BMJ Open

Music Interventions for Dementia and Depression in ELderly care (MIDDEL): Protocol and statistical analysis plan for a multinational cluster-randomised trial

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2018-023436
Article Type:	Protocol
Date Submitted by the Author:	12-Apr-2018
Keywords:	Dementia < NEUROLOGY, Depression & mood disorders < PSYCHIATRY, Care homes, Music therapy, Music interventions, Old age psychiatry < PSYCHIATRY

Complete List of Authors:
Gold, Christian; Uni Research Health Bergen, Uni Research Health
Eickholt, Jasmin; Hochschule fur angewandte Wissenschaften Wurzburg-Schweinfurt
Assmus, Jörg; Uni Research Health Bergen, Uni Research Health
Stige, Brynjulf; Uni Research Health Bergen, Uni Research Health
Wake, Jo; Uni Research Health Bergen, Uni Research Health
Baker, Felicity; Melbourne Conservatorium of Music
Tamplin, Jeanette; Melbourne Conservatorium of Music
Clark, Imogen; Melbourne Conservatorium of Music
Lee, Young-Eun; Melbourne Conservatorium of Music
Jacobsen, Stine; Alborg Universitet Institut for Kommunikation
Ridder, Hanne Mette; Alborg Universitet Institut for Kommunikation
Kreutz, Gunter; Carl von Ossietzky Universitat Oldenburg Institut fur Musik
Muthesius, Dorothea; Universitat der Kunste
Wosch, Thomas; Hochschule fur angewandte Wissenschaften Wurzburg-Schweinfurt
Ceccato, Enrico; Universita degli Studi di Verona Dipartimento di Sanita Pubblica e Medicina di Comunita
Raglio, Alfredo; Universita degli Studi di Pavia Sezione di Medicina Legale e Scienze Forensi Antonio Fornari
Ruggeri, Mirella; University of Verona, Public Health and Community Medicine
Vink, Annemieke; University Medical Centre Groningen, dept. of General Practice
Zuidema, Sytse; University Medical Centre Groningen, dept. of General Practice
Odell-Miller, Helen; Anglia Ruskin University - Cambridge Campus
Orrell, Martin; Mental Health Sciences
Schneider, Justine; University of Nottingham, Sociology & Social Policy
Kubiak, Christine; ECRIN
Romeo, Renee; KCL,
Geretsegger, Monika; Uni Research Health Bergen, Uni Research Health

Music Interventions for Dementia and Depression in ELderly care (MIDDEL): Protocol and statistical analysis plan for a multinational cluster-randomised trial

Authors

Christian Gold, GAMUT – The Grieg Academy Music Therapy Research Centre, Uni Research Health, Uni Research AS, Bergen, Norway;

Jasmin Eickholt, Faculty of Applied Social Sciences, University of Applied Sciences Würzburg-Schweinfurt, Würzburg, Germany;

Jörg Assmus, Uni Research Health, Uni Research AS, Bergen, Norway;

Brynjulf Stige, GAMUT – The Grieg Academy Music Therapy Research Centre, Uni Research Health, Uni Research AS, Bergen, Norway;

Jo Dugstad Wake, Uni Research Health, Uni Research AS, Bergen, Norway;

Felicity Anne Baker, National Music Therapy Research Unit, Faculty of Fine Arts and Music, University of Melbourne, Melbourne, Australia

Jeanette Tamplin, National Music Therapy Research Unit, Faculty of Fine Arts and Music, University of Melbourne, Melbourne, Australia;

Imogen Clark, National Music Therapy Research Unit, Faculty of Fine Arts and Music, University of Melbourne, Melbourne, Australia;

Young-Eun Claire Lee, National Music Therapy Research Unit, Faculty of Fine Arts and Music, University of Melbourne, Melbourne, Australia;

Stine Lindahl Jacobsen, Department of Communication and Psychology, Aalborg University, Aalborg, Denmark;

Hanne Mette Ochsner Ridder, Department of Communication and Psychology, Aalborg University, Aalborg, Denmark;

Gunter Kreutz, Department of Music, Carl von Ossietzky University, Oldenburg, Germany;

Dorothea Muthesius, Berlin Career College, Universität der Künste, Berlin, Germany;

Thomas Wosch, Faculty of Applied Social Sciences, University of Applied Sciences Würzburg-Schweinfurt, Würzburg, Germany;

Enrico Ceccato, Mental Health Department, Hospital of Montecchio Maggiore, ULSS8 Berica, Vicenza, Italy;

Alfredo Raglio, Music Therapy Research Laboratory, Istituti Clinici Scientifici Maugeri, Pavia, Italy;

Mirella Ruggeri, Section of Psychiatry, University of Verona, Verona, Italy;

Annemiek Vink, Music therapy department, Academy of Music, ArtEZ University of the Arts, Enschede, The Netherlands;

Sytse Zuidema, Department of General Practice and Elderly Care Medicine, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands;

Helen Odell-Miller, CIMTR – The Cambridge Institute for Music Therapy Research at Anglia Ruskin University, Cambridge, UK;

Martin Orrell, School of Medicine and Institute for Mental Health, University of Nottingham, Nottingham, UK;

Justine Schneider, School of Sociology and Social Policy and Institute for Mental Health, University of Nottingham, Nottingham, UK;

Christine Kubiak, Management Office, European Clinical Research Infrastructure Network (ECRIN-ERIC), Paris, France;

Renee Romeo, Health Services and Population Research Department, King's College London, London, UK;

Monika Geretsegger, GAMUT – The Grieg Academy Music Therapy Research Centre, Uni Research Health, Uni Research AS, Bergen, Norway.

Address correspondence to: Christian Gold, GAMUT – The Grieg Academy Music Therapy Research Centre, Uni Research Health, Uni Research AS, Postbox 7805, 5020 Bergen, Norway. Email: christian.gold@uni.no. Phone: +47-97501757.

Acknowledgements: Stephen Clift, Evan Dawson, Jarle Flemvåg, Julene Johnson, Orii McDermott, Valentina Cabral Iversen, Inger Hilde Nordhus, Paolo Pizziolo, Sabine Ruths, Teppo Särkämö, Geir Selbæk, and Tanara Sousa provided valuable comments on an earlier version of this protocol. We would like to thank the following patient advisers for feedback on the study purpose and design: Betty Hamilton, a person with dementia living in a care home; Sue Mountain, her daughter; and various people from the Danish Alzheimer Association and Dementia UK. Feedback from the perspective of professional carers was received from Sarah Yeates and various other people at Caladenia Dementia Care Day Centre; Bupa; and BlueCross.

Autors' contributions:

CG took the initiative for the study; CG, JE, JA, BS, HMOR, SZ and RR developed the concept and design; FAB, JT, IC, YCL, GK, DM, TW, EC, AR, MR, AV, HOM, MO, JS, and MG helped to revise the concept and design. CG drafted the manuscript; JE, JA, RR, and MG helped drafting the manuscript. BS, JDW, FAB, JT, IC, YCL, SLJ, HMOR, GK, DM, TW, EC, AR, MR, AV, SZ, HOM, MO, JS, CK revised the manuscript for important intellectual content. CG, FAB, JT, IC, HMOR, and SZ obtained funding. BS, FAB, JT, IC, YCL, SLJ, HMOR, GK, DM, TW, EC, AR, MR, AV, SZ, HOM, and MG were involved in setting up the study conduct in each site. JDW, FAB, JT, IC, YCL, CK, and MG provided administrative, technical, or material support. All authors approved the final version of the manuscript.

Funding: This work was supported by the National Health and Medical Research Council (grant number APP1137853) and by intramural support from Uni Research AS and University Medical Center Groningen.

Competing interests: None declared.

Word count: 7632 (excluding title page, abstract, references, figures and tables)

Abstract

Introduction: In older adults, dementia and depression are associated with individual distress and high societal costs. Music interventions such as group music therapy (GMT) and recreational choir singing (RCS) have shown promising effects, but their comparative effectiveness across clinical subgroups is unknown. This trial aims to determine effectiveness of GMT, RCS, and their combination for care home residents and to examine heterogeneity of treatment effects across subgroups.

Methods and analysis: This large, pragmatic, multinational cluster-randomised controlled trial with a 2x2 factorial design will compare the effects of GMT, RCS, both, or neither, for care home residents aged 65 years or older with dementia and depressive symptoms. We will randomise 100 care home units with ≥ 1000 residents in total across 8 countries. Each intervention will be offered for 6 months (3 months 2x/week followed by 3 months 1x/week), with extension allowed if locally available. The primary outcome will be the change in the Montgomery-Åsberg Depression Rating Scale score at 6 months. Secondary outcomes will include depressive symptoms, cognitive functioning, neuropsychiatric symptoms, psychotropic drug use, caregiver burden, quality of life, mortality, and costs over at least 12 months. The study has 90% power to detect main effects and is also powered to determine interaction effects with gender, severity, and socio-economic status.

Ethics and dissemination: Ethical approval has been obtained for one country and will be obtained for all countries. Results will be presented at national and international conferences and published in scientific journals.

Trial registration numbers: NCT03496675, ACTRN12618000156280

Keywords: group music therapy, recreational choir singing, depression, dementia, non-pharmacological interventions, psychosocial interventions, randomised controlled trial

Strengths and limitations of this study

- As a multinational trial, this study will provide internationally generalisable results concerning the effects of music interventions in older adults with dementia and depression.
- Based on previous small-scale studies, this trial will have adequate power to determine clinical effects as well as to explain variation in treatment effects in relation to patient characteristics.
- A comprehensive set of core outcomes will be measured, including long-term effects in key variables, with assessor blinding where relevant.
- The trial will also enable modelling of trajectories of change and will thereby contribute to an improved understanding of the mechanisms of music interventions.
- Limitations include the potential bias inherent in cluster-randomised studies if recruitment within clusters is incomplete. Due to the nature of the intervention, care providers and participants cannot be blinded, which may bias measures that rely on their reports.

Glossary of terms

- Site: an organisational or geographical entity containing several units, for example a care home/residential care facility.
- Unit (or care home unit; also ‘cluster’): the smallest organisational unit within a site, where residents live together and are cared for together by staff; each unit is randomised.
- Participant: staff or residents within units who have consented to participate.

INTRODUCTION

Dementia and depression are highly prevalent and comorbid conditions in older adults and are associated with individual distress and high and rising societal costs. Globally, around 50 million people were living with dementia in 2017; this number is predicted to reach 82 million in 2030 and 152 million in 2050.[1] The societal costs of dementia are increasing from a total estimated worldwide amount of US$ 818 billion in 2015, about 1.1% of global gross domestic product,[1] to US$ 1 trillion in 2018.[2] Further, the disease's ramifications for families and carers are significant with respect to financial outlay and carer burden.[3] Dementia is highly prevalent among care home residents; more than half of all Australian care home residents in 2016-2017 had dementia.[4]

Depression is the leading cause of disability worldwide.[5] In older adults, it co-occurs and interacts with dementia in complex ways. Depression can cause cognitive impairment and may increase the risk of developing dementia;[6 7] conversely, depression is very common in the early stages of dementia[6] and often exacerbated by admission to a long-term care facility.[8] Psychotropic medication is only a second-line intervention due to limited efficacy and severe adverse effects, including increased mortality from antipsychotics,[9] but is in practice often used to reduce challenging behaviours in later stages of dementia. Non-pharmacological interventions are available and have some supporting evidence, but further research is needed.[10] Among the most promising non-pharmacological approaches to depression and dementia are music interventions, and in the following section we scope out this evidence.

Music interventions for older adults are based on the notion that music elicits emotional responses and helps to retrieve memories,[11] with recent support from research suggesting that brain regions responsible for processing music, particularly known familiar songs, may be spared even in late-stage dementia.[12 13] They are offered in individual,[14 15] group,[16 17], and community settings[18]

and range from targeted clinical interventions offered by trained music therapists to broader recreational activities, which may be facilitated by choir leaders or nursing staff. However, overlaps between the levels of targeting and training do exist. The most common group-based music interventions may be described as group music therapy (GMT) and recreational choir singing (RCS), where GMT is offered by a music therapist and may use a variety of activities ranging from singing through instrumental music making to music listening, whereas RCS is often facilitated by a choir leader and focuses centrally on singing. Both GMT and RCS rely on a combination of biological, psychological (cognitive and emotional), and social mechanisms (Figure 1, left part):

- Among the *psychological mechanisms*, *emotional processing*, such as using musical interactions to reflect on biographical or current relationships, may be most important in GMT, but is also present to some extent in RCS. *Cognitive processing*, for example through learning and memorising music pieces, is a central mechanism in RCS and less pronounced in GMT, although this may vary between cases, groups, or therapists.[16 19 20]
- *Social mechanisms* are important in both GMT and RCS. Meeting as a group may be important in itself. The function of the group in itself may be relatively more important in RCS, whereas GMT to a greater extent also relies on the one-to-one relationship between the therapist and each group member. Another important part of the social mechanisms is developing a shared sense of mastery and achievement through learning and performing music pieces, which is more central in RCS than in GMT but may again vary from case to case.[21]
- *Biological mechanisms* include the physical training effects of singing and other music-related activities, which may include movement. They are important in both GMT and RCS but may be more central in RCS.[18 22]

Systematic reviews of the clinical effects of GMT and RCS have reported mixed results, [10 17 19 20 23 24 25 26 27] possibly owing to the heterogeneity of treatment effects across types of participants and music interventions. One small trial comparing GMT and RCS directly suggested that the comparative effects of these music interventions may depend on the comorbidity of dementia and depressive symptoms.[16] Process-outcome relations of music interventions may be described as follows (Figure 1, right part):

- *Emotional processing* in a therapist-client relationship may lead to finding meaning and regaining orientation, and thereby to reduced *agitation* and related *neuropsychiatric symptoms*. Such effects have been suggested in some systematic reviews,[17 26] but not in others.[19] Reduced agitation may in consequence reduce *burden on staff*[15] and consequently reduce *sick leave*. This may also help to reduce inappropriate use of *medication*,[21] which is a concern in care homes.[28]

- *Cognitive processing* through practicing music may promote or maintain *cognitive functioning* in older adults. Such effects have been shown for active music therapy, but not music listening, for people with dementia.[20]

- *Emotional processing*, but also social, biological, and cognitive mechanisms may be associated with improved mood and reduced *depressive symptoms*. Systematic reviews have suggested effects of music therapy on depressive symptoms in older adults in general[29] and in people with dementia.[19]

- As downstream outcomes of all four mechanisms and of the intermediate outcomes above, one may expect improved *quality of life*, and possibly reduced *mortality*, although these effects may be small[30 31] and indirect.[9] Music interventions may also reduce costs by reducing time spent on treating and alleviating neuropsychiatric symptoms and reducing absence by staff.

Hypotheses

Through these different pathways of GMT and RCS, one may hypothesise differential effects for different outcomes, and therefore for different subgroups of care home residents. Specifically:

- GMT may be more effective than no GMT, and RCS may be more effective than no RCS, with respect to reducing depression symptoms and other outcomes shown in Figure 1.
- GMT and RCS may differ in the pattern of effects across outcome domains, which may be explained by their different mechanisms. For example, GMT may be more effective than RCS for reducing aggression and agitation and may therefore be more beneficial for people with late-stage dementia who often present with these neuropsychiatric symptoms.[32 33] RCS may be more effective than GMT with respect to cognitive functioning. Effects on depression symptoms may be achieved through different pathways (Figure 1), and the strength of those effects may therefore depend on severity or comorbidity.[16]
- When offered together, synergistic effects of GMT and RCS may occur through activation of different pathways.
- Cost-effectiveness may differ accordingly across interventions and subgroups. As RCS is likely to be associated with lower intervention costs, it may have better cost-effectiveness ratio in areas where clinical effects are similar; however, this will depend also on each intervention's effects on use of other treatments and services.

METHODS AND ANALYSIS

Design

This large, multinational cluster-randomised controlled trial will be conducted in care homes in Australia, Denmark, Germany, Italy, Netherlands, Norway, Poland, and the UK. The list of study

sites is provided in the trial registration record. We will use a 2x2 factorial design to examine the effects of GMT, RCS, both, or neither, for elderly care home residents with dementia and depressive symptoms (Figure 2). This design enables investigating the effects of two music interventions as well as potential synergy effects between them. These may occur between intervention providers on the cluster level (GMT and RCS providers learning from each other) and through residents on the individual or cluster level (participants gaining in different ways from the combination). We will randomise 100 or more care home units (clusters) in eight countries for a total of 1000 or more participants.

Block randomisation (block size 4 clusters) will be used to ensure that each site will have a balanced distribution between the interventions. The computer-generated randomisation list will be created and kept concealed at the central study office. Only after the eligibility of a care home unit is confirmed and eligible participants (residents and staff) within that unit have formally consented and completed baseline assessment, will site investigators be informed of the randomisation result for that unit. Where possible, a number of care home units will be randomised at the same time, which will further ensure allocation concealment.

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 and RCS. Plain language summaries and consent forms will use neutral wording to maintain equipoise and to avoid expectancy effects. Blinding of assessors 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 well) 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.

GMT and RCS may entail "ripple effects" beyond the individual participants by leading to changes of the local milieu/culture at the care home unit. These will be assessed by measuring objective and perceived burden on care staff. The cluster design is ideally suited for that situation because it facilitates application in a naturalistic setting and avoids some of the problems of individually randomised trials (such as treatment contamination); it also minimises the additional workload for care staff. Trial procedures will be tested in the Australian cohort before applying them in the other countries. The trial will be conducted and reported in accordance with relevant legal frameworks and research guidelines.[34 35 36 37]

Participants

Eligibility is defined on two levels, care home units and individual participants. Participating care home units will be those that are expected to have at least 10 eligible and consenting residents. Care home units that are currently providing music-based interventions as part of their usual care programme will be excluded. Eligible participants will meet all of the following inclusion criteria:

- aged 65 years or older, resident (full-time, 24h/day) at a participating care home;
- dementia as indicated by a Clinical Dementia Rating score of 0.5 to 2 and a Mini-Mental State Examination (MMSE) score of 26 or less;
- at least mild depressive symptoms, as indicated by a Montgomery-Åsberg Depression Rating Scale (MADRS) score of at least 8;
- a clinical diagnosis of dementia according to ICD-10 research criteria;
- have given written informed consent (may be assent by proxy for those unable to provide consent themselves).

Clinical diagnosis will be ascertained by a clinician or researcher, based on the ICD-10 dementia criteria of memory decline; decline in other cognitive abilities; impairment in activities of daily

living; preserved awareness of the environment; decline in emotional control or motivation or change in social behaviour; and more than 6-month duration of memory decline and other cognitive symptoms.[38] People with a known diagnosis of schizophrenia or Parkinson's disease or those who are known to be severely hearing-impaired, in short-term care, or unable to tolerate sitting in a chair for at least part of the sessions, will be excluded. People may however have other clinical diagnoses such as pre-morbid substance use disorders or anxiety disorders. The list of exclusion criteria is intentionally short to ensure generalisability.[37]

Interventions

General aspects. Units in all intervention arms will continue with standard care as locally available. In the units allocated to music interventions, GMT, RCS, or both will be provided twice weekly for the first three months, followed by weekly sessions for the next three months. Continuation of GMT and RCS is allowed after that period, depending on local availability. GMT and RCS sessions will be 45 minutes each. In line with usual practice, GMT may be divided into smaller groups (e.g. around 5 participants, but this may differ across local contexts), whereas RCS may be conducted in larger groups (e.g. with all residents of the unit in one group). GMT and RCS providers will receive training and implement intervention guidelines developed in the initial phase of the study. Regular exchange and peer supervision for GMT and RCS providers will be organized in conjunction with guidelines and training. This will include monthly online or in-person meetings between researchers and intervention providers to ensure intervention quality and fidelity, to discuss potential threats that might undermine study quality, and to refine the guidelines accordingly. Intervention providers will also attend weekly staff meetings at intervention sites where possible, to maximise local knowledge transfer and benefit. Data on the resources related to the interventions will be measured (number of sessions attended by each participant, duration of each session, non-contact time spent by the intervention provider to prepare or follow up a session, recorded by the provider). The components of standard care provided will also be recorded (see Outcomes).

GMT. The core principle of GMT is affect regulation through active, reciprocal music making with the use of singing and musical instruments. This facilitates the relationship between the music therapist and the person living with dementia, and between participants in the group. The approach takes into account the level of dementia severity and symptoms that can vary from resident to resident and from session to session. A core intention of GMT is to meet the psychosocial needs of each individual resident, which in turn is thought to reduce depressive symptoms and anxiety and to stimulate overall social and emotional wellbeing.[39 40 41] GMT aims to work in the "here and now" by responding to participants' immediate emotional expressions, containing them, and incorporating them into meaningful musical expressions for therapeutic gain.[21] GMT is provided by a trained music therapist, who is registered with the appropriate professional association or registration body in his or her country and should also be skilled as a musician. To facilitate individual relationship-building, the music therapist will offer each resident an initial 20-minute assessment with the aim of determining their musical preferences and starting to build individual rapport. The music therapist will also use other sources to determine the participants' musical biography, cultural background, history, personal strengths, resources, and disabilities, and any other information that could be useful to bring into GMT sessions.

RCS. The core principle of RCS is to sing familiar songs and to provide a familiar musical environment for participants. Choral singing involves a combination of cognitive, physical, and psychosocial engagement components.[42] Drawing on the psychosocial aspects of a choir setting, RCS in this trial aims to foster connectedness in a group either with other older adults residing in the care homes or family caregivers; emotional wellbeing; and enjoyment of music-making in a group. Where participants have engaged in music activities in their past, this may also enable the continuation, as far as possible, of the familiar social experience of music-making in everyday life. RCS in health settings can be characterised as a ritualistic musical practice, in which biographically and culturally grounded song materials are used with the central goal of stimulating positive

experiences shared by groups of individuals. Sections of RCS sessions may vary in their focus; for example, sessions may focus on developing familiarity with well-known songs; learning and developing new material as a group; singing rounds to encourage listening to each other; or offering space for solo singing.[43] The materials can be familiar songs from a range of repertoires, including but not restricted to festive songs (e.g. birthday songs, Christmas carols), folk songs, traditional, classical, or popular songs. The selection of songs can vary from country to country, within and between choir leaders, and may also depend on seasonal and other circumstantial factors. RCS may be specifically useful in the context of mild to moderately severe dementia, as vocal singing expressivity is often spared in the presence of cognitive decline. RCS is provided by a skilled musician with choir leading skills.

Training and assessment of treatment fidelity. A challenge in designing manuals for complex interventions such as GMT and RCS is to standardize the quality of interventions to avoid unwarranted variation between therapists and countries while preserving the possibility for meaningful tailoring to local contexts and individuals.[44] This will be addressed by focusing on general principles rather than fixed behaviours. One of the strengths of these interventions is that they can be applied in a way that is tailored to fit the current situation/status of the group and its individual members, and the therapist is able to adapt the therapy. A tentative comparison showing different and similar principles of GMT and RCS is shown in Table 1. Interventionists will be trained at all sites, both through local in-person meetings with all intervention providers at each site and through remote online training across sites. The purpose of this training is to supplement rather than replace the existing training and expertise of intervention providers. For assessment of adherence and competence, providers of GMT and RCS will be video-recorded in 3-4 randomly selected sessions per unit. We will record and analyse the entire session. To avoid performance bias due to the awareness of being videotaped in a selected session, we will use sham video monitoring in other sessions where possible. Videos will be uploaded and stored on a secure central server and

will be available only to those who check treatment fidelity. Two independent researchers will assess the different components used by intervention providers and the degree of person-centeredness (i.e. tailoring of the intervention to the current situation/needs of the group and its members). This process-related data will help us to understand the mechanisms or effective ingredients of each intervention.

Further development. While the description above provides general guidance and will form the basis for fidelity assessment in this study, no consensus guidelines exist for GMT and RCS. Descriptions in the literature vary in many aspects such as: theoretical frame; session structure; specific therapeutic goals; types of musical instruments and materials; inclusion of music listening in addition to active music-making; structured versus improvisational techniques in active music-making; and adaptation/tailoring to reach each person individually. Therefore, flexible manuals, including sets of detailed principles and techniques for GMT and RCS, will be developed and agreed upon by scientific and clinical experts from different countries using a modified Delphi consensus procedure.

Outcomes

The study uses a broad array of resident-, staff-, and unit-level outcomes measured at 3 months, 6 months (primary), and 12 months after randomisation (Figure 3). A long-term extension with later follow-ups is planned but will be beyond the present project. As described above, blinding of assessors will be attempted but may be incomplete; most residents will be unable to self-rate their status so that assessors will have to rely on staff who know the resident well as proxy informants. Where possible, core outcomes (www.comet-initiative.org) for psychosocial intervention research in dementia care were selected.[45]

The primary endpoint will be change in the total score of the Montgomery-Åsberg Depression Rating Scale (MADRS). The MADRS is a 10-item scale where each item is rated from 0 (no

abnormality to 6 (severe).[46] In the total sum score ranging from 0 to 60, higher scores indicate higher severity of depressive symptoms. Assessment is based on an interview with the resident where possible, but where definite answers cannot be elicited from them, all relevant clues as well as information from other sources should be used as a basis for the rating, in line with usual clinical practice.[47] The total time of administration is approximately 20 minutes. The MADRS has been used successfully in previous studies of music interventions[16 48] and has shown higher sensitivity to change in this population than other scales evaluating depression severity, such as the Cornell Scale for Depression in Dementia (CSDD).[49]

Secondary outcomes will include the following:

- Dementia severity including cognitive and functional performance – Clinical Dementia Rating (CDR), a standard assessment of dementia severity.[50] The CDR is used widely in clinical settings. Its score is derived from a semi-structured interview with the person living with dementia and an appropriate caregiver/relative. It rates impairment in each of 6 cognitive categories (memory, orientation, judgment and problem solving, community affairs, home and hobbies, and personal care). Its score is useful for characterising and tracking a person's level of impairment or dementia: 0 = normal; 0.5 = very mild or questionable dementia; 1 = mild dementia; 2 = moderate dementia; 3 = severe dementia.

- Neuropsychiatric symptoms – Neuropsychiatric Inventory (NPI), "a *de facto* standard for measuring neuropsychiatric symptoms in clinical trials".[45] Developed to assess behaviour in people living with dementia, the NPI has substantial evidence of validity and reliability and has been translated into more than 40 languages.[51 52] The NPI uses a screening approach to minimise administration time, examining and scoring only the domains with positive responses to screening questions. In this study, the NPI – Questionnaire (NPI-Q)[53] will be used; another version specific for nursing homes (NPI-NH) was considered but rejected because it is not

available across all languages. The NPI-Q includes 12 domains where if a symptom is present, both its severity (from 1= mild to 3=severe) and the associated distress on caregivers (from 0=Not distressing at all to 5=Extreme or very severe) are assessed by the professional carer who is most familiar with the resident's behaviour. Item scores across the 12 domains are summed, leading to a total severity score from 0 to 36, where higher values represent higher severity. The additional total distress score can range from 0 to 60, also with higher values representing higher distress.[53]

- Generic quality of life – EuroQol (EQ-5D-5L), a generic health utility measure.[45] The standardized, non-disease-specific instrument for evaluating health-related quality of life was developed by the international EuroQol group and is used to derive quality-adjusted life-years (QALYs). It is based on a descriptive system that defines health in the five dimensions mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Each dimension has five response categories from "no problems" to "extreme problems", which are combined using preference weights to form an overall quality of life score ranging from lower than 0 (worse than death) to 1 (best possible). An additional visual analogue scale indicates today's health on a scale from 0 ("The worst health you can imagine") to 100 ("The best health you can imagine"). As most residents will be unable to self-rate the EQ-5D-5L, the rating will rely on the judgment of the carer as a proxy. Careful selection of assessment mode (self/proxy/both) and choice of appropriate proxies is important to ensure the measure's validity in studies of people with dementia.[54]

- Disease-specific quality of life – Quality of Life in Alzheimer's Dementia (QOL-AD).[45 55] This 13-item scale with a self-rating and proxy version has demonstrated sensitivity to psychosocial intervention, correlates with health-utility measures, is widely translated and used internationally and can be used by people with very low MMSE scores. Items such as "Physical

health", "Memory", or "Ability to do things for fun" are scored on a scale from 1 (poor) to 4 (excellent), resulting in a total score ranging from 13 (worst) to 52 (best).

- All-cause mortality (time to death), as recorded in official electronic registries.
- Any increase in psychotropic drug use: Data on type (ATC Codes N065, N06) of psychotropic medication used and any increase or decrease over time will be collected from care staff using the 'medication profile' section of a tailored version of the Client Socio-Demographic and Service Receipt Inventory (CSSRI).[56] Available electronic health registry data will be used where possible. Psychotropic medications are sometimes used inappropriately to manage behavioural symptoms of dementia.[57 58 59] An earlier study suggested that music therapy may help prevent increase in medication.[14]
- Costs: Total and component costs of the interventions will be assessed from a societal perspective, including the cost of the intervention as well as statutory health and social care services used, using a tailored version of the Client Socio-Demographic and Service Receipt Inventory (CSSRI).[56]
- Any adverse events (safety): No adverse effects of music interventions are known from earlier trials. Intervention providers are trained to work closely with and adapt their interventions to the needs of participants in order to avoid adverse reactions. Because little knowledge exists about what the potential adverse events could be, all types of adverse events and serious adverse events (e.g. unexpected worsening of symptoms), whether related or unrelated to the interventions, will be reported.

Staff-level outcomes will be as follows:

- Subjective perceived burden of care staff: Professional Care Team Burden Scale.[60] The 10-item scale provides a valid and reliable means of obtaining ratings of burden from formal care teams working in care homes in order to evaluate different interventions targeted at the reduction of

burden in care teams. Items are scored on a 5-point scale from 0 (strongly disagree) to 4 (strongly agree), yielding a total sum score from 0 to 40, with higher scores indicating higher burden.

- Days on sick leave of care staff, as recorded monthly by the employer.

Sample size and test power

There is no consensus on the minimal clinically important difference (MCID)[61] on the MADRS. Generally, effect sizes in the small- to medium range (i.e. between d = 0.20 and 0.50) may be considered relevant.[62] Effect sizes in that range were also found in a previous trial on GMT and RCS (d = 0.33 at 6 weeks and 0.49 at 12 weeks).[16] Studies of other depression scales have used anchor-based approaches to determine clinically important percent reductions;[63] we will not use such approaches for the primary analyses, but will include an additional responder analysis.[61]

The trial has the multiple aim of identifying main effects of GMT versus no GMT and RCS versus no RCS, interaction effects of GMT and RCS, and predictive effects of clinical characteristics including severity of dementia; severity of depression; gender; and socio-economic differences. (Although individual socio-economic differences tend to become more equal amongst residents in a given care home, they may still exist at the cluster level, as different homes may have different standards; we will use the average cost of living in each care home unit as a cluster-based proxy measure for socio-economic status.) Power for interaction effects and subgroup analyses is difficult to determine because of the unknown distributions and effect sizes of the different variables. Therefore, the power calculation for the primary outcome was based on the main comparisons of GMT versus no GMT and RCS versus no RCS. This approach maximises power by fully exploiting the factorial design. A general two-sided significance level of 5% will be used, leading with Bonferroni adjustment to a marginal two-sided level of 2.5%. The power calculation was adjusted for cluster effects using the intraclass correlation coefficient (ICC, between 0.01 and 0.10, Figure

4), assuming average cluster size 10. It was further assumed that attrition, which may occur due to death, moving to another care home, or withdrawal from the study, will be no higher than 20% overall. With 100 clusters and 1000 participants randomised, 90% power is reached for effect sizes between 0.25 and 0.35 (Figure 4). Any further increase beyond this sample size will serve heterogeneity of treatment effects analyses.

Statistical analyses

The statistical analyses will use multivariate longitudinal statistical models, which make optimal use of the data by using data from all time points at once and can account for the effects of clustering within care home units and sites. We will use a modified intention-to-treat (ITT) approach using all available data from all participants as randomised, regardless of the intervention actually received. Sensitivity analyses using multiple imputation for missing data will enable a full ITT analysis. Additional per-protocol analysis will address the effects of treatments as actually received and will complement the ITT analyses. All tests in the study will be two-sided. The general significance level is set to 0.05. Since there are two comparisons in the primary analysis (GMT vs. no GMT, RCS vs. no RCS), we will use a marginal Bonferroni level of 0.025. Continuous variables will be screened for normality. All computations will be done using R.[64]

Sociodemographic and clinical baseline properties for the groups will be characterized by descriptive methods (mean (SD), median [range], n (%)) and presented in a table. A similar table will compare those who dropped out versus those who completed the primary outcome.

The primary outcome, change of MADRS score from baseline to 6 months, will be assessed by a linear mixed-effects model (LME).[65] We will fit the unadjusted model for each treatment (RCS vs. no RCS) as well as the multivariate model containing both treatments as predictors both unadjusted and adjusted for the interaction between the treatments.

Secondary analyses of MADRS scores will include the following:

- The development of MADRS in the treatment groups over the entire study period will be assessed by a LME including time, treatment type and the interaction of time and treatment type as fixed effects, and participant nested within cluster as random effects. We will use both linear and simple contrasts in the time domain because it is not known whether there is a linear association in time. This will be illustrated by a figure showing the predicted mean of MADRS for each treatment type at each time point with confidence intervals.
- The synergy of the two treatments will be assessed by the LME containing both treatments as well as their interaction as predictors. The interaction in the model will estimate the synergy effect.
- The predictive effect of several covariates (severity of dementia; severity of depression; gender; and socio-economic differences) will be assessed as odds ratios using LMEs for each covariate containing time, treatment type and their interaction as well as the covariate and the interaction between the covariate and the treatment type as predictors. The interaction in the model will estimate the predictive effect.

Secondary endpoints will be analysed as for the primary analysis, using LMEs for continuous outcomes (both resident-level and staff-level). Special considerations apply for the following variables:

- Binary outcomes, including response rates (the proportion of residents improved by at least 50% from their baseline MADRS score), prevalence of medication use, and adverse events, will be assessed as odds ratios using generalised linear mixed models (GLMMs) with a logit link function. The predictive effect of several covariates (severity of dementia; severity of depression; gender; and socio-economic differences) will be assessed by GLMMs for each covariate containing time, treatment type and their interaction as well as the covariate and

the interaction between the covariate and the treatment type as predictors. The interaction in the model will estimate the predictive effect.

- Count data (days of sick leave) and cost data are more likely to follow a Poisson distribution than a normal distribution and will be analysed using the respective GLMMs.
- Time-to-event data include mortality (time to death of any cause) and will be assessed by Kaplan-Meier and log-rank- or Breslow tests for differences between the treatment types and the hazard ratios at 12 months.
- Loss to follow-up in all other outcomes can be influenced by mortality. Thus, if the survival analysis shows differences between the groups, it will be meaningful to use a joint modelling approach which combines the longitudinal models and the survival analysis.[66]

In addition to analysing effects of interventions as randomised, we will conduct mediator analyses to examine relations between elements of the therapy approach (mechanisms), direct and downstream outcomes, as depicted in Figure 1, using structural equation modelling (SEM).

Cost-effectiveness analysis

Total and component costs of the interventions and the cost-effectiveness of alternative interventions will be assessed from a societal perspective. This perspective will cover three categories of costs: the cost of the intervention, statutory health and social care (and voluntary sector) service costs, and costs of unpaid carer support. The cost per session for each of the interventions will be derived employing established approaches used in a compendium of costs and in published studies.[67 68 69] Information on the time inputs by GMT and RCS providers (for running sessions and for other activities) will be obtained and valued using information on the midpoint of the salary scale and employer's national insurance as well as superannuation contributions. The sum of the staffing contributions and allocations for overheads for each session will then be summed, to derive a cost per session. To this cost per session, the average number of sessions delivered as part

of the intervention will be multiplied to derive a cost per intervention. As there is no clear agreement on how the costs of group interventions should be allocated, we will calculate the cost per session of each of the interventions on the basis of the participants allocated to each of the groups, regardless of whether or not the participant attended, because participants who miss a session are not replaced.

Data on statutory services used will be collected using a tailored version of the Client Socio-Demographic and Service Receipt Inventory (CSSRI),[56] which contains data on the use of health and other formal care resources and unpaid care. To service and support data we will attach unit costs reflecting the long-run marginal opportunity costs drawn from available public sources. Costs per unit of measurement for each service type will be taken from country-specific sources. We will adjust country-specific costs to Euros using purchasing power parity methods. Costs and outcomes will be compared for the comparators using extended dominance approaches. In this approach, the four treatment combinations (GMT, RCS, GMT and RCS, no GMT or RCS) will be ranked by cost, and if one is dominated (more expensive and less effective than another), it will be excluded from further analysis, until two therapeutic groups are left on which to explore which of the two groups is most cost-effective. The cost-effectiveness of one arm over another will be compared by calculating incremental cost-effectiveness ratios (ICERs) defined as difference in mean costs (Euros spent) divided by difference in mean effects (QALYs using the EQ-5D-5L; points improved on MADRS and QOL-AD). Cost-effectiveness acceptability curves will be plotted for each cost-outcome combination to show the likelihood of one treatment being seen as cost-effective relative to another for a range of values placed on incremental outcome improvements. Using the net benefit approach, monetary values of incremental effects and incremental costs will be combined, and net benefit (NB) derived as: $NB = \lambda * (effect_b - effect_a) – (cost_b – cost_a)$, where λ is the willingness-to-pay for a unit improvement in effectiveness, and subscript ‘a’ and ‘b’ denote two candidate treatment arms. This approach allows costs and outcomes to be considered on the same monetary scale, taking

account of sampling uncertainty and adjusting for baseline covariates and clustering. The cost-effectiveness threshold which represents a society's willingness to pay for an additional unit of outcome is used to determine if an intervention is cost-effective. However, this is problematic in multinational trials as there is no agreed cross-national threshold, and in some countries there is no established threshold at all. Other studies have used a threshold of Euros 50,000 per QALY, and we will consider this in the discussion of the results.

Sensitivity analyses will be conducted to assess the robustness of the results to changes in key parameters. One of the possible concerns is likely to be the sample size. If the sample size in some participating countries is too small, their cost-effectiveness estimates are likely to be unreliable. We shall therefore consider the added value of pooling the information on costs and outcomes in sensitivity analyses.

Patient and public involvement

The development of the research question and study design was informed by the priorities, experience and preferences of residents and carers. Co-authors in Australia, Denmark, and the UK have been actively involved in user and advocacy organisations in their countries for a long time and have discussed interventions, outcomes and the need for research with them. Relatives and caregivers spoke to the importance of music interventions as a help for carers and people with dementia, and to the need for high-quality evidence on their effects. Relatives and caregivers are important for giving persons with dementia a voice when they cannot speak for themselves.

Co-authors in Australia had significant involvement with residents, as well as with care staff and care home managers, in discussing and piloting aspects of the study design. While the interventions were generally perceived as pleasurable rather than burdening, some of the outcome measures were felt to be burdening and too demanding due to their length or complexity. As a consequence, the longer Cornell Scale for Depression in Dementia was replaced with the shorter MADRS, and a

more extensive quality of life scale was removed. Recruitment strategies were discussed and adapted in dialogue with care home staff.

User representatives will continue to be actively involved throughout the conduct of the trial (see next section). Results will be disseminated to residents, relatives, and care staff via care homes. Results will also be disseminated to national user and advocacy organisations.

Monitoring and oversight

One representative of each recruiting institution will be a member in the *Trial Steering Committee (TSC)*. They will be supplemented by other members who are independent of the investigators, their organisations, funders and sponsors. The TSC will include service users or their relatives and representatives of stakeholder organisations such as Alzheimer Europe and Dementia Australia. The TSC will have regular meetings to closely supervise all aspects of the study, including any protocol amendments, progress of recruitment, and publication plan.

Data quality monitoring will require a risk-based monitoring approach including remote monitoring activities performed centrally and on-site monitoring as needed. The monitoring will be performed according to the monitoring manual to be developed at the beginning of the project. Recruitment and retention rates will be monitored closely to mitigate the risk of slow recruitment. The number of participating care home units in total and in relation to care home units screened; the number of participating care home residents in total and in relation to residents screened of potential participants; and the retention of participants in the study over the trial period will be closely monitored.

A *Data and Safety Monitoring Committee (DSMC)*, consisting of three people with strong methodological and clinical expertise who are not otherwise affiliated with the project or its institutions, will be appointed early in the international trial. The DSMC will receive regular updates on recruitment, uptake of interventions, any unforeseen events, adverse events, and

immediate information on serious adverse events from the trial statistician. It will have unblinded access to study data. Meetings with the DSMC will be on a biannual basis and will consist of an open and a closed part. In the open part, the general progress of the trial will be discussed; in the closed part, the DSMC will discuss any safety signals with the trial statistician. If issues arise, the DSMC will recommend to the TSC on appropriate action.

All aspects of the study, from intervention fidelity through recruitment, outcome assessment, database and data quality management, to data and safety monitoring, will be pilot-tested in one country (Australia) before being rolled out internationally. The data of the pilot cohort will be included in the main trial; no statistical adjustments are made because the decision depends only on feasibility, not on an interim efficacy analysis. Patient-related documents such as the consent form will be tested because they may influence how the study is perceived by potential participants, relatives and staff.

To ensure data quality, a trial database will be set up and maintained using a safe server hosted by Uni Research (UHEADS) and OpenClinica software. UHEADS is a system for safely storing health research data developed by Uni Research AS, that accommodates the safe upload, storage and retrieval of any sensitive research data. OpenClinica is a web-based system for electronic data capture and clinical data management for multicenter clinical trials, which conforms to relevant international standards for health research. Uni Research AS runs an open source version of OpenClinica.

ETHICS AND DISSEMINATION

Ethical aspects

Ethical approval has been obtained from the Medicine and Dentistry Human Ethics Sub-Committee at the University of Melbourne, Australia (approval date: January 12, 2018) and will be obtained

from the relevant local institutional human research ethics committee at each international site. Local clinical investigators will work on adaptation of study- and patient-related documentation to meet national ethical requirements.

Risk management

Although this is a project with ambitious hypotheses and goals, the known risks to its implementation and completion are manageable. The main risks are: slow recruitment; low fidelity of interventions; and low reliability of outcome measurements. To mitigate the first risk, we will rely on clinical investigators with a track record of successful recruitment as well as relevant experience and expertise. Slow recruitment at some sites can be compensated by other sites. Regarding fidelity of interventions, we will develop clear guidance, selection, and ongoing monitoring of fidelity, as described above. To prevent low reliability of outcomes, we have chosen a range of outcome measures. Most have been used extensively and will be familiar in most countries, taking account of recommended core outcome sets. We will rely on highly qualified staff. We will also conduct tests of inter-rater reliability early in the process to identify and correct any potential problems.

Publication plan

The report on the main, pre-planned analyses of the primary endpoint and up until the 12-month follow-up will be submitted to a leading medical journal. The report on the long-term extension will also be submitted to a leading medical journal.

Further publications may focus on the following additional aspects and results:

- Recruitment and retention strategies and recommendations in international, cluster-randomised multicentre trials of complex interventions in non-medical settings.
- Development of an MCID for the MADRS based on an existing anchor question.
- Inter-relations between outcomes and predictive value of early outcomes for later outcomes

- A comprehensive, integrated systematic review of effects of music interventions addressing heterogeneity of treatment effects systematically, using meta-regression and/or individual participant data meta-analysis methods.
- Clinical descriptions and qualitative research of therapy processes, for example comparing successful with less successful clinical strategies, interactions between clinical characteristics and effective strategies, or qualitative influences on care home staff, their perception of GMT and RCS and how "ripple effects" may influence the overall atmosphere in the home.
- An analysis of barriers and facilitators for implementation, using a combination of qualitative interviews and surveys with questionnaires of known barriers and facilitators.
- Mixed-methods research combining qualitative and quantitative analyses to gain further insights on process-outcome relations and heterogeneity of treatment effects.
- A consensus guideline to indicate the elements of GMT and RCS.
- Translations of key instruments into different languages.

Data deposition and curation: The data and meta-data will be stored in a public repository, such as that of the Norwegian Centre for Research Data (NSD).

Relevance and benefit to society

Music interventions are widely used in care homes for both therapeutic and recreational purposes. Although smaller trials have shown promising results, systematic reviews and one head-to-head trial have suggested heterogeneity of treatment effects. This large multinational trial will provide reliable and broadly generalisable knowledge about the effectiveness, mechanisms and heterogeneity of effects of music interventions. It is designed to overcome limitations of previous studies in this field, including small sample sizes, limited numbers of institutions and therapists, and lack of standard care control. By including longer intervention and follow-up periods, it will also

fill knowledge gaps about potential long-term benefits and preconditions for achieving such sustained benefits. The results will drive changes in aged care policy and practice and will contribute to our understanding of the relation between music and health more generally.

Implications for practice

In terms of implementation, GMT and RCS differ with respect to scalability. GMT requires extensive, specialised music therapy training and is typically provided in small groups to facilitate interaction in a flexible approach. The number of qualified music therapists varies from country to country, but fluctuates around 1 in 100 000 (about 6000 in Europe, http://emtc-eu.com; 5000 in the USA, www.cbmt.org; 500 in Australia, www.austmta.org.au). In the UK, a 2017 survey found that, of 900 qualified music therapists, fewer than 200 are working with older people; most work with children and adolescents. Thus, there are insufficient numbers of trained music therapists to accommodate the needs of all older people who might benefit . However, not all older adults with dementia and depressive symptoms may need the clinical expertise of trained music therapists. RCS is more easily scalable as it can be provided by trained musicians and also in larger groups. With about 1 million choirs and 37 million choir singers in Europe (www.singingeurope.org), choral singing is one of the most popular arts activities. It is therefore a promising approach to meeting the needs of many older people, subject to the findings of the present study.

As complex interventions, both GMT and RCS are applied in a range of care homes with different methods and probably different effects. To ensure best practice, approaches for improvement and standardisation are embedded into the interventions under investigation. Internationally applicable guidelines with an international consensus will be defined, applied further developed during the trial through process evaluation to improve both interventions.

Based on previous studies, one may expect this pragmatic trial to confirm clinically relevant effects of both GMT and RCS, but with strong heterogeneity of these effects depending on clinical

characteristics. The highly person-centred approach of GMT may be most beneficial to those with neuropsychiatric symptoms, which are typical at late-stage dementia. In contrast, social engagement is more emphasised in RCS and may help those at earlier stages. The combination of both GMT and RCS may be best for another subset of residents who need both the social cohesion of a group and the opportunity to cope with disease, conflicts and untreated trauma in GMT. This knowledge will increase the impact of music interventions in care homes with a new best practice model on the expected largest effect of GMT or RCS, and will inform and provide improved education of future music therapists and choir leaders working with older adults in care homes and related contexts, such as day care centres for people still living at home.

Implications for future research

As a strongly interdisciplinary project building on contributions from medicine, social sciences, and humanities, this trial will contribute to strengthening the collaborations between these fields, which is likely to stimulate new cross-disciplinary investigations. The knowledge constructed from this trial will inform GMT and RCS service providers of the contextual factors and conditions that support effectiveness. The recommendations derived from our project will inform intervention delivery across countries, leading to increased safety, effectiveness and cost-effectiveness, and improved quality of life for care home residents. The study is unique in that it examines the interaction of depressive symptoms, cognitive impairment, and dementia in an international sample of participants. A critical feature of MIDDEL is its attention to interventions as applied within different health systems. Results will be valid internationally and will contribute to establishing a model for future research within different health systems.

In conclusion, MIDDEL will provide essential knowledge that will inform treatment guidelines aimed at improving the lives of the rapidly rising number of people living with dementia across countries. Building on previous small-scale randomised controlled trials, this large pragmatic

effectiveness trial will enhance the use of health technology assessment methodology in the area of non-pharmacological (psychosocial) interventions in this area. By improving existing interventions and providing evidence-based guidance on their application or discontinuation, it is anticipated to have a significant positive impact on people living with dementia, their caregivers, and the health system. Furthermore, it will also open several new lines of research and development of personalised psychosocial interventions in an area of high and rising public health relevance.

REFERENCES

1. World Health Organization (WHO). Dementia fact sheet 2017 [internet]. Geneva: WHO; 2017. Available from: http://www.who.int/mediacentre/factsheets/fs362/en/
2. Alzheimer's Disease International. World Alzheimer report 2016. London: Alzheimer's Disease International; 2016.
3. Australian Institute of Health and Welfare (AIHW). Dementia in Australia. Canberra: AIHW; 2012.
4. National Centre for Social and Economic Modelling. The economic cost of dementia in Australia 2016-2056. Canberra: University of Canberra; 2017.
5. World Health Organization (WHO). Depression fact sheet 2018 [internet]. Geneva: WHO; 2018. Available from: http://www.who.int/mediacentre/factsheets/fs369/en/
6. Byers AL, Yaffe K. Depression and risk of developing dementia. *Nat Rev Neurol* 2011;7:323–31.
7. Diniz BS, Butters MA, Albert SM, Dew MA, Reynolds CF. Late-life depression and risk of vascular dementia and Alzheimer's disease: systematic review and meta-analysis of community-based cohort studies. *Br J Psychiatry* 2013;202:329–35.
8. Koritsas S, Davidson S, Clarke D, O'Connor D. Diagnosing and treating depressions in nursing home residents: challenges for GPs. *Aust J Prim Health* 2006;12:104–8.
9. Maust DT, Kim HM, Seyfried LS, *et al.* Antipsychotics, other psychotropics, and the risk of death in patients with dementia number needed to harm. *JAMA Psychiatry* 2015;72(5):438–45.
10. National Institute for Health and Care Excellence (NICE). Clinical guideline 42 – Dementia: supporting people with dementia and their carers in health and social care. London, UK: NICE; 2016. Available from: https://www.nice.org.uk/guidance/cg42
11. Koelsch S. Brain correlates of music-evoked emotions. *Nat Rev Neurosci* 2014;15:170–80.
12. Jacobsen J, Stelzer J, Fritz TH, Chetelat G, La Joie R, Turner R. (2015). Why musical memory can be preserved in advanced Alzheimer's disease. *Brain* 2015;138:2438–50.
13. Janata P. The neural architecture of music-evoked autobiographical memories. *Cereb Cortex* 2009;19:2579–94.
14. Ridder H-M, Stige B, Qvale LG, Gold C. Individual music therapy for agitation in dementia: an exploratory randomized controlled trial. *Aging Ment Health* 2013;17:667–78.
15. 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 Geriatr* 2015;15:84.
16. 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 Ment Health* 2017;21:147–55.
17. Pedersen SKA, Andersen PN, Lugo RG, Andreassen M, Sutterlin S. Effects of music on agitation in dementia: a meta-analysis. *Front Psychol* 2017;8.
18. Stige B. Practicing music as mutual care. In: Stige B, Ansdell G, Elefant C, Pavlicevic M, editors. Where music helps. Community music therapy in action and reflection. Farnham, UK: Ashgate; 2010. p. 254–75.
19. 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 Syst Rev* 2017;5:CD003477.
20. Fusar-Poli L, Bieleninik L, Brondino N, Chen XJ, Gold C. The effect of music therapy on cognitive functions in patients with dementia: a systematic review and meta-analysis. *Aging Ment Health* 2017:1–10.

21. Ridder H-M. How can singing in music therapy influence social engagement for people with dementia: insights from polyvagal theory. In: Baker F, Uhlig S., editors. Voicework in music therapy: research and practice. London, UK: Jessica Kingsley; 2011. p. 130–46
22. Theorell T. Psychological health effects of musical experiences: theories, studies and reflections in music health science. London, UK: Springer; 2014.
23. Aalbers S, Fusar-Poli L, Freeman RE, Spreen M, Ket JCF, Vink AC, Maratos A, Crawford M, Chen X, Gold C. Music therapy for depression. *Cochrane Database Syst Rev* 2017;11:CD004517.
24. Zhang Y, Cai J, Hui F, Ren T, Ma H, Zhao Q. Does music therapy enhance behavioral and cognitive function in elderly dementia patients? A systematic review and meta-analysis. *Ageing Res Rev* 2017;35:1–11.
25. Livingston G, Kelly L, Lewis-Holmes E, *et al.* A systematic review of the clinical effectiveness and cost-effectiveness of sensory, psychological and behavioural interventions for managing agitation in older adults with dementia. *Health Technol Assess* 2014;18(39):1–226.
26. Ueda T, Suzukamo Y, Sato M, Izumi S-I. Effects of music therapy on behavioural and psychological symptoms of dementia: a systematic review and meta-analysis. *Ageing Res Rev* 2013;12:628–41.
27. Vasionyte I, Madison G. Musical intervention for patients with dementia: a meta-analysis. *J Clin Nurs* 2013;22:1203–16.
28. Iden KR, Hjorleifsson S, Ruths S. (2011). Treatment decisions on antidepressants in nursing homes: a qualitative study. *Scand J Prim Health Care* 2011;29:252–6.
29. Zhao K, Bai ZG, Bo A, Chi I. A systematic review and meta-analysis of music therapy for the older adults with depression. *Int J Geriatr Psychiatry* 2016;31:1188–98.
30. Bonde LO, Juel K, Ekholm O. Associations between music and health-related outcomes in adult non-musicians, amateur musicians and professional musicians – results from a nationwide Danish study. *Nord J Music Ther.* Forthcoming 2018.
31. Bygren LO, Konlaan BB, Johansson SE. Attendance at cultural events, reading books or periodicals, and making music or singing in a choir as determinants for survival: Swedish interview survey of living conditions. *BMJ* 1996;313:1577–80.
32. Cerejeira J, Lagarto L and Mukaetova-Ladinska E. Behavioral and Psychological Symptoms of Dementia. *Front Neurol* 2012;3:73.
33. Raglio A, Bellelli G, Traficante D, Gianotti M, Ubezio MC, Gentile S, Villani D, Trabucchi M. Efficacy of music therapy treatment based on cycles of sessions: a randomised controlled trial. *Aging Ment Health* 2010;14:900–4.
34. Campbell MK, Piaggio G, Elbourne DR, Altman DG, Grp C. Consort 2010 statement: extension to cluster randomised trials. *BMJ* 2012;345:e5661.
35. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ* 2008;337:a1655.
36. Boutron I, Altman DG, Moher D, Schulz KF, Ravaud P, Grp CN. CONSORT statement for randomized trials of nonpharmacologic treatments: a 2017 update and a CONSORT extension for nonpharmacologic trial abstracts. *Ann Intern Med* 2017;167:40–7.
37. Loudon K, Treweek S, Sullivan F, Donnan P, Thorpe KE, Zwarenstein M. The PRECIS-2 tool: designing trials that are fit for purpose. *BMJ* 2015;350:h2147.
38. World Health Organization (WHO). ICD-10: International statistical classification of diseases and related health problems. 10th revision. Geneva: WHO; 1992.
39. Kitwood T, Bredin K. Towards a theory of dementia care: personhood and well-being. *Ageing Soc* 1992;12:269–87.
40. McDermott O, Orrell M, Ridder H-M. (2014). The importance of music for people with dementia: the perspectives of people with dementia, family carers, staff and music therapists. *Ageing Ment Health* 2014;18:706–16.

41. Eickholt J, Geretsegger M, Gold C. Perspectives on research and clinical practice in music therapy for older people with depression. In: Zubala A, Karkou V, editors. Arts therapies in the treatment of depression. Abingdon, UK: Routledge; 2018.
42. Johnson JK, Nápoles AM, Stewart AL, *et al.* Study protocol for a cluster randomized trial of the Community of Voices choir intervention to promote the health and well-being of diverse older adults. *BMC Public Health* 2015;15:1049.
43. McDermott O. The development and preliminary evaluation of CHORD (CHOrus Research in Dementia) Manual. Alzheimer Europe Annual Conference; 2016 2 November; Copenhagen, Denmark.
44. Collins FS, Fleming R. Sound health: An NIH-Kennedy Center initiative to explore music and the mind. *JAMA* 2017;317:2470–1.
45. Moniz-Cook E, Vernooij-Dassen M, Woods R, *et al.* A European consensus on outcome measures for psychosocial intervention research in dementia care. *Aging Ment Health* 2008;12:14–29.
46. Montgomery SA, Åsberg M. A new depression scale designed to be sensitive to change. *Br J Psychiatry* 1979;134:3829.
47. Leontjevas R, Gerritsen DL, Vernooij-Dassen MFJ, Smalbrugge M, Koopmans RTCM. Comparative validation of proxy-based Montgomery-Asberg Depression Rating Scale and Cornell Scale for Depression in Dementia in nursing home residents with dementia. *Am J Geriatr Psychiatry* 2012;20:985–93.
48. Erkkilä J, Punkanen M, Fachner J, Ala-Ruona E, Pöntiö I, Tervaniemi M, Vanhala M, Gold C. Individual music therapy for depression: randomised controlled trial. *Brit J Psychiatry* 2011;*199:*132–9.
49. Colasanti V, Marianetti M, Micacchi F, Amabile GA, Mina C. (2010). Tests for the evaluation of depression in the elderly: a systematic review. *Arch Gerontol Geriatr* 2010;50:227–30.
50. Morris JC. The Clinical Dementia Rating (CDR): Current version and scoring rules. *Neurology* 1993;43:2412–4.
51. Cummings JL. The neuropsychiatric inventory: assessing psychopathology in dementia patients. *Neurology* 1997;48:S10–6.
52. Cummings JL, Ihl R, Herrschaft H, Hoerr R, Tribanek M. Sensitivity to change of composite and frequency scores of the Neuropsychiatric Inventory in mild to moderate dementia. *Int Psychogeriatr* 2013;25:431–8.
53. Kaufer DI, Cummings JL, Ketchel P, Smith V, MacMillan A, Shelley T, Lopez OL, DeKosky ST. Validation of the NPI-Q, a brief clinical form of the Neuropsychiatric Inventory. *J Neuropsychiatry Clin Neurosci* 2000;12:233–9.
54. Orgeta V, Edwards RT, Hounsome B, Orrell M, Woods B. The use of the EQ-5D as a measure of health-related quality of life in people with dementia and their carers. *Qual Life Res* 2015;24:315–24.
55. Logsdon RG, Gibbons LE, McCurry SM, Teri L. Assessing quality of life in older adults with cognitive impairment. *Psychosom Med* 2002;64:510–9.
56. Chisholm D, Knapp MRJ, Knudsen HC, *et al.* Client Socio-Demographic and Service Receipt Inventory - European Version: development of an instrument for international research - EPSILON Study 5. *Brit J Psychiat* 2000;177:S28–33.
57. Cool C, Cestac P, Laborde C, Lebaudy C, Rouch L, Lepage B, *et al.* Potentially inappropriate drug prescribing and associated factors in nursing homes. *J Am Med Dir Assoc* 2014;15:850.e1–9.
58. Halvorsen, KH, Selbæk G, Ruths S. Trends in potentially inappropriate medication prescribing to nursing home patients: comparison of three cross-sectional studies. *Pharmacoepidemiol Drug Saf* 2017;26:192–200.

59. Lester P, Kohen I, Stefanacci RG, Feuerman M. Antipsychotic drug use since the FDA black box warning: Survey of nursing home policies. *J Am Med Dir Assoc* 2011;12:573–7.
60. Auer S, Graessel E, Viereckl C, Kienberger U, Span E, Luttenberger K. Professional Care Team Burden (PCTB) scale – reliability, validity and factor analysis. *Health Qual Life Outcomes* 2015;13:17.
61. McGlothlin AE, Lewis RJ. Minimal Clinically Important Difference defining what really matters to patients. *JAMA* 2014;312:1342–3.
62. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum; 1988.
63. Button KS, Kounali D, Thomas L, et al. Minimal clinically important difference on the Beck Depression Inventory - II according to the patient's perspective. *Psychol Med* 2015;45:3269–79.
64. R Foundation for Statistical Computing. R: a language and environment for statistical computing [computer program]. Version 3.4.1. Vienna, Austria: R Foundation for Statistical Computing; 2017.
65. Twisk JWR. Applied longitudinal data analysis for epidemiology. 2nd ed. Cambridge, UK: Cambridge University Press; 2013.
66. Rizopoulos D. Joint models for longitudinal and time-to-event data: with applications in R. Boca Raton, FL: Chapman & Hall/CRC Biostatistics Series; 2012.
67. Romeo R, Knapp M, McIntosh B, *et al.* Training costs of person centred planning. In: Curtis L, Netten A, editors. Unit costs of health and social care 2006. Canterbury, UK: Personal Social Services Research Unit, University of Kent; 2006. p. 29–33
68. Beecham J. Collecting and estimating costs. In: Knapp MRJ, editor. The economic evaluation of mental health care. Aldershot: Arena; 1995. p. 157–174.
69. Curtis L. Unit costs of health and social care 2013. Canterbury, UK: Personal Social Services Research Unit, University of Kent; 2013.
70. Ukoumunne OC, Gulliford MC, Chinn S, Sterne JAC, Burney PGJ. Methods for evaluating area-wide and organisation-based interventions in health and health care: a systematic review. *Health Technol Assess* 1999;3(5).

Tables

Table 1. Differences and similarities of group music therapy and recreational choir singing

	Group music therapy	Recreational choir singing
Core principles	Meet psychosocial needs Affect regulation and attunement Empathic relationship	Cognitive activation Sing familiar songs, learn new songs Focus on melody, lyrics, and rhythm
Core intentions	Facilitate and improve communication Reduce behavioural and psychological symptoms through regulation of emotions	Positive experience of self and others Stimulate expression, semantic autobiographic memory, and positive affect
Shared principles and intentions	Use and support remaining faculty of musical reminiscence Tailor to individuals Support social experience, stimulate social and emotional wellbeing	
Proscribed	Push participants to achieve goals	Instrumental improvisation
Dementia inclusion criteria	All levels of dementia, but may be divided to form homogeneous groups	All levels of dementia, but primarily mild to moderately severe dementia; mixed groups possible (inclusiveness)
Group size	Approx. 5	Approx. 10
Qualification of intervention provider	Music therapy degree; skilled musician; member of professional music therapy association or registration body	Skilled musician, choir leading skills and relevant further training

Figure captions

Figure 1. Mechanisms and outcomes of GMT and RCS

Note. GMT – group music therapy; RCS – recreational choir singing.

Figure 2. Flow of participants through the study: Illustration of the study design

Note. CDR – Clinical Dementia Rating; MADRS – Montgomery-Åsberg Depression Rating Scale; GMT – group music therapy; MMSE – Mini-Mental State Examination; MT – music therapy; RCS – recreational choir singing.

Figure 3. Schedule of enrolment, interventions, and assessments

Note. CDR – Clinical Dementia Rating; CSSRI – Client Socio-Demographic and Service Receipt Inventory; d – day; ICD – International Classification of Diseases and Related Health Problems; m – month; MADRS – Montgomery-Åsberg Depression Rating Scale; MMSE – Mini-Mental State Examination; NPI – Neuropsychiatric Inventory; PCTB – Professional Care Team Burden Scale; QOL-AD – Quality of Life-Alzheimer Disease.

Figure 4. Test power as a function of effect size and ICC.

Note. The intraclass correlation coefficient (ICC) describes the relative similarity of participants within units and is typically as low as 0.05 or 0.01;[70] we have added the pessimistic scenario of ICC = 0.10 for completeness only.

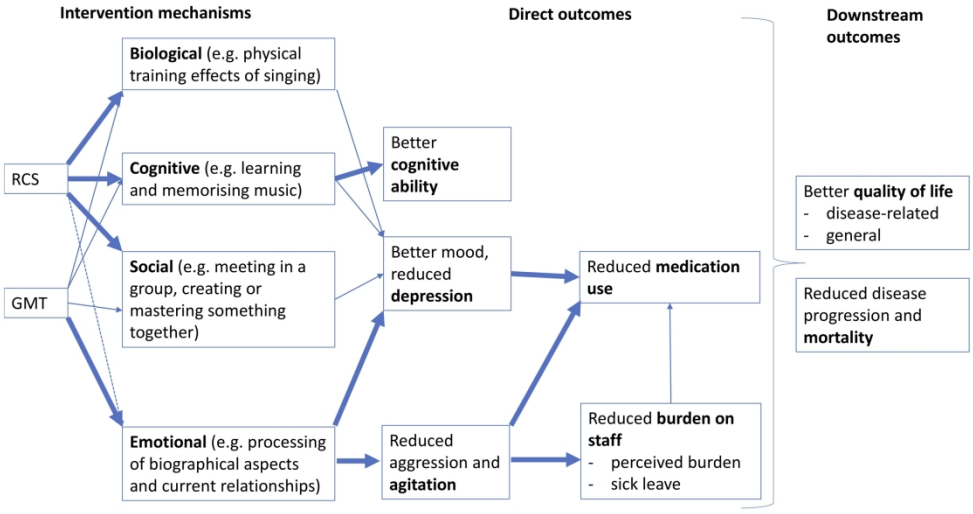

Figure 1. Mechanisms and outcomes of GMT and RCS
Note. GMT – group music therapy; RCS – recreational choir singing.

209x148mm (300 x 300 DPI)

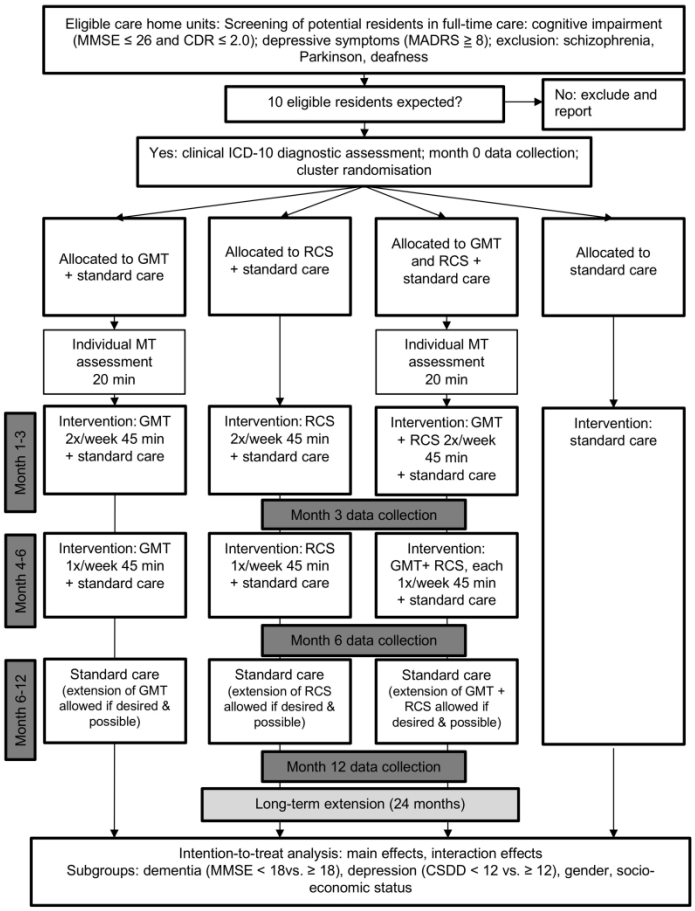

Figure 2. Flow of participants through the study: Illustration of the study design
Note. CDR – Clinical Dementia Rating; MADRS – Montgomery-Åsberg Depression Rating Scale; GMT – group music therapy; MMSE – Mini-Mental State Examination; MT – music therapy; RCS – recreational choir singing.

297x420mm (300 x 300 DPI)

	STUDY PERIOD							
	Enrol-ment	Alloca-tion	Post-allocation					Close-out
TIMEPOINT	-1d	0	1d	3m	6m	12m	...	
ENROLMENT:								
Unit, residents: Eligibility screen	X							
Residents, staff: informed consent (or assent)	X							
Unit: allocation		X						
INTERVENTIONS:								
Group music therapy			←	—	→			
Recreational choir singing			←	—	→			
Standard care			←	—	—	—	→	
ASSESSMENTS:								
Unit baseline: Geographical area Size and costs (CSSRI part 3)	X							
Residents baseline: sociodemographic information (CSSRI part 1); MMSE; dementia diagnosis (ICD-10 code)	X							
Staff baseline: Age, sex	X							
Unit outcomes: Sick leave days			---	---	---	---	---	---------
Residents outcomes: MADRS; CDR; NPI; EQ-5D; QOL-AD; medication and service use (CSSRI part 2)	X			X	X	X	(X)	
Residents outcomes: Adverse events; death			←	—	—	—	—	→
Staff outcomes: PCTB; sick leave days	X			X	X	X	(X)	X

Figure 3. Schedule of enrolment, interventions, and assessments
Note. CDR – Clinical Dementia Rating; CSSRI – Client Socio-Demographic and Service Receipt Inventory; d – day; ICD – International Classification of Diseases and Related Health Problems; m – month; MADRS – Montgomery-Åsberg Depression Rating Scale; MMSE – Mini-Mental State Examination; NPI – Neuropsychiatric Inventory; PCTB – Professional Care Team Burden Scale; QOL-AD – Quality of Life-Alzheimer Disease.

297x420mm (300 x 300 DPI)

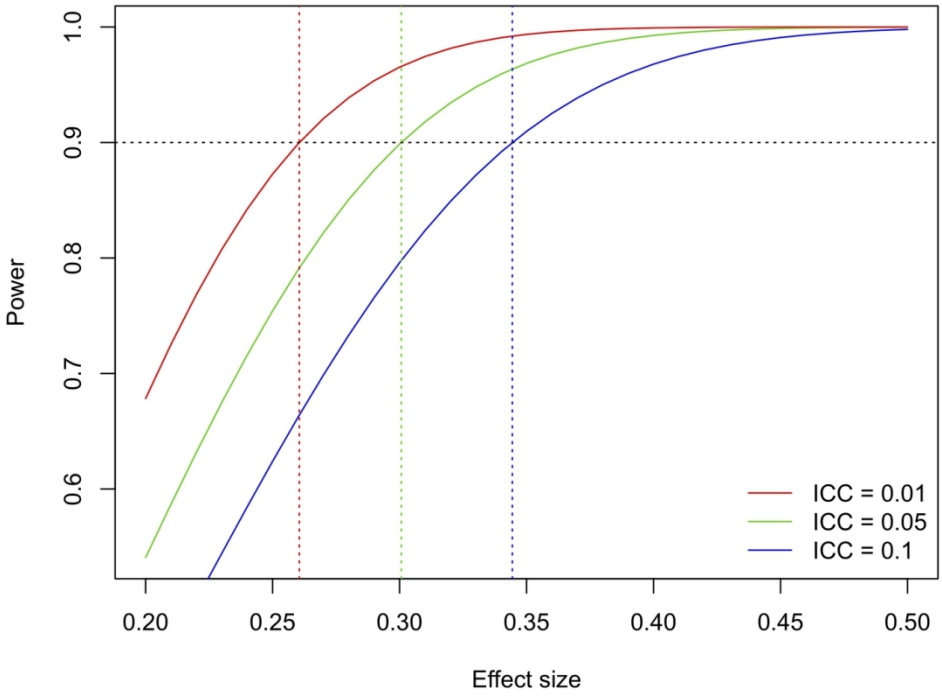

Figure 4. Test power as a function of effect size and ICC.
Note. The intraclass correlation coefficient (ICC) describes the relative similarity of participants within units and is typically as low as 0.05 or 0.01;70 we have added the pessimistic scenario of ICC = 0.10 for completeness only.

125x88mm (300 x 300 DPI)

SPIRIT 2013 Checklist: Recommended items to address in a clinical trial protocol and related documents*

Section/item	Item No	Description	Addressed on page number
Administrative information			
Title	1	Descriptive title identifying the study design, population, interventions, and, if applicable, trial acronym	_____1
Trial registration	2a	Trial identifier and registry name. If not yet registered, name of intended registry	_____3
	2b	All items from the World Health Organization Trial Registration Data Set	_____3
Protocol version	3	Date and version identifier	_____NA (original)
Funding	4	Sources and types of financial, material, and other support	_____2
Roles and responsibilities	5a	Names, affiliations, and roles of protocol contributors	_____1-2
	5b	Name and contact information for the trial sponsor	_____2
	5c	Role of study sponsor and funders, if any, in study design; collection, management, analysis, and interpretation of data; writing of the report; and the decision to submit the report for publication, including whether they will have ultimate authority over any of these activities	_____NA/2
	5d	Composition, roles, and responsibilities of the coordinating centre, steering committee, endpoint adjudication committee, data management team, and other individuals or groups overseeing the trial, if applicable (see Item 21a for data monitoring committee)	_____24

Introduction			
Background and rationale	6a	Description of research question and justification for undertaking the trial, including summary of relevant studies (published and unpublished) examining benefits and harms for each intervention	5-7
	6b	Explanation for choice of comparators	NA
Objectives	7	Specific objectives or hypotheses	8
Trial design	8	Description of trial design including type of trial (eg, parallel group, crossover, factorial, single group), allocation ratio, and framework (eg, superiority, equivalence, noninferiority, exploratory)	8-10
Methods: Participants, interventions, and outcomes			
Study setting	9	Description of study settings (eg, community clinic, academic hospital) and list of countries where data will be collected. Reference to where list of study sites can be obtained	10
Eligibility criteria	10	Inclusion and exclusion criteria for participants. If applicable, eligibility criteria for study centres and individuals who will perform the interventions (eg, surgeons, psychotherapists)	10-11
Interventions	11a	Interventions for each group with sufficient detail to allow replication, including how and when they will be administered	11-14
	11b	Criteria for discontinuing or modifying allocated interventions for a given trial participant (eg, drug dose change in response to harms, participant request, or improving/worsening disease)	11-14
	11c	Strategies to improve adherence to intervention protocols, and any procedures for monitoring adherence (eg, drug tablet return, laboratory tests)	13-14
	11d	Relevant concomitant care and interventions that are permitted or prohibited during the trial	10-11
Outcomes	12	Primary, secondary, and other outcomes, including the specific measurement variable (eg, systolic blood pressure), analysis metric (eg, change from baseline, final value, time to event), method of aggregation (eg, median, proportion), and time point for each outcome. Explanation of the clinical relevance of chosen efficacy and harm outcomes is strongly recommended	14-18
Participant timeline	13	Time schedule of enrolment, interventions (including any run-ins and washouts), assessments, and visits for participants. A schematic diagram is highly recommended (see Figure)	Fig. 3

Sample size	14	Estimated number of participants needed to achieve study objectives and how it was determined, including clinical and statistical assumptions supporting any sample size calculations	18
Recruitment	15	Strategies for achieving adequate participant enrolment to reach target sample size	24-26
Methods: Assignment of interventions (for controlled trials)			
Allocation:			
Sequence generation	16a	Method of generating the allocation sequence (eg, computer-generated random numbers), and list of any factors for stratification. To reduce predictability of a random sequence, details of any planned restriction (eg, blocking) should be provided in a separate document that is unavailable to those who enrol participants or assign interventions	9
Allocation concealment mechanism	16b	Mechanism of implementing the allocation sequence (eg, central telephone; sequentially numbered, opaque, sealed envelopes), describing any steps to conceal the sequence until interventions are assigned	9
Implementation	16c	Who will generate the allocation sequence, who will enrol participants, and who will assign participants to interventions	9
Blinding (masking)	17a	Who will be blinded after assignment to interventions (eg, trial participants, care providers, outcome assessors, data analysts), and how	9
	17b	If blinded, circumstances under which unblinding is permissible, and procedure for revealing a participant's allocated intervention during the trial	NA
Methods: Data collection, management, and analysis			
Data collection methods	18a	Plans for assessment and collection of outcome, baseline, and other trial data, including any related processes to promote data quality (eg, duplicate measurements, training of assessors) and a description of study instruments (eg, questionnaires, laboratory tests) along with their reliability and validity, if known. Reference to where data collection forms can be found, if not in the protocol	14-18
	18b	Plans to promote participant retention and complete follow-up, including list of any outcome data to be collected for participants who discontinue or deviate from intervention protocols	NA

Data management	19	Plans for data entry, coding, security, and storage, including any related processes to promote data quality (eg, double data entry; range checks for data values). Reference to where details of data management procedures can be found, if not in the protocol	24-25
Statistical methods	20a	Statistical methods for analysing primary and secondary outcomes. Reference to where other details of the statistical analysis plan can be found, if not in the protocol	19-23
	20b	Methods for any additional analyses (eg, subgroup and adjusted analyses)	20
	20c	Definition of analysis population relating to protocol non-adherence (eg, as randomised analysis), and any statistical methods to handle missing data (eg, multiple imputation)	20
Methods: Monitoring			
Data monitoring	21a	Composition of data monitoring committee (DMC); summary of its role and reporting structure; statement of whether it is independent from the sponsor and competing interests; and reference to where further details about its charter can be found, if not in the protocol. Alternatively, an explanation of why a DMC is not needed	24-25
	21b	Description of any interim analyses and stopping guidelines, including who will have access to these interim results and make the final decision to terminate the trial	NA/24
Harms	22	Plans for collecting, assessing, reporting, and managing solicited and spontaneously reported adverse events and other unintended effects of trial interventions or trial conduct	17, 24-25
Auditing	23	Frequency and procedures for auditing trial conduct, if any, and whether the process will be independent from investigators and the sponsor	24
Ethics and dissemination			
Research ethics approval	24	Plans for seeking research ethics committee/institutional review board (REC/IRB) approval	25-26
Protocol amendments	25	Plans for communicating important protocol modifications (eg, changes to eligibility criteria, outcomes, analyses) to relevant parties (eg, investigators, REC/IRBs, trial participants, trial registries, journals, regulators)	24-25

Consent or assent	26a	Who will obtain informed consent or assent from potential trial participants or authorised surrogates, and how (see Item 32)	10
	26b	Additional consent provisions for collection and use of participant data and biological specimens in ancillary studies, if applicable	NA
Confidentiality	27	How personal information about potential and enrolled participants will be collected, shared, and maintained in order to protect confidentiality before, during, and after the trial	24-25
Declaration of interests	28	Financial and other competing interests for principal investigators for the overall trial and each study site	2
Access to data	29	Statement of who will have access to the final trial dataset, and disclosure of contractual agreements that limit such access for investigators	NA
Ancillary and post-trial care	30	Provisions, if any, for ancillary and post-trial care, and for compensation to those who suffer harm from trial participation	11
Dissemination policy	31a	Plans for investigators and sponsor to communicate trial results to participants, healthcare professionals, the public, and other relevant groups (eg, via publication, reporting in results databases, or other data sharing arrangements), including any publication restrictions	24, 26-27
	31b	Authorship eligibility guidelines and any intended use of professional writers	NA
	31c	Plans, if any, for granting public access to the full protocol, participant-level dataset, and statistical code	27
Appendices			
Informed consent materials	32	Model consent form and other related documentation given to participants and authorised surrogates	NA
Biological specimens	33	Plans for collection, laboratory evaluation, and storage of biological specimens for genetic or molecular analysis in the current trial and for future use in ancillary studies, if applicable	NA

*It is strongly recommended that this checklist be read in conjunction with the SPIRIT 2013 Explanation & Elaboration for important clarification on the items. Amendments to the protocol should be tracked and dated. The SPIRIT checklist is copyrighted by the SPIRIT Group under the Creative Commons "Attribution-NonCommercial-NoDerivs 3.0 Unported" license.

BMJ Open

Music Interventions for Dementia and Depression in ELderly care (MIDDEL): Protocol and statistical analysis plan for a multinational cluster-randomised trial

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2018-023436.R1
Article Type:	Protocol
Date Submitted by the Author:	12-Oct-2018
Complete List of Authors:	Gold, Christian; Uni Research Health Bergen, Uni Research Health Eickholt, Jasmin; Hochschule fur angewandte Wissenschaften Wurzburg-Schweinfurt Assmus, Jörg; Uni Research Health Bergen, Uni Research Health Stige, Brynjulf; Uni Research Health Bergen, Uni Research Health Wake, Jo; Uni Research Health Bergen, Uni Research Health Baker, Felicity; Melbourne Conservatorium of Music Tamplin, Jeanette; Melbourne Conservatorium of Music Clark, Imogen; Melbourne Conservatorium of Music Lee, Young-Eun; Melbourne Conservatorium of Music Jacobsen, Stine; Alborg Universitet Institut for Kommunikation Ridder, Hanne Mette; Alborg Universitet Institut for Kommunikation Kreutz, Gunter; Carl von Ossietzky Universitat Oldenburg Institut fur Musik Muthesius, Dorothea; Universitat der Kunste Wosch, Thomas; Hochschule fur angewandte Wissenschaften Wurzburg-Schweinfurt Ceccato, Enrico; Universita degli Studi di Verona Dipartimento di Sanita Pubblica e Medicina di Comunita Raglio, Alfredo; Universita degli Studi di Pavia Sezione di Medicina Legale e Scienze Forensi Antonio Fornari Ruggeri, Mirella; University of Verona, Public Health and Community Medicine Vink, Annemieke; University Medical Centre Groningen, dept. of General Practice Zuidema, Sytse; University Medical Centre Groningen, dept. of General Practice Odell-Miller, Helen; Anglia Ruskin University - Cambridge Campus Orrell, Martin; Mental Health Sciences Schneider, Justine; University of Nottingham, Sociology & Social Policy Kubiak, Christine; ECRIN Romeo, Renee; KCL, Geretsegger, Monika; Uni Research Health Bergen, Uni Research Health
Primary Subject Heading:	Geriatric medicine
Secondary Subject Heading:	Mental health, Neurology, Patient-centred medicine, Rehabilitation medicine

Keywords:	Dementia < NEUROLOGY, Depression & mood disorders < PSYCHIATRY, Care homes, Music therapy, Music interventions, Old age psychiatry < PSYCHIATRY

Music Interventions for Dementia and Depression in ELderly care (MIDDEL): Protocol and statistical analysis plan for a multinational cluster-randomised trial

Authors

Christian Gold, GAMUT – The Grieg Academy Music Therapy Research Centre, Uni Research Health, Uni Research AS, Bergen, Norway;

Jasmin Eickholt, Faculty of Applied Social Sciences, University of Applied Sciences Würzburg-Schweinfurt, Würzburg, Germany;

Jörg Assmus, Uni Research Health, Uni Research AS, Bergen, Norway;

Brynjulf Stige, GAMUT – The Grieg Academy Music Therapy Research Centre, Uni Research Health, Uni Research AS, Bergen, Norway;

Jo Dugstad Wake, Uni Research Health, Uni Research AS, Bergen, Norway;

Felicity Anne Baker, National Music Therapy Research Unit, Faculty of Fine Arts and Music, University of Melbourne, Melbourne, Australia

Jeanette Tamplin, National Music Therapy Research Unit, Faculty of Fine Arts and Music, University of Melbourne, Melbourne, Australia;

Imogen Clark, National Music Therapy Research Unit, Faculty of Fine Arts and Music, University of Melbourne, Melbourne, Australia;

Young-Eun Claire Lee, National Music Therapy Research Unit, Faculty of Fine Arts and Music, University of Melbourne, Melbourne, Australia;

Stine Lindahl Jacobsen, Department of Communication and Psychology, Aalborg University, Aalborg, Denmark;

Hanne Mette Ochsner Ridder, Department of Communication and Psychology, Aalborg University, Aalborg, Denmark;

Gunter Kreutz, Department of Music, Carl von Ossietzky University, Oldenburg, Germany;

Dorothea Muthesius, Berlin Career College, Universität der Künste, Berlin, Germany;

Thomas Wosch, Faculty of Applied Social Sciences, University of Applied Sciences Würzburg-Schweinfurt, Würzburg, Germany;

Enrico Ceccato, Mental Health Department, Hospital of Montecchio Maggiore, ULSS8 Berica, Vicenza, Italy;

Alfredo Raglio, Music Therapy Research Laboratory, Istituti Clinici Scientifici Maugeri, Pavia, Italy;

Mirella Ruggeri, Section of Psychiatry, University of Verona, Verona, Italy;

Annemiek Vink, Music therapy department, Academy of Music, ArtEZ University of the Arts, Enschede, The Netherlands;

Sytse Zuidema, Department of General Practice and Elderly Care Medicine, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands;

Helen Odell-Miller, CIMTR – The Cambridge Institute for Music Therapy Research at Anglia Ruskin University, Cambridge, UK;

Martin Orrell, School of Medicine and Institute for Mental Health, University of Nottingham, Nottingham, UK;

Justine Schneider, School of Sociology and Social Policy and Institute for Mental Health, University of Nottingham, Nottingham, UK;

Christine Kubiak, Management Office, European Clinical Research Infrastructure Network (ECRIN-ERIC), Paris, France;

Renee Romeo, Health Services and Population Research Department, King's College London, London, UK;

Monika Geretsegger, GAMUT – The Grieg Academy Music Therapy Research Centre, Uni Research Health, Uni Research AS, Bergen, Norway.

Address correspondence to: Christian Gold, GAMUT – The Grieg Academy Music Therapy Research Centre, Uni Research Health, Uni Research AS, Postbox 7805, 5020 Bergen, Norway. Email: christian.gold@uni.no. Phone: +47-97501757.

Acknowledgements: Stephen Clift, Evan Dawson, Jarle Flemvåg, Julene Johnson, Orii McDermott, Valentina Cabral Iversen, Inger Hilde Nordhus, Paolo Pizziolo, Sabine Ruths, Teppo Särkämö, Geir Selbæk, Tanara Sousa, and Antje Timmer provided valuable comments on an earlier version of this protocol. We would like to thank the following patient advisers for feedback on the study purpose and design: Betty Hamilton, a person with dementia living in a care home; Sue Mountain, her daughter; and various people from the Danish Alzheimer Association, Alzheimer Gesellschaft Würzburg/Unterfranken (Germany), the POLYFON user panel (Norway) Dementia UK, Together in Sound/Saffron Hall, and MHA Care Homes (UK). Feedback from the perspective of professional carers was received from Sarah Yeates and various other people at Caladenia Dementia Care Day Centre; Bupa; and BlueCross.

Autors' contributions:

CG took the initiative for the study; CG, JE, JA, BS, HMOR, SZ and RR developed the concept and design; FAB, JT, IC, YCL, GK, DM, TW, EC, AR, MR, AV, HOM, MO, JS, and MG helped to revise the concept and design. CG drafted the manuscript; JE, JA, RR, and MG helped drafting the manuscript. BS, JDW, FAB, JT, IC, YCL, SLJ, HMOR, GK, DM, TW, EC, AR, MR, AV, SZ, HOM, MO, JS, CK revised the manuscript for important intellectual content. CG, FAB, JT, IC, HMOR, and SZ obtained funding. BS, FAB, JT, IC, YCL, SLJ, HMOR, GK, DM, TW, EC, AR, MR, AV, SZ, HOM, and MG were involved in setting up the study conduct in each site. JDW, FAB, JT, IC, YCL, CK, and MG provided administrative, technical, or material support. All authors approved the final version of the manuscript.

Funding: This work was supported by the National Health and Medical Research Council (grant number APP1137853) and by intramural support from Uni Research AS and University Medical Center Groningen.

Competing interests: None declared.
Word count: 7250 (excluding title page, abstract, references, figures and tables)

Abstract

Introduction: In older adults, dementia and depression are associated with individual distress and high societal costs. Music interventions such as group music therapy (GMT) and recreational choir singing (RCS) have shown promising effects, but their comparative effectiveness across clinical subgroups is unknown. This trial aims to determine effectiveness of GMT, RCS, and their combination for care home residents and to examine heterogeneity of treatment effects across subgroups.

Methods and analysis: This large, pragmatic, multinational cluster-randomised controlled trial with a 2x2 factorial design will compare the effects of GMT, RCS, both, or neither, for care home residents aged 65 years or older with dementia and depressive symptoms. We will randomise 100 care home units with ≥ 1000 residents in total across 8 countries. Each intervention will be offered for 6 months (3 months 2x/week followed by 3 months 1x/week), with extension allowed if locally available. The primary outcome will be the change in the Montgomery-Åsberg Depression Rating Scale score at 6 months. Secondary outcomes will include depressive symptoms, cognitive functioning, neuropsychiatric symptoms, psychotropic drug use, caregiver burden, quality of life, mortality, and costs over at least 12 months. The study has 90% power to detect main effects and is also powered to determine interaction effects with gender, severity, and socio-economic status.

Ethics and dissemination: Ethical approval has been obtained for one country and will be obtained for all countries. Results will be presented at national and international conferences and published in scientific journals.

Trial registration numbers: NCT03496675, ACTRN12618000156280

Keywords: group music therapy, recreational choir singing, depression, dementia, non-pharmacological interventions, psychosocial interventions, randomised controlled trial

Strengths and limitations of this study

- As a multinational trial, this study will provide internationally generalisable results concerning the effects of music interventions in older adults with dementia and depression.
- Based on previous small-scale studies, this trial will have adequate power to determine clinical effects as well as to explain variation in treatment effects in relation to patient characteristics.
- A comprehensive set of core outcomes will be measured, including long-term effects in key variables, with assessor blinding where relevant.
- The trial will also enable modelling of trajectories of change and will thereby contribute to an improved understanding of the mechanisms of music interventions.
- Limitations include the potential bias inherent in cluster-randomised studies if recruitment within clusters is incomplete. Due to the nature of the intervention, care providers and participants cannot be blinded, which may bias measures that rely on their reports.

Glossary of terms

- Site: an organisational or geographical entity containing several units, for example a care home/residential care facility.
- Unit (or care home unit; also ‘cluster’): the smallest organisational unit within a site, where residents live together and are cared for together by staff; each unit is randomised.
- Participant: staff or residents within units who have consented to participate.

INTRODUCTION

Dementia and depression are highly prevalent and comorbid conditions in older adults and are associated with individual distress and high and rising societal costs. Globally, around 50 million people were living with dementia in 2017; this number is predicted to reach 82 million in 2030 and 152 million in 2050.[1] The societal costs of dementia are increasing from a total estimated worldwide amount of US$ 818 billion in 2015, about 1.1% of global gross domestic product,[1] to US$ 1 trillion in 2018.[2] Further, the disease's ramifications for families and carers are significant with respect to financial outlay and carer burden.[3] Dementia is highly prevalent among care home residents; more than half of all Australian care home residents in 2016-2017 had dementia.[4]

Depression is the leading cause of disability worldwide.[5] In older adults, it co-occurs and interacts with dementia in complex ways. Depression can cause cognitive impairment and may increase the risk of developing dementia;[6 7] conversely, depression is very common in the early stages of dementia[6] and often exacerbated by admission to a long-term care facility.[8] Psychotropic medication is only a second-line intervention due to limited efficacy and severe adverse effects, including increased mortality from antipsychotics,[9] but is in practice often used to reduce challenging behaviours in later stages of dementia. Non-pharmacological interventions are available and have some supporting evidence, but further research is needed.[10] Among the most promising non-pharmacological approaches to depression and dementia are music interventions, and in the following section we scope out this evidence.

Music interventions for older adults are based on the notion that music elicits emotional responses and helps to retrieve memories,[11] with recent support from research suggesting that brain regions responsible for processing music, particularly known familiar songs, may be spared even in late-stage dementia.[12 13] They are offered in individual,[14 15] group,[16 17], and community settings[18]

and range from targeted clinical interventions offered by trained music therapists to broader recreational activities, which may be facilitated by choir leaders or nursing staff. However, overlaps between the levels of targeting and training do exist. The most common group-based music interventions may be described as group music therapy (GMT) and recreational choir singing (RCS), where GMT is offered by a music therapist and may use a variety of activities ranging from singing through instrumental music making to music listening, whereas RCS is often facilitated by a choir leader and focuses centrally on singing. Putative mechanisms of GMT and RCS can be described as a combination of biological, psychological (cognitive and emotional), and social mechanisms (Figure 1, left part), however with strong overlaps:

- Among the *psychological mechanisms*, *emotional processing*, such as using musical interactions to regulate affects and to reflect on biographical or current relationships, may be most important in GMT, but is also present to some extent in RCS. *Cognitive processing*, for example through learning and memorising music pieces, is a central mechanism in RCS and less pronounced in GMT, although this may vary between cases, groups, or therapists.[16 19 20]

- *Social mechanisms* are important in both GMT and RCS. Meeting as a group may be important in itself. The function of the group in itself may be relatively more important in RCS, whereas GMT to a greater extent also relies on the one-to-one relationship between the therapist and each group member. Another important part of the social mechanisms is developing a shared sense of mastery and achievement through learning and performing music pieces, which is more central in RCS than in GMT but may again vary from case to case.[21]

- *Biological mechanisms* include the physical training effects of singing and other music-related activities, which may include movement. They are important in both GMT and RCS but may be more central in RCS.[18 22]

Systematic reviews of the clinical effects of GMT and RCS have reported mixed results, [10 17 19 20 23 24 25 26 27] possibly owing to the heterogeneity of treatment effects across types of participants and music interventions. One small trial comparing GMT and RCS directly suggested that the comparative effects of these music interventions may depend on the comorbidity of dementia and depressive symptoms.[16]

Process-outcome relations of music interventions may be described as follows (Figure 1, right part):

- *Emotional processing* in a therapist-client relationship may lead to finding meaning and regaining orientation, and thereby to reduced *agitation* and related *neuropsychiatric symptoms*. Such effects have been suggested in some systematic reviews,[17 26] but not in others.[19] Reduced agitation may in consequence reduce *burden on staff*[15] and consequently reduce *sick leave*. This may also help to reduce inappropriate use of *medication*,[14] which is a concern in care homes.[28]
- *Cognitive processing* through practicing music may promote or maintain *cognitive functioning* in older adults. Such effects have been shown for active music therapy, but not music listening, for people with dementia.[20]
- *Emotional processing*, but also social, biological, and cognitive mechanisms may be associated with improved mood and reduced *depressive symptoms*. Systematic reviews have suggested effects of music therapy on depressive symptoms in older adults in general[29] and in people with dementia.[19]
- As downstream outcomes of all four mechanisms and of the intermediate outcomes above, one may expect improved *quality of life*, and possibly reduced *mortality*, although these effects may be small[30 31] and indirect.[9] Music interventions may also reduce costs by reducing time spent on treating and alleviating neuropsychiatric symptoms and reducing absence by staff.

Hypotheses

Through these different pathways of GMT and RCS, one may hypothesise differential effects for different outcomes, and therefore for different subgroups of care home residents. Specifically:

- GMT may be more effective than no GMT, and RCS may be more effective than no RCS, with respect to reducing depression symptoms and other outcomes shown in Figure 1.
- GMT and RCS may differ in the pattern of effects across outcome domains, which may be explained by their different mechanisms. For example, GMT may be more effective than RCS for reducing aggression and agitation and may therefore be more beneficial for people with late-stage dementia who often present with these neuropsychiatric symptoms.[32 33] RCS may be more effective than GMT with respect to cognitive functioning. Effects on depression symptoms may be achieved through different pathways (Figure 1), and the strength of those effects may therefore depend on severity or comorbidity.[16]
- When offered together, synergistic effects of GMT and RCS may occur through activation of different pathways.
- Cost-effectiveness may differ accordingly across interventions and subgroups. As RCS is likely to be associated with lower intervention costs, it may have better cost-effectiveness ratio in areas where clinical effects are similar; however, this will depend also on each intervention's effects on use of other treatments and services.

METHODS AND ANALYSIS

Design

This large, multinational cluster-randomised controlled trial will be conducted in care homes in Australia, Denmark, Germany, Italy, Netherlands, Norway, Poland, and the UK. The list of study sites is provided in the trial registration record. MIDDEL uses a 2x2 factorial design to examine the effects of GMT, RCS, both, or neither, for elderly care home residents with dementia and depressive symptoms (Figure 2). This design enables investigating the effects of two music interventions as well as potential synergy effects between them. These may occur between intervention providers on the cluster level (GMT and RCS providers learning from each other) and through residents on the individual or cluster level (participants gaining in different ways from the combination). We will randomise 100 or more care home units (clusters) in eight countries for a total of 1000 or more participants. Recruitment started in July 2018, and primary completion is anticipated for April 2020.

Block randomisation (block size 4 clusters) will be used to ensure that each site will have a balanced distribution between the interventions. The computer-generated randomisation list will be created and kept concealed at the central study office. Only after the eligibility of a care home unit is confirmed and eligible participants (residents and staff) within that unit have formally consented and completed baseline assessment, will site investigators be informed of the randomisation result for that unit. Where possible, a number of care home units will be randomised at the same time, which will further ensure allocation concealment.

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 and RCS. Plain language summaries and consent forms will use neutral wording to maintain equipoise and to avoid expectancy effects. 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 well) 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.

GMT and RCS may entail "ripple effects" beyond the individual participants by leading to changes of the local milieu/culture at the care home unit.[15 34] These will be assessed by measuring objective and perceived burden on care staff. The cluster design is ideally suited for that situation because it facilitates application in a naturalistic setting and avoids some of the problems of individually randomised trials (such as treatment contamination); it also minimises the additional workload for care staff. Trial procedures will be tested in the Australian cohort before applying them in the other countries. The trial will be conducted and reported in accordance with relevant legal frameworks and research guidelines.[35 36 37 38]

Participants

Eligibility is defined on two levels, care home units and individual participants. Participating care home units will be those that are expected to have at least 10 eligible and consenting residents. Care home units that are currently providing music-based interventions as part of their usual care programme will be excluded. Eligible participants will meet all of the following inclusion criteria:

- aged 65 years or older, resident (full-time, 24h/day) at a participating care home;
- dementia as indicated by a Clinical Dementia Rating score of 0.5 to 2 and a Mini-Mental State Examination (MMSE) score of 26 or less;
- at least mild depressive symptoms, as indicated by a Montgomery-Åsberg Depression Rating Scale (MADRS) score of at least 8;

- a clinical diagnosis of dementia according to ICD-10 research criteria;
- have given written informed consent (may be assent by proxy for those unable to provide consent themselves).

Clinical diagnosis will be ascertained by a clinician or researcher, based on the ICD-10 dementia criteria of memory decline; decline in other cognitive abilities; impairment in activities of daily living; preserved awareness of the environment; decline in emotional control or motivation or change in social behaviour; and more than 6-month duration of memory decline and other cognitive symptoms.[39] People with a known diagnosis of schizophrenia or Parkinson's disease or those who are known to be severely hearing-impaired, in short-term care, or unable to tolerate sitting in a chair for at least part of the sessions, will be excluded. People may however have other clinical diagnoses such as pre-morbid substance use disorders or anxiety disorders. The list of exclusion criteria is intentionally short to ensure generalisability.[38] Residents will always be provided information about the study, and their ability for consent will be assessed directly, before turning to proxies (next of kin/legal representative/carer) for written informed assent. In case of doubt, consent/assent will be provided by both resident and proxy. Residents unable to provide written consent will still be asked if they agree to the interventions and assessments when these begin.[40]

Interventions

Units in all intervention arms will continue with standard care as locally available. In the units allocated to music interventions, GMT, RCS, or both will be provided twice weekly for the first three months, followed by weekly sessions for the next three months. Continuation of GMT and RCS is allowed after that period, depending on local availability. Data on the resources related to the interventions will be measured (number of sessions attended by each participant, duration of each

session, non-contact time spent by the intervention provider to prepare or follow up a session, recorded by the provider). The components of standard care provided will also be recorded (see Outcomes).

GMT and RCS sessions will be 45 minutes each. In line with usual practice, GMT may be divided into smaller groups (e.g. around 5 participants, but this may differ across local contexts), whereas RCS may be conducted in larger groups (e.g. with all residents of the unit in one group).

GMT. A core principle of GMT is affect regulation through active, reciprocal music making with the use of singing and musical instruments (Table 1). This facilitates the relationship between the music therapist and the person living with dementia, and between participants in the group. Another core principle of GMT is to meet the psychosocial needs of each individual resident, which in turn is thought to reduce depressive symptoms and anxiety and to stimulate overall social and emotional wellbeing.[41 42 43] GMT aims to work in the "here and now" by responding to participants' immediate emotional expressions, containing them, and incorporating them into meaningful musical expressions for therapeutic gain.[21] GMT is provided by a trained music therapist, who is registered with the appropriate professional association or registration body in his or her country and should also be skilled as a musician. To facilitate individual relationship-building, the music therapist will offer each resident an initial 20-minute assessment with the aim of determining their musical preferences and starting to build individual rapport. The music therapist will also use other sources to determine the participants' musical biography, cultural background, history, personal strengths, resources, and disabilities, and any other information that could be useful to bring into GMT sessions.

RCS. A core principle of RCS is to sing familiar songs and to provide a familiar musical environment for participants (Table 1). Choral singing involves a combination of cognitive, physical, and psychosocial engagement components.[44] Drawing on the psychosocial aspects of a choir setting, RCS in this trial aims to foster connectedness in a group either with other older adults residing in the care

homes or family caregivers; emotional wellbeing; and enjoyment of music-making in a group. Biographically and culturally grounded song materials are used with the central goal of stimulating positive experiences shared by groups of individuals. Where participants have engaged in music activities in their past, this may also enable the continuation, as far as possible, of the familiar social experience of music-making in everyday life. Sections of RCS sessions may vary in their focus; for example, sessions may focus on developing familiarity with well-known songs; learning and developing new material as a group; singing rounds to encourage listening to each other; or offering space for solo singing.[45] The materials can be familiar songs from a range of repertoires, including but not restricted to festive songs (e.g. birthday songs, Christmas carols), folk songs, traditional, classical, or popular songs. The selection of songs can vary from country to country, within and between choir leaders, and may also depend on seasonal and other circumstantial factors. RCS is provided by a skilled musician with choir leading skills.

Training and assessment of treatment fidelity. GMT and RCS providers will receive training and implement intervention guidelines developed in the initial phase of the study. Regular exchange and peer supervision for GMT and RCS providers will be organized in conjunction with guidelines and training. This will include monthly online or in-person meetings between researchers and intervention providers to ensure intervention quality and fidelity, to discuss potential threats that might undermine study quality, and to refine the guidelines accordingly. Intervention providers will also attend weekly staff meetings at intervention sites where possible, to maximise local knowledge transfer and benefit. Manuals for complex interventions need to standardize the quality of interventions to avoid unwarranted variation between therapists and countries while preserving the possibility for meaningful tailoring to local contexts and individuals.[46] This will be addressed by focusing on general principles rather than fixed behaviours. Both GMT and RCS should be tailored to fit the current situation/status of

the group and its individual members. Interventionists will be trained at all sites, both through local in-person meetings with all intervention providers at each site and through remote online training across sites. The purpose of this training is to supplement rather than replace the existing training and expertise of intervention providers. For assessment of adherence and competence, providers of GMT and RCS will be video-recorded in 3-4 randomly selected sessions per unit. We will record and analyse the entire session. To avoid performance bias due to the awareness of being videotaped in a selected session, we will use sham video monitoring in other sessions where possible. Videos will be uploaded and stored on a secure central server and will be available only to those who check treatment fidelity. Two independent researchers will assess the different components used by intervention providers and the degree of person-centeredness (i.e. tailoring of the intervention to the current situation/needs of the group and its members). This process-related data will help us to understand the mechanisms or effective ingredients of each intervention.

Further development. While the description above provides general guidance and will form the basis for fidelity assessment in this study, no consensus guidelines exist for GMT and RCS. Descriptions in the literature vary in many aspects such as: theoretical frame; session structure; specific therapeutic goals; types of musical instruments and materials; inclusion of music listening in addition to active music-making; structured versus improvisational techniques in active music-making; and adaptation/tailoring to reach each person individually. Therefore, flexible manuals, including sets of detailed principles and techniques for GMT and RCS, will be developed and agreed upon by scientific and clinical experts from different countries using a modified Delphi consensus procedure.

Outcomes

The study uses a broad array of resident-, staff-, and unit-level outcomes measured at 3 months, 6 months (primary), and 12 months after randomisation (Figure 3). A long-term extension with later

follow-ups is planned separately. Where possible, core outcomes (www.comet-initiative.org) for psychosocial intervention research in dementia care, that are widely used and available across the languages of the trial, were selected.[47]

The primary endpoint will be change in the total score of the Montgomery-Åsberg Depression Rating Scale (MADRS). The MADRS is a 10-item scale where each item is rated from 0 (no abnormality to 6 (severe).[48] In the total sum score ranging from 0 to 60, higher scores indicate higher severity of depressive symptoms. Assessment is based on an interview with the resident where possible, but where definite answers cannot be elicited from them, all relevant clues as well as information from other sources should be used as a basis for the rating, in line with usual clinical practice.[49] The total time of administration is approximately 20 minutes. The MADRS has been used successfully in previous studies of music interventions.[16 50] It has shown high reliability and validity, and its sensitivity to change compares favourably to other scales evaluating depression severity in this population, such as the Cornell Scale for Depression in Dementia (CSDD).[49 51]

Secondary outcomes will include the following:

- Dementia severity including cognitive and functional performance – Clinical Dementia Rating (CDR), a standard assessment of dementia severity.[52] The CDR is used widely in clinical settings. Its score is derived from a semi-structured interview with the person living with dementia and an appropriate caregiver/relative. It rates impairment in each of 6 cognitive categories (memory, orientation, judgment and problem solving, community affairs, home and hobbies, and personal care). Its score is useful for characterising and tracking a person's level of impairment or dementia: 0 = normal; 0.5 = very mild or questionable dementia; 1 = mild dementia; 2 = moderate dementia; 3 = severe dementia.

- Neuropsychiatric symptoms – Neuropsychiatric Inventory (NPI), "a *de facto* standard for measuring neuropsychiatric symptoms in clinical trials".[47] Developed to assess behaviour in people living with dementia, the NPI has substantial evidence of validity and reliability and has been translated into more than 40 languages.[53 54] The NPI uses a screening approach to minimise administration time, examining and scoring only the domains with positive responses to screening questions. In this study, the NPI – Questionnaire (NPI-Q)[55] will be used; another version specific for nursing homes (NPI-NH) was considered but rejected because it is not available across all languages. The NPI-Q includes 12 domains where if a symptom is present, both its severity (from 1= mild to 3=severe) and the associated distress on caregivers (from 0=Not distressing at all to 5=Extreme or very severe) are assessed by the professional carer who is most familiar with the resident's behaviour. Item scores across the 12 domains are summed, leading to a total severity score from 0 to 36, where higher values represent higher severity. The additional total distress score can range from 0 to 60, also with higher values representing higher distress.[55]
- Generic quality of life – EuroQol (EQ-5D-5L), a generic health utility measure.[47] The standardized, non-disease-specific instrument for evaluating health-related quality of life was developed by the international EuroQol group and is used to derive quality-adjusted life-years (QALYs). It is based on a descriptive system that defines health in the five dimensions mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Each dimension has five response categories from "no problems" to "extreme problems", which are combined using preference weights to form an overall quality of life score ranging from lower than 0 (worse than death) to 1 (best possible). An additional visual analogue scale indicates today's health on a scale from 0 ("The worst health you can imagine") to 100 ("The best health you can imagine"). As most residents will be unable to self-rate the EQ-5D-5L, the rating will rely on the judgment of the carer as a proxy. Careful selection of

assessment mode (self/proxy/both) and choice of appropriate proxies is important to ensure the measure's validity in studies of people with dementia.[56]

- Disease-specific quality of life – Quality of Life in Alzheimer's Dementia (QOL-AD).[47 57] This 13-item scale with a self-rating and proxy version has demonstrated sensitivity to psychosocial intervention, correlates with health-utility measures, is widely translated and used internationally and can be used by people with very low MMSE scores. Items such as "Physical health", "Memory", or "Ability to do things for fun" are scored on a scale from 1 (poor) to 4 (excellent), resulting in a total score ranging from 13 (worst) to 52 (best).
- All-cause mortality (time to death), as recorded in official electronic registries.
- Any increase in psychotropic drug use: Data on type (ATC Codes N065, N06) of psychotropic medication used and any increase or decrease over time will be collected from care staff using the 'medication profile' section of a tailored version of the Client Socio-Demographic and Service Receipt Inventory (CSSRI).[58] Available electronic health registry data will be used where possible. Psychotropic medications are sometimes used inappropriately to manage behavioural symptoms of dementia.[59 60 61] An earlier study suggested that music therapy may help prevent increase in medication.[14]
- Costs: Total and component costs of the interventions will be assessed from a societal perspective, including the cost of the intervention as well as statutory health and social care services used, using a tailored version of the CSSRI.[58]
- Any adverse events (safety): No adverse effects of music interventions are known from earlier trials. Intervention providers are trained to work closely with and adapt their interventions to the needs of participants in order to avoid adverse reactions. Because little knowledge exists about

what the potential adverse events could be, all types of adverse events and serious adverse events (e.g. unexpected worsening of symptoms), whether related or unrelated to the interventions, will be reported.

Staff-level outcomes will be as follows:

- Subjective perceived burden of care staff: Professional Care Team Burden Scale.[62] The 10-item scale provides a valid and reliable means of obtaining ratings of burden from formal care teams working in care homes in order to evaluate different interventions targeted at the reduction of burden in care teams. Items are scored on a 5-point scale from 0 (strongly disagree) to 4 (strongly agree), yielding a total sum score from 0 to 40, with higher scores indicating higher burden.
- Days on sick leave of care staff, as recorded monthly by the employer.

Sample size and test power

There is no consensus on the minimal clinically important difference (MCID)[63] on the MADRS. Generally, effect sizes in the small- to medium range (i.e. between d = 0.20 and 0.50) may be considered relevant.[64] Effect sizes in that range were also found in a previous trial on GMT and RCS (d = 0.33 at 6 weeks and 0.49 at 12 weeks).[16] Studies of other depression scales have used anchor-based approaches to determine clinically important percent reductions;[65] we will not use such approaches for the primary analyses, but will include an additional responder analysis.[63]

The trial has the multiple aim of identifying main effects of GMT versus no GMT and RCS versus no RCS, interaction effects of GMT and RCS, and predictive effects of clinical characteristics including severity of dementia; severity of depression; gender; and socio-economic differences. (Although individual socio-economic differences tend to become more equal amongst residents in a given care home, they may still exist at the cluster level, as different homes may have different standards; we will

use the average cost of living in each care home unit as a cluster-based proxy measure for socio-economic status.) Power for interaction effects and subgroup analyses is difficult to determine because of the unknown distributions and effect sizes of the different variables. Therefore, the power calculation for the primary outcome was based on the main comparisons of GMT versus no GMT and RCS versus no RCS. This approach maximises power by fully exploiting the factorial design. A general two-sided significance level of 5% will be used, leading with Bonferroni adjustment to a marginal two-sided level of 2.5%. The power calculation was adjusted for cluster effects using the intraclass correlation coefficient (ICC, between 0.01 and 0.10, Figure 4), assuming average cluster size 10. It was further assumed that attrition, which may occur due to death, moving to another care home, or withdrawal from the study, will be no higher than 20% overall. With 100 clusters and 1000 participants randomised, 90% power is reached for effect sizes between 0.25 and 0.35 (Figure 4). Any further increase beyond this sample size will serve heterogeneity of treatment effects analyses.

Statistical analyses

The statistical analyses will use multivariate longitudinal statistical models, which make optimal use of the data by using data from all time points at once and can account for the effects of clustering within care home units and sites. We will use a modified intention-to-treat (ITT) approach using all available data from all participants as randomised, regardless of the intervention actually received. Sensitivity analyses using multiple imputation for missing data will enable a full ITT analysis. Additional per-protocol analysis will address the effects of treatments as actually received and will complement the ITT analyses. All tests in the study will be two-sided. The general significance level is set to 0.05. Since there are two comparisons in the primary analysis (GMT vs. no GMT, RCS vs. no RCS), we will use a marginal Bonferroni level of 0.025. Continuous variables will be screened for normality. All computations will be done using R.[66]

Sociodemographic and clinical baseline properties for the groups will be characterized by descriptive methods (mean (SD), median [range], n (%)) and presented in a table. A similar table will compare those who dropped out versus those who completed the primary outcome.

The primary outcome, change of MADRS score from baseline to 6 months, will be assessed by a linear mixed-effects model (LME).[67] We will fit the unadjusted model for each treatment (RCS vs. no RCS) as well as the multivariate model containing both treatments as predictors both unadjusted and adjusted for the interaction between the treatments.

Secondary analyses of MADRS scores will include the following:

- The development of MADRS in the treatment groups over the entire study period will be assessed by a LME including time, treatment type and the interaction of time and treatment type as fixed effects, and participant nested within cluster as random effects. We will use both linear and simple contrasts in the time domain because it is not known whether there is a linear association in time. This will be illustrated by a figure showing the predicted mean of MADRS for each treatment type at each time point with confidence intervals.
- The synergy of the two treatments will be assessed by the LME containing both treatments as well as their interaction as predictors. The interaction in the model will estimate the synergy effect.
- The predictive effect of several covariates (severity of dementia; severity of depression; gender; and socio-economic differences) will be assessed as odds ratios using LMEs for each covariate containing time, treatment type and their interaction as well as the covariate and the interaction between the covariate and the treatment type as predictors. The interaction in the model will estimate the predictive effect.

Secondary endpoints will be analysed as for the primary analysis, using LMEs for continuous outcomes (both resident-level and staff-level). Special considerations apply for the following variables:

- Binary outcomes, including response rates (the proportion of residents improved by at least 50% from their baseline MADRS score), prevalence of medication use, and adverse events, will be assessed as odds ratios using generalised linear mixed models (GLMMs) with a logit link function. The predictive effect of several covariates (severity of dementia; severity of depression; gender; and socio-economic differences) will be assessed by GLMMs for each covariate containing time, treatment type and their interaction as well as the covariate and the interaction between the covariate and the treatment type as predictors. The interaction in the model will estimate the predictive effect.
- Count data (days of sick leave) and cost data are more likely to follow a Poisson distribution than a normal distribution and will be analysed using the respective GLMMs.
- Time-to-event data include mortality (time to death of any cause) and will be assessed by Kaplan-Meier and log-rank- or Breslow tests for differences between the treatment types and the hazard ratios at 12 months.
- Loss to follow-up in all other outcomes can be influenced by mortality. Thus, if the survival analysis shows differences between the groups, it will be meaningful to use a joint modelling approach which combines the longitudinal models and the survival analysis.[68]

In addition to analysing effects of interventions as randomised, we will conduct mediator analyses to examine relations between elements of the therapy approach (mechanisms), direct and downstream outcomes, as depicted in Figure 1, using structural equation modelling (SEM).

Cost-effectiveness analysis

Total and component costs of the interventions and the cost-effectiveness of alternative interventions will be assessed from a societal perspective. This perspective will cover three categories of costs: the cost of the intervention, statutory health and social care (and voluntary sector) service costs, and costs of unpaid carer support. The cost per session for each of the interventions will be derived employing established approaches used in a compendium of costs and in published studies.[69 70 71] Information on the time inputs by GMT and RCS providers (for running sessions and for other activities) will be obtained and valued using information on the midpoint of the salary scale and employer's national insurance as well as superannuation contributions. The sum of the staffing contributions and allocations for overheads for each session will then be summed, to derive a cost per session. To this cost per session, the average number of sessions delivered as part of the intervention will be multiplied to derive a cost per intervention. As there is no clear agreement on how the costs of group interventions should be allocated, we will calculate the cost per session of each of the interventions on the basis of the participants allocated to each of the groups, regardless of whether or not the participant attended, because participants who miss a session are not replaced.

Data on statutory services used will be collected using a tailored version of the Client Socio-Demographic and Service Receipt Inventory (CSSRI),[58] which contains data on the use of health and other formal care resources and unpaid care. To service and support data we will attach unit costs reflecting the long-run marginal opportunity costs drawn from available public sources. Costs per unit of measurement for each service type will be taken from country-specific sources. We will adjust country-specific costs to Euros using purchasing power parity methods. Costs and outcomes will be compared for the comparators using extended dominance approaches. In this approach, the four treatment combinations (GMT, RCS, GMT and RCS, no GMT or RCS) will be ranked by cost, and if one is dominated (more expensive and less effective than another), it will be excluded from further

analysis, until two therapeutic groups are left on which to explore which of the two groups is most cost-effective. The cost-effectiveness of one arm over another will be compared by calculating incremental cost-effectiveness ratios (ICERs) defined as difference in mean costs (Euros spent) divided by difference in mean effects (QALYs using the EQ-5D-5L; points improved on MADRS and QOL-AD). Cost-effectiveness acceptability curves will be plotted for each cost-outcome combination to show the likelihood of one treatment being seen as cost-effective relative to another for a range of values placed on incremental outcome improvements. Using the net benefit approach, monetary values of incremental effects and incremental costs will be combined, and net benefit (NB) derived as: $NB = \lambda * (effect_b - effect_a) – (cost_b – cost_a)$, where λ is the willingness-to-pay for a unit improvement in effectiveness, and subscript 'a' and 'b' denote two candidate treatment arms. There is no agreed cross-national willingness-to-pay threshold, and in some countries there is no established threshold at all. Other studies have used a threshold of Euros 50,000 per QALY, and we will consider this in the discussion of the results.

Sensitivity analyses will be conducted to assess the robustness of the results to changes in key parameters. One of the possible concerns is likely to be the sample size. If the sample size in some participating countries is too small, their cost-effectiveness estimates are likely to be unreliable. We shall therefore consider the added value of pooling the information on costs and outcomes in sensitivity analyses.

Patient and public involvement

The development of the research question and study design was informed by the priorities, experience and preferences of residents and carers. Co-authors in Australia, Denmark, and the UK have been actively involved in user and advocacy organisations in their countries for a long time and have discussed interventions, outcomes and the need for research with them. Relatives and caregivers spoke

to the importance of music interventions as a help for carers and people with dementia, and to the need for high-quality evidence on their effects. Relatives and caregivers are important for giving persons with dementia a voice when they cannot speak for themselves.

Co-authors in Australia had significant involvement with residents, as well as with care staff and care home managers, in discussing and piloting aspects of the study design. While the interventions were generally perceived as pleasurable rather than burdening, some of the outcome measures were felt to be burdening and too demanding due to their length or complexity. As a consequence, the longer Cornell Scale for Depression in Dementia was replaced with the shorter MADRS, and a more extensive quality of life scale was removed. Recruitment strategies were discussed and adapted in dialogue with care home staff.

User representatives will continue to be actively involved throughout the conduct of the trial (see next section). Results will be disseminated to residents, relatives, and care staff via care homes. Results will also be disseminated to national user and advocacy organisations.

Monitoring and oversight

One representative of each recruiting institution will be a member in the *Trial Steering Committee (TSC)*. They will be supplemented by other members who are independent of the investigators, their organisations, funders and sponsors. The TSC will include service users or their relatives and representatives of stakeholder organisations such as Alzheimer Europe and Dementia Australia. The TSC will have regular meetings to closely supervise all aspects of the study, including any protocol amendments, progress of recruitment, and publication plan.

Data quality monitoring will require a risk-based monitoring approach including remote monitoring activities performed centrally and on-site monitoring as needed. The monitoring will be performed

according to the monitoring manual to be developed at the beginning of the project. Recruitment and retention rates will be monitored closely to mitigate the risk of slow recruitment. The number of participating care home units in total and in relation to care home units screened; the number of participating care home residents in total and in relation to residents screened of potential participants; and the retention of participants in the study over the trial period will be closely monitored.

A *Data and Safety Monitoring Committee (DSMC)*, consisting of three people with strong methodological and clinical expertise who are not otherwise affiliated with the project or its institutions, will be appointed early in the international trial. The DSMC will receive regular updates on recruitment, uptake of interventions, any unforeseen events, adverse events, and immediate information on serious adverse events from the trial statistician. It will have unblinded access to study data. Meetings with the DSMC will be on a biannual basis and will consist of an open and a closed part. In the open part, the general progress of the trial will be discussed; in the closed part, the DSMC will discuss any safety signals with the trial statistician. If issues arise, the DSMC will recommend to the TSC on appropriate action.

All aspects of the study, from intervention fidelity through recruitment, outcome assessment, database and data quality management, to data and safety monitoring, will be pilot-tested in one country (Australia) before being rolled out internationally. The data of the pilot cohort will be included in the main trial; no statistical adjustments are made because the decision depends only on feasibility, not on an interim efficacy analysis. Patient-related documents such as the consent form will be tested because they may influence how the study is perceived by potential participants, relatives and staff.

To ensure data quality, a trial database will be set up and maintained using a safe server hosted by Uni Research (UHEADS) and OpenClinica software. UHEADS is a system for safely storing health research data developed by Uni Research AS, that accommodates the safe upload, storage and retrieval

of any sensitive research data. OpenClinica is a web-based system for electronic data capture and clinical data management for multicenter clinical trials, which conforms to relevant international standards for health research. Uni Research AS runs an open source version of OpenClinica.

ETHICS AND DISSEMINATION

Ethical aspects

Ethical approval has been obtained from the Medicine and Dentistry Human Ethics Sub-Committee at the University of Melbourne, Australia (approval date: January 12, 2018) and will be obtained from the relevant local institutional human research ethics committee at each international site. Local clinical investigators will work on adaptation of study- and patient-related documentation to meet national ethical requirements.

Risk management

The main risks are: slow recruitment; low fidelity of interventions; and low reliability of outcome measurements. Regarding recruitment, we will rely on clinical investigators with a track record of successful recruitment. Slow recruitment at some sites can be compensated by other sites. Fidelity of interventions will be ensured through clear guidance and ongoing monitoring. Reliability of outcomes will be facilitated by the selection of widely used, recommended core outcome measures, and assessed through tests of inter-rater reliability.

Publication plan

The report on the main, pre-planned analyses of the primary endpoint and up until the 12-month follow-up will be submitted to a leading medical journal. The report on the long-term extension will also be submitted to a leading medical journal. Further publications may focus on recruitment and

retention strategies for international cluster-randomised multicentre trials of complex interventions in non-medical settings; development of an MCID for the MADRS based on an existing anchor question; inter-relations between outcomes and predictive value of early outcomes for later outcomes; clinical descriptions and qualitative research of therapy processes, including qualitative influences on care home staff, their perception of GMT and RCS and their potential "ripple effects"; barriers and facilitators for implementation, using qualitative interviews and surveys; and consensus guidelines for GMT and RCS. The data and meta-data will be stored in a public repository, such as that of the Norwegian Centre for Research Data (NSD).

Relevance and benefit to society

Music interventions are widely used in care homes, and their effects are likely heterogeneous. MIDDEL is designed to provide reliable and generalisable knowledge about effectiveness, mechanisms, and heterogeneity of effects of music interventions. It will also fill knowledge gaps about potential long-term benefits and preconditions for achieving such sustained benefits. The results will drive changes in aged care and will contribute to our understanding of the relation between music and health.

Implications for practice

If MIDDEL shows beneficial effects, differences in scalability need to be considered for successful implementation. GMT requires extensive, specialised music therapy training and is typically provided in small groups. The number of qualified music therapists is limited; it varies from country to country, but fluctuates around 1 in 100 000 (about 6000 in Europe, http://emtc-eu.com; 5000 in the USA, www.cbmt.org; 500 in Australia, www.austmta.org.au). RCS is more easily scalable as it can be provided by trained musicians and also in larger groups. There are about 1 million choirs and 37

million choir singers in Europe (www.singingeurope.org). MIDDEL will provide the knowledge needed to identify the best targeting of both approaches, as well as contributing to their improvement and standardisation. For example, GMT with its highly person-centred approach may be most beneficial to those with neuropsychiatric symptoms, which are typical at late-stage dementia, whereas the social engagement in RCS may help those at earlier stages, and the combination of both may be best for another subset of residents with more complex needs. The knowledge generated by MIDDEL will thus increase the impact of music interventions in care homes and potentially in related contexts, such as day care centres for people still living at home.

Implications for future research

As a strongly interdisciplinary project building on contributions from medicine, social sciences, and humanities, this trial will contribute to strengthening the collaborations between these fields, which is likely to stimulate new cross-disciplinary investigations. The study is unique in that it examines the interaction of depressive symptoms, cognitive impairment, and dementia in an international sample of participants. A critical feature of MIDDEL is its attention to interventions as applied within different health systems. Results will be valid internationally and will contribute to establishing a model for future research within different health systems.

In conclusion, MIDDEL will provide essential knowledge that will inform treatment guidelines aimed at improving the lives of the rapidly rising number of people living with dementia across countries. Building on previous small-scale randomised controlled trials, this large pragmatic effectiveness trial will enhance the use of health technology assessment methodology in the area of non-pharmacological interventions in this area. It is anticipated to have a significant positive impact on people living with dementia, their caregivers, and the health system. Furthermore, it will also open several new lines of

research and development of personalised psychosocial interventions in an area of high and rising public health relevance.

REFERENCES

1. World Health Organization (WHO). Dementia fact sheet 2017 [internet]. Geneva: WHO; 2017. Available from: http://www.who.int/mediacentre/factsheets/fs362/en/
2. Alzheimer's Disease International. World Alzheimer report 2016. London: Alzheimer's Disease International; 2016.
3. Australian Institute of Health and Welfare (AIHW). Dementia in Australia. Canberra: AIHW; 2012.
4. National Centre for Social and Economic Modelling. The economic cost of dementia in Australia 2016-2056. Canberra: University of Canberra; 2017.
5. World Health Organization (WHO). Depression fact sheet 2018 [internet]. Geneva: WHO; 2018. Available from: http://www.who.int/mediacentre/factsheets/fs369/en/
6. Byers AL, Yaffe K. Depression and risk of developing dementia. *Nat Rev Neurol* 2011;7:323–31.
7. Diniz BS, Butters MA, Albert SM, Dew MA, Reynolds CF. Late-life depression and risk of vascular dementia and Alzheimer's disease: systematic review and meta-analysis of community-based cohort studies. *Br J Psychiatry* 2013;202:329–35.
8. Koritsas S, Davidson S, Clarke D, O'Connor D. Diagnosing and treating depressions in nursing home residents: challenges for GPs. *Aust J Prim Health* 2006;12:104–8.
9. Maust DT, Kim HM, Seyfried LS, *et al.* Antipsychotics, other psychotropics, and the risk of death in patients with dementia number needed to harm. *JAMA Psychiatry* 2015;72(5):438–45.
10. National Institute for Health and Care Excellence (NICE). Clinical guideline 42 – Dementia: supporting people with dementia and their carers in health and social care. London, UK: NICE; 2016. Available from: https://www.nice.org.uk/guidance/cg42
11. Koelsch S. Brain correlates of music-evoked emotions. *Nat Rev Neurosci* 2014;15:170–80.
12. Jacobsen J, Stelzer J, Fritz TH, Chetelat G, La Joie R, Turner R. (2015). Why musical memory can be preserved in advanced Alzheimer's disease. *Brain* 2015;138:2438–50.
13. Janata P. The neural architecture of music-evoked autobiographical memories. *Cereb Cortex* 2009;19:2579–94.
14. Ridder H-M, Stige B, Qvale LG, Gold C. Individual music therapy for agitation in dementia: an exploratory randomized controlled trial. *Aging Ment Health* 2013;17:667–78.
15. 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 Geriatr* 2015;15:84.
16. 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 Ment Health* 2017;21:147–55.
17. Pedersen SKA, Andersen PN, Lugo RG, Andreassen M, Sutterlin S. Effects of music on agitation in dementia: a meta-analysis. *Front Psychol* 2017;8.
18. Stige B. Practicing music as mutual care. In: Stige B, Ansdell G, Elefant C, Pavlicevic M, editors. Where music helps. Community music therapy in action and reflection. Farnham, UK: Ashgate; 2010. p. 254–75.
19. van der Steen JT, Smaling HJA, van der Wouden JC, Bruinsma MS, Scholten RJ, Vink AC. Music-based therapeutic interventions for people with dementia. *Cochrane Database Syst Rev* 2018;7:CD003477.

20. Fusar-Poli L, Bieleninik L, Brondino N, Chen XJ, Gold C. The effect of music therapy on cognitive functions in patients with dementia: a systematic review and meta-analysis. *Aging Ment Health* 2017:1–10.
21. Ridder H-M. How can singing in music therapy influence social engagement for people with dementia: insights from polyvagal theory. In: Baker F, Uhlig S., editors. Voicework in music therapy: research and practice. London, UK: Jessica Kingsley; 2011. p. 130–46
22. Theorell T. Psychological health effects of musical experiences: theories, studies and reflections in music health science. London, UK: Springer; 2014.
23. Aalbers S, Fusar-Poli L, Freeman RE, Spreen M, Ket JCF, Vink AC, Maratos A, Crawford M, Chen X, Gold C. Music therapy for depression. *Cochrane Database Syst Rev* 2017;11:CD004517.
24. Zhang Y, Cai J, Hui F, Ren T, Ma H, Zhao Q. Does music therapy enhance behavioral and cognitive function in elderly dementia patients? A systematic review and meta-analysis. *Ageing Res Rev* 2017;35:1–11.
25. Livingston G, Kelly L, Lewis-Holmes E, *et al.* A systematic review of the clinical effectiveness and cost-effectiveness of sensory, psychological and behavioural interventions for managing agitation in older adults with dementia. *Health Technol Assess* 2014;18(39):1–226.
26. Ueda T, Suzukamo Y, Sato M, Izumi S-I. Effects of music therapy on behavioural and psychological symptoms of dementia: a systematic review and meta-analysis. *Ageing Res Rev* 2013;12:628–41.
27. Vasionyte I, Madison G. Musical intervention for patients with dementia: a meta-analysis. *J Clin Nurs* 2013;22:1203–16.
28. Iden KR, Hjorleifsson S, Ruths S. (2011). Treatment decisions on antidepressants in nursing homes: a qualitative study. *Scand J Prim Health Care* 2011;29:252–6.
29. Zhao K, Bai ZG, Bo A, Chi I. A systematic review and meta-analysis of music therapy for the older adults with depression. *Int J Geriatr Psychiatry* 2016;31:1188–98.
30. Bonde LO, Juel K, Ekholm O. Associations between music and health-related outcomes in adult non-musicians, amateur musicians and professional musicians – results from a nationwide Danish study. *Nord J Music Ther.* Forthcoming 2018.
31. Bygren LO, Konlaan BB, Johansson SE. Attendance at cultural events, reading books or periodicals, and making music or singing in a choir as determinants for survival: Swedish interview survey of living conditions. *BMJ* 1996;313:1577–80.
32. Cerejeira J, Lagarto L and Mukaetova-Ladinska E. Behavioral and Psychological Symptoms of Dementia. *Front Neurol* 2012;3:73.
33. Raglio A, Bellelli G, Traficante D, Gianotti M, Ubezio MC, Gentile S, Villani D, Trabucchi M. Efficacy of music therapy treatment based on cycles of sessions: a randomised controlled trial. *Aging Ment Health* 2010;14:900–4.
34. Pavlicevic M, Tsiris G, Wood S, et al. The "ripple effect': Towards researching improvisational music therapy in dementia care homes. *Dementia-International Journal of Social Research and Practice* 2015;14(5):659-79.
35. Campbell MK, Piaggio G, Elbourne DR, Altman DG, Grp C. Consort 2010 statement: extension to cluster randomised trials. *BMJ* 2012;345:e5661.
36. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ* 2008;337:a1655.
37. Boutron I, Altman DG, Moher D, Schulz KF, Ravaud P, Grp CN. CONSORT statement for randomized trials of nonpharmacologic treatments: a 2017 update and a CONSORT extension for nonpharmacologic trial abstracts. *Ann Intern Med* 2017;167:40–7.

38. Loudon K, Treweek S, Sullivan F, Donnan P, Thorpe KE, Zwarenstein M. The PRECIS-2 tool: designing trials that are fit for purpose. *BMJ* 2015;350:h2147.
39. World Health Organization (WHO). ICD-10: International statistical classification of diseases and related health problems. 10th revision. Geneva: WHO; 1992.
40. Diazordaz K, Slowther AM, Potter R, et al. Consent processes in cluster-randomised trials in residential facilities for older adults: a systematic review of reporting practices and proposed guidelines. *BMJ Open* 2013;3(7) doi: 10.1136/bmjopen-2013-003057.
41. Kitwood T, Bredin K. Towards a theory of dementia care: personhood and well-being. *Ageing Soc* 1992;12:269–87.
42. McDermott O, Orrell M, Ridder H-M. (2014). The importance of music for people with dementia: the perspectives of people with dementia, family carers, staff and music therapists. *Ageing Ment Health* 2014;18:706–16.
43. Eickholt J, Geretsegger M, Gold C. Perspectives on research and clinical practice in music therapy for older people with depression. In: Zubala A, Karkou V, editors. Arts therapies in the treatment of depression. Abingdon, UK: Routledge; 2018.
44. Johnson JK, Nápoles AM, Stewart AL, *et al.* Study protocol for a cluster randomized trial of the Community of Voices choir intervention to promote the health and well-being of diverse older adults. *BMC Public Health* 2015;15:1049.
45. McDermott O. The development and preliminary evaluation of CHORD (CHOrus Research in Dementia) Manual. Alzheimer Europe Annual Conference; 2016 2 November; Copenhagen, Denmark.
46. Collins FS, Fleming R. Sound health: An NIH-Kennedy Center initiative to explore music and the mind. *JAMA* 2017;317:2470–1.
47. Moniz-Cook E, Vernooij-Dassen M, Woods R, *et al.* A European consensus on outcome measures for psychosocial intervention research in dementia care. *Aging Ment Health* 2008;12:14–29.
48. Montgomery SA, Åsberg M. A new depression scale designed to be sensitive to change. *Br J Psychiatry* 1979;134:3829.
49. Leontjevas R, Gerritsen DL, Vernooij-Dassen MFJ, Smalbrugge M, Koopmans RTCM. Comparative validation of proxy-based Montgomery-Asberg Depression Rating Scale and Cornell Scale for Depression in Dementia in nursing home residents with dementia. *Am J Geriatr Psychiatry* 2012;20:985–93.
50. Erkkilä J, Punkanen M, Fachner J, Ala-Ruona E, Pöntiö I, Tervaniemi M, Vanhala M, Gold C. Individual music therapy for depression: randomised controlled trial. *Brit J Psychiatry* 2011;*199:*132–9.
51. Colasanti V, Marianetti M, Micacchi F, Amabile GA, Mina C. (2010). Tests for the evaluation of depression in the elderly: a systematic review. *Arch Gerontol Geriatr* 2010;50:227–30.
52. Morris JC. The Clinical Dementia Rating (CDR): Current version and scoring rules. *Neurology* 1993;43:2412–4.
53. Cummings JL. The neuropsychiatric inventory: assessing psychopathology in dementia patients. *Neurology* 1997;48:S10–6.
54. Cummings JL, Ihl R, Herrschaft H, Hoerr R, Tribanek M. Sensitivity to change of composite and frequency scores of the Neuropsychiatric Inventory in mild to moderate dementia. *Int Psychogeriatr* 2013;25:431–8.
55. Kaufer DI, Cummings JL, Ketchel P, Smith V, MacMillan A, Shelley T, Lopez OL, DeKosky ST. Validation of the NPI-Q, a brief clinical form of the Neuropsychiatric Inventory. *J Neuropsychiatry Clin Neurosci* 2000;12:233–9.

56. Orgeta V, Edwards RT, Hounsome B, Orrell M, Woods B. The use of the EQ-5D as a measure of health-related quality of life in people with dementia and their carers. *Qual Life Res* 2015;24:315–24.
57. Logsdon RG, Gibbons LE, McCurry SM, Teri L. Assessing quality of life in older adults with cognitive impairment. *Psychosom Med* 2002;64:510–9.
58. Chisholm D, Knapp MRJ, Knudsen HC, *et al.* Client Socio-Demographic and Service Receipt Inventory - European Version: development of an instrument for international research - EPSILON Study 5. *Brit J Psychiat* 2000;177:S28–33.
59. Cool C, Cestac P, Laborde C, Lebaudy C, Rouch L, Lepage B, *et al.* Potentially inappropriate drug prescribing and associated factors in nursing homes. *J Am Med Dir Assoc* 2014;15:850.e1–9.
60. Halvorsen, KH, Selbæk G, Ruths S. Trends in potentially inappropriate medication prescribing to nursing home patients: comparison of three cross-sectional studies. *Pharmacoepidemiol Drug Saf* 2017;26:192–200.
61. Lester P, Kohen I, Stefanacci RG, Feuerman M. Antipsychotic drug use since the FDA black box warning: Survey of nursing home policies. *J Am Med Dir Assoc* 2011;12:573–7.
62. Auer S, Graessel E, Viereckl C, Kienberger U, Span E, Luttenberger K. Professional Care Team Burden (PCTB) scale – reliability, validity and factor analysis. *Health Qual Life Outcomes* 2015;13:17.
63. McGlothlin AE, Lewis RJ. Minimal Clinically Important Difference defining what really matters to patients. *JAMA* 2014;312:1342–3.
64. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum; 1988.
65. Button KS, Kounali D, Thomas L, et al. Minimal clinically important difference on the Beck Depression Inventory - II according to the patient's perspective. *Psychol Med* 2015;45:3269–79.
66. R Foundation for Statistical Computing. R: a language and environment for statistical computing [computer program]. Version 3.4.1. Vienna, Austria: R Foundation for Statistical Computing; 2017.
67. Twisk JWR. Applied longitudinal data analysis for epidemiology. 2nd ed. Cambridge, UK: Cambridge University Press; 2013.
68. Rizopoulos D. Joint models for longitudinal and time-to-event data: with applications in R. Boca Raton, FL: Chapman & Hall/CRC Biostatistics Series; 2012.
69. Romeo R, Knapp M, McIntosh B, *et al.* Training costs of person centred planning. In: Curtis L, Netten A, editors. Unit costs of health and social care 2006. Canterbury, UK: Personal Social Services Research Unit, University of Kent; 2006. p. 29–33
70. Beecham J. Collecting and estimating costs. In: Knapp MRJ, editor. The economic evaluation of mental health care. Aldershot: Arena; 1995. p. 157–174.
71. Curtis L. Unit costs of health and social care 2013. Canterbury, UK: Personal Social Services Research Unit, University of Kent; 2013.
72. Ukoumunne OC, Gulliford MC, Chinn S, Sterne JAC, Burney PGJ. Methods for evaluating area-wide and organisation-based interventions in health and health care: a systematic review. *Health Technol Assess* 1999;3(5).

Tables

Table 1. Differences and similarities of group music therapy and recreational choir singing

	Group music therapy	Recreational choir singing
Core principles	Affect regulation and attunement Meet psychosocial needs Empathic relationship	Sing familiar songs, learn new songs Cognitive activation Focus on melody, lyrics, and rhythm
Core intentions	Facilitate and improve communication Reduce behavioural and psychological symptoms through regulation of emotions	Facilitate positive experience of self and others Stimulate expression, semantic autobiographic memory, and positive affect
Shared principles and intentions	Use and support remaining faculty of musical reminiscence Tailor to individuals Support social experience, stimulate social and emotional wellbeing	
Proscribed	Push participants to achieve goals	Instrumental improvisation
Dementia inclusion criteria	All levels of dementia, but may be divided to form homogeneous groups	All levels of dementia, but primarily mild to moderately severe dementia; mixed groups possible (inclusiveness)
Group size	Approx. 5	Approx. 10
Qualification of intervention provider	Music therapy degree; skilled musician; member of professional music therapy association or registration body	Skilled musician, choir leading skills and relevant further training

Figure captions

Figure 1. Mechanisms and outcomes of GMT and RCS

Note. GMT – group music therapy; RCS – recreational choir singing.

Figure 2. Flow of participants through the study: Illustration of the study design

Note. CDR – Clinical Dementia Rating; MADRS – Montgomery-Åsberg Depression Rating Scale; GMT – group music therapy; MMSE – Mini-Mental State Examination; MT – music therapy; RCS – recreational choir singing.

Figure 3. Schedule of enrolment, interventions, and assessments

Note. CDR – Clinical Dementia Rating; CSSRI – Client Socio-Demographic and Service Receipt Inventory; d – day; ICD – International Classification of Diseases and Related Health Problems; m – month; MADRS – Montgomery-Åsberg Depression Rating Scale; MMSE – Mini-Mental State Examination; NPI – Neuropsychiatric Inventory; PCTB – Professional Care Team Burden Scale; QOL-AD – Quality of Life-Alzheimer Disease.

Figure 4. Test power as a function of effect size and ICC.

Note. The intraclass correlation coefficient (ICC) describes the relative similarity of participants within units and is typically as low as 0.05 or 0.01;[72] we have added the pessimistic scenario of ICC = 0.10 for completeness only.

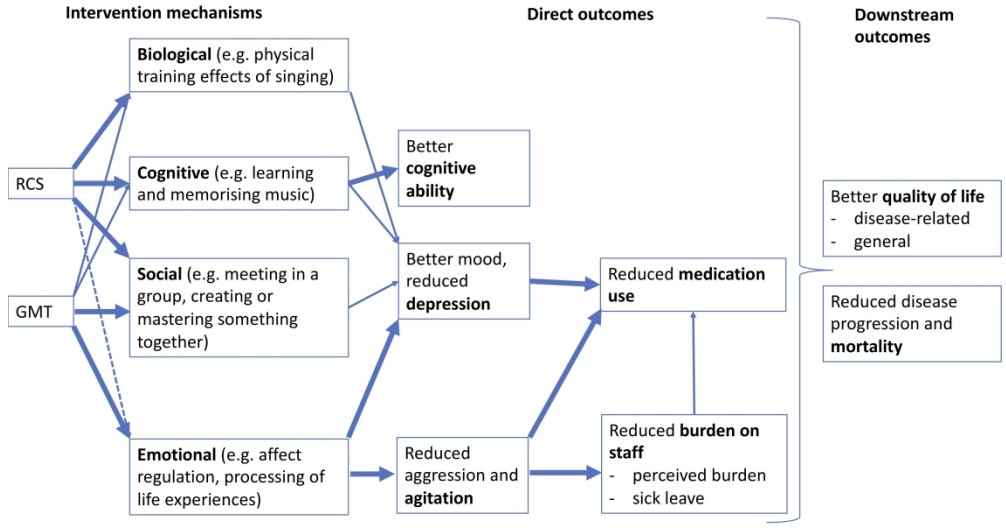

Figure 1. Mechanisms and outcomes of GMT and RCSNote. GMT – group music therapy; RCS – recreational choir singing.

338x190mm (300 x 300 DPI)

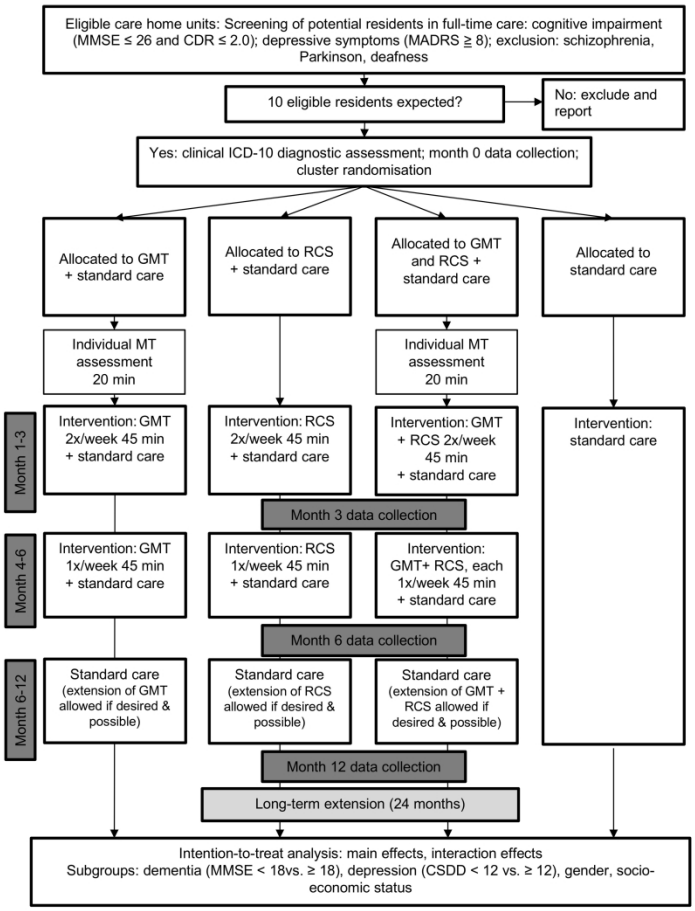

Figure 2. Flow of participants through the study: Illustration of the study design
Note. CDR – Clinical Dementia Rating; MADRS – Montgomery-Åsberg Depression Rating Scale; GMT – group music therapy; MMSE – Mini-Mental State Examination; MT – music therapy; RCS – recreational choir singing.

297x420mm (300 x 300 DPI)

	STUDY PERIOD							
	Enrol-ment	Alloca-tion	Post-allocation					Close-out
TIMEPOINT	-1d	0	1d	3m	6m	12m	...	
ENROLMENT:								
Unit, residents: Eligibility screen	X							
Residents, staff: informed consent (or assent)	X							
Unit: allocation		X						
INTERVENTIONS:								
Group music therapy			←	—	→			
Recreational choir singing			←	—	→			
Standard care			←	—	—	—	→	
ASSESSMENTS:								
Unit baseline: Geographical area Size and costs (CSSRI part 3)	X							
Residents baseline: sociodemographic information (CSSRI part 1); MMSE; dementia diagnosis (ICD-10 code)	X							
Staff baseline: Age, sex	X							
Unit outcomes: Sick leave days			---	---	---	---	---	---------
Residents outcomes: MADRS; CDR; NPI; EQ-5D; QOL-AD; medication and service use (CSSRI part 2)	X			X	X	X	(X)	
Residents outcomes: Adverse events; death			←	—	—	—	—	→
Staff outcomes: PCTB; sick leave days	X			X	X	X	(X)	X

Figure 3. Schedule of enrolment, interventions, and assessments
Note. CDR – Clinical Dementia Rating; CSSRI – Client Socio-Demographic and Service Receipt Inventory; d – day; ICD – International Classification of Diseases and Related Health Problems; m – month; MADRS – Montgomery-Åsberg Depression Rating Scale; MMSE – Mini-Mental State Examination; NPI – Neuropsychiatric Inventory; PCTB – Professional Care Team Burden Scale; QOL-AD – Quality of Life-Alzheimer Disease.

297x420mm (300 x 300 DPI)

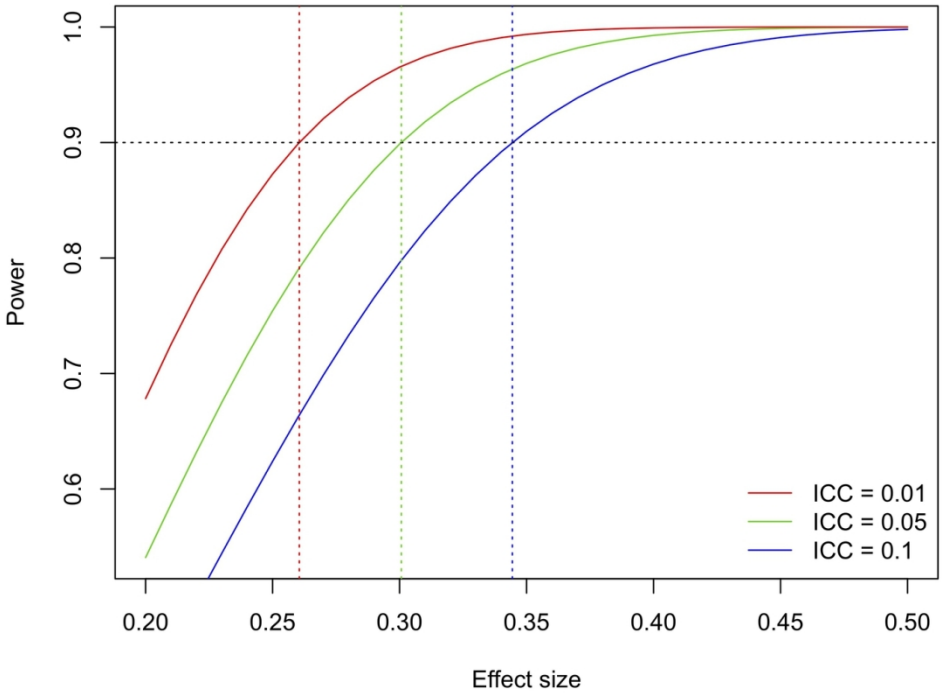

Figure 4. Test power as a function of effect size and ICC.
Note. The intraclass correlation coefficient (ICC) describes the relative similarity of participants within units and is typically as low as 0.05 or 0.01;70 we have added the pessimistic scenario of ICC = 0.10 for completeness only.

125x88mm (300 x 300 DPI)

SPIRIT 2013 Checklist: Recommended items to address in a clinical trial protocol and related documents*

Section/item	Item No	Description	Addressed on page number
Administrative information			
Title	1	Descriptive title identifying the study design, population, interventions, and, if applicable, trial acronym	_____1
Trial registration	2a	Trial identifier and registry name. If not yet registered, name of intended registry	_____3
	2b	All items from the World Health Organization Trial Registration Data Set	_____3
Protocol version	3	Date and version identifier	_____NA (original)
Funding	4	Sources and types of financial, material, and other support	_____2
Roles and responsibilities	5a	Names, affiliations, and roles of protocol contributors	_____1-2
	5b	Name and contact information for the trial sponsor	_____2
	5c	Role of study sponsor and funders, if any, in study design; collection, management, analysis, and interpretation of data; writing of the report; and the decision to submit the report for publication, including whether they will have ultimate authority over any of these activities	_____NA/2
	5d	Composition, roles, and responsibilities of the coordinating centre, steering committee, endpoint adjudication committee, data management team, and other individuals or groups overseeing the trial, if applicable (see Item 21a for data monitoring committee)	_____24

Introduction			
Background and rationale	6a	Description of research question and justification for undertaking the trial, including summary of relevant studies (published and unpublished) examining benefits and harms for each intervention	5-7
	6b	Explanation for choice of comparators	NA
Objectives	7	Specific objectives or hypotheses	8
Trial design	8	Description of trial design including type of trial (eg, parallel group, crossover, factorial, single group), allocation ratio, and framework (eg, superiority, equivalence, noninferiority, exploratory)	8-10
Methods: Participants, interventions, and outcomes			
Study setting	9	Description of study settings (eg, community clinic, academic hospital) and list of countries where data will be collected. Reference to where list of study sites can be obtained	10
Eligibility criteria	10	Inclusion and exclusion criteria for participants. If applicable, eligibility criteria for study centres and individuals who will perform the interventions (eg, surgeons, psychotherapists)	10-11
Interventions	11a	Interventions for each group with sufficient detail to allow replication, including how and when they will be administered	11-14
	11b	Criteria for discontinuing or modifying allocated interventions for a given trial participant (eg, drug dose change in response to harms, participant request, or improving/worsening disease)	11-14
	11c	Strategies to improve adherence to intervention protocols, and any procedures for monitoring adherence (eg, drug tablet return, laboratory tests)	13-14
	11d	Relevant concomitant care and interventions that are permitted or prohibited during the trial	10-11
Outcomes	12	Primary, secondary, and other outcomes, including the specific measurement variable (eg, systolic blood pressure), analysis metric (eg, change from baseline, final value, time to event), method of aggregation (eg, median, proportion), and time point for each outcome. Explanation of the clinical relevance of chosen efficacy and harm outcomes is strongly recommended	14-18
Participant timeline	13	Time schedule of enrolment, interventions (including any run-ins and washouts), assessments, and visits for participants. A schematic diagram is highly recommended (see Figure)	Fig. 3

Sample size	14	Estimated number of participants needed to achieve study objectives and how it was determined, including clinical and statistical assumptions supporting any sample size calculations	18
Recruitment	15	Strategies for achieving adequate participant enrolment to reach target sample size	24-26
Methods: Assignment of interventions (for controlled trials)			
Allocation:			
Sequence generation	16a	Method of generating the allocation sequence (eg, computer-generated random numbers), and list of any factors for stratification. To reduce predictability of a random sequence, details of any planned restriction (eg, blocking) should be provided in a separate document that is unavailable to those who enrol participants or assign interventions	9
Allocation concealment mechanism	16b	Mechanism of implementing the allocation sequence (eg, central telephone; sequentially numbered, opaque, sealed envelopes), describing any steps to conceal the sequence until interventions are assigned	9
Implementation	16c	Who will generate the allocation sequence, who will enrol participants, and who will assign participants to interventions	9
Blinding (masking)	17a	Who will be blinded after assignment to interventions (eg, trial participants, care providers, outcome assessors, data analysts), and how	9
	17b	If blinded, circumstances under which unblinding is permissible, and procedure for revealing a participant's allocated intervention during the trial	NA
Methods: Data collection, management, and analysis			
Data collection methods	18a	Plans for assessment and collection of outcome, baseline, and other trial data, including any related processes to promote data quality (eg, duplicate measurements, training of assessors) and a description of study instruments (eg, questionnaires, laboratory tests) along with their reliability and validity, if known. Reference to where data collection forms can be found, if not in the protocol	14-18
	18b	Plans to promote participant retention and complete follow-up, including list of any outcome data to be collected for participants who discontinue or deviate from intervention protocols	NA

Data management	19	Plans for data entry, coding, security, and storage, including any related processes to promote data quality (eg, double data entry; range checks for data values). Reference to where details of data management procedures can be found, if not in the protocol	24-25
Statistical methods	20a	Statistical methods for analysing primary and secondary outcomes. Reference to where other details of the statistical analysis plan can be found, if not in the protocol	19-23
	20b	Methods for any additional analyses (eg, subgroup and adjusted analyses)	20
	20c	Definition of analysis population relating to protocol non-adherence (eg, as randomised analysis), and any statistical methods to handle missing data (eg, multiple imputation)	20
Methods: Monitoring			
Data monitoring	21a	Composition of data monitoring committee (DMC); summary of its role and reporting structure; statement of whether it is independent from the sponsor and competing interests; and reference to where further details about its charter can be found, if not in the protocol. Alternatively, an explanation of why a DMC is not needed	24-25
	21b	Description of any interim analyses and stopping guidelines, including who will have access to these interim results and make the final decision to terminate the trial	NA/24
Harms	22	Plans for collecting, assessing, reporting, and managing solicited and spontaneously reported adverse events and other unintended effects of trial interventions or trial conduct	17, 24-25
Auditing	23	Frequency and procedures for auditing trial conduct, if any, and whether the process will be independent from investigators and the sponsor	24
Ethics and dissemination			
Research ethics approval	24	Plans for seeking research ethics committee/institutional review board (REC/IRB) approval	25-26
Protocol amendments	25	Plans for communicating important protocol modifications (eg, changes to eligibility criteria, outcomes, analyses) to relevant parties (eg, investigators, REC/IRBs, trial participants, trial registries, journals, regulators)	24-25

Consent or assent	26a	Who will obtain informed consent or assent from potential trial participants or authorised surrogates, and how (see Item 32)	10
	26b	Additional consent provisions for collection and use of participant data and biological specimens in ancillary studies, if applicable	NA
Confidentiality	27	How personal information about potential and enrolled participants will be collected, shared, and maintained in order to protect confidentiality before, during, and after the trial	24-25
Declaration of interests	28	Financial and other competing interests for principal investigators for the overall trial and each study site	2
Access to data	29	Statement of who will have access to the final trial dataset, and disclosure of contractual agreements that limit such access for investigators	NA
Ancillary and post-trial care	30	Provisions, if any, for ancillary and post-trial care, and for compensation to those who suffer harm from trial participation	11
Dissemination policy	31a	Plans for investigators and sponsor to communicate trial results to participants, healthcare professionals, the public, and other relevant groups (eg, via publication, reporting in results databases, or other data sharing arrangements), including any publication restrictions	24, 26-27
	31b	Authorship eligibility guidelines and any intended use of professional writers	NA
	31c	Plans, if any, for granting public access to the full protocol, participant-level dataset, and statistical code	27
Appendices			
Informed consent materials	32	Model consent form and other related documentation given to participants and authorised surrogates	NA
Biological specimens	33	Plans for collection, laboratory evaluation, and storage of biological specimens for genetic or molecular analysis in the current trial and for future use in ancillary studies, if applicable	NA

*It is strongly recommended that this checklist be read in conjunction with the SPIRIT 2013 Explanation & Elaboration for important clarification on the items. Amendments to the protocol should be tracked and dated. The SPIRIT checklist is copyrighted by the SPIRIT Group under the Creative Commons "Attribution-NonCommercial-NoDerivs 3.0 Unported" license.

BMJ Open

Music Interventions for Dementia and Depression in ELderly care (MIDDEL): Protocol and statistical analysis plan for a multinational cluster-randomised trial

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2018-023436.R2
Article Type:	Protocol
Date Submitted by the Author:	26-Dec-2018
Complete List of Authors:	Gold, Christian; NORCE Norwegian Research Centre, GAMUT Eickholt, Jasmin; Hochschule fur angewandte Wissenschaften Wurzburg-Schweinfurt Assmus, Jörg; NORCE Norwegian Research Centre, GAMUT Stige, Brynjulf; NORCE Norwegian Research Centre, GAMUT Wake, Jo; NORCE Norwegian Research Centre Baker, Felicity; Melbourne Conservatorium of Music Tamplin, Jeanette; Melbourne Conservatorium of Music Clark, Imogen; Melbourne Conservatorium of Music Lee, Young-Eun; Melbourne Conservatorium of Music Jacobsen, Stine; Alborg Universitet Institut for Kommunikation Ridder, Hanne Mette; Alborg Universitet Institut for Kommunikation Kreutz, Gunter; Carl von Ossietzky Universitat Oldenburg Institut fur Musik Muthesius, Dorothea; Universitat der Kunste Wosch, Thomas; Hochschule fur angewandte Wissenschaften Wurzburg-Schweinfurt Ceccato, Enrico; Universita degli Studi di Verona Dipartimento di Sanita Pubblica e Medicina di Comunita Raglio, Alfredo; Universita degli Studi di Pavia Sezione di Medicina Legale e Scienze Forensi Antonio Fornari Ruggeri, Mirella; University of Verona, Public Health and Community Medicine Vink, Annemieke; University Medical Centre Groningen, dept. of General Practice Zuidema, Sytse; University Medical Centre Groningen, dept. of General Practice Odell-Miller, Helen; Anglia Ruskin University - Cambridge Campus Orrell, Martin; Mental Health Sciences Schneider, Justine; University of Nottingham, Sociology & Social Policy Kubiak, Christine; ECRIN Romeo, Renee; KCL, Geretsegger, Monika; NORCE Norwegian Research Centre, GAMUT
Primary Subject Heading:	Geriatric medicine
Secondary Subject Heading:	Mental health, Neurology, Patient-centred medicine, Rehabilitation medicine

Keywords:	Dementia < NEUROLOGY, Depression & mood disorders < PSYCHIATRY, Care homes, Music therapy, Music interventions, Old age psychiatry < PSYCHIATRY

Music Interventions for Dementia and Depression in ELderly care (MIDDEL): Protocol and statistical analysis plan for a multinational cluster-randomised trial

Authors

Christian Gold, GAMUT – The Grieg Academy Music Therapy Research Centre, Uni Research Health, Uni Research AS, Bergen, Norway;

Jasmin Eickholt, Faculty of Applied Social Sciences, University of Applied Sciences Würzburg-Schweinfurt, Würzburg, Germany;

Jörg Assmus, Uni Research Health, Uni Research AS, Bergen, Norway;

Brynjulf Stige, GAMUT – The Grieg Academy Music Therapy Research Centre, Uni Research Health, Uni Research AS, Bergen, Norway;

Jo Dugstad Wake, Uni Research Health, Uni Research AS, Bergen, Norway;

Felicity Anne Baker, National Music Therapy Research Unit, Faculty of Fine Arts and Music, University of Melbourne, Melbourne, Australia

Jeanette Tamplin, National Music Therapy Research Unit, Faculty of Fine Arts and Music, University of Melbourne, Melbourne, Australia;

Imogen Clark, National Music Therapy Research Unit, Faculty of Fine Arts and Music, University of Melbourne, Melbourne, Australia;

Young-Eun Claire Lee, National Music Therapy Research Unit, Faculty of Fine Arts and Music, University of Melbourne, Melbourne, Australia;

Stine Lindahl Jacobsen, Department of Communication and Psychology, Aalborg University, Aalborg, Denmark;

Hanne Mette Ochsner Ridder, Department of Communication and Psychology, Aalborg University, Aalborg, Denmark;

Gunter Kreutz, Department of Music, Carl von Ossietzky University, Oldenburg, Germany;

Dorothea Muthesius, Berlin Career College, Universität der Künste, Berlin, Germany;

Thomas Wosch, Faculty of Applied Social Sciences, University of Applied Sciences Würzburg-Schweinfurt, Würzburg, Germany;

Enrico Ceccato, Mental Health Department, Hospital of Montecchio Maggiore, ULSS8 Berica, Vicenza, Italy;

Alfredo Raglio, Music Therapy Research Laboratory, Istituti Clinici Scientifici Maugeri, Pavia, Italy;

Mirella Ruggeri, Section of Psychiatry, University of Verona, Verona, Italy;

Annemiek Vink, Music therapy department, Academy of Music, ArtEZ University of the Arts, Enschede, The Netherlands;

Sytse Zuidema, Department of General Practice and Elderly Care Medicine, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands;

Helen Odell-Miller, CIMTR – The Cambridge Institute for Music Therapy Research at Anglia Ruskin University, Cambridge, UK;

Martin Orrell, School of Medicine and Institute for Mental Health, University of Nottingham, Nottingham, UK;

Justine Schneider, School of Sociology and Social Policy and Institute for Mental Health, University of Nottingham, Nottingham, UK;

Christine Kubiak, Management Office, European Clinical Research Infrastructure Network (ECRIN-ERIC), Paris, France;

Renee Romeo, Health Services and Population Research Department, King's College London, London, UK;

Monika Geretsegger, GAMUT – The Grieg Academy Music Therapy Research Centre, Uni Research Health, Uni Research AS, Bergen, Norway.

Address correspondence to: Christian Gold, GAMUT – The Grieg Academy Music Therapy Research Centre, Uni Research Health, Uni Research AS, Postbox 7805, 5020 Bergen, Norway. Email: christian.gold@norceresearch.no. Phone: +47-97501757.

Acknowledgements: Stephen Clift, Evan Dawson, Jarle Flemvåg, Julene Johnson, Orii McDermott, Valentina Cabral Iversen, Inger Hilde Nordhus, Paolo Pizziolo, Sabine Ruths, Teppo Särkämö, Geir Selbæk, Tanara Sousa, and Antje Timmer provided valuable comments on an earlier version of this protocol. We would like to thank the following patient advisers for feedback on the study purpose and design: Betty Hamilton, a person with dementia living in a care home; Sue Mountain, her daughter; and various people from the Danish Alzheimer Association, Alzheimer Gesellschaft Würzburg/Unterfranken (Germany), the POLYFON user panel (Norway) Dementia UK, Together in Sound/Saffron Hall, and MHA Care Homes (UK). Feedback from the perspective of professional carers was received from Sarah Yeates and various other people at Caladenia Dementia Care Day Centre; Bupa; and BlueCross.

Autors' contributions:

CG took the initiative for the study; CG, JE, JA, BS, HMOR, SZ and RR developed the concept and design; FAB, JT, IC, YCL, GK, DM, TW, EC, AR, MR, AV, HOM, MO, JS, and MG helped to revise the concept and design. CG drafted the manuscript; JE, JA, RR, and MG helped drafting the manuscript. BS, JDW, FAB, JT, IC, YCL, SLJ, HMOR, GK, DM, TW, EC, AR, MR, AV, SZ, HOM, MO, JS, CK revised the manuscript for important intellectual content. CG, FAB, JT, IC, HMOR, and SZ obtained funding. BS, FAB, JT, IC, YCL, SLJ, HMOR, GK, DM, TW, EC, AR, MR, AV, SZ, HOM, and MG were involved in setting up the study conduct in each site. JDW, FAB, JT, IC, YCL, CK, and MG provided administrative, technical, or material support. All authors approved the final version of the manuscript.

Funding: This work was supported by the National Health and Medical Research Council (grant number APP1137853) and by intramural support from Uni Research AS and University Medical Center Groningen.

Competing interests: None declared.
Word count: 7208 (excluding title page, abstract, references, figures and tables)

Abstract

Introduction: In older adults, dementia and depression are associated with individual distress and high societal costs. Music interventions such as group music therapy (GMT) and recreational choir singing (RCS) have shown promising effects, but their comparative effectiveness across clinical subgroups is unknown. This trial aims to determine effectiveness of GMT, RCS, and their combination for care home residents and to examine heterogeneity of treatment effects across subgroups.

Methods and analysis: This large, pragmatic, multinational cluster-randomised controlled trial with a 2x2 factorial design will compare the effects of GMT, RCS, both, or neither, for care home residents aged 65 years or older with dementia and depressive symptoms. We will randomise 100 care home units with ≥ 1000 residents in total across 8 countries. Each intervention will be offered for 6 months (3 months 2x/week followed by 3 months 1x/week), with extension allowed if locally available. The primary outcome will be the change in the Montgomery-Åsberg Depression Rating Scale score at 6 months. Secondary outcomes will include depressive symptoms, cognitive functioning, neuropsychiatric symptoms, psychotropic drug use, caregiver burden, quality of life, mortality, and costs over at least 12 months. The study has 90% power to detect main effects and is also powered to determine interaction effects with gender, severity, and socio-economic status.

Ethics and dissemination: Ethical approval has been obtained for one country and will be obtained for all countries. Results will be presented at national and international conferences and published in scientific journals.

Trial registration numbers: NCT03496675, ACTRN12618000156280

Keywords: group music therapy, recreational choir singing, depression, dementia, non-pharmacological interventions, psychosocial interventions, randomised controlled trial

Strengths and limitations of this study

- As a multinational trial, this study will provide internationally generalisable results concerning the effects of music interventions in older adults with dementia and depression.
- Based on previous small-scale studies, this trial will have adequate power to determine clinical effects as well as to explain variation in treatment effects in relation to patient characteristics.
- A comprehensive set of core outcomes will be measured, including long-term effects in key variables, with assessor blinding where relevant.
- The trial will also enable modelling of trajectories of change and will thereby contribute to an improved understanding of the mechanisms of music interventions.
- Limitations include the potential bias inherent in cluster-randomised studies if recruitment within clusters is incomplete. Due to the nature of the intervention, care providers and participants cannot be blinded, which may bias measures that rely on their reports.

Glossary of terms

- Site: an organisational or geographical entity containing several units, for example a care home/residential care facility.
- Unit (or care home unit; also ‘cluster’): the smallest organisational unit within a site, where residents live together and are cared for together by staff; each unit is randomised.
- Participant: staff or residents within units who have consented to participate.

INTRODUCTION

Dementia and depression are highly prevalent and comorbid conditions in older adults and are associated with individual distress and high and rising societal costs. Globally, around 50 million people were living with dementia in 2017; this number is predicted to reach 82 million in 2030 and 152 million in 2050.[1] The societal costs of dementia are increasing from a total estimated worldwide amount of US$ 818 billion in 2015, about 1.1% of global gross domestic product,[1] to US$ 1 trillion in 2018.[2] Further, the disease's ramifications for families and carers are significant with respect to financial outlay and carer burden.[3] Dementia is highly prevalent among care home residents; more than half of all Australian care home residents in 2016-2017 had dementia.[4]

Depression is the leading cause of disability worldwide.[5] In older adults, it co-occurs and interacts with dementia in complex ways. Depression can cause cognitive impairment and may increase the risk of developing dementia;[6 7] conversely, depression is very common in the early stages of dementia[6] and often exacerbated by admission to a long-term care facility.[8] Psychotropic medication is only a second-line intervention due to limited efficacy and severe adverse effects, including increased mortality from antipsychotics,[9] but is in practice often used to reduce challenging behaviours in later stages of dementia. Non-pharmacological interventions are available and have some supporting evidence, but further research is needed.[10] Among the most promising non-pharmacological approaches to depression and dementia are music interventions, and in the following section we scope out this evidence.

Music interventions for older adults are based on the notion that music elicits emotional responses and helps to retrieve memories,[11] with recent support from research suggesting that brain regions responsible for processing music, particularly known familiar songs, may be spared even in late-stage dementia.[12 13] They are offered in individual,[14 15] group,[16 17], and community settings[18]

and range from targeted clinical interventions offered by trained music therapists to broader recreational activities, which may be facilitated by choir leaders or nursing staff. However, overlaps between the levels of targeting and training do exist. The most common group-based music interventions may be described as group music therapy (GMT) and recreational choir singing (RCS), where GMT is offered by a music therapist and may use a variety of activities ranging from singing through instrumental music making to music listening, whereas RCS is often facilitated by a choir leader and focuses centrally on singing. Putative mechanisms of GMT and RCS can be described as a combination of biological, psychological (cognitive and emotional), and social mechanisms (Figure 1, left part), however with strong overlaps:

- Among the *psychological mechanisms*, *emotional processing*, such as using musical interactions to regulate affects and to reflect on relationships, may be most important in GMT, but is also present to some extent in RCS. *Cognitive processing*, for example through learning and memorising music pieces, is a central mechanism in RCS and less pronounced in GMT, although this may vary between cases, groups, or therapists.[16 19 20]
- *Social mechanisms* are important in both GMT and RCS. The function of the group in itself may be relatively more important in RCS, whereas GMT to a greater extent also relies on the one-to-one relationship between the therapist and each group member. Another important part of the social mechanisms is developing a shared sense of mastery and achievement through learning and performing music pieces, which is more central in RCS than in GMT but may again vary from case to case.[21]
- *Biological mechanisms* include physical training effects of singing and other music-related activities, which may include movement. They are important in both GMT and RCS but may be more central in RCS.[18 22]

Systematic reviews of the clinical effects of GMT and RCS have reported mixed results, [10 17 19 20 23 24 25 26 27] possibly owing to the heterogeneity of treatment effects across types of participants and music interventions. One small trial comparing GMT and RCS directly suggested that the comparative effects of these music interventions may depend on the comorbidity of dementia and depressive symptoms.[16] Process-outcome relations of music interventions may be described as follows (Figure 1, right part):

- *Emotional processing* in a therapist-client relationship may lead to finding meaning and regaining orientation, and thereby to reduced *agitation* and related *neuropsychiatric symptoms*. Such effects have been suggested in some systematic reviews,[17 26] but not in others.[19] Reduced agitation may in consequence reduce *burden on staff*[15] and consequently *sick leave*. This may also help to reduce inappropriate use of *medication*,[14] which is a concern in care homes.[28]
- *Cognitive processing* through practicing music may promote or maintain *cognitive functioning* in older adults. Such effects have been shown for active music therapy, but not music listening, for people with dementia.[20]
- *Emotional processing*, but also social, biological, and cognitive mechanisms may be associated with improved mood and reduced *depressive symptoms*. Systematic reviews have suggested effects of music therapy on depressive symptoms in older adults in general[29] and in people with dementia.[19]
- As downstream outcomes of all four mechanisms and of the intermediate outcomes above, one may expect improved *quality of life*, and possibly reduced *mortality*, although these effects may be small[30 31] and indirect.[9] Music interventions may also reduce costs by reducing time spent on treating neuropsychiatric symptoms and reducing absence by staff.

Hypotheses

Through these different pathways of GMT and RCS, one may hypothesise differential effects for different outcomes, and therefore for different subgroups of care home residents. Specifically:

- GMT may be more effective than no GMT, and RCS may be more effective than no RCS, with respect to reducing depression symptoms and other outcomes shown in Figure 1.
- GMT and RCS may differ in the pattern of effects across outcome domains, which may be explained by their different mechanisms. For example, GMT may be more effective than RCS for reducing aggression and agitation and may therefore be more beneficial for people with late-stage dementia who often present with these neuropsychiatric symptoms.[32 33] RCS may be more effective than GMT with respect to cognitive functioning. Effects on depression symptoms may be achieved through different pathways (Figure 1), and the strength of those effects may therefore depend on severity or comorbidity.[16]
- When offered together, synergistic effects of GMT and RCS may occur through activation of different pathways.
- Cost-effectiveness may differ accordingly across interventions and subgroups. As RCS is likely to be associated with lower intervention costs, it may have better cost-effectiveness ratio in areas where clinical effects are similar; however, this will depend also on each intervention's effects on use of other treatments and services.

METHODS AND ANALYSIS

Design

This large, multinational cluster-randomised controlled trial will be conducted in care homes in Australia, Denmark, Germany, Italy, Netherlands, Norway, Poland, and the UK. The list of study sites is provided in the trial registration record. MIDDEL uses a 2x2 factorial design to examine the

effects of GMT, RCS, both, or neither, for elderly care home residents with dementia and depressive symptoms (Figure 2). This design enables investigating the effects of two music interventions as well as potential synergy effects between them. These may occur between intervention providers on the cluster level (GMT and RCS providers learning from each other) and through residents on the individual or cluster level (participants gaining in different ways from the combination). We will randomise 100 or more care home units (clusters) in eight countries for a total of 1000 or more participants. Recruitment started in July 2018, and primary completion is anticipated for April 2020.

Block randomisation (block size 4 clusters) will be used to ensure that each site will have a balanced distribution between the interventions. The computer-generated randomisation list will be created and kept concealed at the central study office. Only after the eligibility of a care home unit is confirmed and eligible participants (residents and staff) within that unit have formally consented and completed baseline assessment, will site investigators be informed of the randomisation result for that unit. Where possible, a number of care home units will be randomised at the same time, which will further ensure allocation concealment.

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 and RCS. Plain language summaries and consent forms will use neutral wording to maintain equipoise and to avoid expectancy effects. 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 well) 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.

GMT and RCS may entail "ripple effects" beyond the individual participants by leading to changes of the local milieu/culture at the care home unit.[15 34] These will be assessed by measuring objective and perceived burden on care staff. The cluster design is ideally suited for that situation because it facilitates application in a naturalistic setting and avoids some of the problems of individually randomised trials (such as treatment contamination); it also minimises the additional workload for care staff. Trial procedures will be tested in the Australian cohort before applying them in the other countries. The trial will be conducted and reported in accordance with relevant legal frameworks and research guidelines.[35 36 37 38]

Participants

Eligibility is defined on two levels, care home units and individual participants. Participating care home units will be those that are expected to have at least 10 eligible and consenting residents. Care home units that are currently providing music-based interventions as part of their usual care programme will be excluded. Eligible participants will meet all of the following inclusion criteria:

- aged 65 years or older, resident (full-time, 24h/day) at a participating care home;
- dementia as indicated by a Clinical Dementia Rating score of 0.5 to 2 and a Mini-Mental State Examination (MMSE) score of 26 or less;
- at least mild depressive symptoms, as indicated by a Montgomery-Åsberg Depression Rating Scale (MADRS) score of at least 8;
- a clinical diagnosis of dementia according to ICD-10 research criteria;
- have given written informed consent (may be assent by proxy for those unable to provide consent themselves).

Clinical diagnosis will be ascertained by a clinician or researcher, based on the ICD-10 dementia criteria of memory decline; decline in other cognitive abilities; impairment in activities of daily

living; preserved awareness of the environment; decline in emotional control or motivation or change in social behaviour; and more than 6-month duration of memory decline and other cognitive symptoms.[39] People with a known diagnosis of schizophrenia or Parkinson's disease or those who are known to be severely hearing-impaired, in short-term care, or unable to tolerate sitting in a chair for at least part of the sessions, will be excluded. People may however have other clinical diagnoses such as pre-morbid substance use disorders or anxiety disorders. The list of exclusion criteria is intentionally short to ensure generalisability.[38] Residents will always be provided information about the study, and their ability for consent will be assessed directly, before turning to proxies (next of kin/legal representative/carer) for written informed assent. In case of doubt, consent/assent will be provided by both resident and proxy. Residents unable to provide written consent will still be asked if they agree to the interventions and assessments when these begin.[40]

Interventions

Units in all intervention arms will continue with standard care as locally available. In the units allocated to music interventions, GMT, RCS, or both will be provided twice weekly for the first three months, followed by weekly sessions for the next three months. Continuation of GMT and RCS is allowed after that period, depending on local availability. Data on the resources related to the interventions will be measured (number of sessions attended by each participant, duration of each session, non-contact time spent by the intervention provider to prepare or follow up a session, recorded by the provider). The components of standard care provided will also be recorded (see Outcomes).

GMT and RCS sessions will be 45 minutes each. In line with usual practice, GMT may be divided into smaller groups (e.g. around 5 participants, but this may differ across local contexts), whereas RCS may be conducted in larger groups (e.g. with all residents of the unit in one group).

GMT. A core principle of GMT is affect regulation through active, reciprocal music making with the use of singing and musical instruments (Table 1). This facilitates the relationship between the music therapist and the person living with dementia, and between participants in the group. Another core principle of GMT is to meet the psychosocial needs of each individual resident, which in turn is thought to reduce depressive symptoms and anxiety and to stimulate overall social and emotional wellbeing.[41 42 43] GMT aims to work in the "here and now" by responding to participants' immediate emotional expressions, containing them, and incorporating them into meaningful musical expressions for therapeutic gain.[21] GMT is provided by a trained music therapist, who is registered with the appropriate professional association or registration body in his or her country and should also be skilled as a musician. To facilitate individual relationship-building, the music therapist will offer each resident an initial 20-minute assessment with the aim of determining their musical preferences and starting to build individual rapport. The music therapist will also use other sources to determine the participants' musical biography, cultural background, history, personal strengths, resources, and disabilities, and any other information that could be useful to bring into GMT sessions.

RCS. A core principle of RCS is to sing familiar songs and to provide a familiar musical environment for participants (Table 1). Choral singing involves a combination of cognitive, physical, and psychosocial engagement components.[44] Drawing on the psychosocial aspects of a choir setting, RCS in this trial aims to foster connectedness in a group either with other older adults residing in the care homes or family caregivers; emotional wellbeing; and enjoyment of music-making in a group. Biographically and culturally grounded song materials are used with the central goal of stimulating positive experiences shared by groups of individuals. Where participants have engaged in music activities in their past, this may also enable the continuation, as far as possible, of the familiar social experience of music-making in everyday life. Sections of RCS sessions may vary in their focus; for example, sessions may focus on developing familiarity with well-known songs;

learning and developing new material as a group; singing rounds to encourage listening to each other; or offering space for solo singing.[45] The materials can be familiar songs from a range of repertoires, including but not restricted to festive songs (e.g. birthday songs, Christmas carols), folk songs, traditional, classical, or popular songs. The selection of songs can vary from country to country, within and between choir leaders, and may also depend on seasonal and other circumstantial factors. RCS is provided by a skilled musician with choir leading skills.

Training and assessment of treatment fidelity. GMT and RCS providers will receive training and implement intervention guidelines developed in the initial phase of the study. Regular exchange and peer supervision for GMT and RCS providers will be organized in conjunction with guidelines and training. This will include monthly online or in-person meetings between researchers and intervention providers to ensure intervention quality and fidelity, to discuss potential threats that might undermine study quality, and to refine the guidelines accordingly. Intervention providers will also attend weekly staff meetings at intervention sites where possible, to maximise local knowledge transfer and benefit. Manuals for complex interventions need to standardize the quality of interventions to avoid unwarranted variation between therapists and countries while preserving the possibility for meaningful tailoring to local contexts and individuals.[46] This will be addressed by focusing on general principles rather than fixed behaviours. Both GMT and RCS should be tailored to fit the current situation/status of the group and its individual members. Interventionists will be trained at all sites, both through local in-person meetings with all intervention providers at each site and through remote online training across sites. The purpose of this training is to supplement rather than replace the existing training and expertise of intervention providers. For assessment of adherence and competence, providers of GMT and RCS will be video-recorded in 3-4 randomly selected sessions per unit. We will record and analyse the entire session. To avoid performance bias due to the awareness of being videotaped in a selected session, we will use sham video monitoring in other sessions where possible. Videos will be uploaded and stored on a secure central server and

will be available only to those who check treatment fidelity. Two independent researchers will assess the different components used by intervention providers and the degree of person-centeredness (i.e. tailoring of the intervention to the current situation/needs of the group and its members). This process-related data will help us to understand the mechanisms or effective ingredients of each intervention.

Further development. While the description above provides general guidance and will form the basis for fidelity assessment in this study, no consensus guidelines exist for GMT and RCS. Descriptions in the literature vary in many aspects such as: theoretical frame; session structure; specific therapeutic goals; types of musical instruments and materials; inclusion of music listening in addition to active music-making; structured versus improvisational techniques in active music-making; and adaptation/tailoring to reach each person individually. Therefore, flexible manuals, including sets of detailed principles and techniques for GMT and RCS, will be developed and agreed upon by scientific and clinical experts from different countries using a modified Delphi consensus procedure.

Outcomes

The study uses a broad array of resident-, staff-, and unit-level outcomes measured at 3 months, 6 months (primary), and 12 months after randomisation (Figure 3). A long-term extension with later follow-ups is planned separately. Where possible, core outcomes (www.comet-initiative.org) for psychosocial intervention research in dementia care, that are widely used and available across the languages of the trial, were selected.[47]

The primary endpoint will be change in the total score of the Montgomery-Åsberg Depression Rating Scale (MADRS). The MADRS is a 10-item scale where each item is rated from 0 (no abnormality to 6 (severe).[48] In the total sum score ranging from 0 to 60, higher scores indicate higher severity of depressive symptoms. Assessment is based on an interview with the resident

where possible, but where definite answers cannot be elicited from them, all relevant clues as well as information from other sources should be used as a basis for the rating, in line with usual clinical practice.[49] The total time of administration is approximately 20 minutes. The MADRS has been used successfully in previous studies of music interventions.[16 50] It has shown high reliability and validity, and its sensitivity to change compares favourably to other scales evaluating depression severity in this population, such as the Cornell Scale for Depression in Dementia (CSDD).[49 51]

Secondary outcomes will include the following:

- Dementia severity including cognitive and functional performance – Clinical Dementia Rating (CDR), a standard assessment of dementia severity.[52] The CDR is used widely in clinical settings. Its score is derived from a semi-structured interview with the person living with dementia and an appropriate caregiver/relative. It rates impairment in each of 6 cognitive categories (memory, orientation, judgment and problem solving, community affairs, home and hobbies, and personal care). Its score is useful for characterising and tracking a person's level of impairment or dementia: 0 = normal; 0.5 = very mild or questionable dementia; 1 = mild dementia; 2 = moderate dementia; 3 = severe dementia.
- Neuropsychiatric symptoms – Neuropsychiatric Inventory (NPI), "a *de facto* standard for measuring neuropsychiatric symptoms in clinical trials".[47] Developed to assess behaviour in people living with dementia, the NPI has substantial evidence of validity and reliability and has been translated into more than 40 languages.[53 54] The NPI uses a screening approach to minimise administration time, examining and scoring only the domains with positive responses to screening questions. In this study, the NPI – Questionnaire (NPI-Q)[55] will be used; another version specific for nursing homes (NPI-NH) was considered but rejected because it is not available across all languages. The NPI-Q includes 12 domains where if a symptom is present, both its severity (from 1= mild to 3=severe) and the associated distress on caregivers (from

0=Not distressing at all to 5=Extreme or very severe) are assessed by the professional carer who is most familiar with the resident's behaviour. Item scores across the 12 domains are summed, leading to a total severity score from 0 to 36, where higher values represent higher severity. The additional total distress score can range from 0 to 60, also with higher values representing higher distress.[55]

- Generic quality of life – EuroQol (EQ-5D-5L), a generic health utility measure.[47] The standardized, non-disease-specific instrument for evaluating health-related quality of life was developed by the international EuroQol group and is used to derive quality-adjusted life-years (QALYs). It is based on a descriptive system that defines health in the five dimensions mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Each dimension has five response categories from "no problems" to "extreme problems", which are combined using preference weights to form an overall quality of life score ranging from lower than 0 (worse than death) to 1 (best possible). An additional visual analogue scale indicates today's health on a scale from 0 ("The worst health you can imagine") to 100 ("The best health you can imagine"). As most residents will be unable to self-rate the EQ-5D-5L, the rating will rely on the judgment of the carer as a proxy. Careful selection of assessment mode (self/proxy/both) and choice of appropriate proxies is important to ensure the measure's validity in studies of people with dementia.[56]

- Disease-specific quality of life – Quality of Life in Alzheimer's Dementia (QOL-AD).[47 57] This 13-item scale with a self-rating and proxy version has demonstrated sensitivity to psychosocial intervention, correlates with health-utility measures, is widely translated and used internationally and can be used by people with very low MMSE scores. Items such as "Physical health", "Memory", or "Ability to do things for fun" are scored on a scale from 1 (poor) to 4 (excellent), resulting in a total score ranging from 13 (worst) to 52 (best).

- All-cause mortality (time to death), as recorded in official electronic registries.
- Any increase in psychotropic drug use: Data on type (ATC Codes N065, N06) of psychotropic medication used and any increase or decrease over time will be collected from care staff using the 'medication profile' section of a tailored version of the Client Socio-Demographic and Service Receipt Inventory (CSSRI).[58] Available electronic health registry data will be used where possible. Psychotropic medications are sometimes used inappropriately to manage behavioural symptoms of dementia.[59 60 61] An earlier study suggested that music therapy may help prevent increase in medication.[14]
- Costs: Total and component costs of the interventions will be assessed from a societal perspective, including the cost of the intervention as well as statutory health and social care services used, using a tailored version of the CSSRI.[58]
- Any adverse events (safety): No adverse effects of music interventions are known from earlier trials. Intervention providers are trained to work closely with and adapt their interventions to the needs of participants in order to avoid adverse reactions. Because little knowledge exists about what the potential adverse events could be, all types of adverse events and serious adverse events (e.g. unexpected worsening of symptoms), whether related or unrelated to the interventions, will be reported.

Staff-level outcomes will be as follows:

- Subjective perceived burden of care staff: Professional Care Team Burden Scale.[62] The 10-item scale provides a valid and reliable means of obtaining ratings of burden from formal care teams working in care homes in order to evaluate different interventions targeted at the reduction of burden in care teams. Items are scored on a 5-point scale from 0 (strongly disagree) to 4 (strongly agree), yielding a total sum score from 0 to 40, with higher scores indicating higher burden.

- Days on sick leave of care staff, as recorded monthly by the employer.

Sample size and test power

There is no consensus on the minimal clinically important difference (MCID)[63] on the MADRS. Generally, effect sizes in the small- to medium range (i.e. between d = 0.20 and 0.50) may be considered relevant.[64] Effect sizes in that range were also found in a previous trial on GMT and RCS (d = 0.33 at 6 weeks and 0.49 at 12 weeks).[16] Studies of other depression scales have used anchor-based approaches to determine clinically important percent reductions;[65] we will not use such approaches for the primary analyses, but will include an additional responder analysis.[63]

The trial has the multiple aim of identifying main effects of GMT versus no GMT and RCS versus no RCS, interaction effects of GMT and RCS, and predictive effects of clinical characteristics including severity of dementia; severity of depression; gender; and socio-economic differences. (Although individual socio-economic differences tend to become more equal amongst residents in a given care home, they may still exist at the cluster level, as different homes may have different standards; we will use the average cost of living in each care home unit as a cluster-based proxy measure for socio-economic status.) Power for interaction effects and subgroup analyses is difficult to determine because of the unknown distributions and effect sizes of the different variables. Therefore, the power calculation for the primary outcome was based on the main comparisons of GMT versus no GMT and RCS versus no RCS. This approach maximises power by fully exploiting the factorial design. A general two-sided significance level of 5% will be used, leading with Bonferroni adjustment to a marginal two-sided level of 2.5%. The power calculation was adjusted for cluster effects using the intraclass correlation coefficient (ICC, between 0.01 and 0.10, Figure 4), assuming average cluster size 10. It was further assumed that attrition, which may occur due to death, moving to another care home, or withdrawal from the study, will be no higher than 20% overall. With 100 clusters and 1000 participants randomised, 90% power is reached for effect sizes

between 0.25 and 0.35 (Figure 4). Any further increase beyond this sample size will serve heterogeneity of treatment effects analyses.

Statistical analyses

The statistical analyses will use multivariate longitudinal statistical models, which make optimal use of the data by using data from all time points at once and can account for the effects of clustering within care home units and sites. We will use a modified intention-to-treat (ITT) approach using all available data from all participants as randomised, regardless of the intervention actually received. Sensitivity analyses using multiple imputation for missing data will enable a full ITT analysis. Additional per-protocol analysis will address the effects of treatments as actually received and will complement the ITT analyses. All tests in the study will be two-sided. The general significance level is set to 0.05. Since there are two comparisons in the primary analysis (GMT vs. no GMT, RCS vs. no RCS), we will use a marginal Bonferroni level of 0.025. Continuous variables will be screened for normality. All computations will be done using R.[66]

Sociodemographic and clinical baseline properties for the groups will be characterized by descriptive methods (mean (SD), median [range], n (%)) and presented in a table. A similar table will compare those who dropped out versus those who completed the primary outcome.

The primary outcome, change of MADRS score from baseline to 6 months, will be assessed by a linear mixed-effects model (LME).[67] We will fit the unadjusted model for each treatment (RCS vs. no RCS) as well as the multivariate model containing both treatments as predictors both unadjusted and adjusted for the interaction between the treatments.

Secondary analyses of MADRS scores will include the following:

- The development of MADRS in the treatment groups over the entire study period will be assessed by a LME including time, treatment type and the interaction of time and treatment type as fixed effects, and participant nested within cluster as random effects. We will use

both linear and simple contrasts in the time domain because it is not known whether there is a linear association in time. This will be illustrated by a figure showing the predicted mean of MADRS for each treatment type at each time point with confidence intervals.

- The synergy of the two treatments will be assessed by the LME containing both treatments as well as their interaction as predictors. The interaction in the model will estimate the synergy effect.
- The predictive effect of several covariates (severity of dementia; severity of depression; gender; and socio-economic differences) will be assessed as odds ratios using LMEs for each covariate containing time, treatment type and their interaction as well as the covariate and the interaction between the covariate and the treatment type as predictors. The interaction in the model will estimate the predictive effect.

Secondary endpoints will be analysed as for the primary analysis, using LMEs for continuous outcomes (both resident-level and staff-level). Special considerations apply for the following variables:

- Binary outcomes, including response rates (the proportion of residents improved by at least 50% from their baseline MADRS score), prevalence of medication use, and adverse events, will be assessed as odds ratios using generalised linear mixed models (GLMMs) with a logit link function. The predictive effect of several covariates (severity of dementia; severity of depression; gender; and socio-economic differences) will be assessed by GLMMs for each covariate containing time, treatment type and their interaction as well as the covariate and the interaction between the covariate and the treatment type as predictors. The interaction in the model will estimate the predictive effect.
- Count data (days of sick leave) and cost data are more likely to follow a Poisson distribution than a normal distribution and will be analysed using the respective GLMMs.

- Time-to-event data include mortality (time to death of any cause) and will be assessed by Kaplan-Meier and log-rank- or Breslow tests for differences between the treatment types and the hazard ratios at 12 months.
- Loss to follow-up in all other outcomes can be influenced by mortality. Thus, if the survival analysis shows differences between the groups, it will be meaningful to use a joint modelling approach which combines the longitudinal models and the survival analysis.[68]

In addition to analysing effects of interventions as randomised, we will conduct mediator analyses to examine relations between elements of the therapy approach (mechanisms), direct and downstream outcomes, as depicted in Figure 1, using structural equation modelling (SEM).

Cost-effectiveness analysis

Total and component costs of the interventions and the cost-effectiveness of alternative interventions will be assessed from a societal perspective. This perspective will cover the cost of the intervention, statutory health and social care (and voluntary sector) service costs, and costs of unpaid carer support. The cost per session for each of the interventions will be derived employing established approaches used in a compendium of costs and in published studies.[69 70 71] Information on the time inputs by GMT and RCS providers (for running sessions and for other activities) will be obtained and valued using information on the midpoint of the salary scale and employer's national insurance as well as superannuation contributions. The sum of the staffing contributions and allocations for overheads for each session will then be summed, to derive a cost per session. The average number of sessions delivered as part of the intervention will be multiplied to derive a cost per intervention. As there is no clear agreement on how the costs of group interventions should be allocated, we will calculate the cost per session of each intervention on the basis of the participants allocated to each of the groups, regardless of whether or not the participant attended.

Data on statutory services used will be collected using a tailored version of the Client Socio-Demographic and Service Receipt Inventory (CSSRI),[58] which contains data on the use of health and other formal care resources and unpaid care. To service and support data we will attach unit costs reflecting the long-run marginal opportunity costs drawn from available public sources. Costs per unit of measurement for each service type will be taken from country-specific sources. We will adjust country-specific costs to Euros using purchasing power parity methods. Costs and outcomes will be compared for the comparators using extended dominance approaches. In this approach, the four treatment combinations (GMT, RCS, GMT and RCS, no GMT or RCS) will be ranked by cost, and if one is dominated (more expensive and less effective than another), it will be excluded from further analysis, until two therapeutic groups are left on which to explore which of the two groups is most cost-effective. The cost-effectiveness of one arm over another will be compared by calculating incremental cost-effectiveness ratios (ICERs) defined as difference in mean costs (Euros spent) divided by difference in mean effects (QALYs using the EQ-5D-5L; points improved on MADRS and QOL-AD). Cost-effectiveness acceptability curves will be plotted for each cost-outcome combination to show the likelihood of one treatment being seen as cost-effective relative to another for a range of values placed on incremental outcome improvements. Using the net benefit approach, monetary values of incremental effects and incremental costs will be combined, and net benefit (NB) derived as: $NB = \lambda * (effect_b - effect_a) – (cost_b – cost_a)$, where λ is the willingness-to-pay for a unit improvement in effectiveness, and subscript ‘a’ and ‘b’ denote two candidate treatment arms. There is no agreed cross-national willingness-to-pay threshold, and in some countries there is no established threshold at all. Other studies have used a threshold of Euros 50,000 per QALY, and we will consider this in the discussion of the results.

Sensitivity analyses will be conducted to assess the robustness of the results to changes in key parameters. One of the possible concerns is likely to be the sample size. If the sample size in some participating countries is too small, their cost-effectiveness estimates are likely to be unreliable. We

shall therefore consider the added value of pooling the information on costs and outcomes in sensitivity analyses.

Patient and public involvement

The development of the research question and study design was informed by the priorities, experience and preferences of residents and carers. Co-authors in Australia, Denmark, and the UK have been actively involved in user and advocacy organisations in their countries for a long time and have discussed interventions, outcomes and the need for research with them. Relatives and caregivers spoke to the importance of music interventions as a help for carers and people with dementia, and to the need for high-quality evidence on their effects. Relatives and caregivers are important for giving persons with dementia a voice when they cannot speak for themselves.

Co-authors in Australia had significant involvement with residents, as well as with care staff and care home managers, in discussing and piloting aspects of the study design. While the interventions were generally perceived as pleasurable rather than burdening, some of the outcome measures were felt to be burdening and too demanding due to their length or complexity. As a consequence, the longer Cornell Scale for Depression in Dementia was replaced with the shorter MADRS, and a more extensive quality of life scale was removed. Recruitment strategies were discussed and adapted in dialogue with care home staff.

User representatives will continue to be actively involved throughout the conduct of the trial (see next section). Results will be disseminated to residents, relatives, and care staff via care homes. Results will also be disseminated to national user and advocacy organisations.

Monitoring and oversight

One representative of each recruiting institution will be a member in the *Trial Steering Committee (TSC)*. They will be supplemented by other members who are independent of the investigators, their organisations, funders and sponsors. The TSC will include service users or their relatives and

representatives of stakeholder organisations such as Alzheimer Europe and Dementia Australia. The TSC will have regular meetings to closely supervise all aspects of the study, including any protocol amendments, progress of recruitment, and publication plan.

Data quality monitoring will require a risk-based monitoring approach including remote monitoring activities performed centrally and on-site monitoring as needed. The monitoring will be performed according to the monitoring manual to be developed at the beginning of the project. Recruitment and retention rates will be monitored closely to mitigate the risk of slow recruitment. The number of participating care home units in total and in relation to care home units screened; the number of participating care home residents in total and in relation to residents screened of potential participants; and the retention of participants in the study over the trial period will be closely monitored.

A *Data and Safety Monitoring Committee (DSMC)*, consisting of three people with strong methodological and clinical expertise who are not otherwise affiliated with the project or its institutions, will be appointed early in the international trial. The DSMC will receive regular updates on recruitment, uptake of interventions, any unforeseen events, adverse events, and immediate information on serious adverse events from the trial statistician. It will have unblinded access to study data. Meetings with the DSMC will be on a biannual basis and will consist of an open and a closed part. In the open part, the general progress of the trial will be discussed; in the closed part, the DSMC will discuss any safety signals with the trial statistician. If issues arise, the DSMC will recommend to the TSC on appropriate action.

All aspects of the study, from intervention fidelity through recruitment, outcome assessment, database and data quality management, to data and safety monitoring, will be pilot-tested in one country (Australia) before being rolled out internationally. The data of the pilot cohort will be included in the main trial; no statistical adjustments are made because the decision depends only on

feasibility, not on an interim efficacy analysis. Patient-related documents such as the consent form will be tested because they may influence how the study is perceived by potential participants, relatives and staff.

To ensure data quality, a trial database will be set up and maintained using a safe server hosted by Uni Research (UHEADS) and OpenClinica software. UHEADS is a system for safely storing health research data developed by Uni Research AS, that accommodates the safe upload, storage and retrieval of any sensitive research data. OpenClinica is a web-based system for electronic data capture and clinical data management for multicenter clinical trials, which conforms to relevant international standards for health research. Uni Research AS runs an open source version of OpenClinica.

ETHICS AND DISSEMINATION

Ethical aspects

Ethical approval has been obtained from the Medicine and Dentistry Human Ethics Sub-Committee at the University of Melbourne, Australia (approval date: January 12, 2018) and will be obtained from the relevant local institutional human research ethics committee at each international site. Local clinical investigators will work on adaptation of study- and patient-related documentation to meet national ethical requirements.

Risk management

The main risks are: slow recruitment; low fidelity of interventions; and low reliability of outcome measurements. Regarding recruitment, we will rely on clinical investigators with a track record of successful recruitment. Slow recruitment at some sites can be compensated by other sites. Fidelity of interventions will be ensured through clear guidance and ongoing monitoring. Reliability of

outcomes will be facilitated by the selection of widely used, recommended core outcome measures, and assessed through tests of inter-rater reliability.

Publication plan

The report on the main, pre-planned analyses of the primary endpoint and up until the 12-month follow-up will be submitted to a leading medical journal. The report on the long-term extension will also be submitted to a leading medical journal. Further publications may focus on recruitment and retention strategies for international cluster-randomised multicentre trials of complex interventions in non-medical settings; development of an MCID for the MADRS based on an existing anchor question; inter-relations between outcomes and predictive value of early outcomes for later outcomes; clinical descriptions and qualitative research of therapy processes, including qualitative influences on care home staff, their perception of GMT and RCS and their potential "ripple effects"; barriers and facilitators for implementation, using qualitative interviews and surveys; and consensus guidelines for GMT and RCS. The data and meta-data will be stored in a public repository, such as that of the Norwegian Centre for Research Data (NSD).

Relevance and benefit to society

Music interventions are widely used in care homes, and their effects are likely heterogeneous. MIDDEL is designed to provide reliable and generalisable knowledge about effectiveness, mechanisms, and heterogeneity of effects of music interventions. It will also fill knowledge gaps about potential long-term benefits and preconditions for achieving such sustained benefits. The results will drive changes in aged care and will contribute to our understanding of the relation between music and health.

Implications for practice

If MIDDEL shows beneficial effects, differences in scalability need to be considered for successful implementation. GMT requires extensive, specialised music therapy training and is typically

provided in small groups. The number of qualified music therapists is limited; it varies from country to country, but fluctuates around 1 in 100 000 (about 6000 in Europe, http://emtc-eu.com; 5000 in the USA, www.cbmt.org; 500 in Australia, www.austmta.org.au). RCS is more easily scalable as it can be provided by trained musicians and also in larger groups. There are about 1 million choirs and 37 million choir singers in Europe (www.singingeurope.org). MIDDEL will provide the knowledge needed to identify the best targeting of both approaches, as well as contributing to their improvement and standardisation. For example, GMT with its highly person-centred approach may be most beneficial to those with neuropsychiatric symptoms, which are typical at late-stage dementia, whereas the social engagement in RCS may help those at earlier stages, and the combination of both may be best for another subset of residents with more complex needs. The knowledge generated by MIDDEL will thus increase the impact of music interventions in care homes and potentially in related contexts, such as day care centres for people still living at home.

Implications for future research

As a strongly interdisciplinary project building on contributions from medicine, social sciences, and humanities, this trial will contribute to strengthening the collaborations between these fields, which is likely to stimulate new cross-disciplinary investigations. The study is unique in that it examines the interaction of depressive symptoms, cognitive impairment, and dementia in an international sample of participants. A critical feature of MIDDEL is its attention to interventions as applied within different health systems. Results will be valid internationally and will contribute to establishing a model for future research within different health systems.

In conclusion, MIDDEL will provide essential knowledge that will inform treatment guidelines aimed at improving the lives of the rapidly rising number of people living with dementia across countries. Building on previous small-scale randomised controlled trials, this large pragmatic

effectiveness trial will enhance the use of health technology assessment methodology in the area of non-pharmacological interventions in this area. It is anticipated to have a significant positive impact on people living with dementia, their caregivers, and the health system. Furthermore, it will also open several new lines of research and development of personalised psychosocial interventions in an area of high and rising public health relevance.

REFERENCES

1. World Health Organization (WHO). Dementia fact sheet 2017 [internet]. Geneva: WHO; 2017. Available from: http://www.who.int/mediacentre/factsheets/fs362/en/
2. Alzheimer's Disease International. World Alzheimer report 2016. London: Alzheimer's Disease International; 2016.
3. Australian Institute of Health and Welfare (AIHW). Dementia in Australia. Canberra: AIHW; 2012.
4. National Centre for Social and Economic Modelling. The economic cost of dementia in Australia 2016-2056. Canberra: University of Canberra; 2017.
5. World Health Organization (WHO). Depression fact sheet 2018 [internet]. Geneva: WHO; 2018. Available from: http://www.who.int/mediacentre/factsheets/fs369/en/
6. Byers AL, Yaffe K. Depression and risk of developing dementia. *Nat Rev Neurol* 2011;7:323–31.
7. Diniz BS, Butters MA, Albert SM, Dew MA, Reynolds CF. Late-life depression and risk of vascular dementia and Alzheimer's disease: systematic review and meta-analysis of community-based cohort studies. *Br J Psychiatry* 2013;202:329–35.
8. Koritsas S, Davidson S, Clarke D, O'Connor D. Diagnosing and treating depressions in nursing home residents: challenges for GPs. *Aust J Prim Health* 2006;12:104–8.
9. Maust DT, Kim HM, Seyfried LS, *et al.* Antipsychotics, other psychotropics, and the risk of death in patients with dementia number needed to harm. *JAMA Psychiatry* 2015;72(5):438–45.
10. National Institute for Health and Care Excellence (NICE). Clinical guideline 42 – Dementia: supporting people with dementia and their carers in health and social care. London, UK: NICE; 2016. Available from: https://www.nice.org.uk/guidance/cg42
11. Koelsch S. Brain correlates of music-evoked emotions. *Nat Rev Neurosci* 2014;15:170–80.
12. Jacobsen J, Stelzer J, Fritz TH, Chetelat G, La Joie R, Turner R. (2015). Why musical memory can be preserved in advanced Alzheimer's disease. *Brain* 2015;138:2438–50.
13. Janata P. The neural architecture of music-evoked autobiographical memories. *Cereb Cortex* 2009;19:2579–94.
14. Ridder H-M, Stige B, Qvale LG, Gold C. Individual music therapy for agitation in dementia: an exploratory randomized controlled trial. *Aging Ment Health* 2013;17:667–78.
15. 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 Geriatr* 2015;15:84.
16. 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 Ment Health* 2017;21:147–55.
17. Pedersen SKA, Andersen PN, Lugo RG, Andreassen M, Sutterlin S. Effects of music on agitation in dementia: a meta-analysis. *Front Psychol* 2017;8.
18. Stige B. Practicing music as mutual care. In: Stige B, Ansdell G, Elefant C, Pavlicevic M, editors. Where music helps. Community music therapy in action and reflection. Farnham, UK: Ashgate; 2010. p. 254–75.
19. van der Steen JT, Smaling HJA, van der Wouden JC, Bruinsma MS, Scholten RJ, Vink AC. Music-based therapeutic interventions for people with dementia. *Cochrane Database Syst Rev* 2018;7:CD003477.
20. Fusar-Poli L, Bieleninik L, Brondino N, Chen XJ, Gold C. The effect of music therapy on cognitive functions in patients with dementia: a systematic review and meta-analysis. *Aging Ment Health* 2017:1–10.

21. Ridder H-M. How can singing in music therapy influence social engagement for people with dementia: insights from polyvagal theory. In: Baker F, Uhlig S., editors. Voicework in music therapy: research and practice. London, UK: Jessica Kingsley; 2011. p. 130–46
22. Theorell T. Psychological health effects of musical experiences: theories, studies and reflections in music health science. London, UK: Springer; 2014.
23. Aalbers S, Fusar-Poli L, Freeman RE, Spreen M, Ket JCF, Vink AC, Maratos A, Crawford M, Chen X, Gold C. Music therapy for depression. *Cochrane Database Syst Rev* 2017;11:CD004517.
24. Zhang Y, Cai J, Hui F, Ren T, Ma H, Zhao Q. Does music therapy enhance behavioral and cognitive function in elderly dementia patients? A systematic review and meta-analysis. *Ageing Res Rev* 2017;35:1–11.
25. Livingston G, Kelly L, Lewis-Holmes E, *et al.* A systematic review of the clinical effectiveness and cost-effectiveness of sensory, psychological and behavioural interventions for managing agitation in older adults with dementia. *Health Technol Assess* 2014;18(39):1–226.
26. Ueda T, Suzukamo Y, Sato M, Izumi S-I. Effects of music therapy on behavioural and psychological symptoms of dementia: a systematic review and meta-analysis. *Ageing Res Rev* 2013;12:628–41.
27. Vasionyte I, Madison G. Musical intervention for patients with dementia: a meta-analysis. *J Clin Nurs* 2013;22:1203–16.
28. Iden KR, Hjorleifsson S, Ruths S. (2011). Treatment decisions on antidepressants in nursing homes: a qualitative study. *Scand J Prim Health Care* 2011;29:252–6.
29. Zhao K, Bai ZG, Bo A, Chi I. A systematic review and meta-analysis of music therapy for the older adults with depression. *Int J Geriatr Psychiatry* 2016;31:1188–98.
30. Bonde LO, Juel K, Ekholm O. Associations between music and health-related outcomes in adult non-musicians, amateur musicians and professional musicians – results from a nationwide Danish study. *Nord J Music Ther.* Forthcoming 2018.
31. Bygren LO, Konlaan BB, Johansson SE. Attendance at cultural events, reading books or periodicals, and making music or singing in a choir as determinants for survival: Swedish interview survey of living conditions. *BMJ* 1996;313:1577–80.
32. Cerejeira J, Lagarto L and Mukaetova-Ladinska E. Behavioral and Psychological Symptoms of Dementia. *Front Neurol* 2012;3:73.
33. Raglio A, Bellelli G, Traficante D, Gianotti M, Ubezio MC, Gentile S, Villani D, Trabucchi M. Efficacy of music therapy treatment based on cycles of sessions: a randomised controlled trial. *Aging Ment Health* 2010;14:900–4.
34. Pavlicevic M, Tsiris G, Wood S, et al. The "ripple effect': Towards researching improvisational music therapy in dementia care homes. *Dementia-International Journal of Social Research and Practice* 2015;14(5):659-79.
35. Campbell MK, Piaggio G, Elbourne DR, Altman DG, Grp C. Consort 2010 statement: extension to cluster randomised trials. *BMJ* 2012;345:e5661.
36. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ* 2008;337:a1655.
37. Boutron I, Altman DG, Moher D, Schulz KF, Ravaud P, Grp CN. CONSORT statement for randomized trials of nonpharmacologic treatments: a 2017 update and a CONSORT extension for nonpharmacologic trial abstracts. *Ann Intern Med* 2017;167:40–7.
38. Loudon K, Treweek S, Sullivan F, Donnan P, Thorpe KE, Zwarenstein M. The PRECIS-2 tool: designing trials that are fit for purpose. *BMJ* 2015;350:h2147.
39. World Health Organization (WHO). ICD-10: International statistical classification of diseases and related health problems. 10th revision. Geneva: WHO; 1992.

40. Diazordaz K, Slowther AM, Potter R, et al. Consent processes in cluster-randomised trials in residential facilities for older adults: a systematic review of reporting practices and proposed guidelines. *BMJ Open* 2013;3(7) doi: 10.1136/bmjopen-2013-003057.
41. Kitwood T, Bredin K. Towards a theory of dementia care: personhood and well-being. *Ageing Soc* 1992;12:269–87.
42. McDermott O, Orrell M, Ridder H-M. (2014). The importance of music for people with dementia: the perspectives of people with dementia, family carers, staff and music therapists. *Ageing Ment Health* 2014;18:706–16.
43. Eickholt J, Geretsegger M, Gold C. Perspectives on research and clinical practice in music therapy for older people with depression. In: Zubala A, Karkou V, editors. Arts therapies in the treatment of depression. Abingdon, UK: Routledge; 2018.
44. Johnson JK, Nápoles AM, Stewart AL, *et al.* Study protocol for a cluster randomized trial of the Community of Voices choir intervention to promote the health and well-being of diverse older adults. *BMC Public Health* 2015;15:1049.
45. McDermott O. The development and preliminary evaluation of CHORD (CHOrus Research in Dementia) Manual. Alzheimer Europe Annual Conference; 2016 2 November; Copenhagen, Denmark.
46. Collins FS, Fleming R. Sound health: An NIH-Kennedy Center initiative to explore music and the mind. *JAMA* 2017;317:2470–1.
47. Moniz-Cook E, Vernooij-Dassen M, Woods R, *et al.* A European consensus on outcome measures for psychosocial intervention research in dementia care. *Aging Ment Health* 2008;12:14–29.
48. Montgomery SA, Åsberg M. A new depression scale designed to be sensitive to change. *Br J Psychiatry* 1979;134:3829.
49. Leontjevas R, Gerritsen DL, Vernooij-Dassen MFJ, Smalbrugge M, Koopmans RTCM. Comparative validation of proxy-based Montgomery-Asberg Depression Rating Scale and Cornell Scale for Depression in Dementia in nursing home residents with dementia. *Am J Geriatr Psychiatry* 2012;20:985–93.
50. Erkkilä J, Punkanen M, Fachner J, Ala-Ruona E, Pöntiö I, Tervaniemi M, Vanhala M, Gold C. Individual music therapy for depression: randomised controlled trial. *Brit J Psychiatry* 2011;*199:*132–9.
51. Colasanti V, Marianetti M, Micacchi F, Amabile GA, Mina C. (2010). Tests for the evaluation of depression in the elderly: a systematic review. *Arch Gerontol Geriatr* 2010;50:227–30.
52. Morris JC. The Clinical Dementia Rating (CDR): Current version and scoring rules. *Neurology* 1993;43:2412–4.
53. Cummings JL. The neuropsychiatric inventory: assessing psychopathology in dementia patients. *Neurology* 1997;48:S10–6.
54. Cummings JL, Ihl R, Herrschaft H, Hoerr R, Tribanek M. Sensitivity to change of composite and frequency scores of the Neuropsychiatric Inventory in mild to moderate dementia. *Int Psychogeriatr* 2013;25:431–8.
55. Kaufer DI, Cummings JL, Ketchel P, Smith V, MacMillan A, Shelley T, Lopez OL, DeKosky ST. Validation of the NPI-Q, a brief clinical form of the Neuropsychiatric Inventory. *J Neuropsychiatry Clin Neurosci* 2000;12:233–9.
56. Orgeta V, Edwards RT, Hounsome B, Orrell M, Woods B. The use of the EQ-5D as a measure of health-related quality of life in people with dementia and their carers. *Qual Life Res* 2015;24:315–24.
57. Logsdon RG, Gibbons LE, McCurry SM, Teri L. Assessing quality of life in older adults with cognitive impairment. *Psychosom Med* 2002;64:510–9.

58. Chisholm D, Knapp MRJ, Knudsen HC, *et al.* Client Socio-Demographic and Service Receipt Inventory - European Version: development of an instrument for international research - EPSILON Study 5. *Brit J Psychiat* 2000;177:S28–33.
59. Cool C, Cestac P, Laborde C, Lebaudy C, Rouch L, Lepage B, *et al.* Potentially inappropriate drug prescribing and associated factors in nursing homes. *J Am Med Dir Assoc* 2014;15:850.e1–9.
60. Halvorsen, KH, Selbæk G, Ruths S. Trends in potentially inappropriate medication prescribing to nursing home patients: comparison of three cross-sectional studies. *Pharmacoepidemiol Drug Saf* 2017;26:192–200.
61. Lester P, Kohen I, Stefanacci RG, Feuerman M. Antipsychotic drug use since the FDA black box warning: Survey of nursing home policies. *J Am Med Dir Assoc* 2011;12:573–7.
62. Auer S, Graessel E, Viereckl C, Kienberger U, Span E, Luttenberger K. Professional Care Team Burden (PCTB) scale – reliability, validity and factor analysis. *Health Qual Life Outcomes* 2015;13:17.
63. McGlothlin AE, Lewis RJ. Minimal Clinically Important Difference defining what really matters to patients. *JAMA* 2014;312:1342–3.
64. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum; 1988.
65. Button KS, Kounali D, Thomas L, et al. Minimal clinically important difference on the Beck Depression Inventory - II according to the patient's perspective. *Psychol Med* 2015;45:3269–79.
66. R Foundation for Statistical Computing. R: a language and environment for statistical computing [computer program]. Version 3.4.1. Vienna, Austria: R Foundation for Statistical Computing; 2017.
67. Twisk JWR. Applied longitudinal data analysis for epidemiology. 2nd ed. Cambridge, UK: Cambridge University Press; 2013.
68. Rizopoulos D. Joint models for longitudinal and time-to-event data: with applications in R. Boca Raton, FL: Chapman & Hall/CRC Biostatistics Series; 2012.
69. Romeo R, Knapp M, McIntosh B, *et al.* Training costs of person centred planning. In: Curtis L, Netten A, editors. Unit costs of health and social care 2006. Canterbury, UK: Personal Social Services Research Unit, University of Kent; 2006. p. 29–33
70. Beecham J. Collecting and estimating costs. In: Knapp MRJ, editor. The economic evaluation of mental health care. Aldershot: Arena; 1995. p. 157–174.
71. Curtis L. Unit costs of health and social care 2013. Canterbury, UK: Personal Social Services Research Unit, University of Kent; 2013.
72. Ukoumunne OC, Gulliford MC, Chinn S, Sterne JAC, Burney PGJ. Methods for evaluating area-wide and organisation-based interventions in health and health care: a systematic review. *Health Technol Assess* 1999;3(5).

Tables

Table 1. Differences and similarities of group music therapy and recreational choir singing

	Group music therapy	Recreational choir singing
Core principles	Affect regulation and attunement Meet psychosocial needs Empathic relationship	Sing familiar songs, learn new songs Cognitive activation Focus on melody, lyrics, and rhythm
Core intentions	Facilitate and improve communication Reduce behavioural and psychological symptoms through regulation of emotions	Facilitate positive experience of self and others Stimulate expression, semantic autobiographic memory, and positive affect
Shared principles and intentions	Use and support remaining faculty of musical reminiscence Tailor to individuals Support social experience, stimulate social and emotional wellbeing	
Proscribed	Push participants to achieve goals	Instrumental improvisation
Dementia inclusion criteria	All levels of dementia, but may be divided to form homogeneous groups	All levels of dementia, but primarily mild to moderately severe dementia; mixed groups possible (inclusiveness)
Group size	Approx. 5	Approx. 10
Qualification of intervention provider	Music therapy degree; skilled musician; member of professional music therapy association or registration body	Skilled musician, choir leading skills and relevant further training

Figure captions

Figure 1. Mechanisms and outcomes of GMT and RCS

Note. GMT – group music therapy; RCS – recreational choir singing.

Figure 2. Flow of participants through the study: Illustration of the study design

Note. CDR – Clinical Dementia Rating; MADRS – Montgomery-Åsberg Depression Rating Scale; GMT – group music therapy; MMSE – Mini-Mental State Examination; MT – music therapy; RCS – recreational choir singing.

Figure 3. Schedule of enrolment, interventions, and assessments

Note. CDR – Clinical Dementia Rating; CSSRI – Client Socio-Demographic and Service Receipt Inventory; d – day; ICD – International Classification of Diseases and Related Health Problems; m – month; MADRS – Montgomery-Åsberg Depression Rating Scale; MMSE – Mini-Mental State Examination; NPI – Neuropsychiatric Inventory; PCTB – Professional Care Team Burden Scale; QOL-AD – Quality of Life-Alzheimer Disease.

Figure 4. Test power as a function of effect size and ICC.

Note. The intraclass correlation coefficient (ICC) describes the relative similarity of participants within units and is typically as low as 0.05 or 0.01;[72] we have added the pessimistic scenario of ICC = 0.10 for completeness only.

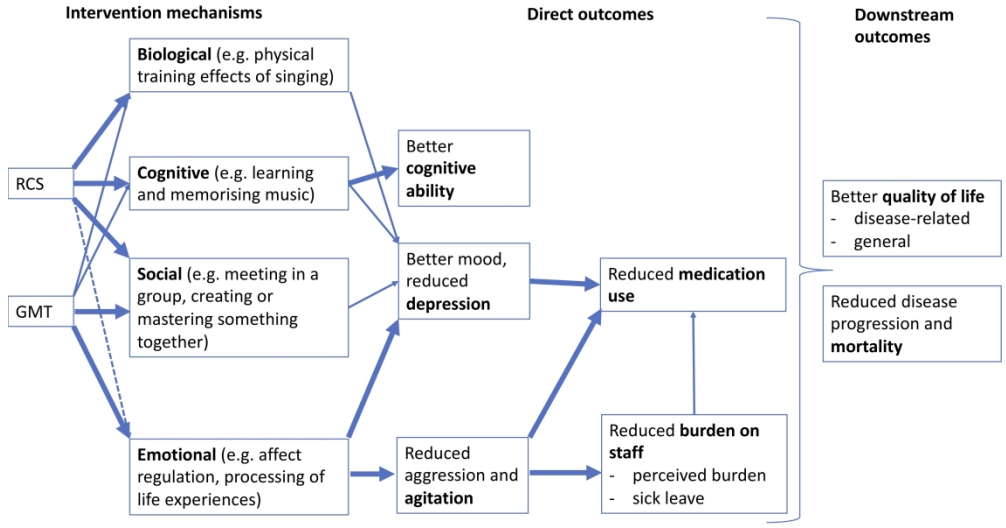

Figure 1. Mechanisms and outcomes of GMT and RCSNote. GMT – group music therapy; RCS – recreational choir singing.

338x190mm (300 x 300 DPI)

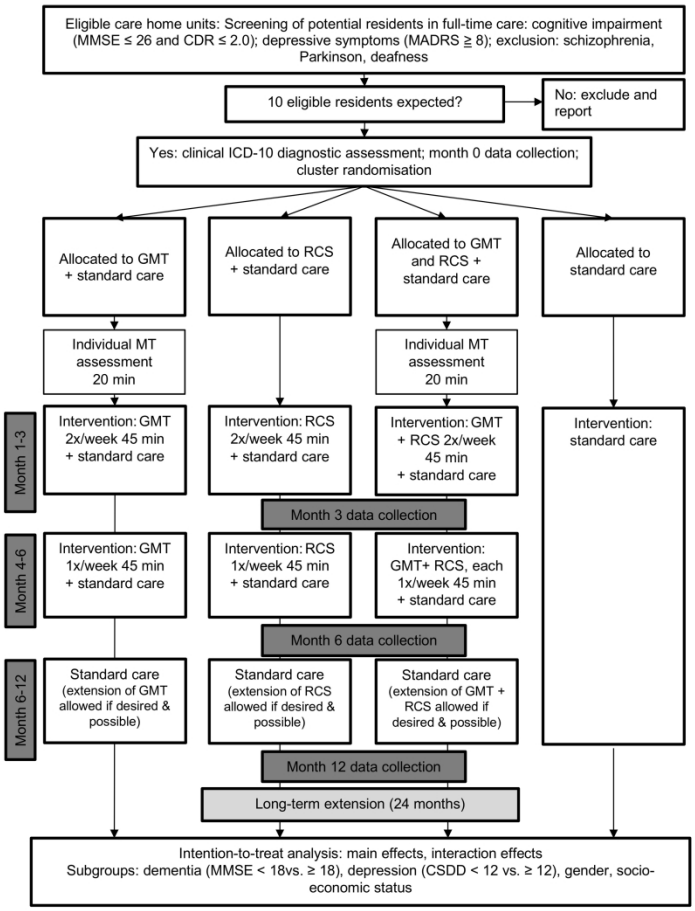

Figure 2. Flow of participants through the study: Illustration of the study design
Note. CDR – Clinical Dementia Rating; MADRS – Montgomery-Åsberg Depression Rating Scale; GMT – group music therapy; MMSE – Mini-Mental State Examination; MT – music therapy; RCS – recreational choir singing.

297x420mm (300 x 300 DPI)

	STUDY PERIOD							
	Enrol-ment	Alloca-tion	Post-allocation					Close-out
TIMEPOINT	*-1d*	0	*1d*	*3m*	*6m*	*12m*	...	
ENROLMENT:								
Unit, residents: Eligibility screen	X							
Residents, staff: informed consent (or assent)	X							
Unit: allocation		X						
INTERVENTIONS:								
Group music therapy			←	—	→			
Recreational choir singing			←	—	→			
Standard care			←	—	—	—	→	
ASSESSMENTS:								
Unit baseline: Geographical area Size and costs (CSSRI part 3)	X							
Residents baseline: sociodemographic information (CSSRI part 1); MMSE; dementia diagnosis (ICD-10 code)	X							
Staff baseline: Age, sex	X							
Unit outcomes: Sick leave days			---	---	---	---	---	---------
Residents outcomes: MADRS; CDR; NPI; EQ-5D; QOL-AD; medication and service use (CSSRI part 2)	X			X	X	X	(X)	
Residents outcomes: Adverse events; death			←	—	—	—	—	→
Staff outcomes: PCTB; sick leave days	X			X	X	X	(X)	X

Figure 3. Schedule of enrolment, interventions, and assessments
Note. CDR – Clinical Dementia Rating; CSSRI – Client Socio-Demographic and Service Receipt Inventory; d – day; ICD – International Classification of Diseases and Related Health Problems; m – month; MADRS – Montgomery-Åsberg Depression Rating Scale; MMSE – Mini-Mental State Examination; NPI – Neuropsychiatric Inventory; PCTB – Professional Care Team Burden Scale; QOL-AD – Quality of Life-Alzheimer Disease.

297x420mm (300 x 300 DPI)

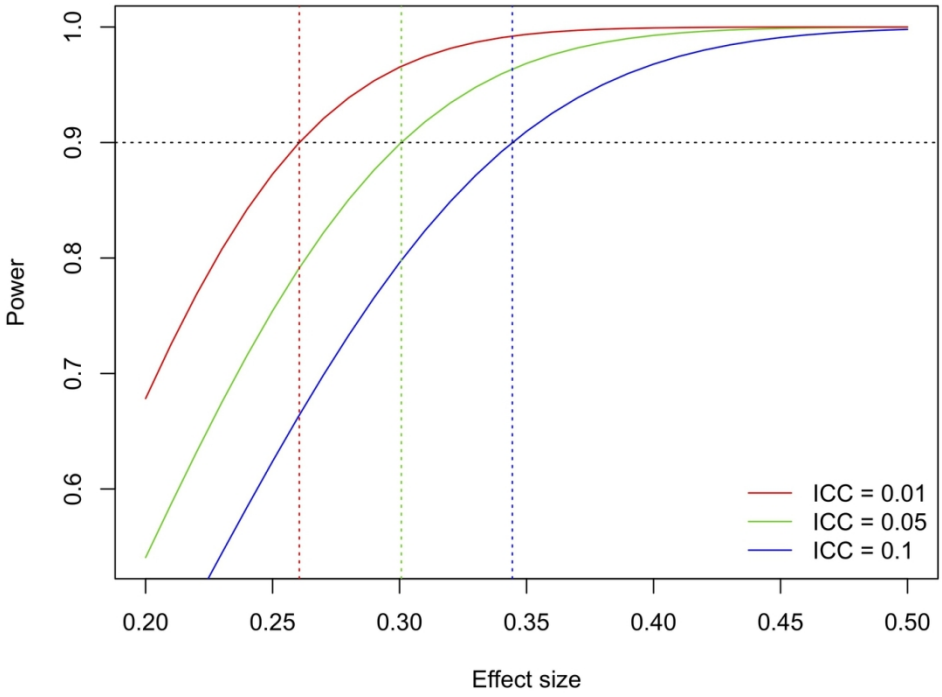

Figure 4. Test power as a function of effect size and ICC.
Note. The intraclass correlation coefficient (ICC) describes the relative similarity of participants within units and is typically as low as 0.05 or 0.01;70 we have added the pessimistic scenario of ICC = 0.10 for completeness only.

125x88mm (300 x 300 DPI)

SPIRIT 2013 Checklist: Recommended items to address in a clinical trial protocol and related documents*

Section/item	Item No	Description	Addressed on page number
Administrative information			
Title	1	Descriptive title identifying the study design, population, interventions, and, if applicable, trial acronym	_____1
Trial registration	2a	Trial identifier and registry name. If not yet registered, name of intended registry	_____3
	2b	All items from the World Health Organization Trial Registration Data Set	_____3
Protocol version	3	Date and version identifier	_____NA (original)
Funding	4	Sources and types of financial, material, and other support	_____2
Roles and responsibilities	5a	Names, affiliations, and roles of protocol contributors	_____1-2
	5b	Name and contact information for the trial sponsor	_____2
	5c	Role of study sponsor and funders, if any, in study design; collection, management, analysis, and interpretation of data; writing of the report; and the decision to submit the report for publication, including whether they will have ultimate authority over any of these activities	_____NA/2
	5d	Composition, roles, and responsibilities of the coordinating centre, steering committee, endpoint adjudication committee, data management team, and other individuals or groups overseeing the trial, if applicable (see Item 21a for data monitoring committee)	_____24

Introduction			
Background and rationale	6a	Description of research question and justification for undertaking the trial, including summary of relevant studies (published and unpublished) examining benefits and harms for each intervention	5-7
	6b	Explanation for choice of comparators	NA
Objectives	7	Specific objectives or hypotheses	8
Trial design	8	Description of trial design including type of trial (eg, parallel group, crossover, factorial, single group), allocation ratio, and framework (eg, superiority, equivalence, noninferiority, exploratory)	8-10
Methods: Participants, interventions, and outcomes			
Study setting	9	Description of study settings (eg, community clinic, academic hospital) and list of countries where data will be collected. Reference to where list of study sites can be obtained	10
Eligibility criteria	10	Inclusion and exclusion criteria for participants. If applicable, eligibility criteria for study centres and individuals who will perform the interventions (eg, surgeons, psychotherapists)	10-11
Interventions	11a	Interventions for each group with sufficient detail to allow replication, including how and when they will be administered	11-14
	11b	Criteria for discontinuing or modifying allocated interventions for a given trial participant (eg, drug dose change in response to harms, participant request, or improving/worsening disease)	11-14
	11c	Strategies to improve adherence to intervention protocols, and any procedures for monitoring adherence (eg, drug tablet return, laboratory tests)	13-14
	11d	Relevant concomitant care and interventions that are permitted or prohibited during the trial	10-11
Outcomes	12	Primary, secondary, and other outcomes, including the specific measurement variable (eg, systolic blood pressure), analysis metric (eg, change from baseline, final value, time to event), method of aggregation (eg, median, proportion), and time point for each outcome. Explanation of the clinical relevance of chosen efficacy and harm outcomes is strongly recommended	14-18
Participant timeline	13	Time schedule of enrolment, interventions (including any run-ins and washouts), assessments, and visits for participants. A schematic diagram is highly recommended (see Figure)	Fig. 3

Sample size	14	Estimated number of participants needed to achieve study objectives and how it was determined, including clinical and statistical assumptions supporting any sample size calculations	18
Recruitment	15	Strategies for achieving adequate participant enrolment to reach target sample size	24-26
Methods: Assignment of interventions (for controlled trials)			
Allocation:			
Sequence generation	16a	Method of generating the allocation sequence (eg, computer-generated random numbers), and list of any factors for stratification. To reduce predictability of a random sequence, details of any planned restriction (eg, blocking) should be provided in a separate document that is unavailable to those who enrol participants or assign interventions	9
Allocation concealment mechanism	16b	Mechanism of implementing the allocation sequence (eg, central telephone; sequentially numbered, opaque, sealed envelopes), describing any steps to conceal the sequence until interventions are assigned	9
Implementation	16c	Who will generate the allocation sequence, who will enrol participants, and who will assign participants to interventions	9
Blinding (masking)	17a	Who will be blinded after assignment to interventions (eg, trial participants, care providers, outcome assessors, data analysts), and how	9
	17b	If blinded, circumstances under which unblinding is permissible, and procedure for revealing a participant's allocated intervention during the trial	NA
Methods: Data collection, management, and analysis			
Data collection methods	18a	Plans for assessment and collection of outcome, baseline, and other trial data, including any related processes to promote data quality (eg, duplicate measurements, training of assessors) and a description of study instruments (eg, questionnaires, laboratory tests) along with their reliability and validity, if known. Reference to where data collection forms can be found, if not in the protocol	14-18
	18b	Plans to promote participant retention and complete follow-up, including list of any outcome data to be collected for participants who discontinue or deviate from intervention protocols	NA

Data management	19	Plans for data entry, coding, security, and storage, including any related processes to promote data quality (eg, double data entry; range checks for data values). Reference to where details of data management procedures can be found, if not in the protocol	24-25
Statistical methods	20a	Statistical methods for analysing primary and secondary outcomes. Reference to where other details of the statistical analysis plan can be found, if not in the protocol	19-23
	20b	Methods for any additional analyses (eg, subgroup and adjusted analyses)	20
	20c	Definition of analysis population relating to protocol non-adherence (eg, as randomised analysis), and any statistical methods to handle missing data (eg, multiple imputation)	20
Methods: Monitoring			
Data monitoring	21a	Composition of data monitoring committee (DMC); summary of its role and reporting structure; statement of whether it is independent from the sponsor and competing interests; and reference to where further details about its charter can be found, if not in the protocol. Alternatively, an explanation of why a DMC is not needed	24-25
	21b	Description of any interim analyses and stopping guidelines, including who will have access to these interim results and make the final decision to terminate the trial	NA/24
Harms	22	Plans for collecting, assessing, reporting, and managing solicited and spontaneously reported adverse events and other unintended effects of trial interventions or trial conduct	17, 24-25
Auditing	23	Frequency and procedures for auditing trial conduct, if any, and whether the process will be independent from investigators and the sponsor	24
Ethics and dissemination			
Research ethics approval	24	Plans for seeking research ethics committee/institutional review board (REC/IRB) approval	25-26
Protocol amendments	25	Plans for communicating important protocol modifications (eg, changes to eligibility criteria, outcomes, analyses) to relevant parties (eg, investigators, REC/IRBs, trial participants, trial registries, journals, regulators)	24-25

Consent or assent	26a	Who will obtain informed consent or assent from potential trial participants or authorised surrogates, and how (see Item 32)	10
	26b	Additional consent provisions for collection and use of participant data and biological specimens in ancillary studies, if applicable	NA
Confidentiality	27	How personal information about potential and enrolled participants will be collected, shared, and maintained in order to protect confidentiality before, during, and after the trial	24-25
Declaration of interests	28	Financial and other competing interests for principal investigators for the overall trial and each study site	2
Access to data	29	Statement of who will have access to the final trial dataset, and disclosure of contractual agreements that limit such access for investigators	NA
Ancillary and post-trial care	30	Provisions, if any, for ancillary and post-trial care, and for compensation to those who suffer harm from trial participation	11
Dissemination policy	31a	Plans for investigators and sponsor to communicate trial results to participants, healthcare professionals, the public, and other relevant groups (eg, via publication, reporting in results databases, or other data sharing arrangements), including any publication restrictions	24, 26-27
	31b	Authorship eligibility guidelines and any intended use of professional writers	NA
	31c	Plans, if any, for granting public access to the full protocol, participant-level dataset, and statistical code	27
Appendices			
Informed consent materials	32	Model consent form and other related documentation given to participants and authorised surrogates	NA
Biological specimens	33	Plans for collection, laboratory evaluation, and storage of biological specimens for genetic or molecular analysis in the current trial and for future use in ancillary studies, if applicable	NA

*It is strongly recommended that this checklist be read in conjunction with the SPIRIT 2013 Explanation & Elaboration for important clarification on the items. Amendments to the protocol should be tracked and dated. The SPIRIT checklist is copyrighted by the SPIRIT Group under the Creative Commons "Attribution-NonCommercial-NoDerivs 3.0 Unported" license.