well-being including quality of life	9	298	Std. Mean Difference (IV, Random, 95% CI)	0.20 [-0.09, 0.49]

Outcome or subgroup title	No. of studies	No. of partici-pants	Statistical method	Effect size
3.2 Mood disturbance or nega-tive affect: depression	10	359	Std. Mean Difference (IV, Random, 95% CI)	-0.14 [-0.36, 0.08]
3.3 Mood disturbance or nega-tive affect: anxiety	10	291	Std. Mean Difference (IV, Random, 95% CI)	-0.75 [-1.27, -0.24]
3.4 Behaviour problems: agita-tion or aggression	6	168	Std. Mean Difference (IV, Random, 95% CI)	0.01 [-0.31, 0.32]
3.5 Behaviour problems: overall	8	292	Std. Mean Difference (IV, Random, 95% CI)	-0.08 [-0.33, 0.17]
3.6 Social behaviour	4	84	Std. Mean Difference (IV, Random, 95% CI)	0.52 [0.08, 0.96]
3.7 Cognition	5	147	Std. Mean Difference (IV, Random, 95% CI)	0.12 [-0.21, 0.45]

Analysis 3.1. Comparison 3: Music-based therapeutic interventions versus other activities: end of treatment, Outcome 1: Emotional well-being including quality of life

Study or Subgroup	Music-based therapeutic interventions Mean	SD	Total	Other activities Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI	Std. Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
Cooke 2010a	3.38	1	23	3.09	0.79781	23	15.2%	0.32 [-0.27, 0.90]		
Narme 2012-study 1b	12.02	38.49	12	-12.9	50.5	10	8.8%	0.54 [-0.32, 1.40]		
Narme 2012-study 2c	22.79	28.42	5	-37.97	20.89	6	2.8%	2.27 [0.59, 3.94]		
Narme 2014d	-9.79	37.2	18	-2.09	31.7	19	13.3%	-0.22 [-0.87, 0.43]		
Raglio 2015e	4.9	6.9	20	5.2	9.9	40	16.8%	-0.03 [-0.57, 0.50]		
Liesk 2015f	92.2	15.5	12	87.9	11.1	12	9.7%	0.31 [-0.50, 1.11]		
Cho 2016g	47.29	6.58	7	41.43	7.09	14	7.5%	0.81 [-0.14, 1.76]		
Cho 2016h	47.29	6.58	7	45.71	6.37	7	6.3%	0.23 [-0.82, 1.28]		
Reschke-Hernández 2019i	64.84	13.319	7	65.633	16.349	6	5.9%	-0.05 [-1.14, 1.04]		
Baker 2022j	26.1	6.5	13	27.4	7.2	37	13.7%	-0.18 [-0.81, 0.45]		
Total (Waldk)			**124**			**174**	**100.0%**	**0.20 [-0.09, 0.49]**		

Test for overall effect: Z = 1.37 (P = 0.17)
Test for subgroup differences: Not applicable
Heterogeneity: Tau² (DLl) = 0.05; Chi² = 12.14, df = 9 (P = 0.21); I² = 26%

-2 -1 0 1 2
Favours other activities   Favours music-based therapy

Footnotes
aHigher scores reflect higher quality of life. SD calculated from 95% CI with t distribution (Cooke et al 2010 J Health Psychol); used data from first period of cross-over only
bStudy 1 data. Balance of positive and negative emotional facial expressions as a percentage of total. Figure 2 data were provided by the author
cStudy 2 data. Balance of positive and negative emotional facial expressions as a percentage of total. Figure 2 data were provided by the author
dEmotional facial expressions, balance of positive and negative facial expressions as a percentage of total expressions
eHigher scores reflect better quality of life. Intervention number of residents reduced by 50% because of two control groups
fHigher scores reflect better quality of life. Both proxy and participant values are being reported; for the analyses we used patient report
gHigher scores reflect better quality of life. Control group: music listening. Intervention number of residents reduced by 50% because of two control groups
hHigher scores reflect better quality of life. Intervention group size reduced by 50% because of two control groups
iExact data obtained from author. Published data not corrected for cluster design, so group sizes adapted (ICC 0.10, design effect 2.5).
jIntervention number of residents reduced by 50% because of two control groups; higher scores: higher self-report quality of life
kCI calculated by Wald-type method.
lTau² calculated by DerSimonian and Laird method.

Risk of bias legend
(A) Random sequence generation (selection bias)
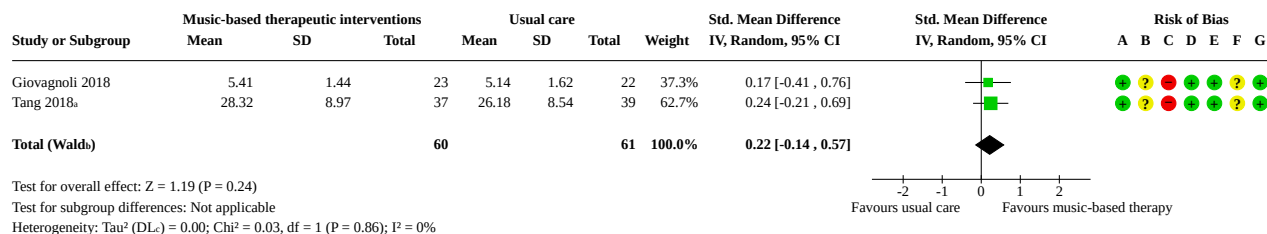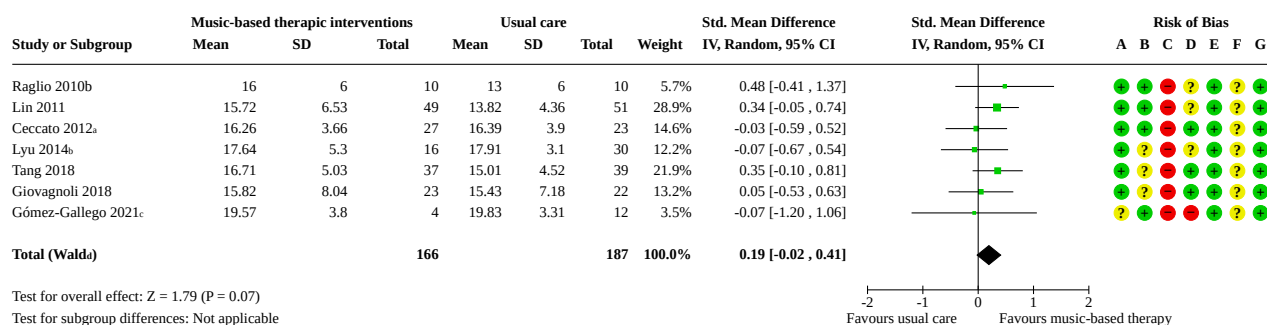(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
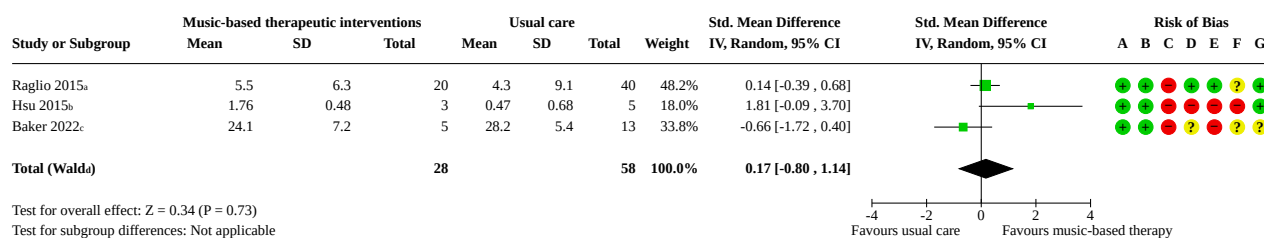(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 3.2. Comparison 3: Music-based therapeutic interventions versus other activities: end of treatment, Outcome 2: Mood disturbance or negative affect: depression

Study or Subgroup	Music-based therapeutic interventions Mean	SD	Total	Other activities Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI	Std. Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
Guétin 2009	8.9	3.3	14	11.2	6.1	12	7.7%	-0.46 [-1.25 , 0.32]		
Cooke 2010a	4.38	2.486	23	4.57	2.87906	23	14.2%	-0.07 [-0.65 , 0.51]		
Vink 2013b	0.14	0.535	14	0.33	0.816	6	5.1%	-0.29 [-1.25 , 0.67]		
Narme 2014c	0.3	0.7	18	0.5	1.5	19	11.3%	-0.17 [-0.81 , 0.48]		
Raglio 2015d	7.7	4.421	20	9.46	8.638	39	16.2%	-0.23 [-0.77 , 0.31]		
Thornley 2016e	0.667	1.155	3	1.6	1.673	5	2.2%	-0.53 [-2.01 , 0.94]		
Werner 2017f	13.37	8.84	10	17.58	9.37	7	4.9%	-0.44 [-1.42 , 0.54]		
Liu 2021	4.44	0.65	25	4.6	0.65	25	15.3%	-0.24 [-0.80 , 0.31]		
Gómez-Gallego 2021g	6.46	1.55	6	4.95	1.28	6	3.1%	0.98 [-0.25 , 2.21]		
Baker 2022h	12	9.7	22	11.4	8.1	62	20.0%	0.07 [-0.42 , 0.56]		
Total (Waldi)			**155**			**204**	**100.0%**	**-0.14 [-0.36 , 0.08]**		

Test for overall effect: Z = 1.28 (P = 0.20)
Test for subgroup differences: Not applicable
Heterogeneity: Tau² (DLj) = 0.00; Chi² = 5.60, df = 9 (P = 0.78); I² = 0%

-2 -1 0 1 2
Favours music-based therapy   Favours other activities

Footnotes

aSD calculated from 95% CI with t distribution
bDepression subscale score of NPI, data about control group provided by the author
cDepression subscale of NPI data provided by the author
dMeans and SD of the Cornell scale were provided by the author. Intervention group size reduced by 50% because of two control groups
eBased on data provided by authors
fSubgroup persons with physician diagnosis of dementia data provided by author. Uncorrected for clustering, so group sizes adapted (ICC 0.10, design effect 3.97)
gTwo control groups so intervention group reduced by 50%. Data uncorrected for clustering, so group sizes adapted (ICC 0.10, design effect 2.53); assuming 2/30 dropped out
hIntervention group size reduced by 50% because of two control groups
iCI calculated by Wald-type method.
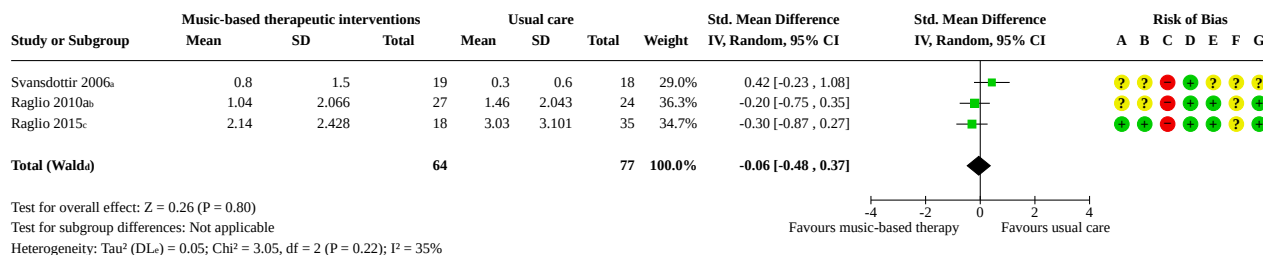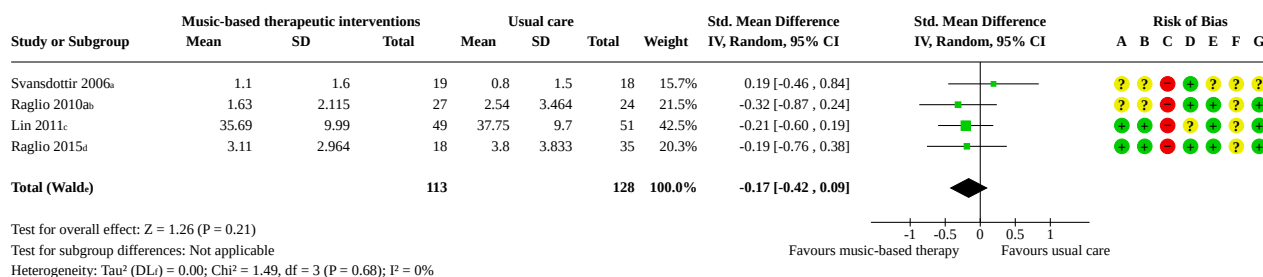jTau² calculated by DerSimonian and Laird method.

Risk of bias legend

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 3.3. Comparison 3: Music-based therapeutic interventions versus other activities: end of treatment, Outcome 3: Mood disturbance or negative affect: anxiety

Study or Subgroup	Music-based therapeutic interventions			Other activities			Weight	Std. Mean Difference IV, Random, 95% CI	Std. Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
	Mean	SD	Total	Mean	SD	Total				
Guétin 2009	8.4	3.7	14	20.8	6.2	12	9.2%	-2.40 [-3.45 , -1.35]		? ? ● ● ? ? ●
Cooke 2010[a]	7.58	7.111	23	11.26	7.65438	23	12.4%	-0.49 [-1.08 , 0.10]		? ? ● ● ● ● ●
Narme 2012-study 2[b]	-17.44	40.54	5	27.72	26.75	6	7.3%	-1.23 [-2.58 , 0.12]		? ? ● ● ● ? ●
Narme 2012-study 1[c]	-10.41	25.43	12	15.34	23.62	10	10.2%	-1.01 [-1.91 , -0.10]		? ? ● ● ? ? ●
Sakamoto 2013[d]	0.3	0.6	6	0.5	0.5	13	9.7%	-0.36 [-1.34 , 0.62]		? ? ● ● ● ● ●
Vink 2013[e]	0.07	0.267	14	0.5	0.837	6	9.5%	-0.83 [-1.83 , 0.17]		● ? ● ● ? ● ●
Narme 2014[f]	0.7	1.5	18	0.6	1.3	19	12.0%	0.07 [-0.58 , 0.71]		? ? ● ● ● ? ●
Raglio 2015[g]	2.64	2.769	18	4.18	3.655	34	12.5%	-0.45 [-1.03 , 0.13]		● ● ● ● ● ? ●
Thornley 2016[h]	8	6.928	3	0.4	0.894	5	5.2%	1.63 [-0.19 , 3.44]		● ● ● ● ● ? ●
Liu 2021	10.2	1.94	25	12.96	1.21	25	12.0%	-1.68 [-2.33 , -1.03]		● ● ● ● ● ? ●
Total (Wald[i])			**138**			**153**	**100.0%**	**-0.75 [-1.27 , -0.24]**		

Test for overall effect: Z = 2.86 (P = 0.004)
Test for subgroup differences: Not applicable
Heterogeneity: Tau² (DL[j]) = 0.47; Chi² = 33.34, df = 9 (P = 0.0001); I² = 73%

-4 -2 0 2 4
Favours music-based therapy   Favours other activities

Footnotes

[a]SD calculated from 95% CI with t distribution
[b]Study 2 data. Figure 2 means and SDs of STAI-A for the two studies provided by the authors. We reversed the scores so higher scores mean greater anxiety
[c]Study 1 data. Figure 2 means and SDs of STAI-A for the two studies provided by the authors. We reversed the scores so higher scores mean greater anxiety
[d]Anxiety and phobia subscale of BEHAVE-AD; total scores included elsewhere. Intervention group size reduced by 50% because of two control groups
[e]Anxiety subscale score of NPI, data about control group provided by the author
[f]Anxiety subscale score of NPI (STAI-A data not used because we preferred the more widely used NPI), data provided by the author
[g]Anxiety subscale score of NPI, data provided by the author. Intervention group size reduced by 50% because of two control groups
[h]Based on data provided by authors
[i]CI calculated by Wald-type method.
[j]Tau² calculated by DerSimonian and Laird method.
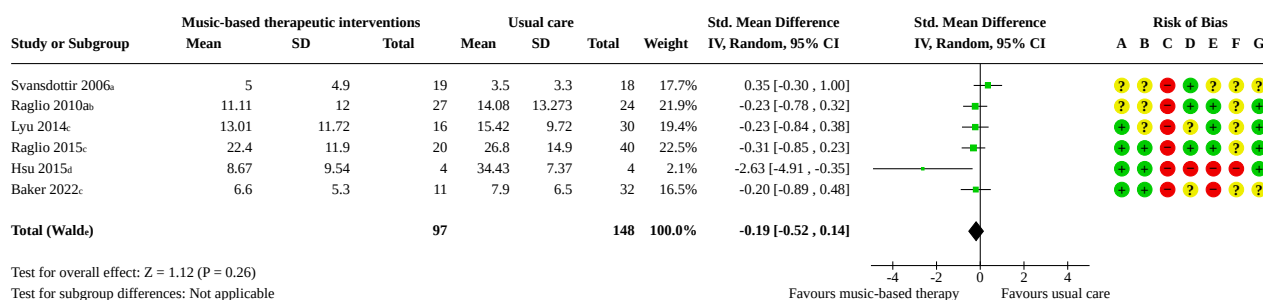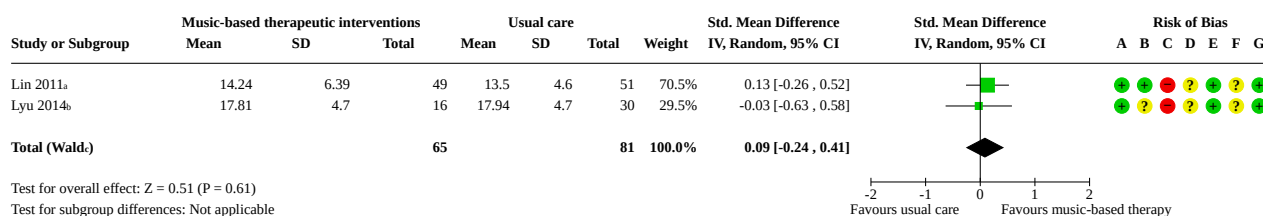
Risk of bias legend

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 3.4. Comparison 3: Music-based therapeutic interventions versus other activities: end of treatment, Outcome 4: Behaviour problems: agitation or aggression

Study or Subgroup	Music-based therapeutic interventions Mean	SD	Total	Other activities Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI	Std. Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
Cooke 2010[a]	1.67	0.416	23	1.66	0.68219	23	29.3%	0.02 [-0.56 , 0.60]		
Sakamoto 2013[b]	0.7	1	6	1.5	0.9	13	9.6%	-0.82 [-1.83 , 0.19]		
Vink 2013[c]	1	1.22	5	0.67	0.58	3	4.7%	0.27 [-1.17 , 1.72]		
Narme 2014	37.5	16.4	18	31.8	5.6	19	22.8%	0.46 [-0.19 , 1.11]		
Raglio 2015[d]	3.78	3.053	18	4.26	3.203	34	29.9%	-0.15 [-0.72 , 0.42]		
Thornley 2016[e]	84.33	28.39	3	78	28.73	3	3.8%	0.18 [-1.43 , 1.79]		
Total (Wald[f])			**73**			**95**	**100.0%**	**0.01 [-0.31 , 0.32]**		

Test for overall effect: Z = 0.04 (P = 0.97)
Test for subgroup differences: Not applicable
Heterogeneity: Tau² (DL[g]) = 0.00; Chi² = 4.89, df = 5 (P = 0.43); I² = 0%

-2 -1 0 1 2
Favours music-based therapy   Favours other activities

Footnotes
[a]SD calculated from 95% CI with t distribution
[b]Aggressiveness subscale of the NPI. Intervention group size reduced by 50% because of two control groups
[c]End-of-treatment data provided by the author
[d]Agitation sub scale score of NPI, data provided by the author. Intervention group size reduced by 50% because of two control groups
[e]Based on data provided by the author
[f]CI calculated by Wald-type method.
[g]Tau² calculated by DerSimonian and Laird method.

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
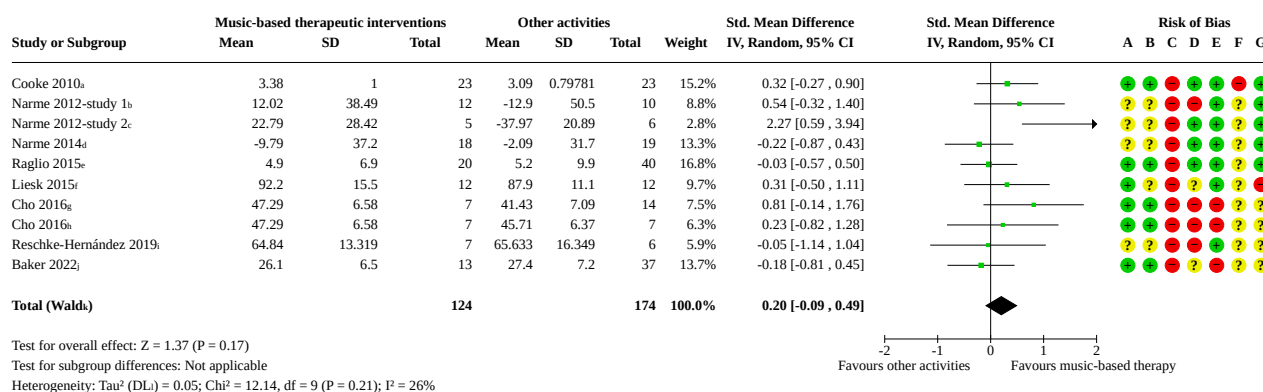(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 3.5. Comparison 3: Music-based therapeutic interventions versus other activities: end of treatment, Outcome 5: Behaviour problems: overall

Study or Subgroup	Music-based therapeutic interventions Mean	SD	Total	Other activities Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI	Std. Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
Sakamoto 2013[a]	0.7	0.6	6	0.8	0.4	13	6.5%	-0.20 [-1.17 , 0.77]		
Vink 2013[b]	3.67	3.31	15	4	2	6	6.8%	-0.10 [-1.05 , 0.84]		
Narme 2014	8.7	16.4	18	3.3	4.7	19	14.4%	0.44 [-0.21 , 1.10]		
Lyu 2014[c]	13.52	11.63	16	12.65	10.17	31	16.8%	0.08 [-0.52 , 0.68]		
Raglio 2015[d]	23.7	10.7	20	29.1	17	40	21.0%	-0.35 [-0.89 , 0.19]		
Thornley 2016[e]	9.33	7.572	3	7.5	16.263	4	2.7%	0.11 [-1.39 , 1.61]		
Gómez-Gallego 2021[f]	11.36	4.01	7	17.62	8.27	10	5.9%	-0.86 [-1.88 , 0.16]		
Baker 2022	6.6	6.4	22	7	6.5	62	25.9%	-0.06 [-0.55 , 0.43]		
Total (Wald[g])			**107**			**185**	**100.0%**	**-0.08 [-0.33 , 0.17]**		

Test for overall effect: Z = 0.63 (P = 0.53)
Test for subgroup differences: Not applicable
Heterogeneity: Tau² (DL[h]) = 0.00; Chi² = 6.08, df = 7 (P = 0.53); I² = 0%

-2 -1 0 1 2
Favours music-based therapy   Favours other activities

Footnotes
[a]Total scores, subscale scores included elsewhere. Intervention group size reduced by 50% because of two control groups
[b]End-of-treatment data provided by the author
[c]Intervention group size reduced by 50% because of two control groups
[d]Total scores, subscale scores included elsewhere. Intervention number of residents reduced by 50% because of two control groups
[e]Based on data provided by authors
[f]Two control groups so intervention group reduced by 50%. Published data uncorrected for clustering, group size adapted (ICC 0.06, design effect 2.02); assuming 2/30 dropped out
[g]CI calculated by Wald-type method.
[h]Tau² calculated by DerSimonian and Laird method.

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 3.6. Comparison 3: Music-based therapeutic interventions versus other activities: end of treatment, Outcome 6: Social behaviour

Study or Subgroup	Music-based therapeutic interventions Mean	SD	Total	Other actitvities Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI	Std. Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
Narme 2012-study 1[a]	17.31	28.89	12	-23.3	66.44	10	25.1%	0.79 [-0.09 , 1.67]		
Narme 2012-study 2[b]	54.76	34.64	5	-0.54	88.23	6	12.5%	0.73 [-0.52 , 1.97]		
Narme 2014[c]	22.69	31.7	18	6.9	53.3	19	45.7%	0.35 [-0.30 , 1.00]		
Reschke-Hernández 2019[d]	1.967	0.113	8	1.848	0.396988664825584	6	16.8%	0.41 [-0.66 , 1.49]		
Total (Wald[e])			**43**			**41**	**100.0%**	**0.52 [0.08 , 0.96]**		

Test for overall effect: Z = 2.31 (P = 0.02)
Test for subgroup differences: Not applicable
Heterogeneity: Tau² (DL$_f$) = 0.00; Chi² = 0.77, df = 3 (P = 0.86); I² = 0%

Favours other activities   Favours music-based therapy

Footnotes
[a]Study 1 data. Figure 2 data were provided by the author
[b]Study 2 data. Figure 2 data were provided by the author
[c]Measured by discourse content, counts of positive and negative words; higher scores mean more positive compared to negative words
[d]Reversed non-engagement Menorah Park Engagement Scale (MPES) data, obtained from author. Uncorrected for clustering, so group sizes adapted (ICC 0.09, design effect 2.35)
[e]CI calculated by Wald-type method.
[f]Tau² calculated by DerSimonian and Laird method.

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
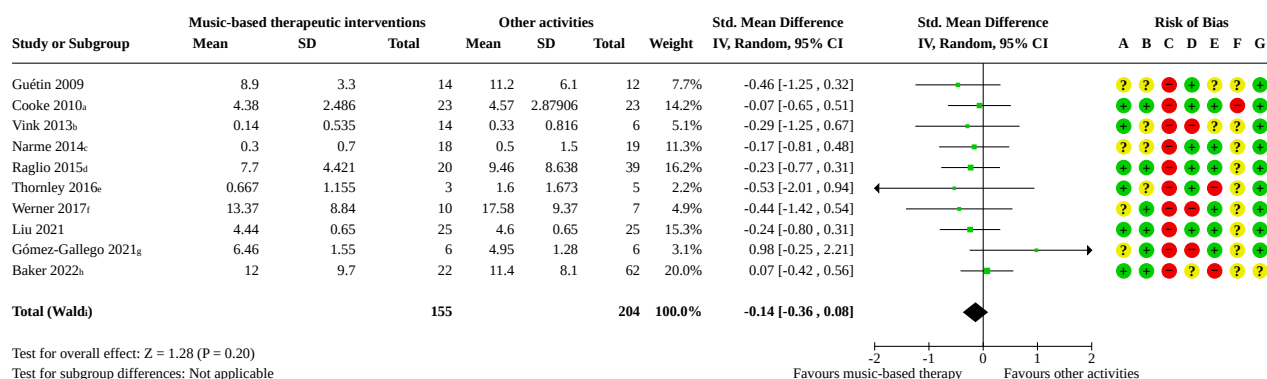(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 3.7. Comparison 3: Music-based therapeutic interventions versus other activities: end of treatment, Outcome 7: Cognition

Study or Subgroup	Music-based therapic interventions Mean	SD	Total	Other activities Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI	Std. Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
Guétin 2009	19.6	4.4	14	19.8	3.3	12	18.4%	-0.05 [-0.82 , 0.72]		
Lyu 2014[a]	17.64	5.3	16	17.57	4.1	31	30.0%	0.02 [-0.59 , 0.62]		
Narme 2014[b]	32.9	16.2	18	27.4	20.7	19	26.0%	0.29 [-0.36 , 0.94]		
Liesk 2015	20.1	3.7	12	19.6	5.9	12	17.0%	0.10 [-0.70 , 0.90]		
Gómez-Gallego 2021[c]	19.57	3.8	5	17.57	6.14	8	8.6%	0.34 [-0.78 , 1.47]		
Total (Wald[d])			**65**			**82**	**100.0%**	**0.12 [-0.21 , 0.45]**		

Test for overall effect: Z = 0.69 (P = 0.49)
Test for subgroup differences: Not applicable
Heterogeneity: Tau² (DL$_e$) = 0.00; Chi² = 0.71, df = 4 (P = 0.95); I² = 0%

Favours other activities   Favours music-based therapy

Footnotes
[a]Intervention group size reduced by 50% because of two control groups
[b]No end-of-treatment assessment with MMSE, included in analysis: results with the SIB with higher scores representing higher cognition same as MMSE
[c]Two control groups so intervention group reduced by 50%; published data uncorrected for clustering, group size adapted (ICC 0.09, design effect 2.62); assuming 2/30 dropped out
[d]CI calculated by Wald-type method.
[e]Tau² calculated by DerSimonian and Laird method.

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 4. Music-based therapeutic interventions versus other activities: long-term effects

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
4.1 Emotional well-being including quality of life	4	130	Std. Mean Difference (IV, Random, 95% CI)	0.10 [-0.29, 0.49]

Outcome or subgroup title	No. of studies	No. of partici-pants	Statistical method	Effect size
4.2 Mood disturbance or negative affect: depression	4	174	Std. Mean Difference (IV, Random, 95% CI)	-0.07 [-0.39, 0.25]
4.3 Mood disturbance or negative affect: anxiety	4	124	Std. Mean Difference (IV, Random, 95% CI)	-0.53 [-1.31, 0.25]
4.4 Behavioural problems: agitation or aggression	2	89	Std. Mean Difference (IV, Random, 95% CI)	0.10 [-0.66, 0.86]
4.5 Behavioural problems: overall	4	197	Std. Mean Difference (IV, Random, 95% CI)	-0.09 [-0.39, 0.22]
4.6 Social behaviour	2	48	Std. Mean Difference (IV, Random, 95% CI)	0.53 [-0.53, 1.60]
4.7 Cognition	1		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only

Analysis 4.1.   Comparison 4: Music-based therapeutic interventions versus other activities: long-term effects, Outcome 1: Emotional well-being including quality of life

	Music-based therapeutic interventions			Other activities				Std. Mean Difference	Std. Mean Difference	Risk of Bias
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI	IV, Random, 95% CI	A B C D E F G
Narme 2012-study 2[a]	-14.1	54.29	5	-41.66	18.25	6	9.6%	0.65 [-0.58 , 1.89]		? ? ● ● + ? +
Narme 2014[b]	-10.27	36.3	18	-31.9	59.7	19	31.7%	0.43 [-0.23 , 1.08]		? ? ● ● + ? +
Raglio 2015[c]	5.5	6.3	20	6.2	8.5	40	44.6%	-0.09 [-0.63 , 0.45]		+ + ● + + ? +
Baker 2022[d]	24.1	7.2	5	26.8	5.8	17	14.2%	-0.43 [-1.43 , 0.58]		+ + ● ? ● ? ?
Total (Wald[e])			**48**			**82**	**100.0%**	**0.10 [-0.29 , 0.49]**		

Test for overall effect: Z = 0.49 (P = 0.62)
Test for subgroup differences: Not applicable
Heterogeneity: Tau² (DL[f]) = 0.01; Chi² = 3.24, df = 3 (P = 0.36); I² = 7%

Favours other activities   Favours music-based therapy

Footnotes
[a]Data for study 2 provided by the study author, and they represent the status four weeks after treatment ended.
[b]The data represent the status four weeks after treatment ended.
[c]Higher scores reflect better quality of life. Intervention group size reduced by 50% because of two control groups
[d]Intervention group size reduced by 50% because of two control groups. Quality of life self-report higher scores: better quality of life
[e]CI calculated by Wald-type method.
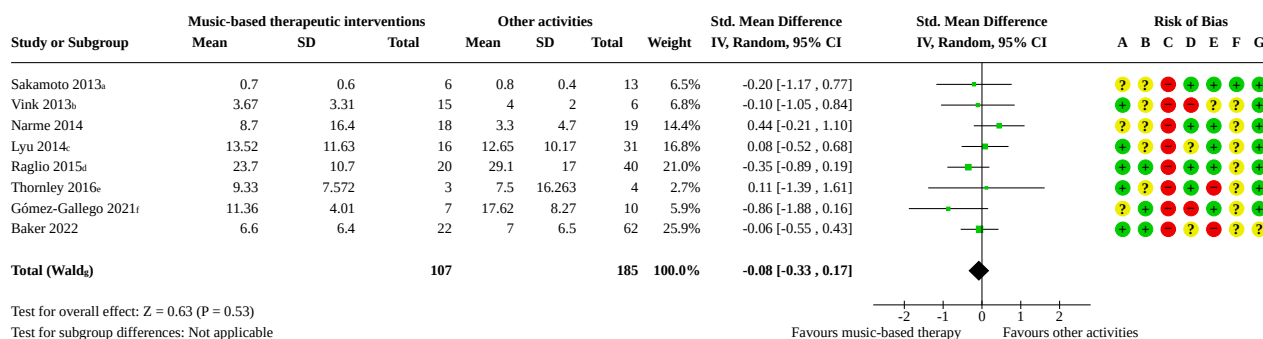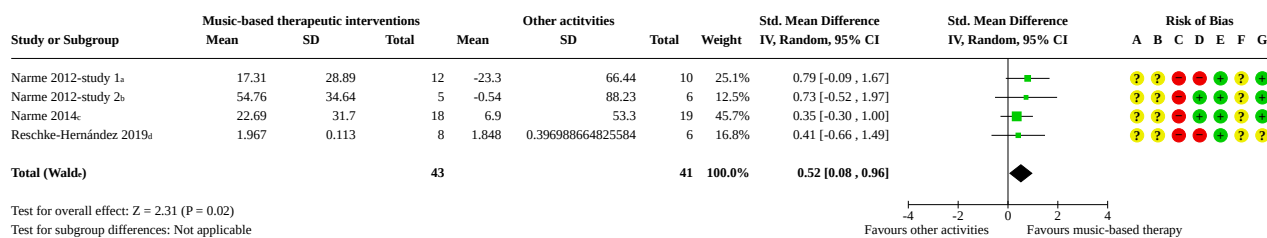[f]Tau² calculated by DerSimonian and Laird method.

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 4.2. Comparison 4: Music-based therapeutic interventions versus other activities: long-term effects, Outcome 2: Mood disturbance or negative affect: depression

Study or Subgroup	Music-based therapeutic interventions			Other activities			Weight	Std. Mean Difference IV, Random, 95% CI	Std. Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
	Mean	SD	Total	Mean	SD	Total				
Guétin 2009[a]	12.5	6.4	13	12.1	7.6	11	16.0%	0.06 [-0.75 , 0.86]		? ? ⊖ ⊕ ? ? ⊕
Narme 2014[b]	0.8	2.9	18	1	3	19	24.8%	-0.07 [-0.71 , 0.58]		? ? ⊖ ⊕ ⊕ ? ⊕
Raglio 2015[c]	8.27	5.449	20	8.62	5.045	40	35.8%	-0.07 [-0.60 , 0.47]		⊕ ⊕ ⊖ ⊕ ⊕ ? ⊕
Baker 2022[d]	9.6	8.3	11	11.1	8.7	42	23.4%	-0.17 [-0.84 , 0.49]		⊕ ⊕ ⊖ ? ⊖ ? ⊕
Total (Wald[e])			**62**			**112**	**100.0%**	**-0.07 [-0.39 , 0.25]**		

Test for overall effect: Z = 0.44 (P = 0.66)
Test for subgroup differences: Not applicable
Heterogeneity: Tau² (DL[f]) = 0.00; Chi² = 0.18, df = 3 (P = 0.98); I² = 0%

Favours music-based therapy   Favours other activities

Footnotes
[a]Geriatric Depression Scale data represent the status four weeks after treatment ended.
[b]Depression subscale of NPI data provided by the study author, and they represent the status four weeks after treatment ended.
[c]Means and SD of the Cornell scale provided by the study author. Intervention group size reduced by 50% because of two control groups
[d]Intervention group size reduced by 50% because of two control groups
[e]CI calculated by Wald-type method.
[f]Tau² calculated by DerSimonian and Laird method.

Risk of bias legend
(A) Random sequence generation (selection bias)
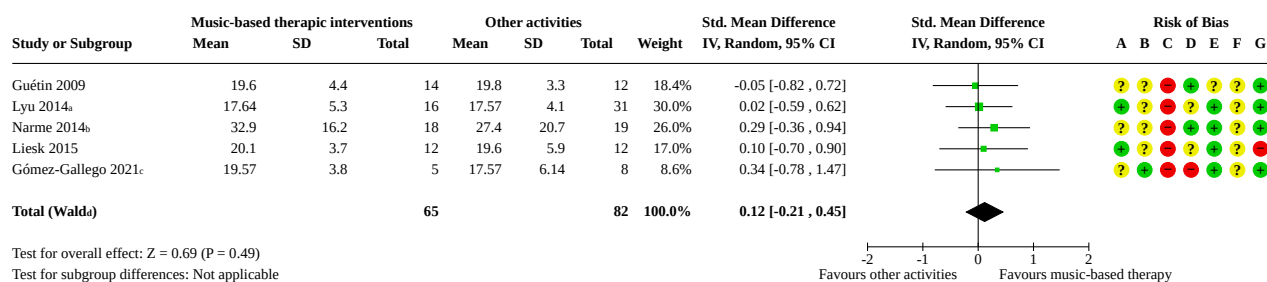(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 4.3. Comparison 4: Music-based therapeutic interventions versus other activities: long-term effects, Outcome 3: Mood disturbance or negative affect: anxiety

Study or Subgroup	Music-based therapeutic interventions			Other activities			Weight	Std. Mean Difference IV, Random, 95% CI	Std. Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
	Mean	SD	Total	Mean	SD	Total				
Guétin 2009[a]	10.6	6.3	13	20.5	5.4	11	23.2%	-1.62 [-2.56 , -0.67]		? ? ⊖ ⊕ ⊕ ? ⊕
Narme 2012-study 2[b]	21.39	29.02	5	34.89	30.4	6	19.1%	-0.41 [-1.62 , 0.79]		? ? ⊖ ⊕ ⊕ ? ⊕
Narme 2014[c]	2.4	4.1	18	1.2	3.2	19	28.3%	0.32 [-0.33 , 0.97]		? ? ⊖ ⊕ ⊕ ? ⊕
Raglio 2015[d]	2.14	2.428	18	4.06	3.651	34	29.4%	-0.58 [-1.16 , 0.01]		⊕ ⊕ ⊖ ⊕ ⊕ ? ⊕
Total (Wald[e])			**54**			**70**	**100.0%**	**-0.53 [-1.31 , 0.25]**		

Test for overall effect: Z = 1.34 (P = 0.18)
Test for subgroup differences: Not applicable
Heterogeneity: Tau² (DL[f]) = 0.45; Chi² = 11.42, df = 3 (P = 0.010); I² = 74%

Favours music-based therapy   Favours other activities

Footnotes
[a]Hamilton anxiety scale. The data represent the status four weeks after treatment ended.
[b]STAI-A data for study 2 provided by the study author, and they represent the status four weeks after treatment ended.
[c]NPI subscale anxiety data provided by the study author, and they represent the status four weeks after treatment ended.
[d]Anxiety subscale score of NPI, data provided by the study author. Intervention group size reduced by 50% because of two control groups
[e]CI calculated by Wald-type method.
[f]Tau² calculated by DerSimonian and Laird method.

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
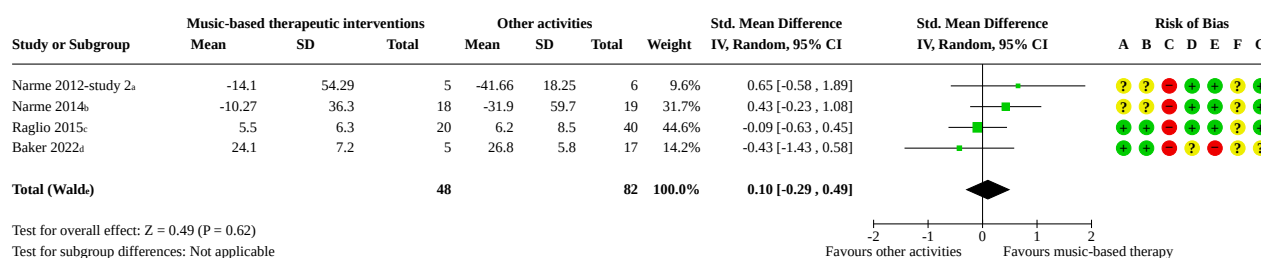(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 4.4.   Comparison 4: Music-based therapeutic interventions versus other activities: long-term effects, Outcome 4: Behavioural problems: agitation or aggression

Study or Subgroup	Music-based therapeutic interventions			Other activities			Weight	Std. Mean Difference IV, Random, 95% CI	Std. Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
	Mean	SD	Total	Mean	SD	Total				
Narme 2014[a]	40.2	15.4	18	34	7.6	19	47.8%	0.50 [-0.15 , 1.16]		? ? ● ● ● ? ●
Raglio 2015[b]	3.11	2.964	18	3.91	2.843	34	52.2%	-0.27 [-0.85 , 0.30]		● ● ● ● ● ? ●
Total (Wald[c])			36			53	100.0%	0.10 [-0.66 , 0.86]		

Test for overall effect: Z = 0.25 (P = 0.80)
Test for subgroup differences: Not applicable
Heterogeneity: Tau² (DL[d]) = 0.20; Chi² = 3.05, df = 1 (P = 0.08); I² = 67%

-4 -2 0 2 4
Favours music-based therapy   Favours other activities

Footnotes
[a]The data represent the status four weeks after treatment ended.
[b]Agitation subscale score of NPI, data provided by the study author. Intervention group size reduced by 50% because of two control groups
[c]CI calculated by Wald-type method.
[d]Tau² calculated by DerSimonian and Laird method.

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
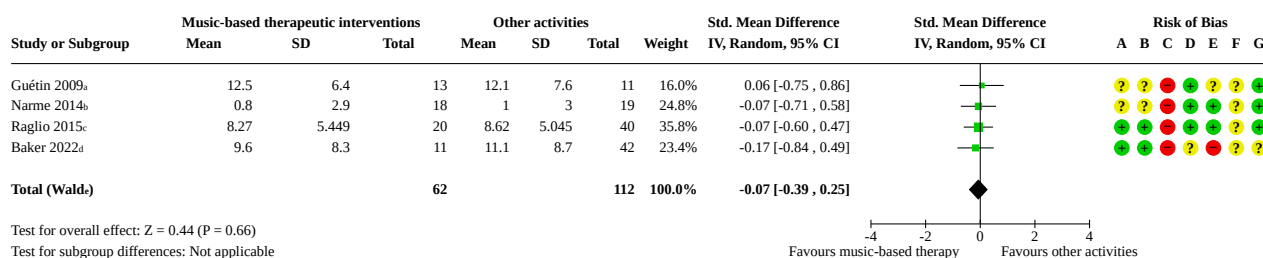(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
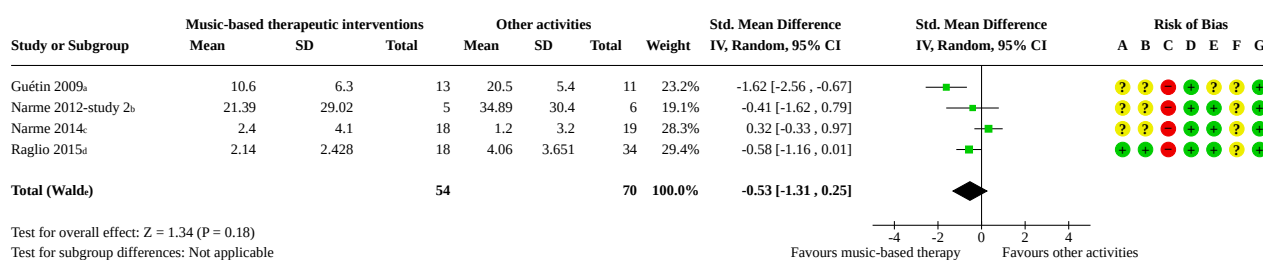(G) Other bias

Analysis 4.5.   Comparison 4: Music-based therapeutic interventions versus other activities: long-term effects, Outcome 5: Behavioural problems: overall

Study or Subgroup	Music-based therapeutic interventions			Other activities			Weight	Std. Mean Difference IV, Random, 95% CI	Std. Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
	Mean	SD	Total	Mean	SD	Total				
Lyu 2014[a]	13.01	11.72	16	12.58	10.03	31	25.4%	0.04 [-0.56 , 0.64]		● ? ● ● ● ? ●
Narme 2014[b]	10.6	12.6	18	8.5	13.5	19	22.1%	0.16 [-0.49 , 0.80]		? ? ● ● ● ? ●
Raglio 2015[a]	22.4	11.9	20	28.4	17.2	40	31.5%	-0.38 [-0.92 , 0.16]		● ● ● ● ● ? ●
Baker 2022[a]	6.6	5.3	11	7	6.5	42	21.0%	-0.06 [-0.73 , 0.60]		● ● ● ? ● ? ?
Total (Wald[c])			65			132	100.0%	-0.09 [-0.39 , 0.22]		

Test for overall effect: Z = 0.56 (P = 0.57)
Test for subgroup differences: Not applicable
Heterogeneity: Tau² (DL[d]) = 0.00; Chi² = 1.84, df = 3 (P = 0.61); I² = 0%

-2 -1 0 1 2
Favours music-based therapy   Favours other activities

Footnotes
[a]Intervention group size reduced by 50% because of two control groups
[b]The data represent the status four weeks after treatment ended.
[c]CI calculated by Wald-type method.
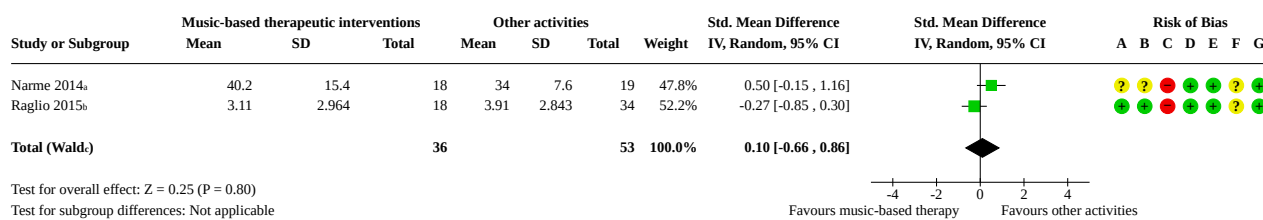[d]Tau² calculated by DerSimonian and Laird method.

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 4.6. Comparison 4: Music-based therapeutic interventions versus other activities: long-term effects, Outcome 6: Social behaviour

Study or Subgroup	Music-based therapeutic interventions			Other activities			Weight	Std. Mean Difference IV, Random, 95% CI	Std. Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
	Mean	SD	Total	Mean	SD	Total				
Narme 2012-study 2[a]	-3.22	29.39	5	-38.5	21.92	6	35.6%	1.26 [-0.09 , 2.62]		? ? ● ● ● ? ●
Narme 2014[b]	4	52.6	18	-2.8	50.1	19	64.4%	0.13 [-0.52 , 0.78]		? ? ● ● ● ? ●
Total (Wald[c])			**23**			**25**	**100.0%**	**0.53 [-0.53 , 1.60]**		

Test for overall effect: Z = 0.98 (P = 0.33)
Test for subgroup differences: Not applicable
Heterogeneity: Tau² (DL[d]) = 0.35; Chi² = 2.19, df = 1 (P = 0.14); I² = 54%

Favours other activities   Favours music-based therapy

Footnotes
[a]Data for study 2 provided by the study author, and they represent the status four weeks after treatment ended.
[b]The data represent the status four weeks after treatment ended.
[c]CI calculated by Wald-type method.
[d]Tau² calculated by DerSimonian and Laird method.
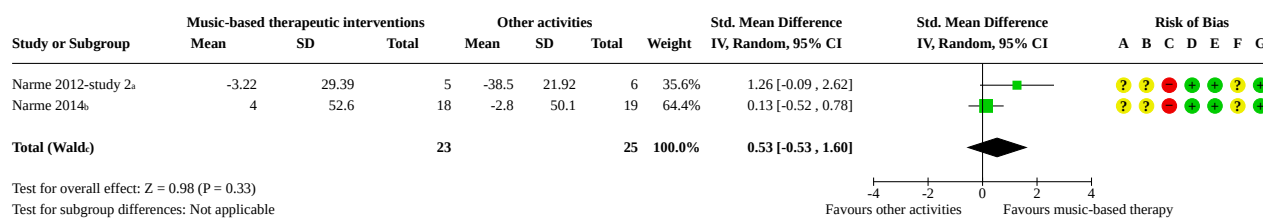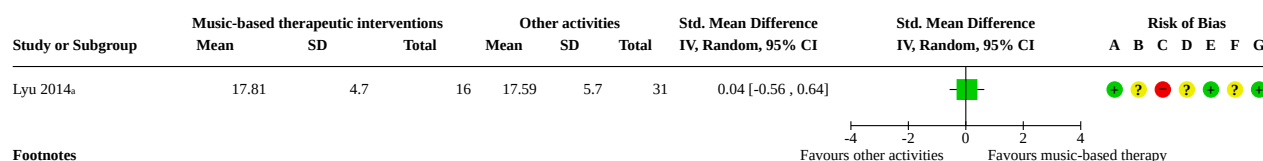
Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 4.7. Comparison 4: Music-based therapeutic interventions versus other activities: long-term effects, Outcome 7: Cognition

Study or Subgroup	Music-based therapeutic interventions			Other activities			Std. Mean Difference IV, Random, 95% CI	Std. Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
	Mean	SD	Total	Mean	SD	Total			
Lyu 2014[a]	17.81	4.7	16	17.59	5.7	31	0.04 [-0.56 , 0.64]		● ? ● ? ● ? ●

Favours other activities   Favours music-based therapy

Footnotes
[a]Intervention group size reduced by 50% because of two control groups

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 5. Subgroup analysis

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
5.1 Music-based therapeutic interventions versus usual care for behaviour problems: agitation or aggression at end of treatment; subgroup analysis: individual therapy versus group therapy	11	503	Std. Mean Difference (IV, Random, 95% CI)	-0.05 [-0.27, 0.17]
5.1.1 Individual therapy	5	195	Std. Mean Difference (IV, Random, 95% CI)	-0.31 [-0.60, -0.02]
5.1.2 Group therapy	6	308	Std. Mean Difference (IV, Random, 95% CI)	0.13 [-0.12, 0.37]

Analysis 5.1.   Comparison 5: Subgroup analysis, Outcome 1: Music-based therapeutic interventions versus usual care for behaviour problems: agitation or aggression at end of treatment; subgroup analysis: individual therapy versus group therapy

Study or Subgroup	Music-based therapeutic interventions			Usual care			Weight	Std. Mean Difference IV, Random, 95% CI	Std. Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
	Mean	SD	Total	Mean	SD	Total				
5.1.1 Individual therapy										
Clark 1998	65.56	58.02	18	121.56	119.23	18	7.9%	-0.58 [-1.25 , 0.08]		
Raglio 2010a[a]	1.41	1.907	27	2.38	3.386	24	10.3%	-0.35 [-0.91 , 0.20]		
Ridder 2013[b]	26.09	13.54	17	28	18.15	18	8.0%	-0.12 [-0.78 , 0.55]		
Sakamoto 2013[c]	0.7	1	7	3.2	3	13	4.4%	-0.95 [-1.93 , 0.02]		
Raglio 2015[d]	3.78	3.053	18	3.77	3.011	35	10.0%	0.00 [-0.57 , 0.57]		
Subtotal (Wald[e])			**87**			**108**	**40.6%**	**-0.31 [-0.60 , -0.02]**		
Test for overall effect: Z = 2.11 (P = 0.04)										
Heterogeneity: Tau² (DL[f]) = 0.00; Chi² = 3.82, df = 4 (P = 0.43); I² = 0%										
5.1.2 Group therapy										
Svansdottir 2006	1.2	1.7	20	1.3	1.6	18	8.5%	-0.06 [-0.70 , 0.58]		
Raglio 2010b	2.5	4.2	10	1.6	2.1	10	5.2%	0.26 [-0.62 , 1.14]		
Lin 2011	36.37	10.64	49	38.55	10.27	51	15.2%	-0.21 [-0.60 , 0.19]		
Ceccato 2012	25.63	15.88	27	22.8	12.73	23	10.2%	0.19 [-0.37 , 0.75]		
Sung 2012	32.7	4.98	27	31	2.96	28	10.8%	0.41 [-0.12 , 0.95]		
Giovagnoli 2018	1.41	2.82	23	0.38	0.97	22	9.4%	0.48 [-0.12 , 1.07]		
Subtotal (Wald[e])			**156**			**152**	**59.4%**	**0.13 [-0.12 , 0.37]**		
Test for overall effect: Z = 1.02 (P = 0.31)										
Heterogeneity: Tau² (DL[f]) = 0.01; Chi² = 5.63, df = 5 (P = 0.34); I² = 11%										
Total (Wald[e])			**243**			**260**	**100.0%**	**-0.05 [-0.27 , 0.17]**		

Test for overall effect: Z = 0.42 (P = 0.67)
Test for subgroup differences: Chi² = 5.14, df = 1 (P = 0.02), I² = 80.6%
Heterogeneity: Tau² (DL[f]) = 0.04; Chi² = 14.62, df = 10 (P = 0.15); I² = 32%

-2 -1 0 1 2
Favours music-based therapy   Favours usual care

Footnotes
[a]Agitation subscale score of NPI, data about control group provided by the author
[b]Adapted CMAI with different range; note that an effect size is reported but based on SD baseline
[c]Aggressiveness subscale of the BEHAVE-AD. Intervention group size reduced by 50% because of two control groups
[d]Agitation subscale score of NPI, data provided by the author. Intervention group size reduced by 50% because of two control groups
[e]CI calculated by Wald-type method.
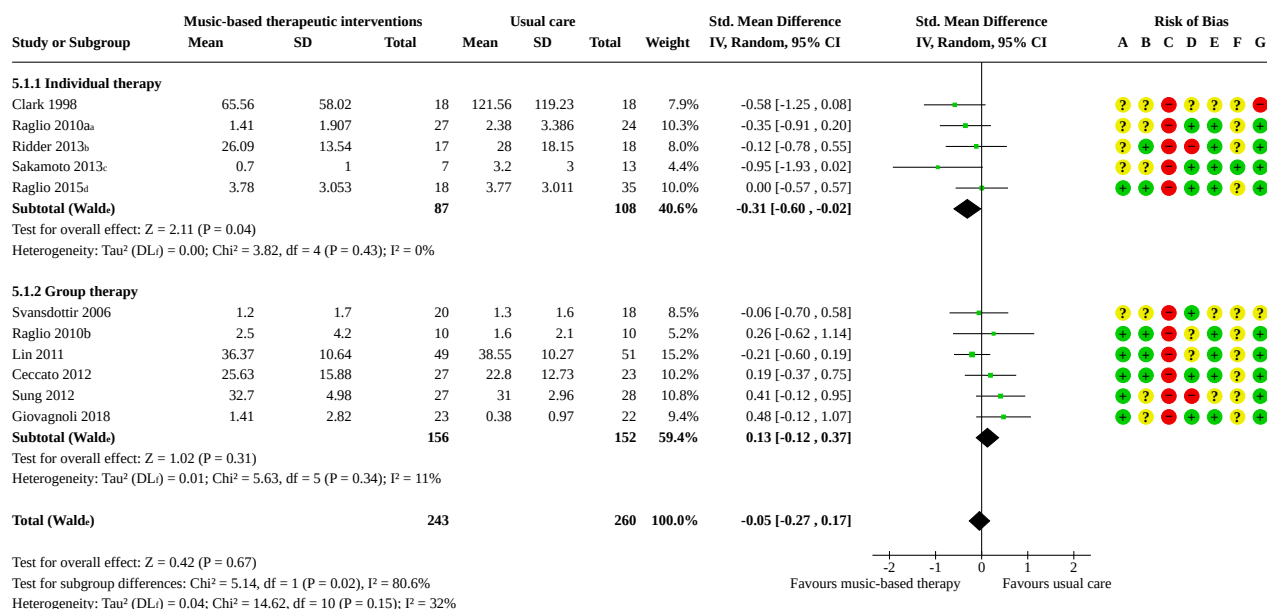[f]Tau² calculated by DerSimonian and Laird method.

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

A P P E N D I C E S

Appendix 1. Sources searched and search strategies used

Source searched	Search strategy	Hits
MEDLINE In-process and other non-indexed citations and MEDLINE 1950 to present	1. exp Dementia/	Apr 2010: 15
	2. Delirium/	Oct 2014: 59
	3. Wernicke Encephalopathy/	Jul 2015: 15
Most recent search performed: 30 November 2023	4. Delirium, Dementia, Amnestic, Cognitive Disorders/	Apr 2016: 36
	5. dement*.mp.	Jun 2017: 47
	6. alzheimer*.mp.	Dec 2021: 175

(Continued)

	7. (lewy* adj2 bod*).mp.	Dec 2023 (up to 30 November 2023): 97
	8. deliri*.mp.	
	9. (chronic adj2 cerebrovascular).mp.	
	10. ("organic brain disease" or "organic brain syndrome").mp.	
	11. ("normal pressure hydrocephalus" and "shunt*").mp.	
	12. "benign senescent forgetfulness".mp.	
	13. (cerebr* adj2 deteriorat*).mp.	
	14. (cerebral* adj2 insufficient*).mp.	
	15. (pick* adj2 disease).mp.	
	16. (creutzfeldt or jcd or cjd).mp.	
	17. huntington*.mp.	
	18. binswanger*.mp.	
	19. korsako*.mp.	
	20. or/1-19	
	21. music*.mp.	
	22. exp Music Therapy/	
	23. singing.mp.	
	24. sing.mp.	
	25. "auditory stimul*".mp.	
	26. piano.mp.	
	27. or/21-26	
	28. 27 and 20	
	29. randomized controlled trial.pt.	
	30. controlled clinical trial.pt.	
	31. random*.ab.	
	32. placebo.ab.	
	33. trial.ab.	
	34. groups.ab.	
	35. or/29-34	
	36. (animals not (humans and animals)).sh.	
	37. 35 not 36	
	38. 28 and 37	
Embase	1. exp dementia/	Apr 2010: 28
1980 to present	2. Lewy body/	Oct 2014: 230

(Continued)

Most recent search performed: 30 November 2023	3. delirium/	Jul 2015: 42
	4. Wernicke encephalopathy/	Apr 2016: 106
	5. cognitive defect/	Jun 2017: 101
	6. dement*.mp.	Dec 2021: 486
	7. alzheimer*.mp.	Dec 2023: 260
	8. (lewy* adj2 bod*).mp.	
	9. deliri*.mp.	
	10. (chronic adj2 cerebrovascular).mp.	
	11. ("organic brain disease" or "organic brain syndrome").mp.	
	12. "supranuclear palsy".mp.	
	13. ("normal pressure hydrocephalus" and "shunt*").mp.	
	14. "benign senescent forgetfulness".mp.	
	15. (cerebr* adj2 deteriorat*).mp.	
	16. (cerebral* adj2 insufficient*).mp.	
	17. (pick* adj2 disease).mp.	
	18. (creutzfeldt or jcd or cjd).mp.	
	19. huntington*.mp.	
	20. binswanger*.mp.	
	21. korsako*.mp.	
	22. CADASIL.mp.	
	23. or/1-22	
	24. music*.mp.	
	25. exp music therapy/	
	26. singing.mp.	
	27. sing.mp.	
	28. exp singing/	
	29. "auditory stimul*".mp.	
	30. exp auditory stimulation/	
	31. piano.mp.	
	32. or/24-31	
	33. 23 and 32	
	34. randomized controlled trial/	
	35. exp controlled clinical trial/	
	36. random*.ab.	

(Continued)

	37. placebo.ab.	
	38. trial.ab.	
	39. groups.ab.	
	40. or/34-39	
	41. 33 and 40	
PsycINFO	1. exp Dementia/	Apr 2010: 26
1806 to present	2. exp Delirium/	Oct 2014: 100
Most recent search performed: 30 November 2023	3. exp Huntingtons Disease/	Jul 2015: 14
	4. exp Kluver Bucy Syndrome/	Apr 2016: 34
	5. exp Wernickes Syndrome/	Jun 2017: 35
	6. exp Cognitive Impairment/	Dec 2021: 166
	7. dement*.mp.	Dec 2023: 76
	8. alzheimer*.mp.	
	9. (lewy* adj2 bod*).mp.	
	10. deliri*.mp.	
	11. (chronic adj2 cerebrovascular).mp.	
	12. ("organic brain disease" or "organic brain syndrome").mp.	
	13. "supranuclear palsy".mp.	
	14. ("normal pressure hydrocephalus" and "shunt*").mp.	
	15. "benign senescent forgetfulness".mp.	
	16. (cerebr* adj2 deteriorat*).mp.	
	17. (cerebral* adj2 insufficient*).mp.	
	18. (pick* adj2 disease).mp.	
	19. (creutzfeldt or jcd or cjd).mp.	
	20. huntington*.mp.	
	21. binswanger*.mp.	
	22. korsako*.mp.	
	23. ("parkinson* disease dementia" or PDD or "parkinson* dementia").mp.	
	24. or/1-23	
	25. music*.mp.	
	26. exp Music Therapy/	
	27. sing.mp.	
	28. singing.mp.	
	29. exp Singing/	

(Continued)

30. "auditory stimul*".mp.

31. *Auditory Stimulation/

32. piano.mp.

33. or/25-32

34. 24 and 33

35. exp Clinical Trials/

36. random*.ti,ab.

37. trial.ti,ab.

38. group.ab.

39. placebo.ab.

40. or/35-39

41. 34 and 40

CINAHL via EBSCO	S42 S28 AND S41	Apr 2010: 18
Most recent search performed: 30 November 2023	S41 S29 OR S30 OR S31 OR S32 OR S33 OR S34 OR S35 OR S36 OR S37 OR S38 OR S39 OR S40	Oct 2014: 53
	S40 MH "Random Assignment"	Jul 2015: 8
	S39 MH "Single-Blind Studies" or MH "Double-Blind Studies" or MH "Triple-Blind Studies"	Apr 2016: 12
	S38 MH "Crossover Design"	Jun 2017: 20
	S37 MH "Factorial Design"	Dec 2021: 110
	S36 MH "Placebos"	Dec 2023: 45
	S35 MH "Clinical Trials"	
	S34 TX "multi-centre study" OR "multi-center study" OR "multicentre study" OR "multicenter study" OR "multi-site study"	
	S33 TX crossover OR "cross-over"	
	S32 AB placebo*	
	S31 TX random*	
	S30 TX trial*	
	S29 TX "latin square"	
	S28 S19 AND S27	
	S27 S20 OR S21 OR S22 OR S23 OR S24 OR S25 OR S26	
	S26 MM "Acoustic Stimulation"	
	S25 TX "auditory stimul*"	
	S24 MM "Singing"	
	S23 TX singing	

(Continued)

S22 TX sing

S21 (MH "Music Therapy") or (MH "Music Therapy (Iowa NIC)")

S20 TX music*

S19 S1 OR S2 OR S3 OR S4 OR S5 OR S6 OR S7 OR S8 OR S9 OR S10 OR S11 OR S12 OR S13 OR S14 OR S15 OR S16 OR S17 OR S18

S18 TX korsako*

S17 TX binswanger*

S16 TX huntington*

S15 TX creutzfeldt or jcd or cjd

S14 TX pick* N2 disease

S13 TX cerebral* N2 insufficient*

S12 TX cerebr* N2 deteriorat*

S11 TX "benign senescent forgetfulness"

S10 TX "normal pressure hydrocephalus" and "shunt*"

S9 TX "organic brain disease" or "organic brain syndrome"

S8 TX chronic N2 cerebrovascular

S7 TX deliri*

S6 TX lewy* N2 bod*

S5 TX alzheimer*

S4 TX dement*

S3 MH "Wernicke's Encephalopathy"

S2 (MH "Delirium") or (MH "Delirium, Dementia, Amnestic, Cognitive Disorders")

S1 MH "Dementia+"

Web of Science via Clarivate (1945 to present) Most recent search performed: 30 November 2023	Topic=(music* OR singing OR sing OR "auditory stimul*") AND Topic=(dement* OR alzheimer* OR "lew* bod*" OR huntington*) AND Topic=(random* OR trial OR placebo OR "double blind*" OR "single blind*" OR groups)	Apr 2010: 33
		Oct 2014: 205
		Jul 2015: 20
		Apr 2016: 76
		Jun 2017: 45
		Dec 2021: 530
		Dec 2023: 643
LILACS Most recent search performed: 30 November 2023	demen$ [Words] and music OR singing [Words]	Apr 2010: 7
		Oct 2014: 12
		Jul 2015: 0
		Apr 2016: 0

(Continued)

		Jun 2017: 0
		Dec 2021: 16
		Dec 2023: 5
CDCIG Register (CRSWEB) Most recent search performed: 30 November 2023	music OR singing OR auditory	Apr 2010: 29
		Oct 2014: 18
		Jul 2015: 0
		Apr 2016: 6
		Jun 2017: 0
		Dec 2021: 403
		Dec 2023: 163
UMIN (Clinical Trial Register of Japan) Most recent search performed:April 2017	Free Keyword: music OR singing OR auditory	Apr 2010: 0
		Oct 2014: 0
		Jul 2015: 0
		Apr 2016: 0
		Jun 2017: 0
CENTRAL via CRSO Most recent search performed: 30 November 2023	#1 MeSH descriptor Dementia explode all trees	Apr 2010: 10
	#2 MeSH descriptor Delirium, this term only	Oct 2014: 53
	#3 MeSH descriptor Wernicke Encephalopathy, this term only	Jul 2015: 11
	#4 MeSH descriptor Delirium, Dementia, Amnestic, Cognitive Disorders, this term only	Apr 2016: 9
	#5 dement*	Jun 2017: 38
	#6 alzheimer*	Dec 2021: 363
	#7 "lewy* bod*"	Dec 2023: 107
	#8 deliri*	
	#9 "chronic cerebrovascular"	
	#10 "organic brain disease" or "organic brain syndrome"	
	#11 "normal pressure hydrocephalus" and "shunt*"	
	#12 "benign senescent forgetfulness"	
	#13 "cerebr* deteriorat*"	
	#14 "cerebral* insufficient*"	
	#15 "pick* disease"	
	#16 creutzfeldt or jcd or cjd	
	#17 huntington*	
	#18 binswanger*	

(Continued)

	#19 korsako*	
	#20 (#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19)	
	#21 MeSH descriptor Music Therapy explode all trees	
	#22 music*	
	#23 singing	
	#24 sing	
	#25 "auditory stimul*"	
	#26 (#21 OR #22 OR #23 OR #24 OR #25)	
	#27 (#20 AND #26)	
ClincalTrials.gov Most recent search performed: 30 November 2023	dementia OR alzheimer OR alzheimers OR alzheimer's | music OR sing OR singing OR auditory	Apr 2010: 2
		Oct 2014: 14
		Jul 2015: 0
		Apr 2016: 0
		Jun 2017: 0
		Dec 2021: 70
		Dec 2023: 37
ICTRP Search Platform (WHO portal) Most recent search performed: 30 November 2023	Advanced search: [condition: Dementia OR alzheimer OR alzheimers] AND [Intervention: music OR singing OR sing OR auditory]	Apr 2010: 20
		Oct 2014: 18
		Jul 2015: 0
		Apr 2016: 3
		Jun 2017: 0
		Dec 2021: 64
		Dec 2023: 5
TOTAL		Apr 2010: 188
		Oct 2014: 761
		Jul 2015: 110
		Apr 2016: 282
		Jun 2017: 286
		Dec 2021: 2383
		Dec 2023: 1438
		TOTAL: 5448
Total after deduplication:		Dec 2021: 1387
		Dec 2023: 986

Appendix 2. History of search results and studies included in each version of the review

First citation (Vink 2002)

The protocol for a new review entitled 'Music therapy in the care of people with dementia' was published in January 2002.

Second citation (Vink 2003)

In the first version of the review, with an abbreviated title of 'Music therapy for people with dementia', five studies met the original criteria for inclusion (Brotons 2000; Clark 1998; Gerdner 2000; Groene 1993; Lord 1993). One ongoing study was listed.

A search in 2006 did not find new trials.

In 2008, an additional three studies met the inclusion criteria (Raglio 2008; Sung 2006; Svansdottir 2006).

In 2010, two more studies met the inclusion criteria (Guétin 2009; Raglio 2010a), so a total of 10 studies were included in that update. One study was listed as 'awaiting classification'.

These updates did not receive new citations but automatically superseded the existing version; the version available on the Cochrane Library as Vink 2003 is the one based on the 14 April 2010 search.

Third citation (Van der Steen 2017)

This update was written with three additional review authors. The title was changed to 'Music-based therapeutic interventions for people with dementia'. Inclusion of studies in the previous version were reconsidered according to the new and more stringent criteria (see Differences between protocol and review). Due to clarified criteria for the eligibility of interventions, randomisation and more stringent application of criteria for analysis of outcomes after a minimum number of sessions, we excluded five of the 10 previously included studies (Brotons 2000; Gerdner 2000; Groene 1993; Raglio 2008; Sung 2006).

We conducted a search in July 2015 that identified 12 new trials for inclusion (Ceccato 2012; Cooke 2010; Liesk 2015; Lin 2011; Narme 2012-study 1; Narme 2012-study 2; Narme 2014; Raglio 2010b; Ridder 2013; Sung 2012; Sakamoto 2013; Vink 2013).

A search on 12 April 2016 identified several studies that were potentially eligible for inclusion, which we did not assess at that time but categorised as 'awaiting classification'.

Therefore, in the 2017 version, we had a total of 17 included studies. In addition, one study was ongoing and 19 were awaiting classification.

Fourth citation (Van der Steen 2018)

This update was written with another new review author who worked on data collection and analysis with the first review author.

Of the studies we had placed in awaiting classification from the 2016 search, we included four (from which we could extract data with the help of collaborators): Hsu 2015; Lyu 2014; Raglio 2015; Thornley 2016.

The next search was performed on 19 June 2017. We identified a new eligible study (Cho 2016). Therefore, we had 22 included studies. There were three ongoing studies and 19 potentially eligible studies awaiting classification.

Fifth citation (Van der Steen 2025)

This update was written with another new review author who worked on data collection.

The next search, on 2 December 2021, identified 39 articles on 32 potentially eligible studies. Of these, we included seven (Baker 2022; Giovagnoli 2018; Gómez-Gallego 2021; Liu 2021; Reschke-Hernández 2019; Tang 2018; Werner 2017), one of which had five articles (Baker 2022). All seven studies contributed data to meta-analysis. We moved the study that was still ongoing in the previous update, NCT02147652, to 'awaiting classification'. The database search identified two other ongoing studies (Baroni Caramel 2024; NCT04666077), and we also presented Gold 2018 as an ongoing study, even though we had already included some of the data from this study, Baker 2022, in this version of the review. From the database search, we categorised three more studies as awaiting classification (NCT00448318; ISRCTN11001662; Prick 2024), while handsearching identified one study, which we also categorised as awaiting classification (Campbell 2022).

Our most recent search was conducted on 30 November 2023. We included Prieto Alvarez 2022, but the study did not contribute to meta-analysis. We added Moreira 2023 to 'Studies awaiting classification' as potentially eligible.

Therefore, in the 2025 update of the review, we have included a total of 30 studies. There are three ongoing studies and 19 studies await classification.

Appendix 3. Description of the interventions

Baker 2022

Music-based therapeutic intervention: group mixed music therapy (versus a control condition of recreational choir singing and a control condition of usual care)

Experimental group

Forty-five-minute, active music therapy sessions, optionally including receptive music therapy elements, delivered by credentialed music therapists to nursing home residents. The sessions were informed by the 12 principles of person-centred care of Kitwood 1997.

In preparation for the sessions, the music therapist offered each participant an initial 20-minute individual assessment to start to build individual rapport and determine their musical preferences. The music therapist also used other sources to determine the participants' musical biography, cultural background, history, personal strengths, resources and disabilities relevant for the sessions. The sessions included a welcome song, the singing of songs that are meaningful to participants, music-evoked reminiscence, improvising on small hand-held percussion instruments, and gentle spontaneous or directed movement to music.

The sessions were delivered in closed groups of 8-10 participants who were often from the same care unit, and, if possible, from the same cultural background or with the same musical preferences. Sessions were delivered twice a week for 3 months, followed by a reduced frequency of once a week for 3 months, resulting in 39 sessions.

Control group 1

Forty-five-minute recreational choir singing sessions were delivered by a community musician with ensemble leading experience in an open group of 15-20 residents. The sessions were informed by the 12 principles of person-centred care of Kitwood 1997. Biographically and culturally grounded song materials were used to stimulate positive experiences shared by the groups.

The sessions were structured around the singing of familiar and preferred songs with lyrics displayed on a screen. Based on the needs and abilities of the group, brief physical and vocal warm-ups and learning new songs were incorporated into the session.

Sessions were delivered twice a week for 3 months, followed by a reduced frequency of once a week for 3 months, resulting in 39 sessions.

Control group 2

The residents received care as usual, meaning they took part in leisure programmes provided by the nursing home facility as usual (group games, attending concerts or movies, arts and crafts classes, exercise classes, gardening, and outings).

Ceccato 2012

Music-based therapeutic intervention: sound training for attention and memory in dementia (STAM-Dem) (versus a control group of usual care)

Experimental group

A 45-minute mixed (active and receptive) group intervention delivered by "professionally trained music therapists trained to administer the STAM-Dem protocol". Highly structured, progressive series music sessions, with a minimum of four and a maximum of five participants per group. The music therapists were instructed to "pay attention to the relational atmosphere" and "maintain the level of motivation as high as possible."

The intervention included "step-by-step exercises aimed at stimulating and checking both attention and memory". Participants were asked to perform specific movements, count, clap hands, alternate clapping hands and tapping the table, repeat sequences of previously recorded sounds (not stated how) after listening to recorded and live music. It was a mixed intervention because the active component was combined with listening to music.

The STAM-Dem protocol comprised four phases, one for each specific cognitive function that was trained (selective attention, sustained attention, alternate attention and working memory). The phases involved: 1. stimulus-movement association, 2. reaction to acoustic stimuli, 3. shifting attention with two exercises, and 4. orderly and inverted repetition. It is not clear from the text if the phases each lasted four sessions, and were progressive, but is described in other sources (not cited in the article) (STAM protocol). Each phase then lasted four sessions and was followed by the next. However, the intervention phase lasted 12 weeks, in which 24 sessions were held.

Control group

Usual care

Cho 2016

Music-based therapeutic intervention: active group singing (versus two control groups, music listening and television)

Experimental group

A 40-minute active group music therapy, which consisted of singing songs that reflected participants' preferences with regard to music genres, songs and musicians. Eight lists of songs for the music therapy-singing group were developed, centred around a different theme for each session (country, rat pack, the moon, World War II, Broadway, 1950s and 1960s, autumn and patriotic). A board-certified music therapist with 15 years of experience in dementia care delivered the intervention in a separate room. The sessions were delivered twice a week for four weeks.

Control group 1

A 40-minute music listening session in which participants listened to a CD that contained almost the same songs and order of the songs sung in the music therapy singing group (but, the latter sessions, for example, always concluded with "Show me the way to go home", which was not on the CD). The nursing home activity assistants who delivered this intervention were instructed to lead the group in the same manner as other activities and to validate and process the participants' responses.

Control group 2

A 40-minute session in which participants watched a DVD of a comedy program ("I Love Lucy"). The intervention was facilitated by nursing home activity assistants who validated any spontaneous responses.

Clark 1998

Music-based therapeutic intervention: preferred, recorded music during bathing episodes with aggressive behaviour (versus a control group with no music during bathing)

Experimental group

A receptive, individual intervention with music, listening through speakers, delivered by nursing staff. Duration followed established nursing routines and varied from 11 to 18 minutes.

Preferred music was recorded and selections played via an audiotape recorder during the bathing episode. Background information on participants' music experiences and preferences was obtained from interviews with the family member or responsible agent. "Bathing times were scheduled for either morning or afternoon" "following established nursing routines". Participants received either a partial bath, which was given in the participant's room, or a full bath, which was given in the shower in the nursing unit.

Nursing staff delivered the bathing session. It was not clear from the text whether nursing staff were responsible for turning on the music, but it is highly probable that this was done by the observer: "Initially, consideration was given to having nursing staff be responsible for turning on the audiotape recorder...However, during pilot testing of the procedures, this proved too cumbersome for already overburdened nursing staff". The sessions were given 10 times over two weeks.

Control group

No music during bathing

Cooke 2010

Music-based therapeutic intervention: active group music sessions with live and recorded music (versus a reading group as the control condition)

Experimental group

An active, structured 40-minute group music session delivered by two musicians. The session consisted of singing and playing on instruments accompanied by live familiar songs and recorded instrumental music. The group had a maximum of 16 participants.

The session covered 30 minutes of musician-led, familiar song-singing with guitar accompaniment, and 10 minutes of prerecorded instrumental music. A set repertoire was established for each of three sessions and this was repeated for eight weeks.

"Residents were encouraged to participate actively through singing/humming, playing instruments and… movement". Choice of the instruments was not described. The repertoire selection was based primarily on participants' musical preferences, musicians' repertoire knowledge and the findings from a practice session (conducted in an alternative aged care setting). The 10 minutes of listening to prerecorded music allowed the musicians and participants to have a short rest from performance and singing and to cater for participants who had a preference for more instrumental music. The sessions were delivered three mornings a week (Monday, Wednesday and Friday) for eight weeks, with a total of 24 sessions.

Control group

An interactive reading session included a range of reading and social activities, such as reading local news stories, short stories, telling jokes and undertaking quiz activities. The sessions were led by one trained research assistant. A maximum number of attendees was not clear from the text. The control sessions took 40 minutes, and were delivered three times a week (Monday, Wednesday and Friday) for eight weeks, totalling 24 sessions.

Giovagnoli 2018

Music-based therapeutic intervention: active group music therapy with memantine treatment (versus a memantine-only control group)

Experimental group

A 40-minute, active, group music therapy session, delivered by a music therapist, with a non-verbal approach was offered to persons with Alzheimer's disease and impaired language who were also treated with memantine 20 mg/day added to treatment as usual with cholinesterase inhibitors.

During the sessions, the participants engaged in free sound-music interactions making use of rhythmic and melodic instruments, including xylophones, glockenspiels, triangles, wind chimes, maracas, small woods, guiros, and ethnic percussions. The sound-music interaction involved cognition and emotions and stimulated interpersonal adaptation. Each session began with musical improvisation inviting participants to choose an instrument and to play using a free technique.

The sessions were delivered twice a week for 24 weeks, totalling 48 sessions.

Control group

Memantine 20 mg/day was added to treatment as usual with cholinesterase inhibitors.

Gómez-Gallego 2021

Music-based therapeutic intervention: active group music therapy (versus a control condition of group music listening and a control condition of group watching nature videos)

Experimental group 1

A music therapist delivered 45-minute active group sessions to nursing home residents. In preparation for the session, a questionnaire of musical preferences (referred to Mercadal-Brotons 2018) was administered to participants by the music therapist. Three lists of 12 songs each were created on the basis of these preferences. The list of songs was revised monthly.

A typical session comprised: (1) a welcome song with residents greeting each other; (2) rhythmic exercises with well-known songs and residents clapping their hands; (3) dance exercises, the therapist encouraging the residents to make free body movements in response to music; (4) a music quiz with teams guessing the names of songs and singer and singing the chorus of the song; (5) a goodbye song. When necessary, the therapist adapted their interventions to the needs of residents to avoid adverse reactions.

The group size was six to nine participants and sessions were delivered in a room spacious enough to be comfortable, "twice a week for 3 months (12 sessions in total)". Note that the trial registry mentions a duration of three months as well, but the frequency and total number of sessions is unclear.

Control group 1 (group music listening)

A music therapist delivered a 45-minute group music listening intervention based on an audio-recording of the same songs as used in the experimental group. "The same lists of songs were used for both the active therapy and the music listening interventions with the exceptions of the welcome and goodbye songs, which were replaced by two different songs". At the beginning and at the end of each session, the therapist briefly assessed the needs of the group and selected a recorded song accordingly. The songs were listened to as a group, seated in comfortable chairs in a medium-sized room. After each song, the therapist allowed residents to share their feelings or memories.

Control group 2 (group watching nature videos)

The participants watched nature videos without music in a large-sized room. The activity was monitored by a music therapist and two nurses. The videos had nearly the same duration as the music interventions.

Guétin 2009

Cochrane
Library

Trusted evidence.
Informed decisions.
Better health.

Cochrane Database of Systematic Reviews

Music-based therapeutic intervention: individual receptive therapy with the 'U' sequence method (versus a reading group as the control condition)

Experimental group

An individual, receptive music therapy method, the 'U-sequence' method, involved listening to music sequences, selected from a limited number of musical styles delivered through headphones, in the participant's room. The musical style was chosen based on participants' personal tastes following an interview or questionnaire. From the suggested different musical styles, a musical sequence was selected. This usual musical sequence, lasting 20 minutes, was broken down into several phases, according to the 'U sequence' method and making use of a computer program especially designed for this method. Musical rhythm, orchestral formation, frequency and volume were reduced. After a phase of sustained reduced musical rhythm, orchestral formation, frequency and volume, a re-enlivening phase followed in which musical rhythm, orchestral formation, frequency and volume increased again, and ended at a moderate level in comparison to the beginning phase. The style of music varied from one session to another for a given patient.

"Patients were either in a supine position or seated in a comfortable armchair and were offered a mask so as to avoid visual stimuli". Details on the 'U sequence' method are retrievable through this external link (not included in the paper): www.music-care.com/en/page/ treatment.

Sessions were extended by a period of time spent listening to the participant. This period of time served "to create a 'psychotherapist'-type of therapeutic relationship and …reinforced the effect triggered by listening to music". Duration of this 'listening' intervention with a therapist was not reported.

Personnel delivering the music and the listening intervention was not clear from the text. Sessions were delivered once a week, lasted 20 minutes (plus time spent listening to patients' responses – duration of which is not stated), and 16 sessions were delivered.

Control group

"Rest and reading under the same conditions and at the same intervals"

Hsu 2015

Music-based therapeutic intervention: active individual music therapy for people with dementia and their carers (versus a control group of usual care)

Experimental group

A 30-minute, individual, active music therapy, which consisted of singing well-known songs, instrumental improvisation, talking to allow reminiscence and expression of feelings, and use of facial and bodily expressions of the music therapists combined with a weekly 15-minute video presentation, to direct care staff as an ongoing training tool focused on improving staff knowledge of their patients and confidence and skills to interact.

A music therapist delivered the intervention in a separate, quiet room on the unit. The two qualified music therapists had at least two years' experience working in this setting and were registered with the Health and Care Professions Council (HCPC). To provide consistency and to maintain the therapeutic relationship, residents received all sessions from the same music therapist. The sessions were delivered once a week for five months, in addition to standard care.

Control group

Received standard care for five months. This consisted of medical and personal care, provision of basic needs and activities carried out as usual within the home such as chaplaincy services, entertainment and leisure activities.

Liesk 2015

Music-based therapeutic intervention: a 'Musikgeragogik' group music programme (versus a cognitive stimulation intervention as the control condition)

Experimental group

A 90-minute, structured, active group music intervention based on the principles of 'Musikgeragogik' by T Hartogh (2005), which was designated as "music education for elders". Sessions consisted of singing folk songs, rounds and playing on instruments (woodblocks, bells, tambourine and maracas). Participants were stimulated to improvise in a structured way according to cues in the song lyrics, alternated with spontaneous expression of individual impressions provoked by the songs that were played or sung. It is probable that the music used was live, as the music intervention was "created as an active therapy form", but this was not explicitly mentioned in the text.

A music recreational therapist ('Musikgeragogin') delivered the intervention. Duration of sessions was 90 minutes and frequency was twice a week, during six weeks, totalling 12 sessions.

Control group

A cognitive stimulation programme in which cognitive function is trained through quiz questions of differing complexity and theme-focused conversations, a cognitive training programme of NEUROvitalis from a group in Cologne, adapted for people with dementia. A gerontologist delivered the intervention. The sessions lasted 90 minutes, twice a week over six weeks, totalling 12 sessions.

Lin 2011

Music-based therapeutic intervention: group music therapy (versus a control group of usual care that "continued to perform their usual daily activities")

Experimental group

This was a 30-minute, structured, mixed group music therapy intervention, based on the protocol developed by Clair 1990. The size of the group is not clear from the text.

The intervention consisted of rhythmic music and slow-tempo instrumental activities (choice of instruments not specified), therapeutic singing, listening to specially selected music, glockenspiel playing and musical activities and traditional holiday and 'music creator' activities. "…before the therapy sessions a subject's fondness for music was evaluated through an interview, and the musical activities in the group sessions were arranged according to the interview findings."

The person delivering the intervention was a researcher schooled in two university music therapy courses. The sessions lasted 30 minutes and were conducted twice a week for six consecutive weeks. The total number of sessions was 12.

Control group

Participants received usual care and "continued to perform their usual daily activities."

Liu 2021

Music-based therapeutic intervention: active group music intervention (versus a control condition of reading and relaxing)

Experimental group

A 60-minute, active, group music therapeutic intervention making use of percussion instruments with movement delivered by a music facilitator trained in providing the music intervention.

The music facilitator aimed to work in the "here and now" by responding to participants' immediate affective feelings and incorporating them into musical expressions. Interpersonal connections between the music facilitator and the participants, and between participants were facilitated through reciprocal music-making.

The sessions included 10 minutes of warm-up (moving fingers and upper-extremities and breathing exercises with music) and 10 minutes of cooling down (return to earlier rhythmic patterns used during the session). During the main part of the session, participants were guided with touch by the facilitators in the use of percussion instruments (including wrist bell, tambourine, maracas, triangle, double-tone woodblock, hand drum, castanet) while listening to music and songs familiar to the participants. Participants' music preferences had been assessed through interviewing the participants, families or nursing staff. The preferred music were Taiwanese and Chinese songs from the 1930s to 1970s with moderate rhythm and tempo.

Sessions were delivered in the morning, once a week for 12 weeks, resulting in 12 sessions.

Control group

Participants participated in a rest and reading session, at the same intervals and under the same conditions as the experimental group. There was no active participation in music interventions in the control group.

Lord 1993

Music-based therapeutic intervention: mixed music programme (versus two control groups, jigsaw puzzle activities and a control group of usual care)

Experimental group

A 30-minute, mixed group music intervention, during which music of the "Big Bands" of the 1920s and 1930s were played. It is not clear if the music used was repeated every session or varied from session to session. The group had a size of 20 participants. Active music making (on triangles and tambourines) and singing was possible. It is not clear to what degree active music-making was stimulated by personnel or depended on participants' initiative only.

Personnel delivering the session was an "activities specialist" and two nurses. Sessions were delivered six times per week and continued for six months, therefore totalling 156 sessions.

Control group 1

Participants were given several puzzle-play activities (cardboard jigsaw cutouts and pegboard puzzles), new puzzles were introduced periodically.

Control group 2

Participants received the usual recreational activities of drawing, painting and watching television.

Lyu 2014

Music-based therapeutic intervention: active group music therapy (versus a reading control condition and a control group of usual care)

Experimental group

A 30-minute, group, active music intervention consisting of the singing of familiar songs. The participants learnt to sing the songs, or sang after the therapists. Classical and soothing old songs familiar to most participants were selected. A qualified music therapist delivered the intervention daily for three months.

Control group 1

The reading of familiar lyrics without music, supervised by a music therapist.

Control group 2

Participants received care as usual.

Narme 2012-study 1; Narme 2012-study 2

Music-based therapeutic interventions: group music programme (versus the control condition of art therapy in study 1, and versus cooking in study 2)

Study 1: experimental group

A two-hour, structured, mixed group intervention, with a maximum of 12 participants. Music selections were chosen independent of participants' preference and were played through a loudspeaker. The selections varied from classical music to songs from the 1950s and included instrumental and vocal music, and varied from 'calming' to 'dynamic' music. Calming music was used at the start and end of each session. The order of the musical selections was the same for every session, and pieces were played twice if participants expressed the wish to hear a song again. Participants were encouraged to play along (on percussion instruments, maracas or bell chains), sing and improvise. Participants were stimulated to express their feeling and memories evoked by the activity.

Study 1: control group

The control intervention in study 1 was another pleasant art therapy intervention. The painting session offered participants the use of wax crayons, colouring pencils, felt pens and gouache painting. They were stimulated to create simple drawings, to make circular movements with different materials and to make drawings based on their imagination. Participants were also encouraged to express their feeling and memories evoked by the activity.

Personnel delivering the two interventions were two psychologists. All sessions lasted two hours and were delivered twice a week during three weeks, totalling 12 hours during six sessions.

Study 2: experimental group

The same two-hour, structured, mixed group intervention was delivered by two psychologists, and the sessions were delivered twice a week, but during four weeks, and therefore totalling 16 hours during eight sessions.

Study 2: control group

The control intervention in study 2 was cooking, because it was a pleasant activity that stimulated a number of senses. There was more interaction compared to the painting control condition. Further, similar to the music therapy intervention, the cooking intervention also involved alternating productive (prepare a recipe) and receptive phases (taste a dessert). The sessions included preparing a different recipe collectively, with roles distributed according to the participants' abilities. Participants were encouraged to taste ingredients and verbalise remembrances.

Narme 2014

Cochrane
Library

Trusted evidence.
Informed decisions.
Better health.

Cochrane Database of Systematic Reviews

Music-based therapeutic intervention: a group music programme (versus cooking as the control condition)

Experimental group

A 60-minute, structured, mixed group intervention, with a maximum of eight participants. Music selections were chosen independent of the participants' preferences, and were played on a CD player (loudspeaker). The selections varied from classical music to songs from the 1950s to 1980s, included minor and major keys) and were 'calming' with slow to moderate tempo and 'arousing' music with a higher tempo. Calming music was used at the start and end of the session. The same playlist was used in the same order for each music session, but pieces were played twice if participants expressed the wish to hear a song again. Participants were asked to listen or to play along (on percussion instruments: clapping or playing hand drums) and sing along. Receptive and active phases were alternated. Participants were encouraged to express their feelings and autobiographical memories evoked by the activity.

The sessions were delivered twice a week, for a period of four weeks, totalling eight one-hour sessions. Personnel delivering the intervention were "two supervisors," including one psychologist, with no prior education in music therapy.

Control group

A cooking intervention, in which participants were asked to make a different recipe for each session (e.g. chocolate cake; French pancakes). Each session commenced with a game about ingredients where participants were asked to collectively prepare a given recipe. Roles were distributed according to participants' abilities (e.g. cutting, peeling, measuring quantities, mixing or cooking). Receptive (tasting) and productive phases were alternated. Participants were encouraged to express their feelings and autobiographical memories evoked by the activity.

The sessions had a duration of one hour and were delivered twice a week, for a period of four weeks, totalling eight one-hour sessions. Personnel delivering the intervention were "two supervisors," including one psychologist, with no prior education in music therapy.

Prieto Alvarez 2022

Music-based therapeutic intervention: group music therapy (versus 2 control groups, an enrichment programme and watching television)

Experimental group

A 50-minute, mixed (primarily active) group intervention delivered by a neurologic music therapist in a closed room. Group size was four to five residents. The therapeutic approach was habilitative, person-centred, focused on active engagement and applied Neurologic Music Therapy (NMT) techniques. The music therapist completed the NMT training and fellowship and was employed by the assisted living community in the USA.

Each session followed a structure: welcome, warm-up, main theme, and conclusion. In the welcome part, cognitive NMT techniques were used to stimulate arousal and increase orientation. The warm-up part made use of sensory-motor NMT techniques and speech/language NMT techniques. Activities included, for example, physical movement, patterned and metered tongue twister exercises and vocal exercises training aspects of voice. The main theme part also covered cognitive NMT techniques, using predetermined themes such as "seasons", "patriotism", "relationships" and improvised and pre-composed music. Activities included, for example, analysing song lyrics, conducting exercises and rhythmic call-response improvisations. The predetermined theme and mood of group members were the main criteria for the music therapist's choice of songs or pre-recorded music. Musical parameters of existing songs (for example, articulation, tempo, or lyrics) were changed when needed.

The sessions were delivered four times a week during two weeks, totalling eight sessions. Sessions were delivered at 11.00 am or at 2.00 pm.

Control group 1

A 50-minute, active, enrichment group intervention delivered by a program coordinator employed by the assisted living facility in a closed room. Group size was four to five residents. Physical and cognitive exercises were offered in a similar session structure and making use of the same predetermined themes as the experimental group. Activities included, for example, exergaming, stretching, ball-tossing games, balloon tennis, trivia questions, word games and reminiscing, making use of a person-centred approach. Session frequency and timing were equal to the experimental group.

Control group 2

TV watching sessions were offered with the same duration and under the same conditions as the experimental group and control group 1. The session was not actively facilitated. A certified nursing assistant was present and interacted when needed. The documentaries and series were chosen parallel to the themes used in the experimental and control 1 group. Session frequency and timing were equal to the experimental group.

Note, in providing the intervention and in data collection, and in any of the three groups, at least five co-ordinators or research assistants were involved. It is unclear which tasks were conducted by the music therapist, a dance/movement therapist, two music therapy interns and a volunteer.

Raglio 2010a

Music-based therapeutic intervention: active individual music therapy based on relationship (versus a control group of usual care)

Experimental group

A 30-minute, active, non-verbal, individual music therapy intervention, in which free musical improvisation was used to build a relationship between participant and music therapist. During the session, the participant and the music therapist had a non-verbal dialogue and expressed their feelings and emotions through non-verbal behaviours (possibly by using voice and tapping, not specified in the text) and by playing musical instruments. Choice of instruments included rhythmic-melodic instruments, percussions, glockenspiels, xylophones, etc. Sharing emotions, raising awareness and the possibility of introducing new ways of expression and communication were a focus of the session and may have led to empathetic processes and mutual calibration.

A music therapist delivered the sessions, which were twice a week for 15 weeks, with a total of 30 sessions.

Control group

Usual care

Raglio 2010b

Music-based therapeutic intervention: active group music therapy based on relationship (versus a control group of usual care)

Experimental group

A 30-minute, active, non-verbal, group music therapy intervention, in which free musical improvisation was used to build a relationship between participant and music therapist. Groups had three participants. The intervention focused on favouring the moments of attunement that helped organise and regulate the participants' behaviours and emotions. Participants and music therapist interacted and expressed their feelings and emotions through non-verbal behaviours and using musical instruments. Note that this approach is inspired by the intersubjective psychology (references provided in the article).

A music therapist delivered the sessions. The sessions were delivered in three non-continuous treatment cycles consisting of four weeks of three sessions per week followed by one month of no treatment (washout; however, not in the context of a cross-over design). The total number of sessions was 36, within six months.

Control group

Usual care

Raglio 2015

Music-based therapeutic intervention: active music therapy (versus music listening and a control group of usual care)

Experimental group

A 30-minute, individual, active music therapy, which consisted of playing and improvising on instruments, focused on promoting 'affect attunement' moments. The music therapist followed the participants' rhythm and music production (also introducing variations) to create nonverbal communication. During the session, the music therapist built a relationship with the participant by singing and using melodic and rhythmic instruments (improvisation), facilitating the expression and modulation of the participant's emotions.

The intervention was delivered by a certified, specifically trained music therapist, twice a week for 10 weeks in a separate, medium-sized room.

Control group 1

Individualised, 30-minute music listening sessions, delivered through speakers in the room of the participant or in a quiet private place.

Control group 2

Participants received standard care, which included daily educational, occupational and physical activities performed under supervision of specialised professionals. Standard care did not include music exposure.

Reschke-Hernández 2019

Music-based therapeutic intervention: active group music therapy (versus a control condition of themed conversation)

Experimental group

Twenty-five-minute, active, group singing sessions were delivered by a credentialed music therapist. The therapist received training in delivering a clinical practice model developed by the author based on literature and expert input. The clinical practice model was based on the integration of the principles of personhood (Kitwood 1997)) with the biopsychosocial model (Engel 1980; Sarafino 2008).

Family members provided information regarding participants' music preferences using a checklist and preferences were confirmed with participants by singing a few song excerpts to them based on what family had reported.

The music intervention protocol consisted of six intervention attributes (cognition, attention, familiarity, audibility, structure, autonomy) and contained detailed practical application instructions for each intervention theme based on the reciprocal response model (Hargreaves 2011).

The sessions consisted of 85% interactive music (singing of themed songs selected on the basis of participants' background and needs and partly on the basis of therapists' repertoire) and 15% verbal interaction (using conversation starters specific to each song and verbal facilitation strategies). Each session began with a welcome song after which themed songs were sung, accompanied by the therapist on guitar and ended with a closing song. The therapist optimised participation by using specific musical facilitation strategies: slowing down the tempo (cognition); chaining songs together and matching participants' current state, as well as varying lyrics and dynamics (in contrast to repetition) to surprise (attention); starting with the most familiar part of the song and playing in the familiar song style (familiarity); chordal guitar and strong vocal skills and articulation (audibility); consistent use of a musical opening, closing and transitions, as well as visual aids for topics, song lyrics and instruments (structure); providing choices and using nonverbal and musical cues to sing (autonomy).

Sessions were delivered in the afternoons, in groups of three to five participants, three times a week for two weeks, resulting in six sessions.

Control group

Twenty-five-minute verbal discussion sessions were delivered by credentialed music therapists. The therapist led the conversation around assigned topics, making use of the clinical practice model developed by the author, but without any music. Conversation starters and prepared trivia were used throughout the session. Images were used (e.g. photograph of a farm) and contained non-musical cues only.

Sessions were delivered in the afternoon, in groups of three to five participants, three times a week for two weeks, resulting in six sessions.

Ridder 2013

Music-based therapeutic intervention: individual mixed music therapy (versus a control group of usual care)

Experimental group

An individual, mixed music therapy intervention, not prestructured, delivered by music therapists with a mean duration of 33.8 (standard deviation 9.91) minutes. The aim of the music therapy was phrased in a more positive way than a goal of reducing (e.g. challenging behaviour ("to facilitate initiative, engagement, self-expression and mutual understanding")). The authors refer to Tom Kitwood for the theoretical basis of a relation-based and person-centred approach in music therapy.

Vocal or instrumental improvisation, singing, dancing/moving, listening and talking/going for a walk could be part of the session. The music accompanying the activities was pre-recorded or live music, and consisted of 'free' improvisation or based on songs/melodies. The overall aim of the music therapy was to facilitate initiative, engagement, self-expression and mutual understanding. Clinicians were instructed to be aware of at least three different ways of applying music in therapy: catching attention and creating a safe setting, regulating arousal level to a point where self-regulation is possible and engaging in social communication to fulfil psychosocial needs. The session was not especially focused on decreasing agitation.

Music therapists with university-level training delivered the intervention, which was twice a week for a period of six weeks, with 12 sessions offered in total. The mean number of sessions received was 10 (standard deviation 2.82, range 0 to 13).

Control group

Usual care

Sakamoto 2013

Music-based therapeutic intervention: an individual mixed music (therapy) intervention (versus 2 control groups)

Experimental group

A 30-minute, individual, mixed music therapy intervention. The selection of music was based on determination of a period of the participant's life that was recalled most frequently, interviews with participants and their family, and links to special memories. Music was selected for probable evoking of positive emotions such as pleasure or joy.

The selected music was played via a CD player (loudspeaker). The participants also participated in activities guided by a music facilitator, including clapping, singing and dancing. The sessions took place in a familiar room.

During the session, participants were monitored to confirm that "the music was suitable in terms of engaging the participants and eliciting a joyful emotional state". Participants' attention was directed to the music, and "an interactive approach that responded to the participants' emotional reactions to the music" was used.

The sessions were delivered by music therapists, occupational therapists and nurses, each trained for 10 days in delivering the intervention. The sessions took place weekly for a period of 10 weeks (10 sessions in total), and were scheduled between 10 am and 11 am.

Control group 1: passive individual music intervention (the music intervention did not meet our criteria for music-based therapeutic interventions)

A 30-minute, individual music intervention. The selection of music was made based on determination of a period of participants' life that was recalled most frequently, interviews with participants and their family, and links to special memories. Music was selected for probable evoking of positive emotions such as pleasure or joy.

The selected music was played via a CD player (loudspeaker). Personnel delivering the intervention was a carer and a music provider, but no interaction took place between personnel and participants during the intervention. The session took place in a familiar room, weekly, for a period of 10 weeks (10 sessions in total), and were scheduled between 10 am and 11 am.

Control group 2: observation

Spending 30 minutes in their own room, as usual, in a silent environment, with a carer observing from a distance and no interaction between carer and participant. The sessions took place weekly for a period of 10 weeks (10 sessions in total), and were scheduled between 10 am and 11 am.

Sung 2012

Music-based therapeutic intervention: active group music intervention (versus a control group of usual care)

Experimental group

A 30-minute, active group music therapy intervention with movement. The sessions included five minutes of warm-up and five minutes of cooling down (stretching major muscle groups and breathing exercise with music). During the main part of the session, participants were guided in the use of percussion instruments (hand bell, tambourine, maracas, guiro tone block, flapper and loop bell) while listening to music and songs familiar to the participants. Participants' music preferences were assessed through interviewing the participants, carers, families or nursing staff. The preferred music was Taiwanese and Chinese songs from the 1950s to 1970s with moderate rhythm and tempo.

Sessions were delivered by a nursing researcher and two research assistants trained in providing the music intervention, twice a week for six weeks, with a total of 12 sessions.

Control group

Usual care

Svansdottir 2006

Music-based therapeutic intervention: mixed group music therapy (versus a control group of usual care)

Experimental group

A 30-minute, mixed music therapy intervention, with three or four participants per group. The sessions were accompanied by guitar playing and consisted of (listening to) singing with the help of songbooks, playing along on various kind of instruments (choice of instruments not specified), instrumental improvisation and moving/dancing, if "patients had an urge to move and dance." The music therapist selected a collection of songs that were familiar to the residents.

A music therapist delivered the sessions three times a week for six weeks, totalling 18 sessions.

Control group

Usual care

Tang 2018

Music-based therapeutic intervention: mixed group music intervention (versus a control condition of care as usual)

Experimental group

Fifty-minute, mixed active and receptive group music sessions were delivered by a trained therapist assisted by a research assistant.

The music intervention aimed at promoting the residents' receptive ability stimulating the senses by listening to music and re-establishing a connection with the environment. The session consisted of three music interventions comprising: (1) sensory-stimulating music and sound listening exercises, (2) singing nostalgic songs and (3) playing an instrument. During the listening part of the session, participants were asked to distinguish the sounds of musical instruments (e.g. drum, gong, mouth organ, flute), sounds from the natural world (for example, wind, rain, thunder), and sounds of animals (for example, birds, goats, horses, chooks). In the singing part of the session, the participants were instructed and supported (in steps) to sing and clap along with nostalgic red songs (revolutionary songs that express love

Cochrane
Library

Trusted evidence.
Informed decisions.
Better health.

Cochrane Database of Systematic Reviews

and praise for the people's motherland), nostalgic nursery rhymes and nostalgic Cantonese opera. In the active music making part of the session, participants were guided (in steps) to play a simple song on a xylophone with the help of digital numbers marked beside the scale.

Sessions were delivered three mornings a week for 12 weeks, resulting in 36 sessions.

Control group

The control group received care as usual.

Thornley 2016

Music-based therapeutic intervention: active individual music therapy (versus a control condition with individual active engagement)

Experimental group

A 60-minute, individual, active music therapy, which consisted of singing and playing simple instruments to music adapted to the participants' preferences.

An accredited music therapist delivered the intervention twice a week for four weeks. The participants were encouraged to actively engage in the musical process and to follow the music therapist's lead. Participants were provided with specific instructions on how to participate by singing or playing simple instruments (or both), including maracas and small drums. The music was selected in accordance with participant preferences and was of a calming nature.

Control group

A 60-minute, individual, active engagement and attention (active engagement) intervention delivered by a social worker, including supportive interviewing, and encouragement of expression through simple occupational activities such as folding towels and browsing magazines. The control intervention was also delivered twice a week for four weeks.

Vink 2013

Music-based therapeutic intervention: mixed group music therapy (versus a control condition with general recreational activities)

Experimental group

A 40-minute, mixed, group music therapy intervention, which consisted of a welcome song; listening to selected music, sung or played by the therapist (Dutch familiar songs, classical and folk music); and singing, dancing or playing along (on simple rhythm instruments). Within the group session, the therapist adjusted the level of each intervention to individual capacities. The music accompanying the session was played live, e.g. on a piano or guitar and was selected with the goal of inciting pleasant memories and reducing agitation. For this, musical parameters were used "such as slow tempo and little instrumentation."

Music therapists delivered the intervention in rooms away from the nursing home ward. The sessions were delivered twice a week for four months, with a total of up to 34 sessions.

Control group

General recreational activities, such as handwork, playing shuffleboard, making flower bouquets and playing games. The sessions also lasted 40 minutes, were delivered twice a week for four months and were also held in rooms away from the nursing home ward.

Werner 2017

Music-based therapeutic intervention: group music therapy (versus a control condition of recreational choir singing)

Experimental group

A 40-minute, mixed (active and receptive) group intervention delivered by a music therapist (who was not employed by the nursing home) in an open group in the lunchroom on the wards. The therapeutic approach was holistic, person-centred and tailored to the biography of each individual nursing home resident. Other residents, nursing staff and relatives were able to join the sessions.

A flexible, process-oriented approach was used. Emotional states of individual group members or of the whole group were the main criteria for the music therapist's decision to choose songs or music to listen to in order to reflect or to help change these states. The authors refer to Muthesius 2010 and Hamberger 2011 for the basis of the interactive music therapeutic approach.

Group singing, receptive music therapy, instrumental improvisation and dance or movement was selected depending on reflections on the participant's individual biography or emotional states of participants individually or as a group. After each musical action or music listening, the participants reflected on it verbally. Each session started with the greeting of each participant personally, and always ended with the same song.

The sessions were delivered twice a week for 10 weeks, totalling 20 sessions.

Control group

A 1.5-hour weekly recreational group singing activity in an open setting, delivered by either a gerontopsychiatric specialist employed by the nursing home (a nurse, occupational therapist or social worker with specialised gerontopsychiatric training), or a musician from outside the nursing home. The participants were instructed to sing songs collectively.

The sessions were delivered according to the same schedule as the music therapy but once a week during 10 weeks (weeks 1-11, therefore probably 10 in total).

Appendix 4. Funnel plots

In order to assess potential funnel plot asymmetry, we drew funnel plots for comparisons with at least 10 included studies in the meta-analysis (see Assessment of reporting biases).

Music-based therapeutic interventions versus usual care: end of treatment

Figure 4: Analysis 1.4 Behaviour problems: agitation or aggression

Figure 4. Funnel plot for Analysis 1.4: Music-based therapeutic interventions compared to usual care, outcome (end of treatment): behaviour problems: agitation or aggression Horizontal axis: direction and size of effect Vertical axis: standard error
Dotted line: mean effect size *Abbreviations*
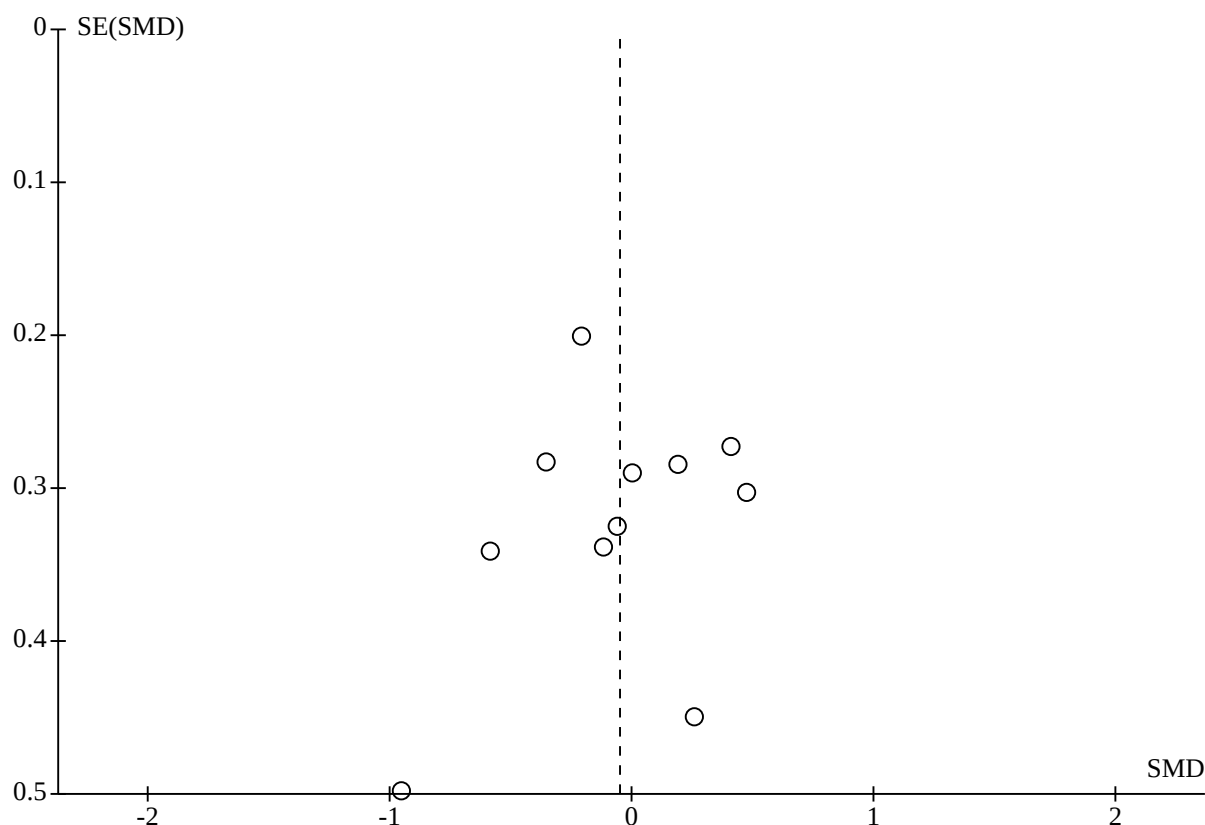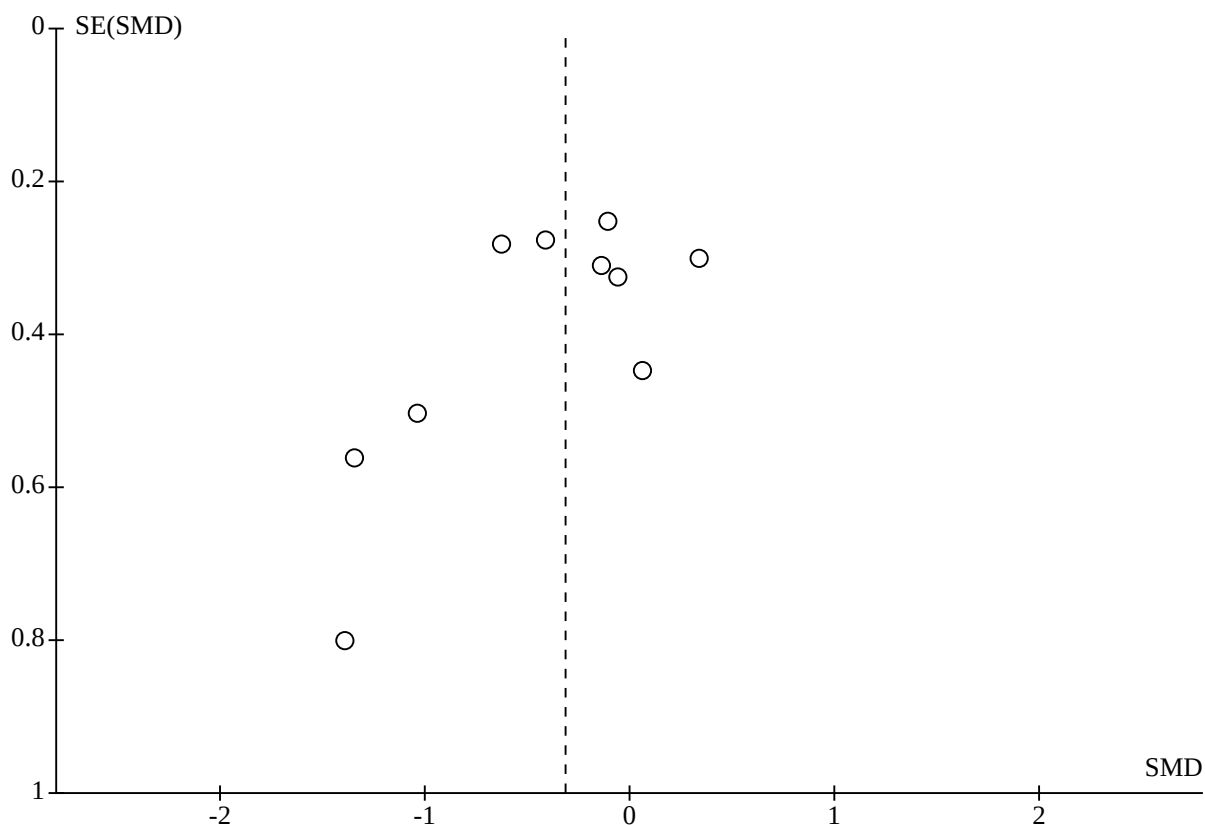**SE: standard error; SMD: standardised mean difference.**

Figure 5: Analysis 1.5 Behaviour problems overall

Figure 5. Funnel plot for Analysis 1.5: Music-based therapeutic interventions compared to usual care; outcome (end of treatment): behaviour problems: overall Horizontal axis: direction and size of effect Vertical axis: standard error Dotted line: mean effect size *Abbreviations* **SE: standard error; SMD: standardised mean difference.**

Music-based therapeutic interventions versus other activities: end of treatment

Figure 6: Analysis 3.2 Mood disturbance or negative affect: depression

Figure 6. Funnel plot for Analysis 3.2: Music-based therapeutic interventions compared to other activities; outcome (end of treatment): mood disturbance or negative affect: depression Horizontal axis: direction and size of effect Vertical axis: standard error
Dotted line: mean effect size *Abbreviations*
SE: standard error; SMD: standardised mean difference.
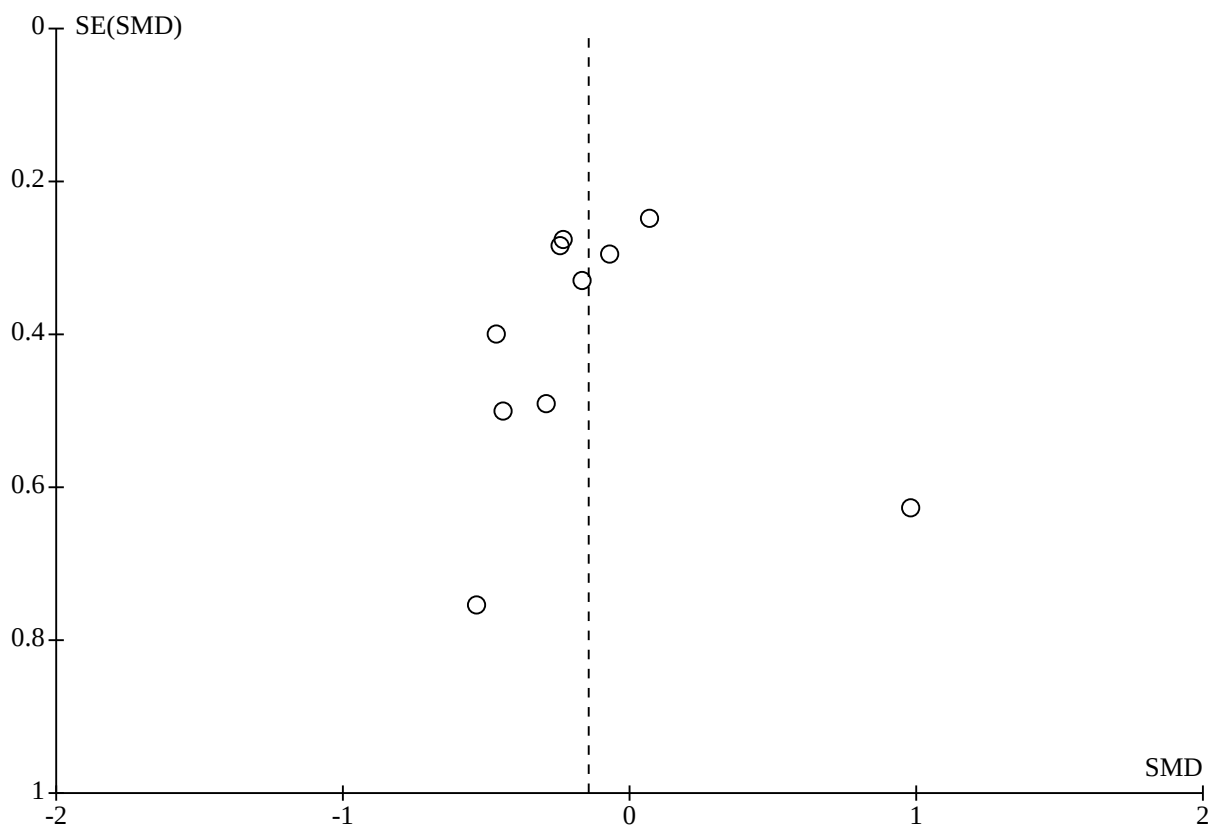
Figure 7: Analysis 3.3 Mood disturbance or negative affect: anxiety

Figure 7. Funnel plot for Analysis 3.3: Music-based therapeutic interventions compared to other activities; outcome (end of treatment): mood disturbance or negative effect: anxiety Horizontal axis: direction and size of effect Vertical axis: standard error
Dotted line: mean effect size *Abbreviations*
SE: standard error; SMD: standardised mean difference.
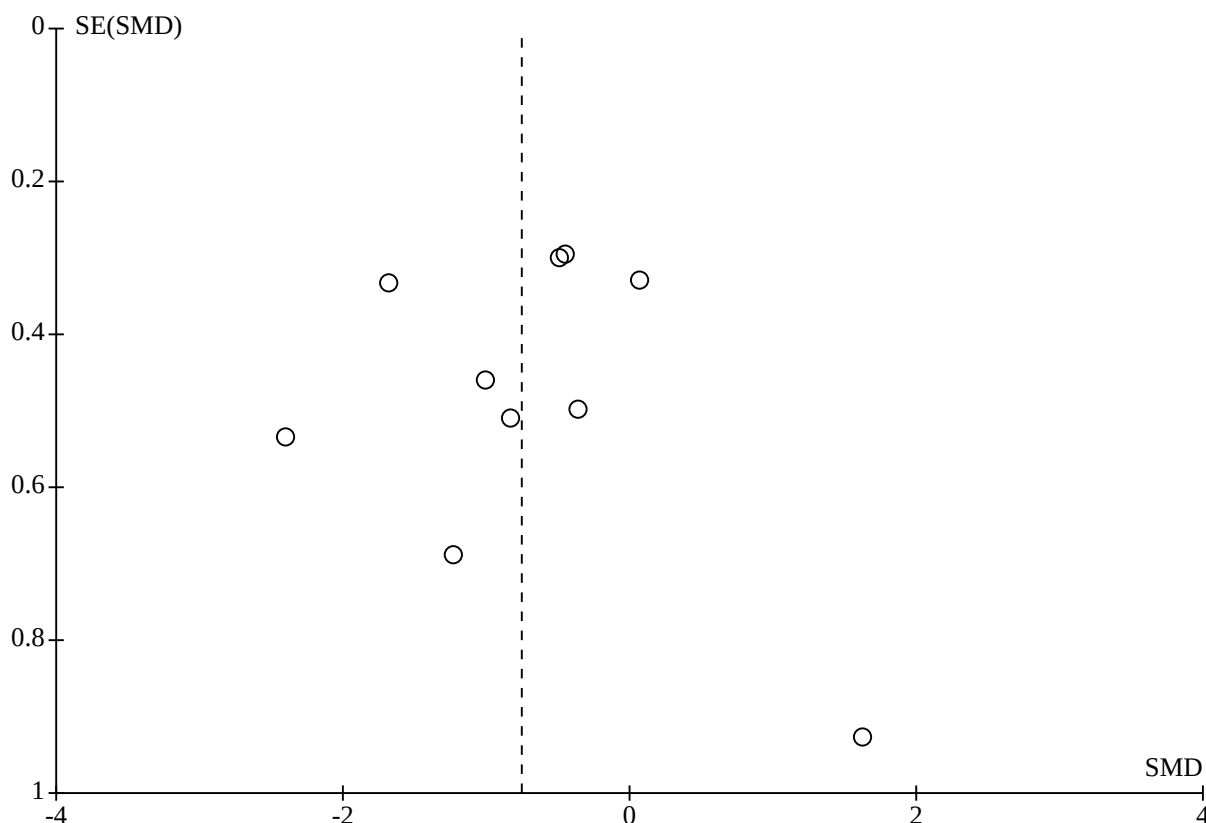
For behaviour problems overall, the funnel plot is skewed (Figure 5), with three studies in the lower left corner. None of the studies reported a positive standardised mean difference (SMD) and a large standard error (SE) to counterbalance the ones in the lower left corner with a negative SMD and a large SE, while sudies with a small SE are all situated close to SMD 0. This could be a sign of publication bias. For the other outcomes, the funnel plots look rather symmetrical.

Appendix 5. Sensitivity analysis: involvement of trained music therapist

Compared to usual care – end of treatment: SMD (95% CI), n of studies

Outcome	Overall	Definite or possible trained music therapists (could be assisted)	Definite music therapist ONLY	Definite music therapists and no potential conflict of interest related to funding (OR no funding reported)
1.1 Emotional well-being, including quality of life	0.14 (−0.29 to 0.56) 4 studies	−0.13 (−0.77 to 0.52) 1 study	SAME	SAME

(Continued)

Outcome	Overall	Definite or possible trained music therapists (could be assisted)	Definite music therapist ONLY	Definite music therapists and no potential conflict of interest related to funding (OR no funding reported)
1.2 Mood disturbance or negative affect: depression	−0.23 (−0.42 to −0.04) 9 studies	SAME	SAME	−0.28 (−0.48 to −0.07) 8 studies
1.3 Mood disturbance or negative affect: anxiety	−0.15 (−0.39 to 0.09) 7 studies	−0.05 (−0.33 to 0.22) 5 studies	SAME	SAME
1.4 Behavioural problems: agitation or aggression	−0.05 (−0.27 to 0.17) 11 studies	−0.07 (−0.27 to 0.13) 9 studies	-0.02 (-0.22 to 0.18) 8 studies	−0.04 (−0.28 to 0.19) 6 studies
1.5 Behavioural problems: overall	−0.31 (−0.60 to −0.02) 10 studies	−0.26 (−0.55 to 0.03) 9 studies	SAME	−0.28 (−0.62 to 0.05) 8 studies
1.6 Social behaviour	0.22 (−0.14 to 0.57) 2 studies	SAME	0.17 (−0.41 to 0.76) 1 study	SAME
1.7 Cognition	0.19 (−0.02 to 0.41) 7 studies	SAME	0.15 (-0.09 to 0.39) 6 studies	0.25 (−0.04 to 0.55) 4 studies

Compared to usual care – long-term effects: SMD (95% CI), n of studies

Outcome	Overall	Definite or possible trained music therapists (could be assisted)	Definite music therapist ONLY	Definite music therapists and no potential conflict of interest related to funding (OR no funding reported)
2.1 Emotional well-being, including quality of life	0.17 (−0.80 to 1.14) 3 studies	SAME	SAME	SAME
2.2 Mood disturbance or negative affect: depression	−0.07 (−0.31 to 0.18) 5 studies	SAME	SAME	SAME
2.3 Mood disturbance or negative affect: anxiety	−0.06 (−0.48 to 0.37) 3 studies	SAME	SAME	SAME
2.4 Behavioural problems: agitation or aggression	−0.17 (−0.42 to 0.09)	SAME	SAME	SAME

(Continued)

	4 studies			
2.5 Behavioural problems: overall	−0.19 (−0.52 to 0.14) 6 studies	SAME	SAME	-0.19 (-0.62 to 0.23) 5 studies
2.6 Social behaviour	No studies	No studies	No studies	No studies
2.7 Cognition	0.09 (−0.24 to 0.41) 2 studies	SAME	SAME	0.13 (−0.26 to 0.52) 1 study

Compared to other interventions – end of treatment: SMD (95% CI), n of studies

Outcome	Overall	Definite or possible trained music therapists (could be assisted)	Definite music therapist ONLY	Definite music therapists and no potential conflict of interest related to funding (OR no funding reported)
3.1 Emotional well-being, including quality of life	0.20 (−0.09 to 0.49) 10 studies	0.12 (−0.17 to 0.41) 6 studies	0.06 (−0.28 to 0.39) 5 studies	SAME
3.2 Mood disturbance or negative affect: depression	−0.14 (−0.36 to 0.08) 10 studies	−0.12 (−0.36 to 0.12) 8 studies	−0.04 (−0.36 to 0.28) 5 studies	SAME
3.3 Mood disturbance or negative affect: anxiety	−0.75 (−1.27 to −0.24) 10 studies	−0.86 (−1.60 to −0.11) 6 studies	−0.18 (−1.17 to 0.80) 3 studies	SAME
3.4 Behavioural problems: agitation or aggression	0.01 (−0.31 to 0.32) 6 studies	−0.03 (−0.41 to 0.35) 4 studies	−0.07 (−0.57 to 0.44) 3 studies	SAME
3.5 Behavioural problems: overall	−0.08 (−0.33 to 0.17) 8 studies	−0.16 (−0.44 to 0.11) 6 studies	SAME	−0.23 (−0.54 to 0.08) 5 studies
3.6 Social behaviour	0.52 (0.08 to 0.96) 4 studies	0.27 (−0.15 to 0.68) 2 studies	0.41 (−0.66 to 1.49) 1 study	SAME
3.7 Cognition	0.12 (−0.21 to 0.45)	0.04 (−0.39 to 0.48)	0.09 (−0.44 to 0.62)	0.34 (−0.78 to 1.47)

(Continued)

	5 studies	3 studies	2 studies	1 study

Compared to other interventions – long term: SMD (95% CI), n of studies

Outcome	Overall	Definite or possible trained music therapists (could be assisted)	**Definite music therapist ONLY**	**Definite music therapists and no potential conflict of interest related to funding (OR no funding reported)**
4.1 Emotional well-being, including quality of life	0.10 (−0.29 to 0.49) 4 studies	−0.16 (−0.64 to 0.31) 2 studies	SAME	SAME
4.2 Mood disturbance or negative affect: depression	−0.07 (−0.39 to 0.25) 4 studies	−0.07 (−0.44 to 0.30) 3 studies	−0.11 (−0.53 to 0.31) 2 studies	SAME
4.3 Mood disturbance or negative affect: anxiety	−0.53 (−1.31 to 0.25) 4 studies	−1.03 (−2.04 to −0.01) 2 studies	−0.58 (−1.16 to 0.01) 1 study	SAME
4.4 Behavioural problems: agitation or aggression	0.10 (−0.66 to 0.86) 2 studies	−0.27 (−0.85 to 0.30) 1 study	SAME	SAME
4.5 Behavioural problems: overall	−0.09 (−0.39 to 0.22) 4 studies	−0.16 (−0.50 to0.19) 3 studies	SAME	−0.25 (−0.67 to 0.17) 2 studies
4.6 Social behaviour	0.53 (−0.53 to 1.60) 2 studies	No studies	No studies	No studies
4.7 Cognition	0.04 (−0.56 to 0.64) 1 study	SAME	SAME	No studies

Abbreviations
CI: confidence interval; n: number; SAME: same studies and results as left-hand cell; SMD: standardised mean difference

Appendix 6. Sensitivity analysis: risk of bias

Compared to usual care – end of treatment: SMD (95% CI), n of studies

Outcome	Overall	No high risk of bias (apart from unmasked therapist and participants)
1.1 Emotional well-being, including quality of life	0.14 (−0.29 to 0.56) 4 studies	0.03 (−0.50 to 0.57) 1 study
1.2 Mood disturbance or negative affect: depression	−0.23 (−0.42 to −0.04) 9 studies	−0.25 (−0.47 to −0.03) 7 studies
1.3 Mood disturbance or negative affect: anxiety	−0.15 (−0.39 to 0.09) 7 studies	−0.05 (−0.33 to 0.22) 5 studies
1.4 Behavioural problems: agitation or aggression	−0.05 (−0.27 to 0.17) 11 studies	−0.01 (−0.22 to 0.20) 7 studies
1.5 Behavioural problems: overall	−0.31 (−0.60 to −0.02) 10 studies	−0.17 (−0.46 to 0.12) 6 studies
1.6 Social behaviour	0.22 (−0.14 to 0.57) 2 studies	SAME
1.7 Cognition	0.19 (−0.02 to 0.41) 7 studies	0.20 (−0.01 to 0.42) 6 studies

Compared to usual care – long-term effects: SMD (95% CI), n of studies

Outcome	Overall	No high risk of bias (apart from unmasked therapist and participants)
2.1 Emotional well-being, including quality of life	0.17 (−0.80 to 1.14) 3 studies	0.14 (−0.39 to 0.68) 1 study
2.2 Mood disturbance or negative affect: depression	−0.07 (−0.31 to 0.18)	−0.02 (−0.28 to 0.24) 4 studies

(Continued)

	5 studies	
2.3 Mood disturbance or negative affect: anxiety	−0.06 (−0.48 to 0.37) 3 studies	SAME
2.4 Behavioural problems: agitation or aggression	−0.17 (−0.42 to 0.09) 4 studies	SAME
2.5 Behavioural problems: overall	−0.19 (−0.52 to 0.14) 6 studies	−0.14 (−0.43 to 0.15) 4 studies
2.6 Social behaviour	No studies	No studies
2.7 Cognition	0.09 (−0.24 to 0.41) 2 studies	SAME

Compared to other intervention – end of treatment: SMD (95% CI), n of studies

Outcome	Overall	No high risk of bias (apart from unmasked therapist and participants)
3.1 Emotional well-being, including quality of life	0.20 (−0.09 to 0.49) 10 studies	0.31 (−0.59 to 1.21) 3 studies
3.2 Mood disturbance or negative affect: depression	−0.14 (−0.36 to 0.08) 10 studies	−0.26 (−0.56 to 0.05) 4 studies
3.3 Mood disturbance or negative affect: anxiety	−0.75 (−1.27 to −0.24) 10 studies	−1.08 (−1.95 to −0.22) 5 studies
3.4 Behavioural problems: agitation or aggression	0.01 (−0.31 to 0.32) 6 studies	0.13 (−0.46 to 0.73) 2 studies
3.5 Behavioural problems: overall	−0.08 (−0.33 to 0.17) 8 studies	0.03 (−0.42 to 0.48) 3 studies
3.6 Social behaviour	0.38	0.43

![Cochrane Library] Trusted evidence. Informed decisions. Better health.

Cochrane Database of Systematic Reviews

(Continued)

	(0.07 to 0.70)	(−0.15 to 1.01)
	5 studies	2 studies
3.7 Cognition	0.12 (−0.21 to 0.45) 5 studies	0.09 (−0.29 to 0.48) 3 studies

Compared to other intervention – long term: SMD (95% CI), n of studies

Outcome	Overall	No high risk of bias (apart from unmasked therapist and participants)
4.1 Emotional well-being, including quality of life	0.10 (−0.29 to 0.49) 4 studies	0.18 (−0.22 to 0.58) 3 studies
4.2 Mood disturbance or negative affect: depression	−0.07 (−0.39 to 0.25) 4 studies	−0.04 (−0.41 to 0.33) 3 studies
4.3 Mood disturbance or negative affect: anxiety	−0.53 (−1.31 to 0.25) 4 studies	SAME
4.4 Behavioural problems: agitation or aggression	0.10 (−0.66 to 0.86) 2 studies	SAME
4.5 Behavioural problems: overall	−0.09 (−0.39 to 0.22) 4 studies	−0.09 (−0.44 to 0.25) 3 studies
4.6 Social behaviour	0.53 (−0.53 to 1.60) 2 studies	SAME
4.7 Cognition	0.04 (−0.56 to 0.64) 1 study	SAME

Abbreviations
CI: confidence interval; n: number; SAME: same studies and results as left-hand cell; SMD: standardised mean difference

WHAT'S NEW

Date	Event	Description
7 March 2025	New citation required and conclusions have changed	Search updated. Eight new studies added. Analysis adapted and conclusions changed. New co-author added to byline
7 March 2025	New search has been performed	New searches for this review update were performed on 2 December 2021 and 30 November 2023.

HISTORY

Protocol first published: Issue 1, 2002
Review first published: Issue 3, 2004

Date	Event	Description
11 April 2017	New citation required and conclusions have changed	New studies included. Conclusions changed. New author
12 April 2016	New search has been performed	Updated search and potentially eligible studies included under studies awaiting classification
14 April 2010	New search has been performed	An update search was performed for this review on 14 April 2010. New studies were retrieved for possible inclusion or exclusion within the review. Two new studies have been included in this update.
26 November 2008	New search has been performed	A new update search was performed on 20 March 2008. New studies were retrieved for possible inclusion or exclusion in the review. Three new studies have been included in this update, and 15 new studies have been excluded. Risk of Bias tables have been completed for all included studies.
23 January 2006	New search has been performed	January 2006: the update searches of 5 December 2005 yielded 4 new trials, which were not suitable for inclusion. The results and conclusions of this review remain unchanged.

CONTRIBUTIONS OF AUTHORS

- JS, JCW, HS and AV contributed to all aspects of the review.
- AM and MB assisted with data extraction and commented on review drafts.

For the 2025 update, HS and MB assisted with data extraction and JCW supported the analysis and grading of the certainty of the evidence. AM screened new records and extracted data for this update. All authors contributed to the assessment of the eligibility of studies and commented on draft versions.

DECLARATIONS OF INTEREST

JS declares that she is involved in a trial on a music-based therapeutic intervention for people with dementia. She is serving as a methodologist and is supervising the PhD student who is conducting the trial, which could qualify for inclusion in a future update of

Cochrane
Library

Trusted evidence.
Informed decisions.
Better health.

Cochrane Database of Systematic Reviews

this review. This study started in 2019 and has concluded recently (results not fully available yet); it was funded by university funds and Alzheimer Nederland.

JW declares that he has no conflicts of interest.

AM is a Senior Clinical Psychologist and sole proprietor of the business Innovative Clinical Psychology Solutions Ltd (from 9 March 2021 to present), a private clinical psychology practice; she can see patients with dementia and neurological conditions due to her training, but this is not her primary client group. Until 29 December 2022, AM was employed by the NHS and saw people with suspected and confirmed cases of dementia clinically within this role. AM declares consultancy fees from Cochrane to Innovative Clinical Psychology Solutions Ltd for payment for contributing to this review (i.e. to assist with the process of the updated systematic search, including abstract screening and data extraction); this was a personal payment.

HS declares that she has no conflicts of interest.

AV is affiliated with the ArtEZ Music Therapy Department, and she has published opinion pieces on the website artez.nl. AV declares that she was involved in a study supported by the national funder ZonMw, an independent organisation that supports research and innovation in health, healthcare and well-being, as part of her dissertation at UMCG The Netherlands, which was eligible for inclusion in the review**. AV is involved in music therapy research and teaching.

MB is involved in teaching and practice of music therapy in persons with dementia. She declares that she has no conflicts of interest.

**We included two studies on which AV was an author. AV was not involved in assessing the eligibility of these studies, extracting data from them, assessing their risk of bias or grading the certainty of evidence involving these studies. These tasks were performed by JS and HS, and for other articles also by JW and AM. The lead review author, JS, and a co-author who is a Cochrane expert (JW; and also RS in previous versions), made the final decisions about analysis, presentation and interpretation of the data. None of the authors has a conflict of interest related to the effectiveness of music therapy.

SOURCES OF SUPPORT

Internal sources

- ArtEZ University of the Arts. Professorship Music-based Therapies and Interventions, Enschede, Netherlands

  AV employed by Higher Education Institute

External sources

- The Netherlands Organisation for Scientific Research (NWO); Innovational Research Incentives Scheme, Netherlands

  Career award to JS from the national funder of scientific research; Vidi grant number 917.11.339
- National Institute for Health Research (NIHR), UK

  This update was supported by the NIHR, via Cochrane Infrastructure funding to the Cochrane Dementia and Cognitive Improvement group, up to March 2023. The views and opinions expressed therein are those of the authors and do not necessarily reflect those of the Systematic Reviews Programme, NIHR, National Health Service or the Department of Health
- Cochrane, UK

  Consulting fees to Innovative Clinical Psychology Solutions to support AM with screening and data extraction
- European Research Council (ERC), Other

  European Union ERC Consolidator grant career award to JS; ID 771483

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

In the 2025 version

Objectives

We modified the objectives slightly to specify that mood disturbance or negative affect included depressive symptoms and anxiety, and to specify that we would assess adverse effects.

Outcomes

In previous versions of this review, we did not specify whether the outcome of adverse effects was primary or secondary. For the 2025 update, we specified it as a secondary outcome, and we assessed it qualitatively. We included anything that was reported in the studies about adverse effects even if adverse effects had not been measured formally as an outcome.

Methods

For the 2025 update, we implemented the following changes.

- As advised by the Central Editorial Service, we performed separate meta-analyses rather than subgroup analyses to distinguish between studies comparing music-based therapeutic interventions to usual care and studies comparing music-based therapeutic interventions to other active interventions.
- Further, in view of recent studies that employed cluster randomisation, we adjusted sample sizes for cluster-randomised trials that we included in the review when this had not been taken into account in the study publication, as advised in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2019).
- We performed subgroup analysis for individual versus group interventions with a minimum of five studies in each group. This replaced the sensitivity analysis on individual or group interventions in the previous updates (Van der Steen 2017; Van der Steen 2018).
- We updated our methods to the latest relevant guidance provided in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2019), such as the risk of bias tool for RCTs (Higgins 2017), as well as GRADE methodology and summary of findings tables (Schünemann 2023).

In previous versions

Inclusion criteria

The following changes were implemented in the Van der Steen 2017 version, and we then reassessed previously included studies using the new criteria, where necessary, consulting the lead author of the earlier version when in doubt (Vink 2003).

Study design

- We included only randomised controlled trials (RCTs) because, unlike at the time the protocol was written, we expected more RCTs to be available, and this study design is associated with a lower risk of bias than other study types. We included studies that used a physician's diagnosis of dementia if no data on formal criteria, such as DSM-IV, DSM-5 (major neurocognitive disorders) or comparable instruments, were available, in order to increase the relevance of the review to clinical practice and known underdiagnosis of dementia.

Intervention

- We assessed whether an article reported a music intervention with an individual therapeutic intent, including – but not limited to – interventions provided by qualified music therapists. In the protocol and first version of the review, we defined music therapy as any type of music therapy of at least five sessions. In the Van der Steen 2017 update, we added: "We defined therapeutic music-based interventions as: therapy provided by a qualified music therapist, or interventions based on a therapeutic relationship and meeting at least two of the following criteria/indicators: (a) therapeutic objective which may include communication, relationships, learning, expression, mobilisation and other relevant therapeutic objectives; (b) music matches individual preferences; (c) active participation of the people with dementia using musical instruments or singing; (d) participants had a clinical indication for the intervention or were referred for the intervention by a clinician. We also required music to be a main element of the intervention (e.g. not merely moving with use of music). Simple participation in a choir would not meet our definition of a therapeutic intervention." Therefore, we focused on therapeutic aspects and elements that are more complex and require special skills, while also targeting the individual, compared with, for example, playing recorded music for a group activity. We did not require a certified music therapist to provide the intervention because the exact qualifications, training, and experience were often unclear, and training programmes vary between countries. Moreover, the importance of requiring a qualification is unclear relative to the importance of having experience with the specific needs of people with dementia (e.g. a trained music therapist with no experience in comparison with a musician with years of experience in providing therapy to people with dementia).

Control

- In the Van der Steen 2017 update, we added: "Control groups could not receive any music-based therapeutic intervention (even if fewer sessions than the intervention group)."

Objectives, outcomes and terminology

The protocol formulated the objectives in terms of problems only ("to assess the effects of music therapy in the treatment of behavioural, social, cognitive and emotional problems of older people with dementia"), but we considered emotions and (social) behaviour to be broader than that. We adapted the objectives in the Van der Steen 2017 version to cover both the original aims and more positive outcomes, and to consider broader inclusion criteria for interventions that were not delivered by a music therapist (see below). In updates of the review, we modified the objectives to: "to assess the effects of music-based therapeutic interventions for people with dementia on emotional well-being including quality of life, mood disturbance or negative affect, behavioural problems, social behaviour and cognition, at the end of therapy and four or more weeks after the end of treatment."

The protocol referred to effects "in older people", but we did not have an exclusion criterion based on age. Therefore, in the Van der Steen 2017 update, we removed any reference to "older" people.

In line with the broader objectives, we adapted terminology for outcomes in the Van der Steen 2017 version. In the updates of the review, we consistently referred to: emotional well-being including quality of life; mood disturbance or negative affect (depression and anxiety); behavioural problems (behaviour overall and agitation or aggression); social behaviour; and cognition. We also searched for any (other) possible adverse effects.

Methods

Since the Van der Steen 2017 version, two rather than three review authors have independently assessed publications for eligibility. Further, we did not analyse by length of treatment, but we analysed end-of-treatment data, accepting variable durations and number of sessions as long as the outcomes were assessed after a minimum of five sessions. We also aimed at assessing long-term effects, analysing data about assessments at a minimum of four weeks after the end of treatment, in order to understand whether any effects are retained as a result of therapy.

In the Van der Steen 2017 version, we used more stringent criteria with respect to analyses referring to outcome assessments after a minimum of five sessions or analyses that included earlier assessments if there was evidence of no different effect over time.

In the Van der Steen 2017 version, we modified our risk of bias assessment. If no research protocol was available, we assessed the risk of reporting bias as unclear - or high if there were specific reasons - and revised our previous low ratings accordingly.

I N D E X   T E R M S

Medical Subject Headings (MeSH)

Aggression; Anxiety [therapy]; Cognition; *Dementia [psychology] [therapy]; Depression [therapy]; Emotions; *Music Therapy [methods]; Psychomotor Agitation [therapy]; *Quality of Life; *Randomized Controlled Trials as Topic; Social Behavior

MeSH check words

Aged; Humans