TITLE: Group Therapy for Adult Psychiatric Inpatients: A Review of Clinical Evidence and Effectiveness

DATE: May 14, 2012

CONTEXT AND POLICY ISSUES

Group therapy is a form of psychotherapy in which patients are treated together in a group format. It was introduced in the early 1900s to treat patients with psychiatric disorders including schizophrenia. In addition to the usual pharmaceutical therapy used in clinical practice, it has become a common treatment model for inpatients as well as outpatients. Compared to outpatients, inpatients have more severe disease and require more intensive medical care; in addition, treatment time may be compressed.

The primary goals for inpatient group therapy are to improve interpersonal skills, to decrease symptoms and to speed up discharge. Diverse approaches have been developed to accomplish these goals. The common group therapy approaches include the cognitive behavioral, educative, interpersonal/interactive, psychodynamic, group-as-a-whole, and creative approaches. An integrative approach that combines educational, psychodynamic and interpersonal theories and techniques have also been introduced by researchers.

Previous research has indicated that not all patients with psychiatric disorders may benefit from group psychotherapy. Patients with paranoid symptoms, acute psychosis, severe depression, or those who were organically damaged may not be appropriate candidates for group therapy.

The purposes of this review are to identify the clinical evidence on group therapy for adult psychiatric inpatients and to examine the clinical effectiveness of group therapy for these patients.

RESEARCH QUESTIONS

1. What is the clinical evidence on group therapy for adult psychiatric inpatients?

2. What is the clinical effectiveness of group therapy for adult psychiatric inpatients?
KEY MESSAGE

Conclusions regarding the clinical effectiveness of group therapy in psychiatric inpatients cannot be drawn due to the limited quantity and quality of the evidence.

METHODS

Literature Search Strategy

A limited literature search was conducted on key resources including MEDLINE, PubMed, PsycINFO, The Cochrane Library (2012, Issue 4 of 12), University of York Centre for Reviews and Dissemination (CRD) databases, Canadian and abbreviated lists of major international health technology agencies, as well as a focused Internet search. No filters were applied to limit the retrieval by study type. Where possible, retrieval was limited to the human population. The search was also limited to English language documents published between January 1, 2007 and April 13, 2012.

Selection Criteria and Methods

One reviewer screened the titles and abstracts of the retrieved publications and evaluated the full-text publications for the final article selection, according to selection criteria presented in Table 1.

Table 1: Selection Criteria

<table>
<thead>
<tr>
<th>Population</th>
<th>Adult psychiatric inpatients in a general psychiatric unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intervention</td>
<td>Group psychotherapy</td>
</tr>
<tr>
<td>Comparator</td>
<td>Pharmacotherapy, standard of care for individual therapy</td>
</tr>
<tr>
<td>Outcomes</td>
<td>Violence, Patient engagement, Patient satisfaction, Self-esteem, Discharge from hospital, Clinical benefit (i.e. alleviating psychiatric symptoms)</td>
</tr>
<tr>
<td>Study Designs</td>
<td>Health technology assessments, systematic reviews, meta-analyses, randomized controlled trials and non-randomized controlled trials.</td>
</tr>
</tbody>
</table>

Exclusion Criteria

Studies were excluded if they did not meet the selection criteria, were duplicate publications, were abstracts/ conference proceedings, or were published prior to 2007. Studies with pre-post-test design but without a control group were excluded. Studies were also excluded when it was unclear whether the subjects were inpatients or outpatients.

Critical Appraisal of Individual Studies

The quality of the included randomized controlled trials (RCTs) was examined using the Jadad scale and Schulz’s concealment of treatment allocation tool. Non-randomized controlled trials (non-RCTs) were evaluated using the SIGN-50 instrument. A numeric score was not calculated for each study. Instead, the strengths and weakness of each study were summarized and described.
SUMMARY OF EVIDENCE:

Quantity of Research Available

The literature search yielded 555 citations. Upon screening titles and abstracts, 521 citations were excluded, and 34 potentially relevant articles were retrieved for full-text review. Of the 34 potentially relevant reports, 29 did not meet the inclusion criteria, and thus five publications were included in this review. The study selection process is outlined in a Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flowchart (Appendix 1). Two RCTs and three non-RCTs met the inclusion criteria. No health technology assessments or systematic reviews were identified.

Summary of Study Characteristics

One Finnish
\(^{10}\) and one German\(^{11}\) RCT were included. One non-RCT was conducted in Taiwan,\(^{12}\) one\(^{13}\) in Australia, and one\(^{14}\) in South Korea. The studies used questionnaires/scales that were administered at baseline and at the end of the trials. A summary of the trial characteristics of these studies are presented in Appendix 2. A list of instruments used to measure psychiatric symptoms and other outcomes from these studies is provided in Appendix 3.

Randomized Controlled Trials

Aho-Mustonen and coworkers examined the effect of group psychoeducation in long-term offender patients with schizophrenia living in a high-security hospital in Finland.\(^{10}\) The patients admitted to this hospital underwent involuntary psychiatric treatment when they were judged to lack criminal responsibility for the offences due to mental illness, primarily schizophrenia. Thirty-nine patients were randomized to receive the group psychoeducation program in addition to treatment as usual (TAU) (19 patients), or TAU alone (20 patients). The intervention consisted of eight group sessions, which were held weekly and lasted 45 to 60 minutes each. The patients were followed for three months after the treatment. The assessors were blinded about the treatment allocation. Facilitators of the group therapy took a two-day training session on leading the group. The TAU approach contained medication, psychosocial rehabilitation efforts and usual ward activities. The Finnish version of *Knowledge about Schizophrenia Questionnaire* (KASQ) and *Scale to Assess Unawareness of Mental Disorder* (SUMD) were used to assess patient’s knowledge of schizophrenia. The *Brief Psychiatric Rating Scale* (BPRS) and the *Beck Depression Inventory-II* (BDI-II) were adopted to measure the change in symptoms before and after treatment. The *Rosenberg Self-Esteem Scale* (RSE) and Sintonen’s 15D instrument were used to address self-esteem and health-related quality of life, respectively. The baseline patient characteristics were comparable with respect to age, gender, education, disease duration and comorbidities. There were no significant differences between the groups regarding the baseline outcome measure scores.

A German study conducted by Ulrich et al. evaluated the effect of group music therapy in schizophrenic patients who were hospitalized for acute care.\(^{11}\) Thirty-seven patients were randomized to the intervention group (21 patients) or the control group (16 patients). In addition to standard care, the intervention group underwent group music therapy consisting of 7.5 sessions in five weeks. Each session lasted 45 minutes. The main activities for the participants in the intervention group were playing together on rhythm instruments and singing, as well as group discussions. There was no information provided regarding the treatment modality in the
control group. The outcomes of interest were interpersonal contact [subscales of the Gießentest, completed by patients (GTS) or nurses (GTFm)], negative symptoms [measured with German version of the Scale for the Assessment of Negative Symptoms (SANS), completed by nurses], and quality of life [measured with the Scales for Mental Health (SPG), completed by patients]. The outcomes were assessed before and after treatment. At baseline, patient characteristics were similar between the two groups in terms of age, gender, time between hospital admission and pre-treatment, as well as the scores of GTS, GTFm, SANS and SPG.

Non-Randomized Controlled Trials

Peng et al. investigated the effect of group music activity in addition to standard care in inpatients with acute schizophrenia. Adult patients admitted to an acute psychiatric ward in Taiwan were screened and 67 of them were allocated to either group music therapy in addition to usual care (32 patients) or usual care alone (35 patients). Both groups received standard medications. The first patient was randomly assigned. For the remainder of the patients, the allocation was carried out by alternating assignment, based on order of admission. The group music therapy was a two-week program comprised of ten 50-minute sessions. The activities included singing and listening to music. Each group consisted of four to five patients. A Chinese version of BPRS was used to assess the severity of psychotic symptoms in each patient before and after the treatment. The BPRS was assessed by two psychiatric nurses with 16 and five years of clinical experience. Six patients in the intervention group and two patients in the control group withdrew from the study due to early discharge or disease exacerbation. Before the study, there were no significant differences in age, gender, and number of previous hospitalization between the two groups.

Dean et al. evaluated the effectiveness of a Motivational Enhancement Therapy (MET) group program in patients admitted to an Australian inpatient eating disorders unit. Eligible patients were treated with MET (23 patients) in addition to TAU, or TAU alone (19 patients). MET was a four-session intervention which was developed specifically for eating disordered patients, with an intention to enhance their motivations to overcome the illness and to increase willingness to engage in the recovery process. The group sessions were held weekly and lasted 1.25 hours each. They were open which means that new participants were allowed to join each week as they entered the hospital program. The MET sessions were completed prior to the TAU control data being collected. The study outcome was the measures related to motivation evaluated using modified Anorexia nervosa stages of change questionnaire (ANSOCQ), Self efficacy scale for anorexia nervosa (SES), Treatment engagement questionnaire, Eating disorder inventory II (EDI-II), Eating disorders examination questionnaire (EDE-Q), and BDI-II. All measures were administered before and after treatment. The patients were followed for six weeks; however, the authors did not specify if the study duration was six weeks in total, or if patients were followed for six weeks after the treatment. All patients were female. The percentages of patients who did not complete the post-treatment assessment were similar between MET and TAU groups [4 patients (17%) versus 3 patients (16%), respectively]. At baseline, there was no statistically significant difference between the two groups for age, marital status, body mass index, disease duration, or self-induced vomiting and laxative abuse. The two groups were comparable on all measures from the questionnaires before the treatment, except for two subscales of the EDI-II: patients in the TAU group reported significantly higher levels of perfectionism and a greater desire to retreat to the securities of childhood, than did the MET group.
Another study was carried out by Choi et al., to investigate the effectiveness of group music therapy for improving depression, anxiety and relationships in psychiatric patients. The patients were admitted to a psychiatric hospital in South Korea due to a range of psychiatric illnesses including schizophrenia, psychotic disorder, bipolar disorder, anxiety disorder and mental retardation. They were non-randomly allocated to the music therapy group or the routine care control group. Patients in the music therapy group attended a total of fifteen 60-minute sessions, once or twice a week. All the certified music therapists had over eight years’ experience in the application of group music intervention for psychiatric inpatients. Depression was measured by BDI, anxiety by the State and Trait Anxiety Inventory (STAI), and relationships by the Relationship Change Scale (RCS), before and after 15 sessions of treatment. Patients in the control group did not receive additional treatment to routine care. In total, 26 patients were enrolled in the study, with 13 in the group music therapy group and 13 in the routine care group. There were no dropouts during the study period. Patients in each treatment group were similar with regards to age, sex, and baseline questionnaire scores for depression, anxiety and relationships.

**Summary of Critical Appraisal**

Details on the critical appraisal of individual studies are presented in Appendix 4.

Two RCTs were included in our report. The methods of randomization were described in both. In the Aho-Mustonen study, the interviewers were blinded to patients’ treatment allocation to diminish the risk of interviewer bias. No patients withdrew from the program during the treatment period. Intention-to-treat analyses were performed to compare the effect of intervention to standard care in the follow-up period of the study. In the Ulrich study, it is unclear if the interviewers were blinded to the treatment allocation. The method of statistical analysis was not provided. Approximately 21% of the originally randomized patients (n=47) refused to take part in the study without giving a reason. Data were analyzed based on the number of patients that received the assigned treatment. Results from this study may not be reliable due to the high patient attrition in a small population.

For non-RCTs, quality of these studies was compromised due to the nature of the study design, in that patients were allocated to different treatment modalities according to the order of hospital admissions, or physician/patient’s desire. Selection bias is likely to be introduced in this manner. The Peng study was labeled as quasi-RCT, where patients were not truly randomized to ensure the comparability between treatment groups. Patient characteristics other than age, gender and number of previous hospitalization were lacking, especially the baseline scale scores. In the Dean study, patients in the control group were more likely to be more educated, be perfectionists, and have greater desire to retreat to the securities of childhood than those in the intervention group. It is unclear what impact of these unbalanced characteristics on the study results. Potential confounders such as disease severity and duration, co-morbidities, previous and concomitant treatment, and experience of the therapists were not elaborated either. In the Choi study, trial and patient characteristics were not reported in sufficient details, such as study duration, length of follow-up, disease status, methods for treatment allocation and methods to adjust for potential confounders. The patients had varied psychiatric diagnosis. A description of the control group was lacking. It is challenging to compare the two treatments based on limited information.

Sample size is an issue for all included RCTs and non-RCTs. The number of participants in these studies ranged from 26 to 67. Power or sample size calculation was not conducted in any
of them. The corresponding large standard deviations for the between-group mean differences reported in most of the studies make it difficult to draw conclusions regarding the effectiveness of group therapy.

None of the included studies were conducted in a Canadian setting. Considering the quantity and quality of the available evidence, generalizability of the study results to Canadian patients is uncertain.

Summary of Findings

Details of the main study findings and authors’ conclusions are presented in Appendix 5.

Results from the various questionnaires in the Aho-Mustonen study\textsuperscript{10} indicated that patients in the group psychoeducation program had significantly better understanding about illness at 3-month follow up (p=0.04) and improved self-esteem immediately after the treatment (p=0.03) comparing those treated with TAU. No significant differences were detected between the intervention and control groups in terms of patients’ insight into the illness immediately after the treatment, change in clinical symptoms and health-related quality of life post-treatment or at 3-month follow up. Within-group comparisons that measured the difference before and after the treatment were not performed.

Ulrich and coworkers\textsuperscript{11} found that five-week group music therapy significantly improved interpersonal contact measured by patients (p=0.03) and diminished negative symptoms (p=0.01) in schizophrenic inpatients, when compared with standard care. Group music therapy along with standard care did not show added clinical benefit over standard care alone in improving interpersonal contact measured by medical professionals or health-related quality of life. Within-group comparisons that measured the difference before and after the treatment were not performed.

In the study by Peng et al.,\textsuperscript{12} patients receiving group music therapy in addition to standard care for acute schizophrenia significantly reduced the disease severity compared to those in the control group, according to the post-treatment scores measured by the total and subscales of the BPRS. In particular, group music therapy improved the symptoms relating to anxiety, conceptual disorganization, hallucinatory behavior, uncooperativeness, and blunted affect.

In the study by Dean et al.,\textsuperscript{13} self-report questionnaires were used to measure patient’s motivation to recover, treatment engagement, and eating disorder psychopathology and behavior. The before-after differences within each study group were reported, without providing p values or confidence interval. There were no significant differences between the MET and TAU groups on any assessment measures, except that patients in the TAU group reported a significantly greater reduction on the Drive For Thinness subscale of the EDI-II, p=0.041.

In the Choi et al. study,\textsuperscript{14} patients receiving group music therapy had significant symptom improvement in all four measures: depression, state anxiety, trait anxiety, and relationship compared to routine care (all p < 0.001).

Limitations

The evidence was limited to two RCTs and three non-RCTs. The investigated interventions include a psychoeducation program, music therapy and a motivational enhancement program
carried out in group formats. The findings of the studies are specific to these group therapies. They would only be applicable to patients with similar psychiatric disorders and may not be generalizable at large.

The quality of the randomized studies was compromised by high patient attrition rate, or uncertainty about investigator blinding. The quality of the included studies was lower due to the non-randomized study design in three studies (bias introduction), small sample sizes (26 to 67 patients), and insufficient data reporting (patient characteristics, therapist experience, and study design).

The duration of the trials varied from two weeks to 15 weeks. Evidence of longer-term benefits of group therapy in the study population is lacking.

CONCLUSIONS AND IMPLICATIONS FOR DECISION OR POLICY MAKING:

In total, two randomized controlled trials and three non-randomized controlled trials have been examined in the current report. The conditions examined include schizophrenia, eating disorder, bipolar disorder, anxiety disorder, psychotic disorder, conduct disorder and mental retardation. Sample size ranged from 26 to 67 inpatients. The group programs were carried out over two to 15 weeks.

Evidence from three studies comparing group music therapy with standard care suggests that music therapy significantly improves clinical symptoms and better interpersonal contact in psychiatric patients, although it has no beneficial effect on patient’s health-related quality of life. A group psychoeducation program examined in forensic patients diagnosed with schizophrenia increased patient’s knowledge of the illness and enhanced their self-esteem, when compared with treatment as usual. A four-week Group Motivational Enhancement Therapy as an adjunct to inpatient treatment did not demonstrate apparent clinical benefits over routine care in patients with eating disorder in a non-randomized study. Conclusions regarding the clinical effectiveness of group psychotherapy cannot be drawn due to the small sample sizes and low quality of these studies.

Conducting clinical trials in inpatients with psychiatric diseases can be challenging because of the disease characteristics, patient’s perception of the treatment, and high patient turnover in hospital. Treatment effect may also be influenced by the working experience of the therapists. However, this should not preclude from well-designed large clinical trials with longer-term follow-up, especially in a Canadian setting, to collect compelling evidence to evaluate the clinical effectiveness of group psychotherapy in inpatients with psychiatric disorders.

PREPARED BY:
Canadian Agency for Drugs and Technologies in Health
Tel: 1-866-898-8439
www.cadth.ca
REFERENCES


APPENDIX 1: Selection of Included Studies

555 citations identified from electronic literature search and screened

→ 521 citations excluded

34 potentially relevant articles retrieved for scrutiny (full text, if available)

→ 0 potentially relevant reports retrieved from other sources (grey literature, hand search)

→ 34 potentially relevant reports

29 reports excluded:
- irrelevant population (5)
- irrelevant intervention (4)
- irrelevant comparator (8)
- irrelevant outcome (3)
- irrelevant study design (9)

→ 5 reports included in this report:
- 2 RCTs
- 3 non-RCTs
## APPENDIX 2. Summary of Study Characteristics

<table>
<thead>
<tr>
<th>First Author, Publication Year, Country</th>
<th>Study Design</th>
<th>Patient Characteristics</th>
<th>Intervention</th>
<th>Comparator</th>
<th>Clinical Outcomes Measured</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Randomized controlled trial</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aho-Mustonen, 2011, Finland¹⁰</td>
<td>RCT</td>
<td>Long-term offender patients with schizophrenia in a high-security hospital</td>
<td>Group psychoeducation 19 patients</td>
<td>TAU 20 patients</td>
<td>Knowledge of schizophrenia (KASQ in Finland and SUMD), Change in symptoms pre- and post-treatment (BPRS and BDI-II), Self-esteem (RSE), and QOL (Sintonen’s 15D)</td>
</tr>
<tr>
<td>Ulrich, 2007, Germany¹¹</td>
<td>RCT</td>
<td>Inpatients with schizophrenia</td>
<td>Group music therapy + standard care 21 patients</td>
<td>Standard care 16 patients</td>
<td>Interpersonal contact (GTS and GTFm), negative symptoms (SANS), and QOL (SPG)</td>
</tr>
<tr>
<td><strong>Non-randomized controlled trials</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peng, 2010, Taiwan¹²</td>
<td>Quasi-RCT</td>
<td>Inpatients with schizophrenia</td>
<td>Group music activity + standard care 32 patients</td>
<td>Standard care 35 patients</td>
<td>BPRS (Chinese version) to assess the severity of clinical symptoms</td>
</tr>
<tr>
<td>Dean, 2008, Australia¹³</td>
<td>Non-RCT, prospective study</td>
<td>Inpatients with eating disorder</td>
<td>MET group program 23 patients</td>
<td>TAU 19 patients</td>
<td>Motivations measured by ANSOCQ, SES, EDI-II, EDE-Q, and BDI-II</td>
</tr>
<tr>
<td>Choi, 2008, South Korea¹⁴</td>
<td>Non-RCT, prospective study</td>
<td>Inpatients with various psychiatric disorders</td>
<td>Group music intervention 13 patients</td>
<td>Routine care 13 patients</td>
<td>Questionnaires assessed depression (BDI), anxiety (STAI), and relationship (RCS)</td>
</tr>
</tbody>
</table>

ANSOCQ=Anorexia nervosa stages of change questionnaire; BDI=The Beck depression inventory; BDI-II=The Beck depression inventory II; BPRS=the Brief Psychiatric Rating Scale; EDE-Q=Eating disorders examination questionnaire; EDI-II=Eating disorder inventory II; GTS=Gießentest (completed by patients); GTFm=Gießentest (completed by medical professionals); KASQ=Knowledge about Schizophrenia; MET=Motivational Enhancement Therapy; OQL=quality of life; RCS=Relationship Change Scale; RCT=randomized controlled trial; RSE=Rosenberg Self-Esteem Scale; SANS=Scale for the Assessment of Negative Symptoms; SES=Self efficacy scale for anorexia nervosa; SPG=Scales for Mental Health; STAI=State and Trait Anxiety Inventory; SUMD=Scale to Assess Unawareness of Mental Disorder; TAU=treatment as usual
APPENDIX 3: Instruments Used in the Studies Reviewed in this Report

ANSOCQ
The Anorexia nervosa stages of change questionnaire is a 20-item self-report measure assessing a broad range of anorexic symptomatology including eating behaviors, aspects of weight, emotional difficulties, weight control methods and interpersonal difficulties. Its internal consistency and reliability have been demonstrated.\(^\text{13}\)

BDI-II
The Beck Depression Inventory II is a widely used 21-item self-report questionnaire of depressive symptoms and attitudes experienced over the preceding two weeks. It shows acceptable levels of internal consistency and test-retest reliability.\(^\text{10,13}\)

BPRS
The Brief Psychiatric Rating Scale is a widely used 18-item psychometric scale for assessing the severity of clinical symptoms of schizophrenic patients. Each item is scored from 0 (not present) to 4 (extremely severe). These items are divided into five subscales: affect subscale, positive symptoms subscale, negative symptoms subscale, resistance subscale, and activation subscale.\(^\text{10,12}\)

EDE-Q
The Eating Disorders Examination Questionnaire is a 36-item self-report questionnaire. It was developed to obtain information regarding eating disorder symptomatology. It contains four subscales (Shape Concern, Weight Concern, Dietary Restraint and eating Concern) using a seven-point forced choice format. They show sound test-retest reliability.\(^\text{13}\)

EDI-II
The Eating Disorder Inventory II is a widely used self-report measure assessing a number of psychological and behavioral traits common in eating disorders. It contains 11 subscales, and higher scores reflect greater levels of pathology. Its reliability has been demonstrated.\(^\text{13}\)

Gießentest
A German test to assess interpersonal contact such as social relations, and closeness with or dependency on other people. It can be completed by patients (GTS), or medical professionals (GTFm).\(^\text{11}\)

KASQ
The Knowledge about Schizophrenia Questionnaire is a 25-item multiple-choice questionnaire designed and validated for assessment of patients’ knowledge of schizophrenia and its management.\(^\text{10}\)

RCS
The Relationship Change Scale provides a measure of the improvement or deterioration in the general quality of the couples’ relationships between testing sessions.\(^\text{15}\)

RSE
The Rosenberg Self-Esteem Scale measures the patients’ global self-esteem (range 10-40).\(^\text{10}\)
SANS
The *Scale for the Assessment of Negative Symptoms* is used to assess negative symptoms in patients with schizophrenia.\textsuperscript{11}

SES
The *Self Efficacy Scale for Anorexia Nervosa* examines the strength of one’s belief that they could successfully do a given task, using a 10-point Likert scale. The questions in this scale corresponded to the items on the ANSOCQ, and this questionnaire shows a high positive correlation with the ANSOCQ.\textsuperscript{13}

SPG
The *Scales for Mental Health* is a German self-reported tool to assess patient's quality of life. It contains 76 items and seven scales.\textsuperscript{11}

STAI
The *State and Trait Anxiety Inventory* is self-report scales to assess both state and trait anxiety in research or clinical practice. Its internal consistency and test-retest stability have been demonstrated.\textsuperscript{16}

Stintonen’s 15D
The *Stintonen’s 15D* instrument measures patients’ health-related quality of life. A single index score can be calculated (range 0=dead to 1=full health).\textsuperscript{10}

SUMD
The *Scale to Assess Unawareness of Mental Disorder* is a semi-structured interview scale to assess insight, in terms of patient’ present state of knowledge concerning 1) basic awareness of illness, 2) the effects of medication upon the illness and 3) a general understanding of the social consequences of the illness. Scoring is from 1 (aware) to 3 (severely unaware). A total SUMD score (range 3 – 6) is created by combining individual scores from three aforementioned SUMD items to reflect a broader measure of insight.\textsuperscript{10}

Treatment Engagement questionnaire
It captures the degree to which participants were interested in taking part in aspects of the hospital treatment program. Five scales ranging from 0 (not at all interested) to 10 (extremely interested) assessed interest in obtaining feedback from medical tests, participating in therapy groups, having individual therapy sessions with clinicians, having individual sessions with a dietician, and the extent to which participants ate meals for themselves or for others. These scales have been found to be significantly positively correlated to the ANSOCQ.\textsuperscript{13}
## APPENDIX 4: Summary of Critical Appraisal

<table>
<thead>
<tr>
<th>First Author, Publication Year</th>
<th>Strengths</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Randomized controlled trials</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Aho-Mustonen, 2010<sup>10</sup> | • Method of randomization described  
• Assessor were blinded to treatment allocation  
• Patient withdrawal described  
• ITT analysis adopted | • No power calculation  
• Small sample |
| Ulrich, 2007<sup>11</sup> | • Method of randomization described  
• Patient withdrawal described | • Not clear if the assessor was blinded to the treatment  
• No power calculation  
• Small sample  
• Statistical methods not described  
• ITT analysis not used |
| **Non-randomized controlled trials** |           |             |
| Peng, 2010<sup>12</sup> | • Objectives, inclusion criteria, and outcome measures clearly described | • Non-randomized study design  
• No sample size calculation  
• Potential confounders not identified  
• Results not reported in details |
| Dean, 2008<sup>13</sup> | • Objectives, inclusion criteria, and outcome measures clearly described  
• Statistical methods to adjust for baseline differences were adopted | • Non-randomized study design led to unbalanced baseline characteristics, even if statistical methods were used to adjust the unbalance  
• No sample size calculation  
• Potential confounders not identified |
| Choi, 2007<sup>14</sup> | • Objectives, inclusion criteria, and outcome measures clearly described | • Non-randomized study design  
• Insufficient reporting on baseline patient characteristics and treatment allocation method  
• No sample size calculation or justification  
• Potential confounders not identified |

ITT=intention-to-treat
### APPENDIX 5: Summary of Findings for Group Therapy

<table>
<thead>
<tr>
<th>First Author, Publication Year</th>
<th>Main Study Findings</th>
<th>Authors’ Conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Randomized controlled trials</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Aho-Mustonen, 2010\textsuperscript{10} | KASQ (change from baseline, mean±SD):  
Psychoeducation:  
Post-treatment: 0.7±2.4  
3-month follow up: 1.4±2.8  
TAU:  
Post-treatment: -0.3±2.7, p=0.24  
3-month follow up: -0.5±2.6, p=0.04  
SUMD total (change from baseline, mean±SD):  
Psychoeducation:  
Post-treatment: -0.4±0.9  
3-month follow up: -0.7±1.1  
TAU:  
Post-treatment: -0.3±1.0, p=0.67  
3-month follow up: -0.2±0.8, p=0.09  
BPRS total (change from baseline, mean±SD):  
Psychoeducation:  
Post-treatment: -3.2±7.5  
3-month follow up: -4.6±8.0  
TAU:  
Post-treatment: -4.4±5.4, p=0.57  
3-month follow up: -4.0±4.6, p=0.76  
BDI-II (change from baseline, mean±SD):  
Psychoeducation:  
Post-treatment: -0.8±5.7  
3-month follow up: -2.5±6.2  
TAU:  
Post-treatment: -2.0±4.5, p=0.46  
3-month follow up: -0.1±7.9, p=0.30  
RSE (change from baseline, mean±SD):  
Psychoeducation:  
Post-treatment: 2.3±3.6  
3-month follow up: 2.0±2.8  
TAU:  
Post-treatment: -0.4±3.9, p=0.03  
3-month follow up: -0.2±4.3, p=0.06  
15D (change from baseline, mean±SD):  
Psychoeducation:  
Post-treatment: 0.00±0.08  
3-month follow up: 0.00±0.06  
TAU:  
Post-treatment: 0.01±0.06, p=0.50  
3-month follow up: 0.04±0.08, p=0.09  | Psychoeducation showed positive effect on knowledge about illness and self-esteem compared with TAU. |
| Ulrich, 2007\textsuperscript{11} | GTS (before and after treatment, mean±SD):  
Group music therapy |                      |

Group Therapy for Adult Psychiatric Inpatients
<table>
<thead>
<tr>
<th>First Author, Publication Year</th>
<th>Main Study Findings</th>
<th>Authors' Conclusions</th>
</tr>
</thead>
</table>
|                               | Music therapy: 0.32±0.94 to 0.70±0.93  
Control: 0.27±1.28 to 0.07±1.04, p=0.03  
GTFm (before and after treatment, mean±SD):  
Music therapy: -0.20±0.83 to -0.05±0.81  
Control: -0.17±0.77 to -0.26±0.85, p=0.44  
SANS (before and after treatment, mean±SD):  
Music therapy: 1.09±0.66 to 0.72±0.74  
Control: 0.70±0.59 to 1.08±0.99, p=0.01  
SPG (before and after treatment, mean±SD):  
Music therapy: 2.90±0.47 to 3.01±0.44  
Control: 2.83±0.33 to 2.99±0.37, p=0.67 | significantly improved interpersonal contact assessed by patients and diminished negative symptoms. |

Non-randomized controlled trials

| Peng, 2010 | Total BPRS scores (post-treatment scores, mean±SD):  
Music: 5.08±2.56  
Control: 8.30±2.56, p<0.001  
BPRS subscale scores (post-treatment scores, mean±SD):  
Affect subscale:  
Music: 0.17±1.03  
Control: 0.89±1.03, p=0.011  
Positive symptoms subscale:  
Music: 3.44±1.03  
Control: 4.44±1.03, p<0.001  
Negative symptoms subscale:  
Music: 0.23±0.77  
Control: 0.75±0.77, p=0.012  
Resistance subscale:  
Music: 0.58±0.65  
Control: 1.01±0.64, p=0.017  
Activation subscale:  
Music: 0.84±0.80  
Control: 1.04±0.81, p=0.338 | Group music therapy significantly reduced the disease severity for patients with acute schizophrenia based on the scores in total and subscales of the BPRS. |

| Dean, 2008 | ANSOCQ scores (post-pre treatment difference, mean±SD):  
MET: 8.98±7.10  
TAU: 10.97±14.61, p=0.604  
BDI-II scores (post-pre treatment difference, mean±SD):  
MET: 4.39±8.50  
TAU: 10.09±14.87, p=0.166 | There were few overall statistical differences between the 2 groups. The TAU group reported a significantly greater reduction on the Drive For Thinness subscale of the EDI-II. |
<table>
<thead>
<tr>
<th>First Author, Publication Year</th>
<th>Main Study Findings</th>
<th>Authors’ Conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>EDE-Q subscales (post-pre treatment difference, mean±SD):</strong></td>
<td><strong>Eating Concern:</strong></td>
<td>Group music therapy may improve depression, anxiety and relationship in psychiatric patients in this small non-RCT. The</td>
</tr>
<tr>
<td></td>
<td>MET: -1.67±1.12</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TAU: -1.49±1.19, p=0.426</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Shape Concern:</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MET: -0.57±0.99</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TAU: -0.53±0.76, p=0.891</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Weight Concern:</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MET: -0.54±1.02</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TAU: -0.68±1.04, p=0.701</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Vomiting episodes:</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MET: -22.00±45.77</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TAU: -35.64±37.36, p=0.403</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Laxative abuse:</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MET: 0.00±0.43</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TAU: -2.00±4.76, p=0.161</td>
<td></td>
</tr>
<tr>
<td><strong>EDI-II subscales (post-pre treatment difference, mean±SD):</strong></td>
<td><strong>Drive For Thinness:</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MET: -1.38±3.78</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TAU: -4.28±4.28, p=0.041</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Body Dissatisfaction:</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MET: -0.37±2.54</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TAU: -1.66±4.25, p=0.276</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Perfectionism:</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MET: 0.00±5.14</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TAU: -0.75±2.65, p=0.602</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Maturity Fears:</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MET: 0.07±3.05</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TAU: -2.06±3.42, p=0.060</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Social Insecurity:</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MET: -1.37±2.91</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TAU: -2.44±4.49, p=0.402</td>
<td></td>
</tr>
<tr>
<td><strong>SES scores (post-pre treatment difference, mean±SD):</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MET: 17.44±19.73</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TAU: 16.40±20.33, p=0.881</td>
<td></td>
</tr>
<tr>
<td><strong>Treatment Engagement scores (post-pre treatment difference, mean±SD):</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MET: 3.22±5.75</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TAU: 3.13±3.28, p=0.953</td>
<td></td>
</tr>
</tbody>
</table>

**Choi, 2008**

Depression (from baseline to post-treatment, mean±SD): 
Music: from 49.3±3.1 to 25.5±2.2 
Control: from 47.7±2.8 to 44.8±3.8 
P<0.001
<table>
<thead>
<tr>
<th>First Author, Publication Year</th>
<th>Main Study Findings</th>
<th>Authors' Conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>State anxiety:</td>
<td>authors suggested</td>
</tr>
<tr>
<td></td>
<td>Music: from 36.5±0.8 to 22.8±1.7</td>
<td>conducting a larger RCT to provide more solid evidence.</td>
</tr>
<tr>
<td></td>
<td>Control: from 36.2±1.2 to 32.5±1.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>P&lt;0.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Trait anxiety:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Music: from 35.9±1.0 to 23.0±1.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Control: from 36.4±1.6 to 34.2±1.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>P&lt;0.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Relationship:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Music: from 72.4±1.2 to 45.8±2.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Control: from 72.5±2.3 to 66.7±2.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>P&lt;0.001</td>
<td></td>
</tr>
</tbody>
</table>

ANSOCQ=Anorexia nervosa stages of change questionnaire; BDI-II=The Beck depression inventory II; BPRS=The Brief Psychiatric Rating Scale; EDE-Q=Eating disorders examination questionnaire; EDI-II=Eating disorder inventory II; GTS=Gießentest (completed by patients); GTFm=Gießentest (completed by medical professionals); KASQ=Knowledge about Schizophrenia Questionnaire; MET=Motivational Enhancement Therapy; RCT=randomized controlled trial; RSE=Rosenberg Self-Esteem Scale; SANS=Scale for the Assessment of Negative Symptoms; SD=standard deviation; SES=Self efficacy scale for anorexia nervosa; SPG=Scales for Mental Health; SUMD=Scale to Assess Unawareness of Mental Disorder; TAU=treatment as usual