Psychoneuroendocrinology 182 (2025) 107640

Contents lists available at ScienceDirect

Psychoneuroendocrinology

journal homepage: www.elsevier.com/locate/psyneuen

The short-term impact of music interventions on stress: Results of a multinational cluster-randomized trial using salivary cortisol and alpha-amylase assessments in care home residents with dementia

Naomi L. Rasing [a,b,*,1], Sarah I.M. Janus [a,b,2], Annemieke C. Vink [c,3], Ulrike A.S. Frischen [d], Johanna Neuser [d], Jo Dugstad Wake [e], Nadine Skoluda [f], Vigdis Sveinsdottir [e,4], Monika Geretsegger [f,5], Elias Langeland [f], Gunter Kreutz [c], Christian Gold [f,6], Urs M. Nater [g,7], Sytse U. Zuidema [a,b,8,†]

[a] Department of Primary and Long-term Care, University of Groningen, University Medical Center Groningen, P.O. Box 196, HPC FA21, Groningen, AD 9700, the Netherlands
[b] Alzheimer Center Groningen, University Medical Center Groningen, the Netherlands
[c] ArtEZ Professorship Music-based Therapies and Interventions, Music therapy Department, Academy of Music, ArtEZ University of the Arts, Van Essengaarde 10, Enschede 7511PN, the Netherlands
[d] Department of Music, Speech and Music Lab, Carl von Ossietzky University Oldenburg, Ammerländer Heerstraße 114-118, Oldenburg 26129, Germany
[e] Health & Social Sciences, NORCE Norwegian Research Centre, Nygårdsgaten 112, Bergen 5008, Norway
[f] GAMUT—The Grieg Academy Music Therapy Research Centre, NORCE Norwegian Research Centre, Nygårdsgaten 112, Bergen 5008, Norway
[g] Department of Clinical and Health Psychology, University of Vienna, Liebiggasse 5, Vienna 1010, Austria

ARTICLE INFO

Keywords:
Salivary alpha-amylase
Salivary cortisol
Music-based interventions
Dementia
Stress
Long-term care

ABSTRACT

Objective: Stress can have a negative impact on well-being and quality of life in people with dementia. Our study explored the effect of music as a potentially stress-reducing strategy in care home residents with dementia and depression of 25 care home units included in the multinational cluster-randomized controlled trial Music Interventions for Dementia and Depression in Elderly care (MIDDEL). *Methods:* Group music therapy consisted of singing well-known songs and reminiscence in small groups, optionally improvising on instruments and moving to music. Choir singing consisted of rehearsing well-known and new songs in larger groups of 10 + participants. During month 1, 3, and 6, saliva samples were collected 10 min before and 15 and 60 min after a session of group music therapy ($n = 55$), recreational choir singing ($n = 72$) or a book reading ($n = 56$) in 183 care home residents, along with a stress visual analogue scale. *Results:* Of 1014 saliva samples collected, 671 (66 %) were valid for alpha-amylase assay and 633 samples (62 %) for cortisol assay. Significant pre/post session changes were found for salivary cortisol, alpha-amylase, and subjective stress within study arms, although no significant difference was found between study arms. *Conclusion:* Group interventions, including music therapy, can have a

* Correspondence to: Naomi Rasing, Department of Primary and Long-term Care, University Medical Center Groningen, P.O. Box 196, Groningen, AD 9700, the Netherlands.

E-mail addresses: n.l.rasing@umcg.nl (N.L. Rasing), s.i.m.janus@umcg.nl (S.I.M. Janus), a.vink@artez.nl (A.C. Vink), frischen@uni-bremen.de (U.A.S. Frischen), johanna.neuser@uni-oldenburg.de (J. Neuser), jowa@norceresearch.no (J.D. Wake), nadine.skoluda@univie.ac.at (N. Skoluda), visv@norceresearch.no (V. Sveinsdottir), moge@norceresearch.no (M. Geretsegger), elil@norceresearch.no (E. Langeland), gunter.kreutz@uni-oldenburg.de (G. Kreutz), chgo@norceresearch.no (C. Gold), urs.nater@univie.ac.at (U.M. Nater).

[1] https://orcid.org/0000–0002-9761–3705
[2] https://orcid.org/0000–0003-0419–4407
[3] https://orcid.org/0000–0003-4242–4314
[4] https://orcid.org/0000–0002-9776–4925
[5] https://orcid.org/0000–0002-6219–0060
[6] https://orcid.org/0000–0002-8654–7474
[7] https://orcid.org/0000–0002-2430–5090
[8] https://orcid.org/0000–0002-4991–9507
[†] Deceased

https://doi.org/10.1016/j.psyneuen.2025.107640
Received 24 April 2025; Received in revised form 29 September 2025; Accepted 29 September 2025
Available online 2 October 2025

positive impact on stress levels in care home residents with dementia. The level or intensity of participation required from the person with dementia may play a role in the impact of music on stress. Findings indicate it is feasible to collect saliva in persons with dementia non-invasively. Future studies in this population should consider strategies to improve validity of data, such as stimulating saliva flow. ClinicalTrials.gov registration NCT03496675.

1. Introduction

Stress can have a negative impact on the well-being and quality of life of people with dementia (Sharp, 2019). Ongoing stress can even cause depressive symptoms and depression (C. Linnemann and Lang, 2020). Previous research suggests that active engagement in music such as singing, playing instruments, or moving to music may have a possible stress-reducing effect in people with dementia (Campbell et al., 2022; de Witte et al., 2019, de Witte et al., 2022; Linnemann et al., 2015; Yu et al., 2022).

Developments in biomarker research, in which body signals and body material are monitored or collected, facilitate the measurement of physiological stress in people with dementia, in addition to observational assessments. Saliva, for example, can be collected non-invasively with few instructions for participants (Emami et al., 2022), which makes it a suitable biomarker for measuring stress in people with dementia. However, stress measurements with saliva may be impeded in people with dementia and depressive symptoms in the case of potential confounders, such as chronological age, somatic health and comorbidity, medication use, smoking, food and drink intake, alcohol and juice consumption, physical activity, sleep, and collection method (Strahler et al., 2017). Various data can be extracted from the saliva samples, such as the hormone salivary cortisol (sCort) as well as enzyme salivary alpha-amylase (sAA) levels. Both have been used in music intervention research to measure physiological responses to music, with the sCort being used most often (Wong et al., 2021). Both biomarkers represent a different physiological stress-response system: cortisol concentrations represent activity of the hypothalamic-pituitary-adrenal (HPA) axis, while alpha-amylase activity represent activity of the autonomic nervous system (ANS) (Ali and Nater, 2020). They have opposite diurnal rhythms, but can be collected from the same saliva sample and may represent an acute stress response, both increasing in response to an acute stressor.

Recently, several studies assessed the effect of different music interventions in people with dementia by measuring sCort and/or sAA, including community-based choir singing, which significantly reduced cortisol levels (Dawudi et al., 2024);, a home-based digital music listening intervention, which led to lower morning cortisol levels in family caregivers (Emami et al., 2023; Theorell et al., 2021); and group music therapy, which resulted in decreased cortisol concentrations (De La Rubia Ortí et al., 2018). A systematic review (Sittler et al., 2021) indicated mixed results, with no effect on alpha-amylase, decreased cortisol in one study, but no change in cortisol in two other studies. In a range of clinical populations, results demonstrating the potential impact of music interventions on short-term stress are mixed. To illustrate, sCort concentrations and sAA activity showed a significant change post-intervention in some studies, but not in others (Campbell et al., 2022). Physiological and psychological stress responses are interconnected, but they do not necessarily move in the same direction in response to a stressor, and patterns can vary between individuals (James et al., 2023). Therefore, p Physiological and psychological stress assessments can complement each other, and both provide valuable insights to form a complete picture of stress (Campbell et al., 2022).

The aims of this study were 1) to establish the feasibility of saliva sample collection in care home residents with (often advanced) dementia and depressive symptoms, 2) to detect changes in stress pre/post session, and 3) to detect differences between intervention groups (group music therapy, recreational choir singing, and a book reading control condition). Given the novelty of this study and the fragile nature of the population, no a priori hypothesis was formulated. Instead, the study was designed in an exploratory manner to assess feasibility and to examine potential stress-related and mood-related effects of active music engagement.

2. Methods

2.1. Study design and participants

The current study employed a repeated-measures between-groups design. It was part of the multinational cluster-randomized controlled trial Music Interventions for Dementia and Depression in ELderly care (MIDDEL) (ClinicalTrials.gov NCT03496675), which is described in detail elsewhere (Gold et al., 2019). The main MIDDEL trial included five European countries, of which study sites in three countries had the necessary resources to participate in the present sub study: Germany, the Netherlands and Norway. To assess the impact of music on stress, saliva samples were collected between late-2021 to mid-2023 in a subsample of participants of the MIDDEL trial allocated to either group music therapy (GMT), recreational choir singing (RCS), or the control group (CONTROL) which received care as usual without structural music interventions. Study participants with a dementia diagnosis and mild, undiagnosed depressive symptoms indicated by a Montgomery-Åsberg Depression Rating scale (MADRS) score of at least 8 were included. Both music interventions were offered twice a week during the first three months and once a week during the next three months of the 6-month intervention period. Saliva samples were collected when sessions took place at three different timepoints: approximately one month (T1), three months (T3), and six months (T6) after the start of the intervention period. On each of these days, samples were collected at three moments: approximately 10 min before the start of the session (saliva *a*), 15 min after the end of the session (saliva *b*) and 60 min after the end of the session (saliva *c*). A visual analogue scale of stress (VAS-S) was also completed with each saliva sample. Design choices, methodological considerations and rationale for the current study have been thoroughly described in a published study protocol (Rasing et al., 2022). We used the CONSORT checklist when writing our report (Schulz et al., 2010).

2.2. Clinical baseline measures

Demographic data of participants were collected as part of the main MIDDEL trial, including age, sex, type of dementia diagnosis, dementia severity, severity of depressive symptoms, psychotropic drug prescription and chronic stress. The Clinical Dementia Rating (CDR) scale was used to assess dementia severity. It consists of ratings of impairment in six categories (memory, orientation, judgement and problem solving, community affairs, home and hobbies, and personal care) derived from a semi-structured interview with the person with dementia and an appropriate caregiver, resulting in a score of 0–3 indicating no, mild, moderate or severe dementia (Morris, 1993). The Mini Mental State Examination (MMSE) is a measurement used to assess five areas of cognitive function (orientation, registration, attention and calculation, recall, and language) and consists of 11 questions answered by the participant, with a maximum score of 30 where a score of 23 or lower indicates cognitive impairment (Folstein et al., 1975). The Severe Impairment Battery (SIB-8) consists of eight questions answered by the person with dementia, resulting in a score ranging from 0 to 16, with a

higher score indicating more severe cognitive impairments (Schmitt et al., 2013). Depression severity was assessed with the Montgomery-Åsberg Depression Rating Scale (MADRS), which consists of 10 items that are discussed in an interview with the participant or a proxy, where each item is rated from 0 =no abnormality to 6 =severe (Montgomery and Asberg, 1979).

2.3. Interventions

Group music therapy (GMT) was conducted by a certified music therapist. The session consisted of an opening with a welcome song, followed by singing familiar songs and reminiscing, optional improvising on instruments, optional movement to music, singing familiar songs and reminiscing again, and closing the session with a goodbye song.

Recreational choir singing (RCS) was conducted by a choir leader, also in groups. The session consisted of an opening with a welcome song, singing familiar songs, practicing new material, and closing the session with a goodbye song.

The control group received a book reading session, which was similar to the music interventions in terms of group size (five to ten participants) and duration (45 min) and was offered only on days when saliva was collected. The session was offered by a care staff member or by a researcher, depending on availability of care staff. The reading material was selected by the person who conducted the reading session and consisted of short stories for older adults, fairy tales, stories with a local setting or local famous people as main characters, collections of poems, and children's books.

To minimize potential confounding effects of the circadian rhythm of sCort and sAA, sessions and saliva collection took place early in the afternoon; ideally between 12 a.m. and 5 p.m., with sessions scheduled around 3 p.m.

2.4. Study assessments

2.4.1. Saliva sample collection

To assess short-term effects of the interventions on stress, saliva samples were collected at three time points, as described above. The timepoint *b*, 15 min post-session, can be interpreted as *reactivity* to the intervention; timepoint *c*, 60 min post-session, as *recovery* after the intervention (Thoma et al., 2013). Based on previous research a reactivity time of 15 and recovery time of 60 min were considered appropriate (Narvaez Linares et al., 2020). Saliva was collected using SalivaBio Children's Swabs (Sarstedt, Sevelen, Switzerland); these are long cotton rolls (Salivettes) where the nurse or researcher can hold one end while the participant has the other end in their mouth under the tongue or in the cheek for 60–90 s. This swab is considered easy to use and reduces choking hazards. A survey was filled out to gain background information on factors that could potentially affect saliva compounds. Specifically, care staff were asked to indicate whether residents ate (1 h); drank juice (18 h); consumed caffeine (18 h); used tobacco (18 h); brushed teeth (1 h); chewed gum (24 h); took calming or psychotropic drugs (6 h); or exercised (24 h) prior to collection of sample *a*, between collection of sample *a* – *b*, and *b* – *c*. Due to the fragility of the study population, it was considered neither desirable nor feasible to apply any dietary or behavioral restrictions before saliva collection. sCort and sAA values vary throughout the day and over longer periods of time (Ali and Nater, 2020; Kanikowska et al., 2019; Matsuda et al., 2012). Hence, this saliva collection procedure was repeated at three different days throughout the intervention period: approximately one (T1), three (T3), and six (T6) months after randomization. To assess feasibility, samples were also collected from participants who chose not to attend the session. After collection, saliva samples were stored locally at −80° C until they were transported on dry ice to the University of Vienna for biochemical analysis. These procedures have been described in detail elsewhere (Rasing et al., 2022).

2.4.2. Stress visual analogue scale

In addition to the objective measurement of biomarkers, a self-reported measure to assess experienced stress should be used (Warth et al., 2022), for example the State-Trait Anxiety Inventory, the Perceived Stress Scale, or a Stress Visual Analogue Scale (Wong et al., 2021). Subjective stress was assessed alongside the collection of each saliva sample, using the VAS-S, filled out by the participant themselves or by proxy (e.g., a care staff member familiar with the participant). Selection of the VAS-S had several advantages over the other examples. First, it consists of just one question, enabling assessment of stress observed or experienced in a particular moment. This made it easy and efficient to fill out alongside the collection of saliva samples. Second, it could be filled out by participant and proxy. The level of stress experienced by or observed in a participant in that moment could be indicated on a vertical line of 100 mm ranging from 0 ('The least stress you can imagine') to 100 ('The most stress you can imagine') (Lesage et al., 2012).

2.5. Statistical analysis

Statistical analyses were performed using SPSS version 28 and R version 4.4.2. Participant *demographics* were described for the entire sample and per study arm (count, percentage; mean, SD; median, range, as applicable). *Feasibility assessment* included (1) time of day of session delivery and saliva collection; (2) session attendance of participants; (3) number of samples collected per participant; (4) number of samples with sufficient saliva fluid to be assayed; and (5) frequency of potential confounders such as eating or drinking before the sample collection. To compare the study groups appropriately for descriptive statistics of randomized controlled trials, key demographic and clinical characteristics were presented in a table. No statistical testing was performed to highlight significant differences between groups at baseline, as randomization already ensured that any differences occurred by chance (Moher et al., 2010).

Outcome variables sCort, sAA, and VAS-S were examined for normality and for statistical outliers using graphical methods (scatterplots, violin plots). We excluded outliers that appeared physiologically implausible (sCort $>$ 200 nmol/l). If outcome variables showed right-skewed distributions on the original scale, we used log transformation ($\log10(x + 1)$) for parametric statistical tests, but used original values for descriptive statistics and visualizations for better interpretability. For all outcomes, change from timepoint *a* to *b* (*reactivity*) and change from timepoint *b* to *c* (*recovery*) was assessed. Participants were included in the analyses regardless of how many of the usual sessions assigned they attended (intention-to-treat principle), but were excluded from analysis of effects if they did not attend the session in question, even if their saliva was collected. We displayed mean (SD) for each group and timepoint graphically and numerically. Pre/post-session changes within groups were tested with paired t-tests. *Differences between groups* were modelled using multiple regression analysis at which post-session values (*b* or *c*) of each intervention group (GMT, RCS) were compared to those of CONTROL, controlled for the previous value (*a* or *b*). P-values $<$ .05 were reported as statistically significant and no formal adjustment for multiple testing was made. Finally, we examined correlation (Spearman, Pearson, with 95 % CIs) between VAS-S and biomarker change.

3. Results

3.1. Participant characteristics

A total of 457 care home residents were included in the MIDDEL trial across Germany (n = 161), the Netherlands (n = 176), and Norway (n = 120), and were eligible to participate in this sub study. Prior to randomization to study arms, a subsample of care home residents of 30 care home units were asked for consent to collect saliva samples. Informed consent for saliva sample collection was provided for 207

(45.3 % of eligible participants) care home residents in Germany (n = 62), the Netherlands (n = 110) and Norway (n = 35). After randomization, 24 care home residents were excluded (Germany, n = 12; the Netherlands, n = 5; Norway, n = 7), since their care home unit was allocated to receive both group music therapy and recreational choir singing. In the flowchart (Fig. 1) inclusion of participants and collection of saliva samples is illustrated. *Baseline characteristics* of the 183 care home residents included for the assessment of short-term stress during the MIDDEL trial are presented in Table 1. More than half of the included participants were from the Netherlands (n = 105, 57.4 %). Most were female (n = 141, 77.0 %), and the mean age was 84.5 years (SD = 7.1). Alzheimer's disease was the most prevalent dementia diagnosis (37.2 %). Half of the participants (50.3 %) had a Clinical Dementia Rating score of 2, corresponding to moderate dementia. About a third of the participants (66/180, 37 %; Table 1) had a psychotropic drug prescription.

3.2. Feasibility assessment

(1) The time of day at which the session took place was similar across groups and timepoints, with sessions typically starting between 12 and 2 p.m. Across all arms, saliva *a* was on average collected 3–13 min before the session started; collection of saliva *b* and *c* varied as session duration and end time of the sessions varied. (2) Session attendance was lowest at T1 (71/110, 65 %) and higher at T3 (102/116, 88 %) and T6 (100/114, 88 %; Table 1). (3) A median of 7 out of 9 possible saliva samples were collected from participants (Table 1). For 53 participants, the complete set of 9 samples was collected (26 GMT participants; 13 RCS; 14 CONTROL). At T1, saliva was collected in 131 participants, compared to 122 participants at T3 and 117 participants at T6. (4) Of 1014 unique saliva samples collected, 633 samples (62.4 %) were valid for cortisol assay and 671 samples (66.2 %) were valid for alpha-amylase assay. For sCort one sample (0.1 %) and for sAA 27 samples (2.7 %) could not be analyzed as values were below or above limit of detection. (5) Across all three timepoints, the most prevalent confounding variable was eating within one hour prior to saliva collection (59.1 %–66.7 %). Almost all participants drank caffeinated beverages (85.5 %–92.9 %) and up to three quarters of participants drank juice in the 18 h prior to saliva collection (69.2 %–77.9 %). In less than a quarter of the cases participants took psychotropic drugs six hours prior to saliva

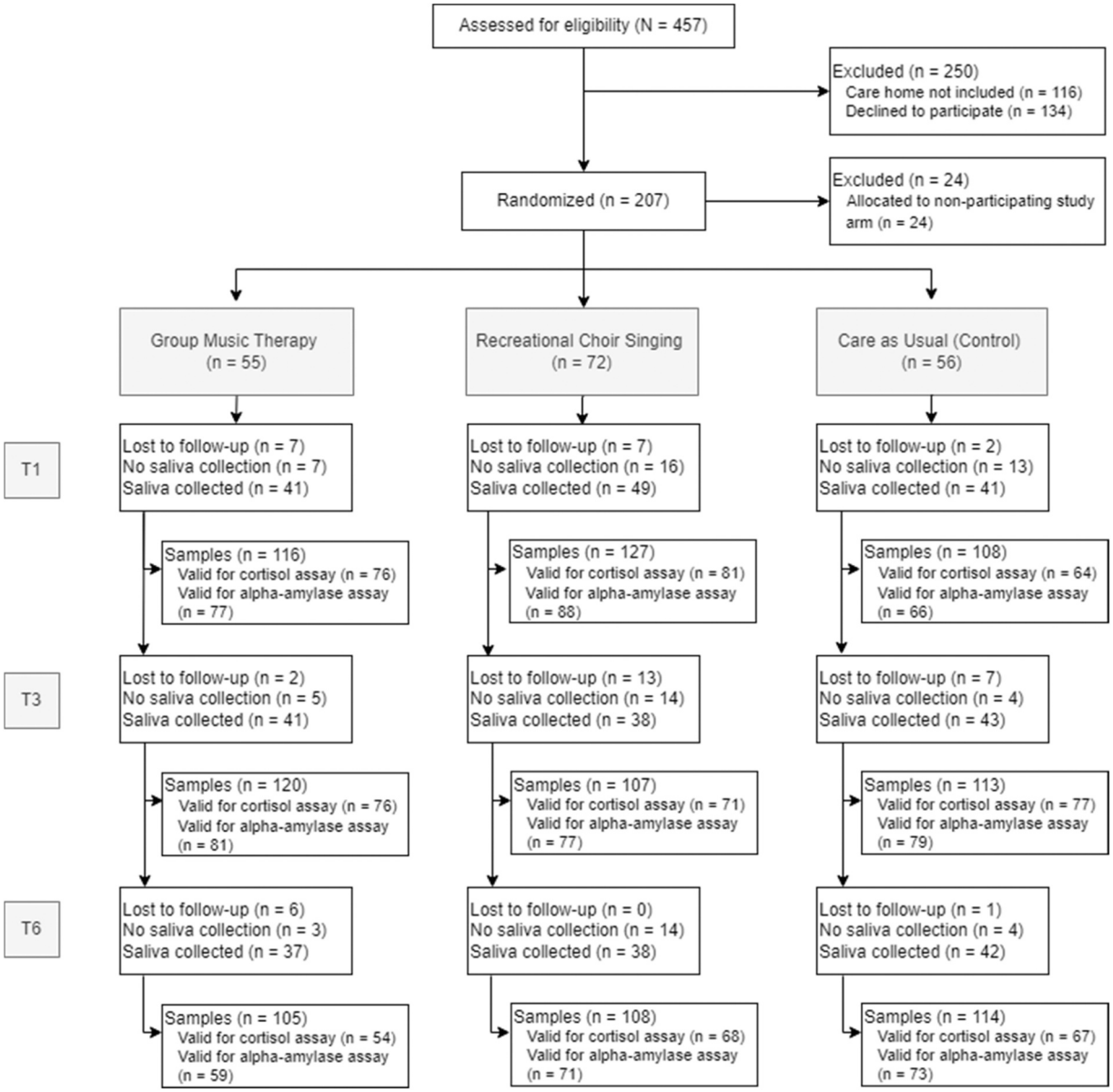

Fig. 1. CONSORT flowchart of care home residents included in the sub study and saliva samples collected throughout the trial.

Table 1
Baseline characteristics of care home residents included in the substudy.

	GMT	RCS	CONTROL	TOTAL
Participant characteristics, *n*	55	72	56	183
Allocated arm, *n* (%)	55 (30.1)	72 (39.3)	56 (30.6)	183 (100.0)
Care home units	9	9	7	25
Age, in years, *n*	55	71	53	179
mean (SD)	84.4 (6.1)	83.4 (7.8)	86.0 (6.4)	84.5 (7.1)
Sex, *n* (female, %)	45 (81.8)	53 (73.6)	43 (76.8)	141 (77.0)
Country, *n*	55	72	56	183
Germany, *n* (%)	16 (29.1)	22 (30.6)	12 (21.4)	50 (27.4)
Netherlands, *n* (%)	30 (54.5)	40 (55.6)	35 (62.5)	105 (57.4)
Norway, *n* (%)	9 (16.4)	10 (13.9)	9 (16.1)	28 (15.3)
Dementia diagnosis, *n*	55	72	53	180
Alzheimer's disease, *n* (%)	27 (49.1)	25 (34.7)	15 (28.3)	67 (37.2)
Vascular dementia, *n* (%)	7 (12.7)	7 (9.7)	7 (13.2)	21 (11.7)
Unspecified dementia, *n* (%)	8 (14.5)	20 (27.8)	9 (17.0)	37 (20.6)
Dementia with Lewy bodies, *n* (%)	2 (3.6)	2 (2.8)	1 (1.9)	5 (2.8)
Other, *n* (%)	11 (20.0)	18 (25.0)	21 (39.6)	50 (27.8)
Clinical Dementia Rating Scale (CDR), *n*	55	69	53	177
CDR = 0.5, *n* (%)			1 (1.9)	1 (0.6)
CDR = 1, *n* (%)	11 (20.0)	16 (23.2)	12 (22.6)	39 (22.0)
CDR = 2, *n* (%)	30 (54.5)	32 (46.2)	27 (50.9)	89 (50.3)
CDR = 3, *n* (%)	14 (25.5)	21 (30.4)	13 (24.5)	48 (27.1)
MMSE score, *n*	55	72	53	180
mean (SD) [range]	9.6 (6.3) [0 – 23]	9.9 (7.1) [0 – 25]	11.5 (6.4) [0 – 22]	10.3 (6.7) [0 – 25]
MADRS score, *n*	54	69	50	173
mean (SD) [range]	17.8 (6.8) [8 – 31]	18.5 (7.7) [8 – 41]	17.5 (4.8) [8 – 38]	18.0 (7.2) [8 – 41]
SIB score, *n*	55	72	53	180
mean (SD) [range]	9.3 (4.8) [0 – 16]	8.6 (5.6) [0 – 16]	9.4 (4.8) [0 – 16]	9.1 (5.1) [0 – 16]
Psychotropic drug prescription, *n* Yes, *n* (%)	55 14 (25.5)	72 31 (43.1)	53 21 (39.6)	180 66 (36.7)
Samples collected, *n*	48	58	49	155
median [IQR]	9 [6 – 9]	6 [3 – 8]	6 [6 – 9]	7 [5 – 9]
Session characteristics	**GMT**	**RCS**	**CONTROL**	**TOTAL**
Session attendance at T1	33	43	34	110
Yes, n (%)	28 (84.8)	21 (48.8)	22 (64.7)	71 (64.5)
No, n (%)	4 (12.1)	13 (30.2)	5 (14.7)	22 (20)
Unknown, n (%)	1 (3.0)	9 (20.9)	7 (20.6)	17 (15.5)
Session attendance at T3	40	38	38	116
Yes, n (%)	38 (95.0)	30 (78.9)	34 (89.5)	102 (87.9)
No, n (%)	2 (5.0)	8 (21.1)	3 (7.9)	13 (11.2)
Unknown, n (%)			1 (2.6)	1 (.86)
Session attendance at T6	36	37	41	114
Yes, n (%)	30 (83.3)	31 (83.8)	39 (95.1)	100 (87.7)
No, n (%)	6 (16.7)	5 (13.5)		11 (9.6)
Unknown, n (%)		1 (2.7)	2 (4.9)	3 (2.6)

collection (12.9 %–23.8 %). Between saliva a and b, right after the session ended, the presence of potential confounders was low. In the 45 min between saliva *b* and *c*, more participants had something to eat (25.9 %–32.4 %) or to drink (34.3 %–47.8 %), such as a caffeinated beverage (11.3 %–23.5 %).

3.3. Differences within intervention groups

Descriptive values of sCort, sAA, and VAS-S before and after the intervention sessions for three timepoints across the six-month intervention period are presented in Table 2. Pre/post-session reactivity (change from timepoint *a* to *b*) and recovery (change from timepoint *b* to *c*) are presented for sCort (Fig. 2), sAA (Fig. 3) and VAS-S (Fig. 4) per study arm (GMT, RCS, book reading control group).

At T1 sCort decreased slightly from *a* to *b* to *c* in GMT and the control group (Fig. 2), in RCS there was no change from *a* to *b*, but sCort decreased during recovery (from *b* to *c*). At T3 sCort decreased in GMT from *a* to *b*, but increased from *b* to *c*. RCS showed a similar pattern, but decreased to a lesser extent from *a* to *b*. The control group showed an opposite pattern, with increased sCort from *a* to *b*, and a decrease from *b* to *c*. At T6 sCort increased slightly in RCS from *a* to *b* to *c*. In GMT there was a steep increase from *b* to *c*. In the control group there is a slight decrease from *a* to *b*, followed by a slight increase from *b* to *c*.

At T1 sAA showed a comparable pattern for GMT and RCS, both increased slightly from *a* to *b*, with a larger increase for GMT than for RCS from *b* to *c*. In contrast, sAA was higher at baseline in the control group compared to the music groups. It decreased from *a* to *b*, with a steep significant decrease from *b* to *c*. At T3 sAA slightly decreased in GMT from *a* to *b*, but increased from *b* to *c*. In RCS sAA showed a steep increase from *a* to *b*, followed by a steep decrease from *b* to *c*. Pre-session sAA was slightly higher in the control group compared to the music groups and showed a decrease from *a* to *b* to *c*. At T6 sAA in GMT decreased from *a* to *b* to *c*. In the control group sAA slightly increased from *a* to *b* and decreased from *b* to *c*. Pre-session sAA was lower in RCS than in GMT and the control group. It increased from *a* to *b*, then decreased slightly from *b* to *c*.

At T1 VAS-S scores decrease in each intervention group from *a* to *b* and increase from *b* to *c*. The decrease in *reactivity* was larger and significant in the control group and the increase from *b* to *c* is lower compared to the music intervention groups. At T3 VAS-S showed a significant decrease in RCS from *a* to *b*, followed by a slight increase from *b* to *c*. In GMT VAS-S scores decreased from *a* to *b* to *c*. VAS-S showed a large increase from *a* to *b* in the control group, followed by a steep decrease during *recovery*. At T6, VAS-S decreased in RCS from a to b to c. In the control group it remained similar from *a* to *b* to *c*. GMT showed a large and significant decrease in VAS-S from *a* to *b*, followed by a steep increase from *b* to *c*.

Pre/post-session *reactivity* (change from timepoint *a* to *b*) within groups showed no significant changes in sCort and sAA (Table 2). For VAS-S *reactivity* three significant within-group reductions were found, at each timepoint in a different study arm: at T1 in the book reading control group $t(16) = -3.47$, $p < .01$; at T3 in RCS $t(26) = -2.13$, $p < .05$; at T6 in GMT $t(26) = -3.64$, $p < .01$ (Fig. 4).

Post-session *recovery* (change from *b* to *c*) within groups showed no significant changes in sCort (Fig. 2). A significant within-group change was found in sAA *recovery* at T1 in the book reading control intervention, $t(11) = -5.66$, $p < .001$, indicating a significant reduction of sAA levels from *b* to *c* (Fig. 3).

3.4. Differences between groups

Pre/post-session *reactivity* between groups showed no significant differences in sCort and sAA (Table 2). For VAS-S two significant between-group differences were found: at T3 reduction from *a* to *b* was significantly more in RCS compared to the control group ($p < .05$). At T6 reduction from *a* to *b* was significantly more in GMT compared to the control group ($p < .01$).

Regarding post-session *recovery* (*b* to *c*) for sCort, sAA, and VAS-S, at T1, T3 and T6, the observed change did not significantly differ between study arms (Table 2).

Table 2
Mean and standard deviation for salivary cortisol (sCort), alpha-amylase (sAA) and VAS-S for three timepoints.

Variable	Pre-session[a]		Post-session[b]				Recovery[c]			
	n	M (SD)	n	M (SD)	p within[d]	p between[e]	n	M (SD)	p within[d]	p between[e]
sCort										
Month 1										
GMT[f]	18	6.71 (5.36)	15	6.19 (5.59)	0.16	0.09	15	6.02 (4.33)	0.06	0.87
RCS[f]	11	6.12 (4.80)	13	6.28 (4.67)	0.66	0.12	9	4.61 (3.07)	0.08	0.82
Book reading[f]	10	6.04 (2.39)	15	5.85 (4.12)	0.16		10	5.09 (2.67)	0.07	
Month 3										
GMT[f]	23	10.73 (18.78)	23	5.98 (4.32)	0.58	0.78	24	7.72 (15.00)	0.84	0.92
RCS[f]	20	6.04 (3.20)	19	4.65 (1.84)	0.43	0.29	21	6.40 (4.68)	0.36	0.28
Book reading[f]	27	6.32 (3.56)	21	9.54 (19.26)	0.56		18	8.36 (11.10)	0.17	
Month 6										
GMT[f]	15	4.92 (3.28)	14	5.92 (4.27)	0.06	0.85	14	12.92 (30.53)	0.31	0.95
RCS[f]	21	4.14 (1.80)	16	5.38 (4.65)	0.67	0.58	22	5.64 (3.59)	0.75	0.51
Book reading[f]	21	6.70 (4.96)	21	5.68 (4.38)	0.23		23	6.39 (9.28)	0.48	
sAA										
Month 1										
GMT[f]	18	277.82 (151.51)	15	298.08 (296.83)	0.33	0.09	16	380.87 (519.14)	0.69	0.01
RCS[f]	13	288.65 (261.86)	14	326.47 (255.45)	0.81	0.44	10	342.90 (250.43)	0.38	0.00
Book reading[f]	9	468.00 (294.42)	16	392.48 (253.91)	0.73		12	146.28 (79.85)	0.00	
Month 3										
GMT[f]	28	304.31 (271.15)	23	293.05 (199.49)	0.47	0.13	24	361.79 (426.02)	0.86	0.54
RCS[f]	21	302.17 (366.61)	19	471.85 (848.66)	0.55	0.58	14	291.50 (299.05)	0.79	0.76
Book reading[f]	26	365.24 (690.98)	21	254.74 (220.89)	0.34		15	243.31 (216.79)	0.41	
Month 6										
GMT[f]	16	367.59 (236.14)	14	347.58 (261.39)	0.83	0.46	17	264.32 (157.17)	0.12	0.21
RCS[f]	21	208.19 (168.83)	19	265.74 (230.16)	0.18	0.40	16	249.42 (237.46)	0.56	0.20
Book reading[f]	26	370.45 (267.11)	20	382.40 (537.04)	0.27		17	266.64 (201.77)	0.83	
VAS-S										
Month 1										
GMT[f]	27	14.63 (19.46)	26	11.54 (19.99)	0.20	0.21	25	15.60 (25.18)	0.37	0.15
RCS[f]	20	11.25 (16.85)	19	7.89 (16.27)	0.36	0.33	20	9.65 (16.45)	0.67	0.64
Book reading[f]	17	13.24 (13.57)	22	3.41 (6.97)	0.00		18	3.89 (8.50)	0.56	
Month 3										
GMT[f]	38	14.08 (15.37)	37	11.49 (18.06)	0.05	0.42	37	11.22 (20.63)	0.44	0.55
RCS[f]	29	14.31 (20.99)	28	7.07 (12.15)	0.04	0.21	29	6.45 (17.40)	0.27	0.93
Book reading[f]	34	8.38 (13.24)	31	14.19 (21.91)	0.60		28	8.04 (14.36)	0.33	
Month 6										
GMT[f]	29	13.62 (17.77)	27	5.00 (9.71)	0.00	0.07	27	12.59 (17.89)	0.15	0.13
RCS[f]	31	10.65 (22.57)	30	7.67 (19.02)	0.11	0.49	30	4.67 (10.74)	0.67	0.72
Book reading[f]	37	10.19 (16.76)	36	9.64 (17.45)	0.26		35	9.57 (20.66)	0.27	

[a] 10 min before start.
[b] 15 min after end.
[c] 1 h after post-session.
[d] p-value from paired *t*-test comparing current time point with previous time point within the group.
[e] p-value from regression model comparing intervention groups with book reading, controlled for the previous time point.
[f] Sessions lasted 45 min.

3.5. Sensitivity analyses

Pearson's product-moment correlation indicated no significant correlation between VAS-S and sCort, $r(620) = .07, p = .082$; between VAS-S and sAA, $r(620) = -.01$, $p = .787$. Similar results were found for Spearman. A moderate positive correlation was found between sAA and sCort, $r(555) = .10, p = .014$. A scatterplot showed no clear patterns. As a sensitivity analysis, participants who ate prior to saliva collection, an important confounding variable, were excluded. This led to similar or lower correlations between sAA and sCort.

4. Discussion

This study assessed the stress-reducing effect of two distinct music interventions, group music therapy (GMT) and recreational choir singing (RCS), in participants with dementia and depression using salivary biomarkers cortisol (sCort) and alpha-amylase (sAA), and a Visual Analogue Scale for stress (VAS-S). Overall, there were no clear patterns in the three outcomes before and after the intervention sessions and across timepoints. Pre/post-session *reactivity* within groups showed no significant differences in sCort and sAA. VAS-S *reactivity* showed a significant reduction within intervention groups on three occasions, at each timepoint in a different group. sCort and VAS-S showed no significant changes in *recovery* within intervention groups. Post-session sAA *recovery* decreased significantly following a book reading session in the control group. No significant differences in pre/post-session *reactivity* in sCort and sAA were found between study arms. VAS-S *reactivity* showed a significant reduction in RCS at T3 and in GMT at T6, compared to the control group. No significant between-group differences were found during *recovery*. Non-invasive collection of saliva samples from people with advanced dementia is feasible, acceptable, and safe. However, VAS-S scores indicated significant changes, even when the assessment of physiological salivary biomarkers of stress did not. These findings indicate a difference between physiological stress measured with salivary biomarkers sAA and sCort on the one hand, and subjective stress as experienced by the person with dementia or observed by care staff (VAS-S) on the other hand. In our sensitivity analysis we found no significant correlation between VAS-S and salivary stress biomarkers. This aligns with the researchers' observations that filling out the VAS-S did not appear consistent across care staff members, participants and timepoints. People with dementia often did not understand the term "stress," and care staff, unable to detect changes, sometimes appeared to complete the scale arbitrarily.

First, it should be considered that a strong correlation between

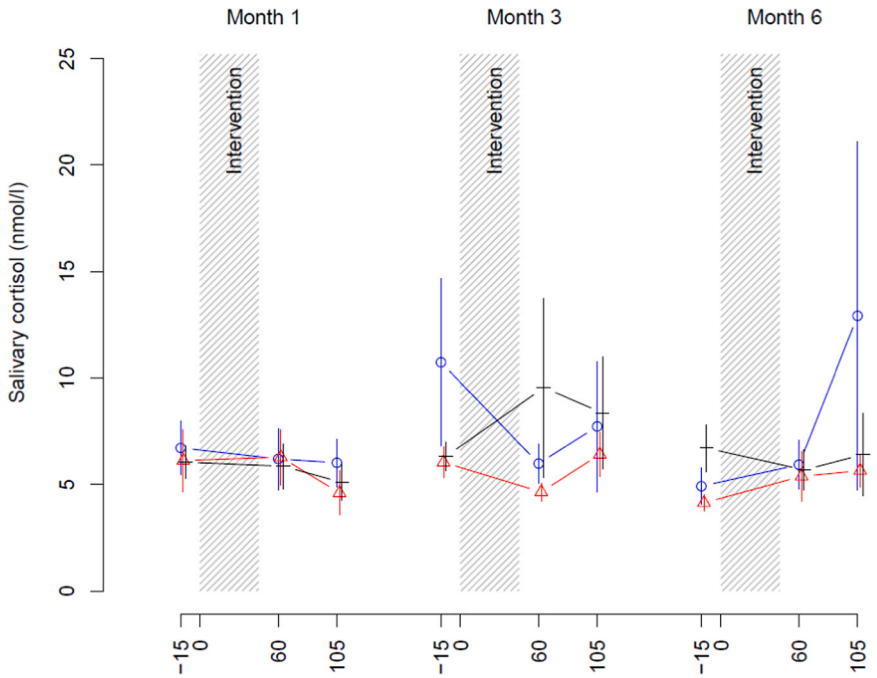

Fig. 2. sCort values pre/post-session and recovery. *Note.* Mean ± SEM. Blue – group music therapy, red – recreational choir singing, black – book reading. Time on x-axis in minutes. *** p < .001, ** p < .01, * p < .05 (left side: within-group change from previous measurement; right side: between-group change compared to book reading, adjusted for previous measurement).

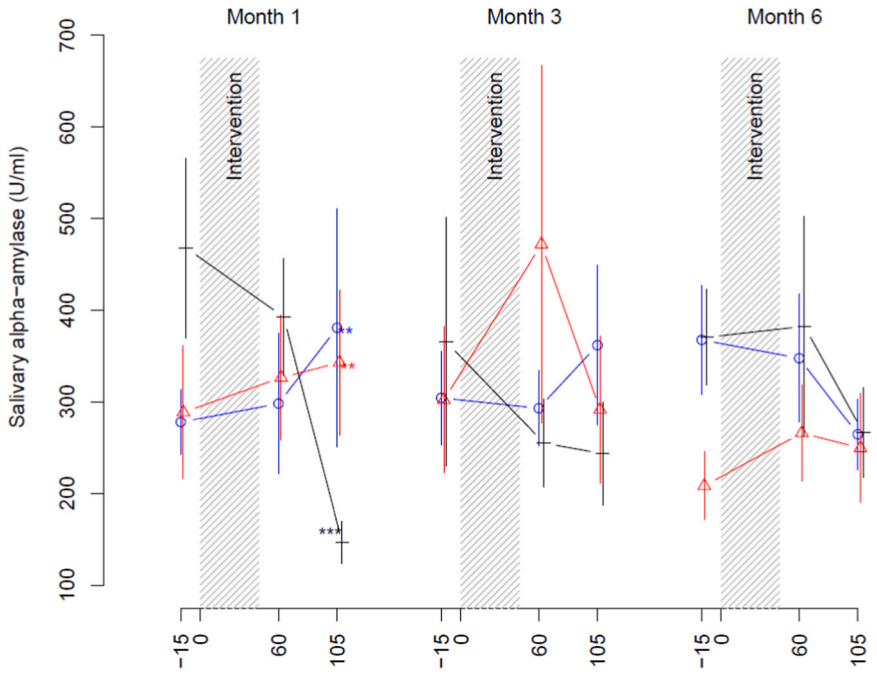

Fig. 3. sAA values pre/post-session and recovery. *Note.* Mean ± SEM. Blue – group music therapy, red – recreational choir singing, black – book reading. Time on x-axis in minutes. *** p < .001, ** p < .01, * p < .05 (left side: within-group change from previous measurement; right side: between-group change compared to book reading, adjusted for previous measurement).

subjectively reported stress and endocrine markers, such as cortisol, may not necessarily be expected. Previous studies (e.g., Fischer et al., 2000; Het et al., 2012) have suggested that cortisol might play a role in dampening the subjective experience of stress, even when physiological activation is present. This potential regulatory effect could partly obscure the relationship between objective stress responses and subjective reports. As such, the absence of a strong correlation in our data may reflect this complex interplay, rather than a lack of physiological stress reactivity. A similar pattern might also influence the relationship with salivary amylase, although this remains speculative.

Individuals with dementia may exhibit a diminished or altered endocrinological response to interventions. As suggested by Emami et al. (2024), this blunted responsiveness could stem from pharmacological factors—such as the use of medications that modulate stress pathways—as well as from the underlying neuropathology of dementia itself. The condition may be associated with increased physiological

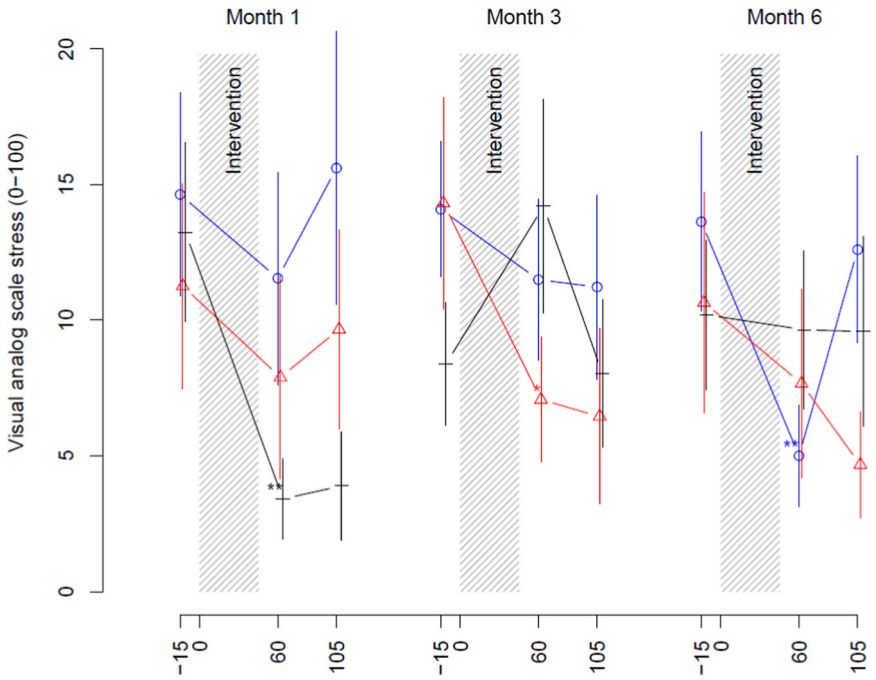

Fig. 4. VAS-S scores pre/post-session and recovery. *Note*. Mean ± SEM. Blue – group music therapy, red – recreational choir singing, black – book reading. Time on x-axis in minutes. *** $p < .001$, ** $p < .01$, * $p < .05$ (left side: within-group change from previous measurement; right side: between-group change compared to book reading, adjusted for previous measurement).

rigidity and a general weakening of adaptive response mechanisms, potentially limiting the capacity to mount typical hormonal responses to external stimuli.

Alpha-amylase decreased during book reading and increased after group music therapy, which may illustrate passive participation in this activity. VAS-scores decreased more rapidly after book reading, illustrating a delay in endocrine reaction. In the first occasion, cortisol reactions (T1) were absent in all groups. During the following occasions the reactions mirrored more complicated (counter)reactions, that have to do with patterns established partly on the first occasion. Alpha-amylase increased significantly during recovery in both music groups, which could mirror an activation that could be beneficial: It can be challenging to activate and mobilize people with dementia, especially when apathetic behaviour is present.

The mixed and inconsistent effects observed across groups and time points, along with the lack of alignment between VAS scores and physiological biomarkers, suggest several possible interpretations. One is that the music intervention does not consistently affect stress in this population. Alternatively, the biomarkers used—such as cortisol and alpha-amylase—may lack the sensitivity to detect subtle changes, or stress may not be the main mechanism influenced by the intervention. Possibly the intervention does not primarily target stress reduction, but instead promotes outcomes such as relaxation or enhanced social connectedness—dimensions that may not exist on the same continuum as stress. In that case, these markers may not be appropriate outcome measures. Given their cost and complexity, this finding is relevant for future studies considering similar methods. Measures more closely related to emotional states relevant to BPSD, or direct observational assessments of BPSD symptoms, could provide a more accurate reflection of the intervention's effects. However, inconsistencies may also stem from sample variability or missing data. Further research with more controlled samples and complete data is needed to clarify these issues. Additionally, future research might explore alternative physiological markers that are more sensitive to changes in arousal or emotional regulation.

4.1. Strengths

This study knows several strengths. First, this was the first multinational cluster-randomized trial in which care home residents with dementia and depressive symptoms were included to assess the short-term stress-reducing impact of group music therapy and recreational choir singing. Second, this study demonstrated willingness to contribute to non-invasive salivary biomarker research among this population and their legal representatives. Third, the assessment was conducted at three different timepoints across the six-month intervention period. Data from timepoint T1, T3 and T6 were deliberately not combined for the analysis, to take into account familiarity of the participants with attending the sessions. Being more familiar with the session structure, intervention provider and setting of the session may play a role in stress response. The fourth strength is the design of the main MIDDEL trial, in which participants were randomly allocated to GMT, RCS or the CONTROL group. The book reading provided to the CONTROL group resembled the music interventions in characteristics such as group size and duration, which ruled out differences based on an attention-based effect. Fifth, insight into non-attendance enabled us to exclude participants for whom saliva was collected, but who did not actually receive the intervention as intended. Finally, embedding this sub study in the MIDDEL trial facilitated a large sample size.

4.2. Limitations

This study knows several limitations as well. Changes in sCort and sAA concentrations could to some extent be due to diurnal variation (Warth et al., 2022), or day-to-day variations (Theorell et al., 2021), which may increase variance. In a previous study, sCort decreased over time in each intervention group (Warth et al., 2022). It is challenging to indicate what changes are attributable the intervention. To this end, the VAS-S was used to complement the physiological findings. However, similar to our findings, a previous study in palliative care also found weak correlations between sCort and self-reported stress levels (Warth et al., 2022). The VAS-S has been previously used in studies to assess subjective experienced stress to complement the objective

measurements of sAA and sCort (Maidhof et al., 2023; Warth et al., 2022), with the important difference that in these target groups the participants could complete the VAS-S themselves. In the current study, VAS-S scores were not consistently reported by either proxies or participants. A limitation of the VAS-S was its difficulty for participants to understand the question. Those who understood frequently reported no stress, whereas participants who did not apprehend what was asked of them became stressed. Hence, solely collection of proxy-reported VAS-S would have been preferred. Moreover, previous studies found differences between patient and proxy-reported outcomes (Kroenke et al., 2022; O'Shea et al., 2020), indicating our findings should be interpreted with caution. Our findings highlight a difference between stress reduction measured in salivary biomarkers compared to stress reduction observed by a caregiver or experienced by a person with dementia themselves. Regarding policy making, potential health insurance reimbursement and targeted use of music interventions – or other group interventions for that matter – the subjective user experience should be leading, especially considering the many confounding variables that may affect sAA and sCort values. Due to the different study arms and timepoints, multiple statistical tests were performed, which may increase the risk of finding significant results by chance. A limitation of the statistical analysis was that many of the saliva samples were not included in the analyses, which may have affected the results. Most importantly, the exclusion of these samples led to reduced test power. Additionally, it may have caused bias, giving more weight to participants who were able to fully adhere to the protocol. Such participants may be more likely to have better levels of functioning and lower severity of dementia than others whose samples had to be excluded.

4.3. Clinical considerations

The interventions in the current study differ in level of active contribution of participants: in the book reading session, no active participation was required, as passive listening sufficed. In choir singing, active participation was possible, but participants could more easily blend into the group by choosing not to sing along. The GMT group was smaller and individual attention was given to each individual participant. Moreover, in GMT improvising on instruments and movement to music were optional session components that required active engagement. Active engagement may be less or different in RCS and even more so in book reading, which required a more passive contribution. In nature, the GMT format may be more arousing for participants than RCS or book reading. The book reading sessions that were organized for the participants in the CONTROL arm were an active control condition. Less active and individual participation in the book reading CONTROL condition, followed by RCS and GMT, may partly explain why pre/post-session changes were observed in all three intervention arms. Indeed, a recent study argued that arousal affects the potential stress-reducing effects of music (Song et al., 2024).

The three interventions we compared were all group-based. A recent study suggested a therapeutic benefit can occur in group-based interventions due to being in a group, irrespective of the intervention (Lamont and Ranaweera, 2020). This could also partly explain why RCS may have a different impact on people than GMT, because RCS usually takes place in a larger group. Another study found larger reduction of stress in a singing group and a discussion group compared to a waiting list control group (Schäfer, 2023). However, group singing showed stronger effects on loneliness, social participation, and satisfaction with life. Previous studies found beneficial effects of RCS in oxytocin and on feelings of social cohesion and belonging in singing groups (Galinha et al., 2022). Other studies also reported effects in active control conditions (Schäfer, 2023), where the role of doing something in a group could create an effect, whether or not through making music or through some other shared activity (Fujioka and Hunt, 2023; Lamont and Ranaweera, 2020; Tsoi et al., 2018; van der Steen et al., 2025).

4.4. Methodological considerations

The proportion of participants for whom the complete set of saliva samples was collected was relatively low. However, being the first study of this kind in care home residents with mild to severe dementia, there was no frame of reference to base expectations of feasibility and acceptability of saliva collection on. In a recent study among people living at home with dementia and their loved ones, in which the family caregiver could assist with saliva sample collection, a higher percentage of possible saliva samples were collected (87.9 % of potential samples) and valid for assay (85.8 %) (Emami et al., 2022). Similar to a study in palliative care, there was a high proportion of missing data (Warth et al., 2022). Several factors could have contributed to samples having insufficient fluid for assay. The use of Salivette Children's Swabs was found to be a practical method to safely collect saliva samples in care home residents in advanced stages of dementia. However, the biochemical analysis indicated that a significant proportion, albeit less than half, of the collected saliva samples were not suitable for sAA and sCort assay. A common reason was insufficient fluid being available for sampling, potentially caused by dry mouth or xerostomia, which is common in older people, with dementia, who may use a variety of medications (Johansson et al., 2022). Saliva flow can be stimulated without affecting saliva compound, for example, by introducing the smell or an image of certain foods such lemon or bacon (Peres et al., 2015). Other factors that may have contributed to insufficient fluid on the cotton swab include lack of understanding saliva collection instructions; requests of residents to collect saliva independently; having many participants in the session; or lack of researchers or care staff available to assist in saliva collection. Nevertheless, the percentage of samples that were valid to include in the analysis of sCort and sAA was similar to a study conducted in palliative care (66.6 % and 69.6 % respectively) (Warth et al., 2022).

Care staff received instructions to consider potential confounders affecting sCort concentrations and sAA activity. As described in the study procedures, participants were not required to adhere to common instructions for saliva collection (e.g., refrain from food or drink intake) as this would potentially cause distress among residents who might not understand the reasons for such limitations, and disrupt everyday routines in the care home. To this end, adherence to instructions in this vulnerable population was not tested, but information about potential confounders was collected. Since sessions usually took place around 1 p.m., people often ate lunch within an hour before collecting the first saliva sample. During the hour after the session, people regularly had a drink or something to eat together in the common room.

4.5. Recommendations for future studies

Saliva collection required collaboration with care staff familiar with the participants. Even though participants may not recognize care staff members personally, the one-to-one interaction in the personal spaces in care home units is routinised. Collection of saliva samples is dependent on these routines, including appearing familiar (e.g., by wearing care staff attire); how to verbally approach the residents; and having expectations about the residents' potential response (e.g., agitation, skepticism). Future studies that plan saliva sampling in care home residents with dementia should also consider the active role of care staff in the data collection.

For follow-up studies it is recommended to explore the presence and influence of potential confounders and to take non-participation of residents in the intervention group into account. Also, the influence of proxy- and patient-reported outcomes for the VAS-S could be assessed, by including them in an embedded/stepwise multilevel model. It is recommended to use a suitable alternative to the visual analogue scale for stress, such as the use of standardized emotional vignettes (Gooch et al., 2020) to self-reported subjective experience. Another alternative could be the Two-dimensional Mood Scale (TDMS) for self-reported mood in older adults with dementia consists of eight items addressing

vitality, stability, pleasure and arousal (Kobayashi et al., 2024).

Future steps for this type of research include 1) assessing the correlation between VAS-S and cognition measures; 2) explore validity of self-report data and comprehension of scales; 3) using pre-post session saliva changes to predict later clinical change in depression and quality of life; 4) take into account country as a predictor, since this was found to be important in the clinical outcomes of the main MIDDEL trial.

In future studies with a similar aim, we would optimize the timing of sessions and saliva collection, which frequently occurred shortly after lunchtime in the current study. Ideally, we would ensure that sessions take place at least an hour after lunchtime, to reduce chances of saliva samples being invalid by food debris. However, our feasibility findings show that in daily practice in many nursing homes the ideal time for conducting sessions is immediately after lunch. To increase the chance of successfully collecting saliva samples, in a large group (such as in choir singing and reading aloud) we would collect saliva from approximately five participants rather than from all of them.

5. Conclusions

This study assessed the short-term impact of group-based music interventions on salivary biomarkers of stress and subjective stress in a large sample of care home residents with dementia and depressive symptoms in three European countries. Pre/post-session changes occurred within study arms, but without clear and consistent group differences. Subjective stress was not per se reflected in salivary cortisol (sCort) concentrations and alpha-amylase (sAA) activity. Findings did however show that the group-based interventions could have a stress-reducing effect. The physiological outcomes may have been influenced by differences in arousal or the degree of active participation between intervention groups. This study also showed that repeated assessments of salivary biomarkers were feasible and acceptable among this population. Meanwhile, a significant proportion of collected data were missing or invalid, warranting recommendations for early identification, prevention and close monitoring of potential confounding variables and stimulation of saliva flow in future studies.

CRediT authorship contribution statement

Naomi L. Rasing: Writing – review & editing, Writing – original draft, Validation, Resources, Project administration, Methodology, Investigation, Formal analysis. **Sarah I.M. Janus:** Writing – review & editing, Validation, Supervision, Project administration, Methodology, Investigation, Formal analysis. **Annemieke C. Vink:** Writing – review & editing. **Ulrike A.S. Frischen:** Writing – review & editing, Investigation. **Johanna Neuser:** Writing – review & editing, Validation, Investigation. **Jo Dugstad Wake:** Writing – review & editing, Investigation. **Nadine Skoluda:** Investigation, Formal analysis. **Vigdis Sveinsdottir:** Writing – review & editing, Validation, Project administration. **Monika Geretsegger:** Writing – review & editing, Project administration. **Elias Langeland:** Writing – review & editing, Formal analysis. **Gunter Kreutz:** Writing – review & editing, Validation, Project administration, Methodology, Funding acquisition, Conceptualization. **Christian Gold:** Writing – review & editing, Validation, Project administration, Methodology, Funding acquisition, Formal analysis, Conceptualization. **Urs M. Nater:** Writing – review & editing, Validation, Software, Resources, Methodology, Formal analysis. **Sytse U. Zuidema:** Writing – review & editing, Validation, Supervision, Project administration, Methodology, Funding acquisition, Conceptualization.

Ethical approval

The research was carried out in accordance with the WMA Declaration of Helsinki. Ethical approval was obtained prior to data collection in Germany: Medical Ethics Committee at the Carl von Ossietzky University Oldenburg (project number 2020–134; approval date: September 9, 2020); in the Netherlands: Medical Ethical Committee (METc) of the University Medical Center Groningen (research register 202000194; approval date: January 12, 2021); and in Norway: Regional Committees for Medical and Health Research Ethics (REC), Norway (project number 110052; approval date: Jun 29, 2020). Study- and patient-related documentation was adapted where needed to adhere to national ethical requirements.

Funding

The EU Joint Programme—Neurodegenerative Disease Research (JPND) project Music Interventions for Dementia and Depression in ELderly care (MIDDEL) was supported by the following funding organizations under the aegis of JPND: The Netherlands Organization for Health Research and Development (ZonMw, project number 733051108), the Federal Ministry of Education and Research (BMBF, project number 01ED2011), and the Research Council of Norway (RCN, project number 311628). The funders did not have any role in the study design, data collection, interpretation or analysis, or decisions to publish. U.M.N. acknowledges funding from the Research Platform, the Stress of Life (SOLE). Funding for additional analyses was provided by Leon Jarners Minnefond (research prize 2024 to C.G.)

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to thank all research participants, care staff, and intervention providers for their efforts and willingness to participate in this sub-study, we gained a lot of practical experiences and collected a lot of valuable data. The authors would like to acknowledge and thank the music therapy bachelor students of ArtEZ University of the Arts who helped with the saliva data collection on multiple occasions. We would like to thank Dr. Burger and Dr. Assmus for their advice regarding statistical analyses. We are grateful for all the efforts, expertise and pleasant collaboration with the team of experts of the Biochemical Laboratory of the University of Vienna.

References

Ali, N., Nater, U.M., 2020. Salivary Alpha-Amylase as a biomarker of stress in behavioral Medicine. Int. J. Behav. Med. 27 (3), 337–342. https://doi.org/10.1007/s12529-019-09843-x.

Campbell, Q., Bodkin-Allen, S., Swain, N., 2022. Group singing improves both physical and psychological wellbeing in people with and without chronic health conditions: a narrative review. J. Health Psychol. 27 (8), 1897–1912. https://doi.org/10.1177/13591053211012778.

Dawudi, M., Schall, A., Tesky, V.A., Pantel, J., 2024. The psychosocial and physiological effects of Choir-Singing in people with dementia a pilot study. J. Gerontopsychol. Geriatr. Psychiatry 37 (1), 3–13. https://doi.org/10.1024/1662-9647/a000323.

De La Rubia Ortí, J.E., García-Pardo, M.P., Iranzo, C.C., Madrigal, J.J.C., Castillo, S.S., Rochina, M.J., Gascó, V.J.P., 2018. Does music therapy improve anxiety and depression in Alzheimer's patients? J. Altern. Complement. Med. 24 (1), 33–36. https://doi.org/10.1089/acm.2016.0346.

van der Steen, J.T., van der Wouden, J.C., Methley, A.M., Smaling, H.J.A., Vink, A.C., Bruinsma, M.S., 2025. Music-based therapeutic interventions for people with dementia. Cochrane Database Syst. Rev. (3). https://doi.org/10.1002/14651858.CD003477.pub5.

Emami, A., Hallinder, H., Theorell, T., Kim, H., Engström, G., 2022. The feasibility and acceptability of in-home saliva collection for stress in persons with dementia and their family caregivers. Biol. Res. Nursing 0 (0), 1–8. https://doi.org/10.1177/10998004221076554Emami.

Emami, A., Jun, J., Theorell, T., Engström, G., Berglund, L., 2024. Day-to-day variability of stress biomarkers during a music intervention in people living with dementia and their family caregivers. Sci. Prog. 107 (3). https://doi.org/10.1177/00368504241263692.

Emami, A., Theorell, T., Kim, H., Berglund, L., Hallinder, H., Engström, G., 2023. Assessing stress using repeated saliva concentration of steroid hormones in dementia

care dyads: results from a controlled pilot care music intervention. Upsala J. Med. Sci. 128. https://doi.org/10.48101/ujms.v128.9340.
Fischer, J.E., Calame, A., Dettling, A.C., Zeier, H., Fanconi, S., 2000. Experience and endocrine stress responses in neonatal and pediatric critical care nurses and physicians. Crit. Care Med. 28 (9), 3281–3288. https://doi.org/10.1097/00003246-200009000-00027.
Folstein, M.F., Folstein, S.E., McHugh, P.R., 1975. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J. Psychiatr. Res. 12 (3), 189–198. https://doi.org/10.1016/0022-3956(75)90026-6.
Fujioka, T., Hunt, A.M., 2023. Mechanisms of music therapy and Music-Based interventions. In: Devlin, In.K., Kang, K., Pantelyat, A. (Eds.), Music Therapy and Music-Based Interventions in Neurology. Current Clinical Neurology, pp. 9–22.
Galinha, I.C., García-Martín, M.Á., Lima, M.L., 2022. Sing4Health: randomised controlled trial of the effects of a singing group program on the subjective and social well-being of older adults. Appl. Psychol. Health WellBeing 14 (1), 176–195. https://doi.org/10.1111/aphw.12297.
Gold, C., Eickholt, J., Assmus, J., Stige, B., Wake, J.D., Baker, F.A., Tamplin, J., Clark, I., Lee, Y.E.C., Jacobsen, S.L., Ridder, H.M.O., Kreutz, G., Muthesius, D., Wosch, T., Ceccato, E., Raglio, A., Ruggeri, M., Vink, A., Zuidema, S., Geretsegger, M., 2019. Music interventions for dementia and depression in ELderly care (MIDDEL): protocol and statistical analysis plan for a multinational cluster-randomised trial. BMJ Open 9 (3). https://doi.org/10.1136/bmjopen-2018-023436.
Gooch, D., Mehta, V., Price, B., McCormick, C., Bandara, A., Bennaceur, A., Bennasar, M., Stuart, A., Clare, L., Levine, M., Cohen, J., & Nuseibeh, B. (2020). How are you feeling? Using Tangibles to Log the Emotions of Older Adults. In Proceedings of the Fourteenth International Conference on Tangible, Embedded, and Embodied Interaction (TEI '20). Association for Computing Machinery, New York, NY, USA, 31–43. https://doi.org/10.1145/3374920.3374922.
Het, S., Schoofs, D., Rohleder, N., Wolf, O.T.T., 2012. Stress-induced cortisol level elevations are associated with reduced negative affect after stress: indications for a mood-buffering cortisol effect. Psychosom. Med. 74 (1), 23–32. https://doi.org/10.1097/PSY.0b013e31823a4a25.
James, K.A., Stromin, J.I., Steenkamp, N., Combrinck, M.I., 2023. Understanding the relationships between physiological and psychosocial stress, cortisol and cognition. Front. Endocrinol. 14, 1085950. https://doi.org/10.3389/fendo.2023.1085950.
Johansson, A.K., Omar, R., Mastrovito, B., Sannevik, J., Carlsson, G.E., Johansson, A., 2022. Prediction of xerostomia in a 75-year-old population: a 25-year longitudinal study. J. Dent. 118, 104056. https://doi.org/10.1016/j.jdent.2022.104056.
Kanikowska, D., Roszak, M., Rutkowski, R., Sato, M., Sikorska, D., Orzechowska, Z., Bre, A., Janusz Witowski, X., 2019. Seasonal differences in rhythmicity of salivary cortisol in healthy adults. J. Appl. Physiol. 126, 764–770. https://doi.org/10.1152/japplphysiol.00972.2018.
Kobayashi, T., Igusa, T., Uchida, H., Tsuchiya, K., Akiba, T., Takahashi, T., Taguchi, M., Okada, C., Yabuki, H., Kanazawa, Y., Kikuchi, S., Hirao, K., 2024. Reliability of the Two-dimensional mood scale for self-reported mood assessment by older adults with dementia. Geriatr. Nurs. 58, 459–465. https://doi.org/10.1016/j.gerinurse.2024.06.016.
Kroenke, K., Stump, T.E., Monahan, P.O., 2022. Agreement between older adult patient and caregiver proxy symptom reports. J. PatientRep. Outcomes 6 (50). https://doi.org/10.1186/s41687-022-00457-8.
Lamont, A., Ranaweera, N.A., 2020. Knit one, play one: comparing the effects of amateur knitting and amateur music participation on happiness and wellbeing. Appl. Res. Qual. Life 15 (5), 1353–1374. https://doi.org/10.1007/s11482-019-09734-z.
Lesage, F.X., Berjot, S., Deschamps, F., 2012. Clinical stress assessment using a visual analogue scale. Occup. Med. 62 (8), 600–605. https://doi.org/10.1093/occmed/kqs140.
Linnemann, A., Ditzen, B., Strahler, J., Doerr, J.M., Nater, U.M., 2015. Music listening as a means of stress reduction in daily life. Psychoneuroendocrinology 60, 82–90. https://doi.org/10.1016/j.psyneuen.2015.06.008.
Linnemann, C., Lang, U.E., 2020. Pathways connecting late-life depression and dementia. Front. Pharmacol. 11 (279), 1–10. https://doi.org/10.3389/fphar.2020.00279.
Maidhof, R.M., Kappert, M.B., Wuttke, A., Schwerdtfeger, A.R., Kreutz, G., Nater, U.M., 2023. Effects of participant-selected versus researcher-selected music on stress and mood – the role of gender. Psychoneuroendocrinology 158 (September), 106381. https://doi.org/10.1016/j.psyneuen.2023.106381.
Matsuda, S., Yamaguchi, T., Okada, K., Gotouda, A., Mikami, S., 2012. Day-to-day variations in salivary cortisol measurements. J. Prosthodont. Res. 56 (1), 37–41. https://doi.org/10.1016/j.jpor.2011.04.004.
Moher, D., Hopewell, S., Schulz, K.F., Montori, V., Gøtzsche, P.C., Devereaux, P.J., Elbourne, D., Egger, M., Altman, D.G., 2010. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. BMJ 340, c869. https://doi.org/10.1136/bmj.c869.
Montgomery, S.A., Asberg, M., 1979. A new depression scale designed to be sensitive to change. Br. J. Psy. J. Mental Sci. 134, 382–389. https://doi.org/10.1192/bjp.134.4.382.
Morris, J.C., 1993. The clinical dementia rating (CDR): current version and scoring rules. Neurology 43 (11), 2412–2414. https://doi.org/10.1212/wnl.43.11.2412-a.
Narvaez Linares, N.F., Charron, V., Ouimet, A.J., Labelle, P.R., Plamondon, H., 2020. A systematic review of the trier social stress test methodology: issues in promoting study comparison and replicable research. Neurobiol. Stress 13, 100235. https://doi.org/10.1016/j.ynstr.2020.100235.
O'Shea, E., Hopper, L., Marques, M., Gonçalves-Pereira, M., Woods, B., Jelley, H., Verhey, F., Kerpershoek, L., Wolfs, C., de Vugt, M., Stephan, A., Bieber, A., Meyer, G., Wimo, A., Michelet, M., Selbaek, G., Portolani, E., Zanetti, O., Irving, K., 2020. A comparison of self and proxy quality of life ratings for people with dementia and their carers: a european prospective cohort study. Aging Ment. Health 24 (1), 162–170. https://doi.org/10.1080/13607863.2018.1517727.
Peres, J.C., Rouquette, J.L., Miočević, O., Warner, M.C., Slowey, P.D., Shirtcliff, E.A., 2015. New techniques for augmenting saliva collection: bacon rules and lozenge drools. Clin. Ther. 37 (3), 515–522. https://doi.org/10.1016/j.clinthera.2015.02.015.
Rasing, N.L., Janus, S.I.M., Kreutz, G., Sveinsdottir, V., Gold, C., Nater, U.M., Zuidema, S. U., 2022. The impact of music on stress biomarkers: protocol of a substudy of the Cluster-Randomized controlled trial music interventions for dementia and depression in ELderly care (MIDDEL). Brain Sci. 12, 485. https://doi.org/10.3390/brainsci12040485.
Schäfer, T., 2023. The positive effects of online group singing on psycho-physiological variables during the COVID-19 pandemic—A pilot randomized controlled trial. Appl. Psychol. Health WellBeing 15 (4), 1254–1270. https://doi.org/10.1111/aphw.12435.
Schmitt, F.A., Saxton, J., Ferris, S.H., Mackell, J., Sun, Y., 2013. Evaluation of an 8-item severe impairment battery (SIB-8) vs. The full SIB in moderate to severe alzheimer's disease patients participating in a donepezil study. Int. J. Clin. Pract. 67 (10), 1050–1056. https://doi.org/10.1111/ijcp.12188.
Schulz, K.F., Altman, D.G., Moher, D., et al., 2010. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMC Med. 8, 18. https://doi.org/10.1186/1741-7015-8-18.
Sharp, B.K., 2019. Stress as experienced by people with dementia: an interpretative phenomenological analysis. Dementia 18 (4), 1427–1445. https://doi.org/10.1177/1471301217713877.
Sittler, M.C., Worschech, F., Wilz, G., Fellgiebel, A., Wuttke-Linnemann, A., 2021. Psychobiological mechanisms underlying the health-beneficial effects of music in people living with dementia: a systematic review of the literature. Physiol. Behav. 233 (113338), 1–12. https://doi.org/10.1016/j.physbeh.2021.113338.
Song, Y., Ali, N., Nater, U.M., 2024. The effect of music on stress recovery. Psychoneuroendocrinology 168, 107137. https://doi.org/10.1016/j.psyneuen.2024.107137.
Strahler, J., Skoluda, N., Kappert, M.B., Nater, U.M., 2017. Simultaneous measurement of salivary cortisol and alpha-amylase: application and recommendations. Neurosci. Biobehav. Rev. 83, 657–677. https://doi.org/10.1016/j.neubiorev.2017.08.015.
Theorell, T., Engström, G., Hallinder, H., Lennartsson, A.K., Kowalski, J., Emami, A., 2021. The use of saliva steroids (cortisol and DHEA-s) as biomarkers of changing stress levels in people with dementia and their caregivers: a pilot study. Sci. Prog. 104 (2), 1–18. https://doi.org/10.1177/00368504211019856.
Thoma, M.V., La Marca, R., Brönnimann, R., Finkel, L., Ehlert, U., Nater, U.M., 2013. The effect of music on the human stress response. PloS One 8 (8), e70156. https://doi.org/10.1371/journal.pone.0070156.
Tsoi, K.K.F., Chan, J.Y.C., Ng, Y.M., Lee, M.M.Y., Kwok, T.C.Y., Wong, S.Y.S., 2018. Receptive music therapy is more effective than interactive music therapy to relieve behavioral and psychological symptoms of dementia: a systematic review and Meta-Analysis. In: In Journal of the American Medical Directors Association, 19. Elsevier Inc, pp. 568–576.e3. https://doi.org/10.1016/j.jamda.2017.12.009.
Warth, M., Stoffel, M., Koehler, F., Bardenheuer, H.J., Kessler, J., Ditzen, B., 2022. Characteristics of salivary cortisol and alpha-amylase as psychobiological study outcomes in palliative care research. BMC Palliat. Care 21 (1), 1–13. https://doi.org/10.1186/s12904-022-01085-1.
de Witte, M., Pinho, A. da S., Stams, G.J., Moonen, X., Bos, A.E.R., van Hooren, S., 2022. Music therapy for stress reduction: a systematic review and meta-analysis. Health Psychol. Rev. 16 (1), 134–159. https://doi.org/10.1080/17437199.2020.1846580.
de Witte, M., Spruit, A., van Hooren, S., Moonen, X., Stams, G.J., 2019. Effects of music interventions on stress-related outcomes: a systematic review and two meta-analyses. Health Psychol. Rev. 14 (2), 1–33. https://doi.org/10.1080/17437199.2019.1627897.
Wong, M.M., Tahir, T., Wong, M.M., Baron, A., Finnerty, R., 2021. Biomarkers of stress in music interventions: a systematic review. J. Music Ther. 58 (3), 241–277. https://doi.org/10.1093/jmt/thab003.
Yu, A.L., Lo, S.F., Chen, P.Y., Lu, S.F., 2022. Effects of group music intervention on depression for elderly people in nursing homes. Int. J. Environ. Res. Public Health 19 (15). https://doi.org/10.3390/ijerph19159291.