TECHNICAL COMMENTARY

Music therapy

Introduction

Music therapy utilises musical experiences and interactions designed to assist people with disorders such as schizophrenia to address issues they may have difficulty with, such as communication and self-regulation[1, 2]. It may be offered through group or individual programmes and does not require a client to have musical skills. Therapists are trained to respond to challenging behaviour using both musical and non-musical strategies[2]. Music therapy can be active (including improvisation, producing music) or receptive (listening to either live or recorded music). The musical therapist can manipulate the rhythmic or harmonic structure to alter therapy intensity. The therapist can also direct the focus of the session to be concentrating on the processes within the music itself, or to focus more on the client's emotional responses to the music[1].

Method

We have included only systematic reviews (systematic literature search, detailed methodology with inclusion/exclusion criteria) published in full text, in English, from the year 2000 that report results separately for people with a diagnosis of schizophrenia, schizoaffective disorder, schizophreniform disorder or first episode schizophrenia. Reviews were identified by searching MEDLINE, EMBASE, CINAHL, Current Contents, PsycINFO and the Cochrane Library databases. Hand searching reference lists of identified reviews was also conducted. When multiple copies of reviews were found, only the most recent version was included. Reviews with pooled data are prioritised for inclusion.

Review reporting assessment was guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist that describes a preferred way to present a meta-analysis[3]. Reviews rated as having less than 50% of items checked have been excluded from the library. The PRISMA flow diagram is a suggested way of providing information about studies included and excluded with reasons for exclusion. Where no flow diagram has been presented by individual reviews, but identified studies have been described in the text, reviews have been checked for this item. Note that early reviews may have been guided by less stringent reporting checklists than the PRISMA, and that some reviews may have been limited by journal guidelines.

Evidence was graded using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) Working Group approach where high quality evidence such as that gained from randomised controlled trials (RCTs) may be downgraded to moderate or low if review and study quality is limited, if there is inconsistency in results, indirect comparisons, imprecise or sparse data and high probability of reporting bias. It may also be downgraded if risks associated with the intervention or other matter under review are high. Conversely, low quality evidence such as that gained from observational studies may be upgraded if effect sizes are large or if there is a dose dependent response. We have also taken into account sample size and whether results are consistent, precise and direct with low associated risks (see end of table for an explanation of these terms)[4]. The resulting table represents an objective summary of the available evidence, although the conclusions are solely the opinion of staff of NeuRA (Neuroscience Research Australia).

Results

We found four systematic reviews that met our inclusion criteria[5-8].

- Moderate quality evidence suggests large effects of improved global state and anxiety,

Margarete Ainsworth Building, Barker Street, Randwick NSW 2031. Phone: 02 9399 1000. Email: info@neura.edu.au

and a medium-sized effect of increased musical engagement with music therapy over standard care.

- Moderate to low quality evidence suggests medium to large effects of improved positive and negative symptoms, mood, social functioning, behaviour (including catatonic behaviour), and attention (PASAT scale).
- No significant benefits were found for quality of life, medication levels, study attrition, patient satisfaction, general functioning, working memory, or abstract thinking.

Margarete Ainsworth Building, Barker Street, Randwick NSW 2031. Phone: 02 9399 1000. Email: info@neura.edu.au

Gold C, Solli HP, Kruger V, Lie SA

Dose-response relationship in music therapy for people with serious mental disorders: systematic review and meta-analysis

Clinical Psychology Review 2009; 29(3): 193-207

View review abstract online

Comparison	**Music therapy (individual or group) including active and receptive components plus standard care vs. standard care.** **66% of the sample included people with schizophrenia spectrum disorders, and the remainder had depression.** **Treatment duration ranged 1-6 months, 1-6 sessions per week.**
Summary of evidence	**Moderate quality evidence (small to medium-sized samples, consistent, direct, precise) suggests a large effect of improved global state and anxiety, and a medium-sized effect of increased musical engagement with music therapy over standard care, with no differences for positive symptoms, quality of life, medication levels, study attrition or patient satisfaction.**

Study retention

No differences between groups;

3 RCTs and 1 non-randomised trial, N = 226, OR = 1.11, 95%CI 0.42 to 2.92, $p > 0.05$, $I^2 = 0\%$

Global state

A significant, medium to large effect of improvements in the following with music therapy;

Global state: 2 RCTs, N = 140, OR = 0.03, 95%CI 0.01 to 0.09, $p < 0.001$, $I^2 = 0\%$

Anxiety: 2 RCTs, N = 100, $g = 1.31$, 95%CI 0.85 to 1.78, $p < 0.001$, $I^2 = 0\%$

Musical engagement: 2 non-randomised trials, N = 107, $g = 0.49$, 95%CI 0.09 to 0.88, $p < 0.05$, $I^2 = 0\%$

No differences between groups for;

Quality of life: 1 RCT and 1 non-randomised trial, 2 studies, N = 103, $g = 0.16$, 95%CI -0.24 to 0.56, $p > 0.05$, $I^2 = 0\%$

Patient satisfaction: 2 RCTs, N = 118, $g = 0.06$, 95%CI -0.57 to 0.68, $p > 0.05$, $I^2 = 52\%$

Reduced medication: 1 RCT and 1 non-randomised trial, N = 142, $g = -0.25$, 95%CI -0.58 to 0.08, $p > 0.05$, $I^2 = 41\%$

Mental state	
No differences between groups; Positive symptoms: 1 RCT, 1 non-randomised trial and 2 non-controlled studies, N = 170, *g* = 0.18, 95%CI -0.12 to 0.48, *p* > 0.05, I^2 = 0% Note: meta-regression found that the number of treatment sessions significantly predicted 78% of the variance in general symptoms (7 studies N = 315, *p* < 0.01), suggesting the success of music therapy was strongly linked to the number of treatment sessions, with a small effect on general symptoms expected after ten sessions, and a large effect after 39 sessions. Similar effects were found when negative symptoms (8 studies), depressive symptoms (7 studies), and functioning (5 studies) were examined separately.	
Consistency in results	Consistent where applicable (>1 study).
Precision in results	Precise for all except quality of life and study attrition.
Directness of results	Direct

Mössler K, Chen X, Heldal TO, Gold C

Music therapy for people with schizophrenia and schizophrenia-like disorders

Cochrane Database of Systematic Reviews 2011, Issue 12. Art. No.: CD004025. DOI: 10.1002/14651858.CD004025.pub3

View review abstract online

Comparison	**Music therapy (either individual or group) including active and receptive components plus standard care vs. standard care.** **Treatment duration ranged 1-4 months.**
Summary of evidence	**Moderate to low quality evidence (small samples, inconsistent, precise, direct) suggests medium to large effects of improved positive and negative symptoms, social functioning, behaviour, attention (PASAT scale), depression, and anxiety with music therapy. No benefits were found over standard care for study attrition, general functioning, quality of life, patient satisfaction, working memory, or abstract thinking.**
Study retention	

No significant difference between groups; 8 RCTs, N = 493, RR = 1.03, 95%CI 0.38 to 2.78, p = 0.95, I^2 = 0%, p = 0.98
Global state
Significant improvements with music therapy for; Global state, clinically important improvement (large effect): 1 RCT, N = 72, RR = 0.10, 95%CI 0.03 to 0.31, p = 0.000063 Social functioning, SDSI score (large effect): 1 RCT, N = 70, d = -0.78, 95%CI -1.27 to -0.28, p = 0.0020 Attention, PASAT score (medium to large effect): 1 RCT, N = 67, d = 0.72, 95%CI 0.22 to 1.21, p = 0.0048 Positive behaviours, NOSIE score (large effect): 1 RCT, N = 60, d = -1.24 95% CI -1.79 to -0.68, p = 0.000013 Reducing negative behaviours, NOSIE score (large effect): 1 RCT, N = 60, d = -2.22 95% CI -2.87 to -1.57, p < 0.00001 *No significant differences between groups for;* General functioning, GAF endpoint score: 1 RCT, N = 69, d = -0.05, 95%CI -0.53 to 0.43, p = 0.85 Quality of life, SPG score: 1 RCT, N = 31, d = 0.05, 95%CI -0.66 to 0.75, p = 0.90 Patient satisfaction, CSQ score: 1 RCT, N = 69, d = 0.32, 95%CI -0.16 to 0.80, p = 0.19 Working Memory, WMS score: 1 RCT, N = 67, d = 0.43, 95% CI -0.06 to 0.92, p = 0.083 Attention, CCPT score: 1 RCT, N = 67, d = 0.25, 95%CI -0.23 to 0.74, p = 0.30 Abstract thinking, BCST score: 1 RCT, N = 67, d = 0.09, 95%CI -0.39 to 0.58, p = 0.70
Mental state
Significant improvements with music therapy for; Total endpoint scores > 20 sessions, BPRS score (large effect): 2 RCTs, N = 100, d = -0.73, 95%CI -1.16 to -0.31, p = 0.00064, I^2 = 91%, p = 0.0008 Negative symptoms endpoint scores all sessions, SANS score (large effect): 4 RCTs, N = 240, d = -0.74, 95%CI -1.00 to -0.47, p < 0.00001, I^2 = 53%, p = 0.09 Depression symptoms, SDS score (medium effect): 2 RCTs, N = 90, d = -0.63, 95%CI -1.06 to -0.21, p = 0.0036, I^2 = 0%, p = 0.73 Anxiety symptoms, SAS score (medium effect): 1 RCT, N = 60, d = -0.61, 95% CI -1.13 to - 0.09, p = 0.021 *No significant differences between groups for;* PANSS < 20 sessions: 1 RCT, N = 69, d = -0.36, 95%CI -0.84 to 0.12, p = 0.14

Depression symptoms, Ham-D: 1 RCT, N = 30, *d* = -0.52, 95%CI -1.25 to -0.21, *p* = 0.16	
Consistency in results	Consistent for study retention, negative symptoms and depression. Inconsistent for total BPRS scores. Not applicable for outcomes with 1 RCT.
Precision in results	Precise except for anxiety, depression, behaviour, quality of life and study retention.
Directness of results	Direct

Silverman MJ

The influence of music on the symptoms of psychosis: a meta-analysis

Journal of Music Therapy 2003; 40(1): 27-40

View review abstract online

Comparison	**Music therapy (either individual or group) including active and receptive components plus routine care vs. routine care.**
Summary of evidence	**Moderate quality evidence (unclear sample size, inconsistent, precise, direct) suggests music therapy may have a medium-sized effect for improved catatonic behaviour.**
Mental state	
A significant, medium-sized effect of improved catatonic behaviour; 8 studies, N not reported, *d* = 0.57, 95%CI 0.40 to 0.88, *p* < 0.05 *No significant differences for;* General symptoms: 4 studies, N not reported, *d* = 0.59, 95%CI 0.30 to 0.88, *p* = 0.79 Cognitive symptoms: 7 studies, N not reported, *d* = 0.92, 95%CI 0.73 to 1.12, *p* = 0.16 Authors report no moderating effects according to type of therapy received (passive listening, live or recorded music, genre) or according to patient's sex.	
Consistency in results	Authors report results are mostly inconsistent.
Precision in results	Precise
Directness of results	Direct

Tseng PT, Chen YW, Lin PY, Tu KY, Wang HY, Cheng YS, Chang YC, Chang CH, Chung W, Wu CK

Significant treatment effect of adjunct music therapy to standard treatment on the positive, negative, and mood symptoms of schizophrenic patients: A meta-analysis

BMC Psychiatry 2016; 16: 16

View review abstract online

Comparison	**Music therapy (individual or group) plus routine care (all patients were on antipsychotic medication) vs. routine care.**
Summary of evidence	**Moderate to low quality evidence (large sample size, direct, inconsistent, imprecise, possible publication bias) suggests a large effect for improved overall symptoms, including positive, negative and mood symptoms, with adjunctive music therapy.**
Mental state	
A significant, large effect of improved overall symptoms with adjunctive music therapy; 12 studies, N = 804, g = 3.25, 95%CI 2.08 to 4.42, $p < 0.001$, I^2 = 97.6%, $p < 0.001$ Moderator analyses revealed longer duration of illness was related to greater improvement in symptoms, while increased number of hospitalisations was related to a less improvement in symptoms. There were no significant moderating effects of trial design (RCTs vs. non-RCTs), age, sex, frequency and duration of music therapy, or antipsychotic dose. Subgroup analyses of symptoms revealed music therapy improved positive symptoms (g = 1.63, 95%CI 0.30 to 2.96, p = 0.017), negative symptoms (g = 4.14, 95%CI 2.54 to 5.74, $p < 0.001$), and mood symptoms (g = 1.00, 95%CI 0.56 to 1.43, $p < 0.001$), but not subscales of general psychopathology (g = 9.30, 95%CI −0.68 to 19.28, p = 0.068). There were also no differences in dropout rates. Authors report possible publication bias.	
Consistency in results	Inconsistent
Precision in results	Imprecise
Directness of results	Direct

Margarete Ainsworth Building, Barker Street, Randwick NSW 2031. Phone: 02 9399 1000. Email: info@neura.edu.au

Explanation of acronyms

BCST = Berg's Card-Sorting Task, BPRS = Brief Psychiatric Rating Scale, CCPT = Conners Continuous Performance Task, CI = Confidence Interval, CSQ = Client Satisfaction Questionnaire, GAF = Global Assessment of Functioning scale, Ham-D = Hamilton Depression rating scale, I^2 = the percentage of the variability in effect estimates that is due to heterogeneity rather than sampling error (chance), N = number of participants, NOSIE = Nurses Observation Scale for Inpatient Evaluation, *p* = statistical probability of obtaining that result ($p < 0.05$ generally regarded as significant), PANSS = Positive and Negative Syndrome Scale, PASAT = Paced Auditory Serial Addition Task, Q = Q statistic for the test of heterogeneity, RCT = randomised controlled trial, RR = relative risk, SANS = Scale for the Assessment of Negative Symptoms, SAS = Self-rating Anxiety scale, SDS = Self-rating Depression Scale, SDS = Self-rating Depression Scale, SDSI = Social Disability Schedule for Inpatients, SPG = German quality of life task (Skalen zur psychischen Gesundheit) vs. = versus, WMS = Wechsler Memory Scale

Margarete Ainsworth Building, Barker Street, Randwick NSW 2031. Phone: 02 9399 1000. Email: info@neura.edu.au

Explanation of technical terms

* Bias has the potential to affect reviews of both RCT and observational studies. Forms of bias include; reporting bias – selective reporting of results; publication bias - trials that are not formally published tend to show less effect than published trials, further if there are statistically significant differences between groups in a trial, these trial results tend to get published before those of trials without significant differences; language bias – only including English language reports; funding bias - source of funding for the primary research with selective reporting of results within primary studies; outcome variable selection bias; database bias - including reports from some databases and not others; citation bias - preferential citation of authors. Trials can also be subject to bias when evaluators are not blind to treatment condition and selection bias of participants if trial samples are small[9].

---

† Different effect measures are reported by different reviews.

Prevalence refers to how many existing cases there are at a particular point in time. Incidence refers to how many new cases there are per population in a specified time period. Incidence is usually reported as the number of new cases per 100,000 people per year. Alternatively some studies present the number of new cases that have accumulated over several years against a person-years denominator. This denominator is the sum of individual units of time that the persons in the population are at risk of becoming a case. It takes into account the size of the underlying population sample and its age structure over the duration of observation.

Reliability and validity refers to how accurate the instrument is. Sensitivity is the proportion of actual positives that are correctly identified (100% sensitivity = correct identification of all actual positives) and specificity is the proportion of negatives that are correctly identified (100% specificity = not identifying anyone as positive if they are truly not).

Weighted mean difference scores refer to mean differences between treatment and comparison groups after treatment (or occasionally pre to post treatment) and in a randomised trial there is an assumption that both groups are comparable on this measure prior to treatment. Standardised mean differences are divided by the pooled standard deviation (or the standard deviation of one group when groups are homogenous) which allows results from different scales to be combined and compared. Each study's mean difference is then given a weighting depending on the size of the sample and the variability in the data. Less than 0.4 represents a small effect, around 0.5 a medium effect, and over 0.8 represents a large effect[9].

Odds ratio (OR) or relative risk (RR) refers to the probability of a reduction (< 1) or an increase (> 1) in a particular outcome in a treatment group, or a group exposed to a risk factor, relative to the comparison group. For example, a RR of 0.75 translates to a reduction in risk of an outcome of 25% relative to those not receiving the treatment or not exposed to the risk factor. Conversely, a RR of 1.25 translates to an increased risk of 25% relative to those not receiving treatment or not having been exposed to a risk factor. A RR or OR of 1.00 means there is no difference between groups. A medium effect is considered if RR > 2 or < 0.5 and a large effect if RR > 5 or < 0.2[10]. lnOR stands for logarithmic OR where a lnOR of 0 shows no difference between groups. Hazard ratios

Margarete Ainsworth Building, Barker Street, Randwick NSW 2031. Phone: 02 9399 1000. Email: info@neura.edu.au

measure the effect of an explanatory variable on the hazard or risk of an event.

Correlation coefficients (eg, r) indicate the strength of association or relationship between variables. They can provide an indirect indication of prediction, but do not confirm causality due to possible and often unforseen confounding variables. An r of 0.10 represents a weak association, 0.25 a medium association and 0.40 and over represents a strong association. Unstandardised (*b*) regression coefficients indicate the average change in the dependent variable associated with a 1 unit change in the independent variable, statistically controlling for the other independent variables. Standardised regression coefficients represent the change being in units of standard deviations to allow comparison across different scales.

---

‡ Inconsistency refers to differing estimates of effect across studies (i.e. heterogeneity or variability in results) that is not explained by subgroup analyses and therefore reduces confidence in the effect estimate. I^2 is the percentage of the variability in effect estimates that is due to heterogeneity rather than sampling error (chance) - 0% to 40%: heterogeneity might not be important, 30% to 60%: may represent moderate heterogeneity, 50% to 90%: may represent considerable heterogeneity and over this is considerable heterogeneity. I^2 can be calculated from Q (chi-square) for the test of heterogeneity with the following formula[9];

$$I^2 = \left(\frac{Q - df}{Q}\right) \times 100\%$$

---

§ Imprecision refers to wide confidence intervals indicating a lack of confidence in the effect estimate. Based on GRADE recommendations, a result for continuous data (standardised mean differences, not weighted mean differences) is considered imprecise if the upper or lower confidence limit crosses an effect size of 0.5 in either direction, and for binary and correlation data, an effect size of 0.25. GRADE also recommends downgrading the evidence when sample size is smaller than 300 (for binary data) and 400 (for continuous data), although for some topics, these criteria should be relaxed[11].

‖ Indirectness of comparison occurs when a comparison of intervention A versus B is not available but A was compared with C and B was compared with C that allows indirect comparisons of the magnitude of effect of A versus B. Indirectness of population, comparator and/or outcome can also occur when the available evidence regarding a particular population, intervention, comparator, or outcome is not available and is therefore inferred from available evidence. These inferred treatment effect sizes are of lower quality than those gained from head-to-head comparisons of A and B.

Margarete Ainsworth Building, Barker Street, Randwick NSW 2031. Phone: 02 9399 1000. Email: info@neura.edu.au

References

1. Gold C, Heldal TO, Dahle T, Wigram T. Music therapy for schizophrenia or schizophrenia-like illnesses. *Cochrane Database of Systematic Reviews*. 2005(2):CD004025.
2. Edwards J. Music therapy in the treatment and management of mental disorders. *Irish Journal of Psychological Medicine*. 2006;23(1):33-5.
3. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMAGroup. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *British Medical Journal* 2009;151(4):264-9.
4. GRADEWorkingGroup. Grading quality of evidence and strength of recommendations. *British Medical Journal*. 2004;328:1490.
5. Gold C, Solli HP, Kruger V, Lie SA. Dose-response relationship in music therapy for people with serious mental disorders: systematic review and meta-analysis. *Clinical Psychology Review*. 2009;29(3):193-207.
6. Silverman MJ. The influence of music on the symptoms of psychosis: a meta-analysis. *Journal of Music Therapy*. 2003;40(1):27-40.
7. Mossler K, Chen X, Heldal TO, Gold C. Music therapy for people with schizophrenia and schizophrenia-like disorders. *Cochrane Database of Systematic Reviews*. 2011;12:CD004025.
8. Tseng P-T, Chen Y-W, Lin P-Y, Tu K-Y, Wang H-Y, Cheng Y-S, Chang Y-C, Chang C-H, Chung W, Wu C-K. Significant treatment effect of adjunct music therapy to standard treatment on the positive, negative, and mood symptoms of schizophrenic patients: A meta-analysis. *BMC Psychiatry Vol 16 2016, ArtID 16*. 2016;16.
9. CochraneCollaboration. Cochrane Handbook for Systematic Reviews of Interventions. 2008: Accessed 24/06/2011.
10. Rosenthal JA. Qualitative Descriptors of Strength of Association and Effect Size. *Journal of Social Service Research*. 1996;21(4):37-59.
11. GRADEpro. [Computer program]. Jan Brozek, Andrew Oxman, Holger Schünemann. *Version 32 for Windows* 2008.