Cochrane Database of Systematic Reviews | Review - Intervention

Humour-based interventions for people with schizophrenia

Yasushi Tsujimoto, Yuri Nakamura, Masahiro Banno, Kunihiro Kohmura, Hiraku Tsujimoto, Yuki Kataoka
Authors' declarations of interest

Version published: 13 October 2021 Version history
https://doi.org/10.1002/14651858.CD013367.pub2

Abstract

Background

Humour-based interventions are defined as any intervention that promotes health and wellness by stimulating a playful discovery, expression, or appreciation of the absurdity or incongruity of life's situations. Humour-based interventions can be implemented in different settings, including hospitals, nursing homes and day care centres. They have been posed as an adjunct to usual care for people with schizophrenia, but a summary of the evidence is lacking.

Objectives

To examine the effects of humour-based interventions as an add-on intervention to standard care for people with schizophrenia.

Search methods

On 31 July 2019 and 10 February 2021 we searched the Cochrane Schizophrenia Group's study-based register of trials, which is based on CENTRAL, CINAHL, ClinicalTrials.Gov, Embase, ISRCTN, MEDLINE, PsycINFO, PubMed, and WHO ICTRP.

Selection criteria

We included all randomised controlled trials comparing humour-based interventions with active controls, other psychological interventions, or standard care for people with schizophrenia. We excluded studies fulfilling our prespecified selection criteria but without useable data from further quantitative synthesis.

Data collection and analysis

Two review authors independently inspected citations, selected studies, extracted data and appraised study quality, following the guidance from the *Cochrane Handbook for Systematic Reviews of Interventions*. For binary outcomes we calculated risk ratios (RRs) and their 95% confidence intervals (CIs). For continuous outcomes we calculated the mean differences (MDs) and their 95% CIs. We assessed risks of bias for included studies and created summary of findings tables using the GRADE approach.

Main results

We included three studies in this review for qualitative synthesis, although one study did not report any relevant outcomes. We therefore include two studies (n = 96) in our quantitative synthesis. No data were available on the following prespecified primary outcomes: clinically-important change in general mental state, clinically-important change in negative symptoms, clinically-important change in overall quality of life, and adverse effects. As compared with active control, humour-based interventions may not improve the average endpoint score of a general mental state scale (Positive and Negative Syndrome Scale (PANSS) total score: MD −1.70, 95% CI −17.01 to 13.61; 1 study, 30 participants; very low certainty of evidence); positive symptoms (PANSS positive symptom score: MD 0.00, 95% CI −2.58 to 2.58; 1 study, 30 participants; low certainty of evidence), negative symptoms (PANSS negative symptom score: MD −0.70, 95% CI −4.22 to 2.82; 1 study, 30 participants; very low certainty of evidence) and anxiety (State-Trait Anxiety Inventory (STAI): MD −2.60, 95% CI −5.76 to 0.56; 1 study, 30 participants; low certainty of evidence). Due to the small sample size, we remain uncertain about the effect of humour-based interventions on leaving the study early as compared with active control (no event, 1 study, 30 participants; very low certainty of evidence). On the other hand, humour-based interventions may reduce depressive symptoms (Beck Depression Inventory (BDI): MD −6.20, 95% CI −12.08 to −0.32; 1 study, 30 participants; low certainty of evidence). Compared with standard care, humour-based interventions may not improve depressive symptoms (BDI second edition: MD 0.80, 95% CI −2.64 to 4.24; 1 study, 59 participants; low certainty of evidence). We are uncertain about the effect of humour-based interventions on leaving the study early for any reason compared with standard care (risk ratio 0.38, 95% CI 0.08 to 1.80; 1 study, 66 participants; very low certainty of evidence).

Authors' conclusions

We are currently uncertain whether the evidence supports the use of humour-based interventions in people with schizophrenia. Future research with rigorous and transparent methodology investigating clinically important outcomes is warranted.

Plain language summary

Available in **English** | Deutsch | Español | فارسی | Français | 日本語 | Bahasa Malaysia | Polski | Русский | 简体中文

Humour-based interventions for people with schizophrenia

Review question

Are humour-based interventions effective in treating people with schizophrenia?

Background

Schizophrenia is a serious mental illness. It is a disorder of thought, namely firm fixed false beliefs despite there being evidence to the contrary, loss of reality ties, and altered perception. These symptoms are further classified as (i) positive symptoms, such as speech without order, illusions or mistaken and persistent ideas; and (ii) negative symptoms, a lack of emotion or restricted quantity of speech; and decline in cognitive function, including attention, memory, and behavior control. The standard treatment for schizophrenia is antipsychotic medications. Treatment with humour-based interventions, such as watching humorous movies, funny videos, or comedies, has been proposed as an add-on treatment that promotes health and wellness by stimulating a playful discovery, expression, or appreciation of the irrationality or inconsistency of life's situations.

Searching for evidence

We ran an electronic search in February 2021 for trials that randomised people with schizophrenia to receive humour-based interventions in addition to usual care, or to receive usual care only, another psychological intervention or a control condition. We found eight records and checked them for suitability to include in our review.

Evidence found

Three trials met the review requirements and two low-quality trials (total number of participants = 96) provided useable data. Compared with active control, humour-based interventions may not improve positive symptoms and anxiety, but may improve depressive symptoms. However, when compared with standard care, humour-based intervention may not improve the depressive symptoms. Current evidence is very limited and is of low to very low quality. We are uncertain as to whether humour-based interventions may lead to clinically-important improvement in mental state or quality of life in people with schizophrenia.

Conclusions

There is insufficient research evidence to support the use of humour-based interventions in people with schizophrenia.

Authors' conclusions

Implications for practice

1. For people with schizophrenia

We remain uncertain as to whether humour-based interventions are effective in improving mental state or quality of life for people with schizophrenia. Limited data from small studies indicate that these interventions could be useful in improving their depressive symptoms, at least over a short period of time against active controls; however, humour-based interventions may have little or no effects on negative symptoms or anxiety.

2. For clinicians

Given the very limited evidence on our prespecified outcome measures, such as clinically-important change in mental state or quality of life, we are unable to draw solid conclusions on the effectiveness of humour-based interventions as an add-on treatment for schizophrenia care planning. However, clinicians may want to consider incorporating elements of humour-based interventions as an addition to standard care to improve depressive symptoms in people with schizophrenia.

3. For policy-makers

The current evidence base is limited and of low quality, offering little support for policy-makers.

Implications for research

1. General

There is certainly room for improvement in study methodology for sample size, follow-up duration, and trial reporting quality. Future research should use rigorous and transparent methods to allow for comprehensive evidence-based assessment. For future trial design perspectives, it is vital for trial investigators to assess important/relevant outcomes such as clinically-important change in general mental state, clinically-important change in negative symptoms, clinically-important change in overall quality of life, and adverse effects.

2. Specific

Based on our review development process and professional knowledge, we present in Table 1 a suggested outline for future trial design considerations when implementing humour-based interventions in people with schizophrenia. Authors of future studies are also encouraged to follow the appropriate reporting guidelines, such as the Consolidated Standards of Reporting Trials (CONSORT) Statement (Schulz 2010). Due to the nature of humour-based interventions, blinding study personnel could be difficult; we advise trialists to focus on the use of objective and clinically-important outcome measures, with careful consideration of the cut-off values for rating scales. Furthermore, a prespecified subgroup analysis stratified by the severity of depressive symptoms at baseline may help to identify the populations who are most likely to benefit from humour-based interventions.

Table 1. Future trial design **Open in table viewer**

Methods	Treatment allocation: randomised (or cluster-randomised), with methods on sequence generation and allocation concealment clearly described. Blinding: blinded participants and outcomes assessors Duration: at least 1 year Design: parallel-group Settings: outpatients and hospital inpatients

Participants	Diagnosis: clinical diagnosis of schizophrenia using DSM-5 or ICD-11 criteria History of schizophrenia: any Sample size: N = 400 (more if cluster-randomised) Age: ≥ 18 years Gender: any
Interventions	Treatment: Humour-based interventions with prespecified protocol offered by trained personnel Comparator: Standard care, or active controls without elements of humour
Outcomes	Mental state (binary outcome) Depressive symptoms (binary outcome) Anxiety (binary outcome) Relapses (binary outcome) Quality of life (binary outcome) Costs: cost of services, cost of care Adverse events related to humour-interventions (any) Service outcomes: days in hospital, time attending outpatient psychiatric clinic
Notes	Participant adherence to interventions should be reported Detail description of intervention according to TiDER checklist (Hoffmann 2014) should be made

Summary of findings

Open in table viewer

Summary of findings 1. Humour-based interventions compared to active controls for people with schizophrenia

Humour-based interventions compared to active controls for people with schizophrenia

Patient or population: People with schizophrenia
Intervention: Humour-based interventions
Comparison: Active controls

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Certainty of the evidence (GRADE)	Comments
	Risk with active controls	Risk with humour-based interventions				
Average endpoint on a general mental state measured by PANSS total score**	The mean average endpoint on a general mental state measured by PANSS total score was 0	MD 1.7 lower (17.01 lower to 13.61 higher)	-	30 (1 RCT)	⊕⊖⊖⊖ Very low[a,b]	-
Average endpoint on negative symptoms measured by PANSS negative symptom score***	The mean average endpoint on negative symptoms measured by PANSS negative symptom score was 0	MD 0.7 lower (4.22 lower to 2.82 higher)	-	30 (1 RCT)	⊕⊖⊖⊖ Very low[a,b]	-
Quality of life: clinically important change in general quality of life - as defined by individual studies (long term).	No study reported the outcome				Not graded	-
Adverse effects/events: at least 1 adverse effect - as defined by individual studies.	No study reported the outcome				Not graded	-
Leaving the study early for any reason	Study population		not estimable	30 (1 RCT)	⊕⊖⊖⊖ Very low[a,b]	No event observed in either group
	0 per 1000	0 per 1000 (0 to 0)				

[a]Downgraded by one level for serious study limitations. The risk of bias in the included study was high due to suboptimal reporting and non-blinding.
[b]Downgraded by two levels for very serious imprecision. The sample size was small and the 95% CI crossed the threshold of both benefit and harm.

Cognitive functioning: clinically-important change in cognitive functioning - as defined by individual studies (long-term).	No study reported the outcome	Not graded	-
Social functioning: clinically-important change in social functioning - as defined by individual studies (long-term).	No study reported the outcome	Not graded	-

*__The risk in the intervention group__ (and its 95% confidence interval) is based on the assumed risk in the comparison group and the __relative effect__ of the intervention (and its 95% CI).

**This outcome is a proxy for one of the primary outcomes; clinically-important change in general mental state, as defined by individual studies

***This outcome is a proxy for one of the primary outcomes; clinically-important change in negative symptoms, as defined by individual studies

CI: Confidence interval; **MD:** Mean difference **PANSS**; Positive and Negative Syndrome Scale; **RCT:** Randomised controlled trial; **RR:** Risk ratio

GRADE Working Group grades of evidence

High certainty: We are very confident that the true effect lies close to that of the estimate of the effect

Moderate certainty: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different

Low certainty: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect

Very low certainty: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

[a]Downgraded by one level for serious study limitations. The risk of bias in the included study was high due to suboptimal reporting and non-blinding.

[b]Downgraded by two levels for very serious imprecision. The sample size was small and the 95% CI crossed the threshold of both benefit and harm.

Open in table viewer

Summary of findings 2. Humour-based interventions compared to standard care for people with schizophrenia

Humour-based interventions compared to standard care for people with schizophrenia

Patient or population: People with schizophrenia
Intervention: Humour-based interventions
Comparison: Standard care

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Certainty of the evidence (GRADE)	Comments
	Risk with standard care	Risk with humour-based interventions				
Clinically-important change in general mental state, as defined by individual studie	No study reported the outcome				Not graded	-
Clinically-important change in negative symptoms, as defined by individual studies	No study reported the outcome				Not graded	-
Quality of life: clinically-important change in general quality of life - as defined by individual studies (long-term).	No study reported the outcome				Not graded	-
Adverse effects/events: at least 1 adverse effect - as defined by individual studies	No study reported the outcome				Not graded	-
Leaving the study early for any reason	Study population		Risk ratio 0.38 (0.08 higher to 1.80 higher)	66 (1 RCT)	⊕⊖⊖⊖ Very low [a,b]	-
	156 per 1000	59 per 1000 (12 to 281)				
Cognitive functioning: clinically-important change in cognitive functioning - as defined by individual studies (long-term).	No study reported the outcome				Not graded	-

[a]Downgraded by one level for serious study limitations. The risk of bias in the included study was high due to suboptimal reporting and non-blinding.
[b]Downgraded by two levels for very serious imprecision. The sample size was small and the 95% CI crossed the threshold of both benefit and harm.

Social functioning: clinically-important change in social functioning - as defined by individual studies (long-term).	No study reported the outcome	Not graded	-

*__The risk in the intervention group__ (and its 95% confidence interval) is based on the assumed risk in the comparison group and the __relative effect__ of the intervention (and its 95% CI).

CI: Confidence interval; MD: Mean difference; **PANSS**; Positive and Negative Syndrome Scale; RCT: Randomised controlled trial; **RR:** Risk ratio

GRADE Working Group grades of evidence

High certainty: We are very confident that the true effect lies close to that of the estimate of the effect

Moderate certainty: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different

Low certainty: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect

Very low certainty: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

[a]Downgraded by one level for serious study limitations. The risk of bias in the included study was high due to suboptimal reporting and non-blinding.

[b]Downgraded by two levels for very serious imprecision. The sample size was small and the 95% CI crossed the threshold of both benefit and harm.

Background

Description of the condition

Schizophrenia is a disabling and costly serious mental illness with a prevalence of about 0.7 % around the world (Murray 1996; McGrath 2008). Characteristics associated with schizophrenia typically include: (i) positive symptoms such as disorganised speech, hallucinations or delusions; (ii) negative symptoms, such as a flat affect or poverty of speech; and (iii) impairments in cognition, including attention, memory, and executive functions. In addition, schizophrenia can have a severe negative impact on social and occupational functioning (APA 2013). Schizophrenia is regarded as a long-term illness, with age of onset typically during adolescence, with only 12% to 38% of patients fully recovering; around half have an undulating course with remissions, and 10% to 30% are treatment-resistant (Modestin 2003; Lehman 2004; Abel 2010; Lally 2017; Volavka 2018).

Antipsychotic medications are the first-line treatment for schizophrenia (Lieberman 1996). They have been shown to be effective in reducing positive symptoms and behaviours associated with schizophrenia, but most antipsychotics have not shown clinically-meaningful effects on negative symptoms (Leucht 2009; Fusar-Poli 2015).

Other interventions such as cognitive behavioural therapy (CBT), social-skills training (SST) and family therapy are often used in conjunction with pharmacotherapy, with an aim to help achieve better recovery (Schooler 1997; Kurtz 2008; Pharoah 2010; Jauhar 2014). However, implementation of these psychosocial interventions is often complex, requiring

considerable resources, and are often not readily available (Van der Gaag 2011), with robust evidence currently lacking to confirm their effectiveness for people with schizophrenia (Pharoah 2010; Almerie 2015; Jones 2018a; Jones 2018b). Humour-based interventions (or humour-focused programmes) are relatively simple and easy to deliver and thus could be an effective add-on treatment for people with schizophrenia, especially for targeting negative symptoms and cognitive impairment (Cai 2014; Bains 2015). In addition, humour-based interventions are fairly cheap in terms of design and conduct, making them an appealing and unique alternative to conventional non-pharmacological therapies such as CBT, SST and family psycho-education.

Description of the intervention

The Oxford English Dictionary defines humour as the "quality of being amusing or comic, especially as expressed in literature or speech, a mood or state of mind, each of the four" (Oxford Dictionaries 2013). Humour has been shown to affect mental health favourably, serving to reduce anxiety, stress, and tension (Dixon 1980). It is thought to improve judgement, understanding and perception, augmenting sociability and personal interaction (Dixon 1980). In general, people who have humour skills and who include humour and laughter more often in their daily lives were found to be better at coping with stressful events (McGhee 1983; Lefcourt 1986). This phenomenon is thought to be due to humour operating as a specific mechanism by which positive emotions reduce undesirable negative emotions involved in a stressful situation (Strotzka 1957). In addition, as a natural stress antagonist, humour is also associated with improvement in cardiovascular, immune, and endocrine systems (Berk 1989; Fredrickson 1998; Bennett 2003).

The use of humour as a therapy (therapeutic humour or humour-based interventions) is defined as "any intervention that promotes health and wellness by stimulating a playful discovery, expression, or appreciation of the absurdity or incongruity of life's situations". Humour-based interventions could be used as a "complementary treatment of illness to facilitate healing or coping, whether physical, emotional, cognitive, social, or spiritual" (Association for Applied and Therapeutic Humor 2000).

Humour-based interventions consist of the following essential elements: (i) humour stimulus (humour); (ii) an emotional response (mirth); and (iii) resulting behaviours such as grinning, laughing, smiling, or giggling (Fry 1992). The use of humour as an intervention for people with a mental illness was first introduced in the 1990s (Saper 1990). It is worth highlighting that humour-based interventions do not necessarily need to be delivered by trained therapists; for example, studies have shown that simple interventions such as showing a humorous movie could lead to a beneficial effect on inpatients with chronic schizophrenia (Gelkopf 2006). In addition, it has been suggested that a duration of more than 10 hours in total is needed for humour-based interventions to be effective (Cai 2014).

How the intervention might work

In the context of neurobiology, several mechanisms for humour are assumed to affect the symptoms of schizophrenia. First, deficits in dopamine transmission at D1, D3, D4 receptors in the prefrontal cortex might be implicated in cognitive impairments and negative symptoms of schizophrenia (Davis 1991; Seeman 1994; Honey 1999; Heinz 2000; Bertolino 2004). The current predominant view is that the dopamine system in schizophrenia might be characterised by an imbalance between subcortical and cortical dopamine systems, where subcortical mesolimbic dopamine projections might be hyperactive (resulting in hyperstimulation of D2 receptors and positive symptoms), while mesocortical dopamine projections to the prefrontal cortex might be hypoactive (resulting in hypostimulation of D1 receptors, negative symptoms, and cognitive impairment) (Davis 1991; Goldman-Rakic 2000). Other studies have endorsed the effects of subcortical dopamine on negative symptoms, in particular dopamine signalling in the striatum and its impact on the reward (Heinz 2002).

Humour, like most good, fun or pleasant stimuli, has been demonstrated to activate the mesolimbic reward circuit (Mobbs 2003). Dopamine neurons within this circuit also respond to food, and positive family and social interactions (Wilkins 1997; Carr 2002). They are what is called “natural” rewards. It is also intriguing to note that both the supplementary motor area proper and the dorsal anterior cingulate cortex receive rich dopamine input via ascending mesocortical projections from the ventral striatum, suggesting that these regions play an extended role in the dopaminergic reward network associated with humour appreciation (Bates 1993; Dum 1993). It could be possible that humour may activate dopamine neurons in the mesocortical and mesolimbic reward circuit, thereby affecting negative symptoms.

Another hypothesis relates to the activation of the hypothalamic-pituitary-adrenal axis and the subsequent release of glucocorticoids by various psychosocial stressors in schizophrenia. When exposure to stressors persists with continual heightened glucocorticoid release, there could be changes in the hypothalamic-pituitary-adrenal axis such that corticosterone release is augmented, high levels of which would have damaging effects on the hippocampus (McEwen 2001). The hippocampus is responsible for memory function, and disturbance could lead to cognitive impairment (McEwen 2001). Hippocampal volume reduction is one of the most consistent structural abnormalities found in schizophrenia (Weinberger 1992).

Laughter is regarded as eustress (a type of stress, which is healthy or good stress), reversing or attenuating the stress responses (Berk 1989), and stress increases cortisol secretion. It has been reported that laughter leads to reduced cortisol levels, but does not increase adrenocorticotropic hormone (ACTH) via positive feedback. Laughter could therefore potentially attenuate the hypothalamic-pituitary-adrenal-axis response (Berk 1989).

Why it is important to do this review

Treatment with antipsychotics for schizophrenia is not always fully effective, particularly for the negative long-term effects of schizophrenia (Fusar-Poli 2015), and additional non-pharmacological treatments are thus recommended (NICE 2014). Determining how effective these additional treatments are, both clinically and financially, is of vital importance in order to improve recovery rates and quality of life for people with schizophrenia (Petersen 2005; Srihari 2015).

Humour-based interventions are therapeutic methods that can be applied in various therapeutic settings such as hospital wards, nursing homes, and day care centres (Falkenberg 2011; Cai 2014). They are considered to be multimodal psychosocial interventions that can be used in conjunction with the more conventional pharmacotherapies and psychotherapies for people with mental illnesses (Jorm 2008). However, to our knowledge, there are no comprehensive systematic reviews assessing the clinical effectiveness of humour-based interventions for people with schizophrenia. We are confident that high-quality evidence synthesised with a robust methodology in the format of a Cochrane Review is much needed and relevant to all stakeholders, including clinicians, consumers and funders.

Objectives

To systematically review the effects of humour-based interventions as an add-on intervention to standard care for people with schizophrenia.

Methods

Criteria for considering studies for this review

Types of studies

We included all relevant randomised controlled trials (RCTs). We applied no restrictions by language, country of publication or observational periods. We excluded quasi-randomised trials.

Types of participants

We included all people with a diagnosis of schizophrenia as reflected by international standards such as the International Classification of Diseases 11th Revision (ICD-11) (WHO 2018), or the Diagnostic and Statistical Manual of Mental Disorders 5th Edition (DSM–5) (APA 2013). We also considered people diagnosed with schizophrenia by all previous versions of the standards, and also people diagnosed with schizophrenia by other types of diagnostic criteria. We did not exclude by age, sex or race. We included participants from both hospital and outpatient settings.

Types of interventions

1. Humour-based interventions

Interventions that were intended to be humorous. These may include watching humorous movies, funny videos, comedies, or any other humour-based interventions as defined by trial investigators.

2. Comparison interventions

2.1 Active controls

Participants in the comparator groups were given additional/add-on interventions similar to humour-based interventions (such as watching movies), but these interventions were not classified or defined as humorous.

2.2 Other psychosocial interventions

These included cognitive behavioural therapy (CBT), supportive therapy or social-skills training (SST).

2.3 Wait-list control group

Participants in the comparator groups were placed on a waiting list to receive the study-specific humour-based interventions.

2.4 Standard care

Participants continued to receive their normal care.

We considered all lengths and frequency of delivery time and, where possible, we planned to pursue subgroup analysis based on time periods or durations (see Subgroup analysis and investigation of heterogeneity).

Types of outcome measures

We aimed to divide all outcome measures into short-term (less than six weeks), medium-term (six to 12 weeks) and long-term (over 12 weeks). For the better interpretation of the effect size of patient-reported outcome measures, we would have attempted to report binary outcomes recording clear and clinically-meaningful degrees of change (e.g. global impression of much improved, or more than 50% improvement on a rating scale (as defined within the trials) before any others (Collister 2021). Thereafter, we listed other binary outcomes and then those that are continuous. Our approach in identifying types of valid rating scales is further elaborated in the section 'Data extraction and management'.

Primary outcomes

1. Mental state

1.1 General
1.1.1 Clinically-important change in general mental state, as defined by individual studies

1.2 Specific
1.2.1 Clinically-important change in negative symptoms, as defined by individual studies

2. Quality of life

2.1 Clinically-important change in overall quality of life, as defined by individual studies

3. Adverse effects/events

3.1 At least one adverse effect

Secondary outcomes

1. Mental state

1.1 General
1.1.1 Any change in general mental state, as defined by individual studies
1.1.2 Average endpoint or change score on a general mental state scale

1.2 Specific
1.2.1 Clinically-important change in specific symptoms (e.g. positive, affective, cognitive symptoms of schizophrenia), as defined by individual studies
1.2.2 Any change in specific symptoms (e.g. positive, negative, affective, cognitive symptoms of schizophrenia), as defined by individual studies
1.2.3 Average endpoint or change score on a specific symptom scale

2. Quality of life

2.1 Overall
2.1.1 Any change in overall quality of life, as defined by individual studies
2.1.2 Average endpoint or change score on a quality-of-life scale

2.2 Specific
2.2.1 Clinically-important change in specific aspects of quality of life, as defined by individual studies
2.2.2 Any change in specific aspects of quality of life, as defined by individual studies
2.2.3 Average endpoint or change score on a specific aspects of quality-of-life scale

3. Adverse effects

3.1 Clinically-important adverse effects, as defined by individual studies

3.2 Serious adverse effects, as defined by individual studies

3.3 Various adverse effects, as defined by individual studies

3.4 Average endpoint or change score on an adverse-effect scale

4. Leaving the study early

4.1 For any reason

4.2 Due to inefficacy

4.3 Due to adverse effect

5. Cognitive functioning

5.1 General

5.1.1 Clinically-important change in cognitive functioning, as defined by individual studies

5.1.2 Any change in cognitive functioning, as defined by individual studies

5.1.3 Average endpoint or change score on a cognitive functioning scale

5.2 Specific

5.2.1 Clinically-important change in specific aspects of cognitive functioning, as defined by individual studies

5.2.2 Any change in specific aspects of cognitive functioning, as defined by individual studies

5.2.3 Average endpoint or change score on a specific aspects of cognitive functioning scale

6. Social functioning

6.1 Clinically-important change in social functioning, as defined by individual studies

6.2 Any change in social functioning, as defined by individual studies

6.3 Average endpoint or change score on a social functioning scale

7. Economic costs

7.1 Direct costs

7.2 Indirect costs

Search methods for identification of studies

Electronic searches

Cochrane Schizophrenia Group's Study-Based Register of Trials

On 31 July 2019 and 10 February 2021, the Information Specialist of the Cochrane Schizophrenia Group searched the register using the following search strategy:

Humor in Intervention Field of STUDY.

In such study-based register, searching the major concept retrieves all the synonyms and relevant studies because all the studies have already been organised based on their interventions and linked to the relevant topics (Shokraneh 2017; Roberts 2021; Shokraneh 2021). This allows rapid and accurate searches that reduce waste in the next steps of systematic reviewing (Shokraneh 2019).

Following Cochrane methods (Lefebvre 2021), this register is compiled by systematic searches of major resources

(CENTRAL, CINAHL, ClinicalTrials.Gov, Embase, ISRCTN, MEDLINE, PsycINFO, PubMed, WHO ICTRP) and their monthly updates, ProQuest Dissertations and Theses A&I and its quarterly update, hand-searches, grey literature, and conference proceedings (Shokraneh 2020; see Group's website). There are no limitations by language, date, document type, or publication status for inclusion of records into the register.

Searching other resources

1. Reference searching

We inspected the references of all included studies for further relevant studies.

2. Personal contact

We contacted the corresponding author of studies for which it was not possible to retrieve a full publication to ask for unpublished data or other trial-level information. We have noted the outcome of this communication in the 'Characteristics of included studies' table.

Data collection and analysis

Selection of studies

Two review authors (KK, MB) independently inspected all citations from the literature searches and identified relevant titles/abstracts. Where disputes arose, we retrieved the full-text reports for further assessment. We obtained full reports of conference proceedings meeting the review criteria and two review authors (YT, YN) independently inspected these before incorporating relevant data into the review. Where it was not possible to resolve the disagreement by discussion, we consulted a third review author (YK). If doubt still remained, we added these trials to the list of 'Studies awaiting classification' and attempted to contact the study authors for clarification.

Data extraction and management

1. Extraction

Two review authors (YT, YN) extracted data from all included studies. We extracted data presented only in graphs and figures whenever possible, but included only data for which the two review authors independently collected the same results. For multicentre studies, where possible we would have extracted separate data relevant to each participating centre. We discussed any disagreement and documented our decisions; a third review author (MB) was consulted for any remaining disputes. Where necessary, we attempted to contact study authors to obtain missing information or for further clarification.

2. Management

2.1 Forms

We extracted data onto prestandardised data collection forms.

2.2 Scale-derived data

We included continuous data from rating scales only if:

1. the psychometric properties of the measuring instrument have been previously reported (Marshall 2000);

2. the measuring instrument has not been written or modified by one of the trialists for that particular trial; and

3. the instrument should be a global assessment of an area of functioning and not sub-scores which are not, in themselves, validated or shown to be reliable. However there are exceptions; we will include sub-scores from mental state scales measuring positive and negative symptoms of schizophrenia.

Ideally, the measuring instrument should either be a self-report or completed by an independent rater or relative (not the therapist). We realised that this was not often reported clearly and thus in Description of studies we provided relevant information.

2.3 Endpoint versus change data

There are advantages in processing/synthesising both endpoint and change data. Change data can remove the component of between-person variability from the analysis; but calculation of change requires two assessments (baseline and endpoint) that can be difficult to obtain in unstable and difficult-to-measure conditions such as schizophrenia. We thus decided to use endpoint data, and only used change data if the former were not available. Where necessary, we would have combined endpoint and change data in the analysis, as we preferred to use mean differences (MDs) rather than standardised mean differences (SMDs) as the effect measure (Deeks 2021).

2.4 Skewed data

Continuous data on clinical and social outcomes are often not normally distributed. To avoid the pitfalls of applying parametric tests to non-parametric data, we would have applied the following standards to relevant continuous data before inclusion.

For endpoint data from studies including fewer than 200 participants:

1. When a scale starts from the finite number 0 (zero), we would have subtracted the lowest possible value from the mean, and divided this by the standard deviation (SD). If this value was less than 1, it would strongly suggest that the data were skewed and we would have excluded these data from analysis. If this value was more than 1 but less than 2, there would be a suggestion that the data were skewed and we would have entered these data and tested whether their inclusion or exclusion would substantially change the overall results. If such data did change the results we would have treated them as 'Other data'. Finally, if the ratio was larger than 2 we would have included these data because it was unlikely that they were skewed (Altman 1996).

2. if a scale starts from a positive value (such as the Positive and Negative Syndrome Scale (PANSS), which can have values from 30 to 210 (Kay 1986)), we would have modified the calculations described above to take the scale starting point into account. In these cases the presence of skewed data is defined as 2 SDs > (S − S min), where S is the mean score and 'S min' is the minimum score.

Please note: we would have entered all relevant data from studies of more than 200 participants in the analysis irrespective of the above rules, because skewed data pose less of a problem in large studies. However, we did not have any studies enrolling more than 200 participants. Additionally, we would have entered all relevant change data, as when continuous data are presented on a scale that includes a possibility of negative values (such as change data), it is often difficult to distinguish whether or not the data are indeed skewed.

2.5 Common measurement

To facilitate comparison between trials, where relevant we would have converted variables that can be reported in different metrics, such as days in hospital (mean days per year, per week or per month), to a common metric (e.g. mean days per month).

2.6 Conversion of continuous to binary outcome data

Where possible, we would have made efforts to convert continuous outcome measures to dichotomous (binary) data. This is achieved by identifying cut-off points on rating scales and dividing participants accordingly into 'clinically improved' or 'not clinically improved'. It is generally assumed that if there is a 50% reduction in a scale-derived score such as the Brief Psychiatric Rating Scale (BPRS) (Overall 1962), or the PANSS (Kay 1986), this could be considered as a clinically-significant response (Leucht 2005a; Leucht 2005b). If data based on these thresholds were found to be unavailable, we would have used the primary cut-off presented by the original authors. In order to calculate responders from continuous outcomes, we would have used an imputation method proposed by Furukawa 2005 and Samara 2013.

2.7 Direction of graphs

Where possible, we would have entered data in such a way that the area to the left of the line of no effect indicates a favourable outcome for humour-based interventions. Should keeping to this makes it impossible to avoid outcome titles with clumsy double-negatives (e.g. 'not un-improved'), we would have reported data where the left of the line indicates an unfavourable outcome and noted this in the relevant graphs.

Assessment of risk of bias in included studies

Two review authors (YT, YN) independently assessed risks of bias in included studies using the Cochrane tool for assessing risk of bias as described in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2021b). This set of criteria is based on evidence of associations between overestimate of effect and high risk of bias of the article such as sequence generation, allocation concealment, blinding, incomplete outcome data and selective reporting.

For disagreements between the two review authors, we reached the final assessment ratings by consensus with the involvement of a third review author (MB). Where details of randomisation methods and other characteristics of trials were found to be inadequate, we attempted to contact authors of the studies in order to obtain further information. We reported non-concurrence in the risk of bias assessment, and resolved disputes about which category a trial should be allocated to by discussion.

We noted the level of risk of bias in the main text of the review and in both the risk of bias graph and risk of bias summary.

Measures of treatment effect

1. Binary data

For binary outcomes, we calculated a standard estimation of the risk ratio (RR) and its 95% confidence interval (CI), as it has been shown that RRs are more intuitive than odds ratios (ORs) (Boissel 1999), and that ORs tend to be interpreted as RRs by clinicians (Deeks 2002). Although the number needed to treat for an additional beneficial outcome (NNTB) and the number needed to treat for an additional harmful outcome (NNTH), with their respective CIs, are intuitively attractive to clinicians, they are problematic to calculate and interpret in meta-analyses (Hutton 2009). For binary data presented in the GRADE summary of findings table/s we would, where possible, have calculated illustrative comparative risks.

2. Continuous data

For continuous outcomes we estimated the mean difference (MD) between groups. We decided not to calculate effect size measures using standardised mean differences (SMDs). However, if very similar scales were used, we would have presumed there was a small difference in measurement, and we would have calculated the effect size and transformed it back to the units of one or more of the specific instruments.

Unit of analysis issues

1. Cluster trials

Studies increasingly employ 'cluster randomisation' (such as randomisation by clinician or practice), but analysis and pooling of clustered data pose problems. Authors often fail to account for intra-class correlation in clustered studies, leading to a unit-of-analysis error whereby P values are spuriously low, CIs unduly narrow and statistical significance overestimated (Divine 1992). This causes type I errors (Bland 1997; Gulliford 1999).

Where clustering had been incorporated into the analysis of primary studies, we would have presented these data as if from a non-cluster randomised study, but adjusted for the clustering effect.

Where clustering was not accounted for in primary studies, we would have presented data in a table, with a (*) symbol to indicate the presence of a probable unit of analysis error. We would have sought to contact study investigators to obtain intra-class correlation coefficients (ICCs) for their clustered data and to adjust for this by using accepted methods (Gulliford 1999).

We have been advised by a statistician that the binary data from cluster trials presented in a report should be divided by a 'design effect'. This is calculated using the mean number of participants per cluster (m) and the ICC: thus design effect = 1 + (m − 1) * ICC (Donner 2002). If the ICC was not reported, we would have assumed it to be 0.1 (Ukoumunne 1999).

If cluster studies were identified and had been appropriately analysed, with ICCs considered and relevant data documented in the reports, synthesis with other studies would have been possible using the generic inverse variance technique.

2. Cross-over trials

A major concern of cross-over trials is the carry-over effect. This occurs if an effect (e.g. pharmacological, physiological or psychological) of the treatment in the first phase is carried over to the second phase. As a consequence, participants can differ significantly from their initial state at entry to the second phase, despite a wash-out phase. For the same reason, cross-over trials are not appropriate if the condition of interest is unstable (Elbourne 2002). As both carry-over and unstable conditions are very likely in severe mental illnesses such as schizophrenia, we would only have used data from the first phase of cross-over studies.

3. Studies with multiple treatment groups

Where a study involves more than two treatment arms, if relevant, we would have presented the additional treatment arms in comparisons. For binary data, we would have simply added and combined them in a two-by-two table. If data were continuous, we would have combined data following the formula in the *Cochrane Handbook for Systematic Reviews of Interventions* (Li 2021). Where additional treatment arms were not relevant, we would not have reproduced these data.

Dealing with missing data

1. Overall loss of credibility

At some degree of loss of follow-up, data must lose credibility (Xia 2009). We decided that, for any particular outcome, should more than 50% of data be unaccounted for we would not reproduce these data or use them within analyses. If, however, more than 50% of those in one arm of a study are lost, but the total loss is less than 50%, we would have addressed this within the summary of findings table/s by down-rating the certainty. Finally, we would also have downgraded the certainty of evidence within the summary of findings table/s should the loss be 25% to 50% in total.

2. Binary

In the case where attrition for a binary outcome was between 0% and 50% and where these data were not clearly described, we would have presented data on a 'once-randomised-always-analyse' basis (an intention-to-treat (ITT) analysis). Those leaving the study early would have been assumed to have the same rates of negative outcome as those who completed. We would have used the rate of those who stayed in the study - in that particular arm of the trial - and applied this also to those who did not. We would have performed a sensitivity analysis to test how prone the primary outcomes were to change when data only from people who completed the study to that point were compared to the ITT analysis, using the above assumptions.

3. Continuous

3.1 Attrition

We would have used data where attrition for a continuous outcome was between 0% and 50%, and data only from people who completed the study to that point were reported.

3.2 Standard deviations

If standard deviations (SDs) were not reported, we would have attempted to obtain the missing values from the study authors. If these were not available, where there were missing measures of variance for continuous data, but an exact standard error (SE) and CIs available for group means, and either a P value or a t value available for differences in mean, we could have calculated the SDs according to the guidelines in the *Cochrane Handbook for Systematic Reviews of Interventions* (Li 2021). When only the SE was reported, we would have calculated SDs by the formula SD = SE * √(n). The *Cochrane Handbook for Systematic Reviews of Interventions* presents detailed formulae for estimating SDs from P, t or F values, CIs, ranges or other statistics (Li 2021). If these formulae did not apply, we would have calculated the SDs according to a validated imputation method which is based on the SDs of the other included studies (Furukawa 2006). Although some of these imputation strategies can introduce error, the alternative would be to exclude a given study's outcome and thus to lose information. Nevertheless, we would have examined the validity of the imputations in a sensitivity analysis that excluded imputed values.

3.3 Assumptions about participants who left the trials early or were lost to follow-up

Various methods are available to account for participants who left the trials early or were lost to follow-up. Some trials simply present the results of study completers; others use the method of last observation carried forward (LOCF). Recently, methods such as multiple imputation or mixed-effects models for repeated measurements (MMRM) have become more of a standard. While the latter methods seem to be somewhat better than LOCF (Leon 2006), we felt that the high percentage of participants leaving the studies early and differences between groups in their reasons for doing so is often the core problem in randomised schizophrenia trials. We would therefore not have excluded studies based on the statistical approach used. However, by preference we would have used the more sophisticated approaches, i.e. we

would have preferred to use MMRM or multiple-imputation to LOCF, and we would have only presented completer analyses if some kind of ITT data were not available at all. Moreover, we would have addressed this issue in the item 'Incomplete outcome data' of the Cochrane tool for assessing risk of bias (Higgins 2021b).

Assessment of heterogeneity

1. Clinical heterogeneity

We considered all included studies initially, without seeing comparison data, to judge clinical heterogeneity. We inspected all studies for participants who were clearly outliers or situations that we had not predicted would arise and, where found, discussed such situations or participant groups.

2. Methodological heterogeneity

We considered all included studies initially, without seeing comparison data, to judge methodological heterogeneity. We inspected all studies for clearly outlying methods which we had not predicted would arise and discussed any such methodological outliers.

3. Statistical heterogeneity

3.1 Visual inspection

We inspected graphs visually to investigate the possibility of statistical heterogeneity.

3.2 Using the I^2 statistic

We investigated heterogeneity between studies by considering the I^2 statistic alongside the Chi^2 P value. The I^2 statistic provides an estimate of the percentage of inconsistency thought to be due to chance (Higgins 2003). The importance of the observed value of I^2 depends on the magnitude and direction of effects as well as the strength of evidence for heterogeneity (e.g. P value from Chi^2 test, or a confidence interval for I^2). We would have interpreted an I^2 estimate greater than or equal to 50% and accompanied by a statistically significant Chi^2 statistic as evidence of substantial heterogeneity (Deeks 2021). If we found substantial levels of heterogeneity in the primary outcomes, we would have explored reasons for heterogeneity (Subgroup analysis and investigation of heterogeneity).

Assessment of reporting biases

Reporting biases arise when the dissemination of research findings is influenced by the nature and direction of results (Egger 1997). These are described in the *Cochrane Handbook for Systematic Reviews of Interventions* (Sterne 2011).

1. Protocol versus full study report

We tried to locate protocols of included randomised trials. If the protocol had been available, we would have compared outcomes in the protocol and in the published study report. Since the protocols were not available, we compared outcomes listed in the 'Methods' section of the study report with the results actually reported .

2. Funnel plot

We are aware that funnel plots may be useful in investigating reporting biases but are of limited power to detect small-study effects. We therefore did not plan to use funnel plots for outcomes where there were 10 or fewer studies, or where all studies were of similar size. In other cases, where funnel plots were deemed possible, we would have sought

statistical advice in their interpretation.

Data synthesis

We understand that there is no closed argument for preference for use of fixed-effect or random-effects models. The random-effects method incorporates an assumption that the different studies are estimating different yet related intervention effects. This often seems to be true to us and the random-effects model takes into account differences between studies, even if there is no statistically significant heterogeneity. There is, however, a disadvantage to the random-effects model: it puts added weight onto small studies, which are often the most biased ones. Depending on the direction of effect, these studies can either inflate or deflate the effect size. We decided to use a random-effect model.

Subgroup analysis and investigation of heterogeneity

1. Subgroup analyses

As there was only one included study in the quantitative analysis for each comparison, we did not conduct any subgroup analysis.

2. Investigation of heterogeneity

As we included only one study in the quantitative analysis for each comparison, we did not further investigate level or sources of heterogeneity.

Sensitivity analysis

As we included only one study in the quantitative analysis for each comparison, we were not able to conduct any sensitivity analysis.

Summary of findings and assessment of the certainty of the evidence

We used the GRADE approach to interpret findings (Schünemann 2011). We created summary of findings tables using the GRADEpto GDT online tool. These tables provide outcome-specific information about the overall certainty of evidence from each included study in the comparison, the magnitude of effect of the interventions examined, and the sum of available data on all outcomes we rate as important to patient care and decision-making.

We selected the following main outcomes for inclusion in the summary of findings table.

1. Mental state: clinically-important change in general mental state, as defined by individual studies (medium term).
2. Mental state: clinically-important change in negative symptoms, as defined by individual studies (medium term).
3. Quality of life: clinically-important change in general quality of life, as defined by individual studies (long term).
4. Adverse effects/events: at least one adverse effect, as defined by individual studies.
5. Leaving the study early: for any reason (by end of trial).
6. Cognitive functioning: clinically-important change in cognitive functioning, as defined by individual studies (long term).
7. Social functioning: clinically-important change in social functioning, as defined by individual studies (long term).

If data were not available for these prespecified outcomes but were available for ones that were similar, we would have presented the closest outcome to the prespecified one in the table and taken this into account when grading the findings.

Due to the limited data, we used average endpoint on a general mental state measured by PANSS total score as a proxy for clinically-important change in general mental state, and average endpoint on negative symptoms measured by PANSS negative symptom score as a proxy for clinically-important change in negative symptoms in Summary of Findings table 1.

The methods of this review are based on the published protocol (Kohmura 2019).

Results

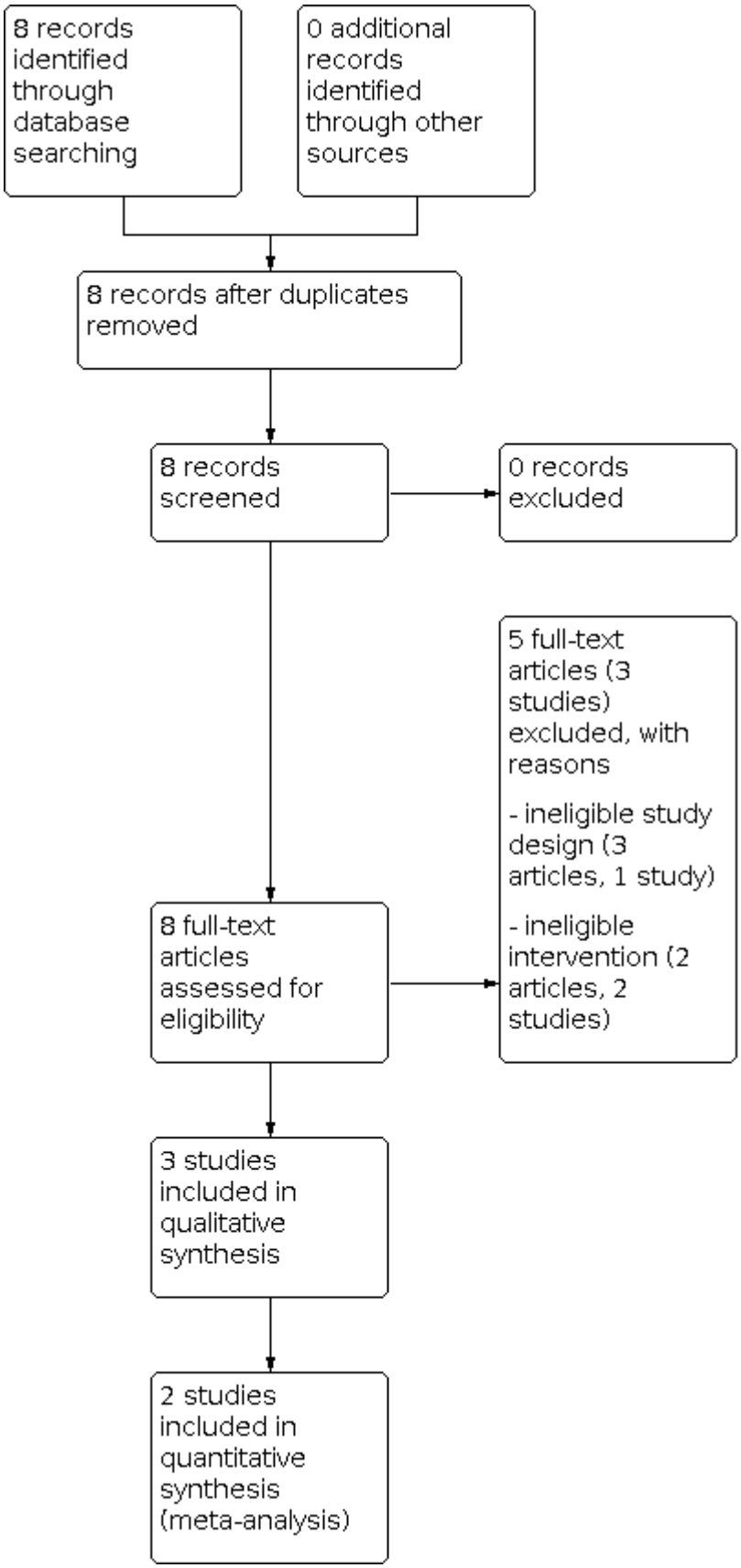
Description of studies

For details see Characteristics of included studies, and Characteristics of excluded studies.

Results of the search

We searched the Cochrane Schizophrenia Group register of trials and identified six eligible studies (in eight reports) for full-text screening (Figure 1). Of these, we included three studies in our review (Cai 2014; Falkenberg 2018; Cheng 2020a). Falkenberg 2018 was only available as a conference abstract, and did not report any of our prespecified outcome measures of interest. We contacted the authors of Falkenberg 2018 and were informed that they were unable to provide the data. We therefore could not include the study in subsequent quantitative synthesis (meta-analysis).

Figure 1 **Open in figure viewer**



Study flow diagram.

Included studies

Design and duration

All studies included in the review were randomised controlled trials. Cai 2014 followed participants to five weeks; Cheng 2020a followed participants to 12 weeks; Falkenberg 2018 did not report the follow-up period.

Participants

Cai 2014 and Cheng 2020a included participants with a diagnosis of schizophrenia according to the DSM-IV criteria. Falkenberg 2018 did not provide information on study eligibility criteria.

Size

Since Falkenberg 2018 did not report the actual study sample size nor any outcomes of interest, our review findings are based on the 96 participants assessed in Cai 2014 and Cheng 2020a.

Setting

Cai 2014 and Cheng 2020a were each conducted in one hospital in China and Taiwan, respectively; Falkenberg 2018 reported the study location to be in Germany but further details were not available.

Interventions

1 Humour-based interventions versus active controls

In Cai 2014, in addition to the usual antipsychotic medication, a humour skill training activity was used, which involved watching a cross-talk video performed by famous Chinese actors after the fun activity at the beginning of each session. The active control group received a recreational therapy programme by doing handwork in addition to their usual antipsychotic medication.

2 Humour-based interventions versus other psychosocial interventions

Falkenberg 2018 compared a humour ability training activity with a training programme of social skills (no data).

3 Humour-based interventions versus usual care

Cheng 2020a compared standard care and a laughing qigong programme which was a combination of qigong techniques and simulated laughter with a focus on the mind and body connection.

Outcomes

1 General

We identified data for the following outcome measures from Cai 2014: average endpoints on general mental state, positive and negative symptoms, depressive symptoms, and anxiety; and from Cheng 2020a: average endpoints on depressive symptoms. However, the study authors did not indicate whether the outcomes were measured by an independent rater. The number of participants leaving the study early for any reason was also reported in both studies.

Falkenberg 2018 reported negative symptoms, stress and psychosocial dysfunction, but no numerical/usable data were available.

2 Outcome scales providing useable data

The scales providing usable data from Cai 2014 and Cheng 2020a were as follows:

General mental state

Positive and Negative Syndrome Scale (PANSS) total score.

PANSS has 30 items, each of which can be defined on a seven-point scoring system varying from one (absent) to seven (extreme) (Kay 1987). This scale can be divided into three subscales for measuring the severity of general psychopathology (16 items), positive symptoms (seven items), and negative symptoms (seven items). PANNS total score is the sum of all items, and ranges from 30 to 210, with high scores indicating greater severity.

Positive symptoms

Positive and Negative Syndrome Scale positive symptom score (PANSS-P).

PANSS-P is the sum of seven items related to positive symptoms in PANSS, and ranges from 7 to 49, with high scores indicating greater severity.

Negative symptoms

Positive and Negative Syndrome Scale negative symptom score (PANSS-N).

PANSS-N is the sum of seven items related to negative symptoms in PANSS, and ranges from 7 to 49, with high scores indicating greater severity.

Depressive symptom

Beck Depression Inventory (BDI)

BDI has 21 items, each of which has a set of four possible responses to rate the intensity from (0) 'I do not feel sad' (absent) to (3) 'I am so sad or unhappy that I can't stand it' (extreme) (Beck 1961). A high score indicates greater severity. BDI 2nd edition (BDI-II) was introduced in 1996 to address DSM-IV criteria (Beck 1996).

Anxiety

State-Trait Anxiety Inventory (STAI)

STAI has 20 items for assessing trait anxiety and 20 for state anxiety (Spielberger 1970). Each item is rated on a four-point scale from 'almost never' (least) to 'almost always' (most). A high score indicates greater anxiety.

3 Missing outcomes

No study reported any of our prespecified primary outcomes, namely clinically-important change in general mental state, clinically-important change in negative symptoms, clinically-important change in overall quality of life, or adverse effects.

Excluded studies

We excluded three studies based on full-text assessment (Characteristics of excluded studies). The reasons for exclusions were a lack of humour elements/components in the study interventions (Gelkopf 1993; Li 2015). We excluded one ongoing study due to inappropriate study design (quasi-experimental study) (Atadokht 2019).

Ongoing studies

We did not find any ongoing study that matched our eligible criteria.

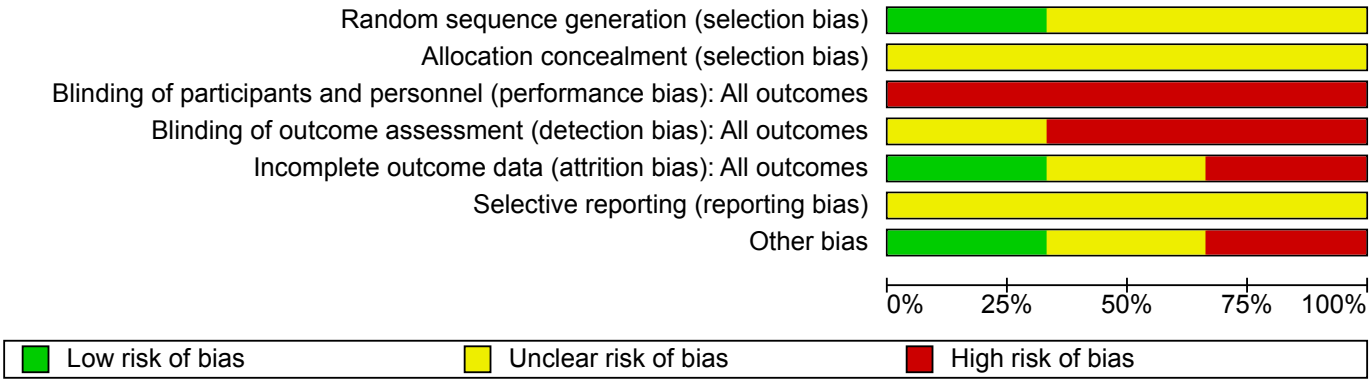
Studies awaiting assessment

There are currently no studies awaiting further assessment.

Risk of bias in included studies

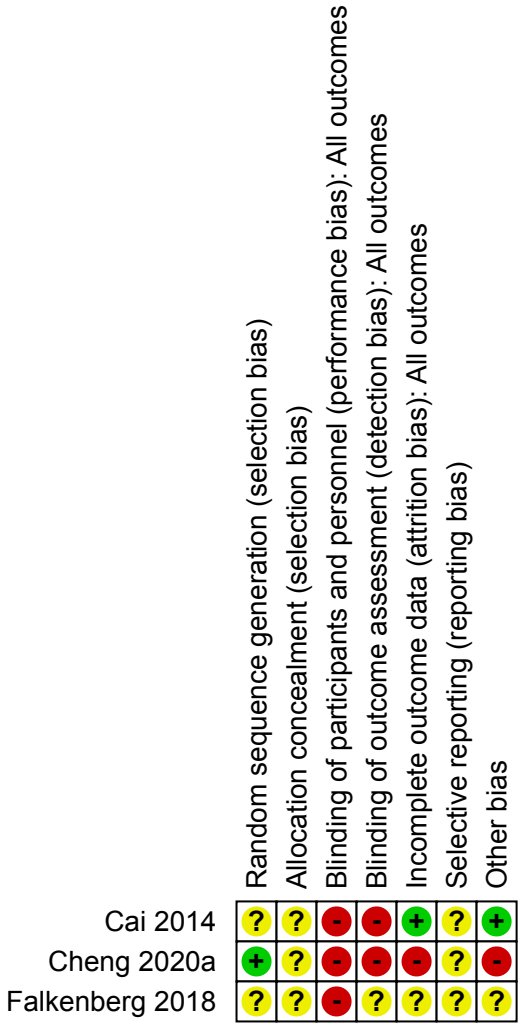
Our assessment of risks of bias in the included studies is illustrated in Figure 2 and Figure 3.

Figure 2 **Open in figure viewer**



Risk of bias graph: review authors' judgements about each risk of bias item presented as percentages across all included studies.

Figure 3 **Open in figure viewer**



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

Allocation

No study described the methods of allocation concealment and thus were judged to be at unclear risk of selection bias.

Blinding

Given the nature of the interventions, it is impossible to blind study participants or personnel. We thus considered all studies to be at high risk of performance bias. The three included studies did not report details of blinding of outcome assessors and thus were rated at unclear risk of detection bias.

Incomplete outcome data

Cai 2014 did not have missing data and thus was judged to be at low risk of attrition bias. In Cheng 2020a, more than 10% of participants had missing outcome data and the study was judged to be at high risk of attrition bias. Falkenberg 2018 did not report any outcomes of interest and we rated it to be at unclear risk of attrition bias.

Selective reporting

We were unable to retrieve study protocols for the three included studies, so we rated all of them at unclear risk of reporting bias.

Other potential sources of bias

There were no substantial baseline differences or imbalance between groups in Cai 2014 and we assessed it to be at low risk of other bias; Cheng 2020a provided the qigong programme as a co-intervention and was thus judged to be at high risk of other bias. Falkenberg 2018 did not report any study participant characteristics and thus was judged to be at unclear risk of other bias.

Effects of interventions

See: **Summary of findings 1** Humour-based interventions compared to active controls for people with schizophrenia; **Summary of findings 2** Humour-based interventions compared to standard care for people with schizophrenia

No included studies reported the following a priori primary outcome measures (Types of outcome measures): clinically-important change in general mental state, clinically-important change in negative symptoms, clinically-important change in overall quality of life, or adverse effects.

Comparison 1: Humour-based interventions versus active controls

We present available data from Cai 2014 as follows.

Primary outcomes

1. Mental state

1.1 General
1.1.1 Clinically-important change in general mental state, as defined by individual studies

No study reported the outcome.

1.2 Specific
1.2.1 Clinically-important change in negative symptoms, as defined by individual studies

No study reported the outcome.

2. Quality of life

2.1 Clinically-important change in overall quality of life, as defined by individual studies

No study reported the outcome

3. Adverse effects/events

3.1 At least one adverse effect

No study reported the outcome

Secondary outcomes

1. Mental state

1.1 General

1.1.1 Any change in general mental state, as defined by individual studies

No study reported the outcome.

1.1.2 Average endpoint or change score on a general mental state scale (PANSS total score, high = poor)

We are uncertain about the effect of humour-based interventions on lowering the average endpoint score for general mental state (PANSS total score: MD −1.70, 95% CI −17.01 to 13.61; 30 participants; very low certainty of evidence) (Analysis 1.1).

1.2 Specific

1.2.1 Clinically-important change in specific symptoms (e.g. positive, affective, cognitive symptoms of schizophrenia), as defined by individual studies

No study reported the outcome

1.2.2 Any change in specific symptoms (e.g. positive, negative, affective, cognitive symptoms of schizophrenia), as defined by individual studies

No study reported the outcome

1.2.3 Average endpoint or change score on a specific symptom scale

Positive symptoms: average endpoint score (PANSS-P, high = poor)

Humour-based interventions may result in little to no difference in positive symptoms (PANSS positive symptom score: MD 0.00, 95% CI −2.58 to 2.58; 30 participants; low certainty of evidence) (Analysis 1.2).

Negative symptoms: average endpoint score (PANSS-N, high = poor)

We are uncertain about the effect of humour-based interventions on negative symptoms (PANSS negative symptom score: MD −0.70, 95% CI −4.22 to 2.82; 30 participants; very low certainty of evidence) (Analysis 1.3).

Depressive symptoms: average endpoint score (BDI, high = poor)

Humour-based interventions may reduce depressive symptoms (BDI: MD −6.20, 95% CI −12.08 to −0.32; 30 participants; low certainty of evidence) (Analysis 1.4).

Anxiety: average endpoint score (STAI, high = poor)

Humour-based interventions may result in little to no difference in anxiety (STAI: MD −2.60, 95% CI −5.76 to 0.56; 30 participants; low certainty of evidence) (Analysis 1.5).

2. Quality of life

2.1 Overall

2.1.1 Any change in overall quality of life, as defined by individual studies

No study reported the outcome.

2.1.2 Average endpoint or change score on a quality-of-life scale

No study reported the outcome.

2.2 Specific

2.2.1 Clinically-important change in specific aspects of quality of life, as defined by individual studies

No study reported the outcome.

2.2.2 Any change in specific aspects of quality of life, as defined by individual studies

No study reported the outcome.

2.2.3 Average endpoint or change score on a specific aspects of quality-of-life scale

No study reported the outcome.

3. Adverse effects

3.1 Clinically-important adverse effects, as defined by individual studies

No study reported the outcome.

3.2 Serious adverse effects, as defined by individual studies

No study reported the outcome.

3.3 Various adverse effects, as defined by individual studies

No study reported the outcome.

3.4 Average endpoint or change score on an adverse-effect scale

No study reported the outcome.

4. Leaving the study early

4.1 For any reason

None of the participants in Cai 2014 were reported to have left the study early for any reason. However, due to the small sample size we remain uncertain about the precise effects of humour-based interventions on impacting participants to leave the study early.

4.2 Due to inefficacy

No study reported the outcome.

4.3 Due to adverse effect

No study reported the outcome.

5. Cognitive functioning

5.1 General

5.1.1 Clinically-important change in cognitive functioning, as defined by individual studies

No study reported the outcome.

5.1.2 Any change in cognitive functioning, as defined by individual studies

No study reported the outcome.

5.1.3 Average endpoint or change score on a cognitive functioning scale

No study reported the outcome.

5.2 Specific

5.2.1 Clinically-important change in specific aspects of cognitive functioning, as defined by individual studies

No study reported the outcome.

5.2.2 Any change in specific aspects of cognitive functioning, as defined by individual studies

No study reported the outcome.

5.2.3 Average endpoint or change score on a specific aspects of cognitive functioning scale

No study reported the outcome.

6. Social functioning

6.1 Clinically-important change in social functioning, as defined by individual studies

No study reported the outcome.

6.2 Any change in social functioning, as defined by individual studies

No study reported the outcome.

6.3 Average endpoint or change score on a social functioning scale

No study reported the outcome.

7. Economic costs

7.1 Direct costs

No study reported the outcome.

7.2 Indirect costs

No study reported the outcome.

Comparison 2: Humour-based interventions versus standard care

We present available data from Cheng 2020a as follows.

Primary outcomes

1. Mental state

1.1 General

1.1.1 Clinically-important change in general mental state, as defined by individual studies

No study reported the outcome.

1.2 Specific

1.2.1 Clinically-important change in negative symptoms, as defined by individual studies

No study reported the outcome.

2. Quality of life

2.1 Clinically-important change in overall quality of life, as defined by individual studies

No study reported the outcome

3. Adverse effects/events

3.1 At least one adverse effect

No study reported the outcome

Secondary outcomes

1. Mental state

1.1 General
1.1.1 Any change in general mental state, as defined by individual studies

No study reported the outcome.

1.1.2 Average endpoint or change score on a general mental state scale
No study reported the outcome.

1.2 Specific
1.2.1 Clinically-important change in specific symptoms (e.g. positive, affective, cognitive symptoms of schizophrenia), as defined by individual studies

No study reported the outcome
1.2.2 Any change in specific symptoms (e.g. positive, negative, affective, cognitive symptoms of schizophrenia), as defined by individual studies

No study reported the outcome
1.2.3 Average endpoint or change score on a specific symptom scale

Depressive symptoms: average endpoint score (BDI, high = poor)

Humour-based interventions may result in little to no difference in depressive symptoms (BDI-II: MD 0.80, 95% CI −2.64 to 4.24; 59 participants; low certainty of evidence) (Analysis 2.1).

2. Quality of life

2.1 Overall
2.1.1 Any change in overall quality of life, as defined by individual studies

No study reported the outcome.
2.1.2 Average endpoint or change score on a quality-of-life scale

No study reported the outcome.

2.2 Specific
2.2.1 Clinically-important change in specific aspects of quality of life, as defined by individual studies

No study reported the outcome.
2.2.2 Any change in specific aspects of quality of life, as defined by individual studies

No study reported the outcome.
2.2.3 Average endpoint or change score on a specific aspects of quality-of-life scale

No study reported the outcome.

3. Adverse effects

3.1 Clinically-important adverse effects, as defined by individual studies

No study reported the outcome.
3.2 Serious adverse effects, as defined by individual studies

No study reported the outcome.
3.3 Various adverse effects, as defined by individual studies

No study reported the outcome.

3.4 Average endpoint or change score on an adverse-effect scale

No study reported the outcome.

4. Leaving the study early

4.1 For any reason

We are uncertain about the effect of humour-based interventions on leaving the study early for any reason (risk ratio (RR) 0.38, 95% CI 0.08 to 1.80; 66 participants; very low certainty of evidence) (Analysis 2.2)

4.2 Due to inefficacy

No study reported the outcome.

4.3 Due to adverse effect

No study reported the outcome.

5. Cognitive functioning

5.1 General

5.1.1 Clinically-important change in cognitive functioning, as defined by individual studies

No study reported the outcome.

5.1.2 Any change in cognitive functioning, as defined by individual studies

No study reported the outcome.

5.1.3 Average endpoint or change score on a cognitive functioning scale

No study reported the outcome.

5.2 Specific

5.2.1 Clinically-important change in specific aspects of cognitive functioning, as defined by individual studies

No study reported the outcome.

5.2.2 Any change in specific aspects of cognitive functioning, as defined by individual studies

No study reported the outcome.

5.2.3 Average endpoint or change score on a specific aspects of cognitive functioning scale

No study reported the outcome.

6. Social functioning

6.1 Clinically-important change in social functioning, as defined by individual studies

No study reported the outcome.

6.2 Any change in social functioning, as defined by individual studies

No study reported the outcome.

6.3 Average endpoint or change score on a social functioning scale

No study reported the outcome.

7. Economic costs

7.1 Direct costs

No study reported the outcome.

7.2 Indirect costs

No study reported the outcome.

Discussion

Summary of main results

Our review suggests that currently available evidence for the effects of humour-based interventions on people with schizophrenia remains limited and of low certainty. Only two small, low-quality randomised trials reported quantitative results for the outcomes of interest. Our findings indicated that humour-based interventions may not have a role in positive symptoms and anxiety, but they might improve depressive symptoms compared with active controls. However, when we compared humour interventions with standard care, they might result in little difference in depressive symptoms. Overall, the certainty of the evidence was low to very low, highlighting the need for more high-quality evidence.

Overall completeness and applicability of evidence

No included studies reported our prespecified primary outcomes, namely clinically-important change in general mental state, clinically-important change in negative symptoms, clinically-important change in overall quality of life, and adverse effects. Since Falkenberg 2018 was only available as a conference abstract with very limited data, we were unable to incorporate its findings in the comparison of humour ability training and social skill training in our quantitative synthesis. We will try to include relevant study findings from Falkenberg 2018 in future review updates, as and when the full study report is released. We were also unable to look at the adverse events, or conduct prespecified subgroup and sensitivity analyses due to insufficient data. Our review is primarily based on findings from two Chinese and Taiwanese studies that recruited hospital inpatients and used short-term interventions (duration of five to eight weeks), and thus the applicability of these findings is limited. The longer-term effects of humour-based interventions in other settings such as non-Asian countries, outpatient clinics, nursing homes or hospices remain unclear.

Quality of the evidence

Overall, the quality of the review evidence was low to very low. The reasons for downgrading the certainty of the evidence using the GRADE approach were serious study-level risk of bias and imprecision. Due to the suboptimal reporting quality of the two included studies, we judged both studies to be at unclear risk of selection, detection and reporting biases. Due to the small study sample size (30 participants), following the GRADE approach this would lead to serious concerns on evidence imprecision (Guyatt 2011). Furthermore, we used the following cut-off scores for clinically-important changes as a way to determine the degree of imprecision of several continuous outcomes, namely general mental state, positive and negative symptoms, depressive symptoms, and anxiety: (i) PANSS total score of 15; (ii) PANSS positive symptoms score of 3.5; (iii) PANSS negative symptoms score of 3.5; (iv) BDI of 5; and (v) STAI of 10. We chose these cut-offs from previous studies (Hermes 2012; Masson 2013; Corsaletti 2014; Leddy-Stacy 2016). For the outcomes

of PANSS total score, PANSS negative symptoms score, and leaving the study early, we downgraded the certainty of evidence by two levels due to very serious imprecision, given that the 95% CIs crossed the thresholds of both benefit and harm (Guyatt 2011).

Potential biases in the review process

Throughout the review process, we followed the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2021a). However, we could not perform Egger's test to assess publication bias since we included fewer than 10 studies in our review. It is worth emphasising that our searches of trial registries did not identify any ongoing trials, although Falkenberg 2018 is, in effect, an ongoing study, as it is currently only available as a conference proceeding with potentially unpublished data.

Agreements and disagreements with other studies or reviews

A recent systematic review investigating the effects of humour-based interventions on depression, anxiety and sleep quality in adults reported conflicting results when compared to ours: quantitative analysis of 10 studies indicated that humour-based interventions did not improve depressive symptom in people with schizophrenia (Zhao 2019). This might be due to the inclusion of Gelkopf 1993, which we excluded from our review due to the exposure of humour-based interventions in both experimental and comparator groups.

On the other hand, our review findings are partially consistent with several small-scale trials that examined the effects of humour-based interventions in people with depressive symptoms (Walter 2007; Hirsch 2010; Konradt 2013; Tagalidou 2019). Depressive symptoms are frequently observed in people with schizophrenia (Conley 2007; Upthegrove 2017; Donde 2018). Indeed, the baseline mean BDI score of the study participants in Cai 2014, one of our two included studies, was 37.8 (SD 8.6), suggesting the participants were suffering from severe depression at the start of the study. Compared with other non-pharmacological interventions such as dance, music, or yoga therapies that involve entertainment elements and have previously shown favourable effects on the general or specific mental state of people with schizophrenia (Ren 2013; Broderick 2015; Broderick 2017a; Broderick 2017b; Broderick 2019; Geretsegger 2017), the effects of humour-based interventions on such outcomes are yet to be ascertained.