Cochrane Library

Cochrane Database of Systematic Reviews

Avatar Therapy for people with schizophrenia or related disorders (Review)

Aali G, Kariotis T, Shokraneh F

Aali G, Kariotis T, Shokraneh F.
Avatar Therapy for people with schizophrenia or related disorders.
Cochrane Database of Systematic Reviews 2020, Issue 5. Art. No.: CD011898.
DOI: 10.1002/14651858.CD011898.pub2.

www.cochranelibrary.com

Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

WILEY

TABLE OF CONTENTS

HEADER 1
ABSTRACT 1
PLAIN LANGUAGE SUMMARY 2
SUMMARY OF FINDINGS 4
BACKGROUND 8
OBJECTIVES 8
METHODS 8
Figure 1. 10
Figure 2. 13
Figure 3. 14
RESULTS 16
DISCUSSION 22
AUTHORS' CONCLUSIONS 23
ACKNOWLEDGEMENTS 24
REFERENCES 25
CHARACTERISTICS OF STUDIES 27
DATA AND ANALYSES 34
Analysis 1.1. Comparison 1: Avatar therapy versus treatment as usual, Outcome 1: Mental state: 1a. Specific – positive – average endpoint score (PANSS-P, high = poor) – short term 36
Analysis 1.2. Comparison 1: Avatar therapy versus treatment as usual, Outcome 2: Mental state: 1b. Specific – insight – average endpoint score (BAVQ-R attitude to voices, high = poor) – short term 36
Analysis 1.3. Comparison 1: Avatar therapy versus treatment as usual, Outcome 3: Mental state: 1c. Specific – insight – average endpoint score (BAVQ-R attitude to voices, high = poor, skewed data) – short term 36
Analysis 1.4. Comparison 1: Avatar therapy versus treatment as usual, Outcome 4: Mental state: 1d. Specific – depression – average endpoint score (various scales, high = poor, skewed data – short term 36
Analysis 1.5. Comparison 1: Avatar therapy versus treatment as usual, Outcome 5: Mental state: 1e. Specific – negative – average endpoint score (PANSS-N, high = poor) – short term 37
Analysis 1.6. Comparison 1: Avatar therapy versus treatment as usual, Outcome 6: Mental state: 2a. General – total – average endpoint score (PANSS total, high = poor) – short term) 37
Analysis 1.7. Comparison 1: Avatar therapy versus treatment as usual, Outcome 7: Mental state: 2b. General – total – average endpoint score (PSYRATS total, high = poor) – short term 37
Analysis 1.8. Comparison 1: Avatar therapy versus treatment as usual, Outcome 8: Mental state: 2c. General – total – average endpoint score (PSRS, high poor, skewed data) – short term 37
Analysis 1.9. Comparison 1: Avatar therapy versus treatment as usual, Outcome 9: Global state: 1. Any change – needing counselling and support – short term 38
Analysis 1.10. Comparison 1: Avatar therapy versus treatment as usual, Outcome 10: Global state: 2. Relapse (rehospitalised) – short term 38
Analysis 1.11. Comparison 1: Avatar therapy versus treatment as usual, Outcome 11: Leaving study early: 1. For any reason – short term 38
Analysis 1.12. Comparison 1: Avatar therapy versus treatment as usual, Outcome 12: Leaving study early: 2. For specific reason – short term 39
Analysis 1.13. Comparison 1: Avatar therapy versus treatment as usual, Outcome 13: Adverse effects: anxiety – short term 39
Analysis 1.14. Comparison 1: Avatar therapy versus treatment as usual, Outcome 14: Quality of life: average endpoint score (QLESQ-SF, high = good) – short term 39
Analysis 2.1. Comparison 2: Avatar therapy versus supportive counselling, Outcome 1: Mental state: 1a. Specific – positive – average endpoint score (SAPS, high = poor, skewed data) – short term 41
Analysis 2.2. Comparison 2: Avatar therapy versus supportive counselling, Outcome 2: Mental state: 1b. Specific – insight – average endpoint score (various scales, high = poor) – short term 41
Analysis 2.3. Comparison 2: Avatar therapy versus supportive counselling, Outcome 3: Mental state: 1c. Specific – depression – average endpoint score (various scales, high = poor, skewed data) – short term 42
Analysis 2.4. Comparison 2: Avatar therapy versus supportive counselling, Outcome 4: Mental state: 1d. Specific – negative – average endpoint score (SANS, high = good, skewed data) – short term 42
Analysis 2.5. Comparison 2: Avatar therapy versus supportive counselling, Outcome 5: Mental state: 2a. General – average endpoint score (PSYRATS total, high = poor) – short term 42

Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 2.6. Comparison 2: Avatar therapy versus supportive counselling, Outcome 6: Leaving study early: 1a. For any reason – short term 42
Analysis 2.7. Comparison 2: Avatar therapy versus supportive counselling, Outcome 7: Leaving study early: 1b. For specific reason – short term 43
Analysis 2.8. Comparison 2: Avatar therapy versus supportive counselling, Outcome 8: Quality of life: average endpoint score (various scales, high = good) – short term 44
ADDITIONAL TABLES 44
HISTORY 45
CONTRIBUTIONS OF AUTHORS 45
DECLARATIONS OF INTEREST 45
SOURCES OF SUPPORT 45
DIFFERENCES BETWEEN PROTOCOL AND REVIEW 45
INDEX TERMS 46

Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

[Intervention Review]

Avatar Therapy for people with schizophrenia or related disorders

Ghazaleh Aali[1], Timothy Kariotis[2], Farhad Shokraneh[3]

[1]Cochrane Schizophrenia Group, Institute of Mental Health, The University of Nottingham, Nottingham, UK. [2]School of Computing and Information Systems, University of Melbourne, Parkville, Melbourne, Australia. [3]Cochrane Schizophrenia Group, Institute of Mental Health, Division of Psychiatry and Applied Psychology, School of Medicine, University of Nottingham, Nottingham, UK

Contact address: Ghazaleh Aali, ghazaleh.aali@nottingham.ac.uk.

Editorial group: Cochrane Schizophrenia Group.
Publication status and date: New, published in Issue 5, 2020.

Citation: Aali G, Kariotis T, Shokraneh F. Avatar Therapy for people with schizophrenia or related disorders. *Cochrane Database of Systematic Reviews* 2020, Issue 5. Art. No.: CD011898. DOI: 10.1002/14651858.CD011898.pub2.

Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

ABSTRACT

Background

Many people with schizophrenia do not achieve satisfactory improvements in their mental state, particularly the symptom of hearing voices (hallucinations), with medical treatment.

Objectives

To examine the effects of Avatar Therapy for people with schizophrenia or related disorders.

Search methods

In December 2016, November 2018 and April 2019, the Cochrane Schizophrenia Group's Study-Based Register of Trials (including registries of clinical trials) was searched, review authors checked references of all identified relevant reports to identify more studies and contacted authors of trials for additional information.

Selection criteria

All randomised clinical trials focusing on Avatar Therapy for people with schizophrenia or related disorders.

Data collection and analysis

We extracted data independently. For binary outcomes, we calculated risk ratio (RR) and 95% confidence intervals (CI), on an intention-to-treat basis. For continuous data, we estimated the mean difference (MD) between groups and 95% CIs. We employed a fixed-effect model for analyses. We assessed risk of bias for included studies and created 'Summary of findings' tables using GRADE. Our main outcomes of interest were clinically important change in; mental state, insight, global state, quality of life and functioning as well as adverse effects and leaving the study early.

Main results

We found 14 potentially relevant references for three studies (participants = 195) comparing Avatar Therapy with two other interventions; treatment as usual or supportive counselling. Both Avatar Therapy and supportive counselling were given in addition (add-on) to the participants' normal care. All of the studies had high risk of bias across one or more domains for methodology and, for other risks of bias, authors from one of the studies were involved in the development of the avatar systems on trial and in another trial, authors had patents on the avatar system pending.

1. Avatar Therapy compared with treatment as usual

Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

When Avatar Therapy was compared with treatment as usual average endpoint Positive and Negative Syndrome Scale – Positive (PANSS-P) scores were not different between treatment groups (MD –1.93, 95% CI –5.10 to 1.24; studies = 1, participants = 19; very low-certainty evidence). A measure of insight (Revised Beliefs about Voices Questionnaire; BAVQ-R) showed an effect in favour of Avatar Therapy (MD –5.97, 95% CI –10.98 to –0.96; studies = 1, participants = 19; very low-certainty evidence). No one was rehospitalised in either group in the short term (risk difference (RD) 0.00, 95% CI –0.20 to 0.20; studies = 1, participants = 19; low-certainty evidence). Numbers leaving the study early from each group were not clearly different – although more did leave from the Avatar Therapy group (6/14 versus 0/12; RR 11.27, 95% CI 0.70 to 181.41; studies = 1, participants = 26; low-certainty evidence). There was no clear difference in anxiety between treatment groups (RR 5.54, 95% CI 0.34 to 89.80; studies = 1, participants = 19; low-certainty evidence). For quality of life, average Quality of Life Enjoyment and Satisfaction Questionnaire-Short Form (QLESQ-SF) scores favoured Avatar Therapy (MD 9.99, 95% CI 3.89 to 16.09; studies = 1, participants = 19; very low-certainty evidence*)*. No study reported data for functioning.

2. Avatar Therapy compared with supportive counselling

When Avatar Therapy was compared with supportive counselling (all short-term), general mental state (Psychotic Symptom Rating Scale (PSYRATS)) scores favoured the Avatar Therapy group (MD –4.74, 95% CI –8.01 to –1.47; studies = 1, participants = 124; low-certainty evidence). For insight (BAVQ-R), there was a small effect in favour of Avatar Therapy (MD –8.39, 95% CI –14.31 to –2.47; studies = 1, participants = 124; low-certainty evidence). Around 20% of each group left the study early (risk ratio (RR) 1.06, 95% CI 0.59 to 1.89; studies = 1, participants = 150; moderate-certainty evidence). Analysis of quality of life scores (Manchester Short Assessment of Quality of Life (MANSA)) showed no clear difference between groups (MD 2.69, 95% CI –1.48 to 6.86; studies = 1, participants = 120; low-certainty evidence). No data were available for rehospitalisation rates, adverse events or functioning.

Authors' conclusions

Our analyses of available data shows few, if any, consistent effects of Avatar Therapy for people living with schizophrenia who experience auditory hallucinations. Where there are effects, or suggestions of effects, we are uncertain because of their risk of bias and their unclear clinical meaning. The theory behind Avatar Therapy is compelling but the practice needs testing in large, long, well-designed, well-reported randomised trials undertaken with help from – but not under the direction of – Avatar Therapy pioneers.

PLAIN LANGUAGE SUMMARY

Avatar Therapy for schizophrenia or related disorders

Review question

Is Avatar Therapy an effective add-on treatment for people with schizophrenia and schizoaffective disorder?

Background

Auditory hallucination is perceiving voices when there is no external stimulus. Around 70% of people with schizophrenia experience these. Medication may help cause these to decrease or disappear. However, some people do not want to take medication and, in a proportion of people, it has little meaningful effect. Avatar Therapy is an experimental technology that uses a visualised avatar face, voice and other sensory input to create an interactive computerised environment. We investigated the effects of Avatar Therapy for improving auditory hallucinations for people with schizophrenia.

Searches

Cochrane Schizophrenia's Information Specialist searched through the Cochrane Schizophrenia Group's database for randomised trials (clinical studies where people are randomly put into one of two or more treatment groups) of people with schizophrenia receiving either Avatar Therapy or treatment as usual. We found 14 reports from four studies.

The evidence is current to April 2020.

Trials

Four studies met our inclusion criteria, three of which provided useable data and one is still underway. The reliability of the evidence for the three included studies was very low to low. All data were limited to six weeks' treatment and one week' follow-up and, in some cases, data were not usable (we contacted the authors). All the reported data are just for the short-term period of trials and it is apparent that we need more studies with medium- to long-term periods to be able to estimate the rate of Avatar Therapy's effectiveness.

Conclusions

Three short-term studies with 195 participants were included in this review. Avatar Therapy was compared with treatment as usual and supportive counselling. Evidence from the trials was not high quality. Although there were some suggestions of positive effects, because of the unclear meaning for front-line care, and the considerable risk of bias of each of the results, we cannot be certain of these effects. There is just as great a risk of Avatar Therapy causing problems to people. More trials are needed, and those undertaking them should collaborate,

Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

work with all involved to choose outcomes that mean something in daily life, and design studies that will provide clear information across months to years for anyone who is interested.

Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

SUMMARY OF FINDINGS

Summary of findings 1. Avatar Therapy compared to treatment as usual (all short-term) for schizophrenia or related disorders

Avatar Therapy compared to treatment as usual (all short-term) for schizophrenia or related disorders

Patient or population: schizophrenia or related disorders

Setting:

Intervention: Avatar Therapy

Comparison: treatment as usual (all short-term)

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Certainty of the evidence (GRADE)	Comments
	Risk with treatment as usual (all short-term)	Risk with Avatar Therapy				
Mental state*: specific – **positive** – average endpoint score (PANSS-P, high = poor) – short term	—	**MD 1.93 lower** (5.10 lower to 1.24 higher)	—	19 (1 RCT)	⊕⊝⊝⊝ **Very low** a,b,c	This was the nearest proxy measure for our prestated binary outcome, clinically important change in mental state.
Mental state: specific – **insight** – average attitude to voices score (BAVQ-R, high = poor) – short term	—	**MD 5.97 lower** (10.98 lower to 0.96 lower)	—	19 (1 RCT)	⊕⊝⊝⊝ **Very low** a,b,c	This was the nearest proxy measure for our prestated binary outcome, clinically important change in insight.
Global state: relapse (rehospitalised) – short term	0 per 1000	0 per 1000 (0 to 0)	**RD 0.0** (–0.2 to 0.2)	19 (1 RCT)	⊕⊕⊝⊝ **Low** a,c	RD reported as no-one in either group was hospitalised.
Leaving study early: for any reason	0 per 1000	0 per 1000 (0 to 0)	**RR 11.27** (0.70 to 181.41)	26 (1 RCT)	⊕⊕⊝⊝ **Low** a,c	Avatar Therapy likely results in little to no difference for the outcome of 'leaving study early for any reason'.
Adverse events – **anxiety** – after first session – short term	0 per 1000	0 per 1000 (0 to 0)	**RR 5.54** (0.34 to 89.80)	19 (1 RCT)	⊕⊕⊝⊝ **Low** a,c	—
Quality of life: average endpoint score (QLESQ-SF, high = good) – short term	—	**MD 9.99 higher** (3.89 higher to 16.09 higher)	—	19 (1 RCT)	⊕⊝⊝⊝ **Very low** a,b,c	This was the nearest proxy measure for our prestated binary outcome, clin-

Cochrane Library
Trusted evidence.
Informed decisions.
Better health.
Cochrane Database of Systematic Reviews

						ically important change in quality of life.
Functioning: clinically important change	See comment	See comment	—	(0 studies)	—	No study reported on this important outcome.

*__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).

BAVQ-R: Revised Beliefs about Voices Questionnaire; **CI:** confidence interval; **MD:** mean difference; **PANSS-P:** Positive and Negative Syndrome Scale – Positive; **QLESQ-SF:** Quality of Life Enjoyment and Satisfaction Questionnaire-Short Form; **RCT:** randomised controlled trial; **RD:** risk difference; **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]Risk of bias: 'serious' (downgraded one level) – poor reporting of randomisation, no substantive discussion of any blinding or consideration of blinding. Authors contacted but limited information available.
[b]Indirectness: 'serious' (downgraded one level) – proxy scale for binary outcome prestated in protocol.
[c]Imprecision: 'serious' (downgraded one level) – wide confidence intervals generated by very small study.

Summary of findings 2. Avatar Therapy compared to supportive counselling (all short-term) for schizophrenia or related disorders

Avatar therapy compared to supportive counselling (all short-term) for schizophrenia or related disorders

Patient or population: schizophrenia or related disorders

Setting:

Intervention: Avatar Therapy

Comparison: supportive counselling (all short-term)

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Certainty of the evidence (GRADE)	Comments
	Risk with supportive counselling (all short-term)	Risk with Avatar Therapy				

Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Mental state*: general – average endpoint score (PSYRATS, high = poor)	—	**MD 4.74 lower** (8.01 lower to 1.47 lower)	—	124 (1 RCT)	⊕⊕⊖⊖ **Low** [a,b]	This was the nearest proxy measure for our prestated binary outcome, clinically important change in positive symptoms. Avatar Therapy may reduce mental state scores slightly but it is unclear how these ratings relate to everyday life and functioning.
Mental state: specific – insight – average attitude to voices score (BAVQ-R, high = poor) – short term	—	**MD 8.39 lower** (14.31 lower to 2.47 lower)	—	124 (1 RCT)	⊕⊕⊖⊖ **Low** [a,b]	Avatar Therapy likely results in a little difference in a measure of insight (average attitude to voices score (BAVQ-R, high score = poor)). This was the nearest proxy measure for our prestated binary outcome, clinically important change in insight. We are unclear of the clinical meaning of this measure.
Global state: relapse (rehospitalised)	See comment	See comment	—	(0 studies)	—	No study reported on this important outcome.
Leaving study early: for any reason – short term	0 per 1000	0 per 1000 (0 to 0)	**RR 1.06** (CI 0.59 to 1.89)	150 (1 RCT)	⊕⊕⊕⊖ **Moderate** [b]	Avatar Therapy results in little to no difference in leaving study early (any reason by up to 6 months).
Adverse events – anxiety – after first session	See comment	See comment	—	(0 studies)	—	No study reported on this important outcome.
Quality of life: average endpoint score (MANSA, high = good) – short term	—	**MD 2.69 higher** (1.48 lower to 6.86 higher)	—	120 (1 RCT)	⊕⊕⊖⊖ **Low** [a,b]	Avatar Therapy likely results in little to no difference in quality of life – as measured by MANSA. This was the nearest proxy measure for our prestated binary outcome, clinically important change in quality of life.
Functioning: clinically important change	See comment	See comment	—	(0 studies)	—	No study reported on this important outcome.

*__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).

BAVQ-R: Revised Beliefs about Voices Questionnaire; **CI:** confidence interval; **MANSA:** Manchester Short Assessment of Quality of Life; **MD:** mean difference; **PSYRATS:** Psychotic Symptom Rating 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.

Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

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]Indirectness: 'serious' (downgraded one level) – proxy scale for binary outcome prestated in protocol.
[b]Imprecision: 'serious' (downgraded one level) – wide confidence intervals generated by small study.

Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

BACKGROUND

Description of the condition

Schizophrenia is a persistent, relapsing and disabling mental disorder that has a global prevalence between 0.4% and 0.7% (Saha 2005). This serious mental illness is characterised by 1. positive symptoms, such as perceptions with no cause (hallucinations) or false beliefs (delusions), or both; and 2. negative symptoms which include catatonic signs, disorganised thoughts and behaviour, apathy and lack of motivation (Carpenter 1994). Hearing voices (auditory hallucinations) is a common symptom of schizophrenia, which is often treatment-resistant. For those people with schizophrenia who suffer from hearing one or more voices, around 70% of people with schizophrenia experience these (Comer 2017), and 30% will have auditory hallucinations that are not alleviated by taking medication (Kane 1996).

Description of the intervention

Antipsychotic medications are the main line of treatment for schizophrenia; however, some of the symptoms of schizophrenia, such as auditory hallucinations, are treatment-resistant (Kane 1996). Non-pharmaceutical psychotherapeutic interventions such as cognitive behavioural therapy (CBT) are often suggested as additional treatments. However, CBT's effect size for treating the positive symptoms is small, and its effect on relapse and the negative symptoms of schizophrenia has not been proved (McKenna 2014).

Virtual reality has also been used for people with schizophrenia for social skills training, to improve thinking and understanding processes, and for treatment support (Kip 2019); however, Avatar Therapy is a relatively recently suggested treatment which could be used as an adjunct to the usual antipsychotic treatment, particularly for treatment-resistant auditory hallucinations (Leff 2013).

Avatar Therapy has been suggested for people who experience distressing voices, hypothesising that it could reduce the severity and rate of auditory hallucinations (Huckvale 2013; ISRCTN65314790 2013). During the therapy, a psychiatrist assists the patient to create an avatar, an audio-visual entity, using a computer program. This avatar is modified so that its characteristics match the voices that bother the patient (Leff 2013). Then the therapist uses this avatar to talk back to the patient in therapy sessions and they can exercise taught methods to cope with the voices. While the person with schizophrenia is encouraged to establish dialogues with the avatar and be resistant to the hallucination, the therapist then manages the avatar so it is gradually controlled by the patient and the avatar's mode changes from persecutory to supportive during the therapy sessions (Leff 2013). Audio recordings of all sessions are made so that the patient can listen to them at home using an MP3 Player (Rus-Calafell 2014). The studies have cited Leff 2013 as the first trial; however, it is unclear if the other studies used the same Avatar Therapy intervention design or made their own intervention.

How the intervention might work

One of the main aspects people with schizophrenia describe about their auditory hallucinations is that they often feel helpless in coping with the voices (Leff 2013). Interacting with the hallucination may help people with schizophrenia. It has been reported that people with schizophrenia who are able to talk back to the voices could have more control over the voices (Nayani 1996); and their suffering gradually is reduced (Leff 2013). However, interacting with a hallucination involves interacting with an invisible character and can be difficult because of lack of mutual interaction and body/face language. Many people with schizophrenia often imagine a voice-associated face when hearing a hallucination, but cannot see the face to interact with it. An avatar is both an audio (voice) and a visual (face) entity, and can be used as an audio-visual interface between the hallucination and the patient to facilitate the establishment of a conversation between patient and the voice (Leff 2013). Also, there might be other advantages in talking to an avatar; first, Avatar Therapy sessions can be recorded and taken home with the patient to use when needed – a psychiatrist is unlikely to be available at all times and places; second, an avatar's voice can be made to be very similar to the hallucination the patient hears; an avatar's voice might help people with schizophrenia to practise how to take control of voices.

Why it is important to do this review

There is currently no systematic review of Avatar Therapy for treatment-resistant auditory hallucinations among people with schizophrenia. We aimed to examine whether Avatar Therapy could have a therapeutic effect on auditory hallucinations in people with schizophrenia. In addition, since there are primary studies to be systematically reviewed to meet our aim and Avatar Therapy is a new treatment, this review found and synthesised the relevant studies on the effects of Avatar Therapy for people with schizophrenia who experience treatment-resistant auditory hallucinations so as to provide clinically useful evidence for informed decision-making among clinicians, people with schizophrenia and health policy makers.

OBJECTIVES

To examine the effects of Avatar Therapy for people with schizophrenia or related disorders.

METHODS

Criteria for considering studies for this review

Types of studies

All relevant randomised controlled trials; if a trial was described as 'double blind' but implied randomisation, we included such trials in a sensitivity analysis (see Sensitivity analysis). If their inclusion did not result in a substantive difference, they remained in the analyses. If their inclusion resulted in important clinically significant but not necessarily statistically significant differences, we did not add the data from these lower-quality studies to the results of the better trials, but presented such data within a subcategory. We found no quasi-randomised studies, such as those allocating by alternate days of the week to exclude. Where people were given additional treatments within Avatar Therapy, we only included data if the adjunct treatment was evenly distributed between groups and it was only the Avatar Therapy that was randomised.

Types of participants

Patients, however defined, with schizophrenia or related disorders, including schizophreniform disorder, schizoaffective disorder and

Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

delusional disorder, again, by any means of diagnosis, of any age, gender or ethnicity.

We were interested in ensuring that information was as relevant to the current care of people with schizophrenia as possible so propose to clearly highlight the current clinical state (acute, early postacute, partial remission, remission) as well as the stage (prodromal, first episode, early illness, persistent) and as to whether the studies primarily focused on people with particular problems (e.g. negative symptoms, treatment-resistant illnesses).

Types of interventions

1. Avatar Therapy

An avatar is a computer-generated audio-visual character utilised to establish a conversation between the person with schizophrenia and the voices that person hears, to cope with the voices.

We assessed all the randomised controlled trials in which Avatar Therapy was one of the study arms either alone or in combination with other interventions. Avatar Therapy is an adjunctive treatment for 'treatment as usual'.

2. Control interventions

Any other add-on (to the participants usual care) intervention or treatment as usual.

3. Supportive counselling

Any face-to-face supportive counselling approach.

Types of outcome measures

We divided all outcomes into short term (less than six months), medium term (six to 12 months) and long term (over 12 months).

We sought 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 defined within the trials). Thereafter, we listed other binary outcomes and then those that were continuous. See Differences between protocol and review.

Primary outcomes

1. Mental state – specific (positive – e.g. auditory hallucinations, delusions)

1.1. Clinically important change in positive symptoms

2. Mental state – specific (insight)

2.1. Clinically important change in insight

3. Adverse effects

3.1. At least one important adverse effect

Secondary outcomes

1. Mental state

1.1. Specific

1.1.1. Clinically important change in specific symptoms (other than positive)
1.1.2. Any change in specific symptoms
1.1.3. Average endpoint or change score on specific mental state scale

1.2 General

1.2.1. Clinically important change in general mental state
1.2.2. Any change in general mental state
1.2.3. Average endpoint or change score on general mental state scale

2. Global state

2.1. Clinically important change in global state
2.2. Any change in global state
2.3. Relapse (including rehospitalisation)
2.4. Average endpoint or change score on global state scale

3. Leaving the study early

3.1. For any reason.
3.2. For specific reason

4. Adverse effects

4.1. Clinically important adverse effects
4.2. Any adverse effect
4.3. Any change in specific adverse effects

5. Quality of life

5.1. Clinically important change in quality of life
5.2. Any change in quality of life
5.3. Average endpoint or change score on quality of life scale

6. Functioning

6.1. Clinically important change in functioning
6.2. Any change in functioning
6.3. Average endpoint or change score on functioning scale

'Summary of findings' table

We used the GRADE approach to interpret findings (Schünemann 2011); and used GRADE profiler (GRADEpro; tech.cochrane.org/revman/gradepro) to import data from Review Manager 5 (tech.cochrane.org/revman) to create 'Summary of findings' tables. These tables provide outcome-specific information concerning the overall certainty of the evidence from each included study in the comparison, the magnitude of effect of the interventions examined, and the sum of available data on all primary outcomes and on selected secondary outcomes. This summary guided our conclusions and recommendations. We selected the following main (short term) outcomes for inclusion in the 'Summary of findings' table:

1. Mental state: clinically important change in positive symptoms (e.g. auditory hallucinations, delusions)
2. Mental state: insight: clinically important change in insight.
3. Adverse effects: at least one important adverse effect.
4. Global state: clinically important change in global state.
5. Leaving the study early: for any reason.
6. Quality of life: clinically important change in quality of life.
7. Functioning: clinically important change in functioning.

If data were not available for these prespecified outcomes but were available for ones that were similar, we presented the closest outcome available, but considered this when grading the finding.

Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Search methods for identification of studies

Electronic searches

1. Cochrane Schizophrenia Group's Study-Based Register of Trials

On 8 December 2016, 9 November 2018 and 14 April 2020, the information specialist searched the register using the following search strategy:

(*Avatar Therapy*) in Intervention Field of STUDY

In such a 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; Shokraneh 2019).

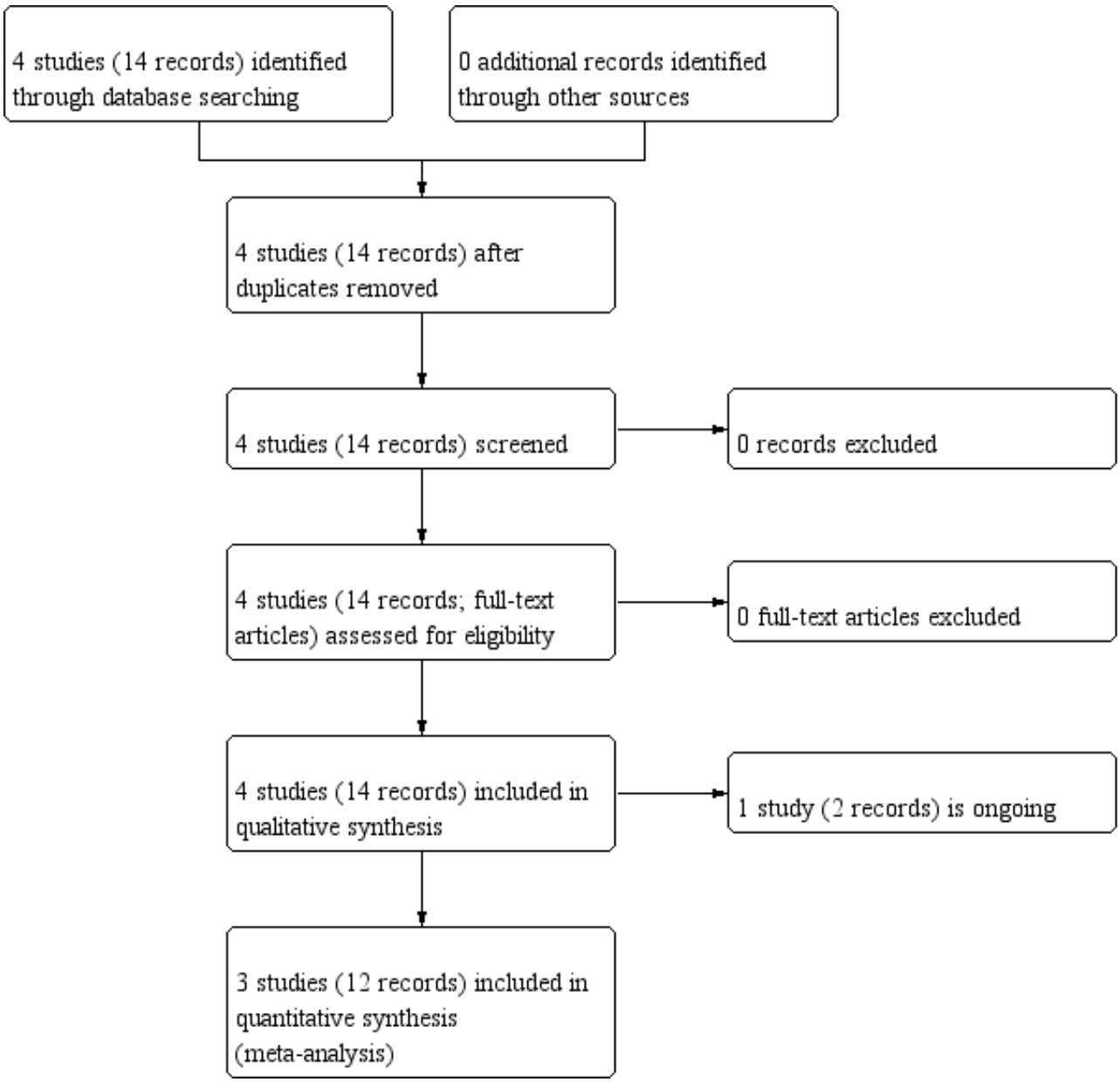
This register is compiled by systematic searches of major resources (AMED, BIOSIS, CENTRAL, CINAHL, ClinicalTrials.Gov, Embase, MEDLINE, PsycINFO, PubMed, WHO ICTRP) and their monthly updates; ProQuest Dissertations and Theses A&I and its quarterly update; Chinese databases (CBM, CNKI, and Wanfang) and their annual updates; handsearches; grey literature and conference proceedings (see Group's website; schizophrenia.cochrane.org/register-trials). There is no language, date, document type, or publication status limitations for inclusion of records into the register.

Searching other resources

1. Reference searching

We inspected references of all included studies for further relevant studies (Figure 1).

Figure 1. Study flow diagram.



Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

2. Personal contact

We contacted the first author of a study for additional information regarding non-reported data (NCT03148639). We noted the outcome of this contact in the Characteristics of included studies table.

Data collection and analysis

Selection of studies

Review authors (GA, TK and FS) independently inspected citations from the searches and identified relevant abstracts; GA independently re-inspected a random 20% sample of these abstracts to ensure reliability of selection. Where disputes arose, we acquired the full report for a more detailed scrutiny. Review authors (GA, TK and FS) then obtained and inspected full reports of the abstracts or reports meeting the review criteria. GA re-inspected a random 20% of these full reports to ensure reliability of selection. Where it was not possible to resolve disagreement by discussion, we contacted the authors of the study concerned for clarification (NCT03148639).

Data extraction and management

1. Extraction

Review authors (GA, TK and FS) independently extracted data from all included studies. In addition, to ensure reliability, GA independently extracted data from a random sample of these studies, comprising 10% of the total. We attempted to extract data presented only in graphs and figures whenever possible, but only included if two review authors independently obtained the same result. We contacted authors through an open-ended request in order to obtain missing information or for clarification whenever necessary. If studies were multicentre, then, where possible, we extracted data relevant to each component centre separately.

2. Management

2.1. Forms

We extracted data onto standard, pre-designed, simple forms.

2.2. Scale-derived data

We included continuous data from rating scales only if:

1. the psychometric properties of the measuring instrument had been described in a peer-reviewed journal (Marshall 2000); and
2. the measuring instrument had not been written or modified by one of the trialists for that particular trial.
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 included 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 realise that this is not often reported clearly: in Description of studies, we noted if this was the case or not.

2.3. Endpoint versus change data

There are advantages of both endpoint and change data. Change data can remove a component of between-person variability from the analysis. Also, calculation of change needs two assessments (baseline and endpoint) which can be difficult in unstable and difficult-to-measure conditions such as schizophrenia. We primarily used endpoint data, and would only use change data if the former were not available. Endpoint and change data would combined in the analysis, as we prefer using mean differences (MD) rather than standardised mean differences throughout (Higgins 2011).

2.4. Skewed data

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

For endpoint data:

- When a scale started from the finite number zero, we subtracted the lowest possible value from the mean, and divided this by the standard deviation (SD). If this value was lower than one, it strongly suggested that the data were skewed and we excluded these data. If this ratio was higher than one but less than two, there wss suggestion that the data were skewed: we entered these data and tested whether their inclusion or exclusion would change the results substantially. Finally, if the ratio was larger than two, we included these data, because it was less likely that they were skewed (Altman 1996; Higgins 2011);
- If a scale started from a positive value (such as the Positive and Negative Syndrome Scale (PANSS), which can have values from 30 to 210 (Kay 1986)), we modified the calculation described above to take the scale starting point into account. In these cases, skewed data are present if $2\ SD > (S - S_{min})$, where S is the mean score and S_{min} is the minimum score.

Please note: we would enter data from studies of at least 200 participants, for example, in the analysis irrespective of the above rules, because skewed data pose less of a problem in large studies. We also would have entered change data as when continuous data were presented on a scale that included a possibility of negative values (such as change data), if it was difficult to determine whether data were skewed or not. We would have presented and entered change data into statistical analyses.

2.5. Common measurement

To facilitate comparison between trials, we would have converted variables that were 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

Where possible, we would have converted outcome measures to dichotomous data. This can be done 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

Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

response (Leucht 2005a; Leucht 2005b). If data based on these thresholds were not available, we would have used the primary cut-off presented by the original authors.

2.7. Direction of graphs

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

Assessment of risk of bias in included studies

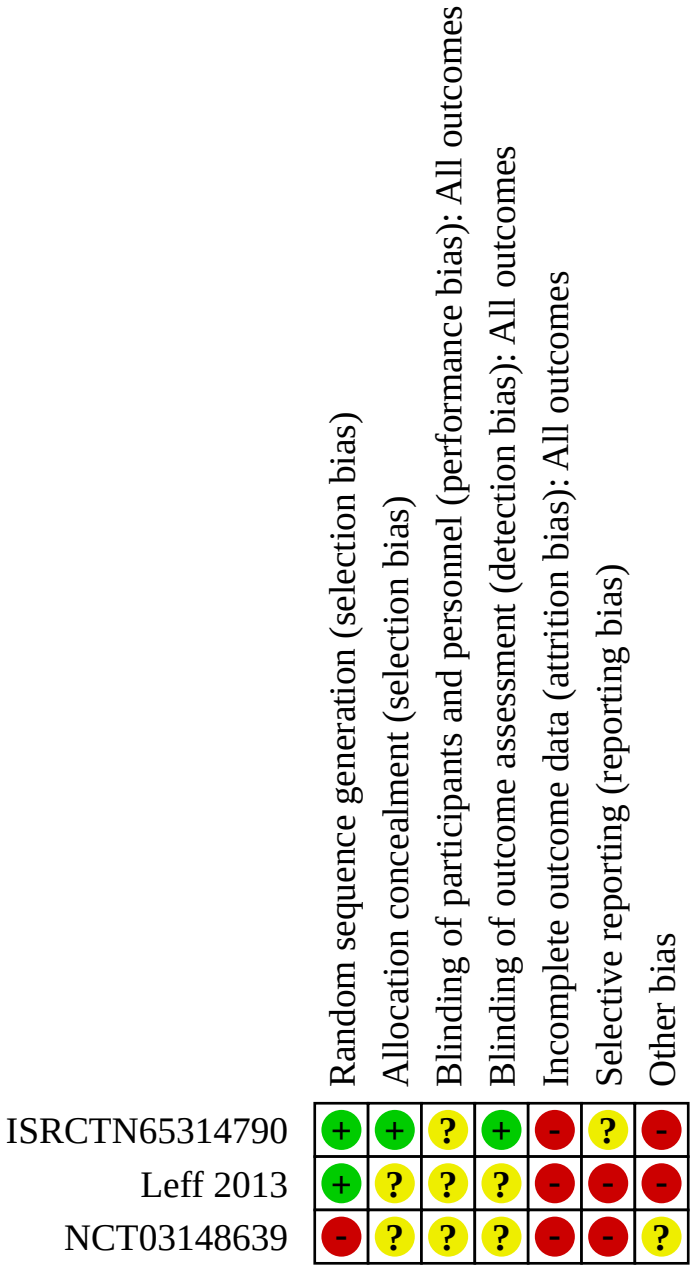
Review authors (GA and FS) independently assessed risk of bias using criteria described in the *Cochrane Handbook for Systemic Reviews of Interventions* to assess trial quality (Higgins 2011). These criteria are based on evidence of associations between overestimate of effect and high risk of bias of the article such as sequence generation, allocation concealment, blinding, incomplete outcome data and selective reporting, or the way in which these 'domains' are reported.

If the raters disagreed, the final rating was made by consensus, with the involvement of another member of the review group. Where inadequate details of randomisation and other characteristics of trials were provided, we attempted to contact authors of the studies to obtain further information. We reported non-concurrence in quality assessment, but if disputes arose regarding the category to which a trial was to be allocated, we resolved this by discussion.

We noted the level of risk of bias in the text of the review, 'Risk of bias' summary (Figure 2), 'Risk of bias' graph (Figure 3), and the Characteristics of included studies tables. (Figure 2)

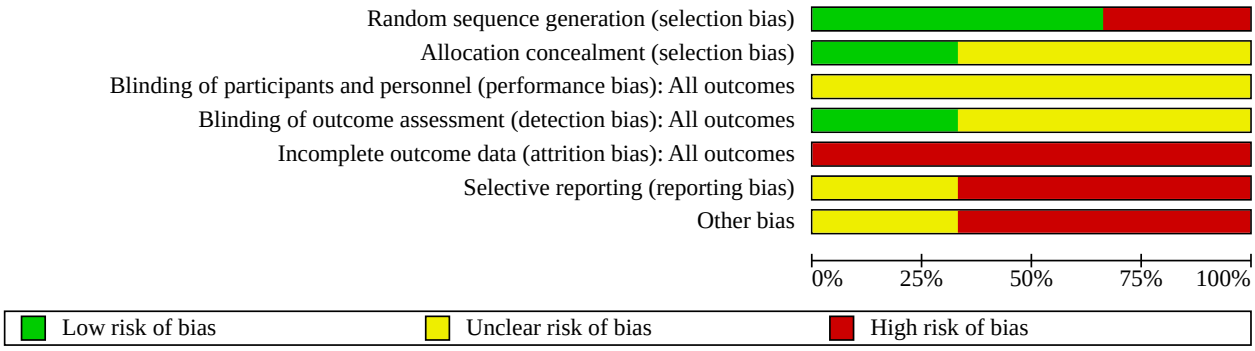
Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Figure 2. Risk of bias summary: review authors' judgements about each risk of bias item for each included study.



Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Figure 3. Risk of bias graph: review authors' judgements about each risk of bias item presented as percentages across all included studies.



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 RR is more intuitive than odds ratios (Boissel 1999); and that odds ratios tend to be interpreted as RR by clinicians (Deeks 2000). 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 CIs, are intuitively attractive to clinicians, they are problematic to calculate and interpret in meta-analyses (Hutton 2009). For binary data presented in the 'Summary of findings' tables, we, where possible, would have calculated illustrative comparative risks.

2. Continuous data

For continuous outcomes, we estimated mean difference (MD) between groups, with 95% CIs. We preferred not to calculate effect size measures (standardised mean difference (SMD)). However, if scales of very considerable similarity were used, we presumed there was a small difference in measurement, and we calculated effect size and transform the effect 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 poses problems. First, authors often fail to account for intraclass 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 was not accounted for in primary studies, we presented data in a table, with a (*) symbol to indicate the presence of a probable unit of analysis error. We sought to contact first authors of studies to obtain intraclass correlation coefficients (ICC) for their clustered data and to adjust for this by using accepted methods (Gulliford 1999). Where clustering had been incorporated into the analysis of primary studies, we presented these data as if from a non-cluster randomised study, but adjust for the clustering effect.

We sought statistical advice and have been advised 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 assumed it to be 0.1 (Ukoumunne 1999).

If cluster studies were appropriately analysed taking into account ICCs and relevant data documented in the report, 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 was 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, on entry to the second phase the participants can differ systematically from their initial state despite a washout phase. For the same reason, cross-over trials are not appropriate if the condition of interest is unstable (Elbourne 2002). As both effects are very likely in severe mental illness, we would only have used data of the first phase of cross-over studies.

3. Studies with multiple treatment groups

Where a study involved more than two treatment arms, if relevant, the additional treatment arms would have been presented in comparisons. If data were binary, these would simply have been added and combined within the two-by-two table. If data were continuous, we would have combined data following the formula in Section 7.7.3.8 (Combining groups) of the *Cochrane Handbook for Systemic Reviews of Interventions* (Higgins 2011). Where the 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 to follow-up, data must lose credibility (Xia 2009). We chose 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 were lost, but the total loss was less than 50%, we would have addressed this within the 'Summary of findings' tables by downgrading the certainty of the evidence.

Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Finally, we also downgraded the certainty of the evidence within the 'Summary of findings' tables should loss have been 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 presented data on a 'once-randomised-always-analyse' basis (an intention-to-treat (ITT) analysis). Those leaving the study early were all assumed to have the same rates of negative outcome as those who completed, with the exception of the outcome of death and adverse effects. For these outcomes, the rate of those who stay in the study – in that particular arm of the trial – would have been used for those who did not. We undertook a sensitivity analysis testing 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 used data where attrition for a continuous outcome was between 0% and 50%, and data only from people who complete the study to that point were reported.

3.2. Standard deviations

If SDs were not reported, we would have tried to obtain the missing values from the 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 P value or t value available for differences in mean, we could have calculated SDs according to the rules described in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011). When only the SE was reported, SDs were calculated 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 (Higgins 2011). 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 have been to exclude a given study's outcome and thus to lose information. Nevertheless, we examined the validity of the imputations in a sensitivity analysis that excluded imputed values.

3.3. Assumptions about participants who left 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 just present the results of study completers; others use the method of last observation carried forward (LOCF); while more 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 feel 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. Therefore, we did not exclude studies based on the statistical approach used. However, by preference we used the more sophisticated approaches, that is, we preferred to use MMRM or multiple-imputation to LOCF, and we only presented completer analyses if some type of ITT data were not available at all. Moreover, we addressed this issue in the item 'Incomplete outcome data' of the 'Risk of bias' tool.

Assessment of heterogeneity

1. Clinical heterogeneity

We considered all included studies initially, without seeing comparison data, to judge clinical heterogeneity. We simply inspected all studies for people or situations 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 simply 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. Employing 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 statistic depends on 1. magnitude and direction of effects and 2. strength of evidence for heterogeneity (e.g. P value from Chi^2 test, or a CI for the I^2 statistic). An I^2 statistic estimate of about 50% or greater accompanied by a statistically significant Chi^2 statistic would be interpreted as evidence of substantial levels of heterogeneity (Section 9.5.2; Higgins 2011). When we found substantial levels of heterogeneity in the primary outcome, we explored reasons for heterogeneity (Subgroup analysis and investigation of heterogeneity).

Assessment of reporting biases

1. Protocol versus full study

Reporting biases arise when the dissemination of research findings is influenced by the nature and direction of results. These are described in Section 10.1 of the *Cochrane Handbook for Systemic Reviews of Interventions* (Higgins 2011). We searched for protocols of included randomised trials. If the protocol was available, we compared outcomes in the protocol and in the published report. If the protocol was not available, we compared outcomes listed in the Methods section of the trial report with actually reported results.

2. Funnel plot

Reporting biases arise when the dissemination of research findings is influenced by the nature and direction of results (Egger 1997). These are described in Section 10 of the *Cochrane Handbook for Systemic Reviews of Interventions* (Higgins 2011). We are aware that funnel plots may be useful in investigating reporting biases but are of limited power to detect small-study effects. We did not use funnel

Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

plots for outcomes where there were 10 or fewer studies, or where all studies were of similar sizes. In other cases, where funnel plots were possible, we 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. However, there is a disadvantage to the random-effects model as it puts added weight onto small studies, which often are the most biased ones. Depending on the direction of effect, these studies can either inflate or deflate the effect size. We chose the random-effects model for all analyses. The reader is, however, able to choose to inspect the data using the fixed-effect model.

Subgroup analysis and investigation of heterogeneity

1. Subgroup analyses

1.1. Primary outcomes

We did not conduct a subgroup analysis as we did not anticipate sufficient power to carry it out.

1.2. Clinical state, stage or problem

We proposed to undertake this review and provide an overview of the effects of Avatar Therapy for people with schizophrenia in general. In addition, we reported data on subgroups of people in the same clinical state, stage and with similar problems.

2. Investigation of heterogeneity

We reported if inconsistency was high. First, we investigated whether data had been entered correctly. Second, if data were correct, we inspected the graph visually and removed outlying studies successively to see if homogeneity was restored. For this review, we decided that should this occur with data contributing to the summary finding of no more than around 10% of the total weighting, we presented data. If not, we would not pool these data and would discuss any issues. We knew of no supporting research for this 10% cut-off but investigated use of prediction intervals as an alternative to this unsatisfactory state.

When unanticipated clinical or methodological heterogeneity were obvious we would have simply stated hypotheses regarding these for future reviews or versions of this review. We did not anticipate undertaking analyses relating to these.

Sensitivity analysis

1. Implication of randomisation

We aimed to include trials in a sensitivity analysis if they were described in some way that implied randomisation. For the primary outcomes, we included these studies and if there was no substantive difference when the implied randomised studies were added to those with better description of randomisation, then all data employed from these studies.

2. Assumptions for lost binary data

Where assumptions had to be made regarding people lost to follow-up (see Dealing with missing data), we would have compared the findings of the primary outcomes when we used our assumption compared with completer data only. If there was a substantial difference, we would have reported results and discuss them, but continued to employ our assumption.

Where assumptions had to be made regarding missing SD data (see Dealing with missing data), we would have compared the findings of primary outcomes when we used our assumption compared with complete data only. We would have undertaken a sensitivity analysis to test how prone results were to change when completer data only were compared to the imputed data using the above assumption. If there was a substantial difference, we would have reported results and discussed them, but continued to employ our assumption.

3. Risk of bias

We analysed the effects of excluding trials that were judged at high risk of bias across one or more of the domains of randomisation (implied as randomised with no further details available), allocation concealment, blinding and outcome reporting for the meta-analysis of the primary outcome (see Assessment of risk of bias in included studies). If the exclusion of trials at high risk of bias did not substantially alter the direction of effect or the precision of the effect estimates, then data from these trials were included in the analysis (Figure 2).

4. Imputed values

We undertook a sensitivity analysis to assess the effects of including data from trials where we used imputed values for ICC in calculating the design effect in cluster-randomised trials.

If there were substantial differences in the direction or precision of effect estimates in any of the sensitivity analyses, we did not pool data from the excluded trials with the other trials contributing to the outcome, but presented them separately.

5. Fixed and random effects

We synthesised data using a fixed-effect model; however, we also synthesised data for the primary outcomes using random-effect model to evaluate whether this altered the significance of the results.

RESULTS

Description of studies

The search identified three studies as eligible for inclusion in this review (ISRCTN65314790; Leff 2013; NCT03148639), and one ongoing study (NCT03585127). For more details, see Characteristics of included studies, Characteristics of excluded studies, and Characteristics of ongoing studies tables.

Results of the search

The search strategy was run on the February 2013, December 2016 and 9 November 2018, which found 14 possibly relevant references. We retrieved all of these for detailed evaluation. Following inspection of the abstracts and, when necessary, full papers, the 14 reports referred to four different studies. We did

Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

not exclude any study because all four of them met our inclusion criteria. However, one study was identified as an ongoing study (NCT03585127). See also Figure 1.

Included studies

The three included trials randomised 195 participants. We extracted data from the full-text publication of these trials, which were published on 21 February 2013 (Leff 2013), 23 November 2017 (ISRCTN65314790), and 18 February 2018 (NCT03148639). The characteristics of this trial are also summarised in the Characteristics of included studies table.

1. Design and duration

All studies were stated to be randomised. Study duration across the three included trials was identical; short-term and for seven weeks. Study design varied: two were partial crossover (Leff 2013; NCT03148639), and one was cross-sectional (ISRCTN65314790). Blinding for two of the trials was single blind (ISRCTN65314790; Leff 2013), and one had no blinding (NCT03148639). Researchers in the UK carried out two of the trials (ISRCTN65314790; Leff 2013), and one took place in Canada (NCT03148639).

2. Participants

All participants were randomly assigned to the treatment groups. In the ISRCTN65314790 study, participants had a *primary* diagnosis of schizophrenia spectrum (International Classification of Diseases 10th Edition – Schizophrenia, schizotypal, delusional, and other non-mood psychotic disorders; ICD10 F20–29) or affective disorder (F30–39) with psychotic symptoms, paranoid schizophrenia, schizoaffective disorder, bipolar disorder, unspecified non-organic psychosis, schizophrenia unspecified and depression with psychotic symptoms. These people also had a history of enduring auditory verbal hallucinations during previous 12 months despite continued treatment. The participants in Leff 2013 had been diagnosed by hearing persecutory voices for at least six months, which had not responded adequately to antipsychotic medication with history of hearing voices for more than one year. In NCT03148639 participants were diagnosed with schizophrenia or schizoaffective disorder with history of treatment-resistant schizophrenia.

The range of age of participants was 14 to 75 years with parental consent obtained for those under 18 years. All studies included males and females and with two studies clearly reporting male to female ratio (ISRCTN65314790; Leff 2013).

3. Size

The number of participants included in each study ranged between 19 and 150 but only one study included over 100 participants (ISRCTN65314790).

4. Setting

Leff 2013 stated that they were focused on inpatients in hospital at community mental health teams in Camden and Islington Mental Health Trust, UK. ISRCTN65314790 took place in south London and Maudsley NHS Trust, UK at sites remote from the clinic, while NCT03148639 took place in the Institut Universitaire en Santé Mentale de Montréal, Canada, and in the community.

5. Interventions

5.1. Avatar Therapy

ISRCTN65314790 reported that during the Avatar Therapy, participants created a computerised representation of the entity that they believed was the source of their main voice and then the team set up the avatar in an introductory session, which included a comprehensive assessment of the voice(s) and included verbatim content. About 10 to 15 minutes of each session involved face-to-face work with the avatar, wherein the therapist facilitated a direct dialogue between the participant and the avatar. Participants sat in one room facing their avatar on a computer monitor. The therapist sat in a second room with a control panel that allowed them to speak in his or her own voice, or as the avatar.

Leff 2013 and ISRCTN65314790 reported that the participants created a representation of the entity they imagines as the source of their avatar using computerised face animation software and in the third study (NCT03148639), they reported that the participants created an avatar best resembling the most distressing person or entity believed to be the source of the malevolent voice, which was designed to closely have both the face and the voice of the 'persecutor'. Over the course of the Avatar Therapy, the avatar's interaction with the participant became gradually less abusive and more supportive.

5.2. Treatment as usual/supportive counselling

Two trials compared Avatar Therapy with treatment as usual (Leff 2013; NCT03148639). In Leff 2013, treatment as usual was the patient's ongoing antipsychotic medication prescribed and supervised by their referring psychiatrist, and in NCT03148639, they offered antipsychotic treatment and usual meetings with their treating clinicians. ISRCTN65314790 used supportive counselling as a control group. In this study, the intervention comprised a manual-based, face-to-face supportive counselling approach adapted with permission from that employed by the SoCRATES (Acute Stroke or Transient Ischemic Attack Treated With Aspirin or Ticagrelor and Patient Outcomes) Trial Group delivered by graduate assistant psychologists who were recruited based on extensive experience of working therapeutically in a psychosis context.

6. Outcomes

6.1. General

The trials reported mental state, insight, global state, leaving the study early, adverse events and quality of life. None of the included studies reported death, relapse or direct economic evaluation of Avatar Therapy. Most reported outcomes were continuous and many were skewed.

6.2. Outcome scales used in included studies

6.2.1. Mental state

- **Calgary Depression Scale –CDS** (Addington 1993)

The CDS is a nine-question scale with a 4-point rating (0 to 3), with a score of zero indicating the absence of depression and a score of 3 indicating severe depression.

- **Depression, Anxiety and Stress Scale –21 Items (DASS-21)** (Henry 2005)

Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

DASS-21 is a set of three self-report scales designed to measure the emotional states of depression, anxiety and stress. Each of the three DASS-21 scales contains seven items, divided into subscales with similar content.

- **Scale for the Assessment of Positive Symptoms (SAPS)** (Andreasen 1984)

SAPS is a 34-item scale where each item is being rated between 0 and 5 by clinicians to measure positive symptoms in schizophrenia across four positive symptom domains: delusions, positive formal thought disorder or disconnected thinking, bizarre behaviour and hallucinations. The higher score is an indicator of more severe positive symptoms.

- **Positive and Negative Syndrome Scale (PANSS)** (Kay 1986)

PANSS consists of 30 items and is rated using a seven-point severity scale from absent to severe. It measures positive symptoms, negative symptoms and an overall total score of general psychopathology. A higher score indicates more severe symptoms. Mean total score from this scale were reported by five studies, three studies collected separate data for both positive and negative symptoms (Ahmed 2015; Bryce 2018; Holzer 2014).

- **Psychotic Symptom Rating Scales (PSYRATS)** (Haddock 1999)

PSYRATS are semi-structured interviews designed to assess the subjective characteristics of hallucinations and delusions. This is used to capture the severity of several dimensions of auditory hallucinations and delusions across two subclass, an auditory hallucination subscale and a delusions subscale. The auditory hallucination subscale includes 11 items while the delusional subscale includes six items, with both subscales using a 5-point ordinal scale to rate symptoms. Scores for the auditory hallucination subscale can range from 0 to 44 and for the delusional subscale can range from 0 to 24, with higher scores indicating more severe symptoms. All items are scored 0 to 4, according to general criteria: 0 = no problem, 1 = minimal or occasional, 2 = minor to moderate, 3 = major and 4 = maximum severity.

- **Scale for the Assessment of Negative Symptoms** (**SANS**) (Andreasen 1989)

SANS was the first instrument developed to provide comprehensive assessment of negative symptoms in schizophrenia. It consists of five scales that evaluate five different aspects of negative symptoms: alogia, affective blunting, avolition-apathy, anhedonia-asociality and attentional impairment. Each of these negative symptoms can be rated globally, but, in addition, detailed observations are made in order to achieve the global rating. It is complemented by SAPS, which permits detailed evaluation and global ratings of hallucinations, delusions, positive formal thought disorder and bizarre behaviour. Together, the two scales provide a comprehensive set of rating scales that measure the symptoms of schizophrenia and to assess their change over time.

- **Perceived Stress Reactivity Scale (PSRS)** (Scholtz 2011)

PSRS is a 23-item questionnaire with five subscales and one overall scale, based on an existing German-language instrument. Perceived stress reactivity and related constructs were assessed in 2040 participants from the UK, the US and Germany. The five-factor structure of the PSRS was found to be similar in the three countries. In the US sample, the questionnaire was applied using two modes of administration (paper-pencil and computerised), and measures were repeated after four weeks. Measurement invariance analyses demonstrated full invariance across mode of administration and partial invariance across gender and countries. Scale scores differed between countries and genders, with women scoring higher on most scales.

6.2.2. Mental state: insight

- **Revised Beliefs about Voices Questionnaire (BAVQ-R)** (Chadwick 2000)

The BAVQ-R was designed to assess these constructs, together with two styles of responding (engagement and resistance). The BAVQ-R is widely used in clinical and research settings, yet it has not received validation of its constructs and factor structure. This is used to measure the beliefs people hold about auditory hallucinations, and both their emotional and behavioural reaction to them. The scale has 35-items across five subscales. The subscales relating to beliefs are a six-item malevolence subscale, a six-item benevolence subscale and a six-item omnipotence subscale. The two subscales relating to emotion and behavioural reactions are a five-item resistance subscale and a four-item engagement subscale. All responses are rated on a 4-point scale from 0 to 3 and the high score means extremely distressed.

6.2.3. Quality of life

- **Manchester Short Assessment of Quality of Life (MANSA)** (Priebe 1999)

The MANSA consists of three sections:

- personal details that are supposed to be consistent over time (date of birth, gender, ethnic origin, and diagnosis);
- personal details that may potentially vary over time and have to be re-documented if change has occurred (education; employment status including type of occupation and working hours per week; monthly income; state benefits; living situation including number of children, people the patient lives with, and type of residence);
- only 16 questions are to be asked every time the instrument is applied. Four of these questions are termed objective and to be answered with yes or no. Twelve questions are strictly subjective.

The subjective questions obtain satisfaction with life as a whole, job (or sheltered employment, or training/education, or unemployment/retirement), financial situation, number and quality of friendships, leisure activities, accommodation, personal safety, people whose the patient lives with (or living alone), sex life, relationship with family, physical health and mental health. Like in the Lancashire Quality of Life Profile, satisfaction is rated on 7-point rating scales (1 = negative extreme, 7 = positive extreme).

- **Rosenberg Self-Esteem Scale (RSES)** (Rosenberg 1965)

The RSES is widely used in social-science research. It uses a scale of 0 to 40 where a score less than 15 may indicate a problematic low self-esteem.

The RSES is designed similar to the social-survey questionnaires. It is a 10-item Likert-type scale with items answered on a 4-

Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

point scale – from strongly agree to strongly disagree. Five of the items have positively worded statements and five have negatively worded statements. The scale measures state self-esteem by asking the respondents to reflect on their current feelings. RRSES is considered a reliable and valid quantitative tool for self-esteem assessment.

Implicit measures of self-esteem began to be used in the 1980s. These rely on indirect measures of cognitive processing thought to be linked to implicit self-esteem, including the Name Letter Task. Such indirect measures are designed to reduce awareness of the process of assessment. When used to assess implicit self-esteem, psychologists feature self-relevant stimuli to the participant and then measure how quickly a person identifies positive or negative stimuli. For example, if a woman were given the self-relevant stimuli of female and mother, psychologists would measure how quickly she identified the negative word, evil, or the positive word, kind.

- **Quality of Life Enjoyment and Satisfaction Questionnaire – Short Form (QLESQ-SF)** (Endicott 1993)

The QLESQ-SF evaluates general activities that are assessed in the longer form of the QLESQ. Each item uses a 5-point scale ranging from 1 (very poor) to 5 (very good). A total score is derived from 14 items with a maximum score of 70 and with higher scores indicating greater life satisfaction and enjoyment. Participants rate their satisfaction with the following domains of activity: physical health, feelings, work, household duties, school/course work, leisure time activities and social relations. Test–retest reliability for the QLESQ-SF has been shown to be.86 (Rapaport 2005) and internal consistency (Cronbach's alpha) has been shown to range from 0.86 to 0.90 (Rapaport 2005; Wyrwich 2009).

7. Missing outcomes

Data were reported on the majority of primary outcomes. However, none of the trials addressed economic outcomes. All three studies reported their data as mean endpoint rather than change.

Several outcomes were in effect missing because of problematic reporting. ISRCTN65314790 did not separate data for each group for the Reactions to Research Participation Questionnaire (RRPQ) rendering data useless. NCT03148639 reported anxiety and fear data only within a simple graph from which numbers could not be extracted.

Finally, all the three studies reported skewed data which are difficult to analyses within systematic reviews.

8. Funding

One study was funded by the National Institute of Health Research (NIHR; RC-PG-0308-10232) and Camden & Islington NHS Foundation Trust (Leff 2013), and one of authors is part funded by the NIHR Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London (ISRCTN65314790). Two of the authors in NCT03148639 are holders of a grant from Otsuka Pharmaceuticals.

9. Conflicts of Interest

Two of the authors of ISRCTN65314790 declared patents pending on the avatar system and in Leff 2013 there also could be conflicts of interest given that the authors involved in the study were the same team that designed Avatar Therapy.

Excluded studies

There were no excluded studies.

Studies awaiting classification

There are no studies awaiting classification.

Ongoing studies

There is currently one ongoing study in this review (NCT03585127). This small study comparing Avatar Therapy with a more cognitive behavioural approach should be complete (due to finish December 2019) and we will include the results in the review update.

Risk of bias in included studies

Overall risk of bias in individual studies are also illustrated in Figure 2; Figure 3; and Characteristics of included studies table.

Allocation

We assessed two studies at low risk for random sequence generation. Leff 2013 randomised using a computer-generated series with 12 blocks by an independent statistician and in ISRCTN65314790, the participants were randomly assigned (1:1) to receive Avatar Therapy or supportive counselling with randomised permuted blocks. NCT03148639 was at high risk as they described participants randomly allocated (1:1 ratio) to either Avatar Therapy or treatment as usual but this ratio does not fit with the actual number of participants in the two groups.

For allocation concealment, ISRCTN65314790 was at low risk as an independent web-based service was provided to maintain allocation concealment. Leff 2013 was at unclear risk as, although participants were randomised into one of two groups by an independent statistician using a computer-generated series with blocks of 12. It is not clear, however, if all of the team were informed of the allocated group or care co-ordinators/case managers, blinded to other allocation decisions, informed the study's participants. We assessed NCT03148639 at unclear risk as they did not provide any clear and detailed information about allocation concealment.

Blinding

We ranked ISRCTN65314790 at unclear risk for both performance and detection bias as the study was blind only for assessors. Leff 2013 had an unclear risk of performance and detection bias as it was not clear whether there was any blinding. NCT03148639 was also at unclear risk as they did not report any information on blinding.

Incomplete outcome data

For incomplete outcome reporting we had to rank ISRCTN65314790 at high risk, as there were no clear data about the adverse events and unusable information about satisfaction (RRPQ). We ranked Leff 2013 at high risk, as there was attrition from the intervention group; however, no further analysis was undertaken to account for these missing data. Also there was difference between protocol and full publication of the trial about the duration of the trials – changing from 11 months to six weeks. We ranked NCT03148639 at high risk, as they reported postcross-over data – amalgamating the Avatar Therapy plus treatment as usual rather than the outcomes for Avatar Therapy and treatment as usual separately. We contacted the authors via email for more information. They supplied data on

Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

loss to follow-up, rehospitalisation and the mean outcome data for Avatar Therapy group before crossover.

Selective reporting

None of the studies were free from selective reporting. ISRCTN65314790 was at unclear risk, as they frequently used "significantly greater improvement" but did not describe how this was measured, they also reported confusing information in their CONSORT flow diagram. Leff 2013 was at high risk for reporting bias as some of the important outcomes such as adverse events were not clearly reported and a protocol was not published prior to the study results being reported. NCT03148639 was at high risk, as there was no clear information in order to allocation concealment and blinding, adverse events and number of participants.

Other potential sources of bias

None of the studies were free from other biases. We had concerns with ISRCTN65314790, as two of authors declared patents pending on the avatar system used in the study. We also assessed Leff 2013 at high risk of other bias, as the study was funded by the NIHR and Bridging Funding from Camden & Islington NHS Foundation Trust but the authors involved in the study were the same team that had designed Avatar Therapy. In addition, information in the published trial was different to the submitted abstract. NCT03148639 was at unclear risk as two of the study authors were holders of a grant from Otsuka Pharmaceuticals.

Effects of interventions

See: **Summary of findings 1** Avatar Therapy compared to treatment as usual (all short-term) for schizophrenia or related disorders; **Summary of findings 2** Avatar Therapy compared to supportive counselling (all short-term) for schizophrenia or related disorders

We identified three randomised trials from which it was possible to extract numerical data. We categorised studies into two comparisons: Avatar Therapy versus treatment as usual and Avatar Therapy versus supportive counselling.

1. Comparison 1: Avatar Therapy versus treatment as usual

1.1. Mental state: 1a. Specific – positive – average endpoint score (PANSS-P, high = poor) – short term

Analysis of reported data showed no clear difference between treatment groups for this outcome (MD –1.93, 95% CI –5.10 to 1.24; studies = 1; participants = 19; very-low certainty evidence; Analysis 1.1).

1.2. Mental state: 1b. Specific – insight – average endpoint score (BAVQ-R attitude to voices, high = poor) – short term

Analysis of insight data showed participants in the Avatar Therapy group had clearly higher average BAVQ-R endpoint scores (MD –5.97, 95% CI –10.98 to –0.96; studies = 1; participants = 19; very-low certainty evidence; Analysis 1.2).

1.3. Mental state: 1c. Specific – insight – average endpoint attitude to voice score (BAVQ-R attitude to voices, high = poor, skewed data) – short term

Data reported for this outcome were skewed and we have presented them as 'other data' (Analysis 1.3).

1.4. Mental state: 1d. Specific – depression – average endpoint score (various scales, high = poor, skewed data) – short term

Data reported for this outcome were skewed and we have presented them as 'other data' (Analysis 1.4).

1.5. Mental state: 1e. Specific – negative – average endpoint score (PANSS-N, high = poor) – short term

Analysis of reported data showed no clear difference between treatment groups for this outcome (MD –1.17, 95% CI –5.70 to 3.36; studies = 1; participants = 19; Analysis 1.5).

1.6. Mental state: 2a. General – total – average endpoint score (PANSS total, high = poor) – short term

Analysis of reported data showed no clear difference between treatment groups for this outcome (MD –8.07, 95% CI –19.49 to 3.35; studies = 1; participants = 19; Analysis 1.6).

1.7. Mental state: 2b. General – total – average endpoint score (PSYRATS total, high = poor) – short term

Analysis of general mental state data showed participants in the Avatar Therapy group had clearly lower average PSYRATS endpoint scores(MD –5.51, 95% CI –9.15 to –1.87; studies = 1; participants = 19; very-low certainty evidence; Analysis 1.7).

1.8. Mental state: 2c. General – total – average endpoint score (PSRS, high = poor, skewed data) – short term

Data reported for this outcome were skewed and we have presented them as 'other data' (Analysis 1.8).

1.9. Global state: 1. Any change – needing counselling and support – short term

Analysis of reported data showed no clear difference between treatment groups for this outcome (RR 1.85, 95% CI 0.09 to 40.05; studies = 1; participants = 19; Analysis 1.9).

1.10. Global state: 2. Rehospitalised – short term

Analysis of reported data showed no clear difference between treatment groups for this outcome (RD 0.00, 95% CI –0.20 to 0.20; studies = 1; participants = 19; low-certainty evidence; Analysis 1.10).

1.11. Leaving study early: 1. For any reason – short term

Analysis of reported data showed no clear difference between treatment groups for this outcome (RR 11.27, 95% CI 0.70 to 181.41; studies = 1; participants = 26; low-certainty evidence; Analysis 1.11).

1.12. Leaving study early: 2. For specific reason – short term

One trial reported specific reasons for attrition (Analysis 1.12).

1.12.1. Change in medication

Analysis of reported data showed no clear difference between treatment groups for this outcome (RR 2.60, 95% CI 0.12 to 58.48; studies = 1; participants = 26).

1.12.2. Fear of voice

Analysis of reported data showed no clear difference between treatment groups for this outcome (RR 9.53, 95% CI 0.58 to 156.49; studies = 1; participants = 26).

Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

1.13. Adverse events: anxiety – short term

One trial reported incidence of anxiety (Analysis 1.13).

1.13.1. After first session

Analysis of reported data showed no clear difference between treatment groups for this outcome (RR 5.54, 95% CI 0.34 to 89.80; studies = 1; participants = 19; low-certainty evidence).

1.13.2. Self-reported

No participants reported self-reported incidence of anxiety.

1.14. Quality of life: average endpoint score (QLESQ-SF, high = good) – short term

Analyses of reported quality of life data showed participants in the Avatar Therapy group had clearly higher average QLWSQ-SF endpoint scores (MD 9.99, 95% CI 3.89 to 16.09; studies = 1; participants = 19; very low-certainty evidence; Analysis 1.14).

2. Comparison 2: Avatar Therapy versus supportive counselling

2.1. Mental state: 1a. Specific – positive – average endpoint score (SAPS, high = good, skewed data)

Data reported for this outcome were skewed and we have presented them as 'other data' (Analysis 2.1).

2.2. Mental state: 1b. Specific – insight – average endpoint score (various scales, high = poor) – short term

Various measures of insight were reported, using various scales (Analysis 2.2).

2.2.1. Attitude to voices (BAVQ-R)

Analysis of insight data showed participants in the Avatar Therapy group had clearly lower average BAVQ-R 'attitude to voice' endpoint scores (MD –8.39, 95% CI –14.31 to –2.47; studies = 1; participants = 124; low-certainty evidence).

2.2.2. Acceptance of voice (Voices Acceptance and Action Scale; VAAS)

Analysis of insight data showed participants in the Avatar Therapy group had clearly higher average VAAS 'acceptance of voice' endpoint scores (MD 4.73, 95% CI 1.40 to 8.06; studies = 1; participants = 124).

2.2.3. Power of voice (Voice Power Differential Scale; VPDS)

Analysis of reported data showed no clear difference between treatment groups for this outcome (MD –0.36, 95% CI –0.89 to 0.17; studies = 1; participants = 115).

2.2.4. Total (VAAS)

Analysis of insight data showed participants in the Avatar Therapy group had clearly lower average VAAS total endpoint scores (MD –2.91, 95% CI –5.71 to –0.11; studies = 1; participants = 115).

2.3. Mental state: 1c. Specific – depression and anxiety – average endpoint score (high = poor, skewed data) – short term

Data reported for this outcome were skewed and we have presented them as 'other data' (Analysis 2.3).

2.4. Mental state: 1d. Specific – negative – average endpoint score (SANS, high = good, skewed data) – short term

Data reported for this outcome were skewed and we have presented them as 'other data' (Analysis 2.4).

2.5. Mental state: 2a. General – total – average endpoint score (PSYRATS, high score = poor) – short term

Analysis of reported general mental state data showed participants in the Avatar Therapy group had clearly lower average PSYRATS endpoint scores (MD –4.74, 95% CI –8.01 to –1.47; studies = 1; participants = 124; low-certainty evidence; Analysis 2.5).

2.6. Leaving study early: 1a. For any reason

Analysis of reported data showed no clear difference between treatment groups for this outcome (RR 1.06, 95% CI 0.59 to 1.89; studies = 1; participants = 150; moderate-certainty evidence; Analysis 2.6).

2.7. Leaving study early: 1b. For specific reason – short term

One trial reported specific reasons for leaving the study early (Analysis 2.7).

2.7.1. Not contactable

Analysis of reported data showed no clear difference between treatment groups for this outcome (RR 3.00, 95% CI 0.63 to 14.39; studies = 1; participants = 150).

2.7.2. Refused

Analysis of reported data showed no clear difference between treatment groups for this outcome (RR 0.83, 95% CI 0.38 to 1.81; studies = 1; participants = 150).

2.7.3. Unwell

Analysis of reported data showed no clear difference between treatment groups for this outcome (RR 0.67, 95% CI 0.11 to 3.88; studies = 1; participants = 150).

2.7.4. Attended no sessions

Analysis of reported data showed no clear difference between treatment groups for this outcome (RR 0.33, 95% CI 0.13 to 0.87; studies = 1; participants = 150).

2.7.5. Discontinued – session 1to 6

Analysis of reported data showed no clear difference between treatment groups for this outcome (RR 1.70, 95% CI 0.83 to 3.47; studies = 1; participants = 150).

2.8. Quality of life: average endpoint score (various scales, high = good) – short term

Quality of life data were reported using various scales (Analysis 2.8).

2.8.1. MANSA

Analysis of reported data showed no clear difference between treatment groups for this outcome (MD 2.69, 95% CI –1.48 to 6.86; studies = 1; participants = 120; low-certainty evidence).

Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

2.8.2. Rosenberg Self-Esteem Scale

Analysis of reported data showed no clear difference between treatment groups for this outcome (MD 0.55, 95% CI –1.55 to 2.65; studies = 1; participants = 121).

DISCUSSION

Summary of main results

1. Avatar Therapy compared to treatment as usual (all short-term) for schizophrenia or related disorders

Two trials compared Avatar Therapy to treatment as usual (Leff 2013; NCT03148639).

1.1. Mental state

It is possible that we, in our protocol, set the expectations too high for measures of mental state. We had hoped to find some binary outcomes such as 'improved to an important extent' and the small trial (participants = 19) only reported continuous measures. It is understandable that a pilot study takes fine-grain measures as outcomes to see if there is a suggestion on a shift on these, hopefully sensitive, measures. There was, however, no suggestion, even on these measures and we found no data for binary outcomes. The study did not make data available so that binary analyses could be undertaken, even if this would have been grossly underpowered. We would like more data sets to be made available so that combination of findings could be achieved in the future as new trials emerge. Currently our certainty about any mental state evidence is very low – but there are very few data to have any degree of certainty over.

1.2. Global state

The same applied to global state measures. There were very few data from one small short-term trial (participants = 19). However, relapse was reported – albeit in the negative. This 'more concrete' outcome is useful when it is impossible to blind outcome measuring.

1.3. Leaving the study early

Six of 14 people left the Avatar Therapy group, and none left the trial from the treatment as usual group – out of 12 people assigned to that group. This could be a function of the trial design – allowing exiting from the avatar group with more ease than the treatment as usual group – and probably this is the explanation. However, there is a worry that this could indicate that the Avatar Therapy approach was, somehow, unacceptable to people with schizophrenia. The evidence is of low certainty (Analysis 1.11), but this is one finding that should be closely monitored in future studies.

1.4. Adverse effects: at least one important adverse effect

1.4.1. Anxiety – after first session

There was no clear difference between Avatar Therapy and treatment as usual for number of participants with anxiety after their first session of treatment but we have low certainty for this finding (Analysis 1.13), and it does seem to concur with the 'leaving the study early' finding reported and discussed above. The therapy could increase anxiety. This, in itself, may not be problematic and could have a positive effect, but this is speculation at this point. We have less evidence of a positive effect than we have of a suggestion of the Avatar Therapy being unacceptable and causing anxiety.

1.5. Quality of life: any change in quality of life

1.5.1. Average endpoint score (QLESQ-SF, high score = good)

Average scores were not what we prestated in our protocol as a preferred outcome. We had high levels of uncertainty about any finding related to quality of life but the one small trial that reported on this found a possible small positive effect for Avatar Therapy (participants = 19; Analysis 1.14). This could be a random finding – a spurious or rogue result. It is impossible to know. Should it be a real effect, it will be possible to replicate in future, larger trials.

1.6. Functioning

Simple recording of functioning, on day-to-day tasks, is not difficult to add to trial design. We look forward to this being included in future studies.

2. Avatar Therapy compared to supportive counselling (all short-term) for schizophrenia or related disorders

One trial compared Avatar Therapy to supportive counselling (ISRCTN65314790).

2.1. Mental state

The trial (participants = 124) did not report binary data – or data in a way that could be made into binary. The nearest proxy measure for a general mental state outcome did, however, show benefit for those allocated to Avatar Therapy (low-certainty evidence; Analysis 2.5). It could be that the supportive counselling was detrimental and that the Avatar Therapy approach better only in contrast to this. However, if this is a real effect, this should be replicated, and, explained. We remain unclear what an average drop of around 5 points on the PSYRATS score would mean for a person with schizophrenia.

When it came to a more specific measure (insight) the finding was not dissimilar (low-certainty evidence; Analysis 2.2). Again, it does seem positive that there is a drop of around 8 points on the BAVQ-R – but we are unclear if this would really make any noticeable difference clinically.

2.2. Global state

The trial did not report this relatively simple outcome. This is a missed opportunity as such a finding would be of interest and of value to future updates as more data become available.

2.3. Leaving the study early

The study had a small proportion of participants leaving early (about 20%), over a long period of time, which was balanced between groups (moderate-certainty evidence; Analysis 2.6).

2.4. Adverse effects: at least one important adverse effect

The trial did not address adverse effects, which is concerning as, at the very least, studies should report them. It is known that supportive therapy can have adverse consequences (Buckley 2015), so not having a clear mechanism for recording these would not seem to be the best design of trial. A drug trial not attempting to record adverse effects would not be allowed to proceed. It is a difficult argument that the therapy trials should have different and lesser standards.

Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

2.5. Quality of life: any change in quality of life

We found very low-certainty evidence for quality of life (Analysis 2.8) – and there was no clear difference between groups. These are important outcomes that should be replicated in future trials.

2.6. Functioning

The trial did not address functioning, which greatly devalued this potentially important trial.

Overall completeness and applicability of evidence

1. Completeness

This systematic review included three studies (participants = 195). Two studies compared Avatar Therapy with treatment as usual (Leff 2013; NCT03148639), and one study compared Avatar Therapy with supportive counselling (ISRCTN65314790). Pooling of data from these studies was not possible. All three studies reported short-term outcomes (seven sessions in seven weeks). Therefore, we have no information about the medium- or long-term effects of the interventions used in these trials and the numerical data we do have are uncertain.

None of the studies reported information related to general functioning, cognitive functioning and economic outcomes (direct, indirect and cost effective). The data that are reported are not fully consistent across the small studies making synthesis problematic. Much of the data reported were skewed. This is not a criticism of these data as it is likely that many of the so-called continuous measures do produce data with considerable skew. The criticism is that these data are not publicly available. Reporting was incomplete.

2. Applicability

The three studies had a wide group of participants and included both men and women between 14 to 75 years old. All studies, however, were conducted in high-income countries, and centres of excellence so, currently, would be difficult to apply outside of that environment. Much more widespread, generalisable, real-world work would have to be done to support wide applicability.

Quality of the evidence

In these decades after CONSORT (www.consort-statement.org), no trial should be unclear in how randomisation was undertaken and then concealed. The three included studies were threatened by risks of inclusion of bias at this most fundamental of stages. With thoughtful design and good conduct and reporting, all future Avatar Therapy studies could have risk of bias minimised. However, at present, the certainty of the evidence is not good and all results have to be treated with considerable caution.

Potential biases in the review process

To avoid introducing our own bias to this review we strictly followed Cochrane methods for conducting reviews and have reported all available processes, methods and data transparently so that they can be checked if needed. We would welcome any comments or additional data to improve this review.

Agreements and disagreements with other studies or reviews

To the best of our knowledge, this is the only systematic review of Avatar Therapy.

AUTHORS' CONCLUSIONS

Implications for practice

1. For people with schizophrenia

Currently, there is no clear evidence for – or against – using Avatar Therapy as a treatment for people with serious mental illness. The intervention is still an experimental treatment. People with schizophrenia could help generate much more evidence by taking part in good evaluative studies but also make this contingent on all data from their trials being made publicly available.

2. For clinicians

This review is unable, at the moment, to provide adequate evidence to inform clinicians about the value of Avatar Therapy for people with schizophrenia. If this therapy is being considered, the person with schizophrenia should be informed of the experimental nature of the treatment and the details should be explained to them in a transparent way so that they could make an informed decision.

3. For policy makers

Currently, there is no clear evidence for encouraging Avatar Therapy as a treatment for people with schizophrenia. Furthermore, there are no data in terms of economic outcomes for using Avatar Therapy in a clinical setting. We believe that more high-quality and long-term studies should be undertaken in this area to explore the effects, safety and costs of this novel intervention.

Implications for research

1. General

Given that there is insufficient research to determine if Avatar Therapy is effective, there is a need for further research to establish its value – or lack of it – in this population. In this fast-moving area, it is likely that additional studies are being planned. All of them should report the standards required by CONSORT (www.consort-statement.org). In addition, the need for studies completed by teams without a 'stake' in Avatar Therapy is absolutely necessary, as is the ensuring of the public release of all data relevant to such studies.

2. Specific

2.1. Trials

We know that the design of trials of any treatment is a painstaking process that takes time and great effort. However, we also have given this topic some thought and suggested an outline design (Table 1). We recognise that there is one ongoing study (NCT03585127), but this, too, is small (participants = 50) and is unlikely to definitively answer clinically meaningful questions. The current available data were extracted from three small short-term studies with fairly high risks of bias. Due to the nature of Avatar Therapy, blinding participants is problematic but the use of suitable outcomes could partly compensate for this. There is a real difficulty that researchers with enthusiasm for a therapy are the ones to undertake the trials on their own invention – and are likely to

Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

continue to be. Every effort should be made – and be seen to be made – to ensure distance of the pioneering inventors of Avatar Therapy from its evaluation without the loss of the necessary expertise.

ACKNOWLEDGEMENTS

Thanks to Clive Adams and Claire Irving at Cochrane Schizophrenia Group's editorial team in the Nottingham University for their support in completing this review. Thanks also to Cochrane Schizophrenia Group Editorial Base in Nottingham that produces and maintains standard text for use in the Methods section of their reviews.

We would like to thank Mahdi Moazzen for preparing and completing the protocol.

We would like to thank Alice Waugh for peer reviewing the protocol and Cochrane Copy Edit Support for copy editing.

We would like to thank Laura Dellazizzo who kindly supplied data on lost to follow-up, rehospitalisation and Avatar Therapy group before crossover (NCT03148639).

Parts of this review were generated using RevMan HAL v 4.2. More information about RevMan HAL is available at schizophrenia.cochrane.org/revman-hal-v4,

Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

REFERENCES

References to studies included in this review

ISRCTN65314790 *{published data only}***Std21739ISRCTN65314790**

Craig T, Garety P, Ward T, Rus-Calafell M, Williams G, Huckvale M, et al. Computer assisted therapy for auditory hallucinations: the avatar clinical trial. *Schizophrenia Research* 2014;**153**:S74. [CSZG: Ref29425]

Craig TK, Rus-Calafell M, Ward T, Fornells-Ambrojo M, McCrone P, Emsley R, et al. The effects of an Audio Visual Assisted Therapy Aid for Refractory auditory hallucinations (AVATAR therapy): study protocol for a randomised controlled trial. *Trials* 2015;**16**(1):349. [CSZG: Ref33036]

Craig TK, Rus-Calafell M, Ward T, Leff JP, Huckvale M, Howarth E, et al. AVATAR therapy for auditory verbal hallucinations in people with psychosis: a single-blind, randomised controlled trial. *Lancet Psychiatry* 2018;**5**(1):31-40. [CSZG: Ref37759]

Hall J, Rus-Calafell M, Omari-Asor L, Ward T, Emsley R, Garety P, et al. Assessing the subjective experience of participating in a clinical trial (AVATAR). *Psychiatry Research* 2018;**263**:82-7. [CSZG: Ref37998]

ISRCTN65314790. Reducing the frequency and severity of auditory hallucinations: the AVATAR clinical trial. isrctn.org/ISRCTN65314790 (first received 15 March 2013). [CSZG: Ref28255]

Leff 2013 *{published data only}***Std22190**

Anonymous. Corrections. *British Journal of Psychiatry* 2013;**203**:313. [CSZG: Ref29824]

Craig T, Garety P, Ward T, Rus-Calafell M, Williams G, Huckvale M, et al. Computer assisted therapy for auditory hallucinations: the avatar clinical trial. *Schizophrenia Research* 2014;**153**:S74. [CSZG: Ref29425]

Leff J, Williams G, Huckvale M, Arbuthnot M, Leff AP. Avatar therapy for persecutory auditory hallucinations: what is it and how does it work? *Psychosis* 2014;**6**(2):166-76. [CSZG: Ref28967]

Leff J, Williams G, Huckvale MA, Arbuthnot M, Leff AP. Computer-assisted therapy for medication-resistant auditory hallucinations: proof-of-concept study. *British Journal of Psychiatry* 2013;**202**:428-33. [CSZG: Ref28706] [DOI: 10.1192/bjp.bp.112.124883] [PMID: 23429202]

Ward T, Jamieson-Craig T, Garety PA, Rus-Calafell M, McCrone P, Emsley R, et al. A randomized clinical trial of a novel audio-visual assisted therapy aid for refractory auditory hallucinations (AVATAR therapy). *Early Intervention in Psychiatry* 2014;**8**:42. [CSZG: Ref29709]

NCT03148639 *{published data only}***Std29448**

NCT03148639. Virtual reality therapy for treatment-resistant auditory hallucinations in schizophrenia. clinicaltrials.gov/show/NCT03148639 (first received 11 May 2017). [CSZG: Ref35873]

du Sert OP, Potvin S, Lipp O, Dellazizzo L, Laurelli M, Breton R, et al. Virtual reality therapy for refractory auditory verbal hallucinations in schizophrenia: a pilot clinical trial. *Schizophrenia Research* 2018;**197**:176-81. [CSZG: Ref37988]

References to ongoing studies

NCT03585127 *{published data only}*

Dellazizzo L, Potvin S, O'Connor K, Dumais A. A randomized controlled trial comparing virtual reality therapy to cognitive behavioral therapy in schizophrenia with treatment refractory hallucinations: preliminary results. *Schizophrenia Bulletin* 2018;**44**(Suppl 1):S346-7.

NCT03585127. Avatar therapy in comparison to cognitive behavioral therapy for treatment-resistant schizophrenia. clinicaltrials.gov/show/NCT03585127 (first received 12 July 2018).

Additional references

Addington 1993

Addington J, Maticka-Tyndale E, Addington D. Assessing depression in schizophrenia: the Calgary Depression Scale. *British Journal of Psychiatry* 1993;**163**(s22):s39-s44.

Altman 1996

Altman DG, Bland JM. Detecting skewness from summary information. *BMJ* 1996;**313**(7066):1200.

Andreasen 1984

Andreasen NC. Scale for the Assessment of Positive Symptoms (SAPS). Iowa City (IA): University of Iowa, 1984.

Andreasen 1989

Andreasen NC. The Scale for the Assessment of Negative Symptoms (SANS): conceptual and theoretical foundations. *British Journal of Psychiatry* 1989;**155**(S7):49-52.

Bland 1997

Bland JM. Statistics notes. Trials randomised in clusters. *BMJ* 1997;**315**:600.

Boissel 1999

Boissel JP, Cucherat M, Li W, Chatellier G, Gueyffier F, Buyse M, et al. The problem of therapeutic efficacy indices. 3. Comparison of the indices and their use [Apercu sur la problematique des indices d'efficacite therapeutique, 3: comparaison des indices et utilisation. Groupe d'Etude des Indices D'efficacite]. *Therapie* 1999;**54**(4):405-11. [PMID: 10667106]

Buckley 2015

Buckley LA, Maayan N, Soares-Weiser K, Adams CE. Supportive therapy for schizophrenia. *Cochrane Database of Systematic Reviews* 2015, Issue 4. [DOI: 10.1002/14651858.CD004716.pub4] [CD004716] [PMID: 25871462]

Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Carpenter 1994

Carpenter WT Jr, Buchanan RW. Schizophrenia. *New England Journal of Medicine* 1994;**330**(10):681-90. [PMID: 8107719]

Chadwick 2000

Chadwick P, Lees S, Birchwood M. The Revised Beliefs About Voices Questionnaire (BAVQ-R). *Birtish Journal of Psychiatry* 2000;**177**(3):229-32. [DOI: 10.1192/bjp.177.3.229]

Comer 2017

Comer RJ. Chapter 15. Treatments for schizophrenia and other severe mental disorders. In: Abnormal Psychology. 8th edition. Moore Park (CA): Content Technologies, 2017:207-18.

Deeks 2000

Deeks J. Issues in the selection for meta-analyses of binary data. In: Proceedings of the 8th International Cochrane Colloquium; 2000 Oct 25-28; Cape Town. Cape Town: The Cochrane Collaboration, 2000.

Divine 1992

Divine GW, Brown JT, Frazier LM. The unit of analysis error in studies about physicians' patient care behavior. *Journal of General Internal Medicine* 1992;**7**(6):623-9.

Donner 2002

Donner A, Klar N. Issues in the meta-analysis of cluster randomized trials. *Statistics in Medicine* 2002;**21**(19):2971-80.

Egger 1997

Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997;**315**(7109):629-34.

Elbourne 2002

Elbourne D, Altman DG, Higgins JP, Curtin F, Worthington HV, Vail A. Meta-analyses involving cross-over trials: methodological issues. *International Journal of Epidemiology* 2002;**31**(1):140-9.

Endicott 1993

Endicott J, Nee J, Harrison W, Blumenthal R. Quality of Life Enjoyment and Satisfaction Questionnaire: a new measure. *Psychopharmacology Bulletin* 1993;**29**(2):321-6. [PMID: 8290681]

Furukawa 2006

Furukawa TA, Barbui C, Cipriani A, Brambilla P, Watanabe N. Imputing missing standard deviations in meta-analyses can provide accurate results. *Journal of Clinical Epidemiology* 2006;**59**(7):7-10.

Gulliford 1999

Gulliford MC. Components of variance and intraclass correlations for the design of community-based surveys and intervention studies: data from the Health Survey for England 1994. *American Journal of Epidemiology* 1999;**149**(9):876-83.

Haddock 1999

Haddock G, McCarron T, Faragher E. Scales to measure dimensions of hallucinations and delusions: the Psychotic Symptom Rating Scales (PSYRATS). *Psychological Medicine* 1999;**29**:879-89.

Henry 2005

Henry JD, Crawford JR. The short-form version of the Depression Anxiety Stress Scales (DASS-21): construct validity and normative data in a large non-clinical sample. *British Journal of Clinical Psychology* 2005;**44**(Pt 2):227-39. [PMID: 16004657]

Higgins 2003

Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;**327**(7414):557-60.

Higgins 2011

Higgins JP, Green S, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from handbook.cochrane.org.

Huckvale 2013

Huckvale M, Leff J, Williams G. Avatar therapy: an audio-visual dialogue system for treating auditory hallucinations. 14th International Speech Communication Conference; 2013 Aug 25-29; Lyon (France).

Hutton 2009

Hutton JL. Number needed to treat and number needed to harm are not the best way to report and assess the results of randomised clinical trials. *British Journal of Haematology* 2009;**146**(1):27-30.

ISRCTN65314790 2013

ISRCTN65314790. Reducing the frequency and severity of auditory hallucinations: A randomised clinical trial of a novel Audio-Visual Assisted Therapy Aid for Refractory auditory hallucinations (AVATAR) compared to supportive counselling. www.isrctn.com/ISRCTN65314790 (first received 15 March 2013). [DOI: 10.1186/ISRCTN65314790]

Kane 1996

Kane JM. Treatment-resistant schizophrenic patients. *Journal of Clinical Psychiatry* 1996;**57 Suppl 9**:35-40. [PMID: 8823348]

Kay 1986

Kay SR, Opler LA, Fiszbein A. Positive and Negative Syndrome Scale (PANSS) Manual. North Tonawanda (NY): Multi-Health Systems, 1986.

Kip 2019

Kip H, Kelders SM, Weerink K, Kuiper A, Bruninghoff I, Bouman YH, et al. Identifying the added value of virtual reality for treatment in forensic mental health: a scenario-based, qualitative approach. *Frontiers in Psychology* 2019;**10**:406. [PMID: 30873093]

Leon 2006

Leon AC, Mallinckrodt CH, Chuang-Stein C, Archibald DG, Archer GE, Chartier K. Attrition in randomized controlled clinical trials: methodological issues in psychopharmacology. *Biological Psychiatry* 2006;**59**(11):1001-5. [PMID: 16905632]

Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Leucht 2005a

Leucht S, Kane JM, Kissling W, Hamann J, Etschel E, Engel R. Clinical implications of Brief Psychiatric Rating Scale scores. *British Journal of Psychiatry* 2005;**187**:366-71. [PMID: 16199797]

Leucht 2005b

Leucht S, Kane JM, Kissling W, Hamann J, Etschel E, Engel RR. What does the PANSS mean? *Schizophrenia Research* 2005;**79**(2-3):231-8. [PMID: 15982856]

Marshall 2000

Marshall M, Lockwood A, Bradley C, Adams C, Joy C, Fenton M. Unpublished rating scales: a major source of bias in randomised controlled trials of treatments for schizophrenia. *British Journal of Psychiatry* 2000;**176**:249-52.

McKenna 2014

McKenna PJ, Mortimer AM. Current and future treatment modalities in schizophrenia: novel antipsychotic drugs and cognitive therapy. *Expert Review of Neurotherapeutics* 2014;**14**(1):67-73. [PMID: 24308281]

Nayani 1996

Nayani TH, David AS. The auditory hallucination: a phenomenological survey. *Psychological Medicine* 1996;**26**(1):177-89. [PMID: 8643757]

Overall 1962

Overall JE, Gorham DR. The Brief Psychiatric Rating Scale. *Psychological Reports* 1962;**10**:799-812.

Priebe 1999

Priebe S, Huxley P, Knight S, Evans S. Application and results of the Manchester Short Assessment of Quality of Life (MANSA). *International Journal of Social Psychiatry* 1999;**45**(1):7-12. [PMID: 10443245]

Rapaport 2005

Rapaport MH, Clary C, Fayyad R, Endicott J. Quality-of-life impairment in depressive and anxiety disorders. *American Journal of Psychiatry* 2005;**162**(6):1171-8. [PMID: 15930066]

Rosenberg 1965

Rosenberg M. Society and the Adolescent Self-Image. Princeton (NJ): Princeton University Press, 1965.

Rus-Calafell 2014

Rus-Calafell M, Jamieson-Craig T, Garety PA, Ward T, McCrone P, Emsley R, et al. Reducing the frequency and severity of voices: the Avatar clinical trial. *European Psychiatry* 2014;**29**(Suppl 1):EPA-0259. [DOI: 10.1016/S0924-9338(14)77707-7]

Saha 2005

Saha S, Chant D, Welham J, McGrath J. A systematic review of the prevalence of schizophrenia. *PLoS Medicine* 2005;**2**(5):e141. [DOI: 10.1371/journal.pmed.0020141] [PMID: 15916472]

Scholtz 2011

Schlotz W, Yim IS, Zoccola PM, Jansen L, Schulz P. The Perceived Stress Reactivity Scale: measurement invariance, stability, and validity in three countries. *Psychological Assessment* 2011;**23**(1):80-94.

Schünemann 2011

Schünemann HJ, Oxman AD, Vist GE, Higgins JP, Deeks JJ, Glasziou P, et al. Chapter 12: Interpreting results and drawing conclusions. In: Higgins JP, Green S, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from handbook.cochrane.org.

Shokraneh 2017

Shokraneh F, Adams CE. Study-based registers of randomized controlled trials: Starting a systematic review with data extraction or meta-analysis. *BioImpacts* 2017;**7**(4):209-17. [DOI: 10.15171/bi.2017.25]

Shokraneh 2019

Shokraneh F, Adams CE. Study-based registers reduce waste in systematic reviewing: discussion and case report. *Systematic Reviews* 2019;**8**:129. [DOI: 10.1186/s13643-019-1035-3]

Ukoumunne 1999

Ukoumunne OC, Gulliford MC, Chinn S, Sterne JA, Burney PG. Methods for evaluating area-wide and organisation-based intervention in health and health care: a systematic review. *Health Technology Assessment* 1999;**3**(5):1-75.

Wyrwich 2009

Wyrwich K, Harnam N, Revicki DA, Locklear JC, Svedsater H, Endicott J. Assessing health-related quality of life in generalized anxiety disorder using the Quality Of Life Enjoyment and Satisfaction Questionnaire. *International Clinical Psychopharmacology* 2009;**24**(6):289-95. [PMID: 19707153]

Xia 2009

Xia J, Adams CE, Bhagat N, Bhagat V, Bhoopathi P, El-Sayeh H, et al. Loss to outcomes stakeholder survey: the LOSS study. *Psychiatric Bulletin* 2009;**33**(7):254-7.

References to other published versions of this review

Moazzen 2015

Moazzen M, Shokraneh F. Avatar therapy for schizophrenia or related disorders. *Cochrane Database of Systematic Reviews* 2015, Issue 9. [DOI: 10.1002/14651858.CD011898]

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

ISRCTN65314790

Study characteristics	
Methods	Allocation: randomised Blindness: single Duration: over 7 sessions (1 introductory session + 6 therapy session) weekly 50 minutes each Design: cross-sectional Country: UK
Participants	Diagnosis: primary diagnosis of schizophrenia spectrum (ICD10 F20–29) or affective disorder (F30–39 with psychotic symptoms N = 150 Age: mean about 43 years, range 18–65 years Sex: men and women Setting: South London and Maudsley NHS Trust, sites remote from the clinic History: enduring auditory verbal hallucinations during previous 12 months despite continued treatment
Interventions	1. Avatar Therapy: participants created a computerised representation of the entity that they believed was the source of their main voice and then the team set up the avatar in an introductory session, which included a comprehensive assessment of the voice(s) and included verbatim content. 10–15 minutes of each session involved face-to-face work with the avatar, where in the therapist facilitated a direct dialogue between the participant and the avatar. Participants sat in 1 room facing their avatar on a computer monitor. The therapist sat in a second room with a control panel that allowed them to speak in his or her own voice, or as the avatar. N= 75. 2. Supportive counselling: a manual-based, face-to-face supportive counselling approach adapted with permission from that employed by the SoCRATES Trial Group delivered by graduate assistant psychologists who were recruited on the basis of extensive experience of working therapeutically in a psychosis context. N= 75.
Outcomes	Mental state – general: PSYRATS total Mental state – specific (insight): BAVQ-R total, VAAS acceptance, VPDS voice power, VPDS total Mental state – specific (depression and anxiety): CDS, DASS-21 total, Psychotic Symptom Rating Mental state – specific (negative symptoms): SANS Leaving study early: any reason, specific reasons Quality of life: MANSA, Rosenberg Self-Esteem Scale **Unable to use** Mental state – specific (positive): SAPS (skewed data) Physiological measures – not protocol outcomes Satisfaction: RRPQ (data not reported by group)
Notes	
Risk of bias	

Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

ISRCTN65314790 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "Participants were randomly assigned (1:1) to receive AVATAR therapy or supportive counselling with randomised permuted blocks (block size randomly varying between two and six". Comment: low risk
Allocation concealment (selection bias)	Low risk	Quote: "independent web-based service provided by the UKCRC [UK Clinical Research Collaboration] Registered Clinical Trials Unit at King's College London (registration number 053) to maintain allocation concealment". Comment: low risk
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Quote: "assessors" Comment: unclear risk
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Quote: "All assessments were done by research assessors who were masked to therapy allocation. To avoid unmasking, we ensured that assessors did not have access to clinical records after the baseline (pre-randomisation) assessment or access to the therapy database at any stage, that all assessments were done at sites remote from the clinic, and that participants were reminded before each assessment not to disclose their allocation." Comment: low risk
Incomplete outcome data (attrition bias) All outcomes	High risk	Quote: no clear information about the adverse events, RRPQ and Air. Comment: high risk
Selective reporting (reporting bias)	Unclear risk	All important outcomes have been reported but a few of the outcomes in registered protocol have not been reported in final report of the trials. Comment: unclear
Other bias	High risk	Quote: "one of authors is part funded by the National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. Two of the authors declared patents pending on the avatar system. The funding source had no input in the preparation of this manuscript". Quote: "significantly greater improvement" used frequently and confusing information in the CONSORT flow diagram. Comment: high risk

Leff 2013

Study characteristics	
Methods	Allocation: randomised Blindness: single Duration: 11 months of treatment stated in protocol but in the final study it stated 6 sessions of 30 minutes + 1 week' follow-up, because there was just 1 therapist, which limited the sessions[a]

Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Leff 2013 *(Continued)*

	Design: partial cross-over[b] Country: UK
Participants	Diagnosis: hearing persecutory voices for ≥ 6 months, which had not responded adequately to antipsychotic medication. N= 26 Age: 14–75 years, no more information[c] Sex: male and female Setting: hospital, community mental health teams in Camden and Islington Mental Health Trust, UK History: 22 participants had complete medical compliance, 2 partial and 2 none. 15 had ≥ 10 years of hearing voices, 3 had 6–10 years, and 8 had 1–5 years. Inclusion criteria: experienced persecutory auditory hallucinations for ≥ 6 months and these hallucinations had not responded adequately to antipsychotic medications. Excluded: organic brain disease and substance misuse
Interventions	1. Avatar Therapy: participants created a representation of the entity they imagined as the source of their auditory hallucinations using computerised face animation software. N= 14. 2. Treatment as usual (delayed therapy group): participant's ongoing antipsychotic medication prescribed and supervised by their referring psychiatrist. N= 12.
Outcomes	Mental state – specific (insight): BAVQ-R attitude to voice Mental state – specific (depression and anxiety): CDS Leaving the study early: for any reason, for specific reason Unable to use: Mental state – general: PSRS (skewed data)
Notes	It is unclear how the participants were approached for recruitment. Participants were recruited from the community mental health team; however it is unclear if all participants in the team were approached or care co-ordinators/case managers selectively informed participants who they thought were appropriate to enter the study. [a]There was a difference between protocol and final study report about the duration of study which has a change from 11 months to 6 weeks + 1 week' follow-up. [b]The immediate therapy group did not crossover into a no-therapy block as we expected the therapy to have carry-over effects. [c]Parental consent was obtained for participants aged under 18 years. The study was approved by the West Kent Research Ethics Committee.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "after baseline assessment, patients were randomised into one of two groups using a computer-generated series with blocks of 12, generated by an independent statistician". Comment: low risk

Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Leff 2013 *(Continued)*

Allocation concealment (selection bias)	Unclear risk	Quote: "patients were randomised into one of two groups using a computer-generated series with blocks of 12, generated by an independent statistician". Comment: it was unclear if all participants in the team were approached or care co-ordinators/case managers selectively informed participants who they thought were appropriate to enter the study. Comment: unclear risk
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Quote: "both groups of patients were assessed at each time point by a user-researcher who knew neither the group to which each patient was assigned, nor the design of the trial." Comment: participants were not blind to the intervention. The authors did not describe this. We rated unclear only because it appeared that Avatar Therapy could not be blinded; however, it may be possible to provide sham therapy. Comment: unclear
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Comment: blinding of assessor at follow-up time points was the only information available. Comment: unclear
Incomplete outcome data (attrition bias) All outcomes	High risk	Comment: no useful and clear information about attrition. Comment: some participants left early from the intervention group; however, no further analysis was undertaken to account for their missing data. Comment: there was a difference between protocol and article about the duration of the study from 11 months (protocol) to 6 weeks (actual length of study). Comment: high risk
Selective reporting (reporting bias)	High risk	Comment: no useful information on adverse events and attrition. Comment: some of the important outcomes such as adverse events were not clearly reported. Furthermore, protocol was not published prior to the study results being reported. Comment: high risk
Other bias	High risk	Quote: "the study was funded by the National Institute of Health Research (RC-PG-0308-10232) and Bridging Funding from Camden & Islington NHS Foundation Trust. Possible conflict of interest given that the authors involved in the study were the same team that designed avatar therapy." Comment: the article information was different from the abstract. Comment: very little information about the referral process, or demographic details such as age. Comment: high risk

NCT03148639

Study characteristics

Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

NCT03148639 *(Continued)*

Methods	Allocation: randomised Blindness: none Duration: 7-week phase II[a] Design: partial cross-over Country: Canada
Participants	Diagnoses: schizophrenia or schizoaffective disorder N= 19 Age: mean about 43 years Sex: male and female[b] Setting: Institut Universitaire en Santé Mentale de Montréal, Canada; institute and in community History: treatment-resistant schizophrenia, aged ≥ 18 years
Interventions	1. Avatar Therapy: participants created an avatar best resembling the most distressing person or entity believed to be the source of the malevolent voice, which was designed to closely have both the face and the voice of the "persecutor". Over the course of therapy, the avatar's interaction with the participant became gradually less abusive and more supportive. 1 avatar creation session + 6 weekly 45-minute further sessions. N= 12. 2. Treatment as usual: antipsychotic treatment and usual meetings with their treating clinicians. N= 7.
Outcomes	Mental state – general: PANSS, PSYRATS total Mental state – specific (depression and anxiety): BDI Mental state – specific (insight): BAVQ-R Mental state – specific (positive): PANSS Global state: rehospitalised Adverse events: various Quality of life: QLESQ-SF Unable to use: Mental state: self-reported anxiety, fear – unable to extract data from graph
Notes	Wrote to authors 24 November 2018 and Dr Laura Dellazizzo on behalf of Dr Alexandre Dumais, kindly replied supplying data on loss to follow-up, rehospitalisation and Avatar Therapy group before crossover. [a]The total amount of weeks was 7 comprising of 1 avatar creation session and 6 therapeutic sessions. [b]There is no clear and detailed information about it from the baseline.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	Quote: "patients fulfilling inclusion criteria were randomly allocated (1:1 ratio) to either VR-assisted therapy (VRT) or treatment-as-usual (TAU)".

Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

NCT03148639 *(Continued)*

		Comment: concern as the number of participants in 2 random groups are different which is not possible in 1:1 ratio.
Allocation concealment (selection bias)	Unclear risk	No information Comment: unclear
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	No information Comment: unclear
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	No information Comment: unclear
Incomplete outcome data (attrition bias) All outcomes	High risk	Comment: there was no information in the article because they had reported the crossover data (therapy +treatment as usual) as outcome report instead of reporting therapy outcome report separately. We contacted them via email for more information, and they kindly replied supplying data on loss to follow-up, rehospitalisation and the mean outcome data for avatar therapy group before crossover, but unfortunately, they did not share the change data for this outcome with us. Comment: high risk
Selective reporting (reporting bias)	High risk	Comment: there was no clear information regarding adverse events and number of participants. Comment: high risk
Other bias	Unclear risk	Quote: "two of authors are holder of a grant from Otsuka Pharmaceuticals." Comment: unclear

BAVQ-R: Revised Beliefs about Voices Questionnaire; BDI: Beck Depression Inventory; CDS: Calgary Depression Scale; CI: confidence interval; DASS-21: Depression, Anxiety and Stress Scale – 21 Items; ICD10 F20–29: International Classification of Diseases 10th Edition – Schizophrenia, schizotypal, delusional, and other non-mood psychotic disorders; MANSA: Manchester Short Assessment of Quality of Life; MD: mean difference; N: number of participants; PANSS-P: Positive and Negative Syndrome Scale – Positive; PSYRATS: Psychotic Symptom Rating Scale; QLESQ-SF: Quality of Life Enjoyment and Satisfaction Questionnaire-Short Form; RCT: randomised controlled trial; RD: risk difference; RR: risk ratio; RRPQ: Reactions to Research Participation Questionnaire; SANS: Scale for the Assessment of Negative Symptoms; SAPS: Scale for the Assessment of Positive Symptoms; SoCRATES: Acute Stroke or Transient Ischemic Attack Treated With Aspirin or Ticagrelor and Patient Outcomes; VAAS: Voices Acceptance and Action Scale; VPDS: Voice Power Differential Scale.

Characteristics of ongoing studies *[ordered by study ID]*

NCT03585127

Study name	Avatar Therapy in comparison to cognitive behavioral therapy for treatment-resistant schizophrenia
Methods	Allocation: randomised Blinding: none Duration: 9 weeks Design: parallel Setting: Institut Universitaire en Santé Mentale de Montréal, Montréal, Quebec, Canada

Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

NCT03585127 *(Continued)*

Participants	Diagnosis: people with a auditory hallucination, verbal and treatment-resistant schizophrenia, DSM-5 diagnosis of schizophrenia or schizoaffective disorder History: participants randomised from waiting list and referred to research staff N = 50 Age: > 18 years Sex: men and women Inclusion criteria: DSM-5 diagnosis of schizophrenia or schizoaffective disorder, aged ≥ 18 years Exclusion criteria: any change in medication within the past 2 months, substance-use disorder within last 12 months, neurological disorder or unstable and serious physical illness, ongoing psychotic episode, cognitive behavioural therapy for psychosis within the last 12 months
Interventions	1. Avatar Therapy: programme for 9 individual and weekly sessions of 1 hour. N= 25 2. Cognitive behavioural therapy: programme for 9 individual and weekly sessions of 1 hour. N= 25
Outcomes	Mental state (binary outcomes) Change in Revised Beliefs About Voices Questionnaire Change in Positive And Negative Syndrome Change in Beck Depression Inventory – II Change in Quality of Life Enjoyment and Satisfaction Questionnaire Change in Empowerment scale (Making Decisions) – Revised Short Form Change in Igroup Presence Questionnaire No information: Relapses (binary outcomes) Quality of life (binary outcomes) Costs: cost of services, cost of care Adverse events related to yoga (number and type of injuries) Service outcomes: days in hospital, time attending outpatient psychiatric clinic
Starting date	1 March 2017 Estimated completion date: December 2019
Contact information	alexandre.dumais@umontreal.ca
Notes	

DSM-5: Diagnostic and Statistical Manual of Mental Disorders, 5th Edition; N: number of participants.

DATA AND ANALYSES

Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Comparison 1. Avatar therapy versus treatment as usual

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1.1 Mental state: 1a. Specific – positive – average endpoint score (PANSS-P, high = poor) – short term	1	19	Mean Difference (IV, Fixed, 95% CI)	-1.93 [-5.10, 1.24]
1.2 Mental state: 1b. Specific – insight – average endpoint score (BAVQ-R attitude to voices, high = poor) – short term	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
1.3 Mental state: 1c. Specific – insight – average endpoint score (BAVQ-R attitude to voices, high = poor, skewed data) – short term	1		Other data	No numeric data
1.4 Mental state: 1d. Specific – depression – average endpoint score (various scales, high = poor, skewed data – short term	2		Other data	No numeric data
1.4.1 BDI	1		Other data	No numeric data
1.4.2 CDS	1		Other data	No numeric data
1.5 Mental state: 1e. Specific – negative – average endpoint score (PANSS-N, high = poor) – short term	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
1.6 Mental state: 2a. General – total – average endpoint score (PANSS total, high = poor) – short term)	1	19	Mean Difference (IV, Fixed, 95% CI)	-8.07 [-19.49, 3.35]
1.7 Mental state: 2b. General – total – average endpoint score (PSYRATS total, high = poor) – short term	1	19	Mean Difference (IV, Fixed, 95% CI)	-5.51 [-9.15, -1.87]
1.8 Mental state: 2c. General – total – average endpoint score (PSRS, high poor, skewed data) – short term	1		Other data	No numeric data
1.9 Global state: 1. Any change – needing counselling and support – short term	1	19	Risk Ratio (M-H, Fixed, 95% CI)	1.85 [0.09, 40.05]
1.10 Global state: 2. Relapse (rehospitalised) – short term	1	19	Risk Difference (M-H, Fixed, 95% CI)	0.00 [-0.20, 0.20]
1.11 Leaving study early: 1. For any reason – short term	1	26	Risk Ratio (M-H, Fixed, 95% CI)	11.27 [0.70, 181.41]
1.12 Leaving study early: 2. For specific reason – short term	1		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
1.12.1 Change in medication	1	26	Risk Ratio (M-H, Fixed, 95% CI)	2.60 [0.12, 58.48]
1.12.2 Fear of voice	1	26	Risk Ratio (M-H, Fixed, 95% CI)	9.53 [0.58, 156.49]

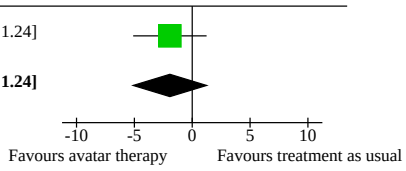
Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1.13 Adverse effects: anxiety – short term	1		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
1.13.1 After first session	1	19	Risk Ratio (M-H, Fixed, 95% CI)	5.54 [0.34, 89.80]
1.13.2 Self-reported	1	19	Risk Ratio (M-H, Fixed, 95% CI)	Not estimable
1.14 Quality of life: average endpoint score (QLESQ-SF, high = good) – short term	1	19	Mean Difference (IV, Fixed, 95% CI)	9.99 [3.89, 16.09]

Analysis 1.1. Comparison 1: Avatar therapy versus treatment as usual, Outcome 1: Mental state: 1a. Specific – positive – average endpoint score (PANSS-P, high = poor) – short term

Study or Subgroup	Avatar therapy Mean	SD	Total	Treatment as usual Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
NCT03148639	17.5	3.21	12	19.43	3.51	7	100.0%	-1.93 [-5.10 , 1.24]
Total (95% CI)			**12**			**7**	**100.0%**	**-1.93 [-5.10 , 1.24]**

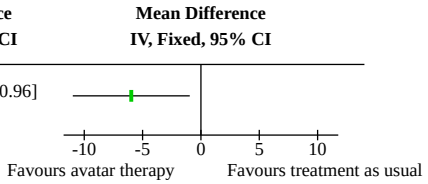
Heterogeneity: Not applicable
Test for overall effect: Z = 1.19 (P = 0.23)
Test for subgroup differences: Not applicable



Analysis 1.2. Comparison 1: Avatar therapy versus treatment as usual, Outcome 2: Mental state: 1b. Specific – insight – average endpoint score (BAVQ-R attitude to voices, high = poor) – short term

Study or Subgroup	Avatar therapy Mean	SD	Total	Treatment as usual Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
NCT03148639	16.5	5.32	12	22.47	5.41	7	-5.97 [-10.98 , -0.96]

Test for subgroup differences: Not applicable



Analysis 1.3. Comparison 1: Avatar therapy versus treatment as usual, Outcome 3: Mental state: 1c. Specific – insight – average endpoint score (BAVQ-R attitude to voices, high = poor, skewed data) – short term

Mental state: 1c. Specific – insight – average endpoint score (BAVQ-R attitude to voices, high = poor, skewed data) – short term

Study	Intervention	Mean	SD	n
Leff 2013	Avatar therapy	18.88	7.24	14
	Treatment as usual	12.37	12.61	12

Analysis 1.4. Comparison 1: Avatar therapy versus treatment as usual, Outcome 4: Mental state: 1d. Specific – depression – average endpoint score (various scales, high = poor, skewed data – short term

Mental state: 1d. Specific – depression – average endpoint score (various scales, high = poor, skewed data – short term

Study	Intervention	Mean	SD	n
BDI				

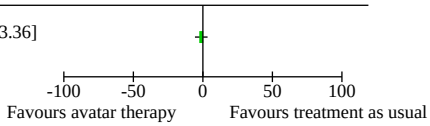
Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

NCT03148639	Avatar therapy	11.13	13.77	12
	Standard care	25.43	8.48	7
CDS				
Leff 2013	Avatar therapy	8.5	5.29	14
	Treatment as usual	4	1.41	12

Analysis 1.5. Comparison 1: Avatar therapy versus treatment as usual, Outcome 5: Mental state: 1e. Specific – negative – average endpoint score (PANSS-N, high = poor) – short term

Study or Subgroup	Avatar therapy Mean	SD	Total	Treatment as usual Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
NCT03148639	18.4	4.17	12	19.57	5.22	7	-1.17 [-5.70 , 3.36]

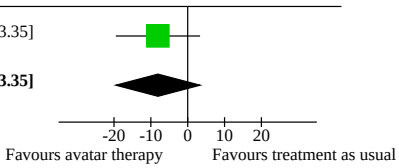
Test for subgroup differences: Not applicable



Analysis 1.6. Comparison 1: Avatar therapy versus treatment as usual, Outcome 6: Mental state: 2a. General – total – average endpoint score (PANSS total, high = poor) – short term)

Study or Subgroup	Avatar therapy Mean	SD	Total	Treatment as usual Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
NCT03148639	73.5	10.86	12	81.57	13	7	100.0%	-8.07 [-19.49 , 3.35]
Total (95% CI)			**12**			**7**	**100.0%**	**-8.07 [-19.49 , 3.35]**

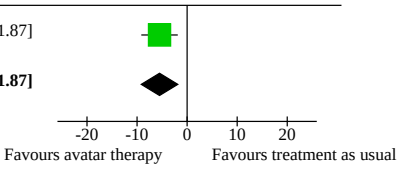
Heterogeneity: Not applicable
Test for overall effect: Z = 1.38 (P = 0.17)
Test for subgroup differences: Not applicable



Analysis 1.7. Comparison 1: Avatar therapy versus treatment as usual, Outcome 7: Mental state: 2b. General – total – average endpoint score (PSYRATS total, high = poor) – short term

Study or Subgroup	Avatar therapy Mean	SD	Total	Treatment as usual Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
NCT03148639	25.63	5.13	12	31.14	2.97	7	100.0%	-5.51 [-9.15 , -1.87]
Total (95% CI)			**12**			**7**	**100.0%**	**-5.51 [-9.15 , -1.87]**

Heterogeneity: Not applicable
Test for overall effect: Z = 2.97 (P = 0.003)
Test for subgroup differences: Not applicable



Analysis 1.8. Comparison 1: Avatar therapy versus treatment as usual, Outcome 8: Mental state: 2c. General – total – average endpoint score (PSRS, high poor, skewed data) – short term

Mental state: 2c. General – total – average endpoint score (PSRS, high poor, skewed data) – short term

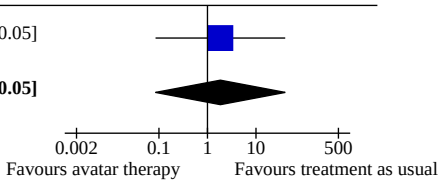
Study	Intervention	Mean	SD	n
Leff 2013	Avatar therapy	23.63	8.03	14
	Treatment as usual	20	13.1	12

Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 1.9. Comparison 1: Avatar therapy versus treatment as usual, Outcome 9: Global state: 1. Any change – needing counselling and support – short term

Study or Subgroup	Avatar therapy Events	Avatar therapy Total	Treatment as usual Events	Treatment as usual Total	Weight	Risk Ratio M-H, Fixed, 95% CI
NCT03148639	1	12	0	7	100.0%	1.85 [0.09 , 40.05]
Total (95% CI)		**12**		**7**	**100.0%**	**1.85 [0.09 , 40.05]**
Total events:	1		0			

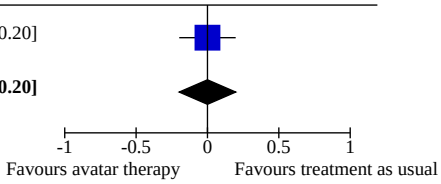
Heterogeneity: Not applicable
Test for overall effect: Z = 0.39 (P = 0.70)
Test for subgroup differences: Not applicable



Analysis 1.10. Comparison 1: Avatar therapy versus treatment as usual, Outcome 10: Global state: 2. Relapse (rehospitalised) – short term

Study or Subgroup	Avatar therapy Events	Avatar therapy Total	Treatment as usual Events	Treatment as usual Total	Weight	Risk Difference M-H, Fixed, 95% CI
NCT03148639	0	12	0	7	100.0%	0.00 [-0.20 , 0.20]
Total (95% CI)		**12**		**7**	**100.0%**	**0.00 [-0.20 , 0.20]**
Total events:	0		0			

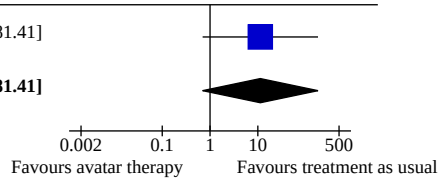
Heterogeneity: Not applicable
Test for overall effect: Z = 0.00 (P = 1.00)
Test for subgroup differences: Not applicable



Analysis 1.11. Comparison 1: Avatar therapy versus treatment as usual, Outcome 11: Leaving study early: 1. For any reason – short term

Study or Subgroup	Avatar therapy Events	Avatar therapy Total	Treatment as usual Events	Treatment as usual Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Leff 2013	6	14	0	12	100.0%	11.27 [0.70 , 181.41]
Total (95% CI)		**14**		**12**	**100.0%**	**11.27 [0.70 , 181.41]**
Total events:	6		0			

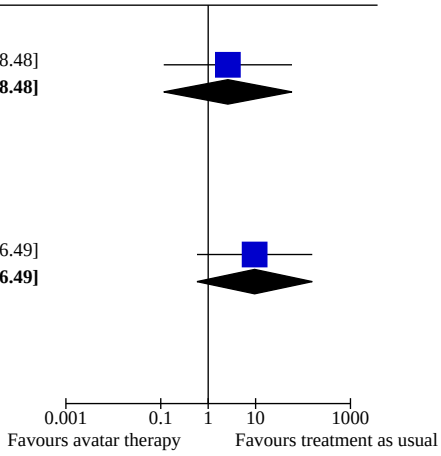
Heterogeneity: Not applicable
Test for overall effect: Z = 1.71 (P = 0.09)
Test for subgroup differences: Not applicable



Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

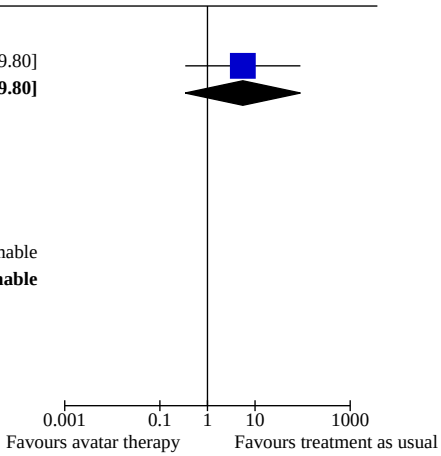
Analysis 1.12. Comparison 1: Avatar therapy versus treatment as usual, Outcome 12: Leaving study early: 2. For specific reason – short term

Study or Subgroup	Avatar therapy Events	Avatar therapy Total	Treatment as usual Events	Treatment as usual Total	Weight	Risk Ratio M-H, Fixed, 95% CI
1.12.1 Change in medication						
Leff 2013	1	14	0	12	100.0%	2.60 [0.12 , 58.48]
Subtotal (95% CI)		**14**		**12**	**100.0%**	**2.60 [0.12 , 58.48]**
Total events:	1		0			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.60 (P = 0.55)						
1.12.2 Fear of voice						
Leff 2013	5	14	0	12	100.0%	9.53 [0.58 , 156.49]
Subtotal (95% CI)		**14**		**12**	**100.0%**	**9.53 [0.58 , 156.49]**
Total events:	5		0			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.58 (P = 0.11)						
Test for subgroup differences: Chi² = 0.37, df = 1 (P = 0.54), I² = 0%						



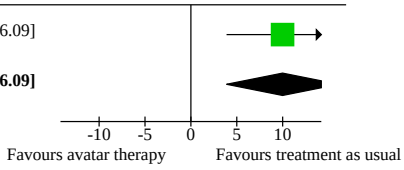
Analysis 1.13. Comparison 1: Avatar therapy versus treatment as usual, Outcome 13: Adverse effects: anxiety – short term

Study or Subgroup	Avatar therapy Events	Avatar therapy Total	Treatment as usual Events	Treatment as usual Total	Weight	Risk Ratio M-H, Fixed, 95% CI
1.13.1 After first session						
NCT03148639	4	12	0	7	100.0%	5.54 [0.34 , 89.80]
Subtotal (95% CI)		**12**		**7**	**100.0%**	**5.54 [0.34 , 89.80]**
Total events:	4		0			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.20 (P = 0.23)						
1.13.2 Self-reported						
NCT03148639	0	12	0	7		Not estimable
Subtotal (95% CI)		**12**		**7**		**Not estimable**
Total events:	0		0			
Heterogeneity: Not applicable						
Test for overall effect: Not applicable						
Test for subgroup differences: Not applicable						



Analysis 1.14. Comparison 1: Avatar therapy versus treatment as usual, Outcome 14: Quality of life: average endpoint score (QLESQ-SF, high = good) – short term

Study or Subgroup	Avatar therapy Mean	Avatar therapy SD	Avatar therapy Total	Treatment as usual Mean	Treatment as usual SD	Treatment as usual Total	Weight	Mean Difference IV, Fixed, 95% CI
NCT03148639	54.13	4.55	12	44.14	7.47	7	100.0%	9.99 [3.89 , 16.09]
Total (95% CI)			**12**			**7**	**100.0%**	**9.99 [3.89 , 16.09]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 3.21 (P = 0.001)								
Test for subgroup differences: Not applicable								



Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Comparison 2. Avatar therapy versus supportive counselling

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
2.1 Mental state: 1a. Specific – positive – average endpoint score (SAPS, high = poor, skewed data) – short term	1		Other data	No numeric data
2.2 Mental state: 1b. Specific – insight – average endpoint score (various scales, high = poor) – short term	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
2.2.1 Attitude to voices (BAVQ-R)	1	124	Mean Difference (IV, Fixed, 95% CI)	-8.39 [-14.31, -2.47]
2.2.2 Acceptance of voice (VAAS)	1	124	Mean Difference (IV, Fixed, 95% CI)	4.73 [1.40, 8.06]
2.2.3 Power of voice (VPDS)	1	115	Mean Difference (IV, Fixed, 95% CI)	-0.36 [-0.89, 0.17]
2.2.4 Total (VAAS)	1	115	Mean Difference (IV, Fixed, 95% CI)	-2.91 [-5.71, -0.11]
2.3 Mental state: 1c. Specific – depression – average endpoint score (various scales, high = poor, skewed data) – short term	1		Other data	No numeric data
2.3.1 CDS	1		Other data	No numeric data
2.3.2 DASS-21 total	1		Other data	No numeric data
2.4 Mental state: 1d. Specific – negative – average endpoint score (SANS, high = good, skewed data) – short term	1		Other data	No numeric data
2.5 Mental state: 2a. General – average endpoint score (PSYRATS total, high = poor) – short term	1	124	Mean Difference (IV, Fixed, 95% CI)	-4.74 [-8.01, -1.47]
2.6 Leaving study early: 1a. For any reason – short term	1	150	Risk Ratio (M-H, Fixed, 95% CI)	1.06 [0.59, 1.89]
2.7 Leaving study early: 1b. For specific reason – short term	1		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
2.7.1 Not contactable	1	150	Risk Ratio (M-H, Fixed, 95% CI)	3.00 [0.63, 14.39]
2.7.2 Refused	1	150	Risk Ratio (M-H, Fixed, 95% CI)	0.83 [0.38, 1.81]
2.7.3 Unwell	1	150	Risk Ratio (M-H, Fixed, 95% CI)	0.67 [0.11, 3.88]
2.7.4 Attended no sessions	1	150	Risk Ratio (M-H, Fixed, 95% CI)	0.33 [0.13, 0.87]

Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

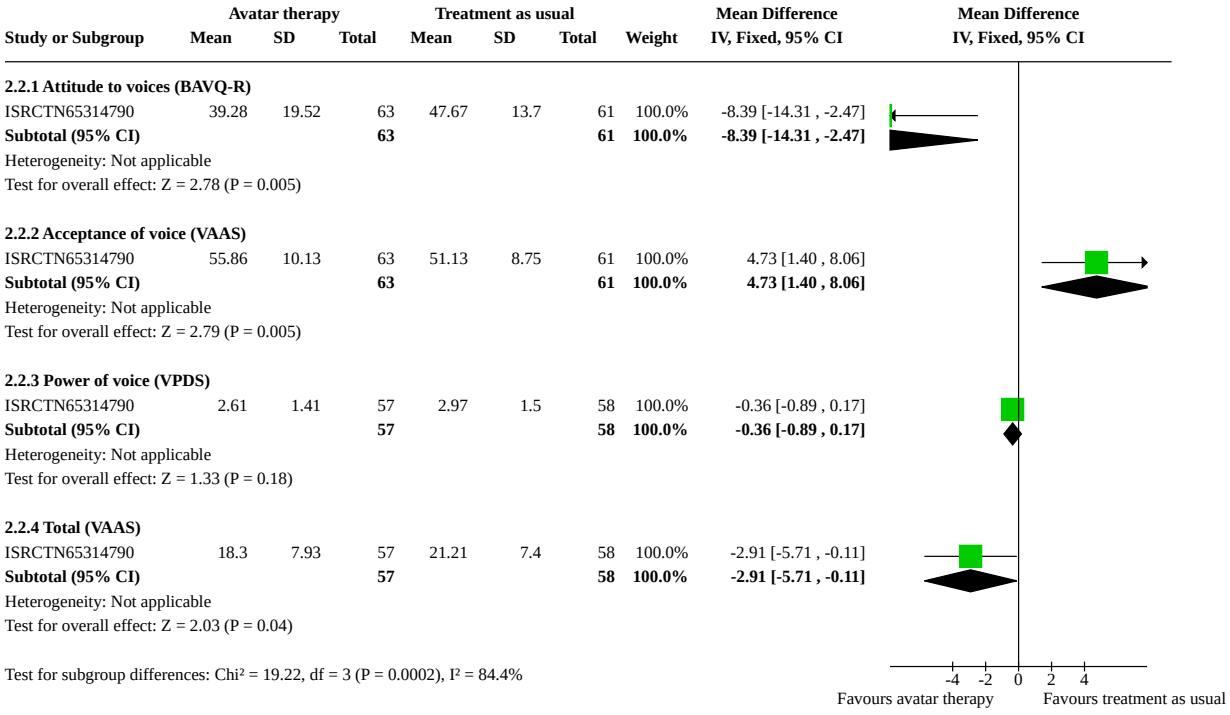
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
2.7.5 Discontinued during early sessions (1- 6)	1	150	Risk Ratio (M-H, Fixed, 95% CI)	1.70 [0.83, 3.47]
2.8 Quality of life: average endpoint score (various scales, high = good) – short term	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
2.8.1 MANSA	1	120	Mean Difference (IV, Fixed, 95% CI)	2.69 [-1.48, 6.86]
2.8.2 Rosenberg Self-Esteem Scale	1	121	Mean Difference (IV, Fixed, 95% CI)	0.55 [-1.55, 2.65]

Analysis 2.1. Comparison 2: Avatar therapy versus supportive counselling, Outcome 1: Mental state: 1a. Specific – positive – average endpoint score (SAPS, high = poor, skewed data) – short term

Mental state: 1a. Specific – positive – average endpoint score (SAPS, high = poor, skewed data) – short term

Study	Intervention	Mean	SD	n
ISRCTN65314790	Avatar therapy	32.1	20.5	62
	Treatment as usual	39.93	20.75	61

Analysis 2.2. Comparison 2: Avatar therapy versus supportive counselling, Outcome 2: Mental state: 1b. Specific – insight – average endpoint score (various scales, high = poor) – short term



Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 2.3. Comparison 2: Avatar therapy versus supportive counselling, Outcome 3: Mental state: 1c. Specific – depression – average endpoint score (various scales, high = poor, skewed data) – short term

Mental state: 1c. Specific – depression – average endpoint score (various scales, high = poor, skewed data) – short term

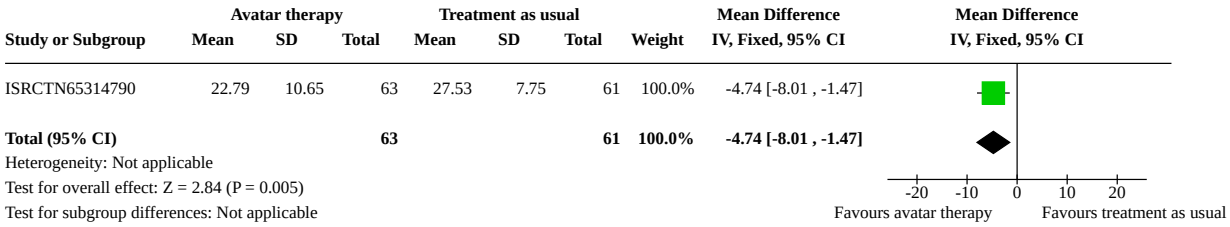
Study	Intervention	Mean	SD	n
CDS				
ISRCTN65314790	Avatar therapy	5.43	5.2	63
	Treatment as usual	7.41	4.82	61
DASS-21 total				
ISRCTN65314790	Avatar therapy	19.46	14.08	62
	Treatment as usual	22.63	15.52	61

Analysis 2.4. Comparison 2: Avatar therapy versus supportive counselling, Outcome 4: Mental state: 1d. Specific – negative – average endpoint score (SANS, high = good, skewed data) – short term

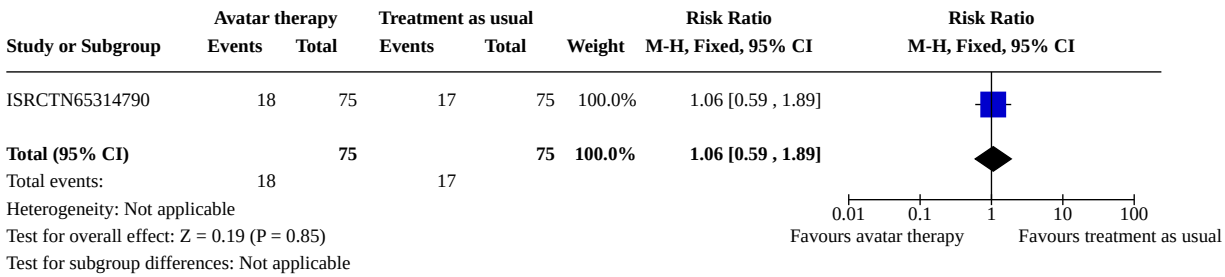
Mental state: 1d. Specific – negative – average endpoint score (SANS, high = good, skewed data) – short term

Study	Intervention	Mean	SD	N
ISRCTN65314790	Avatar Therapy	26.67	19.79	63
	Supportive counselling	25.61	17.66	—

Analysis 2.5. Comparison 2: Avatar therapy versus supportive counselling, Outcome 5: Mental state: 2a. General – average endpoint score (PSYRATS total, high = poor) – short term

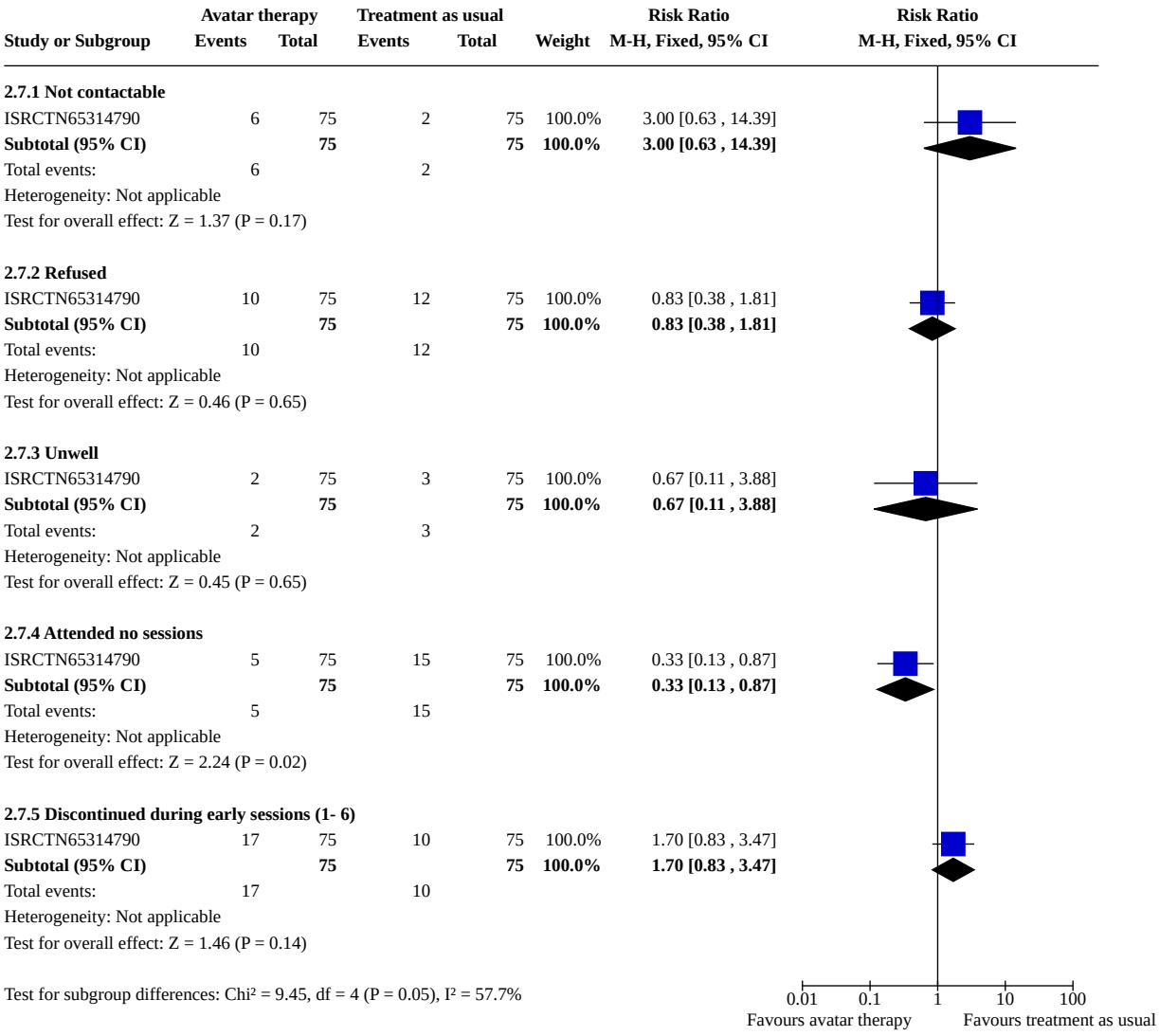


Analysis 2.6. Comparison 2: Avatar therapy versus supportive counselling, Outcome 6: Leaving study early: 1a. For any reason – short term



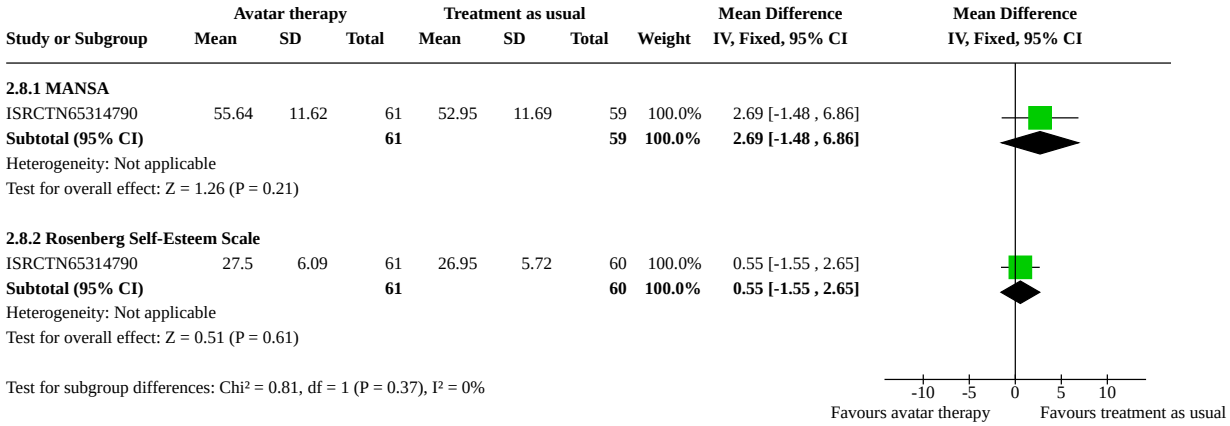
Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 2.7. Comparison 2: Avatar therapy versus supportive counselling, Outcome 7: Leaving study early: 1b. For specific reason – short term



Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 2.8. Comparison 2: Avatar therapy versus supportive counselling, Outcome 8: Quality of life: average endpoint score (various scales, high = good) – short term



ADDITIONAL TABLES

Table 1. Design of a future study

Methods	Allocation: randomised (clearly described)
	Blinding: single blind (outcomes assessor)
	Duration: minimum 1 year
	Design: parallel, crossover
	Setting: outpatient and inpatient settings
Participants	Diagnosis: people with a clinical diagnosis of schizophrenia using DSM-IV criteria
	History: adults randomised from waiting list and referred to research staff
	N= 300
	Age: ≥ 18 years
	Sex: men and women
	Inclusion criteria: DSM-IV diagnosis of schizophrenia, age ≥ 18 years
	Exclusion criteria: any other disease or alcohol and drug problem
Interventions	1. Avatar Therapy. N= 150
	Programme for 12 weeks, once a week, follow-up at 6 months and 1 year
	2. Waiting list. N= 150
	They will receive Avatar Therapy sessions after phase one is finished.
Outcomes	Mental state (binary outcomes)
	Relapses (binary outcomes)
	Quality of life (binary outcomes)
	Costs: cost of services, cost of care

Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Table 1. Design of a future study *(Continued)*

	Adverse events related to yoga (number and type of injuries) Service outcomes: days in hospital, time attending outpatient psychiatric clinic
Notes	Adherence should be logged with participants expected to adhere to 70–75% of scheduled sessions.

DSM-IV: Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition; n: number of participants.

HISTORY

Protocol first published: Issue 9, 2015
Review first published: Issue 5, 2020

CONTRIBUTIONS OF AUTHORS

GA: data extraction, data entry, review writing and data analyses.

TK: data extraction, data entry.

FS: protocol development, search, screening and final check before submission.

DECLARATIONS OF INTEREST

GA: none.

TK: none.

FS: none.

SOURCES OF SUPPORT

Internal sources

- No source of support, Other

External sources

- No sources of support supplied

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

1. General

Since protocol publication (Moazzen 2015), we have updated the methods section of the review using the latest template provided by Cochrane Schizophrenia Group. These amendments, slightly affected the review criteria and objectives.

2. Specific

2.1. Objectives

We have altered the wording of the objectives to more accurately reflect the review title.

2.2. Types of outcome measures

We have made changes to the layout and format of Types of outcome measures in line with latest Cochrane Schizophrenia Group presentation of outcomes.

For example – significant response is now clinically important change.

There were too many mental state outcomes listed as primary outcomes in the protocol. We have specified two (positive symptoms and insight) and added at least one adverse effect, in line with MECIR requirements to include an adverse effect outcome in primary outcomes.

Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

2.3. Summary of findings outcomes

There were only five SOF outcomes in the protocol. We have added two further outcomes to the 'Summary of findings' tables (adverse effects and leaving the study early).

INDEX TERMS

Medical Subject Headings (MeSH)

Anxiety [etiology]; Bias; Confidence Intervals; Hallucinations [*therapy]; Intention to Treat Analysis; Mental Health; Odds Ratio; Patient Dropouts [psychology] [statistics & numerical data]; Psychotic Disorders [therapy]; Quality of Life; Randomized Controlled Trials as Topic; Recurrence; Schizophrenia [*therapy]; Schizophrenic Psychology; Therapies, Investigational [adverse effects] [*methods]; Therapy, Computer-Assisted [methods]; Virtual Reality Exposure Therapy [*methods]

MeSH check words

Adult; Female; Humans; Male

Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.