Cochrane Library

Cochrane Database of Systematic Reviews

Peer support for people with schizophrenia or other serious mental illness (Review)

Chien WT, Clifton AV, Zhao S, Lui S

Chien WT, Clifton AV, Zhao S, Lui S.
Peer support for people with schizophrenia or other serious mental illness.
Cochrane Database of Systematic Reviews 2019, Issue 4. Art. No.: CD010880.
DOI: 10.1002/14651858.CD010880.pub2.

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

[Intervention Review]

Peer support for people with schizophrenia or other serious mental illness

Wai Tong Chien[1], Andrew V Clifton[2], Sai Zhao[3], Steve Lui[4]

[1]Nethersole School of Nursing, The Chinese University of Hong Kong, Shatin, Hong Kong. [2]Faculty of Health and Life Sciences, De Montfort University, Leicester, UK. [3]Systematic Review Solutions Ltd, The Ingenuity Centre, The University of Nottingham, Nottingham, UK. [4]School of Human and Health Sciences, University of Huddersfield, Huddersfield, UK

Contact: Steve Lui, School of Human and Health Sciences, University of Huddersfield, Harold Wilson Building, Queensgate, Huddersfield, HD1 3DH, UK. S.Lui@hud.ac.uk.

Editorial group: Cochrane Schizophrenia Group.
Publication status and date: Edited (no change to conclusions), published in Issue 6, 2019.

Citation: Chien WT, Clifton AV, Zhao S, Lui S. Peer support for people with schizophrenia or other serious mental illness. *Cochrane Database of Systematic Reviews* 2019, Issue 4. Art. No.: CD010880. DOI: 10.1002/14651858.CD010880.pub2.

ABSTRACT

Background

Peer support provides the opportunity for peers with experiential knowledge of a mental illness to give emotional, appraisal and informational assistance to current service users, and is becoming an important recovery-oriented approach in healthcare for people with mental illness.

Objectives

To assess the effects of peer-support interventions for people with schizophrenia or other serious mental disorders, compared to standard care or other supportive or psychosocial interventions not from peers.

Search methods

We searched the Cochrane Schizophrenia Group's Study-Based Register of Trials on 27 July 2016 and 4 July 2017. There were no limitations regarding language, date, document type or publication status.

Selection criteria

We selected all randomised controlled clinical studies involving people diagnosed with schizophrenia or other related serious mental illness that compared peer support to standard care or other psychosocial interventions and that did not involve 'peer' individual/group(s). We included studies that met our inclusion criteria and reported useable data. Our primary outcomes were service use and global state (relapse).

Data collection and analysis

The authors of this review complied with the Cochrane recommended standard of conduct for data screening and collection. Two review authors independently screened the studies, extracted data and assessed the risk of bias of the included studies. Any disagreement was resolved by discussion until the authors reached a consensus. We calculated the risk ratio (RR) and 95% confidence interval (CI) for binary data, and the mean difference and its 95% CI for continuous data. We used a random-effects model for analyses. We assessed the quality of evidence and created a 'Summary of findings' table using the GRADE approach.

Main results

This review included 13 studies with 2479 participants. All included studies compared peer support in addition to standard care with standard care alone. We had significant concern regarding risk of bias of included studies as over half had an unclear risk of bias for

the majority of the risk domains (i.e. random sequence generation, allocation concealment, blinding, attrition and selective reporting). Additional concerns regarding blinding of participants and outcome assessment, attrition and selective reporting were especially serious, as about a quarter of the included studies were at high risk of bias for these domains.

All included studies provided useable data for analyses but only two trials provided useable data for two of our main outcomes of interest, and there were no data for one of our primary outcomes, relapse. Peer support appeared to have little or no effect on hospital admission at medium term (RR 0.44, 95% CI 0.11 to 1.75; participants = 19; studies = 1, very low-quality evidence) or all-cause death in the long term (RR 1.52, 95% CI 0.43 to 5.31; participants = 555; studies = 1, very low-quality evidence). There were no useable data for our other prespecified important outcomes: days in hospital, clinically important change in global state (improvement), clinically important change in quality of life for peer supporter and service user, or increased cost to society.

One trial compared peer support with clinician-led support but did not report any useable data for the above main outcomes.

Authors' conclusions

Currently, very limited data are available for the effects of peer support for people with schizophrenia. The risk of bias within trials is of concern and we were unable to use the majority of data reported in the included trials. In addition, the few that were available, were of very low quality. The current body of evidence is insufficient to either refute or support the use of peer-support interventions for people with schizophrenia and other mental illness.

PLAIN LANGUAGE SUMMARY

Peer support for schizophrenia and other serious mental illnesses

Background

Schizophrenia and other serious mental illnesses are chronic disruptive mental disorders with disturbing psychotic, affective and cognitive symptoms such as delusions, hallucinations, depression, anxiety, insomnia, difficulty in concentration, suspiciousness and social withdrawal. The primary treatment is antipsychotic medicine, but these are not always fully effective.

Peer support provides the opportunity for both a service user and a provider of care to share knowledge, direct experience of their illness and to help each other along the path to recovery. The support is given alongside antipsychotic treatment. Through interpersonal sharing, modelling and assistance within or outside of group sessions, it is believed that these supportive strategies can help combat feelings of hopelessness and behavioural problems that may result from having an illness and empower people to continue their treatment and help them to resume key roles in real life. However, findings from research have been inconsistent regarding the effectiveness of peer support for people with schizophrenia and other serious mental illnesses.

Review aims

This review aimed to find high-quality evidence from relevant randomised clinical trials (studies where people are randomly put into one of two or more treatment groups) so we could assess the effects of peer-support interventions for people with serious mental illness in comparison to standard care or other supportive or psychosocial interventions not from peers. We were interested in finding clinically meaningful data that could provide information regarding the effect peer support has on hospital admission, relapse, global state, quality of life, death and cost to society for people with schizophrenia.

Searches

We searched Cochrane Schizophrenia's specialised register of trials (up to 2017) and found 13 trials that randomised 2479 people with schizophrenia or other similar serious mental illnesses to receive either peer support plus their standard care, clinician-led support plus their standard care or standard care alone.

Key results

Thirteen trials were available but the evidence was very low quality. Useable data were reported for only two of our prespecified outcomes of importance and showed adding peer support to standard care appeared to have little or no clear impact on hospital admission or death for people with schizophrenia and other serious mental illnesses. One of these trials (participants = 156) also compared peer support with clinician-led support (where a health professional provided support). However, there were no useable data for this comparison reported for the main outcomes.

Conclusions

We have little confidence in the above findings. Currently, there is no high-quality evidence available to either support or refute the effectiveness of peer-support interventions for people with schizophrenia or other serious mental illnesses.

SUMMARY OF FINDINGS

Summary of findings for the main comparison. Peer support plus standard care versus standard care for people with schizophrenia or similar serious mental illness

Peer support + standard care vs standard care for people with schizophrenia or similar serious mental illness

Patient or population: people with schizophrenia or other serious mental illness

Settings: inpatients and outpatients

Intervention: peer support + standard care vs standard care

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Peer-support vsstandard care				
Service use: hospital admission – medium term Follow-up: 5 months	**Study population**		**RR 0.44** (0.11 to 1.75)	19 (1 study)	⊕⊝⊝⊝ **Very low**[a,b,c]	—
	500 per 1000	**220 per 1000** (55 to 875)				
	Moderate					
	500 per 1000	**220 per 1000** (55 to 875)				
Service use: days in hospital – medium term Follow-up: 5 months	See comments	See comments	See comments	See comments	—	Data were skewed and could not be use in analyses. See Analysis 1.2.
Global state: relapse	See comments	See comments	See comments	See comments	See comments	No data.
Global state: clinically important change in global state	See comments	See comments	See comments	See comments	See comments	No data
Peer outcomes: clinically important change in quality of life for service user and peer supporter	See comments	See comments	See comments	See comments	—	No study reported data for clinically important change in quality of life. 4 studies mea-

						sured quality of life in the medium term by using different scales; see Analysis 1.37.
Adverse events: all cause – long term Follow-up: 40 weeks	**Study population**		**RR 1.52** (0.43 to 5.31)	555 (1 study)	⊕⊝⊝⊝ **Very low**[a,b,c]	—
	14 per 1000	**22 per 1000** (6 to 76)				
	Moderate					
	14 per 1000	**21 per 1000** (6 to 74)				
Economic: indirect costs (cost to society)	See comments	See comments	See comments	See comments	See comments	No useable data.

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).

CI: confidence interval; **RR:** risk ratio.

GRADE Working Group grades of evidence
High quality: further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: we are very uncertain about the estimate.

[a]Risk of bias downgraded one level due to high risk of performance and detection bias.
[b]Indirectness downgraded one level due to participants having mental illnesses other than schizophrenia.
[c]Imprecision downgraded one level due to very small sample size or low incidence of events.

Summary of findings 2. Peer support plus standard care versus clinician-led support plus standard care for people with schizophrenia or similar serious mental illness

Peer support + standard care vs clinician-led support + standard care for people with schizophrenia or similar serious mental illness
Patient or population: people with schizophrenia or other serious mental illness
Settings: inpatients and outpatients
Intervention: peer support + standard care vs clinician-led support + standard care

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of partici-pants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Peer support vsclinician-led support				
Service use: hospital admission	Data not available for this outcome					
Service use: days in hospital	Data not available for this outcome					
Global state: relapse	Data not available for this outcome					
Global state: clinically important change in global state	Data not available for this outcome					
Peer outcomes: clinically important change in quality of life for service user and peer supporter	Data not available for this outcome					
Adverse events: all cause	Data not available for this outcome					
Economic: indirect costs (cost to society)	Data not available for this outcome					

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).

CI: confidence interval.

GRADE Working Group grades of evidence
High quality: further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: we are very uncertain about the estimate.

BACKGROUND

Description of the condition

The definition of serious mental illness with the widest consensus is that of the US National Institute of Mental Health (NIMH) (Schinnar 1990), and is based on diagnosis, duration and disability (NIMH 1987). People with serious mental illness have conditions such as schizophrenia or bipolar disorder, which last over a protracted period resulting in the erosion of functioning in day-to-day life. Schizophrenia is a chronic, disruptive, mental illness that frequently contributes to a wide variety of functional disabilities, especially within social and occupational domains (Harvey 2012). The worldwide estimate for the life-time prevalence of schizophrenia ranges from 1.4 per 1000 people to 4.6 per 1000 people; the annual incidence rate lies between 0.16 per 1000 people and 0.42 per 1000 people, with onset often occurring in adolescence and early adulthood (Jablensky 2000). The psychopathology of schizophrenia is often described in terms of the severity of positive (e.g. hallucinations and disorganised speech) and negative (e.g. blunted affect and social withdrawal) symptoms. While antipsychotic medications remain the core treatment for controlling the symptoms of schizophrenia, they are associated with a range of undesirable adverse effects on cardiovascular, endocrine and other bodily systems, resulting in poor treatment adherence (Kane 2010).

About 30% of people with schizophrenia have persistent and severe negative symptoms that tend to be resistant to medication. Termed 'deficit syndrome', persistent negative symptoms are characterised by lack of initiative, interests and social fluency; poor verbal communication and concentration; and loss of interpersonal function (Nasrallah 2011; Tandon 2009). Together with progressive deterioration in various cognitive functions (e.g. problems in working memory and information processing, reasoning and problem solving, and social cognition), there are considerable and wide varieties of functional impairments which can severely compromise overall psychosocial functioning, social integration and quality of life (Mohamed 2008). These factors may all eventually reduce treatment efficacy in people with schizophrenia.

The total societal costs of schizophrenia, including treatment, rehabilitation, community care services and loss of productivity, were estimated at more than USD 60 billion per annum in the USA, UK and other high-income countries in the 20th century (Mangalore 2007; Wu 2005). People with schizophrenia have severe social and occupational disability (30%) and are at higher risks of other mental health (e.g. 25% to 30% have depression) and physical health (e.g. 20% to 25% have cardiovascular disease) problems (De Hert 2009), have a two- to three-times higher all-cause mortality rate and are 12 times more likely to die by suicide than the general population (Goff 2005; Wildqust 2010).

Description of the intervention

Peer support is broadly defined as "a system of giving and receiving help founded on key principles of respect, shared responsibility, and mutual agreement of what is helpful" (Mead 2001). Dennis 2003 defined 'peer support' within a healthcare context as ".... the provision of emotional, appraisal and informational assistance by a created social network member who possesses experiential knowledge of a specific behaviour or stressor and similar characteristics as the target population" (Dennis 2003). Peers can be referred to those people who share common characteristics with a specific individual or group, affiliating and empathising with and supporting each other to promote health and deal with life problems. The emphasis is on the idea that 'peers' are considered to be equal (Dennis 2003); in contrast to the traditional healthcare system of mental health services, which distinguishes between providers (i.e. trained professionals) and consumers (e.g. people with schizophrenia and families/friends), peer-support programmes are built on collaborative, mutual and equal partnerships of participants who share their experiences (or expertise) in different stages of recovery (Repper 2010).

Peer-support programmes for people with schizophrenia are mainly classified into two main categories, according to how they run the services and the roles played by their co-ordinators or facilitators (Ahmed 2012).

One type of peer support programme is the mutual/self-help group led by professionals/clinicians. The group members have similar life issues or situations such as care giving to a chronically ill relative. The clinician or professional facilitates the group members to come together for sharing and establishing coping strategies, feeling more empowered and obtaining a sense of community. The clinician or professional acts as a facilitator to assist the group members to get help during the process of relating personal experiences, listening to and accepting others' experiences, providing sympathetic understanding and establishing social networks.

The other type of peer support programme is the consumer-led programme, in which consumers provide supportive services to other patients and their families and offer advice to the mental healthcare team. The consumer-led service is a more structured programme in terms of its system, structure and group sessions. It involves consumers more with leadership of the co-ordinators or facilitators, or both. The consumers are often peer volunteers or the peer specialists who are employed in the healthcare setting to advocate for other consumers.

However, both categories of peer-support programme emphasise interactive mutual peer or social learning. In response to individual groups' and group members' needs, their content can range from psychoeducation about schizophrenia and its symptom management, medication adherence, stress reduction and coping strategies, to problem-solving approaches, and the strengthening of family and community support resources, as well as vocational and social skills training (Chien 2009).

How the intervention might work

Peer support has become an increasingly important strategy in healthcare systems that are encountering limited manpower and resources on one the hand and, on the other hand, continuously increasing costs of managing complex and chronic illnesses such as severe mental disorders (Bradstreet 2010). Peer support has been widely used to improve physical and psychosocial health and enhance behavioural change and self-care in diverse chronic conditions, as well as in population groups in need of support (Cheah 2001). A peer-support programme can provide a platform where fellow patients and those already recovered or on their way to recovery from schizophrenia, or another mental illness, can share their individual experiences of the illness and management strategies in everyday life in a way that is not commonly offered in

traditional healthcare settings where mental-health professionals may often dominate services (Chien 2009). In contrast to traditional healthcare settings, often stigmatised by the general public, the environment of a peer-support group fosters a sense of emotional support, information exchange, companionship, reassurance and appraisals among group members (Ahmed 2012; Dennis 2003). Through interpersonal sharing, modelling and assistance within or outside of group sessions, it is believed that these supportive strategies can effectively combat hopelessness and behavioural problems relating to mental illness and specifically schizophrenia, and empower participants to continue treatment and resume key roles in real life (Chien 2009; Davidson 1999). However, research has shown inconsistent findings on whether social or peer support enhances self-care ability and medication adherence in people with mental illness (Pistrang 2008), and other chronic illnesses such as diabetic mellitus (Toljamo 2001).

While most peer-support groups mainly target those who are in the early stages of recovery, the benefits of these group programmes are not limited only to those who receive the peer-support service, but also extend to those who provide peer support to others (Miyamoto 2012). The peer-support providers who are assigned the roles of co-ordinator or facilitator of the group can successfully rebuild their self-efficacy through having the chance to serve other people with similar conditions. They may even collaborate with professionals to deliver appropriate services to other group members in need. Through active participation in service provision, they themselves increase their knowledge of disease management and enhance various skills that are important to daily functioning (Arnstein 2002).

Why it is important to do this review

Systematic reviews and practice guidelines have recommended that, in adjunction to psychopharmacological treatment, psychosocial interventions designed to support people with schizophrenia and their families should also be used to improve the person's rehabilitation, reintegration into the community and recovery from the illness (NICE 2009; Pharoah 2010). There is now an increasing body of evidence concerning the effects of a range of psychosocial interventions for schizophrenia, including psychoeducation (Xia 2011), cognitive-behavioural therapy (Morrison 2009; Turkington 2004), and family intervention (Pharoah 2010). While psychosocial interventions have indicated significant positive effects on reducing relapse and readmission rates, and enhancing medication compliance, most have not demonstrated consistent and conclusive results in improving psychosocial health conditions of people with schizophrenia. Moreover, research has shown inconsistent findings on whether social or peer support enhances self-care ability and medication adherence in people with mental illness (Pistrang 2008), and other chronic illnesses such as diabetic mellitus (Toljamo 2001). Therefore, the design or testing of alternative approaches to psychosocial intervention for these people should be considered. Guided by the consumer movement and recovery model in mental health care, peer support is one such approach to psychosocial intervention that places emphasis on promoting the overall wellness and empowerment of people with schizophrenia through establishing partnerships between those with the condition throughout the whole journey of recovery (Ahmed 2012).

With its emphasis on the experiences of people with schizophrenia, their needs and perspectives in treatment planning, peer-support programmes have led to growing interest in the role that those who are experiencing difficulties with recovery can play in enlightening the social reintegration and enhancing the rehabilitation process of others with similar mental health problems (Ahmed 2012). The number of peer-support programmes for schizophrenia care has increased rapidly in high-income countries such as the USA and Canada.(REF) Nevertheless, there is no systematic review on the impetus for this alternative treatment approach and its effects on mental condition; relapse; medication adherence; and a wide variety of outcomes such as psychosocial and occupational functioning, social skills, self-efficacy, overall wellness and quality of life in people with schizophrenia (Miyamoto 2012).

This review focused on peer-support programmes and their use varies across cultures. There are no systematic reviews on this topic in the area of schizophrenia and only a few reviews have been published on the effects of support groups for various kinds of mental health problems (e.g. LIoyd-Evans 2014; Pistrang 2008). The findings of this review will enhance our knowledge of the effectiveness of peer-support interventions and the various models for the delivery of peer-support interventions across cultures. The costs and benefits of these programmes can then be systematically evaluated.

OBJECTIVES

To assess the effects of peer-support interventions for people with schizophrenia or other serious mental disorders, compared to standard care or other supportive or psychosocial interventions not from peers.

METHODS

Criteria for considering studies for this review

Types of studies

We included all relevant randomised controlled trials (RCTs), including cluster randomised trials, that evaluated the effects of peer support for people with schizophrenia or similar serious mental illness. We excluded studies that did not include a control or comparison group. Where the participants were given additional types of treatments within peer support, we only included data if the adjunct treatment was applied equally to all study groups and it was only peer support that was randomised and allocated to the treatment or intervention group(s).

If a trial had been described as 'double blind' but only implied randomisation, we would have included such trials in a sensitivity analysis (see Sensitivity analysis). We excluded quasi-randomised studies, such as those allocating participants by alternate days of the week.

Types of participants

We required:

- the majority of participants to be aged 18 to 65 years;
- the majority of participants to have a serious mental illness preferably as defined by NIMH criteria (NIMH 1987), but, in the absence of that, from illness such as schizophrenia, schizophrenia-like disorders, bipolar disorder or serious affective disorders;

- if a trial included participants with a range of serious mental illnesses we included it only if at least 20% of the participants had schizophrenia or schizophrenia-like disorders.

We did not consider substance abuse to be a serious mental illness in its own right; however, studies were eligible if they dealt with people with both diagnoses (i.e. those with serious mental illnesses plus substance abuse). Dementia and mental retardation are not considered to be a serious mental disorder, hence we excluded studies focusing on these populations. Despite the fact that personality disorder was now included in the NIMH definition of serious mental illnesses, we excluded this from our review on the basis that the diagnosis of personality disorders had low inter-rater reliability (Zimmerman 1994), the duration of treatment can be assessed much more precisely than duration of illness (Schinnar 1990), and that insufficient information was given on how to diagnose disability criterion in both the original NIMH definition (NIMH 1987), and in the further work of Schinnar 1990.

Types of interventions

1. Intervention

1.1 Peer support

We defined a 'peer' as someone selected to provide support because they had similar or relevant health experience (Dale 2008). See also Description of the intervention.

2. Comparators

2.1 Standard care

Care that a participant would normally receive in the area in which the trial took place. This normally includes biological, psychological and social approaches to care including antipsychotic medication, and utilisation of services including hospital stay, day hospital attendance and community psychiatric nursing involvement.

2.2 Other psychosocial intervention

Any psychosocial intervention or any supportive intervention (e.g. cognitive-behavioural therapy, psychoeducation programmes, family interventions, social skills training programmes) that did not involve a 'peer' individual/group(s).

Types of outcome measures

We divided outcomes into short term (up to one month), medium term (one or more to six months) and long term (more than six months).

Primary outcomes

1. Service use

1.1 Hospital admission
1.2 Duration of hospital stay (days)

2. Global state

2.1 Relapse – as defined by each of the studies
2.2 Clinically important change in global state (e.g. improved/not improved to an important extent) – as defined by each of the studies

3. Adverse event

3.1 Death: all cause

Secondary outcomes

1. Service use

1.1 Clinically important engagement with all services
1.2 Any contact with services
1.3 Any contact with specialist community services (i.e. early intervention teams, assertive outreach teams and crisis teams)
1.4 Time to hospitalisation

2. Global state

2.1 Any change in global state (improved/not improved) – as defined by each of the studies
2.2 Mean change or endpoint score on global state scale
2.3 Time to relapse
2.4 Compliance with treatment

3. Mental state

3.1 Overall

3.1.1 Clinically important change in overall mental state (improved/not improved to an important extent) – as defined by each of the studies
3.1.2 Any change in mental state (improved/not improved) – as defined by each of the studies
3.1.3 Mean endpoint or change score on mental state scale

3.2 Specific

3.2.1 Clinically important change in specific symptoms (e.g. positive, negative, affective) – as defined by each of the studies
3.2.2 Any change in specific symptoms (e.g. positive, negative, affective) – as defined by each of the studies
3.2.3 Mean endpoint or change score on specific mental state scale

4. Behaviour

4.1 General

4.1.1 Clinically important change in general behaviour – as defined by each study
4.1.2 Any change in general behaviour – as defined by each study
4.1.3 Mean endpoint or change score on general behaviour scale

4.2 Specific

4.1.1 Clinically important change in specific behaviour (e.g. aggression) – as defined by each study
4.1.2 Any change in specific behaviour – as defined by each study
4.1.3 Mean endpoint or change score on specific behaviour scale

5. Leaving the study early

5.1 For any reason
5.2 For specific reason

6. Functioning

6.1 General

6.1.1 Clinically important change in general functioning – as defined by each study
6.1.2 Any change in general functioning – as defined by each study
6.1.3 Mean endpoint or change score on general functioning scale

6.2 Specific (e.g. social, cognitive, psychological, life skills)

6.2.1 Clinically important change in specific functioning – as defined by each study

6.2.2 Any change in specific functioning – as defined by each study
6.2.3 Mean endpoint or change score on specific functioning scales
6.2.4 Employment status or work-related activities
6.2.5 Independent living
6.2.6 Imprisonment/contact with police/justice system

7. Peer outcomes

7.1 Impact on the service user and peer supporter (e.g. anxiety and perceived social support)
7.2 Coping ability/self-efficacy of service user and peer supporter
7.3 Expressed emotion of family, peer supporter or both

7.4 Quality of life for service user and peer supporter

7.4.1 Clinically important change in quality of life for service user and peer supporter

7.4.2 Any change in quality of life for service user and peer supporter

7.4.3 Mean endpoint or change score on quality of life scale

7.5 Satisfaction with care for service user and peer supporter

7.5.1 Clinically important change in satisfaction of life for service user and peer supporter

7.5.2 Any change in satisfaction for service user and peer supporter

7.5.3 Mean endpoint or change score on satisfaction scale

8. Adverse effects

8.1 General adverse effects

8.1.1 At least one adverse effect
8.1.2 Any incidence of clinically important adverse effect
8.1.3 Mean endpoint or change score on adverse effect scale

8.2 Specific adverse effects

8.2.1 Incidence of various specific effects

9. Economic outcomes

9.1 Cost of care
9.2 Direct costs
9.3 Indirect costs

'Summary of findings' table

We used the GRADE approach to interpret findings (Schünemann 2011) and GRADEpro GDT to export data from our review to create the 'Summary of findings' tables. These tables provided outcome-specific information concerning the overall quality of evidence from each included study in the comparison, the magnitude of effect of the interventions examined and the sum of available data on all outcomes we rated as important to the care of people with schizophrenia and to decision making. We aimed to select the following main outcomes for inclusion in the 'Summary of findings' table.

- Service use: hospital admission.
- Service use: duration of hospital stay (days).
- Global state: relapse – as defined by each of the studies.
- Global state: clinically important change in global state.
- Adverse events: death – all cause.
- Peer outcomes: clinically important change in quality of life for service user and peer supporter.
- Economic outcomes: indirect costs (increased cost to society).

Search methods for identification of studies

Electronic searches

Cochrane Schizophrenia Group's Study-Based Register of Trials

On 27 July 2016 and 4 July 2017, the information specialist searched the register using the following search strategy which were developed based on literature review and consulting with the authors of the review:

(*Peer* OR *Self-Help* OR *Social Support* OR *Social Network*) 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.

This register is compiled by systematic searches of major resources (including MEDLINE, Embase, AMED, BIOSIS, CINAHL, PsycINFO, PubMed and registries of clinical trials) and their monthly updates, handsearches, grey literature and conference proceedings (see Group's Module). There is no language, date, document type or publication status limitations for inclusion of records into the register. See Appendix 1 for previous search terms.

Searching other resources

1. Reference searching

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

2. Personal contact

We contacted the first author of each included study for information regarding unpublished trials. However, no unpublished trial was identified through this method.

Data collection and analysis

Selection of studies

Two review authors (SL, WTC) screened the results of the electronic search, a third review author (AC) checked the screening. WTC inspected all abstracts of studies identified through screening and identify potentially relevant reports. Once identified, to ensure reliability, AC inspected a random sample of these abstracts, comprising 10% of the total. Where disagreement occurred, we resolved this by discussion, and where there was still doubt, we acquired the full article for further inspection. We then requested the full articles of relevant reports for reassessment and carefully inspect them for a final decision on inclusion. Two review authors (WTC, SL) independently inspected all full reports and decided whether they met the inclusion criteria. We were not blinded to the names of the authors, institutions or journal of publication. Where difficulties or disputes arose, we asked one review author (AC) for help; if it was impossible to decide, we added these studies to those awaiting assessment and contacted the authors of the papers for clarification.

Data extraction and management

1. Extraction

Two review authors (SL, WTC) independently extracted data from included studies. We discussed any disagreement, documented our

decisions and, if necessary, we contacted the authors of studies for clarification. We had planned to extract data presented only in graphs and figures whenever possible, but would have only included such data only if two review authors independently reached the same result. We attempted to contact authors through an open-ended request to obtain any missing information or for clarification whenever necessary. Where applicable, we extracted data relevant to each component centre of multi-centre studies separately (see the Cochrane Schizophrenia Group Module).

2. Management

2.1 Forms

We extracted data onto standard, predesigned simple forms.

2.2 Scale-derived data

We included continuous data from rating scales only if:

- the psychometric properties of the measuring instrument had been described in a peer-reviewed journal (Marshall 2000); and
- the measuring instrument had not been written or modified by one of the trialists for that particular trial.

Ideally, the measuring instrument should have either been a self-report or completed by an independent rater or relative (not the therapist). We realised that this is not often reported clearly; we noted if this is the case or not in the Description of studies section.

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; however, calculation of change needs two assessments (baseline and endpoint) that can be difficult to obtain in unstable and difficult-to-measure conditions such as schizophrenia. We have decided primarily to use endpoint data, and only use change data if the former are not available. If necessary, we will combine endpoint and change data in the analysis, as we prefer to use mean differences (MDs) rather than standardised mean differences (SMDs) throughout (Deeks 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 relevant continuous data before inclusion.

For endpoint data from studies including fewer than 200 participants:

- when a scale started from the finite number zero, we subtracted the lowest possible value from the mean, and divide this by the standard deviation (SD). If this value was lower than one, it strongly suggested that the data were skewed and we would have excluded these data. If this ratio was higher than one but less than two, there was suggestion that the data were skewed: we would have entered these data and tested whether their inclusion or exclusion would change the results substantially. If such data changed results, we would have entered them as 'other data'. Finally, if the ratio was larger than two, we would have included these data, because it was less likely that they were skewed (Altman 1996);
- 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 would have modified the calculation described above to take the scale starting point into account. In these cases, skewed data were present if $2\ SD > (S - S_{min})$, where S was the mean score and S_{min} was the minimum score.

Note: we would have entered all relevant data from studies of more than 200 participants in the analysis irrespective of the above rules, because skewed data pose less of a problem in large studies. We would also have entered all relevant change data, as when continuous data were presented on a scale that included a possibility of negative values (such as change data), it was difficult to determine whether or not data were skewed.

2.5 Common measure

To facilitate comparison between trials, we converted variables that could have been 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, efforts were made to convert outcome measures to dichotomous data. This was done by identifying cut-off points on rating scales and dividing participants accordingly into 'clinically improved' or 'not clinically improved'. It was generally assumed that if there was a 50% reduction in a scale-derived score such as the Brief Psychiatric Rating Scale (Overall 1962) or the PANSS (Kay 1986), this could be considered a clinically significant response (Leucht 2005a; Leucht 2005b). If data based on these thresholds were not available, we 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 peer support. Where keeping to this made it impossible to avoid outcome titles with clumsy double-negatives (e.g. 'not improved') we reported data in such a way that the area to the left of the line indicated an unfavourable outcome. This was noted in the relevant graphs.

Assessment of risk of bias in included studies

Two review authors (SL, AVC) independently assessed risk of bias using criteria described in the *Cochrane Handbook for Systematic Reviews of Interventions* to assess trial quality (Higgins 2011a). This set of criteria was based on evidence of associations between an overestimation of effect and high risk of bias in an article, such as due to sequence generation, allocation concealment, blinding, incomplete outcome data and selective reporting. 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 contacted authors of the studies to request further information. We reported non-concurrence in quality assessment but, if disputes arose as to which category a trial was to be allocated to, again resolution was made by discussion. We noted the level of risk of bias in both the text of the review and in the 'Summary of findings' tables.

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). It was shown that the RR was more intuitive (Boissel 1999) than the odds ratio, and that odds ratios tended to be interpreted as RR by clinicians (Deeks 2000). The number required to treat for an additional harmful outcome statistic with its 95% CI was intuitively attractive to clinicians but was problematic both in its accurate calculation in meta-analyses and its interpretation (Hutton 2009). For binary data presented in the 'Summary of findings' tables, we calculated illustrative comparative risks where possible.

2. Continuous data

For continuous outcomes, we estimated MD and its 95% CI between groups. We preferred not to calculate effect size measures (standardised mean difference). However, if scales of very considerable similarity had been used, we would have presumed there was a small difference in measurement, and would have calculated effect size and transformed the effect back to the units of one or more of the specific instruments.

Unit of analysis issues

1. Cluster trials

Studies increasingly employed 'cluster randomisation' (such as randomisation by clinician or practice), but analysis and pooling of clustered data posed problems. First, authors often failed to account for intraclass correlation in clustered studies, leading to a 'unit of analysis' error (Divine 1992), whereby P values were spuriously low, CI unduly narrow and statistical significance overestimated. This caused type I errors (Bland 1997; Gulliford 1999).

If clustering had not been accounted for in primary studies, we would have presented data in a table, using a symbol (*) to indicate the presence of a probable unit of analysis error (Table 1). We would have contacted first authors of studies to obtain intraclass correlation coefficients (ICC) for their clustered data and if authors replied, adjusted for this using accepted methods (Gulliford 1999). If clustering had been incorporated into the analysis of primary studies, we would have presented these data as if from a non-cluster randomised study, but adjusted for the clustering effect.

We have sought statistical advice and been advised that binary data presented in a report should be divided by a 'design effect'. This can be calculated using the mean number of participants per cluster (m) and the ICC (design effect = 1 + (m – 1) * ICC) (Donner 2002). If the ICC had not been reported, it would be assumed to be 0.1 (Ukoumunne 1999).

If cluster studies had been appropriately analysed, taking into account ICC and relevant data documented in the report, synthesis with other studies would be possible using the generic inverse variance technique.

2. Cross-over trials

A major concern of cross-over trials is the carry-over effect. This occurs if an effect (e.g. pharmacological or physiological) of the treatment in the first phase of a trial is carried over to the second phase. As a consequence, on entry to the second phase, participants differ systematically from their initial state in spite of a washout phase. For the same reason, cross-over trials are also not appropriate if the condition of interest is unstable (Elbourne 2002). As both effects were very likely in severe mental illness, we would only have used data from the first phase of cross-over studies.

3. Studies with multiple treatment groups

Where a study involved more than two treatment arms, we presented the additional treatment arms in comparisons where relevant. If data were binary, we simply added these and combined them within the two-by-two table. If data were continuous, we combined data following the formula in *Cochrane Handbook for Systemic reviews of Interventions* (Higgins 2011b). Where the additional treatment arms were not relevant, we would not use these data.

Dealing with missing data

1. Overall loss of credibility

At some degree of loss of follow-up, data must lose credibility (Xia 2009). For any particular outcome, if more than 50% of data be unaccounted for, we did not reproduce these data or use them within analyses. However, if more than 50% of data in one arm of a study were lost, but the total loss was less than 50%, we addressed this within the 'Summary of findings' tables by downgrading quality. Finally, we would have downgraded quality within the 'Summary of findings' tables should data loss have been 25% to 50% in total.

2. Binary

In cases where the 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). Participants leaving the study early were all assumed to have the same rates of negative outcome as those who completed, with the exception of the outcomes of death and adverse effects. For these outcomes, the rate of those who stayed in the study – in that particular arm of the trial – was used for those who did not. Sensitivity analysis was undertaken to test how prone the primary outcomes were to change when data from only people who completed the study to that point were compared to the ITT analysis using the above assumptions.

3. Continuous

3.1 Attrition

In cases where the attrition for a continuous outcome was between 0% and 50%, and data only from people who completed the study to that point were reported, we reproduced these.

3.2 Standard deviations

If SD were not reported, we first tried to obtain the missing values from the authors. If not available, where there were missing measures of variance for continuous data, but an exact standard error (SE) and CI available for group means, and either a P value or t value available for differences in mean, we calculated SD according to the rules described in the *Cochrane Handbook for Systemic reviews of Interventions* (Higgins 2011b). When only the SE was reported, SD would have been calculated using the formula SD = SE * square root (n). Sections 7.7.3 and 16.1.3 of the *Cochrane Handbook for Systemic reviews of Intervention* presented detailed

formulae for estimating SD from P values, t or F values, CI, ranges or other statistics *s* (Higgins 2011b). If these formulae did not apply, we calculated the SD according to a validated imputation method which was based on the SD of the other included studies (Furukawa 2006). Although some of these imputation strategies can introduce error, the alternative would had been to exclude a given study's outcome and thus to lose information. We nevertheless would have examined the validity of the imputations in a sensitivity analysis excluding imputed values.

3.3 Last observation carried forward

We anticipated that in some studies the method of last observation carried forward (LOCF) would be employed within the study report. As with all methods of imputation to deal with missing data, LOCF introduces uncertainty about the reliability of the results (Leucht 2007). Therefore, where LOCF data have been used in the trial, if less than 50% of the data had been assumed, we would have presented and used these data, and indicated that they were the product of LOCF assumptions. 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 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 would not have excluded studies based on the statistical approach used. However, by preference we would have used the more sophisticated approaches, that is, we preferred to use MMRM or multiple-imputation to LOCF, and we would have only presented completer analyses if some type of ITT data were not available. Moreover, we would have 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 comparative data, to judge clinical heterogeneity. We simply inspected all studies for clearly outlying people or situations that we had not predicted would arise. When such situations or participant groups arose, we discussed these in the text.

2. Methodological heterogeneity

We considered all included studies initially, without seeing comparative data, to judge methodological heterogeneity. We simply inspected all studies for clearly outlying methods that we had not predicted would arise. When such methodological outliers arose, we discussed these in the text.

3. Statistical heterogeneity

3.1 Visual inspection

We visually inspected graphs 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 method alongside the Chi^2 statistic P value. The I^2 statistic provided an estimate of the percentage of inconsistency thought to be due to chance (Higgins 2003). The importance of the observed value of the I^2 statistic depends on magnitude and direction of effects; and strength of evidence for heterogeneity (e.g. P value from the Chi^2 test, or a CI for the I^2 statistic). I^2 statistic estimates of 50% or greater, accompanied by a statistically significant Chi^2 statistic ($P < 0.1$), were interpreted as evidence of substantial levels of heterogeneity (Deeks 2011). When there were substantial levels of heterogeneity in the primary outcomes, we explored reasons for heterogeneity (see Subgroup analysis and investigation of heterogeneity).

Assessment of reporting biases

Reporting biases arise when the dissemination of research findings is influenced by the nature and direction of results (Egger 1997). These are described in Section 10.1 of the *Cochrane Handbook for Systemic Reviews of Interventions* (Sterne 2011).

1. Protocol versus full study

We tried to locate 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

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 plots for outcomes where there were 10 or fewer studies, or where all studies were of similar size. In other cases, where funnel plots are possible, we will seek statistical advice in their interpretation.

Data synthesis

We understood that there was no closed argument regarding a preference for the use of fixed-effect or random-effects models. The random-effects method incorporated an assumption that the different studies were estimating different yet related intervention effects. To us, this often seemed to be true and the random-effects model took into account differences between studies even if there was no statistically significant heterogeneity. There was, however, a disadvantage to the random-effects model as it put added weight onto small studies, which were often those most biased. Depending on the direction of effect, these studies can either inflate or deflate the effect size. We chose a random-effects model for analyses.

Subgroup analysis and investigation of heterogeneity

1. Subgroup analyses

1.1 Clinical state, stage or problem

We aimed to provide an overview of the effects of peer support for people with schizophrenia in general. In addition, however, we tried to report data on subgroups of people in similar clinical state and stage, and with similar problems.

2. Investigation of heterogeneity

If inconsistency was high, this was reported. First, we investigated whether data had been entered correctly. Second, if data were correct, the graph was visually inspected, and outlying studies was successively removed to see whether 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, data were presented. If not, issues were discussed. We knew of no supporting research for this 10% cut-off but were investigating the use of prediction intervals as an alternative to this unsatisfactory state.

When unanticipated clinical or methodological heterogeneity was obvious, we 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 as to imply randomisation. For the primary outcomes, we would have included these studies; and if there was no substantive difference when the implied randomised studies were added to those with a better description of randomisation, we would have used all relevant data 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 compared the findings of the primary outcomes when we implemented our assumptions, or when we used data only from people who completed the study to that point. If there was a substantial difference, we would have reported and discussed the results but continued to employ our assumption.

Where assumptions had to be made regarding missing SDs (see Dealing with missing data), we would have compared the findings of the primary outcomes when we implemented our assumptions, or when we used data only from people who completed the study to that point. A sensitivity analysis would have been undertaken to test how prone the results were to change when completer-only data were compared with the imputed data using the above assumption. If there was a substantial difference, we would have reported and discussed the results 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 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 we used relevant data from these trials in the analysis.

4. Imputed values

We would have undertaken a sensitivity analysis to assess the effects of including data from trials where we used imputed values for the 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 listed above, we would not have pooled data from the excluded trials with the other trials contributing to the outcome, but would have presented them separately.

5. Fixed and random effects

We synthesised data using a random-effects model. However, we also synthesised data for the primary outcomes using a fixed-effect model to evaluate whether the greater weights assigned to larger trials with greater event rates altered the significance of the results, compared with the more evenly distributed weights in the random-effects model. If we had found differences, we would have reported them.

6. At least 20% of participants with schizophrenia and unclear proportion of people with schizophrenia

We intended to included studies where at least 20% of the participants were diagnosed with schizophrenia or schizophrenia-like disorders in a sensitivity analyses. If a paper had not reported the proportion of various diagnoses, we would have included it, but conducted a sensitivity analysis to test whether such a trial would influence the pooled results of primary outcomes. If inclusion did influence the results, we would not have included this trial but presented it separately.

RESULTS

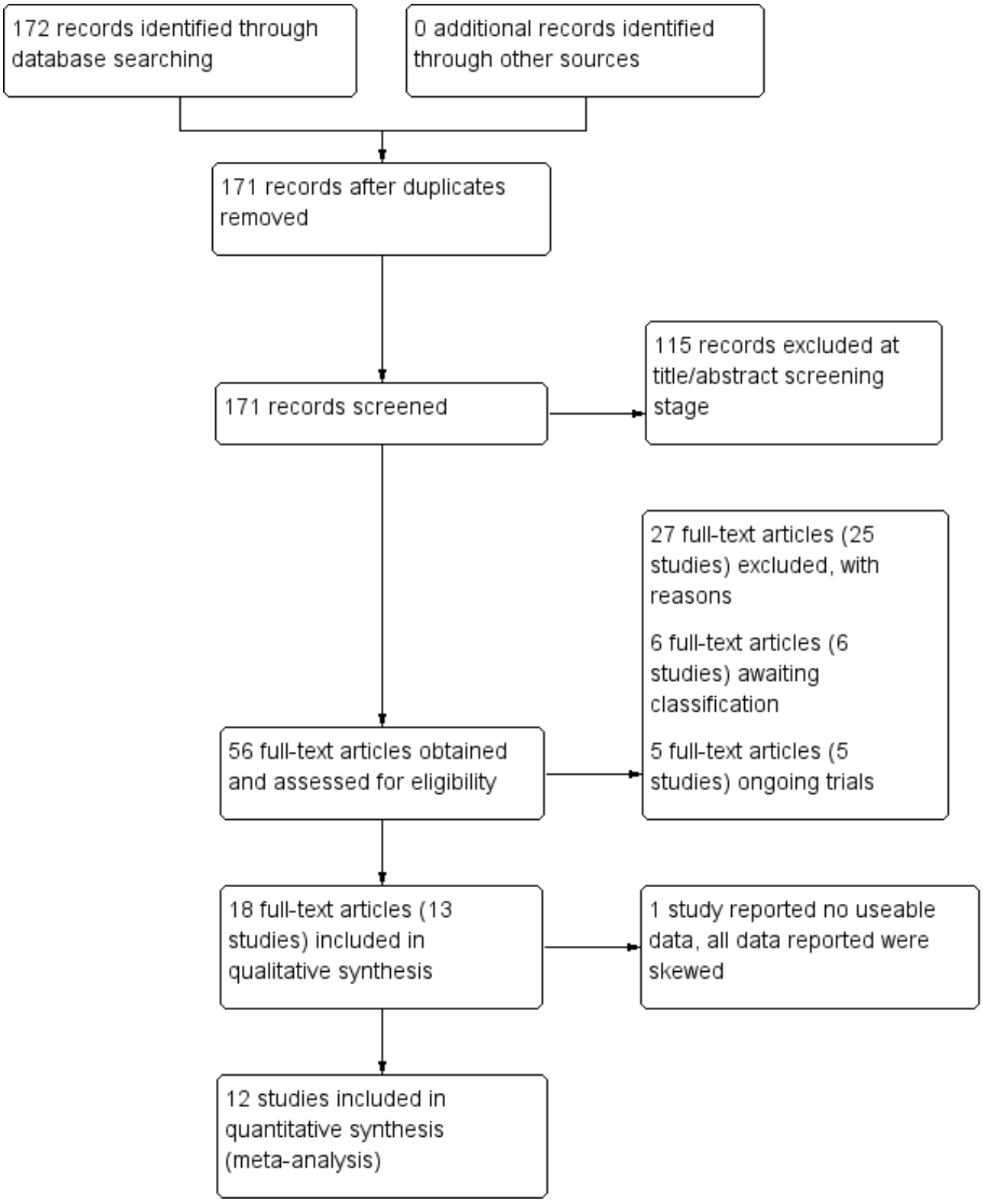
Description of studies

For a substantive description of studies, see the Characteristics of included studies; Characteristics of excluded studies; Characteristics of studies awaiting classification and Characteristics of ongoing studies tables.

Results of the search

The electronic search (4 July 2017) yielded 172 records of potentially eligible studies, after removal of duplicates, we screened 171 records. After checking titles and abstracts, we excluded 115 records and obtained 56 full-text papers for a second assessment. These publications consisted of 13 included studies with 18 references (Castelein 2008; Cook 2012b; Cook 2012a; Druss 2010; Eisen 2012; Goldberg 2013; Kelly 2014; Mahlke 2017; Qian 2015; Reynolds 2004; Rowe 2007; Sells 2008; Van Gestel-Timmermans 2012), 25 excluded studies with 27 references (Buchkremer 1995; Chen 2016; Chinman 2015; Corrigan 2017a; Corrigan 2017b; Craig 2004; Forchuk 2005; Gunter 1983; Hazell 2016; ISRCTN14282228; Kaplan 2011; Kaufmann 1995; Killackey 2013; Klein 1998; NCT02974400; O'Connell 2017; Rivera 2007; Rogers 2012; Salyers 2010; Segal 2010; Shahar 2006; Streicker 1984; Verhaegh 2006; Weissman 2005; Zhou 2016), six studies waiting classification (Robinson 2010; Daumit 2010; Kroon 2011; NCT00458094; NTR1166; Tondora 2010), and five ongoing studies (ACTRN1261200097; Chinman 2017; NCT01566513; NCT02958007; NCT02989805). We contacted authors of the following studies: Castelein 2008, Chinman 2015, Eisen 2012, Goldberg 2013, O'Connell 2017, Salyers 2010, Weissman 2005, and ACTRN1261200097 to clarify some obscure information. See Figure 1.

Figure 1. Study flow diagram.



Included studies

This review included 13 studies with 2479 participants. Comprehensive details are provided in the Characteristics of included studies table.

1. Design

1.1 Duration

The duration of the studies ranged from five weeks (Qian 2015) to 12 months (Mahlke 2017; Rowe 2007; Sells 2008). In seven studies, the study durations were medium term (one to six months) (Druss 2010; Eisen 2012; Goldberg 2013; Kelly 2014; Qian 2015; Reynolds 2004; Van Gestel-Timmermans 2012). The other studies were long term (longer than six months).

1.2 Unit of analysis

One study had three treatment groups (Eisen 2012). None of the studies were cross-over or cluster RCTs. The remaining studies were parallel randomised trials with two arms.

2. Participants

2.1 Age

All studies recruited adults (aged over 18 years). One study reported an age range between 30 and 60 years (Eisen 2012). Eleven studies reported the mean ages of participants, which were between 25.23 and 49.5 years. One study did not report ages of participants (Reynolds 2004).

2.2 Sex

Around half of the participants in the trials were men (1160/2479; 46.8%). Reynolds 2004 did not report gender of participants.

2.3 Diagnosis

Twelve studies recruited participants with a range of serious mental illness including bipolar disorder, major depression, depressive disorder, alcohol-use disorder, drug-use disorder, mood disorder or other disorders, but more than 20% of participants in these studies were diagnosed with schizophrenia or schizophrenia-like disorders. One study recruited only participants with schizophrenia (Qian 2015).

2.4 Exclusion criteria

Reported exclusion criteria of participants included: people aged less than 18 years old (Castelein 2008); people with drug or alcohol (or both) dependency or substance abuse (Castelein 2008; Mahlke 2017; Van Gestel-Timmermans 2012); possible language difficulties (Castelein 2008; Mahlke 2017; Van Gestel-Timmermans 2012); suicidal ideation (Van Gestel-Timmermans 2012); severe psychotic symptoms or not being psychiatrically stable (Castelein 2008; Qian 2015; Van Gestel-Timmermans 2012); unable to give informed consent or be hospitalised at start of the study (Kelly 2014); and people with dementia (Reynolds 2004). Other studies did not report the exclusion criteria (Cook 2012b; Cook 2012a; Druss 2010; Eisen 2012; Goldberg 2013; Rowe 2007; Sells 2008). For other details, see the Characteristics of included studies table.

2.5 Duration of illness

Five studies reported the duration of the illness (Castelein 2008; Cook 2012b; Cook 2012a; Mahlke 2017; Qian 2015), which ranged from 12 months to 13 years (Qian 2015). Other studies did not report the duration of illness.

2.6 Setting

Two studies recruited 323 participants from hospitals (Eisen 2012; Reynolds 2004), in which one study recruited participants from Veterans Hospital (Eisen 2012). The participants in Reynolds 2004 had been discharged from an inpatient facility. Four studies involved 1126 outpatients recruited from mental healthcare centres/administrations (Cook 2012b; Cook 2012a; Druss 2010; Goldberg 2013). Qian 2015 recruited their participants from community settings. Participants in Van Gestel-Timmermans 2012 and Mahlke 2017 were a mix of inpatients from hospital and outpatients from psychiatric care services and mental healthcare providers. The other four studies did not report the setting for participants (Castelein 2008; Kelly 2014; Rowe 2007; Sells 2008).

2.7 Country

Participants were recruited from Netherlands (439 participants) (Castelein 2008; Van Gestel-Timmermans 2012), USA (1699 participants) (Cook 2012b; Cook 2012a; Druss 2010; Eisen 2012; Goldberg 2013; Kelly 2014; Rowe 2007; Sells 2008;), UK (25 participants) (Reynolds 2004), Germany (216 participants) (Mahlke 2017), and China (100 participants) (Qian 2015).

3. Interventions

Of the 13 included studies, all compared peer support in addition to standard care versus standard care alone. For some of these studies, participants in the control group were assigned to a 'waiting-list' where they received standard care (Castelein 2008; Cook 2012b; Cook 2012a). Standard care in all included studies referred to continuation of the participants' usual medical or mental healthcare services. One study involved three arms in which they compared peer support with clinician support and with standard care separately (Eisen 2012). Details of studies are listed in the Characteristics of included studies table and the details of peer-support interventions are listed in Table 1.

4. Outcomes

4.1 General

Data were reported for service use, global state, mental state, behaviour, leaving the study early, functioning, peer outcomes, quality of life and economics. Details of scales used by the included trials to measure outcomes are given below.

4.2 Scales providing useable data

4.2.1 Global state scales

- Veterans RAND 12-Item Health Survey (VR-12) (Kazis 2017)

VR-12 assesses physical and mental health status rated on a 5-point response scale, ranging from 1, none of the time, to 5, all of the time. Total score ranges from 12 to 60 with higher score indicating better health status.

- Clinical Global Impression scale (CGI) (Busner 2007)

This a three-item scale used to measure the global severity and improvement of illness condition with two items (severity and improvement index) rated on a 7-point scale and one item (efficacy index) rated on a 4-pont scale. A higher score in severity and improvement indicates higher severity or more worsening of the clinical condition.

- Brief Symptom Inventory (BSI) (self-reported) (Derogatis 1993)

The BSI's Global Severity Index is designed to quantify a patient's severity of illness and provides a single composite score for

measuring the outcome of a treatment programme based on reducing symptom severity. Respondents are asked how much they were bothered in the past week by 53 symptoms with a 5-point response scale ranging from 'not at all' to 'extremely'.

4.2.2 Mental state scales

- Rogers Empowerment Scale (RES) (Rogers 1997)

The RES comprises 28 items encompassing self-efficacy, self-esteem, perceived power, community activism, righteous anger and optimism. The scores range from 28 to 112 with high score indicating more empowerment.

- Dutch Empowerment Scale (DES) (Boevink 2009)

Th DES consists of 40 items with a 5-point Likert scale ranging from 1, strongly disagree, to 5, strongly agree.

- State Hope Scale (SHS) (Snyder 1991)

The SHS is an instrument designed to measure hope as a cross-situational long-term trait in general populations. Twelve items are rated on a 4-point response scale ranging from 'definitely false' to 'definitely true' and summed to produce a total score. Two subscales measure belief in one's capacity to initiate and sustain actions (agency) and ability to generate routes by which goals may be reached (pathways).

- Herth Hope Index (HHI) (Herth 1992)

The HHI consists of 12 items rated on a 4-point linked scale ranging from 1, strongly disagree, to 4, strongly agree. The total score ranges from 12 to 48 with higher score indicating high level of hope.

- Rosenberg Scale (RS) (Rosenberg 1965)

The RS is used to assess self-esteem and has two subscales: positive and negative self-esteem. The total score ranges from 10 to 40 with higher score indicating higher level of self-esteem.

4.2.3 Behaviour scales

- Patient Activation Scale (PAS) (Hibbard 2004)

The PAS reflects an person's perceived ability to manage his or her illness and to act as an effective patient. It includes two subscales: activation levels and approach to health care. Higher scores reflect greater activation. This construct is measured using the 13-item Patient Activation Measure and is calculated on a 0 to 100 score, with 100 as the highest possible degree of activation.

- Recovery Assessment Scale (RAS) (Giffort 1995)

The RAS comprises 41 items rated on a 5-point scale from 'strongly agree' to 'strongly disagree', the RAS conceptualises recovery along multiple components. In addition to a total score, subscales measure personal confidence, willingness to ask for help, goal orientation, reliance on others and having tolerable levels of symptoms.

- Instrument to Measure Self-Management (IMSM) (Lorig 1996)

The IMSM includes six subscales: healthy eating, physical activity, accessing social support, behavioural and cognitive symptom management, making better use of health care and general self-management behaviours. The subscale scores range from 0 to 5, with higher scores indicating greater frequency.

- Brashers' Patient-Self-Advocacy Scale (PSA, self-reported) (Brashers 1999)

The Brashers' PSA is an instrument designed to measure a person's propensity to engage in self-activism during healthcare encounters. The study employs the 18-item instrument in which statements are rated on a 5-point response scale ranging from 'strongly agree' to 'strongly disagree', and meaned to produce a total score and three subscale scores.

- Self-Management/Self-Efficacy Scale (SMSES) (self-reported) (Lorig 1996)

The SMSES is an 18-item scale and includes six subscales: healthy eating, physical activity, accessing social support, behavioural and cognitive symptom management, making better use of health care (including preparing questions for medical providers to discuss medication concerns) and general self-management behaviours (use of action planning, brainstorming and problem solving). Items are scored on a Likert scale reflecting frequency; scores range from 1, never, to 5, always.

- Mental Health Confidence Scale (MHCS) (Carpinello 2000)

The MHCS is used to assess self-efficacy and is a 16-item scale with three factors: optimism, coping and advocacy. The sum of the items provides the total score, ranging from 16 to 96 with higher scores indicating more empowerment.

- General Self-Efficacy Scale (GSE) (Schwarzer 1995)

The GSE is a 10-item psychometric scale that is designed to assess optimistic self-beliefs to cope with a variety of difficult demands in life. Higher score indicates better self-efficacy.

4.2.4 Functioning

- Global Assessment of Functioning (GAF) (Aas 2010)

The GAF scale is used to rate how serious the mental illness affects a person's day-to-day life functioning on a scale of 0 to 100. Scores range from 1, severely impaired, to 100, extremely high functioning, with higher score indicating better functioning in daily activities.

- Colorado Client Assessment Record (CCAR) (Ellis 1984)

The CCAR is used for people with chronic mental illness and programme evaluation. It measures cognitive, social and role function, which is frequently impaired by chronic mental illness in diverse psychiatric diagnostic groups.

- Short-Form Health Survey (SF-12) (Ware 1996)

The 12-item Short-Form Health Survey is used to assess general health functioning, physical functioning and emotional well-being. Higher scores indicated better functioning. Possible subscale scores range from 0 to 100. The SF-36 was also used to measure health-related quality of life (McHorney 1993).

4.2.5 Peer support scales

- Personal Network Questionnaire (PNQ, self-reported) (Castelein 2008)

The PNQ is used to measure the size and content of the social network asking for information on the frequency of contacts with named family, friends and members of the peer support group.

- Barrett-Lennard Relationship Inventory (BLRI, self-reported) (Barrettlennard 1962)

The BLRI is a 64-item client questionnaire designed to gauge dimensions of the client–provider relationship relevant to favourable therapeutic change. Respondents rate agreement with items on a 6-point scale, ranging from 1, definitely false, to 6, definitely true.

- The Social Support List (SSL) (Bridges 2002)

The SSL measures positive social interactions and discrepancies between the support people want and what they actually receive. The SSL consisted of six subscales: everyday emotional support, emotional support with problems, esteem support, instrumental support, social companionship and informative support. The total score for positive interactions ranged from 34 to 136 and the total score for discrepancies ranged from 34 to 201. HIgher scores on interaction indicated more support. Higher scores on discrepancy indicated a greater deficit in desired support. The 'negative interactions' on a 7-item subscale ranged from 7 to 32 with higher scores indicating more negative interaction.

- The Medical Outcomes Study Social Support Survey (MOSSSS) (Sherbourne 1991)

The MOSSSS includes 19 items measuring emotional and informational support, tangible support, affectionate support and positive social interaction.

4.2.6 Quality of life scales

- EuroQol: Five Dimensions (EQ5D)/EQ-VAS (Balestroni 2012)

The EQ5D is a standardised instrument developed by the EuroQol group as a measure of health-related quality of life in different health conditions. It consists of a descriptive system of health status and EQ-VAS (0 to 100). The descriptive system comprises five dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression, rating on a 3-level response scale from 3, no problem, to 1, extreme problem. The EQ-VAS identifies one's self-rated health on a vertical, visual analogue scale (VAS) with the endpoints from 100, the best imaginable health state, to 0, the worst imaginable health state; and a higher score indicates better health status.

- General Quality of Life Inventory (GQOLI) (this scale is in data analyses but not described here) (Li 1997)

The GQOLI measures the perceived quality of life of people with different health conditions (Li 1997). This scale consists of 74 items, assessing four dimensions of quality of life: physical health, psychological health, social functioning and living conditions. Each item is rated on a 5-point scale, with high score indicating a better quality of life.

- Manchester Short Assessment of Quality of Life (MSAQOL)

Quality of life is assessed with 12 subjective items of the MSAQOL (Priebe 1999). The items use a 7-point Likert scale, from 1, could not be worse, to 7, could not be better. Higher scores indicate higher quality of life.

- World Health Organization Quality of Life (WHOQOL) (WHOQOL Group 1998)

The WHOQOL is a widely used quality of life instrument, with 100 items measuring four domains of well-being: physical, psychological, social and environment. Two additional items focus on the overall 'quality of life' and 'general health'. Scores on these four domains and the additional items can be combined to create an overall score of quality of life (ranging from 18 to 90). Higher scores indicating higher quality of life.

WHOQOL-BREF has been modified from the WHOQOL (WHOQOL Group 1998) to provide a short-form quality of life assessment with 26 items measuring four domains: physical health, psychological health, social relationships and environment, one item measuring overall quality of life, and another item measuring general health. The items use a 5-point Likert scale, from 1, not at all/very poor/very dissatisfied/never, to 5, completely/very good/very satisfied/extremely. Possible score range from 0 to 100 for each domain, with higher scores indicating high quality of life.

- Quality of Life Brief Version (QOLI-BREF) (Lehman 1994)

QOLI-BREF is derived from the QOLI-Full Version and measures both objective quality of life (what people do and experience) and subjective quality of life (what people feel about these experiences). It consists of 45 items, measuring eight domains: living situation, daily activities and functioning, family relations, social relations, finances, work and school, legal and safety issues, and health. Higher scores indicating higher quality of life.

4.3 Other scales

Other scales were used to measure outcomes but data reported from these scales were skewed and we could not use in data analyses.

- Addiction Severity Index (ASI) (Mclellan 1980)

The ASI is a structured interview to assess the degree of potential treatment barriers across domains typically affected by alcohol- and drug-use disorders. Higher score indicates greater problem.

- Behaviour and Symptom Identification Scale (BASIS-24) (Cameron 2007)

The revised 24-item BASIS assesses depression and functioning, difficulty in interpersonal relationships, self-ham, emotional lability, psychotic symptoms, substance abuse and overall mental health. The score ranges from 0 to 4, with higher values indicating greater symptom severity.

- Loneliness scale (Jonggierveld 1985)

The Loneliness Scale consists of 11 items with a 5-point Likert scale ranging from 1, yes, for sure, to 5, no, certainly not.

Studies awaiting classification

Six studies are awaiting classification due to insufficient characteristics information. We contacted authors of these studies for clarification, however, only one study author replied our email.

See Characteristics of studies awaiting classification for more details.

Ongoing studies

We identified five ongoing studies. Two started in 2012 (ACTRN1261200097; NCT01566513), we contacted the authors and both replied stating that they are analysing the results and working on the publication. Chinman 2017 started in 2016, NCT02989805 started in 2017 and NCT02958007 is not yet recruiting. Participants recruited in three studies were aged over 18 years (ACTRN1261200097; NCT01566513; NCT02989805). Chinman 2017 recruited some participants aged under 18 years and NCT02958007 recruited participants aged over 50 years. Diagnoses of participants include serious mental illness (NCT01566513); mental or physical illness (Chinman 2017; NCT02958007; NCT02989805);, or a range of disorders/ auditory verbal hallucination, schizophrenia, psychosis (ACTRN1261200097). The intervention groups in these studies all included a peer specialist who had personal live experience of hearing voices themselves or was trained in Intentional Peer Support.

Excluded studies

We excluded 25 studies with reasons listed in the Characteristics of excluded studies table.

Risk of bias in included studies

The details of the assessments are available in the 'Risk of bias' table corresponding to each study in the Characteristics of included studies table, and are also presented in the 'Risk of Bias' graph in Figure 2 and 'Risk of Bias' Summary in Figure 3.

Figure 2. Risk of bias graph: review authors' judgements about each risk of bias item presented as percentages across all included studies.

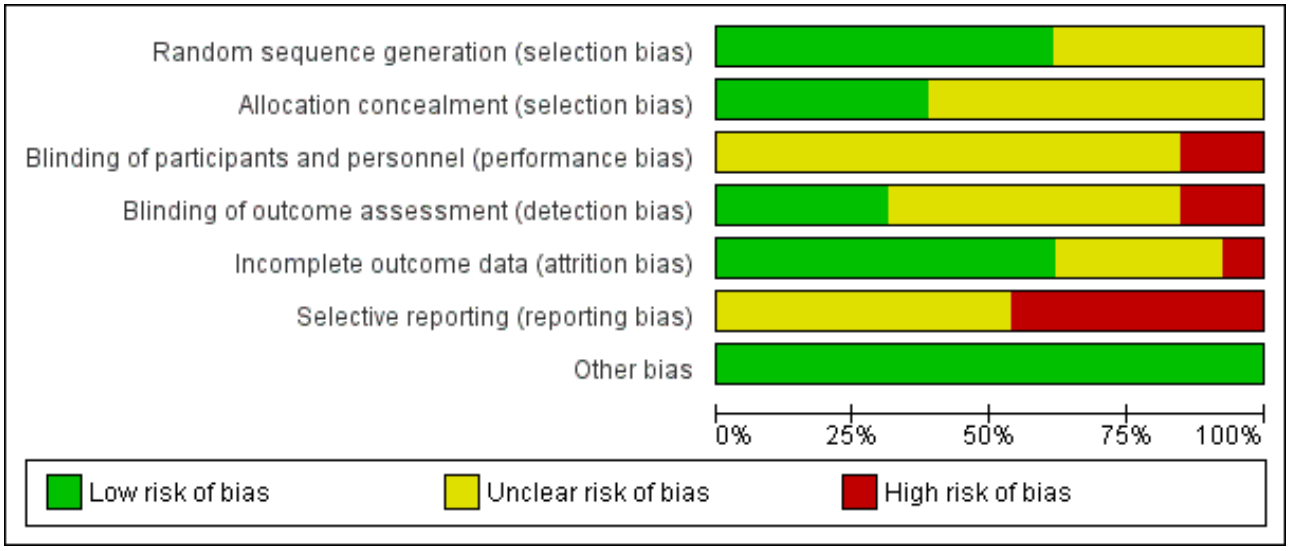
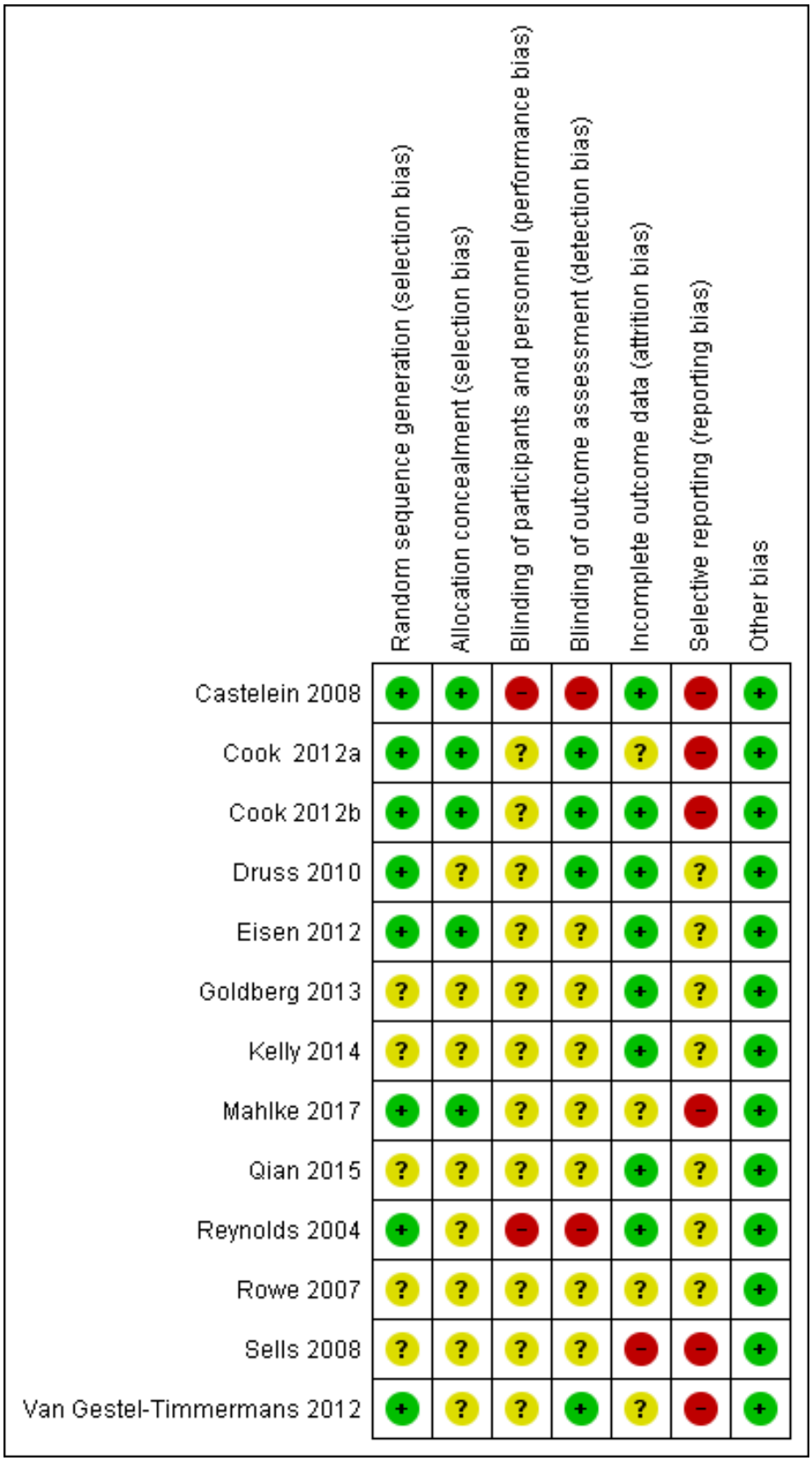


Figure 3. Risk of bias summary: review authors' judgements about each risk of bias item for each included study.



Allocation

All 13 included studies reported some form of randomisation. Eight of 13 studies (61.5%) were at low risk of selection bias as they reported adequate sequence generation (Castelein 2008; Cook 2012b; Cook 2012a; Druss 2010; Eisen 2012; Mahlke 2017; Reynolds 2004; Van Gestel-Timmermans 2012). The method used to generate the allocation sequence were such as drawing lots (Van Gestel-Timmermans 2012), using random block number (Castelein 2008; Mahlke 2017), or computerised randomisation program (Cook 2012b; Cook 2012a; Druss 2010; Reynolds 2004). The remaining studies proving insufficient information to rate this bias ('unclear').

Five studies ensured allocation concealment by using sealed envelopes (Castelein 2008; Eisen 2012; Mahlke 2017), computerised program (Cook 2012b; Cook 2012a). The remaining studies provided insufficient information and were at unclear risk of bias.

Blinding

Blinding of participants and personnel

None of the studies were at low risk of bias for blinding of participants and personnel. Two studies claimed that the personnel or participants were not blinded to the assignments, and thus were assessed at high risk for this bias (Castelein 2008; Reynolds 2004). Other studies did not provide adequate information to assess how blinding of participants and outcome assessors was maintained ('unclear'). Due to the nature of the intervention, it is understood that it would not be possible to blind recipients and providers of peer support services.

Blinding of outcome assessment

Four studies were at low risk of bias for blinding of outcome assessors (Cook 2012b; Cook 2012a; Druss 2010; Van Gestel-Timmermans 2012). An independent investigator who was blinded to the treatment performed the measurements. Two studies claimed that the assessors were not blinded to the treatment sequence or participants, and thus were at high risk (Castelein 2008; Reynolds 2004). All other studies were at unclear risk for this bias.

Incomplete outcome data

Seven studies were at low risk because the authors did an analysis for attrition or the numbers leaving early were small and balanced to groups (Castelein 2008; Cook 2012b; Druss 2010; Eisen 2012; Goldberg 2013; Kelly 2014; Reynolds 2004). Another study was at low risk of bias on this domain in that all participants completed the trial (Qian 2015). One study was at high risk of attrition bias due to a high attrition rate (more than 40%) (Sells 2008). Neither of the studies undertook analysis to account for this attrition (Qian 2015; Sells 2008). Other studies were at unclear risk because there was a moderate attrition rate, reasons for loss were not reported or not enough information was reported for us to assess.

Selective reporting

Six studies were at high risk for reporting bias because some protocol outcomes were not reported (Castelein 2008; Cook 2012b; Cook 2012a; Mahlke 2017; Sells 2008; Van Gestel-Timmermans 2012). The other studies were at unclear risk for reporting bias because the protocols of the studies were not available.

Other potential sources of bias

We identified no other potential sources of bias.

Effects of interventions

See: **Summary of findings for the main comparison** Peer support plus standard care versus standard care for people with schizophrenia or similar serious mental illness; **Summary of findings 2** Peer support plus standard care versus clinician-led support plus standard care for people with schizophrenia or similar serious mental illness

1. Peer support plus standard care versus standard care alone

1.1 Service use: 1a. Hospital admission – medium term

There was no clear difference in hospital admission data between peer support and standard care (RR 0.44, 95% CI 0.11 to 1.75; participants = 19; studies = 1, very low-quality evidence; Analysis 1.1) (Reynolds 2004).

1.2 Service use: 1b. Hospital admission – duration of hospital stay (days) – long term

These data were skewed and are presented as 'Other data' (Analysis 1.2).

1.3 Service use: 2a. Clinically important engagement with services – medium term

Druss 2010 observed the number of participants with one or more primary care visits in each group. Goldberg 2013 reported the use of the emergency department for medical services (Analysis 1.3).

1.3.1 Use of emergency care

There was no clear difference between the peer support and standard care groups, with similar number of participants from each group using emergency care services at medium term (RR 0.39, 95% CI 0.11 to 1.32; participants = 57; studies = 1) (Goldberg 2013).

1.3.2 One or more primary care visit

There was a clear difference between peer support and standard care, with fewer people in the standard care group using primary care services at least once compared to participants in the peer support group in the medium term (RR 1.77, 95% CI 1.09 to 2.85; participants = 80; studies = 1) (Druss 2010).

1.4 Service use: 2b. Contact with services – medium term (skewed data)

Kelly 2014 reported medium-term data for mean number of visits to emergency care and mean number of routine care visits. These data were skewed and are presented as 'Other data' (Analysis 1.4).

1.5 Global state: 3a. General health – mean endpoint score (VR-12, high = good)

1.5.1 Medium term

There was no clear difference in global state endpoint scores measured by the VR-12 scale between the peer support and standard care groups (MD –0.02, 95% CI –3.96 to 3.92; participants = 158; studies = 1; Analysis 1.5) (Eisen 2012).

1.6 Global state: 3b. Severity of illness – mean endpoint score (BSI, high = good)

Cook 2012b (555 participants) measured severity of illness using BSI scale (Analysis 1.6).

1.6.1 Medium term

There was a difference between peer support and standard care groups at medium term with clear positive effect for peer support (MD –0.13, 95% CI –0.25 to –0.01; participants = 458; studies = 1).

1.6.2 Long term

At long term, however, there was no clear difference in endpoint scores on the BSI (MD 0.00, 95% CI –0.11 to 0.11; participants = 440; studies = 1).

1.7 Global state: 3c. Severity of illness – mean endpoint score (CGI, high = poor)

Mahlke 2017 (216 participants) measured severity of illness by the CGI scale (Analysis 1.7).

1.7.1 Medium term

There was a difference between peer support and standard care groups at medium term with clear positive effect for peer support (MD –0.30, 95% CI –0.53 to –0.07; participants = 216; studies = 1).

1.7.2 Long term

However, at long term, there was a clear difference between the treatment groups with positive effect for standard care (MD 0.40, 95% CI 0.15 to 0.65; participants = 216; studies = 1).

1.8 Global state: 4. Compliance with medication (skewed data)

Data reported for this outcome were skewed and presented as 'Other data' tables (Analysis 1.8).

1.9 Adverse event: 1. Death: all cause – long term

There was no clear difference in number of death between peer support and standard care in the long term (RR 1.52, 95% CI 0.43 to 5.31; participants = 555; studies = 1) (Cook 2012b).

1.10 Mental state 1a. Specific: various aspects – mean endpoint score (various scales, high = good) – medium term

Several studies reported medium-term data for empowerment and hope using a range of scales (Analysis 1.10).

1.10.1 Empowerment (RES)

Mean empowerment endpoint scores on the RES showed no clear difference between peer support and standard care for assertiveness at medium term (MD –0.95, 95% CI –3.30 to 1.40; participants = 158; studies = 1) (Eisen 2012).

1.10.2 Empowerment (DES)

Medium-term empowerment scores measured by the DES showed a clear difference in participants 'assertiveness' between the treatment groups, with a positive effect for peer support (MD 0.19, 95% CI 0.05 to 0.33; participants = 220; studies = 1) (Van Gestel-Timmermans 2012).

1.10.3 Hope (SHS)

There was no clear difference in 'hope' scores by the SHS between peer support and standard care (MD 0.37, 95% CI –0.22 to 0.96; participants = 789; studies = 2) (Cook 2012b; Cook 2012a).

1.10.4 Hope (HHI)

There was a clear difference in 'hope' scores measured by the HHI between the treatment groups, favouring peer support (MD 0.24, 95% CI 0.11 to 0.37; participants = 217; studies = 1) (Van Gestel-Timmermans 2012).

1.11 Mental state 1b. Specific: various aspects – mean endpoint score (various scales, high = good) – long term

Four studies reported long-term 'hope' and 'self-esteem' scores . (Analysis 1.11) (Castelein 2008; Cook 2012b; Cook 2012a; Eisen 2012).

1.11.1 Hope (SHS)

There was no clear difference in 'hope' measured by the SHS between peer support and standard care at long term (MD 0.41, 95% CI –0.15 to 0.97; participants = 908; studies = 3) (Cook 2012b; Cook 2012a; Eisen 2012).

1.11.2 Self-esteem (RS)

There was no clear difference in self-esteem measured by the RS between the two treatment groups (MD 0.50, 95% CI –1.22 to 2.22; participants = 106; studies = 1) (Castelein 2008).

1.12 Mental state 1c. Specific: various aspects – mean endpoint score (SHS subscale, high = good)

Two studies reported subscale scores of the SHS scale for Hope agency and Hope pathways (Analysis 1.12) (Cook 2012b; Cook 2012a).

1.12.1 Hope agency – medium term

There was no clear difference in 'hope agency' measured by the SHS between peer support and standard care (MD 0.28, 95% CI –0.06 to 0.63; participants = 796; studies = 2) (Cook 2012b; Cook 2012a).

1.12.2 Hope pathways – medium term

There was no clear difference in 'hope pathways' measured by the SHS between peer support and standard care (MD 0.09, 95% CI –0.22 to 0.40; participants = 792; studies = 2) (Cook 2012b; Cook 2012a).

1.12.3 Hope agency – long term

There was a clear difference in 'hope agency' scores measured by the SHS, favouring peer support over standard care (MD 0.45, 95% CI 0.07 to 0.83; participants = 757; studies = 2) (Cook 2012b; Cook 2012a).

1.12.4 Hope pathways – long term

There was no clear difference in 'hope pathway' scores measured by the SHS between peer support and standard care (MD 0.17, 95% CI –0.14 to 0.48; participants = 755; studies = 2) (Cook 2012b; Cook 2012a).

1.13 Mental state: 1d. Specific: various aspects – mean endpoint score (various subscales) – skewed data

The studies reported a wide range of various other aspects of mental state using a range of scales. However, the reported data were skewed and we were unable to use them in meta-analyses. They are presented as 'Other data' (Analysis 1.13).

1.14 Behaviour: 1a. Specific: self-efficacy – mean endpoint score (various scales, high = good) – medium term

Three studies reported medium-term self-efficacy scores using different scales (Analysis 1.14) (Castelein 2008; Cook 2012b; Mahlke 2017).

1.14.1 PSA

There was no clear difference in self-efficacy scores measured by the PSA between peer support and standard care (MD 0.08, 95% CI –0.02 to 0.18; participants = 458; studies = 1) (Cook 2012b).

1.14.2 SMSES

At medium term, there was a clear difference in self-efficacy scores measured by the SMSES favouring peer support over standard care (MD 1.20, 95% CI 0.11 to 2.29; participants = 57; studies = 1) (Goldberg 2013).

1.14.3 MHCS

There was a positive effect in self-efficacy scores measured by the MHCS favouring peer support over standard care in the medium term (MD 0.31, 95% CI 0.07 to 0.55; participants = 221; studies = 1) (Van Gestel-Timmermans 2012).

1.14.4 GSE

Medium-term data found no clear difference in self-efficacy scores by the GSE between standard care and peer support groups in the medium term (MD 0.90, 95% CI –1.04 to 2.84; participants = 216; studies = 1) (Mahlke 2017).

1.15 Behaviour: 1b. Specific: self-efficacy – mean endpoint score (various scales, high = good) – long term

Three studies reported long-term self-efficacy scores using various scales (Analysis 1.15) (Castelein 2008; Cook 2012b; Mahlke 2017).

1.15.1 PSA

There was a positive effect in long-term endpoint scores by the PSA favouring peer support over standard care (MD 0.10, 95% CI 0.01 to 0.19; participants = 447; studies = 1) (Cook 2012b).

1.15.2 MHCS

There was no clear difference in long-term self-efficacy scores measured by the MHCS between peer support and standard care (MD 2.70, 95% CI –2.40 to 7.80; participants = 106; studies = 1) (Castelein 2008).

1.15.3 GSE

There was a positive effect for peer support with a clear difference in 'self-efficacy' endpoint scores on the GSE (MD 2.20, 95% CI 0.35 to 4.05; participants = 216; studies = 1) (Mahlke 2017).

1.16 Behaviour: 2. Specific: self-management – mean endpoint score (SMS, high = good)

1.16.1 Medium term

There was no clear difference in 'self-management behaviours' measured by the SMS between peer support and standard care in the medium term (MD 0.60, 95% CI –0.10 to 1.30; participants = 57; studies = 1; Analysis 1.16) (Goldberg 2013).

1.17 Behaviour 3. Specific: recovery – mean endpoint score (RAS, high = good)

Three studies used the RAS to measure 'recovery' (Analysis 1.17) (Cook 2012b; Eisen 2012; Goldberg 2013).

1.17.1 Medium term

There was no clear difference between peer support and standard care groups in the medium term (MD 2.69, 95% CI –0.82 to 6.20; participants = 557; studies = 3) (Cook 2012b; Eisen 2012; Goldberg 2013).

1.17.2 Long term

There was a clear difference between peer support and standard care groups with a positive effect for peer support in the long term (MD 4.16, 95% CI 1.16 to 7.16; participants = 318; studies = 1) (Cook 2012b).

1.18 Behaviour: 4a. Specific: various behaviours – mean endpoint score (PAS subscales, high = good) – medium term

Four studies used PAS subscales to measure participant's 'activation', 'approach to healthcare' and 'assertiveness' (Analysis 1.18).

1.18.1 Activation (patient)

There was no clear difference in 'patient activation' scores between peer support and standard care at medium term (MD 3.68, 95% CI –1.85 to 9.22; participants = 295; studies = 3) (Druss 2010; Eisen 2012; Goldberg 2013). Heterogeneity was high (Chi^2 = 10.16; degrees of freedom (df) = 2.0; P = 0.01; I^2 = 80%) with Eisen 2012 as the outlier, but source of this heterogeneity remained unclear.

1.18.2 Approach to healthcare

There was no clear difference in 'approach to healthcare' scores between peer support and standard care scores in the medium term (MD 2.10, 95% CI –0.83 to 5.03; participants = 57; studies = 1) (Goldberg 2013).

1.18.3 Assertiveness

There was no clear difference in 'assertiveness' scores measured by the PAS between peer support and standard care in the medium term (MD 0.08, 95% CI –0.06 to 0.22; participants = 458; studies = 1) (Cook 2012b).

1.19 Behaviour: 4b. Specific: various behaviours – mean endpoint score (PAS subscales, high = good) – medium term

1.19.1 Assertiveness

There was no clear difference in 'assertiveness' scores measured by the PAS subscale between peer support and standard care in the long term (MD 0.07, 95% CI –0.06 to 0.20; participants = 447; studies = 1; Analysis 1.19) (Cook 2012b).

1.20 Behaviour: 4c. Specific: various behaviours – mean endpoint score (various scales) – medium term

Three studies reported endpoint subscale scores at medium term for various behaviours using a range of scales (Analysis 1.20) (Cook 2012b; Cook 2012a; Goldberg 2013).

1.20.1 Goal orientation (RAS, high = good)

There was no clear difference in 'goal orientation' scores measured by the RAS between peer support and standard care (MD 0.72, 95% CI –0.09 to 1.53; participants = 343; studies = 1) (Cook 2012a).

1.20.2 Healthy eating (IMSM, high = good)

There was no clear difference in 'healthy eating' scores measured by the IMSM between peer support and standard care (MD 0.40, 95% CI –0.15 to 0.95; participants = 57; studies = 1) (Goldberg 2013).

1.20.3 Internal locus of control for health (MHLC, high = greater control)

Endpoint scores for 'internal locus of control for health' by the MHLC at medium term were clearly different, with a positive effect for peer support compared with standard care (MD 3.60, 95% CI 0.99 to 6.21; participants = 57; studies = 1) (Goldberg 2013).

1.20.4 Mindful non-adherence (PSA, high = non-adherence)

There was no clear difference in 'mindful non-adherence' scores measured by the PSA between peer support and standard care (MD 0.09, 95% CI –0.05 to 0.23; participants = 456; studies = 1) (Cook 2012b).

1.20.5 No symptom domination (RAS, high = good)

There was no clear difference in 'no symptom domination' scores measured by the RAS between peer support and standard care (MD 0.29, 95% CI –0.31 to 0.89; participants = 342; studies = 1) (Cook 2012a).

1.20.6 Personal confidence (RAS, high = good)

There was a clear difference in 'personal confidence' scores by the RAS between the treatment groups, favouring peer support over standard care (MD 1.59, 95% CI 0.30 to 2.88; participants = 343; studies = 1) (Cook 2012a).

1.20.7 Reliance on other (RAS, high = strong reliance)

Participants in the peer support groups had clearly 'higher reliance on others' scores measured by the RAS compared to those in the standard care group (MD 0.80, 95% CI 0.17 to 1.43; participants = 343; studies = 1) (Cook 2012a).

1.20.8 Willingness to ask for help (RAS, high = strong willingness)

The mean endpoint scores of the participants in the peer support group were clearly higher for 'willingness to ask for help' scores measured by the RAS (MD 0.44, 95% CI 0.01 to 0.87; participants = 343; studies = 1) (Cook 2012a).

1.21 Behaviour: 4d. Specific: various behaviours – mean endpoint score (various sub scales) – long term

Two studies reported long-term data from various behaviour sub scales (Analysis 1.21) (Cook 2012b; Cook 2012a).

1.21.1 Goal orientation (RAS, high = good)

There was no clear difference in endpoint scores for 'goal orientation' measured by the RAS between peer support and standard care groups (MD 0.61, 95% CI –0.19 to 1.41; participants = 320; studies = 1) (Cook 2012a).

1.21.2 Mindful non-adherence (PSA, high = non-adherence)

The mean endpoint scores for 'mindful non-adherence measured by the PSA of the participants in the peer support group were clearly higher compared with standard care (MD 0.17, 95% CI 0.03 to 0.31; participants = 447; studies = 1) (Cook 2012b).

1.21.3 No symptom domination (RAS, high = good)

The mean endpoint scores for 'no symptom domination' measured by the RAS of the participants in the peer support group were clearly higher compared with standard care (MD 0.77, 95% CI 0.15 to 1.39; participants = 319; studies = 1) (Cook 2012b).

1.21.4 Personal confidence (RAS, high = good)

There was a clear difference in endpoint scores for 'personal confidence' measured by the RAS between peer support and standard care groups with a positive effect for peer support (MD 1.90, 95% CI 0.61 to 3.19; participants = 319; studies = 1) (Cook 2012a).

1.21.5 Reliance on others (RAS, high = strong reliance)

There was no clear difference in endpoint scores for 'reliance on others' by the RAS between peer support and standard care groups (MD 0.41, 95% CI –0.21 to 1.03; participants = 320; studies = 1) (Cook 2012a).

1.21.6 Willingness to ask for help (RAS, high = stronger willingness to seek help)

The mean endpoint scores for 'willingness to ask for help' measured by the RAS of the participants in the peer support group were clearly higher for this measure (MD 0.53, 95% CI 0.06 to 1.00; participants = 320; studies = 1) (Cook 2012a).

1.22 Behaviour: 5. Specific: alcohol or drug use (various scales) (skewed data)

Two studies reported skewed data for alcohol or drug use. These are presented as 'Other data' (Analysis 1.22) (Eisen 2012; Rowe 2007).

1.23 Leaving the study early

Eight studies reported data for numbers leaving the study early (Analysis 1.23) (Castelein 2008; Cook 2012b; Druss 2010; Goldberg 2013; Kelly 2014; Mahlke 2017; Reynolds 2004; Van Gestel-Timmermans 2012).

1.23.1 Medium term

At medium term, data from six studies showed clearly more people left the studies early from standard care groups compared with numbers leaving from peer support groups (RR 0.66, 95% CI 0.51 to 0.87; participants = 741; studies = 6) (Druss 2010; Goldberg 2013; Kelly 2014; Mahlke 2017; Reynolds 2004; Van Gestel-Timmermans 2012).

1.23.2 Long term

At long term, three studies provided data and the positive effect for peer support was no longer evident with no clear difference in numbers of participants leaving early (RR 1.34, 95% CI 0.19 to 9.22; participants = 877; studies = 3). This subgroup had important levels of heterogeneity (Chi^2 = 53.42; df = 2.0; P = 0.001; I^2 = 96%). The heterogeneity was due to Cook 2012b, and may be due to different levels of facilitation of peer support provided for the intervention group by different study sites and much varied attendance to peer support groups. The control group also reported participation in similar support groups in routine care (Cook 2012b).

When Cook 2012b was removed, homogeneity was restored and there was a positive effect for peer support with clearly fewer participants leaving early from the peer support groups (RR 0.53, 95% CI 0.37 to 0.75; participants = 322; studies = 2).

1.24 Functioning: 1a. General: mean total endpoint score (various scales, high = good) – medium term

Three studies reported endpoint scores for general functioning at medium term. They used three different scales to measure general functioning (Analysis 1.24) (Goldberg 2013; Mahlke 2017; Reynolds 2004).

1.24.1 CCAR

There was no clear difference in general functioning measured by the CCAR between treatment groups at medium term (MD 0.59, 95% CI –0.93 to 2.11; participants = 19; studies = 1) (Reynolds 2004).

1.24.2 GAF

There was a clear difference in endpoint scores measured by the GAF favouring peer support, between the treatment groups (MD 4.10, 95% CI 0.34 to 7.86; participants = 216; studies = 1) (Mahlke 2017).

1.24.3 SF-12

There was no clear difference in general functioning measured by the SF-12 between treatment groups at medium term (MD 2.60, 95% CI –3.19 to 8.39; participants = 57; studies = 1) (Goldberg 2013).

1.25 Functioning: 1b. General: mean total endpoint score (various scales, high = good) – long term

1.25.1 GAF

At long term, there was no difference in general functioning measured by the GAF between peer support and standard care groups (MD –3.90, 95% CI –7.81 to 0.01; participants = 216; studies = 1) (Mahlke 2017).

1.26 Functioning: 2a. Specific: various aspects – mean endpoint score (CCAR subscales, high = good) – medium term

One study reported medium-term data for various aspects of functioning, measured by the CCAR (Analysis 1.26) (Reynolds 2004).

1.26.1 Cognitive functioning

There was no clear difference in 'cognitive' scores between peer support and standard care groups (MD 0.68, 95% CI –0.83 to 2.19; participants = 25; studies = 1) (Reynolds 2004).

1.26.2 Interpersonal functioning

There was no clear difference in 'interpersonal' scores between peer support and standard care groups (MD 0.62, 95% CI –0.65 to 1.89; participants = 25; studies = 1) (Reynolds 2004).

1.26.3 Physical functioning

There was no clear difference in 'physical' scores between peer support and standard care groups (MD 0.38, 95% CI –1.05 to 1.81; participants = 19; studies = 1) (Reynolds 2004).

1.26.4 Societal role functioning

There was no clear difference in 'societal role' scores between peer support and standard care groups (MD 1.02, 95% CI –0.44 to 2.48; participants = 25; studies = 1) (Reynolds 2004).

1.27 Functioning: 2b. Specific: various aspects – mean endpoint score (SF-12 subscales, high = good) – medium term

Goldberg 2013 reported medium-term data for emotional well-being and physical functioning measured by the SF-12 (Analysis 1.27).

1.27.1 Emotional well-being

There was no clear difference in 'emotional well-being' scores between peer support and standard care groups (MD 3.00, 95% CI –2.76 to 8.76; participants = 57; studies = 1).

1.27.2 Physical functioning

There was no clear difference in 'physical' scores between peer support and standard care groups (MD 3.00, 95% CI –2.82 to 8.82; participants = 57; studies = 1).

1.28 Functioning: 3. Specific: daily living – mean endpoint score (CCAR, high = good) – medium term (skewed data)

One study reported skewed data for daily living, which are presented as 'Other data' (Analysis 1.28) (Reynolds 2004).

1.29 Functioning: 4. Specific: self-management – mean endpoint score (IMSM, high = good) (skewed data)

One study reported skewed data for 'self-management', which are presented as 'Other data' (Analysis 1.29) (Goldberg 2013).

1.30 Functioning: 5. Specific: contact with justice system – criminal justice charges (skewed data)

One study reported skewed data for 'criminal justice charges', which are presented as 'Other data' (Analysis 1.30) (Rowe 2007).

1.31 Peer outcomes: 1a. Impact on participant and peer supporter: improved peer contact – mean endpoint score (PNQ, high = good) – long term

There was a clear effect for 'improved peer contact', favouring peer support for (RR 1.85, 95% CI 1.14 to 3.00; participants = 106; studies = 1; Analysis 1.31) (Castelein 2008).

1.32 Peer outcomes: 1b. Impact on participant and peer supporter: negative aspects – mean endpoint score (BLR subscales, high = true) – medium term

One study reported endpoint BLR subscale scores (Analysis 1.32) (Sells 2008).

1.32.1 Negative empathy

There was no difference between treatment groups for negative empathy (MD –0.32, 95% CI –0.66 to 0.02; participants = 105; studies = 1) (Sells 2008).

1.32.2 Negative regard

There was no clear difference between treatment groups for negative regard (MD –0.27, 95% CI –0.65 to 0.11; participants = 105; studies = 1) (Sells 2008).

1.32.3 Negative overall relationship

There was no clear difference between treatment groups for negative overall relationship (MD –0.19, 95% CI –0.48 to 0.10; participants = 105; studies = 1) (Sells 2008).

1.32.4 Negative unconditionality

There was no clear difference between treatment groups for negative unconditionality (MD 0.01, 95% CI –0.32 to 0.34; participants = 105; studies = 1) (Sells 2008).

1.33 Peer outcomes: 1c. Impact on participant and peer supporter: positive aspects – mean endpoint score (BLRI, high = true) – medium term

One study reported endpoint BLR subscale scores (Analysis 1.33) (Sells 2008).

1.33.1 Positive empathy

There was a clear difference, favouring peer support, for positive empathy (MD 0.49, 95% CI 0.13 to 0.85; participants = 105; studies = 1).

1.33.2 Positive regard

There was a clear difference, favouring peer support, for positive regard (MD 0.44, 95% CI 0.08 to 0.80; participants = 105; studies = 1).

1.33.3 Positive overall relationship

There was a clear difference, favouring peer support, for positive overall relationship (MD 0.43, 95% CI 0.16 to 0.70; participants = 105; studies = 1).

1.33.4 Positive unconditionality

There was a clear difference, favouring peer support, for positive unconditionality (MD 0.33, 95% CI 0.05 to 0.61; participants = 105; studies = 1).

1.34 Peer outcomes: 1d. Impact on participant and peer supporter: various aspects – mean endpoint score (SSL subscales, high = increased need for support) – long term

One study reported SSL subscale scores (Analysis 1.34) (Castelein 2008).

1.34.1 Negative interaction for esteem

There was a clear difference, favouring peer support, for negative interaction for esteem (MD –1.20, 95% CI –2.38 to –0.02; participants = 106; studies = 1).

1.34.2 Social support for positive interactions

There was no clear difference between treatment groups for social support for positive interactions (MD –1.50, 95% CI –7.58 to 4.58; participants = 106; studies = 1).

1.34.3 Social support for discrepancies

There was no clear difference between treatment groups for social support for discrepancies (MD 5.60, 95% CI –0.51 to 11.71; participants = 106; studies = 1)

1.35 Peer outcomes: 1e. Impact on participant and peer supporter: social support – mean endpoint score (MOSSSS, high = good)

1.35.1 Medium term

There was no clear difference between treatment groups for social support (MD –1.12, 95% CI –6.26 to 4.02; participants = 158; studies = 1) (Analysis 1.35) (Eisen 2012).

1.36 Peer outcomes: 1f. impact on the service user and peer supporter: accessing social support (IMSM, high = greater amount of support obtained) – medium term

These data were skewed and were presented as 'other data' (Analysis 1.36) (Goldberg 2013).

1.37 Peer outcomes: 2a. Quality of life for participant and peer supporter: overall – mean total endpoint score (various scales, high = good) – medium term

Five trials reported overall quality of life scores at medium-term follow-up, using a variety of scales (Analysis 1.37) (Castelein 2008; Cook 2012b; Mahlke 2017; Qian 2015; Van Gestel-Timmermans 2012).

1.37.1 EQ5D-Index

There was no clear difference between treatment groups for overall quality of life measured by the EQ5D-Index in the medium term (MD 0.40, 95% CI –4.52 to 5.32; participants = 216; studies = 1) (Mahlke 2017).

1.37.2 EQ5D-VAS

There was no clear difference between treatment groups for overall quality of life measured by the EQ5D-VAS in the medium term (MD 3.20, 95% CI –2.77 to 9.17; participants = 216; studies = 1) (Mahlke 2017).

1.37.3 GQOLI-74

There was no clear difference between treatment groups for overall quality of life measured by the GQOLI-74 in the medium term (MD 40.34, 95% CI 32.70 to 47.98; participants = 100; studies = 1) (Qian 2015).

1.37.4 MSAQOL

There was no clear difference between treatment groups for overall quality of life measured by the MSAQOL in the medium term (MD 0.24, 95% CI –0.04 to 0.52; participants = 208; studies = 1) (Van Gestel-Timmermans 2012).

1.37.5 WHOQOL

There was no clear difference between treatment groups for overall quality of life measured by the WHOQOL in the medium term (MD

1.00, 95% CI –2.82 to 4.82; participants = 106; studies = 1) (Castelein 2008).

1.37.6 WHOQOL-BREF

There was no clear difference between treatment groups for overall quality of life measured by the WHOQOL-BREF in the medium term (MD 0.20, 95% CI –0.33 to 0.73; participants = 458; studies = 1) (Cook 2012b).

1.38 Peer outcomes: 2b. Quality of life for participant and peer supporter: overall – mean total endpoint score (various scales, high = good) – long term

Three trials reported overall quality of life scores at long term (Analysis 1.38) (Castelein 2008; Cook 2012b; Mahlke 2017).

1.38.1 EQ5D-Index

There was no clear difference between treatment groups for overall quality of life measured by the EQ5D-Index in the long term (MD 3.30, 95% CI –1.83 to 8.43; participants = 216; studies = 1) (Mahlke 2017).

1.38.2 EQ5D-VAS

There was no clear difference between treatment groups for overall quality of life measured by the EQ5D-VAS in the long term (MD 5.00, 95% CI –0.67 to 10.67; participants = 216; studies = 1) (Mahlke 2017).

1.38.3 WHOQOL-BREF

There was a clear difference, favouring peer support, for overall quality of life measured by the WHOQOL-BREF in the long term (MD 0.70, 95% CI 0.15 to 1.25; participants = 431; studies = 1) (Cook 2012b).

1.38.4 WHOQOL

There was no clear difference between treatment groups for overall quality of life measured by the WHOQOL in the long term (MD 1.70, 95% CI –2.32 to 5.72; participants = 106; studies = 1) (Castelein 2008).

1.39 Peer outcomes: 3a. Quality of life for participant and peer supporter: specific aspects – mean endpoint score (GQLI-74 subscales, high = good) – medium term

One study reported on specific aspects of quality of life using GQLI-74 (Analysis 1.39) (Qian 2015).

1.39.1 Mental health

There was a clear difference, favouring peer support, for mental health measured by GQLI-74 in the medium term (MD 16.95, 95% CI 13.34 to 20.56; participants = 100; studies = 1).

1.39.2 Physical quality of life

There was no clear difference between treatment groups for physical quality of life measured by GQLI-74 in the medium term (MD 1.43, 95% CI –2.31 to 5.17; participants = 100; studies = 1).

1.39.3 Physical health

There was a clear difference, favouring peer support, for physical health measured by GQLI-74 in the medium term (MD 15.08, 95% CI 11.29 to 18.87; participants = 100; studies = 1).

1.39.4 Social function

There was a clear difference, favouring peer support, for social function measured by GQLI-74 in the medium term (MD 15.87, 95% CI 12.66 to 19.08; participants = 100; studies = 1).

1.40 Peer outcomes: 3b. Quality of life for participant and peer supporter: specific aspects – mean endpoint score (QOLI-BREF) subscales, high = good) – medium term

One study reported specific aspects of quality of life using QOLI-BREF (Analysis 1.40) (Reynolds 2004).

1.40.1 Amount of time spent with others

There was no clear difference between treatment groups for amount of time spent with others measured by QOLI-BREF in the medium term (MD 0.04, 95% CI –1.24 to 1.32; participants = 19; studies = 1).

1.40.2 General life satisfaction

There was no clear difference between treatment groups for general life satisfaction measured by QOLI-BREF in the medium term (MD –0.04, 95% CI –1.25 to 1.17; participants = 19; studies = 1).

1.40.3 Life in general

There was no clear difference between treatment groups for life in general measured by QOLI-BREF in the medium term (MD –0.49, 95% CI –1.73 to 0.75; participants = 19; studies = 1).

1.40.4 Living arrangements

There was no clear difference between treatment groups for living arrangements measured by QOLI-BREF in the medium term (MD –0.32, 95% CI –1.58 to 0.94; participants = 19; studies = 1).

1.40.5 Privacy

There was no clear difference between treatment groups for privacy measured by QOLI-BREF in the medium term (MD –0.58, 95% CI –1.40 to 0.24; participants = 19; studies = 1).

1.40.6 Relax

There was no clear difference between treatment groups for relax measured by QOLI-BREF in the medium term (MD –0.28, 95% CI –1.66 to 1.10; participants = 19; studies = 1).

1.41 Peer outcomes: 3c. Quality of life for participant and peer supporter: specific aspects – mean endpoint score (SF-36 subscales, high = good) – medium term

One study used the SF-36 to measure specific aspects of quality of life (Analysis 1.41) (Druss 2010).

1.41.1 Mental health

There was no clear difference between treatment groups for mental health measured by SF-36 in the medium term (MD –0.20, 95% CI –5.00 to 4.60; participants = 80; studies = 1).

1.41.2 Physical health

There was no clear difference between treatment groups for physical health measured by SF-36 in the medium term (MD 2.90, 95% CI –3.21 to 9.01; participants = 80; studies = 1).

1.42 Peer outcomes: 3d. Quality of life for participant and peer supporter: specific aspects – mean endpoint score (QOL-brief sub scales, high = good) – medium term (skewed data)

These data were skewed and are presented as 'other data' (Analysis 1.42) (Reynolds 2004).

1.43 Economic cost: 1. Direct and indirect costs (Euro): total costs (high = poor)

One study reported total costs (Eur) (Analysis 1.43) (Castelein 2008).

1.43.1 Medium term

There was no clear difference between treatment groups for total costs in the medium term (MD Eur 2092.00, 95% CI −74.00 to 4258.00; participants = 0; studies = 1).

1.43.2 Long term

There was no clear difference between treatment groups for total costs in the long term (MD Eur 775.00, 95% CI −1610.00 to 3160.00; participants = 0; studies = 1).

1.44 Economic outcomes: 2. Direct costs (Euro): for minimally guided peer support (high = poor) – long term (skewed data)

These data were skewed and are presented as 'other data' (Analysis 1.44) (Castelein 2008).

1.45 Economic outcomes: 3a. Indirect costs (Euro): for inpatient and semi-inpatient care (high = poor) – long term (skewed data)

These data were skewed and are presented as 'other data' (Analysis 1.45) (Castelein 2008).

1.46 Economic outcomes: 3b. Indirect costs (Euro): for outpatient and community care (high = poor) – long term (skewed data)

These data were skewed and are presented as 'other data' (Analysis 1.46) (Castelein 2008).

1.47 Economic outcomes: 3c. Indirect costs (Euro): for general healthcare (high = poor) – long term (skewed data)

These data were skewed and are presented as 'other data' (Analysis 1.47) (Castelein 2008).

1.48 Economic outcomes: 3d. Indirect costs (Euro): of day activity institutions (high = poor) – long term (skewed data)

These data were skewed and are presented as 'other data' (Analysis 1.48) (Castelein 2008).

1.49 Economic outcomes: 3e. Indirect costs (Euro): of medication (high = poor) – long term (skewed data)

These data were skewed and are presented as 'other data' (Analysis 1.49) (Castelein 2008).

Missing outcomes

None of the studies reported data for use of specialist community services, time to hospitalisation, relapse or adverse events.

2. Peer support plus standard care versus clinician-led support plus standard care

One study compared peer support with clinician-led support and reported useable data for global state, mental state and impact on participant and peer supporter (Eisen 2012).

2.1 Global state: 1. General health – mean total endpoint score (VR-12, high = good) – medium term

There was no clear difference between treatment groups for general health in the medium term (MD 2.59, 95% CI −1.45 to 6.63; participants = 156; studies = 1) (Analysis 2.1).

2.2 Mental state: 1a. Specific: various aspects – mean endpoint score (various scales, high = good) – medium term

2.2.1 Hope (SHS)

There was no clear difference between treatment groups for mental state measured by SHS in the medium term (MD −0.59, 95% CI −1.80 to 0.62; participants = 156; studies = 1).

2.2.2 Recovery (RAS)

There was no clear difference between treatment groups for recovery measured by RAS in the medium term (MD −0.50, 95% CI −7.13 to 6.13; participants = 156; studies = 1).

2.2.3 Empowerment

There was no clear difference between treatment groups for empowerment (MD −0.65, 95% CI −2.95 to 1.65; participants = 156; studies = 1).

2.3 Mental state: 1b. Specific: various aspects – mean endpoint score (Patient Activation Scale (PAS) subscale, high = good) – medium term

2.3.1 Activation (patient)

There was no clear difference between treatment groups for activation measured by PAS in the medium term (MD 0.30, 95% CI −1.64 to 2.24; participants = 156; studies = 1; Analysis 2.3).

2.4 Mental state: 1c. Specific: various aspects – mean endpoint score (BASIS subscale, high = poor) – medium term (skewed data)

These data were skewed and are presented as 'other data' (Analysis 2.4).

2.5 Behaviour: 1. Specific: drug/alcohol use – mean endpoint score (BASIS subscale, high = poor) – medium term (skewed data)

These data were skewed and are presented as 'other data' (Analysis 2.5).

2.6 Peer outcomes: 1. Impact on the service user and peer supporter: social support (MOSSSS, high = good) – medium term

There was no clear difference between treatment groups for social support measured by MOSSSS in the medium term (MD 4.97, 95% CI −0.62 to 10.56; participants = 156; studies = 1; Analysis 2.6).

3. Sensitivity analysis

Data were reported for two of our primary outcomes: service use and death. However, for each of these outcomes, only one study

contributed data and it was not possible to carry out sensitivity analyses for implication of randomisation, risk of bias and unclear proportion of schizophrenia, neither were imputed values used.

We could carry out sensitivity analyses for assumptions for lost binary data and fixed-effect model.

3.1 Service use: 1. Hospital admission – medium term

For this primary outcome, we analysed the effect of using ITT assumption from information regarding attrition in Reynolds 2004 (Analysis 3.1).

3.1.1 Without intention to treat

There was no clear difference between treatment groups when not using assumptions for ITT (RR 0.44, 95% CI 0.11 to 1.75; participants = 19; studies = 1).

3.1.2 With intention to treat

There was no clear difference between treatment groups when using assumption for ITT (RR 0.55, 95% CI 0.18 to 1.64; participants = 25; studies = 1).

3.2 Fixed-effect model

For the primary outcomes, the direction of estimated effect was not changed when we used a fixed-effect model.

4. Subgroup analysis

We did not undertake any subgroup analysis as the populations between studies were in similar clinical state, stage or problem. The sources of heterogeneity for some outcomes were not identified.

DISCUSSION

Summary of main results

Our primary outcomes were hospital admission, duration of hospital stay, relapse, clinically important change in global state (improved) and death. Other outcomes of importance to us were clinically important change in quality of life for peer supporter and service user as well as cost to society. The trials reported only data for hospital admission and death. The trials did report data for other secondary outcomes and we have summarised results below.

1. Service use

There were limited data for service use. When comparing peer support with standard care, only one study) reported useable data for hospital admission and found no clear difference between groups (RR 0.44, 95% CI 0.11 to 1.75; very low-quality evidence) (Reynolds 2004). The same study also found no difference between treatment groups for use of emergency services (RR 0.39, 95% CI 0.11 to 1.32), while another study found peer support may have led to increased use of primary care services in the medium term (RR 1.77, 95% CI 1.09 to 2.85) (Druss 2010). There were no data for other service outcomes such as specialist community services, time to hospitalisation or number of admissions. When comparing peer-support interventions with clinician-led support, there were no data for any service outcomes.

2. Global state

Thirteen studies compared peer support with standard care but none of these studies reported on relapse, time to relapse or clinically important improvement in global state. The only useable data for global state were endpoint scores on various global state scales, results varied and meta-analyses were not possible. Eisen 2012 used the VR-12 and found no difference in mean endpoint scores between treatment groups at medium term (MD –0.02, 95% CI –3.96 to 3.92). Cook 2012b used BSI endpoint scores and found a favourable difference between scores for peer support at medium term follow-up (MD –0.13, 95% CI –0.25 to –0.01) but no difference at long-term follow-up (MD 0.00, 95% CI –0.11 to 0.11). Mahlke 2017 used the CGI and also found a favourable effect for peer support at medium term but then a favourable effect for standard care at long term.

Eisen 2012 also compared peer support with clinician-led support. At medium-term follow-up, there was no clear difference in global state as measured by mean endpoint scores on the VR-12 (MD 2.59, 95% CI –1.45 to 6.63).

Three studies reported compliance with medication data, but these data were skewed.

3. Mental state

None of the studies reported clinically important changes in mental state. The evaluation of participants' mental state was based on scores from various mental state scales or subscales. Results were inconsistent. For example, one study measured 'hope' by the HHI and found participants in the peer support groups had better scores than those in the standard care group (medium term: MD 0.24, 95% CI 0.11 to 0.37), but when hope was measured by two other studies using SHS, there was no difference in scores (medium term: MD 0.37, 95% CI –0.22 to 0.96).

For behaviour outcomes, three studies measured recovery using RAS at medium term. There were no differences between treatment groups at medium term (MD 2.69, 95% CI –0.82 to 6.20). One study reported long-term data for recovery (also using RAS) and found a difference favouring the peer support group (MD 4.16, 95% CI 1.16 to 7.16). Data were reported for a wide variety of 'behaviours', most results showed no differences between peer support and standard care. However, there were positive effects for peer support at medium term for 'internal locus of control', personal confidence, reliance on others, willingness to ask for help and at long term for 'mindful non-adherence', 'no symptom domination', personal confidence and willingness to ask for help. It should be noted all these above results were based on data from single studies.

When comparing peer-support intervention versus clinician-led support, there was no clear difference between the groups in terms of patient activation, hope, recovery and empowerment. This finding was based on the results from a single study with a very small sample size (Eisen 2012).

4. Functioning

When comparing peer-support intervention with standard care, the evaluation of psychosocial functioning was based on outcomes such as general functioning and specific functioning and encompassed emotional, physical, social, physical, interpersonal areas and cognitive functioning. The findings on general function were inconsistent. One study found that peer-support intervention was associated with higher general function (measured by GAF) compared with standard care (Mahlke 2017). However, when general function was measured using other scales, there was no

difference between the groups. In addition, there were no clear differences for any specific function.

When comparing peer support versus clinician-led support, no study reported data on functioning.

5. Leaving the study early

In the medium term, fewer people left the studies early in the peer support group than in the standard care group (RR 0.66, 95% CI 0.51 to 0.87; participants = 741; studies = 6), though no reason for early leaving was given and this benefit was not observed in the long term (RR 1.34, 95% CI 0.19 to 9.22; participants = 877; studies = 3).

6. Peer outcomes

When comparing peer support with standard care, no study reported data for clinically important change in quality of life for the peer supporter and service user. The only useable data reported were scale scores for 'impact' (which included domains such as negative aspects, positive aspects, empathy, overall relationship and unconditionality) and quality of life (which included aspects such as overall quality of life, mental health, physical health, social function, amount of time with others and living arrangements). For impact, data did not show clear differences between groups for most measures but indicated differences favouring peer support for positive relationships for the service user in the medium term and peer contact in the long term. For quality of life, data showed no clear difference in various subscale scores of quality of life, except for data from one small study that showed clear differences favouring peer support in the medium term in the areas of physical and mental health and social function (Qian 2015). It should be noted that we could not pool data for any peer outcomes.

When comparing peer support versus clinician-led support, only one study measured the social support status between the groups. There were no differences.

Other outcomes such as coping ability, expressed emotion of family and expressed emotion of peer supporter were not reported by the included studies for both groups.

7. Adverse events

Adverse event reporting was also limited, for studies comparing peer support with standard care the only adverse event data reported were for all-cause mortality. There were no clear differences between the groups.

Eisen 2012 compared peer-support intervention with clinician-led support and did not report adverse event data.

8. Economic outcomes

One study comparing peer-support intervention with standard care reported economic data (Castelein 2008). However, because these data were skewed, we were unable to perform standard analyses.

Overall completeness and applicability of evidence

The completeness and applicability of the evidence available for effects of peer support for people with schizophrenia and other serious mental illness is currently inadequate.

Useable data were sparse and limited. The main outcomes we planned to assess were hospital admission relapse, global state, quality of life for peer supporter and service user, adverse events and economic costs. Only one study reported the number of hospital admissions, no relapse data were reported and only scale scores were reported for other outcomes. We could not pool data for many outcomes and in addition, there were high usage of subscale data to represent and compare the results between studies.

Only one study enrolled only participants with schizophrenia (Qian 2015). The other studies enrolled participants with various mental illnesses including schizophrenia but only three studies clearly stated the proportion of people with schizophrenia (Cook 2012b; Cook 2012a; Druss 2010).

The structure, format and components of the peer-support interventions varied among the studies with most studies referring to peer-support intervention as any intervention (such as education, case management) that was delivered by peers. This could have reduced consistency or homogeneity of intervention protocols between studies and validity of their pooled effects.

Quality of the evidence

1. Limitations in study design

The current systematic review included 13 studies involving 2479 participants. Eleven studies did not clearly address the methods employed for randomisation, and the authors from only five studies stated how they concealed the allocation. Consequently, there was a potential risk of selection bias. None of the studies clearly stated that participants/personnel were blinded, and nine studies (70%) did not clearly state that they blinded the outcome assessors. These omissions pointed to a serious risk of performance and detection bias. Lastly, six of the 13 included studies (54%) were rated at high or unclear risk of attrition bias. The outcome data were poorly reported by six of the included studies.

2. Indirectness of the evidence

The indirectness of the evidence was supported by the fact that the participants in the included studies had a variety of mental illnesses besides schizophrenia, such as major depression and mood disorder. High variations of types and percentages of mental illnesses found in the included studies might have reduced the specificity and accuracy of the estimated treatment effects of the peer support group for individual types or diagnoses of mental illnesses.

3. Inconsistency of the results

Because most of the outcomes were based on data from a single study, we could not assess inconsistency between studies. The review included 13 studies, but meta-analyses could only be performed for a few outcomes and where meta-analyses were possible, data were pooled from one to three studies only.

4. Imprecision of the results

Most of the outcomes were imprecise due to small number of included studies and the very small sample sizes. Clinically important change data were not reported and most results were based mean endpoint scores.

5. Publication bias

The assessment of publication bias was not feasible, because none of the comparisons included more than 10 studies. For this reason, we did not create any funnel plots, as they would not have provided any meaningful information.

Potential biases in the review process

We conducted a comprehensive search of the literature for potentially eligible studies to limit the bias in the review process. We employed strict measures to improve screening accuracy and consulted a search specialist. The data screening and extraction process strictly adhered to the Cochrane recommended procedures and standards. However, since we only included published data, it is possible that there is publication bias. In an attempt to minimise potential bias during data extraction, two review authors independently screened studies and extracted data.

Agreements and disagreements with other studies or reviews

Our review showed peer support had no apparent effect on hospital admission, global state or death. This is in line with a previous systematic review assessing the effectiveness of peer support for people with severe mental illness where the authors found that there was little evidence that peer support positively impacted hospitalisations, overall symptoms, satisfaction with services or a combination of these (Lloyd-Evans 2014). Another systematic review summarised existing evidence and addressed certain concerns, namely whether the mental health settings would be too stressful for peer staff, whether they would relapse, could peer staff handle the administrative demands of the job, would they potentially harm clients, or would they make the jobs of other staff more difficult (Davidson 2012). Davidson and colleagues concluded peer support can provide an opportunity for transformation for individuals as they transition from being a service recipient to becoming a service provider and may contribute to much-needed, broader cultural and service-related changes, as well as improve individual outcomes (Davidson 2012). Miyamoto and Sono conducted a review to describe the principles, effects and benefits of peer support that have been documented in the published literature (Miyamoto 2012). They found that the main challenge for peer-support interventions was related to the 'role' and 'relationship' between peer support providers and the recipients (Miyamoto 2012). To redefine the service provider and service user relationship, and better define concepts of helping and support, it is important to gain more knowledge about the factors that influence peer support relationships, such as mutual responsibility and interdependence (Miyamoto 2012). These potential therapeutic components or mechanisms have not been examined in this review and thus should be investigated in future studies.

AUTHORS' CONCLUSIONS

Implications for practice

1. For people with schizophrenia or other serious mental illness

Low-quality evidence shows that adding peer support to standard care does not affect hospital admissions and all-cause mortality compared to standard care for people with schizophrenia and other serious mental illnesses. Limited data from a few small studies shows some differences, favouring peer support, in scale scores for global state and specific mental state and behavioural outcomes such as recovery, empowerment, personal confidence, willingness to ask for help and reliance on others. However, more evidence from high-quality trials is needed before we can make firm conclusions about these results.

2. For clinicians

A comparison of peer-support intervention with standard care found no clear difference between groups on hospital admissions or all-cause mortality for people with schizophrenia and other serious mental illnesses. However, these findings are based on *low-* or *very low-quality evidence*. When compared with standard care, peer support may also improve participants' global state and some specific mental states and behavioural domains such as hope agency, recovery, and empowerment and personal confidence. However, these data are mostly derived from single small trials with relatively low precision, and thus foster very limited confidence in the findings. Currently, the data are insufficient to draw any firm conclusions about the impact of peer-support interventions in people with schizophrenia and other serious mental illnesses.

3. For policy makers

Weak evidence demonstrates that peer support may have some benefits when added to standard care; however, the direct and indirect costs of peer support remain unclear due to insufficient data and more research is needed.

Implications for research

1. General

We found a lack of high-quality evidence from randomised trials to fully evaluate the effect of peer-support interventions. The included randomised controlled trials (RCTs) carried a considerable risk of bias. Large RCTs with attempts to lower selection and attrition bias and follow the standard and high-quality reporting of RCTs such as the CONSORT statement are therefore required to produce more valid effects of peer support group intervention for people with severe mental illnesses.

2. Specific

While peer support groups have been conducted in a wide variety of mental illnesses, future research of this group intervention can be conducted in specific illness groups such as schizophrenia. More conclusive evidence on the benefits of this group intervention in schizophrenia or its subtypes can be established before applying or generalising this intervention to other severe mental illnesses.

Of note, the treatment protocol employed in peer-support interventions varies considerably in current studies and should be further standardised in future studies. The current outcome data are insufficient to draw any conclusions. Future RCTs that will test the effects of different types of peer-support interventions for people with schizophrenia should focus on factors such as measuring participants' usage of specialist community services (e.g. interventions, assertive outreach and crisis teams), hospital admissions, relapse, global state and cost. Adequate reporting of outcome data are also required in future studies.

ACKNOWLEDGEMENTS

The Cochrane Schizophrenia Group Editorial Base in Nottingham produces and maintains standard text for use in the Methods section of their reviews. We have used this text as the basis of what appears here and adapted it as required.

We would like to thank John Lally for peer reviewing the protocol and Masahiro Banno, Laia Briones-Buixassa, Judit Bort-Roig, Anna M Puig-Ribera, Mercè Sitjà-Rabert and Sulafa Ahmad for peer reviewing the full review.

Parts of this review were generated using RevMan HAL v 4.3. You can find more information about RevMan HAL here.

REFERENCES

References to studies included in this review

Castelein 2008 *{published data only}*

Castelein MS. The effectiveness of support groups for people suffering from psychosis: a randomised controlled trial. www.isrctn.com/ISRCTN02457313 (first registered 22 November 2006).

Castelein S, Bruggeman R, van Busschbach JT, van Der Gaag M, Stant AD, Knegtering H, et al. The effectiveness of minimally guided peer support groups for people suffering from psychosis: a randomized controlled trial. Schizophrenia Research 2007:52-3.

Castelein S, Bruggeman R, van Busschbach JT, van der Gaag M, Stant AD, Knegtering H, et al. The effectiveness of peer support groups in psychosis: a randomised controlled trial. *Acta Psychiatrica Scandinavica* 2008;**118**:64-72. [CSzG: 16589]

Stant AD, Castelein S, Bruggeman R, van Busschbach JT, van der Gaag M, Knegtering H, et al. Economic aspects of peer support groups for psychosis. *Community Mental Health Journal* 2011;**47**:99-105. [CSzG: 22722]

Cook 2012a *{published data only}*

Cook JA, Steigman P, Pickett S, Diehl S, Fox A, Shipley P, et al. Randomized controlled trial of peer-led recovery education using building recovery of individual dreams and goals through education and support (BRIDGES). *Schizophrenia Research* 2012;**136**(1-3):36-42. [CSzG: 24471]

Cook 2012b *{published data only}*

Cook JA, Copeland ME, Jonikas JA, Hamilton MM, Razzano LA, Grey DD, et al. Results of a randomised controlled trial of mental illness self-management using Wellness Recovery Action Planning. *Schizophrenia Bulletin* 2012;**38**(4):881-91. [CSzG: 28000]

Jonikas JA, Grey DD, Copeland ME, Razzano LA, Hamilton MM, Floyd CB, et al. Improving propensity for patient self-advocacy through wellness recovery action planning: results of a randomised controlled trial. *Community Mental Health Journal* 2013;**49**(3):260-9.

Druss 2010 *{published data only}*

Druss BG, Zhao L, von Esenwein SA, Bona JR, Fricks L, Jenkins-Tucker S, et al. The Health and Recovery Peer (HARP) Program: a peer-led intervention to improve medical self-management for persons with serious mental illness. *Schizophrenia Research* 2010;**118**(1-3):264-70. [CSzG: 23163]

Eisen 2012 *{published data only}*

Eisen SV, Schultz MR, Mueller LN, Degenhart C, Clark JA, Resnick SG, et al. Outcome of a randomized study of a mental health peer education and support group in the VA. *Psychiatric Services* 2012;**63**(12):1243-6. [CSzG: 27999]

Goldberg 2013 *{published data only}*

Goldberg RW, Dickerson F, Lucksted A, Brown CH, Weber E, Tenhula WN, et al. Living Well: an intervention to improve self-management of medical illness for individuals with serious mental illness. *Psychiatric Services (Washington, D.C.)* 2013;**64**(1):51-7. [CSzG: 28752]

Kelly 2014 *{published data only}*

Kelly E, Fulginiti A, Pahwa R, Tallen L, Duan L, Brekke JS. A pilot test of a peer navigator intervention for improving the health of individuals with serious mental illness. *Community Mental Health Journal* 2014;**50**(4):435-46. [CSzG: 29206]

Mahlke 2017 *{published data only}*

Mahlke CI, Priebe S, Heumann K, Daubmann A, Wegscheider K, Bock T. Effectiveness of one-to-one peer support for patients with severe mental illness – a randomised controlled trial. *European Psychiatry* 2017;**42**:103-10. [CSzG: 35851]

Qian 2015 *{published data only}*

Qian Y. Peer support for the quality of life in patients with schizophrenia [同伴教育对精神分裂症患者生活质量的影响作用] . *Health for Everyone* 2015;**23**:7. [CSzG: 34512]

Reynolds 2004 *{published data only}*

Reynolds W, Lauder W, Sharkey S, Maciver S, Veitch T, Cameron D. The effects of a transitional discharge model for psychiatric patients. *Journal of Psychiatric and Mental Health Nursing* 2004;**11**:82-8. [CSzG: 11252]

Rowe 2007 *{published data only}*

Rowe M, Bellamy C, Baranoski M, Wieland M, O'Connell MJ, Benedict P, et al. A peer-support, group intervention to reduce substance use and criminality among persons with severe mental illness. *Psychiatric Services* 2007;**58**(7):955-61. [CSzG: 15324]

Sells 2008 *{published data only}*

Sells D, Black R, Davidson L, Rowe M. Beyond generic support: incidence and impact of invalidation in peer services for clients with severe mental illness. *Psychiatric Services* 2008;**59**:1322-7. [CSzG: 17088]

Sells D, Davidson L, Jewell C, Falzer P, Rowe M. The treatment relationship in peer-based and regular case management for clients with severe mental illness. *Psychiatric Services* 2006;**57**:1179-84. [CSzG: 13866]

Van Gestel-Timmermans 2012 *{published data only}*

Van Gestel-Timmermans H, Brouwers EP, van Assen MA, van Nieuwenhuizen C. Effects of a peer-run course on recovery from serious mental illness: a randomised controlled trial. *Psychiatric Services* 2012;**63**:54-60. [CSzG: 23608]

References to studies excluded from this review

Buchkremer 1995 *{published data only}*

Buchkremer G, Schulze Monking H, Holle R, Hornung WP. The impact of therapeutic relatives' groups on the course of illness of schizophrenic patients. *European Psychiatry* 1995;**10**:17-27.

Chen 2016 *{published data only}*

Chen Q, Liu L, Zhang D, Li J, Li Y. Impact of social support on quality of life and rehabilitation in patients with schizophrenia [社会支持对精神分裂症患者生活质量和康复的影响]. *Chinese Journal of Clinical Psychology* 2016;**24**(1):185-7.

Chinman 2015 *{published data only}*

Chinman M, Oberman RS, Hanusa BH, Cohen AN, Salyers MP, Twamley EW, et al. A cluster randomized trial of adding peer specialists to intensive case management teams in the Veterans Health Administration. *Journal of Behavioral Health Services & Research* 2015;**42**(1):109-21.

NCT00781079. Do Consumer Providers Enhance Recovery? (PEER). clinicaltrials.gov/ct2/show/NCT00781079 (first received 28 October 2008).

Corrigan 2017a *{published data only}*

Corrigan PW, Kraus DJ, Pickett SA, Schmidt A, Stellon E, Hantke E, et al. Using peer navigators to address the integrated health care needs of homeless African Americans with serious mental illness. *Psychiatric Services* 2017;**68**(3):264-70.

Corrigan 2017b *{published data only}*

Corrigan PW, Pickett S, Schmidt A, Stellon E, Hantke E, Kraus D, et al. Peer navigators to promote engagement of homeless African Americans with serious mental illness in primary care. *Psychiatry Research* 2017;**255**:101-3.

Craig 2004 *{published data only}*

Craig T, Doherty I, Jamiesson-Craig R, Boocock A, Attafua G. The consumer-employee as a member of a Mental Health Assertive Outreach Team. I. Clinical and social outcomes. *Journal of Mental Health* 2004;**13**(1):59-69.

Forchuk 2005 *{published data only}*

Forchuk C, Martin ML, Chan YL, Jensen E. Therapeutic relationships: from psychiatric hospital to community. *Journal of Psychiatric & Mental Health Nursing* 2005;**12**(5):556-64.

Gunter 1983 *{published data only}*

Gunter NC, Bedell JR. The peer-managed small group versus the rehabilitation model of treatment of chronic patients. *Hospital and Community Psychiatry* 1983;**34**(8):724-8.

Hazell 2016 *{published data only}*

Hazell CM, Hayward M, Cavanagh K, Jones AM, Strauss C. Guided self-help cognitive behavioral intervention for VoicEs (GiVE): study protocol for a pilot randomized controlled trial. *Trials* 2016;**17**(1):351.

ISRCTN14282228 *{published data only}*

ISRCTN14282228. Hospitality project (HY): combining peer support and home-based skill training in people with schizophrenia. www.isrctn.com/ISRCTN14282228 (first received 31 January 2017).

Kaplan 2011 *{published data only}*

Kaplan K, Salzer MS, Solomon P, Brusilovskiy E, Cousounis P. Internet peer support for individuals with psychiatric disabilities: a randomised controlled trial. *Social Science & Medicine* 2011;**72**:54-62.

Kaufmann 1995 *{published data only}*

Kaufmann CL, Schulberg HC, Schooler NR. Self-help group participation among people with severe mental illness. *Prevention in Human Services* 1995;**11**:315-31.

Killackey 2013 *{published data only}*

Killackey E, Allott K, Cotton SM, Jackson H, Scutella R, Tseng YP, et al. A randomised controlled trial of vocational intervention for young people with first-episode psychosis: method. *Perspectives in Early Intervention* 2013;**7**:329-37.

Klein 1998 *{published data only}*

Klein AR, Cnaan RA, Whitecraft J. Significance of peer social support with dually diagnosed clients: findings from a pilot study. *Research on Social Work Practice* 1998;**8**(5):529-51.

NCT02974400 *{published data only}*

NCT02974400. Evaluation of internet-based cognitive behavioral self-help treatments for people with psychosis. clinicaltrials.gov/ct2/show/NCT02974400 (first received 28 November 2016).

O'Connell 2017 *{published data only}*

O'Connell MJ, Flanagan EH, Delphin-Rittmon ME, Davidson L. Enhancing outcomes for persons with co-occurring disorders through skills training and peer recovery support. Journal of Mental Health (Abingdon, England) 2017; Vol. 2017:1-6.

Rivera 2007 *{published data only}*

Rivera JJ, Sullivan AM, Valenti SS. Adding consumer-providers to intensive case management: does it improve outcome?. *Psychiatric Services* 2007;**58**(6):802-9.

Rogers 2012 *{published data only}*

Rogers E. Initiatives in researching peer delivered services. Proceedings of the 11th World Congress of the World Association of Psychosocial Rehabilitation; 2012 Nov 10-13; Milan, Italy. 2012:42941.

Salyers 2010 *{published data only}*

Salyers MP, McGuire AB, Rollins AL, Bond GR, Mueser KT, Macy VR. Integrating assertive community treatment and illness management and recovery for consumers with severe mental illness. *Community Mental Health Journal* 2010;**46**:319-29.

Segal 2010 *{published data only}*

Segal SP, Silverman CJ, Temkin TL. Self-help and community mental health agency outcomes: a recovery-focused randomised controlled trial. *Psychiatric Services* 2010;**61**:905-10.

Shahar 2006 *{published data only}*

Shahar G, Kidd S, Styron TH, Davidson L. Consumer support and satisfaction with mental health services in severe mental illness: the moderating role of morale. *Journal of Social and Clinical Psychology* 2006;**25**:945-62.

Streicker 1984 *{published data only}*

Streicker SK. Evaluation of a medication education program for psychiatric patients [Thesis]. Evanston, Illinois (USA): Northwestern University, 1984:1-75.

Verhaegh 2006 *{published data only}*

Verhaegh MJ, Gezondheidszorg G. Researching the effect of Assertive Community Treatment (ACT) for first episode psychosis: the use of peer-interviewers in a quasi-experimental study. *Schizophrenia Research* 2006;**86**:S170.

Weissman 2005 *{published data only}*

Weissman EM, Covell NH, Kushner M, Irwin J, Essock SM. Implementing peer-assisted case management to help homeless veterans with mental illness transition to independent housing. *Community Mental Health Journal* 2005;**41**:267-76.

Zhou 2016 *{published data only}*

Zhou W, Wang J. Peer support for self-care ability and self efficacy in patients with schizophrenia [同伴教育对精神分裂症患者自理能力和自我效能的影响]. *DANGDAIHUSHI* 2016;**2016**(01):96-8.

References to studies awaiting assessment

Daumit 2010 *{published data only}*

Daumit G, Appel L, Leatherman E, Latkin C, Dalcin A, Goggins B, et al. Randomised trial of peer-supported physical activity for persons with severe mental illness in community. *Journal of General Internal Medicine* 2010;**25(3 Suppl)**:S378-9.

Kroon 2011 *{published data only}*

Kroon H, Boevink W. Effectiveness of recovery programs. IXth International Conference of the European Network for Mental Health Service Evaluation (ENMESH); Ulm, Germany; 23–25 June 2011. 2011.

NCT00458094 *{published data only}*

NCT00458094. Peer support for increasing physical activity in people with serious mental illnesses. clinicaltrials.gov/ct2/show/NCT00458094 (first received 9 April 2007).

NTR1166 *{published data only}*

Ruchlewska A, Mulder CL, Smulders R, Roosenschoon BJ, Koopmans G, Wierdsma A. The effects of crisis plans for patients with psychotic and bipolar disorders: a randomised controlled trial. BMC Psychiatry 2009; Vol. 9:41.

Robinson 2010 *{published data only}*

Robinson J, Bruxner A, Harrigan S, Bendall S, Killackey E, Tonin V, et al. Study protocol: the development of a pilot study employing a randomised controlled design to investigate the feasibility and effects of a peer support program following discharge from a specialist first-episode psychosis treatment centre. BMC Psychiatry 2010; Vol. 10:1-7.

Tondora 2010 *{published data only}*

Tondora J, O'Connell M, Miller R, Dinzeo T, Bellamy C, Andres-Hyman R, et al. A clinical trial of peer-based culturally responsive person-centered care for psychosis for African Americans and Latinos. *Clinical Trials* 2010;**7**:368-79.

References to ongoing studies

ACTRN1261200097 *{published data only}*

ACTRN1261200097. Peer delivered support intervention for people who hear voices: pilot randomised controlled trial. Australian New Zealand Clinical Trials Registry 2012.

Chinman 2017 *{published data only}*

Chinman M, Daniels K, Smith J, McCarthy S, Medoff D, Peeples A, et al. Provision of peer specialist services in VA patient aligned care teams: protocol for testing a cluster randomized implementation trial. *Implementation Science* 2017;**12**(1):57.

NCT01566513 *{published data only}*

NCT01566513. Effectiveness and cost effectiveness of peer mentors in reducing hospital use (Project PEP). clinicaltrials.gov/ct2/show/NCT01566513 (first received 29 March 2012).

NCT02958007 *{published data only}*

NCT02958007. Peer support for exercise in older veterans with psychotic disorders. clinicaltrials.gov/ct2/show/NCT02958007 (first received 8 November 2016).

NCT02989805 *{published data only}*

NCT02989805. Engaging patients with mental disorders from the ED in outpatient care. clinicaltrials.gov/ct2/show/NCT02989805 (first received 12 December 2016).

Additional references

Aas 2010

Aas IH. Global Assessment of Functioning (GAF): properties and frontier of current knowledge. *Annals of General Psychiatry* 2010;**9**:20. [PUBMED: 20459646]

Ahmed 2012

Ahmed AO, Doane NJ, Mabe PA, Buckley PF, Birgenheir D, Goodrum NM. Peers and peer-led interventions for people with schizophrenia. *Psychiatric Clinic of North America* 2012;**35**:699-715.

Altman 1996

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

Arnstein 2002

Arnstein P, Vidal M, Wells-Federman C, Morgan B, Caudill M. From chronic pain patient to peer: benefits and risks of volunteering. *Pain Management Nursing* 2002;**3**:94-103.

Balestroni 2012

Balestroni G, Bertolotti G. EuroQol-5D (EQ-5D): an instrument for measuring quality of life [L'EuroQol-5D (EQ-5D): uno strumento per la misura della qualita della vita]. *Monaldi*

Archives for Chest Disease 2012;**78**(3):155-9. [PUBMED: 23614330]

Barrettlennard 1962

Barrettlennard GT. Dimensions of therapist response as causal factors in therapeutic change. *Psychological Monographs* 1962;**76**:1-36.

Bland 1997

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

Boevink 2009

Boevink W, Kroon H, Giesen F. Empowerment: Construction and Validation of a Questionnaire. Utrecht (the Netherlands): Trimbos Instituut, 2009.

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. [PUBMED: 10667106]

Bradstreet 2010

Bradstreet S, Pratt R. Developing peer support worker roles: reflecting on experiences in Scotland. *Mental Health and Social Inclusion* 2010;**14**(3):36-41.

Brashers 1999

Brashers DE, Neidig SM, Judith L. The patient self-advocacy scale: measuring patient involvement in health care decision-making interactions. *Health Communication* 1999;**11**:97.

Bridges 2002

Bridges KR, Sanderman R, Van Sonderen E. An English language version of the Social Support List: preliminary reliability. *Psychological Reports* 2002;**90**:1055-8.

Busner 2007

Busner J, Targum SD. The Clinical Global Impressions Scale: applying a research tool in clinical practice. *Psychiatry (Edgmont (Pa. : Township))* 2007;**4**(7):28-37.

Cameron 2007

Cameron IM, Cunningham L, Crawford JR, Eagles JM, Eisen SV, Lawton K, et al. Psychometric properties of the BASIS-24© (Behaviour and Symptom Identification Scale-Revised) mental health outcome measure. *International Journal of Psychiatry in Clinical Practice* 2007;**11**:36-43.

Carpinello 2000

Carpinello SE, Knight EL, Markowitz FE, Pease EA. The development of the Mental Health Confidence Scale: a measure of self-efficacy in individuals diagnosed with mental disorders. *Psychiatric Rehabilitation Journal* 2000;**23**:236-43.

Cheah 2001

Cheah J. Chronic disease management: a Singapore perspective. *BMJ* 2001;**323**:990-3.

Chien 2009

Chien WT, Norman I. The effectiveness and active ingredients of mutual support groups for family caregivers of people with psychotic disorders: a literature review. *International Journal of Nursing Studies* 2009;**46**(12):1604-23.

Dale 2008

Dale J, Caramlau IO, Lindenmeyer A, William SM. Peer support telephone calls for improving health. *Cochrane Database of Systematic Reviews* 2008, Issue 4. [DOI: 10.1002/14651858.CD006903.pub2]

Davidson 1999

Davidson L, Chinman M, Kloos B, Weingarten R, Stayner D, Tebes J. Peer support among individuals with severe mental illness: a review of the evidence. *Clinical Psychology: Science and Practice* 1999;**6**:1651-87.

Davidson 2012

Davidson L, Bellamy C, Guy K, Miller R. Peer support among persons with severe mental illnesses: a review of evidence and experience. *World Psychiatry* 2012;**11**(2):123-8.

De Hert 2009

De Hert M, Dekker JM, Wood D, Kahl KG, Holt RI, Möller HJ. Cardiovascular disease and diabetes in people with severe mental illness position statement from the European Psychiatric Association, supported by the European Association for the study of Diabetes (EASD) and the European Society of Cardiology (ESC). *European Psychiatry* 2009;**24**(6):412-24.

Deeks 2000

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

Deeks 2011

Deeks JJ, Higgins JP, Altman DG. Chapter 9: Analysing data and undertaking meta-analyses. 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.

Dennis 2003

Dennis CL. Peer support within a health care context: a concept analysis. *International Journal of Nursing Studies* 2003;**40**:321-32.

Derogatis 1993

Derogatis LR. Brief Symptom Inventory: Administration Scoring and Procedures Manual. 3rd Edition. Minneapolis (MN): National Computer Systems, 1993.

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**:2971-80.

Doull 2005

Doull M, O'Connor AM, Welch V, Tugwell P, Wells GA. Peer support strategies for improving the health and well being of individuals with chronic diseases. *Cochrane Database of Systematic Reviews* 2005, Issue 3. [DOI: 10.1002/14651858.CD005352]

Egger 1997

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

Elbourne 2002

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

Ellis 1984

Ellis RH, Wilson NZ, Foster FM. Statewide treatment outcome assessment in Colorado: the Colorado Client Assessment Record (CCAR). *Community Mental Health Journal* 1984;**20**:72-89.

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.

Giffort 1995

Giffort D, Schmook A, Woody C, et al. Construction of a Scale to Measure Consumer Recovery. Springfield, IL: Illinois Office of Mental Health, 1995.

Goff 2005

Goff DC, Cather C, Henderson DC, Freudenreich O, Copeland PM, Bierer M, et al. Medical morbidity and mortality in schizophrenia: guidelines for psychiatrists. *Journal of Clinical Psychiatry* 2005;**66**(2):183-94.

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**:876-83.

Harvey 2012

Harvey PD, Strassnig M. Predicting the severity of everyday functional disability in people with schizophrenia: cognitive deficits, functional capacity, symptoms, and health status. *World Psychiatry* 2012;**11**:73-9.

Herth 1992

Herth K. Abbreviated instrument to measure hope: development and psychometric evaluation. *Journal of Advanced Nursing* 1992;**17**:1251-9.

Hibbard 2004

Hibbard JH, Stockard J, Mahoney ER, Tusler M. Development of the Patient Activation Measure (PAM): conceptualizing and measuring activation in patients and consumers. *Health Services Research* 2004;**39**:1005-26.

Higgins 2003

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

Higgins 2011a

Higgins JP, Altman DG, Sterne JA. Chapter 8: Assessing risk of bias in included studies. 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.

Higgins 2011b

Higgins JP, Green S. Chapter 7: Selecting studies and collecting data. 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.

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.

Jablensky 2000

Jablensky A. Prevalence and incidence of schizophrenia spectrum disorders: implications for prevention. *The Australian and New Zealand Journal of Psychiatry* 2000;**34 (Suppl)**:S26-34.

Jonggierveld 1985

Jonggierveld J, DeKamphuls F. The development of a Rasch-type loneliness scale. *Applied Psychological Measurement* 1985;**9**:289-99.

Kane 2010

Kane JM, Correll CU. Past and present progress in the pharmacologic treatment of schizophrenia. *Journal of Clinical Psychiatry* 2010;**71**:1115-24.

Kay 1986

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

Kazis 2017

Kazis LE, Rogers WE, Rothendler J, Qian S, Selim A, Edelen MO, et al. Outcome performance measure development for persons with multiple chronic conditions, 2017. www.rand.org/pubs/research_reports/RR1844.html (accessed prior to 13 March 2019). [www.rand.org/pubs/research_reports/RR1844.html]

Lehman 1994

Lehman A, Kernan E, Postrado L. Toolkit for Evaluating Quality of Life for Persons with Severe Mental Illness. The Evaluation Centre, Cambridge, MA 1994.

Leucht 2005a

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. [PUBMED: 15982856]

Leucht 2005b

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. [PUBMED: 16199797]

Leucht 2007

Leucht S, Engel RR, Bauml J, Davis JM. Is the superior efficacy of new generation antipsychotics an artifact of LOCF?. *Schizophrenia Bulletin* 2007;**33**(1):183-91. [PUBMED: 16905632]

Li 1997

Li L, Young D, Hao W, Xiao S. Development and validation of the generic quality of life inventory in China. *Quality of Life Research* 1997;**6**:681-4.

Lloyd-Evans 2014

Lloyd-Evans B, Mayo-Wilson E, Harrison B, Istead H, Brown E, Pilling S, et al. A systematic review and meta-analysis of randomised controlled trials of peer support for people with severe mental illness. *BMC Psychiatry* 2014;**14**:39.

Lorig 1996

Lorig K, Steward A, Ritter P, Gonzalez V, Laurent D, Lynch J. Outcome Measures for Health Education and Other Health Care Interventions. Thousand Oaks (CA): Sage, 1996.

Mangalore 2007

Mangalore R, Knapp M. Cost of schizophrenia in England. *Journal of Mental Health Policy and Economics* 2007;**10**(1):23-41.

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.

McHorney 1993

McHorney CA, Ware JE, Raczek AE. Psychometric and clinical tests of validity in measuring physical and mental health constructs. *Medical Care* 1993;**31**:247-63.

Mclellan 1980

McLellan AT, Lubarsky L, Woody GE, O'Brien CP, et al. An improved evaluation instrument for substance abuse patients: the Addiction Severity Index. *Journal of Nervous and Mental Disease* 1980;**168**:26-33.

Mead 2001

Mead S, Hilton D, Curtis L. Peer support: a theoretical perspective. *Psychiatric Rehabilitation Journal* 2001;**25**:134-41.

Miyamoto 2012

Miyamoto Y, Sono T. Lessons from peer support among individuals with mental health difficulties: a review of the literature. *Clinical Practice & Epidemiology in Mental Health* 2012;**8**:822-9.

Mohamed 2008

Mohamed S, Rosenheck R, Swartz M, Stroup S, Lieberman JA, Keefe RS. Relationship of cognition and psychopathology to functional impairment in schizophrenia. *American Journal of Psychiatry* 2008;**165**(8):978-87.

Morrison 2009

Morrison AK. Cognitive behaviour therapy for people with schizophrenia. *Psychiatry (Edgmont (Pa. : Township))* 2009;**6**(12):32-9.

Nasrallah 2011

Nasrallah HA. The primary and secondary symptoms of schizophrenia: current and future management. *Current Psychiatry* 2011;**10**(9):S5-9.

NICE 2009

The National Collaborating Centre for Mental Health. Schizophrenia: core interventions in the treatment and management of schizophrenia in adults in primary and secondary care (NICE guideline), 2009. www.nice.org.uk/guidance/cg82 (accessed prior to 13 March 2019).

NIMH 1987

National Institute of Mental Health. Towards a Model for a Comprehensive Community-Based Mental Health System. Washington (DC): National Institute of Mental Health, 1987.

Overall 1962

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

Pharoah 2010

Pharoah F, Mari JJ, Rathbone J, Wong W. Family intervention for schizophrenia. *Cochrane Database of Systematic Reviews* 2012, Issue 11. [DOI: 10.1002/14651858.CD000088.pub3]

Pistrang 2008

Pistrang N, Barker C, Humphreys K. Mutual help groups for mental health problems: a review of effectiveness studies. *American Journal of Community Psychology* 2008;**42**(1):110-21.

Priebe 1999

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

Repper 2010

Repper J, Carter T, Together UK. Using personal experience to support others with similar difficulties. www.together-uk.org/wp-content/uploads/downloads/2011/11/usingpersexperience.pdf (access prior to 13 March 2019).

Rogers 1997

Rogers ES, Chamberlin JE, Crean T. A consumer-constructed scale to measure empowerment among users of mental health services. *Psychiatric Services* 1997;**48**:1042-7.

Rosenberg 1965

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

Schinnar 1990

Schinnar AP, Rothbard AB, Kanter R, Jung YS. An empirical literature review of definitions of severe and persistent mental illness. *American Journal of Psychiatry* 1990;**147**(12):1602-8.

Schwarzer 1995

Schwarzer R, Jerusalem M. Generalized Self-Efficacy scale. In: Weinman J, Wright S, Johnston M editor(s). Measures in Health Psychology: a User's Portfolio. Causal and Control Beliefs. Windsor (UK): National Foundation for Educational Research, 1995:35-7.

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.

Sherbourne 1991

Sherbourne CD, Stewart AL. The MOS Social Support Survey. *Social Science & Medicine* 1991;**32**:705-14.

Snyder 1991

Snyder CR, Harris C, Anderson JR. The will and the ways: development and validation of an individual-differences measure of hope. *Journal of Personality and Social Psychology* 1991;**60**:570-85.

Sterne 2011

Sterne JA, Egger M, Moher D. Chapter 10: Addressing reporting biases. In: Higgins JP, Green S, editor(s). Cochrane Handbook for Systematic Reviews of Intervention. Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from handbook.cochrane.org.

Tandon 2009

Tandon R, Nasrallah HA, Keshavan S. Schizophrenia, "just the facts" 4. Clinical features and conceptualization. *Schizophrenia Research* 2009;**110**(1-3):1-23.

Toljamo 2001

Toljamo M, Hentinen M. Adherence to self-care and social support. *Journal of Clinical Nursing* 2001;**10**:618-27.

Turkington 2004

Turkington D, Dudley R, Warman DM, Beck AT. Cognitive behavioral therapy for schizophrenia: a review. *Journal of Psychiatric Practice* 2004;**10**(1):5-16.

Ukoumunne 1999

Ukoumunne OC, Gulliford MC, Chinn S, Sterne JAC, Burney PGJ. 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.

Ware 1996

Ware J Jr, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Medical Care* 1996;**34**:220-33.

WHOQOL Group 1998

WHOQOL Group. Development of the World Health Organization WHOQOL-BREF quality of life assessment. *Psychological Medicine* 1998;**28**:551-8.

Wildqust 2010

Wildqust HJ, Beary M. Are there modifiable risk factors which will reduce the excess mortality in schizophrenia?. *Journal of Psychopharmacology* 2010;**24**(4 Suppl):37-50.

Wu 2005

Wu EQ, Birnbaum HG, Shi L, Ball DE, Kessler RC, Moulis M, et al. The economic burden of schizophrenia in the United States in 2002. *Journal of Clinical Psychiatry* 2005;**66**(9):1122-9.

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.

Xia 2011

Xia J, Merinder LB, Belgarmar MR. Psychoeducation for schizophrenia. *Cochrane Database of Systematic Reviews* 2011, Issue 6. [DOI: 10.1002/14651858.CD002831.pub2]

Zimmerman 1994

Zimmerman M. Diagnosing personality disorders. A review of issues and research methods. *Archives of General Psychiatry* 1994;**51**(3):225-45.

References to other published versions of this review

Chien 2013

Chien WT, Lui S, Clifton AV. Peer support for schizophrenia. *Cochrane Database of Systematic Reviews* 2013, Issue 12. [DOI: 10.1002/14651858.CD010880]

CHARACTERISTICS OF STUDIES

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

Castelein 2008

Methods	Allocation: randomised Blindness: non-blinded Study duration: 8 months Location: multicentre Design: parallel Setting: not stated Country: Netherlands Consent: written
Participants	Diagnosis: schizophrenia or related psychotic disorders N = 106 History: mean 9.5, SD 8.6 years Sex: men 70, women 36 Age: mean 37.8, SD 10.5 years Exclusion criteria: aged < 18 years; people with drug or alcohol (or both) dependency, possible language difficulties and severe psychotic symptoms.
Interventions	Group 1: peer-support + standard care (n = 56). Content: GPSG + standard care. Participants decided the topic of each session; each session had the same structure discussing daily life experiences in pairs; it was to provide peer-to-peer interaction. Standard care included medication monitoring, psycho-education and supportive counselling. Delivered by: patients with schizophrenia or related psychotic disorder. Frequency: 16 sessions of 90 minutes delivered biweekly over 8 months. Treatment duration: 8 months. Group 2: standard care (n = 50). Content: WLC consisting of standard care alone which included medication monitoring, psycho-education and supportive counselling. Treatment duration: 8 months.
Outcomes	Mental state: self-efficacy, self-esteem Leaving the study early Peer outcomes: impact on participant and peer supporter, quality of life for participant and peer supporter Economic costs: total costs, direct and indirect costs *Unable to use* Hospital admission rates (only P values was reported) Negative symptoms (only P values) Destress from negative symptoms (only P values)

Castelein 2008 *(Continued)*

	Peer outcomes: social network - PNQ (data not reported)
Notes	Funding source: Zon Mw (the Netherlands Organisation for Health Research and Development), the Rob Giel Research Center, and The Roos Foundation. We contacted study authors and got replied.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "Randomised by computer generated random block number." Comment: adequate sequence generation.
Allocation concealment (selection bias)	Low risk	Quote: "In total, 106 participants were randomly allocated per centre to the GPSG or WLC condition by a person not involved in the study or recruitment using numbered, sealed envelopes." Comment: participants and investigators enrolling participants could not foresee assignment.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Quote: "The design of the study did not allow for masking researchers to service assignment. However, we expect this to interfere only minimally with the study results as all questionnaires used were of the self-report type." Comment: blinding of the participants was not ensured.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Quote: "The design of the study did not allow for masking researchers to service assignment…" Comment: blinding of assessors was not ensured.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Quote: "Nine participants (8%) did not complete the follow-up, but these participants did not differ significantly at baseline from those in the study with regard to age, gender, psychotic episodes, duration of illness, educational level, occupational status, or self-reported quality of life scores." Comment: low attrition rate and number of participants leaving early were balanced in groups.
Selective reporting (reporting bias)	High risk	Comment: author did not report the data for the following outcomes: hospital admission rates, negative symptoms, distress from negative symptoms (only P values) and social network – PNQ (data not reported).
Other bias	Low risk	None noted.

Cook 2012a

Methods	Allocation: randomised Blindness: single blinded Study duration: 40 weeks Location: multicentre Design: parallel Setting: outpatient

Cook 2012a *(Continued)*

	Country: USA Consent: written
Participants	Diagnosis: schizophrenia 15.4%, schizoaffective 5.4%, bipolar 39.5%, depressive 18% and other 18.6% N = 428 History: ≥ 12 months Sex: men 190, women 238 Age: mean 42.8, SD 10.9 years Exclusion criteria: not stated
Interventions	Group 1: peer-support + standard care (n = 212). Content: peer-led, mental illness education intervention called Building Recovery of Individual Dreams and Goals through Education and Support (BRIDGES). Classes were delivered interactively, and included group discussion, illustrative anecdotes and structured exercises designed to apply information to everyday situations. Course topics included recovery principles and stages, strategies for building interpersonal and community support systems, brain biology and psychiatric medications, diagnoses and related symptom complexes, traditional and non-traditional treatments and relapse prevention and coping skills. Delivered by: peers who were certified BRIDGES instructors in recovery from severe mental illness. Frequency: 2.5-hour classes delivered weekly for 8 weeks. Treatment duration: 8 weeks. Group 2: standard care (n = 216). Content: participants were assigned to a course waiting list and guaranteed an opportunity to receive BRIDGES once their final interview ended. Otherwise, they continued to receive services as usual. Treatment duration: 8 weeks.
Outcomes	Mental state: hope, other specific aspects Behaviour: recovery, other specific aspects *Unable to use* Global state: leaving the study early (author did not report data by each group separately) Mental state: depression – BSI, personal empowerment, self-advocacy and coping style (data not reported)
Notes	Funding source: US Department of Education, National Institute on Disability and Rehabilitation Research; and the Substance Abuse & Mental Health Services Administration, Center for Mental Health Services, Cooperative Agreement (H133B050003B). We contacted the author to clarify whether peer support group received standard care; however, we received no reply. Therefore, from a prospective of a clinician, the peer support group should have received standard care.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "random assignment occurred using computer-assisted block randomisation stratified according to centre."

Cook 2012a *(Continued)*

		Comment: adequate sequence generation.
Allocation concealment (selection bias)	Low risk	Quote: "A random allocation sequence that was programmed into the Computer Assisted Personal Interviewing (CAPI) administration software guaranteed allocation concealment up to the point of assignment." Comment: participants could not foresee the assignment.
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Quote: "interviewers were blinded to subjects' study condition assignment." Comment: blinding of personnel was likely to be broken, no blinding information for participants.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Quote: "interviewers were blinded to subjects' study condition assignment." Comment: blinding of assessors ensured.
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Quote: "of 212 experimental subjects, 161 (76%) received the intervention and 51 (24%) did not." Comment: moderate attrition rate.
Selective reporting (reporting bias)	High risk	Comment: study protocol registered on ClinicalTrials.gov (NCT01297985). However, the personal empowerment, self-advocacy and coping style data were not reported.
Other bias	Low risk	None noted.

Cook 2012b

Methods	Allocation: randomised Blindness: single blinded Study duration: 40 weeks Location: multicentre Design: parallel Setting: outpatient Country: USA Consent: written
Participants	Diagnosis: schizophrenia (11.7%), schizoaffective (9.5%), bipolar (38.1%) and depressive (25.3%), other (15.4%) N = 555 History: ≥ 12 months Sex: men 177, women 378 Age: mean 45.8, SD 9.8 years Exclusion criteria: not stated
Interventions	Group 1: peer-support + standard care (n = 276).

Cook 2012b *(Continued)*

	Content: peer-led illness self-management intervention called Wellness Recovery Action Planning (WRAP). Course work included lectures, group discussions, personal examples from the lives of the educators and participants, individual and group exercises, and voluntary homework assignments. Session 1: introduction of key concepts of WRAP; sessions 2 and 3: development of personalised wellness strategies; session 4: introduction of a daily maintenance plan to use every day to stay emotionally and physically healthy; session 5: educating of early warning signs; session 6 and 7: creation of a crisis plan specifying signs of impending crisis, names of people willing to help and types of assistance preferred; session 8: post crisis support. Delivered by: peer instructors. Frequency: 2.5-hour sessions delivered weekly Treatment duration: 8 weeks. Group 2: standard care (n = 279). Content: participants were assigned to a waiting list and guaranteed an opportunity to receive WRAP from the study once their interview ended. Otherwise, they continued to receive services as usual. Treatment duration: 8 weeks.
Outcomes	Global state: total endpoint BSI Mental state: hope, positive symptoms, self-efficacy, other specific aspects Leaving the study early Peer outcomes: quality of life for participant and peer supporter Adverse events: death *Unable to use* Personal empowerment, social support, satisfaction (not reported)
Notes	Funding source: US Department of Education, National Institute on Disability and Rehabilitation Research; and the Substance Abuse & Mental Health Services Administration, Center for Mental Health Services, Cooperative Agreement (H133B050003 and H133B100028). We contacted the author to clarify whether peer support group received standard care; however, we received no reply. Therefore, from a prospective of a clinician, the peer support group should have received standard care.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "Randomization was performed by SRL staff at the end of each interview using a random allocation sequence programmed into Computer Assisted Personal Interviewing (CAPI) administration software that allowed for complete allocation concealment up to the point of assignment." Comment: adequate sequence generation.
Allocation concealment (selection bias)	Low risk	Quote: "Randomization was performed by SRL staff at the end of each interview using a random allocation sequence programmed into Computer Assisted Personal Interviewing (CAPI) administration software that allowed for complete allocation concealment up to the point of assignment." Comment: participants could not foresee the assignment.

Cook 2012b *(Continued)*

Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Quote: "Researchers administered structured telephone interviews, and interviewers were blinded to respondents' study condition." Comment: blinding of personnel ensured, no blinding information for participants.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Quote: "interviewers were blinded to respondents' study condition." Comment: blinding of assessors was ensured properly.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Quote: "Eleven control subjects and 25 intervention subjects were lost to follow-up with reasons including death (control=4, intervention= 6) or ill health (control=1, intervention= 3), moving away from the area (control=1, intervention= 3), formal withdrawal from the study (control=4, intervention= 7) and prior intervention exposure (control=1, intervention= 6)." Comment: low attrition rate.
Selective reporting (reporting bias)	High risk	Comment: study registered at ClinicalTrials.gov (NCT01024569). Outcomes such as 'satisfaction' were not reported in the study.
Other bias	Low risk	None noted.

Druss 2010

Methods	Allocation: randomised Blindness: single blinded Study duration: 6 months Location: single centre Design: parallel Setting: outpatients Country: USA Consent: written
Participants	Diagnosis: schizophrenia (28.8%), bipolar disorder (32.5%), major depression (26.3%), PTSD (11.3%) N = 80 History: not stated Sex: men 24, women 56 Age: mean 47.8, SD 10.1 years Exclusion criteria: not stated
Interventions	Group 1: peer-support + standard care (n = 41). Content: 6 group sessions led by peer specialists, discussed the following topics: overview of self-management; exercise and physical activity; pain and fatigue management; healthy eating on a limited budget; medication management; finding and working with a regular doctor. Delivered by: trained peer specialists.

Druss 2010 *(Continued)*

	Frequency: peer support specialists. Treatment duration: 6 months. Group 2: standard care (n = 39). Content: receiving all medical, mental health and peer-based services that they were otherwise receiving prior to entry into the study. Treatment duration: 6 months.
Outcomes	Service use: clinically important engagement Behaviour: patient activation Leaving the study early Peer outcomes: quality of life for participant and peer supporter *Unable to use* Global state: compliance with medication (see What's new)
Notes	Funding source: NIMH R34MH078583. We contacted the author to clarify whether peer support group received standard care; however, we received no reply. Usually outpatients in the usual care setting normally receive usual psychiatric care and thus are not deprived of any standard care or service in the community.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "Using a computerized algorithm, patients were randomised to the intervention or standard care group by the project manager." Comment: adequate sequence generation.
Allocation concealment (selection bias)	Unclear risk	Comment: author did not describe allocation concealment. Insufficient information to permit judgement of low risk or high risk.
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Quote: "The interviewers were blinded to subjects' randomisation status." Comment: blinding of personnel ensured. No blinding information for participants.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Quote: "The interviewers were blinded to subjects' randomisation status." Comment: blinding of assessors ensured
Incomplete outcome data (attrition bias) All outcomes	Low risk	Comment: 4 participants lost from intervention group and 11 participants lost from standard care group (data from Figure 1 of the publication), with reasons such as unable to locate, deceased and withdrawn. Analysis was included for attrition.
Selective reporting (reporting bias)	Unclear risk	Comment: study protocol not available. Insufficient information to make judgement.
Other bias	Low risk	None noted.

Eisen 2012

Methods	Allocation: randomised Blindness: not stated Study duration: 3 months Location: multicentre Design: parallel Setting: inpatients Country: USA Consent: written
Participants	Diagnosis: psychotic disorders, depressive disorder, alcohol-use disorder or substance-use disorder N = 298 History: not stated Sex: men 220, women 78 Age: range 30–60 years Exclusion criteria: not stated
Interventions	Group 1: peer-support + standard care (n = 74). Content: peer facilitators used written recovery material such as the Spanior Recovery Workbook available from the Boston University. Peer leaders also shared their personal experiences as veterans with mental illness. Delivered by: 2 peer facilitators. Frequency: group met for 45 minutes weekly. Treatment duration: 12 weeks. Group 2: clinician-led recovery + standard care group (n = 82). Content: clinician-led recovery group. Delivered by: 1 Master's-level clinician. Treatment duration: 45 minutes weekly. Group 3: standard care (n = 84). Content: treatment as usual group. Details not reported.
Outcomes	Global state: general health (VR-12) Mental state: empowerment, hope, mental health Behaviour: recovery, activation Peer outcomes: social support *Unable to use* Global state: leaving the study early (data were not reported by each group).

Eisen 2012 *(Continued)*

	Mental state: depression, self-harm, emotional liability, interpersonal relationship, psychotic symptom (skewed data) Behaviour: alcohol use (skewed data)
Notes	Funding source: study was supported by the VA Rehabilitation Research and Development Service grant D4464R. We contacted author to clarify the proportion of schizophrenia but received no reply.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "Veterans were randomly assigned to...", "the random assignment was in an envelope." Comment: trials were randomised with allocation concealment. Under this condition, we assumed the author did the random sequence generation adequately.
Allocation concealment (selection bias)	Low risk	Quote: "the random assignment was in an envelope." Comment: allocation concealment ensured.
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Comment: the author did not describe the blinding of participants and personnel. Insufficient information to make judgement.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Comment: the author did not describe the blinding of outcome assessment. Insufficient information to make judgement.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Quote: "Of these (298 veterans), 240 (81%) completed the three-month follow-up and were included in the analyses." Comment: low attrition rate and participants leaving early were balanced in groups. Analysis of attrition was included in the study.
Selective reporting (reporting bias)	Unclear risk	Comment: protocol not available. Insufficient information to make judgement.
Other bias	Low risk	None noted.

Goldberg 2013

Methods	Allocation: randomised Blindness: not stated Study duration: 5 months Location: multicentre Design: parallel Setting: outpatients

Goldberg 2013 *(Continued)*

	Country: USA Consent: written
Participants	Diagnosis: bipolar disorder, schizophrenia, major depression, or post-traumatic stress disorder N = 63 History: not stated Sex: men 30, women 33 Age: mean 49.5, SD 9.1 years Exclusion criteria: not stated
Interventions	Group 1: peer support + standard care (n = 32). Content: Living Well, the adapted programme. An advisory panel comprising a mental health consumer and study investigators met every other week for 3 months (July–September 2007) to consider modifications of the original CDSMP [Chronic Disease Self-Management Program] intervention for outpatients with serious mental illness. Delivered by: peers. Frequency: 60- to 75-minute sessions delivered weekly. Treatment duration: 13 weeks followed by a 2-month booster. Group 2: standard care (n = 31). Content: not stated. Treatment duration: 5 months.
Outcomes	Service use: clinically important engagement, use of emergency services Behaviour: activation, approach to health care, self-efficacy, recovery, healthy eating, self-management, behaviours, internal locus of control for health Leaving the study early Functioning: general, physical, emotional well-being Peer outcomes: social support *Unable to use* Global state: compliance with medication (see What's new) Mental state: behavioural and cognitive symptoms (skewed data) Behaviour: physical activity (skewed data) Functioning: Instrument to Measure Self-Management, skewed data
Notes	Funding source: Grant MH078168. We contacted author to clarify the proportion of schizophrenia but received no reply.

Risk of bias

Bias	Authors' judgement	Support for judgement

Goldberg 2013 *(Continued)*

Random sequence generation (selection bias)	Unclear risk	Quote: "Of 63 participants, 32 were randomly assigned to Living Well program and 31 to standard care." Comment: insufficient information to make judgement.
Allocation concealment (selection bias)	Unclear risk	Comment: author did not describe allocation concealment. Insufficient information to make judgement.
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Comment: the author did not describe the blinding of participants and personnel. Insufficient information to make judgement.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Comment: the author did not describe the blinding of outcome assessment. Insufficient information to make judgement.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Quote: "Of the 63 participants in the total sample, 58 (92%) completed the postintervention assessment and 57 (90%) completed the two-month follow-up assessment. Follow-up rates did not differ significantly between conditions." Comment: low attrition rate. Attrition rates were balanced in groups.
Selective reporting (reporting bias)	Unclear risk	Comment: study protocol not available. Insufficient information to make judgement.
Other bias	Low risk	None noted

Kelly 2014

Methods	Allocation: randomised Blindness: not stated Study duration: 6 months Location: multicentre Design: parallel Setting: not stated Country: USA Consent: written
Participants	Diagnosis: serious mental illness N = 24 History: not stated Sex: men 13, women 11 Age: mean 46.78, SD 8.45 years

Kelly 2014 *(Continued)*

	Exclusion criteria: participants could not be under conservatorship (a legal concept in the US) (required by the County Department of Mental Health), unable to give informed consent or be hospitalised at the start of the study.
Interventions	Group 1: peer-support + standard care (n = 12). Content: manualised intervention. Navigators encouraged development of self-management of health-care through a series of psychoeducation and behavioural strategies. Delivered by: peers. Frequency: not stated. Treatment duration: 6 months. Group 2: standard care (n = 12). Content: usual mental health services. Treatment duration: 6 months.
Outcomes	Leaving the study early *Unable to use* Service use: contact with services (skewed data) Global state: compliance to medication (skewed data) Health issues: bodily pain, bodily pain interference, total number of health problem, health lack of efficacy, preferred location of care, number of physical medications (not predefined outcomes for this review).
Notes	Funding sources: funded with support from the UniHealth Foundation.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: "Participants were recruited in group of six and then randomized (by the project manager) using a random numbers table."
Allocation concealment (selection bias)	Unclear risk	Comment: author did not describe allocation concealment. Insufficient information to make judgement.
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Comment: the author did not describe the blinding of participants and personnel. Insufficient information to make judgement.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Comment: the author did not describe the blinding of outcome assessment. Insufficient information to make judgement.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Comment: 3 participants in the control group left the study early. Low attrition rate.
Selective reporting (reporting bias)	Unclear risk	Comment: study protocol not available. Insufficient information to make judgement.

Kelly 2014 *(Continued)*

Other bias	Low risk	None noted.

Mahlke 2017

Methods	Allocation: randomised Blindness: not stated Study duration: 12 months Location: multicentre Design: parallel Setting: inpatients and outpatients Country: Germany Consent: written
Participants	Diagnosis: severe mental illness (28% schizophrenia and schizoaffective disorders, other diagnose including bipolar disorder, unipolar depression, or personality disorder) N = 216 History: > 2 years Sex: men 92, women 124 Age: mean 41.48, SD 12.28 years Exclusion criteria: primary diagnosis of drug or alcohol abuse and insufficient command of German to communicate with the peer supporters.
Interventions	Group 1: peer-support group + standard care (n = 114). Content: 1-to-1 peer support in addition to treatment as usual. Peer supporters contacted patients within the first week after randomisation and then established 1-to-1 meetings. The minimum number of meetings required to build a supporting relationship and be effective for the patient, based on the experiences in delivering support by the peers themselves. Delivered by: peers and staff, who trained by a peer worker and a psychologist. Frequency: 4 and 26 times for 1 hour over a 6-month period. Treatment duration: 6 months. Group 2: standard care (n = 102). Content: inpatient and outpatient care with infrequent meetings with outpatient psychiatrists, and access to community-based groups and separate psychological treatments. Treatment duration: 6 months.
Outcomes	Service use: duration of hospital stay Global state: severity of illness (Clinical Global Impression scale) Behaviour: self-efficacy Leaving the study early

Mahlke 2017 *(Continued)*

	Functioning: general Peer outcomes: quality of life for participant and peer supporter
Notes	Funding sources: part of the 'psychenet' project (www.psychenet.de) and received funding from the Federal Ministry of Education and Research in Germany from 2011 to 2015 (BMBFNr: O1KQ1002B).

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "The participants were randomly allocated to either one-to-one peer support or the control group in a 1:1 ratio, stratified by hospital, in blocks of 20." Comments: adequate sequence generation.
Allocation concealment (selection bias)	Low risk	Quote: "An independent statistician, ... produced randomly generated treatment allocations ... within sealed, numbered, opaque envelopes that were stored and inaccessible to the trial team." Comments: allocation concealment ensured.
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Comment: the author did not describe the blinding of participants and personnel. Insufficient information to permit judgement of low risk or high risk.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Comment: the author did not describe the blinding of outcome assessment. Insufficient information to permit judgement of low risk or high risk.
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Quote: "Whilst the dropout rate on the primary outcome was 31%, it was substantially higher on the clinician-rated secondary outcomes." Comments: moderate attrition rate.
Selective reporting (reporting bias)	High risk	Comment: study protocol registered on ClinicalTrials.gov (NCT02276469). Illness of management and satisfaction of the client was not reported.
Other bias	Low risk	None noted.

Qian 2015

Methods	Allocation: randomised Blindness: not stated Study duration: 5 weeks Location: single centre Design: parallel Setting: community Country: China Consent: written

Qian 2015 *(Continued)*

Participants	Diagnosis: schizophrenia N = 100 History: 4–13 years Sex: men 69, women 31 Age: mean 25.23, SD 8.51 years Exclusion criteria: serious physical disorder, brain organic disease
Interventions	Group 1: peer-support + standard care group (n = 50). Content: peer support and psychoeducation. Delivered by: trained peer. Frequency: 5 weekly sessions. Treatment duration: 5 weeks. Group 2: standard care (n = 50). Content: psychoeducation. Treatment duration: 5 weeks.
Outcomes	Peer outcomes: quality of life for participant and peer supporter
Notes	Funding sources: not stated.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Comments: randomised.
Allocation concealment (selection bias)	Unclear risk	Comments: the author did not describe the blinding of participants and personnel. Insufficient information to make judgement.
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Comment: the author did not describe the blinding of participants and personnel. Insufficient information to make judgement.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Comment: the author did not describe the blinding of outcome assessment. Insufficient information to make judgement.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Comments: all participants completed the trial.
Selective reporting (reporting bias)	Unclear risk	Comments: study protocol available. Insufficient data to make judgement.
Other bias	Low risk	None noted.

Reynolds 2004

Methods	Allocation: randomised Blindness: non-blinded Study duration: 5 months Location: single centre Design: parallel Setting: discharged inpatients Country: UK Consent: written
Participants	Diagnosis: range of mental illnesses, including bipolar disorder, schizophrenia and depression N = 25 History: not stated Sex: not stated Age: not stated Exclusion criteria: people with dementia, people who were discharged from hospital before having had the opportunity to develop a relationship with their transitional nurse
Interventions	Group 1: peer-support + standard care (n = 11). Content: peer support, which was assistance from former patients who provide friendship, understanding and encouragement; and overlap of inpatient and community staff in which the inpatient staff continue to work with the discharged patient until a working relationship was established with a community care provider. Delivered by: previous service user of the mental health system. Frequency: type and intensity of assistance provided by the peer supporter varied according to individual preference Treatment duration: 5 months. Group 2: standard care (n = 14). Content: usual treatment, comprised the standard discharge arrangements normally provided to patients and included referral to locality-based community psychiatric nurses. Treatment duration: 5 months.
Outcomes	Service use: hospital admission Leaving study early Functioning: general, physical, societal role, interpersonal functioning, cognitive Peer outcomes: quality of life for participant and peer supporter *Unable to use* Mental state: aggressiveness, anxiety, attention problems, depression, emotional withdrawal, family problems, hyperaffect, interpersonal problems, resistiveness, suicide feelings, thought process difficulties (skewed data)

Reynolds 2004 *(Continued)*

	Functioning: daily living (skewed data)
Notes	Funding source: Chief Scientist Office, Scottish Executive.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "Subjects were randomly assigned to groups by a computerized random number facility." Comment: adequate sequence generation.
Allocation concealment (selection bias)	Unclear risk	Comment: author did not describe allocation concealment. Insufficient information to make judgement.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Quote: "The researchers were not blinded to the intervention status of participants." Comment: personnel were not blinded. No information for blinding of participants.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Quote: "The researchers were not blinded to the intervention status of participants." Comment: the blinding of assessors was not ensured.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Quote: "A small number of patients were lost to study (control n= 3; experimental n=3) and consequently data on 19 subjects were included in the final analysis." Comment: low attrition rate, rates were balanced in groups.
Selective reporting (reporting bias)	Unclear risk	Comment: study protocol not available. Insufficient information to make judgement.
Other bias	Low risk	None noted.

Rowe 2007

Methods	Allocation: randomised Blindness: not stated Study duration: 12 months Location: single centre Design: parallel Setting: not stated Country: USA Consent: written
Participants	Diagnosis: psychotic disorder, major mood disorder, alcohol-use disorder, drug-use disorder, other disorder

Rowe 2007 *(Continued)*

	N = 114 History: not stated Sex: men 78, women 36 Age: mean 39.8, SD 8.8 years Exclusion criteria: not stated
Interventions	Group 1: peer-support + standard care group (n = 73). Content: standard service and peer support which included citizenship intervention plus valued-roles projects. Consist of classes with topics related to social participation and community integration (citizenship classes), followed by projects designed to foster participants' acquisition of valued social roles (valued-roles projects). Delivered by: peer mentor. Frequency: mean once weekly. Treatment duration: 4 months. Group 2: standard care (n = 41). Content: standard service, individual and group treatment with medication management, case management and jail diversion services Treatment duration: 4 months.
Outcomes	*Unable to use* Functioning: criminal justice involvement, alcohol use, drug use (skewed data)
Notes	Funding source: Yale University Institution for social and policy studies.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: "A total of 114 adults with serious mental illness participated in a 2×3 prospective longitudinal, randomised clinical trial with two levels of intervention." Comment: insufficient information to make judgement.
Allocation concealment (selection bias)	Unclear risk	Comment: author did not describe allocation concealment. Insufficient information to make judgement.
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Comment: author did not describe blinding of participants and personnel. Insufficient information to make judgement.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Comment: author did not describe blinding of outcome assessment. Insufficient information to make judgement.
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Quote: "the overall sample showed 23% attrition from time 1, with 20 participants missing the time 2 (six-month) interview but returning for the time 3 (12-month) interview and 19 participants missing the time 3 interview."

Rowe 2007 *(Continued)*

		Comment: moderate attrition rate.
Selective reporting (reporting bias)	Unclear risk	Comment: study protocol not available. Insufficient information to make judgement.
Other bias	Low risk	None noted.

Sells 2008

Methods	Allocation: randomised Blindness: not stated Study duration: 12 months Location: single centre Design: parallel Setting: not stated Country: USA Consent: written
Participants	Diagnosis: psychotic disorder, major mood disorder, substance use disorder, co-occurring disorders N = 137 History: not stated Sex: men 84, women 53 Age: mean 41, SD 9 years Length of illness: not stated Exclusion criteria: not stated
Interventions	Group 1: peer-based intensive case management group (n = 68). Content: peers used past experiences with recovery as a tool for understanding, role modelling and hope building for others. Participants received 1 year of service from intensive case management teams that included peer providers as primary contacts. Delivered by: peer providers who had severe mental illness history. Frequency: not stated. Treatment duration: 12 months. Group 2: traditional intensive case management group (n = 69). Content: traditional intensive case management. Treatment duration: 12 months.
Outcomes	Peer outcomes: impact on participant and peer supporter *Unable to use* Leaving the study early (only missing data in scale)

Sells 2008 *(Continued)*

	Behaviour: drug and alcohol use (no data were reported) Peer outcomes: favourable therapeutic relationship change (no SD data were reported), quality of life for participant and peer supporter (no data reported)
Notes	Funding source: Yale Institution for Social and Policy Studies; the peer-based treatment option was sponsored by the Connecticut Department of Mental Health and Addiction Services.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: "investigators randomly assigned participants to either the experimental (peer provider) or control (regular treatment) condition." Comment: insufficient information to make judgement.
Allocation concealment (selection bias)	Unclear risk	Comment: author did not describe allocation concealment. Insufficient information to make judgement.
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Comment: author did not describe blinding of participants and personnel. Insufficient information to make judgement.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Comment: author did not describe blinding of outcome assessment. Insufficient information to make judgement.
Incomplete outcome data (attrition bias) All outcomes	High risk	Comment: 26 participants left early from the intervention group and 38 participants left early from the control group (data extracted from Table 2); total attrition rate in control group was higher than 50%. Reasons for missing outcome data not reported.
Selective reporting (reporting bias)	High risk	Comment: study protocol not available. Author did not report adequate data for favourable therapeutic relationship change, quality of life, and drug and alcohol use.
Other bias	Low risk	Not noted.

Van Gestel-Timmermans 2012

Methods	Allocation: randomised Blindness: single blinded Study duration: 6 months Location: multicentre Design: parallel Setting: inpatients or outpatients Country: Netherlands Consent: written
Participants	Diagnosis: psychotic disorder, affective disorder, anxiety disorder, personality disorder

Van Gestel-Timmermans 2012 *(Continued)*

	N = 333 Sex: men 113, women 220 Age: mean 43, SD 11 years Length of illness: not stated Inclusion criteria: psychosis, personality disorder, affective disorder, anxiety disorder, addiction problems, eating disorders or other psychiatric problems; self-report of having experienced disruptive periods in life from which the person was recovering Exclusion criteria: illiteracy, inability to speak Dutch, suicidal ideation, florid psychotic symptoms or substance abuse during the peer-run course
Interventions	Group 1: peer-support group + standard care (n = 168). Content: instructors closely followed a standardised manual, which precisely described the goals of each session and the steps to attain the goals. Each session had the same structure and was organised around a specific, recovery-related theme, such as the meaning of recovery to participants, personal experiences of recovery, personal desires for the future, making choices, goal setting, participation in society, roles in daily life, personal values, how to get social support, abilities and personal resources, and empowerment and assertiveness. Delivered by: people in an advanced state of their recovery process. Frequency: 2-hour sessions delivered weekly. Treatment duration: 12 weeks. Group 2: standard care (n = 165). Content: participants received treatment as usual. Treatment duration: 12 weeks.
Outcomes	Mental state: hope, self-efficacy, empowerment, loneliness Leaving the study early Peer outcomes: quality of life for participant and peer supporter
Notes	Funding source: supported by grant 100003017 from the Netherlands Organization for Health Research and Development (ZonMw).

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "Participants were randomly assigned to the experimental or control condition by a research assistant who drew lots." Comment: adequate sequence generation.
Allocation concealment (selection bias)	Unclear risk	Comment: author did not describe allocation concealment. Insufficient information to make judgement.
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Quote: "participants were assigned numbers so that researchers and research assistants were blind to their condition." Comment: blinding of personnel ensured, but no information for blinding of participants.

Van Gestel-Timmermans 2012 *(Continued)*

Blinding of outcome assessment (detection bias) All outcomes	Low risk	Quote: "researchers and research assistants were blind to their condition." Comment: blinding of assessors was ensured.
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Quote: "Overall rates of dropout from the study were 20% and 30% at three and six months, respectively, with significantly more dropout in the control condition than in the experimental condition (35% versus 25% at six months, P=.01)." Comment: moderate attrition rate. Attrition rate was not balanced in groups.
Selective reporting (reporting bias)	High risk	Comment: trial registration number ISRCTN47331661. However, social support, coping and goal-setting skills were not reported.
Other bias	Low risk	None noted.

BSI: Brief Symptom Inventory; GPSG: guided peer support group; ICC: intraclass correlation coefficients; n: number of participants; PNQ: Personal Network Questionnaire; SD: standard deviation; WLC: waiting-list condition.

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

Study	Reason for exclusion
Buchkremer 1995	Allocation: randomised Participants: people with schizophrenia Interventions: not peer support. Therapeutic relatives' group intervention vs Initiated relatives' group intervention
Chen 2016	Allocation: randomised Participants: people with schizophrenia Interventions: not a peer support. Integrated intervention includes psychoeducation led by professionals, patient group discussion, psychoeducation to the families of patients
Chinman 2015	Allocation: randomised Participants: people with schizophrenia Interventions: Mental Health Intensive Case Management + peer-support group vs Mental Health Intensive Case Management Outcome: no usable data
Corrigan 2017a	Allocation: randomised Participants: not a majority of people diagnosed with schizophrenia, only 9.0%
Corrigan 2017b	Allocation: randomised Participants: not a majority of people diagnosed with schizophrenia, only 10.0%
Craig 2004	Allocation: randomised Participants: chronic psychotic illnesses, with paranoid schizophrenia as the most common diagnosis

Study	Reason for exclusion
	Interventions: not peer support, but standard case management vs standard case management + healthcare assistant
Forchuk 2005	Allocation: randomised Participants: schizophrenia, mood disorder, substance related, personality disorder, anxiety disorder, developmental delay, organic disorder Interventions: peer-support + standard care group vs standard care Outcome: no usable data
Gunter 1983	Allocation: quasi-randomised RCT
Hazell 2016	Allocation: RCT protocol Participants: schizophrenia Interventions: CBT vs control
ISRCTN14282228	Allocation: randomised Participants: schizophrenia Interventions: not peer support but an nurse-led intervention that combined home-based skill training with nurse-guided peer-support intervention
Kaplan 2011	Allocation: randomised Participants: 22% with schizophrenia spectrum disorder and 78% affective disorder Interventions: peer support via listserv + standard care group vs peer support via bulletin board + standard care group vs waiting list + standard care Outcome: no usable data
Kaufmann 1995	Allocation: not randomised; because of low rate of participation, first randomised experiment was ended and the second analysis compared participating and non-participating participants in previous intervention group.
Killackey 2013	Allocation: not randomised. Methodology study
Klein 1998	Allocation: not randomised
NCT02974400	Allocation: randomised Participants: schizophrenia, hallucinations, persecutory delusion Interventions: CBT vs wait list
O'Connell 2017	Allocation: randomised Participants: schizophrenia-spectrum disorders and affective disorders with psychotic features Interventions: peer-support + skills training group vs skills training Outcome: no usable data
Rivera 2007	Allocation: randomised Participants: mental illness

Study	Reason for exclusion
	Interventions: strength-based intensive case management with peer enhancement vs strength based intensive case management without peer enhancement vs clinic-based care. The peer enhancement intervention did not focus on peer support.
Rogers 2012	Allocation: not randomised. Study report discussed 1 review, 1 non-completed RCT, 1 non-randomised study and 1 ongoing study.
Salyers 2010	Allocation: randomised Participants: DSM-IV diagnosis on Axis I of 295-296 (schizophrenia, bipolar disorder, and other major mood disorders) Interventions: peer-support + assertive community treatment group vs assertive community treatment group Outcome: no usable data
Segal 2010	Allocation: randomised Participants: people with serious mental illness (76% diagnosis of major depression)
Shahar 2006	Allocation: randomised Participants: psychotic disorder, affective disorder, comorbid substance use disorder. diagnoses were based on the Structured Clinical interview for DSM-III-R. Interventions: peer-support + standard care group vs non-consumer partner + standard care group vs standard care group Outcome: no usable data
Streicker 1984	Allocation: randomised Participants: psychiatric patients Interventions: medication education vs control
Verhaegh 2006	Allocation: quasi-randomised RCT
Weissman 2005	Allocation: randomised Participants: veterans with severe mental illness; clinically diagnosed Axis I psychiatric disorder Interventions: peer-support + usual case management group vs standard care Outcome: no usable data
Zhou 2016	Allocation: not randomised, randomisation based on the admission sequence

CBT: cognitive-behavioural therapy; DSM-III-R: Diagnostic and Statistical Manual of Mental Disorders, 3rd edition, Revised; DSM-IV: Diagnostic and Statistical Manual of Mental Disorders, 4th edition; RCT: randomised controlled trial.

Characteristics of studies awaiting assessment *[ordered by study ID]*

Daumit 2010

Methods	Allocation: randomised Blindness: not stated

Daumit 2010 *(Continued)*

	Duration: 4 months Location: urban adult community psychiatry clinics
Participants	Diagnosis: people with severe mental illness n = 93 Age: mean 47 years Sex: both History: not stated Exclusion criteria: not stated
Interventions	Group 1: group exercise + peer support Group 2: group exercise alone: fitness instructors led exercise classes
Outcomes	Primary and secondary outcomes: none reported Unable to use: cardiorespiratory fitness, walk test and exercise self-efficacy: not in protocol, data not available
Notes	Abstract presented at a conference and published in a supplementary issue of a journal. We have contacted the authors to enquire whether there are any available data and are awaiting a response.

Kroon 2011

Methods	Allocation: randomised Blindness: not stated Duration: 2 years Location: not stated
Participants	Diagnosis: not stated n = 175 Age: not stated Sex: not stated History: not stated Exclusion: not stated
Interventions	Group 1: user-led recovery group Group 2: short recovery courses, added to standard care
Outcomes	Not stated
Notes	Conference proceeding, full characteristics and outcome data not reported. We have contacted the authors to enquire whether there are any available data and are awaiting a response.

NCT00458094

Methods	Allocation: randomised Blindness: single Duration: 4 months Location: not stated
Participants	Diagnosis: people with serious mental illnesses n = 100 Age: 18–70 years Sex: both History: not stated Exclusion: any condition that would make weight loss medically inadvisable; diagnosis of or treatment for cancer (except non-melanoma skin cancer) within 2 years prior to study entry; liver failure; history of anorexia nervosa; pregnant or planning to become pregnant during the study; inability to walk or participate in an exercise class; consumes > 14 alcoholic drinks per week; symptoms of angina or a cardiovascular event within 6 months prior to study entry.
Interventions	Group 1: physical activity intervention with peer support: 3 exercise sessions each week for 4 months and meeting with a peer educator once a week for 15 minutes Group 2: physical activity intervention without peer support: 3 weekly exercise sessions for 4 months
Outcomes	Primary outcome: cardiorespiratory fitness Secondary outcome: weight, waist circumference, physical activity, health status, centre for epidemiology depression scale, exercise-related self-efficacy, general perceived efficacy, participation
Notes	This study has been completed, however no data reported. We have contacted the authors to enquire whether there are any available data and are awaiting a response.

NTR1166

Methods	Allocation: randomised Blindness: not stated Duration: 18 months Location: Netherlands
Participants	Diagnosis: outpatients with psychotic or bipolar disorders and at risk of psychiatric crises n = not stated Age: 18–65 years Sex: both History: experienced ≥ 1 psychiatric crisis during the previous 2 years

NTR1166 *(Continued)*

	Exclusion criteria: having a somatic disease causing a psychotic disorder, inability to give informed consent because of mental incapacity, insufficient command of the Dutch language and already having a 'relapse prevention plan' or a 'crisis plan'
Interventions	Group 1: patients who create a crisis plan with a patient's advocate Group 2: patients create a crisis plan with their clinician only Group 3: patients do not create a crisis plan
Outcomes	Primary outcomes: number of emergency (after hour) visits, (involuntary) admissions and the length of stay in hospital Secondary outcomes: psychosocial functioning and treatment satisfaction
Notes	Protocol, full characteristics and outcome data not reported. We have contacted the authors to enquire whether there are any available data and are awaiting a response.

Robinson 2010

Methods	Allocation: randomised Blindness: the research assistant, who carries out the assessments, will be blind to group allocation. Location: Australia and New Zealand Duration: 18 months
Participants	Diagnosis: not stated n = 36 History: first-episode psychosis Age: 15–24 years Sex: both Exclusion criteria: not stated
Interventions	Group 1: 6-month peer-support intervention delivered to young people with first-episode psychosis over the period of discharge Group 2: treatment as usual
Outcomes	Primary outcomes: levels of engagement and treatment adherence, perceived social support, quantity and quality of service-related information received and service satisfaction. Secondary outcomes: suicide risk (presence of current or recent suicidal ideation or suicidal behaviour including deliberate self-harm, or both).
Notes	Protocol, full characteristics and outcome data not reported. We have contacted author to enquire what the population of schizophrenia is and are awaiting a response.

Tondora 2010

Methods	Allocation: randomised

Tondora 2010 *(Continued)*

	Blindness: not stated Duration: 6 months Location: Community Mental Health Center, USA
Participants	Diagnosis: with a current or past diagnosis consistent with the DSM-IV-TR schizophrenia or schizoaffective disorder, or a current or past diagnosis of psychosis as a part of another Axis I disorder (e.g. bipolar affective disorder with psychotic features) n = 360 History: duration not stated Age: ≥ 18 years Sex: both Exclusion criteria: presence of an organic brain syndrome or dementia
Interventions	Group 1: standard care incorporating illness management Group 2: standard care + facilitation of person-centred care Group 3: illness management/person-centred care + community inclusion
Outcomes	Primary outcome: none Secondary outcomes: community engagement, satisfaction with treatment, symptom distress, ethnic identity, personal empowerment and quality of life
Notes	Study protocol, full characteristics and outcome data not reported. Contacted author to enquire whether there are any available data.

DSM-IV-TR: Diagnostic and Statistical Manual of Mental Disorders, 4th Edition, Text Revision; n: number of studies.

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

ACTRN1261200097

Trial name or title	Peer delivered support intervention for people who hear voices: pilot randomised controlled trial
Methods	Allocation: randomised by computerised sequence generation (treatment allocation made independently via email) Blindness: blinded Duration: 6 months Location: not stated
Participants	Diagnosis: schizophrenia, psychotic disorders, auditory verbal hallucinations n = 35 Age: 18–65 years Sex: both History: not stated

ACTRN1261200097 *(Continued)*

	Exclusion: recent (past 8 weeks) or planned change in antipsychotic medication; currently receiving individual psychological therapy; insufficient English or intellectual functioning to meaningfully participate
Interventions	Group 1: intervention: 12 weekly 1-hour 1-to-1 sessions, from a peer mental health worker who has had experience of hearing voices themselves + treatment as usual Group 2: receive the intervention after a 3-month treatment as usual wait list period
Outcomes	Primary outcome: subjective Experiences of Psychosis Scale Secondary outcome: RAS
Starting date	2012
Contact information	Neil Thomas: neilthomas@swin.edu.au
Notes	Protocol, full characteristics and outcome data not reported. Contacted author for more data, but was told the trial is still ongoing.

Chinman 2017

Trial name or title	Provision of peer specialist services in VA patient-aligned care teams
Methods	Allocation: cluster-randomised Blindness: open-label Duration: 1 year Location: US
Participants	Diagnosis: mental illness, physical illness n = 25 Age: child, adult, senior Sex: both History: not stated Exclusion criteria: non-VA patient aligned care teams, VA sites without an existing peer specialists, and VA patient aligned primary care teams that cannot commit a peer specialist to primary care for a minimum of 10 hours per week.
Interventions	Group 1: facilitated implementation: facilitated implementation sites will receive 1 year of support based on the i-PARIHS implementation model which includes training, implementation planning, ongoing external facilitation, feedback and consultation Group 2: standard implementation: standard Implementation sites will receive written guidance and limited consultation by the investigators' team
Outcomes	Primary outcome: patient activation measure change Secondary outcomes: team development measure change, organisational readiness for change, peer fidelity measure change, the satisfaction Index-Mental health change
Starting date	1 January 2016

Chinman 2017 *(Continued)*

Contact information	Chinman@rand.org
Notes	Funding sources: all the authors are funded by a grant from the Department of Veterans Affairs (QUERI): QUERI for Team-Based Behavioral Health (1IP1HX001979-01): Evaluation of Peer Specialists on VA PACT. Protocol, full characteristics and outcome data not reported.

NCT01566513

Trial name or title	Effectiveness and cost effectiveness of peer mentors in reducing hospital use (Project PEP).
Methods	Allocation: randomised Blindness: open label Duration: 9 months Location: not stated
Participants	Diagnosis: serious mental illness n = 320 Age: ≥ 18 years Sex: both History: not stated Exclusion criteria: dementia or other organic condition limiting ability to provide informed consent
Interventions	Group 1: no intervention, treatment as usual Group 2: behavioural: community connector Group 3: behavioural: peer recovery mentor Group 4: behavioural: peer case manager
Outcomes	Primary outcome measures: service use Secondary outcome measure: psychiatric symptoms, quality of life, community inclusion, psychiatric symptoms, quality of life, community inclusion
Starting date	August 2011
Contact information	larry.davidson@yale.edu
Notes	Protocol, full characteristics and outcome data not reported.

NCT02958007

Trial name or title	Peer support for exercise in older veterans with psychotic disorders
Methods	Allocation: randomised

NCT02958007 *(Continued)*

	Blindness: blinding of outcomes assessor Duration: 12 weeks Location: US
Participants	Diagnosis: psychotic disorder n = not stated Age: ≥ 50 years Sex: both History: not stated Exclusion criteria: • current participation in a supervised exercise programme; • medical conditions which would preclude exercise participation including: unstable angina, proliferative diabetic retinopathy, open wounds poorly controlled type 2 diabetes (HbA1c > 9%), current treatment for active cancer, New York Heart Association Stage II–IV heart failure, dialysis for chronic kidney disease, myocardial infarction in the previous 3 months; • inability to complete the Graded Exercise Treadmill Test; • positive cardiac stress test, unless symptomatic coronary artery disease is ruled out by imaging studies; • problematic substance abuse/dependence; • imminent risk of suicidal or homicidal behaviour; • lack of capacity to consent.
Interventions	Group 1: PEER: 24-week group-based peer coaching intervention delivered by a VA peer specialist, to promote participation in a supervised fitness training programme and general physical activity. Group 2: enhanced supervised fitness training: 24-week intervention to promote participation in a supervised fitness training programme and general physical activity, which includes individual support from non-peer staff.
Outcomes	Primary outcomes • Percent of participants randomised to PEER who attend ≥ 3 group sessions • Percent of sampled PEER group sessions in which the peer coaches were adequately adherent (i.e. mean score equal to 'acceptable' and no items scored as 'unacceptable') on the PEER fidelity measure • Attendance: mean number of supervised fitness training sessions attended • Change from baseline in Ambulatory Physical Activity • Change from baseline in maximal aerobic capacity (VO_{2max})
Starting date	June 2018
Contact information	Anjana.Muralidharan2@va.gov
Notes	Not yet recruiting

NCT02989805

Trial name or title	Engaging patients with mental disorders from the emergency department in outpatient care (EPIC).

NCT02989805 *(Continued)*

Methods	Allocation: randomised Blindness: open-label Duration: 12 months Location: US
Participants	Diagnosis: mental disorder n = 1000 Age: ≥ 18 years Sex: both History: not stated Exclusion criteria: cognitive impairment, unable to speak English
Interventions	Group 1: peer specialist care manager: each participating site will have a peer specialist to provide care management. Peer specialists will have a minimum of a high school education, history of a mental illness, be self-described as 'in recovery,' and have reliable transportation to the study site. All certified peer specialists will receive training in a curriculum that supports identifying and pursuing goals for recovery; developing and documenting recovery-focused treatment plans; and supporting linkages with community-based services. Peers will learn to help other people with mental health conditions to facilitate mental health dialogues; explore mental health choices and options; identify and work with a clinician; and obtain access to community health supports. Group 2: professional care manager: each participating site will have a nurse or social worker to provide care management. Training activities will include modules for each of the key domains covered in the intervention: shared decision making, action planning; motivational interviewing; and mental health as a cornerstone of recovery, working effectively within the mental health system; and self-care and stress management.
Outcomes	Primary outcome: outpatient treatment engagement after emergency department discharge Secondary outcomes: outpatient engagement, change in Patient-Reported Outcomes Measurement Information System (PROMIS) scores, change in RAS score, change in Barriers to Care Survey score
Starting date	3 April 2017
Contact information	bdruss@emory.edu
Notes	Recruiting

HbA1c: glycated haemoglobin; i-PARIHS: integrated – Promoting Action on Research Implementation in Health Services; PEER: Peer Education on Exercise for Recovery; RAS: Recovery Assessment Scale; VA: Veterans Affairs.

DATA AND ANALYSES

Comparison 1. Peer support + standard care versus standard care alone

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Service use: 1a. Hospital admission – medium term	1	19	Risk Ratio (M-H, Random, 95% CI)	0.44 [0.11, 1.75]
2 Service use: 1b. Hospital admission – duration of hospital stay (days) – long term (skewed data)			Other data	No numeric data
3 Service use: 2a. Clinically important engagement with services – medium term	2		Risk Ratio (M-H, Random, 95% CI)	Subtotals only
3.1 Use of emergency care	1	57	Risk Ratio (M-H, Random, 95% CI)	0.39 [0.11, 1.32]
3.2 ≥ 1 primary care visit	1	80	Risk Ratio (M-H, Random, 95% CI)	1.77 [1.09, 2.85]
4 Service use: 2b. Contact with services – medium term (skewed data)			Other data	No numeric data
4.2 Mean number of emergency visits			Other data	No numeric data
4.3 Mean number of routine care visits			Other data	No numeric data
5 Global state: 3a. General Health – mean total endpoint score (Veterans RAND 12-Item Health Survey (VR-12), high = good)	1	158	Mean Difference (IV, Random, 95% CI)	-0.02 [-3.96, 3.92]
5.1 Medium term	1	158	Mean Difference (IV, Random, 95% CI)	-0.02 [-3.96, 3.92]
6 Global state: 3b. Severity of illness – mean total endpoint score (Brief Symptom Inventory (BSI), high = poor)	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
6.1 Medium term	1	458	Mean Difference (IV, Random, 95% CI)	-0.13 [-0.25, -0.01]
6.2 Long term	1	440	Mean Difference (IV, Random, 95% CI)	0.0 [-0.11, 0.11]
7 Global state: 3c. Severity of illness – mean total endpoint score (Clinical Global Impression scale (CGI), high = poor)	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
7.1 Medium term	1	216	Mean Difference (IV, Random, 95% CI)	-0.30 [-0.53, -0.07]
7.2 Long term	1	216	Mean Difference (IV, Random, 95% CI)	0.40 [0.15, 0.65]
8 Global state: 4. Compliance with medication (skewed data)			Other data	No numeric data
8.1 Number of medication			Other data	No numeric data

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
9 Adverse event: 1. Death – all cause (long term)	1	555	Risk Ratio (M-H, Random, 95% CI)	1.52 [0.43, 5.31]
10 Mental state: 1a. Specific: various aspects – mean endpoint score (various scales, high = good) – medium term	4		Mean Difference (IV, Random, 95% CI)	Subtotals only
10.1 Empowerment (Rogers Empowerment Scale (RES))	1	158	Mean Difference (IV, Random, 95% CI)	-0.95 [-3.30, 1.40]
10.2 Empowerment (Dutch Empowerment Scale (DES))	1	220	Mean Difference (IV, Random, 95% CI)	0.19 [0.05, 0.33]
10.3 Hope (State Hope Scale (SHS))	2	789	Mean Difference (IV, Random, 95% CI)	0.37 [-0.22, 0.96]
10.4 Hope (Herth Hope Index (HHI))	1	217	Mean Difference (IV, Random, 95% CI)	0.24 [0.11, 0.37]
11 Mental state: 1b. Specific: various aspects – mean endpoint score (various scales, high = good) – long term	4	1014	Mean Difference (IV, Random, 95% CI)	0.42 [-0.11, 0.95]
11.1 Hope (SHS)	3	908	Mean Difference (IV, Random, 95% CI)	0.41 [-0.15, 0.97]
11.2 Self-esteem (Rosenberg Scale (RS))	1	106	Mean Difference (IV, Random, 95% CI)	0.5 [-1.22, 2.22]
12 Mental state: 1c. Specific: various aspects – mean endpoint score (SHS subscales, high = good)	2		Mean Difference (IV, Random, 95% CI)	Subtotals only
12.1 Hope agency – medium term	2	796	Mean Difference (IV, Random, 95% CI)	0.28 [-0.06, 0.63]
12.2 Hope agency – long term	2	757	Mean Difference (IV, Random, 95% CI)	0.45 [0.07, 0.83]
12.3 Hope pathways – medium term	2	792	Mean Difference (IV, Random, 95% CI)	0.09 [-0.22, 0.40]
12.4 Hope pathways – long term	2	755	Mean Difference (IV, Random, 95% CI)	0.17 [-0.14, 0.48]
13 Mental state: 1d. Specific: various aspects – mean endpoint score (various subscales) (skewed data)			Other data	No numeric data
13.1 Aggressiveness (Colorado Client Assessment Record (CCAR), high = greater severity) – medium term			Other data	No numeric data
13.2 Anxiety (CCAR, high = greater severity) – medium term			Other data	No numeric data

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
13.3 Attention problem (CCAR, high = greater severity) – medium term			Other data	No numeric data
13.4 Behavioural and cognitive symptom (Instrument to Measure Self-Management (IMSM), high = greater frequency) – medium term			Other data	No numeric data
13.5 Cognitive problem (CCAR, high = greater severity) – medium term			Other data	No numeric data
13.6 Depression (Behaviour and Symptom Identification Scale (BASIS-24), high= greater severity) – medium term			Other data	No numeric data
13.7 Depression (CCAR, high = greater severity) – medium term			Other data	No numeric data
13.8 Emotional lability (BASIS-24, high = greater severity) – medium term			Other data	No numeric data
13.9 Emotional withdrawal (CCAR, high = greater severity) – medium term			Other data	No numeric data
13.10 Family problems (CCAR, high = greater severity) – medium term			Other data	No numeric data
13.11 Hyperaffect (CCAR, high = greater severity) – medium term			Other data	No numeric data
13.12 Interpersonal relationship (BASIS-24, high = greater severity) –medium term			Other data	No numeric data
13.13 Interpersonal problems (CCAR, high = greater severity) – medium term			Other data	No numeric data
13.14 Loneliness (Loneliness Scale, high = greater loneliness) – medium term			Other data	No numeric data
13.15 Physical activity (IMSM, high = greater frequency) – medium term			Other data	No numeric data
13.16 Psychotic symptoms (BASIS-24, high = greater severity) – medium term			Other data	No numeric data
13.17 Positive symptoms (BSI, high = greater severity) – medium term			Other data	No numeric data
13.18 Positive symptom (BSI, high = greater severity) – long term			Other data	No numeric data
13.19 Resistiveness (CCAR, high = greater severity) – medium term			Other data	No numeric data

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
13.20 Self-harm (BASIS-24, high = greater severity) -- medium term			Other data	No numeric data
13.21 Suicide feelings (CCAR, high = greater severity) – medium term			Other data	No numeric data
13.22 Thought process difficulties (CCAR, high = greater severity) – medium term			Other data	No numeric data
14 Behaviour : 1a. Specific: self-efficacy – mean endpoint score (various scales, high = good) – medium term	4		Mean Difference (IV, Random, 95% CI)	Subtotals only
14.1 Patient-Self-Advocacy (PSA)	1	458	Mean Difference (IV, Random, 95% CI)	0.08 [-0.02, 0.18]
14.2 Self-Management/Self-Efficacy Scale (SMSES)	1	57	Mean Difference (IV, Random, 95% CI)	1.20 [0.11, 2.29]
14.3 Mental Health Confidence Scale (MHCS)	1	221	Mean Difference (IV, Random, 95% CI)	0.31 [0.07, 0.55]
14.4 General Self-Efficacy Scale (GSE)	1	216	Mean Difference (IV, Random, 95% CI)	0.90 [-1.04, 2.84]
15 Behaviour: 1b. Specific: self-efficacy – mean endpoint score (various scales, high = good) – long term	3	769	Mean Difference (IV, Random, 95% CI)	1.10 [-0.71, 2.91]
15.1 PSA	1	447	Mean Difference (IV, Random, 95% CI)	0.10 [0.01, 0.19]
15.2 MHCS	1	106	Mean Difference (IV, Random, 95% CI)	2.70 [-2.40, 7.80]
15.3 GSE	1	216	Mean Difference (IV, Random, 95% CI)	2.20 [0.35, 4.05]
16 Behaviour: 2. Specific: self-management – mean endpoint score (SMS, high = good)	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
16.1 Medium term	1	57	Mean Difference (IV, Random, 95% CI)	0.60 [-0.10, 1.30]
17 Behaviour: 3. Specific: recovery – mean endpoint score (Recovery Assessment Scale (RAS), high = good)	3		Mean Difference (IV, Random, 95% CI)	Subtotals only
17.1 Medium term	3	557	Mean Difference (IV, Random, 95% CI)	2.69 [-0.82, 6.20]
17.2 Long term	1	318	Mean Difference (IV, Random, 95% CI)	4.16 [1.16, 7.16]
18 Behaviour: 4a. Specific: various behaviours – mean endpoint score (Patient Activation	4	810	Mean Difference (IV, Random, 95% CI)	1.58 [-0.33, 3.49]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
Scale (PAS) subscales, high = good) – medium term				
18.1 Activation (patient)	3	295	Mean Difference (IV, Random, 95% CI)	3.68 [-1.85, 9.22]
18.2 Approach to healthcare	1	57	Mean Difference (IV, Random, 95% CI)	2.10 [-0.83, 5.03]
18.3 Assertiveness	1	458	Mean Difference (IV, Random, 95% CI)	0.08 [-0.06, 0.22]
19 Behaviour: 4b. Specific: various behaviours – mean endpoint score (PAS subscales, high = good) – long term	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
19.1 Assertiveness	1	447	Mean Difference (IV, Fixed, 95% CI)	0.07 [-0.06, 0.20]
20 Behaviour: 4c. Specific: various behaviours – mean endpoint score (various subscales) – medium term	3		Mean Difference (IV, Random, 95% CI)	Subtotals only
20.1 Goal orientation (RAS, high = good)	1	343	Mean Difference (IV, Random, 95% CI)	0.72 [-0.09, 1.53]
20.2 Healthy eating (IMSM, high = good)	1	57	Mean Difference (IV, Random, 95% CI)	0.40 [-0.15, 0.95]
20.3 Internal locus of control for health (Multidimensional Health Locus of Control Scale (MHLC), high = greater control)	1	57	Mean Difference (IV, Random, 95% CI)	3.60 [0.99, 6.21]
20.4 Mindful non-adherence (PSA, high = non-adherence)	1	456	Mean Difference (IV, Random, 95% CI)	0.09 [-0.05, 0.23]
20.5 No symptom domination (RAS, high = good)	1	342	Mean Difference (IV, Random, 95% CI)	0.29 [-0.31, 0.89]
20.6 Personal confidence (RAS, high = good)	1	343	Mean Difference (IV, Random, 95% CI)	1.59 [0.30, 2.88]
20.7 Reliance on others (RAS, high = strong reliance)	1	343	Mean Difference (IV, Random, 95% CI)	0.80 [0.17, 1.43]
20.8 Self-management behaviours (SMS, high = good)	1	57	Mean Difference (IV, Random, 95% CI)	0.60 [-0.10, 1.30]
20.9 Willingness to ask for help (RAS subscale, high = strong willingness)	1	343	Mean Difference (IV, Random, 95% CI)	0.44 [0.01, 0.87]
21 Behaviour: 4d. Specific: various behaviours – mean endpoint score (various subscales) – long term	2		Mean Difference (IV, Random, 95% CI)	Subtotals only

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
21.1 Goal orientation (RAS, high = good)	1	320	Mean Difference (IV, Random, 95% CI)	0.61 [-0.19, 1.41]
21.2 Mindful non-adherence (PSA, high = non-adherence)	1	447	Mean Difference (IV, Random, 95% CI)	0.17 [0.03, 0.31]
21.3 No symptom domination (RAS, high = good)	1	319	Mean Difference (IV, Random, 95% CI)	0.77 [0.15, 1.39]
21.4 Personal confidence (RAS, high = good)	1	319	Mean Difference (IV, Random, 95% CI)	1.90 [0.61, 3.19]
21.5 Reliance on others (RAS, high = strong reliance)	1	320	Mean Difference (IV, Random, 95% CI)	0.41 [-0.21, 1.03]
21.6 Willingness to ask for help (RAS, high = strong willingness)	1	320	Mean Difference (IV, Random, 95% CI)	0.53 [0.06, 1.00]
22 Behaviour: 5. Specific: alcohol or drug use (various subscales) (skewed data)			Other data	No numeric data
22.1 Alcohol/drug use (BASIS-24, high = strong) – medium term			Other data	No numeric data
22.2 Alcohol use (Addiction Severity Index (ASI), high = strong) – medium term			Other data	No numeric data
22.3 Alcohol use (ASI, high = strong) – long term			Other data	No numeric data
22.4 Drug use (ASI, high = strong) – medium term			Other data	No numeric data
22.5 Drug use (ASI, high = strong) – long term			Other data	No numeric data
23 Leaving the study early – for any reason	8		Risk Ratio (M-H, Random, 95% CI)	Subtotals only
23.1 Medium term	6	741	Risk Ratio (M-H, Random, 95% CI)	0.66 [0.51, 0.87]
23.2 Long term	3	877	Risk Ratio (M-H, Random, 95% CI)	1.34 [0.19, 9.22]
24 Functioning: 1a. General: mean total endpoint score (various scales, high = good) – medium term	3		Mean Difference (IV, Random, 95% CI)	Subtotals only
24.1 CCAR	1	19	Mean Difference (IV, Random, 95% CI)	0.59 [-0.93, 2.11]
24.2 GAF	1	216	Mean Difference (IV, Random, 95% CI)	4.10 [0.34, 7.86]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
24.3 12-item Short Form (SF-12)	1	57	Mean Difference (IV, Random, 95% CI)	2.60 [-3.19, 8.39]
25 Functioning: 1b. General: mean total endpoint score (various scales, high = good) – long term	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
25.1 Global Assessment of Functioning (GAF)	1	216	Mean Difference (IV, Random, 95% CI)	-3.90 [-7.81, 0.01]
26 Functioning: 2a. Specific: various aspects – mean endpoint score (CCAR subscales, high = good) – medium term	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
26.1 Cognitive functioning	1	25	Mean Difference (IV, Random, 95% CI)	0.68 [-0.83, 2.19]
26.2 Interpersonal functioning	1	25	Mean Difference (IV, Random, 95% CI)	0.62 [-0.65, 1.89]
26.3 Physical functioning	1	19	Mean Difference (IV, Random, 95% CI)	0.38 [-1.05, 1.81]
26.4 Societal role functioning	1	25	Mean Difference (IV, Random, 95% CI)	1.02 [-0.44, 2.48]
27 Functioning: 2b. Specific: various aspects – mean endpoint score (SF-12 subscales, high = good) – medium term	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
27.1 Emotional well-being	1	57	Mean Difference (IV, Random, 95% CI)	3.00 [-2.76, 8.76]
27.2 Physical functioning	1	57	Mean Difference (IV, Random, 95% CI)	3.00 [-2.82, 8.82]
28 Functioning: 3. Specific: daily living – mean endpoint score (CCAR, high = good) – medium term (skewed data)			Other data	No numeric data
29 Functioning: 4. Specific: self-management – mean endpoint score (IMSM, high = good) (skewed data)			Other data	No numeric data
29.1 IMSM			Other data	No numeric data
30 Functioning: 5. Specific: contact with justice system – criminal justice charges (skewed data)			Other data	No numeric data
30.1 Felony (counts of criminal justice charges, high = more criminal charges) -- medium term			Other data	No numeric data

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
30.2 Felony (counts of criminal justice charges, high = more criminal charges) – long term			Other data	No numeric data
30.3 Infraction (counts of criminal justice charges, high = more criminal charges) – medium term			Other data	No numeric data
30.4 Infraction (counts of criminal justice charges, high = more criminal charges) – long term			Other data	No numeric data
30.5 Misdemeanour (counts of criminal justice charges, high = more criminal charges) – medium term			Other data	No numeric data
30.6 Misdemeanour (counts of criminal justice charges, high = more criminal charges) – long term			Other data	No numeric data
30.7 Total charges (counts of criminal justice charges, high = more criminal charges) – medium term			Other data	No numeric data
30.8 Total charges (counts of criminal justice charges, high = more criminal charges) – long term			Other data	No numeric data
30.9 Violation (counts of criminal justice charges, high = more criminal charges) -– medium term			Other data	No numeric data
30.10 Violation (counts of criminal justice charges, high = more criminal charges) – long term			Other data	No numeric data
31 Peer outcomes: 1a. Impact on the participant and peer supporter: improved peer contact – mean endpoint score (Personal Network Questionnaire (PNQ), high = good) – long term	1	106	Risk Ratio (M-H, Random, 95% CI)	1.85 [1.14, 3.00]
32 Peer outcomes: 1b. Impact on participant and peer supporter: negative aspects – mean endpoint score (BLR subscales, high = true) – medium term	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
32.1 Negative empathy	1	105	Mean Difference (IV, Random, 95% CI)	-0.32 [-0.66, 0.02]
32.2 Negative regard	1	105	Mean Difference (IV, Random, 95% CI)	-0.27 [-0.65, 0.11]
32.3 Negative overall relationship	1	105	Mean Difference (IV, Random, 95% CI)	-0.19 [-0.48, 0.10]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
32.4 Negative unconditionality	1	105	Mean Difference (IV, Random, 95% CI)	0.01 [-0.32, 0.34]
33 Peer outcomes: 1c. Impact on participant and peer supporter: positive aspects – mean endpoint score (Barrett-Lennard Relationship Inventory (BLRI) subscales, high = true) – medium term	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
33.1 Positive empathy	1	105	Mean Difference (IV, Random, 95% CI)	0.49 [0.13, 0.85]
33.2 Positive regard	1	105	Mean Difference (IV, Random, 95% CI)	0.44 [0.08, 0.80]
33.3 Positive overall relationship	1	105	Mean Difference (IV, Random, 95% CI)	0.43 [0.16, 0.70]
33.4 Positive unconditionality	1	105	Mean Difference (IV, Random, 95% CI)	0.33 [0.05, 0.61]
34 Peer outcomes: 1d. Impact on participant and peer supporter: various aspects – mean endpoint score (Social Support List (SSL) subscales, high = increased need for support) – long term	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
34.1 Negative interaction esteem support	1	106	Mean Difference (IV, Random, 95% CI)	-1.20 [-2.38, -0.02]
34.2 Social support for discrepancies	1	106	Mean Difference (IV, Random, 95% CI)	-1.5 [-7.58, 4.58]
34.3 Social support for positive interactions	1	106	Mean Difference (IV, Random, 95% CI)	5.60 [-0.51, 11.71]
35 Peer outcomes: 1e. Impact on participant and peer supporter: social support – mean endpoint score (Medical Outcomes Study Social Support Survey (MOSSSS), high = good)	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
35.1 Medium term	1	158	Mean Difference (IV, Random, 95% CI)	-1.12 [-6.26, 4.02]
36 Peer outcomes: 1f. Impact on participant and peer supporter: accessing social support (IMSM, high = greater amount of support obtained) – medium term (skewed data)			Other data	No numeric data
37 Peer outcomes: 2a. Quality of life for participant and peer supporter: overall – mean total endpoint (various scales, high = good) – medium term	5		Mean Difference (IV, Random, 95% CI)	Subtotals only
37.1 EuroQol: Five Dimensions (EQ5D)-Index	1	216	Mean Difference (IV, Random, 95% CI)	0.40 [-4.52, 5.32]

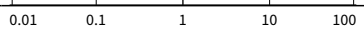
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
37.2 EuroQol: Five Dimensions-Visual Analogue Scale (EQ5D-VAS)	1	216	Mean Difference (IV, Random, 95% CI)	3.20 [-2.77, 9.17]
37.3 General Quality of Life Inventory (GQOLI-74)	1	100	Mean Difference (IV, Random, 95% CI)	40.34 [32.70, 47.98]
37.4 Manchester Short Assessment of Quality of Life (MSAQOL)	1	208	Mean Difference (IV, Random, 95% CI)	0.24 [-0.04, 0.52]
37.5 World Health Organisation Quality of Life (WHOQOL)	1	106	Mean Difference (IV, Random, 95% CI)	1.0 [-2.82, 4.82]
37.6 Quality of Life Brief Version (WHO-QOL-BREF)	1	458	Mean Difference (IV, Random, 95% CI)	0.20 [-0.33, 0.73]
38 Peer outcomes: 2b. Quality of life for participant and peer supporter: overall – mean total endpoint (various scales, high = good) – long term	3		Mean Difference (IV, Random, 95% CI)	Subtotals only
38.1 EQ5D-Index	1	216	Mean Difference (IV, Random, 95% CI)	3.30 [-1.83, 8.43]
38.2 EQ5D-VAS	1	216	Mean Difference (IV, Random, 95% CI)	5.0 [-0.67, 10.67]
38.3 WHOQOL-BREF	1	431	Mean Difference (IV, Random, 95% CI)	0.70 [0.15, 1.25]
38.4 WHOQOL	1	106	Mean Difference (IV, Random, 95% CI)	1.70 [-2.32, 5.72]
39 Peer outcomes: 3a. Quality of life for participant and peer supporter: specific aspects – mean endpoint score (GQOLI-74 subscales, high = good) – medium term	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
39.1 Mental health	1	100	Mean Difference (IV, Random, 95% CI)	16.95 [13.34, 20.56]
39.2 Physical quality of life	1	100	Mean Difference (IV, Random, 95% CI)	1.43 [-2.31, 5.17]
39.3 Physical health	1	100	Mean Difference (IV, Random, 95% CI)	15.08 [11.29, 18.87]
39.4 Social function	1	100	Mean Difference (IV, Random, 95% CI)	15.87 [12.66, 19.08]
40 Peer outcomes: 3b. Quality of life for participant and peer supporter: specific aspects – mean endpoint score (QOLI-BREF subscales, high = good) – medium term	1		Mean Difference (IV, Random, 95% CI)	Subtotals only

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
40.1 Amount of time spent with others	1	19	Mean Difference (IV, Random, 95% CI)	0.04 [-1.24, 1.32]
40.2 General life satisfaction	1	19	Mean Difference (IV, Random, 95% CI)	-0.04 [-1.25, 1.17]
40.3 Life in general	1	19	Mean Difference (IV, Random, 95% CI)	-0.49 [-1.73, 0.75]
40.4 Living arrangements	1	19	Mean Difference (IV, Random, 95% CI)	-0.32 [-1.58, 0.94]
40.5 Privacy	1	19	Mean Difference (IV, Random, 95% CI)	-0.58 [-1.40, 0.24]
40.6 Relax	1	19	Mean Difference (IV, Random, 95% CI)	-0.28 [-1.66, 1.10]
41 Peer outcomes: 3c. Quality of life for participant and peer supporter: specific aspects – mean endpoint score (36-item Short Form (SF-36) subscales, high = good) – medium term	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
41.1 Mental health	1	80	Mean Difference (IV, Random, 95% CI)	-0.20 [-5.00, 4.60]
41.2 Physical health	1	80	Mean Difference (IV, Random, 95% CI)	2.90 [-3.21, 9.01]
42 Peer outcomes: 3d. Quality of life for participant and peer supporter: specific aspects – mean endpoint score (QOL-BREF subscale, high = good) – medium term (skewed data)			Other data	No numeric data
43 Economic cost: 1. Direct and indirect costs (Euro): total cost (high = poor)	1		Mean Difference (Random, 95% CI)	Subtotals only
43.1 Medium term	1		Mean Difference (Random, 95% CI)	2092.0 [-72.00, 4258.00]
43.2 Long term	1		Mean Difference (Random, 95% CI)	775.00 [-1610.00, 3160.00]
44 Economic outcomes: 2. Direct costs (Euro): for minimally guided peer support (high = poor) – long term (skewed data)			Other data	No numeric data
45 Economic outcomes: 3a. Indirect cost of care (Euro): for inpatient and semi-inpatient care (high = poor) – long term (skewed data)			Other data	No numeric data
45.1 Hospital admission			Other data	No numeric data
45.2 Day care			Other data	No numeric data

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
45.3 Sheltered living			Other data	No numeric data
46 Economic outcomes: 3b. Indirect cost of care (Euro): for outpatient and community care (high = poor) – long term (skewed data)			Other data	No numeric data
46.1 Psychiatrist			Other data	No numeric data
46.2 Psychologist			Other data	No numeric data
46.3 Social psychiatric nurse			Other data	No numeric data
46.4 Social worker			Other data	No numeric data
46.5 Crisis intervention			Other data	No numeric data
46.6 Psychiatric home care			Other data	No numeric data
46.7 Consultation clinic for alcohol and drug addiction			Other data	No numeric data
46.8 Other outpatient care			Other data	No numeric data
47 Economic outcomes: 3c. Indirect cost of care (Euro): for general healthcare (high = poor) – long term (skewed data)			Other data	No numeric data
47.1 General practitioner			Other data	No numeric data
47.2 Alternative health care			Other data	No numeric data
47.3 Emergency care			Other data	No numeric data
47.4 Other general health care			Other data	No numeric data
48 Economic outcomes: 3d. Indirect costs (Euro): of day activity institutions (high = poor) – long term (skewed data)			Other data	No numeric data
48.1 Day activity centre			Other data	No numeric data
48.2 Drop-in centre			Other data	No numeric data
48.3 Recreation/activity centre			Other data	No numeric data
48.4 Other institutions			Other data	No numeric data
49 Economic outcomes: 3e. Indirect cost (Euro): of medication (high = poor) – long term (skewed data)			Other data	No numeric data
49.1 Prescribed			Other data	No numeric data
49.2 Non-prescribed			Other data	No numeric data

Analysis 1.1. Comparison 1 Peer support + standard care versus standard care alone, Outcome 1 Service use: 1a. Hospital admission – medium term.

Study or subgroup	Peer support n/N	Standard care n/N	Risk Ratio M-H, Random, 95% CI	Weight	Risk Ratio M-H, Random, 95% CI
Reynolds 2004	2/9	5/10		100%	0.44[0.11,1.75]
Total (95% CI)	**9**	**10**		**100%**	**0.44[0.11,1.75]**
Total events: 2 (Peer support), 5 (Standard care)					
Heterogeneity: Tau^2=0; Chi^2=0, df=0(P<0.0001); I^2=100%					
Test for overall effect: Z=1.16(P=0.25)					



Favours peer support — Favours standard care

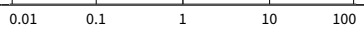
Analysis 1.2. Comparison 1 Peer support + standard care versus standard care alone, Outcome 2 Service use: 1b. Hospital admission – duration of hospital stay (days) – long term (skewed data).

Service use: 1b. Hospital admission – duration of hospital stay (days) – long term (skewed data)

Study	Interventions	Mean	SD	N
Mahlke 2017	Peer support	24.9	41.6	114
Mahlke 2017	Standard care	30.3	57.6	102

Analysis 1.3. Comparison 1 Peer support + standard care versus standard care alone, Outcome 3 Service use: 2a. Clinically important engagement with services – medium term.

Study or subgroup	Peer support n/N	Standard care n/N	Risk Ratio M-H, Random, 95% CI	Weight	Risk Ratio M-H, Random, 95% CI
1.3.1 Use of emergency care					
Goldberg 2013	3/28	8/29		100%	0.39[0.11,1.32]
Subtotal (95% CI)	**28**	**29**		**100%**	**0.39[0.11,1.32]**
Total events: 3 (Peer support), 8 (Standard care)					
Heterogeneity: Not applicable					
Test for overall effect: Z=1.52(P=0.13)					
1.3.2 ≥ 1 primary care visit					
Druss 2010	26/41	14/39		100%	1.77[1.09,2.85]
Subtotal (95% CI)	**41**	**39**		**100%**	**1.77[1.09,2.85]**
Total events: 26 (Peer support), 14 (Standard care)					
Heterogeneity: Not applicable					
Test for overall effect: Z=2.33(P=0.02)					
Test for subgroup differences: Chi^2=5.12, df=1 (P=0.02), I^2=80.48%					



Favours standard care — Favours peer support

Analysis 1.4. Comparison 1 Peer support + standard care versus standard care alone, Outcome 4 Service use: 2b. Contact with services – medium term (skewed data).

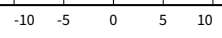
Service use: 2b. Contact with services – medium term (skewed data)

Study	Intervention	Mean	SD	N
Mean number of emergency visits				

Service use: 2b. Contact with services – medium term (skewed data)				
Study	**Intervention**	**Mean**	**SD**	**N**
Kelly 2014	Peer support	1.42	1.78	12
Kelly 2014	Standard care	2.00	1.50	11
Mean number of routine care visits				
Kelly 2014	Peer support	2.5	1.45	12
Kelly 2014	Standard care	2.11	1.45	11

Analysis 1.5. Comparison 1 Peer support + standard care versus standard care alone, Outcome 5 Global state: 3a. General Health – mean total endpoint score (Veterans RAND 12-Item Health Survey (VR-12), high = good).

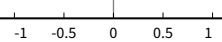
Study or subgroup	Peer support N	Peer support Mean(SD)	Standard care N	Standard care Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
1.5.1 Medium term							
Eisen 2012	74	47.2 (12.4)	84	47.2 (12.8)		100%	-0.02[-3.96,3.92]
Subtotal ***	**74**		**84**			**100%**	**-0.02[-3.96,3.92]**
Heterogeneity: Not applicable							
Test for overall effect: Z=0.01(P=0.99)							
Total ***	**74**		**84**			**100%**	**-0.02[-3.96,3.92]**
Heterogeneity: Not applicable							
Test for overall effect: Z=0.01(P=0.99)							



Favours standard care | Favours peer support

Analysis 1.6. Comparison 1 Peer support + standard care versus standard care alone, Outcome 6 Global state: 3b. Severity of illness – mean total endpoint score (Brief Symptom Inventory (BSI), high = poor).

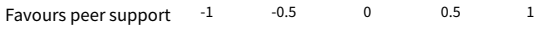
Study or subgroup	Peer support N	Peer support Mean(SD)	Standard care N	Standard care Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
1.6.1 Medium term							
Cook 2012b	224	0.7 (0.6)	234	0.9 (0.7)		100%	-0.13[-0.25,-0.01]
Subtotal ***	**224**		**234**			**100%**	**-0.13[-0.25,-0.01]**
Heterogeneity: Not applicable							
Test for overall effect: Z=2.08(P=0.04)							
1.6.2 Long term							
Cook 2012b	220	0.4 (0.6)	220	0.4 (0.6)		100%	0[-0.11,0.11]
Subtotal ***	**220**		**220**			**100%**	**0[-0.11,0.11]**
Heterogeneity: Not applicable							
Test for overall effect: Not applicable							
Test for subgroup differences: Chi^2=2.31, df=1 (P=0.13), I^2=56.77%							



Favours peer support | Favours standard care

Analysis 1.7. Comparison 1 Peer support + standard care versus standard care alone, Outcome 7 Global state: 3c. Severity of illness – mean total endpoint score (Clinical Global Impression scale (CGI), high = poor).

Study or subgroup	Peer support N	Peer support Mean(SD)	Standard care N	Standard care Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
1.7.1 Medium term							
Mahlke 2017	114	4.5 (0.9)	102	4.8 (0.9)		100%	-0.3[-0.53,-0.07]
Subtotal ***	**114**		**102**			**100%**	**-0.3[-0.53,-0.07]**
Heterogeneity: Not applicable							
Test for overall effect: Z=2.52(P=0.01)							
1.7.2 Long term							
Mahlke 2017	114	4.6 (0.9)	102	4.2 (1)		100%	0.4[0.15,0.65]
Subtotal ***	**114**		**102**			**100%**	**0.4[0.15,0.65]**
Heterogeneity: Not applicable							
Test for overall effect: Z=3.12(P=0)							
Test for subgroup differences: Chi^2=15.99, df=1 (P<0.0001), I^2=93.75%							



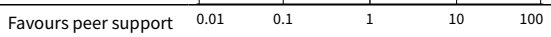
Analysis 1.8. Comparison 1 Peer support + standard care versus standard care alone, Outcome 8 Global state: 4. Compliance with medication (skewed data).

Global state: 4. Compliance with medication (skewed data)

Study	Intervention	MD	SD	N
		Number of medication		
Kelly 2014	peer support	2.83	1.80	12
Kelly 2014	standard care	3.5	2.68	11

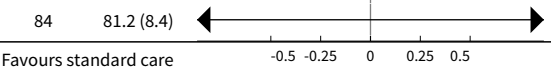
Analysis 1.9. Comparison 1 Peer support + standard care versus standard care alone, Outcome 9 Adverse event: 1. Death – all cause (long term).

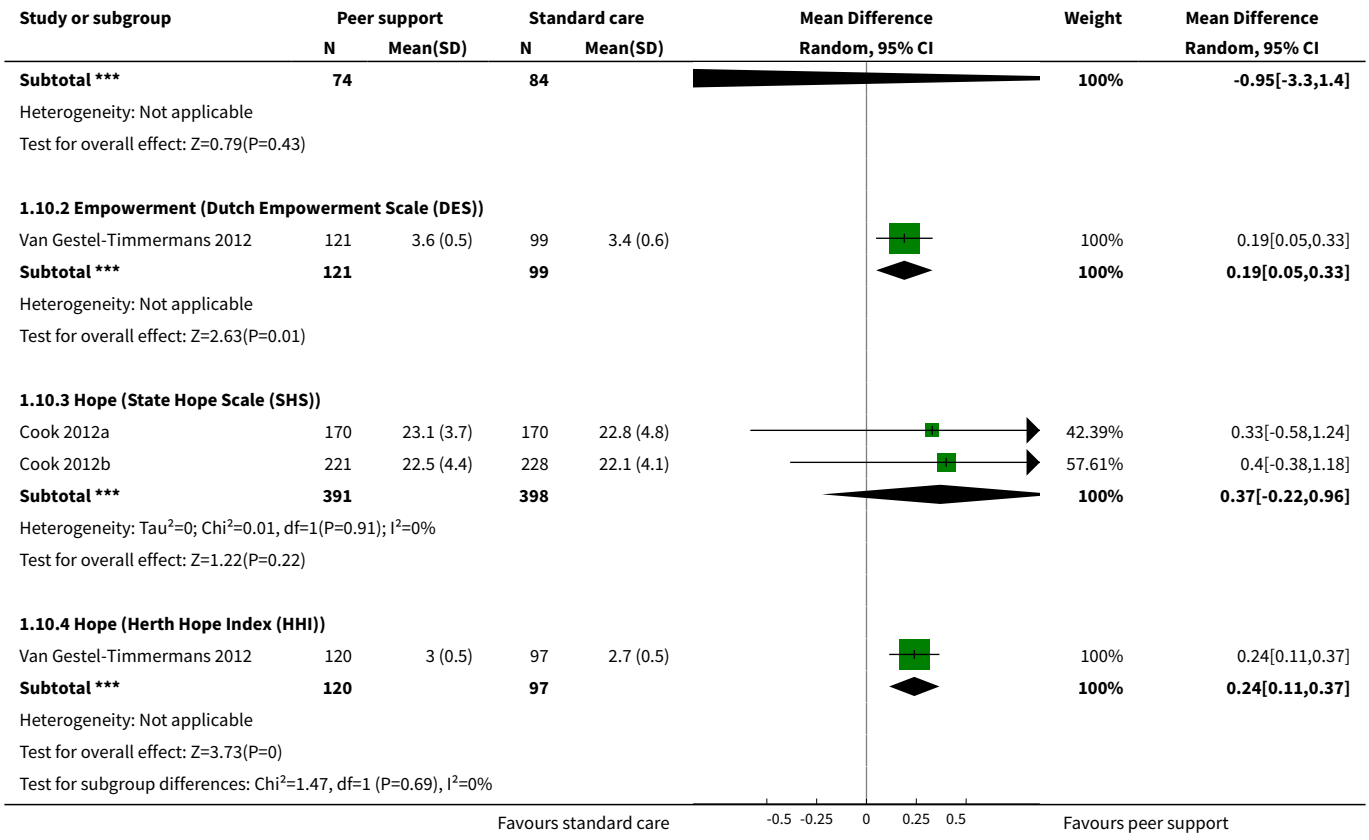
Study or subgroup	Peer support n/N	Standard care n/N	Risk Ratio M-H, Random, 95% CI	Weight	Risk Ratio M-H, Random, 95% CI
Cook 2012b	6/276	4/279		100%	1.52[0.43,5.31]
Total (95% CI)	**276**	**279**		**100%**	**1.52[0.43,5.31]**
Total events: 6 (Peer support), 4 (Standard care)					
Heterogeneity: Not applicable					
Test for overall effect: Z=0.65(P=0.52)					



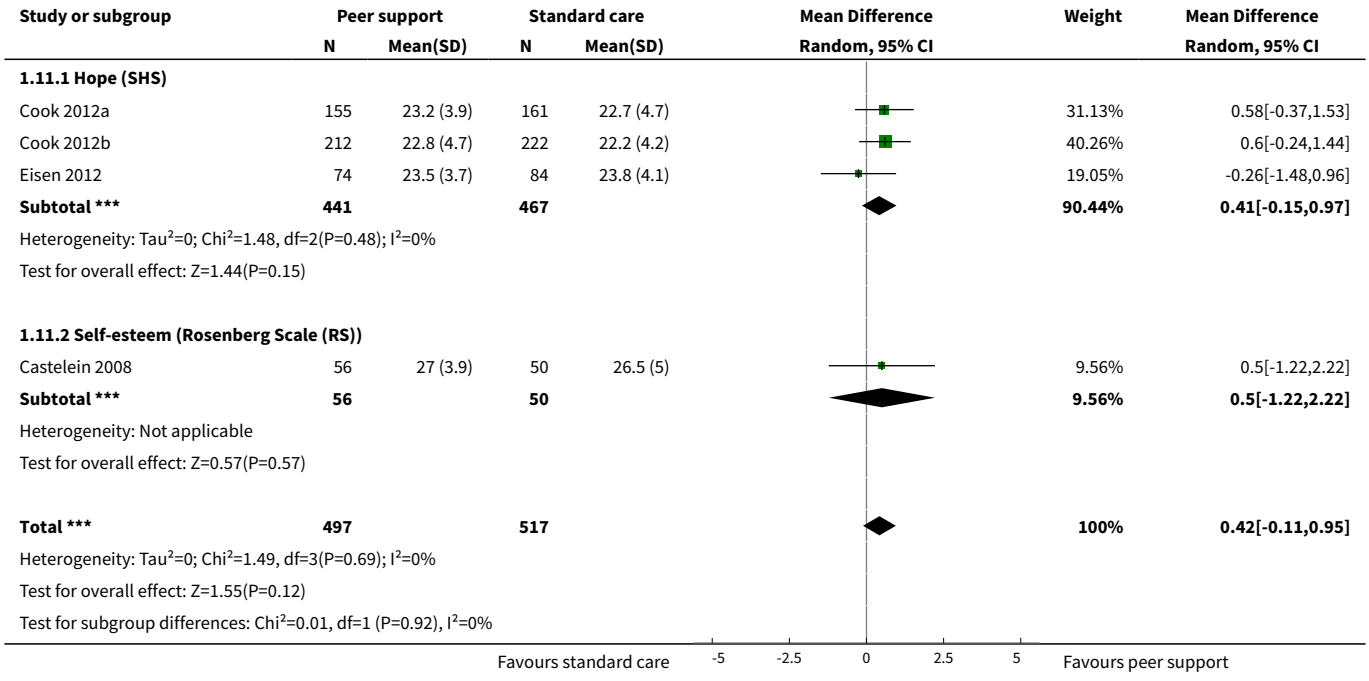
Analysis 1.10. Comparison 1 Peer support + standard care versus standard care alone, Outcome 10 Mental state: 1a. Specific: various aspects – mean endpoint score (various scales, high = good) – medium term.

Study or subgroup	Peer support N	Peer support Mean(SD)	Standard care N	Standard care Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
1.10.1 Empowerment (Rogers Empowerment Scale (RES))							
Eisen 2012	74	80.2 (6.7)	84	81.2 (8.4)		100%	-0.95[-3.3,1.4]



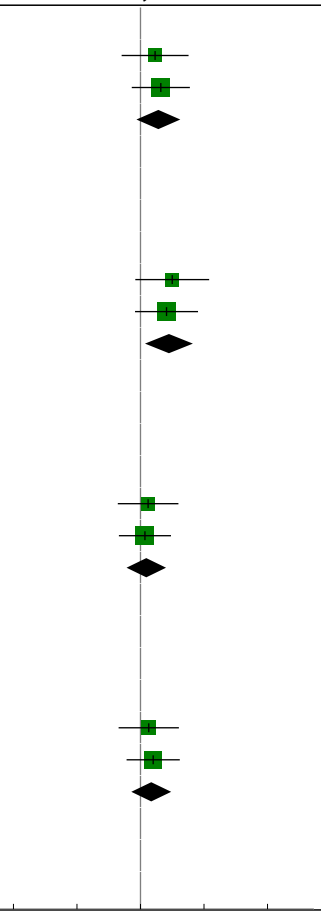


Analysis 1.11. Comparison 1 Peer support + standard care versus standard care alone, Outcome 11 Mental state: 1b. Specific: various aspects – mean endpoint score (various scales, high = good) – long term.



Analysis 1.12. Comparison 1 Peer support + standard care versus standard care alone, Outcome 12 Mental state: 1c. Specific: various aspects – mean endpoint score (SHS subscales, high = good).

Study or subgroup	Peer support N	Peer support Mean(SD)	Standard care N	Standard care Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
1.12.1 Hope agency – medium term							
Cook 2012a	170	11.5 (2.2)	172	11.3 (2.8)		42.97%	0.23[-0.3,0.76]
Cook 2012b	223	11.2 (2.5)	231	10.9 (2.5)		57.03%	0.32[-0.14,0.78]
Subtotal ***	**393**		**403**			**100%**	**0.28[-0.06,0.63]**
Heterogeneity: Tau²=0; Chi²=0.06, df=1(P=0.8); I²=0%							
Test for overall effect: Z=1.6(P=0.11)							
1.12.2 Hope agency – long term							
Cook 2012a	157	11.7 (2.5)	162	11.2 (2.8)		42.12%	0.5[-0.08,1.08]
Cook 2012b	215	11.3 (2.7)	223	10.9 (2.6)		57.88%	0.41[-0.09,0.91]
Subtotal ***	**372**		**385**			**100%**	**0.45[0.07,0.83]**
Heterogeneity: Tau²=0; Chi²=0.05, df=1(P=0.82); I²=0%							
Test for overall effect: Z=2.33(P=0.02)							
1.12.3 Hope pathways – medium term							
Cook 2012a	171	11.6 (2)	170	11.5 (2.5)		42.46%	0.12[-0.36,0.6]
Cook 2012b	222	11.3 (2.3)	229	11.2 (2.1)		57.54%	0.07[-0.34,0.48]
Subtotal ***	**393**		**399**			**100%**	**0.09[-0.22,0.4]**
Heterogeneity: Tau²=0; Chi²=0.02, df=1(P=0.88); I²=0%							
Test for overall effect: Z=0.58(P=0.57)							
1.12.4 Hope pathways – long term							
Cook 2012a	155	11.6 (2)	162	11.4 (2.3)		43.81%	0.13[-0.34,0.6]
Cook 2012b	213	11.4 (2.4)	225	11.2 (2.1)		56.19%	0.2[-0.22,0.62]
Subtotal ***	**368**		**387**			**100%**	**0.17[-0.14,0.48]**
Heterogeneity: Tau²=0; Chi²=0.05, df=1(P=0.83); I²=0%							
Test for overall effect: Z=1.06(P=0.29)							
Test for subgroup differences: Chi²=2.27, df=1 (P=0.52), I²=0%							



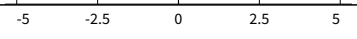
Analysis 1.13. Comparison 1 Peer support + standard care versus standard care alone, Outcome 13 Mental state: 1d. Specific: various aspects – mean endpoint score (various subscales) (skewed data).

Mental state: 1d. Specific: various aspects – mean endpoint score (various subscales) (skewed data)				
Study	**Intervention**	**Mean**	**SD**	**N**
Aggressiveness (Colorado Client Assessment Record (CCAR), high = greater severity) – medium term				
Reynolds 2004	Peer support	1.20	0.63	8
Reynolds 2004	Standard care	1.10	0.32	11
Anxiety (CCAR, high = greater severity) – medium term				
Reynolds 2004	Peer support	3.33	1.73	8
Reynolds 2004	Standard care	3.00	1.88	11
Attention problem (CCAR, high = greater severity) – medium term				
Reynolds 2004	Peer support	1.89	1.62	8
Reynolds 2004	Standard care	2.50	1.84	11
Behavioural and cognitive symptom (Instrument to Measure Self-Management (IMSM), high = greater frequency) – medium term				
Goldberg 2013	Peer support	1.9	1.0	28
Goldberg 2013	Standard care	1.8	1.2	29

Mental state: 1d. Specific: various aspects – mean endpoint score (various subscales) (skewed data)				
Study	**Intervention**	**Mean**	**SD**	**N**
Cognitive problem (CCAR, high = greater severity) – medium term				
Reynolds 2004	Peer support	1.89	1.45	8
Reynolds 2004	Standard care	1.80	1.32	11
Depression (Behaviour and Symptom Identification Scale (BASIS-24), high= greater severity) – medium term				
Eisen 2012	peer support	1.3	0.9	74
Eisen 2012	standard care	1.21	0.87	84
Depression (CCAR, high = greater severity) – medium term				
Reynolds 2004	peer support	2.44	1.42	8
Reynolds 2004	standard care	3.4	1.43	11
Emotional lability (BASIS-24, high = greater severity) – medium term				
Eisen 2012	Peer support	1.32	1.06	74
Eisen 2012	Standard care	1.49	0.95	84
Emotional withdrawal (CCAR, high = greater severity) – medium term				
Reynolds 2004	Peer support	2.56	1.94	8
Reynolds 2004	Standard care	2.20	1.75	11
Family problems (CCAR, high = greater severity) – medium term				
Reynolds 2004	Peer support	1.67	1.41	8
Reynolds 2004	Standard care	1.40	0.97	11
Hyperaffect (CCAR, high = greater severity) – medium term				
Reynolds 2004	Peer support	1.56	1.33	8
Reynolds 2004	Standard care	1.40	0.70	11
Interpersonal relationship (BASIS-24, high = greater severity) –medium term				
Eisen 2012	peer support	1.28	0.76	74
Eisen 2012	standard care	1.26	0.85	84
Interpersonal problems (CCAR, high = greater severity) – medium term				
Reynolds 2004	peer support	2.78	1.3	8
Reynolds 2004	standard care	2.5	1.43	11
Loneliness (Loneliness Scale, high = greater loneliness) – medium term				
Van Gestel-Timmermans 2012	Peer support	5.45	3.87	125
Van Gestel-Timmermans 2012	Standard care	6.49	3.68	102
Physical activity (IMSM, high = greater frequency) – medium term				
Goldberg 2013	Peer support	3.2	1.2	28
Goldberg 2013	Standard care	2.2	1.4	29
Psychotic symptoms (BASIS-24, high = greater severity) – medium term				
Eisen 2012	peer support	0.58	0.87	74
Eisen 2012	standard care	0.66	0.87	84
Positive symptoms (BSI, high = greater severity) – medium term				
Cook 2012b	peer support	19.52	13.74	224
Cook 2012b	standard care	21.38	13.68	234
Positive symptom (BSI, high = greater severity) – long term				
Cook 2012b	peer support	12.2	outlier	220
Cook 2012b	standard care	12.65	15	228
Resistiveness (CCAR, high = greater severity) – medium term				
Reynolds 2004	Peer support	1.67	1.41	8
Reynolds 2004	Standard care	1.40	0.84	11
Self-harm (BASIS-24, high = greater severity) -– medium term				
Eisen 2012	Peer support	0.18	0.50	74
Eisen 2012	Standard care	0.18	0.46	84
Suicide feelings (CCAR, high = greater severity) – medium term				
Reynolds 2004	Peer support	2.22	2.44	8
Reynolds 2004	Standard care	1.70	1.06	11
Thought process difficulties (CCAR, high = greater severity) – medium term				
Reynolds 2004	peer support	2.56	2	8
Reynolds 2004	standard care	2.1	1.91	11

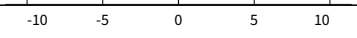
Analysis 1.14. Comparison 1 Peer support + standard care versus standard care alone, Outcome 14 Behaviour : 1a. Specific: self-efficacy – mean endpoint score (various scales, high = good) – medium term.

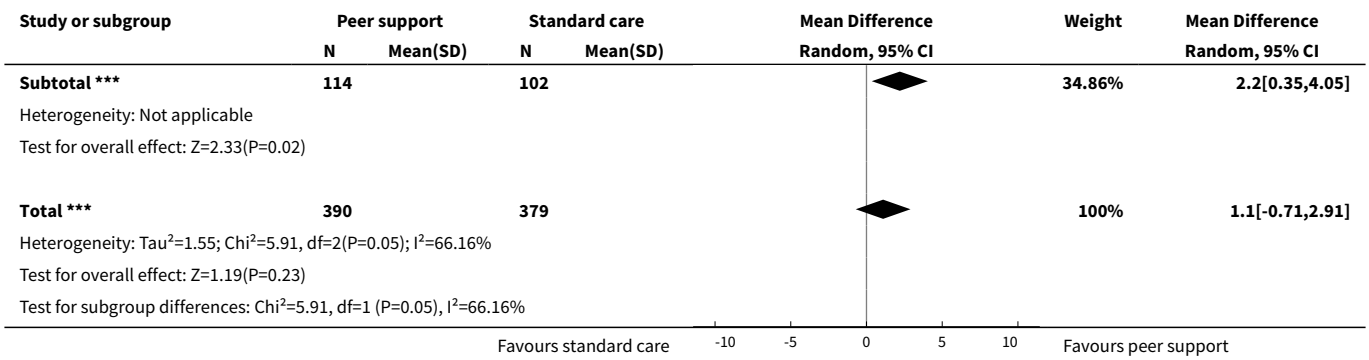
Study or subgroup	Peer support N	Peer support Mean(SD)	Standard care N	Standard care Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
1.14.1 Patient-Self-Advocacy (PSA)							
Cook 2012b	224	3.6 (0.5)	234	3.5 (0.5)		100%	0.08[-0.02,0.18]
Subtotal ***	**224**		**234**			**100%**	**0.08[-0.02,0.18]**
Heterogeneity: Not applicable							
Test for overall effect: Z=1.63(P=0.1)							
1.14.2 Self-Management/Self-Efficacy Scale (SMSES)							
Goldberg 2013	28	7.1 (2.1)	29	5.9 (2.1)		100%	1.2[0.11,2.29]
Subtotal ***	**28**		**29**			**100%**	**1.2[0.11,2.29]**
Heterogeneity: Not applicable							
Test for overall effect: Z=2.16(P=0.03)							
1.14.3 Mental Health Confidence Scale (MHCS)							
Van Gestel-Timmermans 2012	121	4.7 (0.9)	100	4.4 (0.9)		100%	0.31[0.07,0.55]
Subtotal ***	**121**		**100**			**100%**	**0.31[0.07,0.55]**
Heterogeneity: Not applicable							
Test for overall effect: Z=2.54(P=0.01)							
1.14.4 General Self-Efficacy Scale (GSE)							
Mahlke 2017	114	25.2 (6.5)	102	24.3 (7.9)		100%	0.9[-1.04,2.84]
Subtotal ***	**114**		**102**			**100%**	**0.9[-1.04,2.84]**
Heterogeneity: Not applicable							
Test for overall effect: Z=0.91(P=0.36)							
Test for subgroup differences: Chi^2=7.48, df=1 (P=0.06), I^2=59.88%							



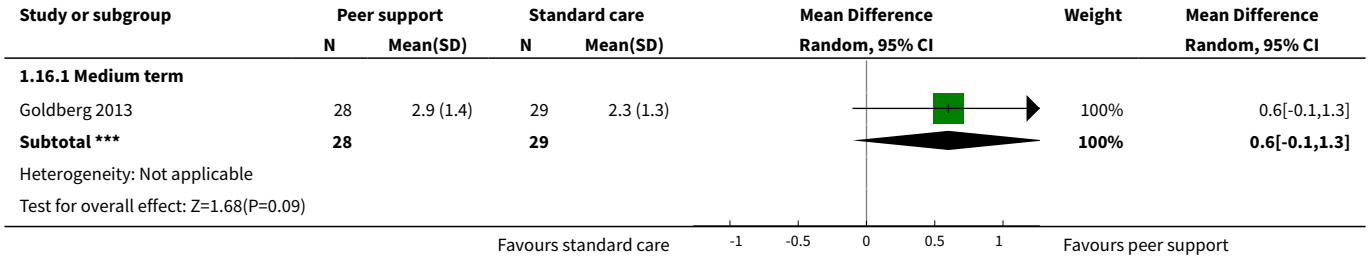
Analysis 1.15. Comparison 1 Peer support + standard care versus standard care alone, Outcome 15 Behaviour: 1b. Specific: self-efficacy – mean endpoint score (various scales, high = good) – long term.

Study or subgroup	Peer support N	Peer support Mean(SD)	Standard care N	Standard care Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
1.15.1 PSA							
Cook 2012b	220	3.7 (0.5)	227	3.6 (0.5)		54.91%	0.1[0.01,0.19]
Subtotal ***	**220**		**227**			**54.91%**	**0.1[0.01,0.19]**
Heterogeneity: Not applicable							
Test for overall effect: Z=2.09(P=0.04)							
1.15.2 MHCS							
Castelein 2008	56	67.5 (12)	50	64.8 (14.5)		10.23%	2.7[-2.4,7.8]
Subtotal ***	**56**		**50**			**10.23%**	**2.7[-2.4,7.8]**
Heterogeneity: Not applicable							
Test for overall effect: Z=1.04(P=0.3)							
1.15.3 GSE							
Mahlke 2017	114	26 (6.7)	102	23.8 (7.1)		34.86%	2.2[0.35,4.05]

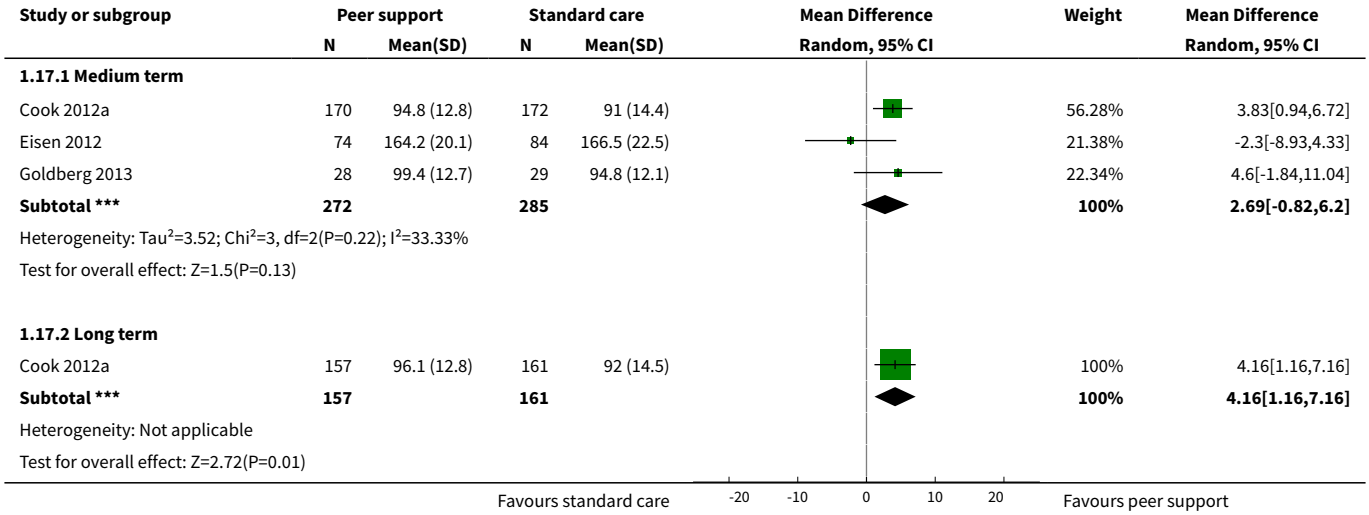




Analysis 1.16. Comparison 1 Peer support + standard care versus standard care alone, Outcome 16 Behaviour: 2. Specific: self-management – mean endpoint score (SMS, high = good).

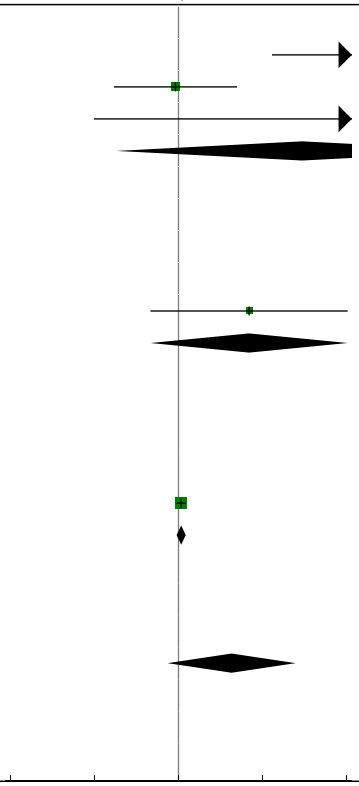


Analysis 1.17. Comparison 1 Peer support + standard care versus standard care alone, Outcome 17 Behaviour: 3. Specific: recovery – mean endpoint score (Recovery Assessment Scale (RAS), high = good).



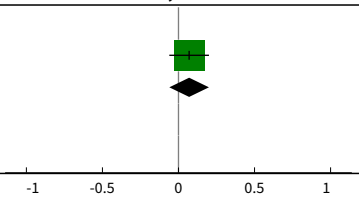
Analysis 1.18. Comparison 1 Peer support + standard care versus standard care alone, Outcome 18 Behaviour: 4a. Specific: various behaviours – mean endpoint score (Patient Activation Scale (PAS) subscales, high = good) – medium term.

Study or subgroup	Peer support N	Peer support Mean(SD)	Standard care N	Standard care Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
1.18.1 Activation (patient)							
Druss 2010	41	52 (10.1)	39	44.9 (9.6)		12.65%	7.1[2.78,11.42]
Eisen 2012	74	29.1 (5.9)	84	29.2 (5.8)		27.02%	-0.09[-1.92,1.74]
Goldberg 2013	28	65.5 (16.2)	29	60.1 (14.2)		4.99%	5.4[-2.52,13.32]
Subtotal ***	**143**		**152**			**44.66%**	**3.68[-1.85,9.22]**
Heterogeneity: Tau^2=18.08; Chi^2=10.16, df=2(P=0.01); I^2=80.32%							
Test for overall effect: Z=1.3(P=0.19)							
1.18.2 Approach to healthcare							
Goldberg 2013	28	42 (5.5)	29	39.9 (5.8)		19.43%	2.1[-0.83,5.03]
Subtotal ***	**28**		**29**			**19.43%**	**2.1[-0.83,5.03]**
Heterogeneity: Not applicable							
Test for overall effect: Z=1.4(P=0.16)							
1.18.3 Assertiveness							
Cook 2012b	224	3.8 (0.8)	234	3.7 (0.7)		35.92%	0.08[-0.06,0.22]
Subtotal ***	**224**		**234**			**35.92%**	**0.08[-0.06,0.22]**
Heterogeneity: Tau^2=0; Chi^2=0, df=0(P<0.0001); I^2=100%							
Test for overall effect: Z=1.15(P=0.25)							
Total ***	**395**		**415**			**100%**	**1.58[-0.33,3.49]**
Heterogeneity: Tau^2=2.63; Chi^2=13.72, df=4(P=0.01); I^2=70.84%							
Test for overall effect: Z=1.62(P=0.1)							
Test for subgroup differences: Chi^2=3.44, df=1 (P=0.18), I^2=41.85%							



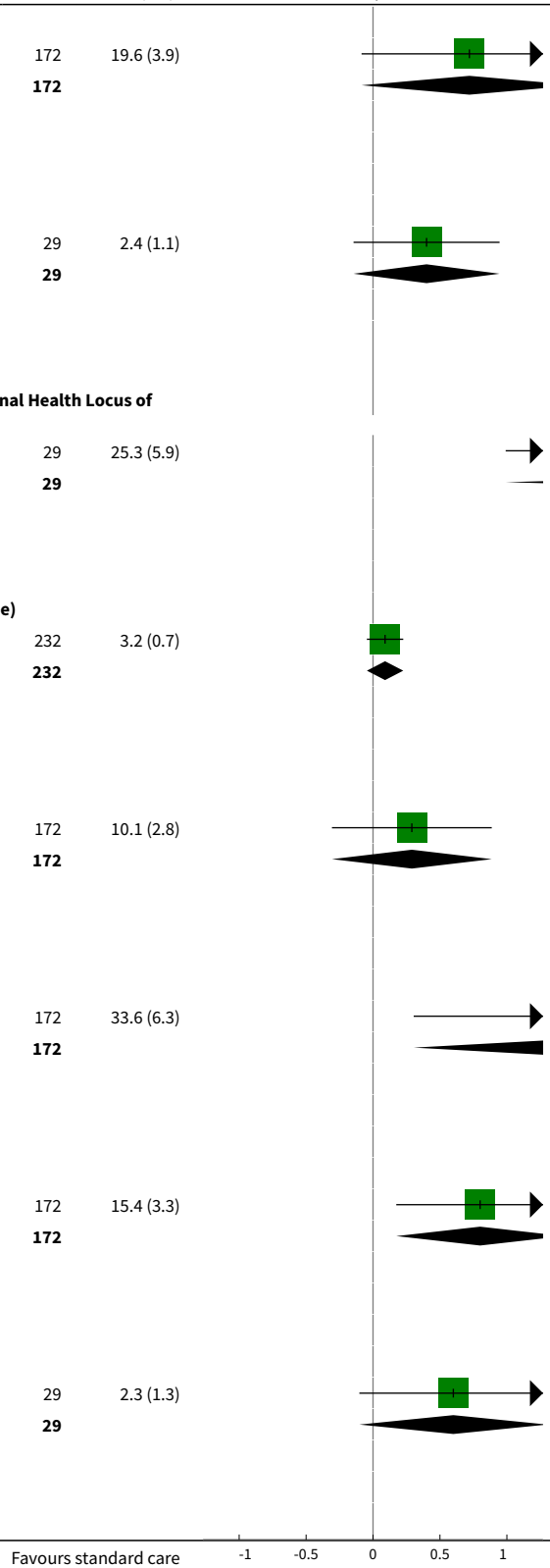
Analysis 1.19. Comparison 1 Peer support + standard care versus standard care alone, Outcome 19 Behaviour: 4b. Specific: various behaviours – mean endpoint score (PAS subscales, high = good) – long term.

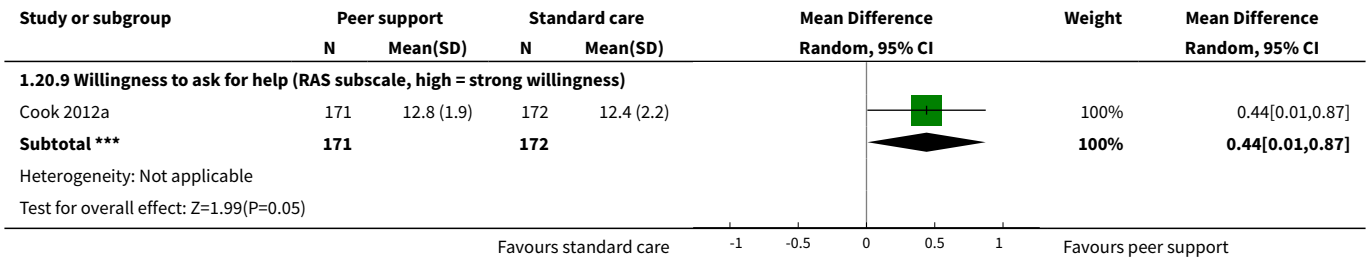
Study or subgroup	Peer support N	Peer support Mean(SD)	Standard care N	Standard care Mean(SD)	Mean Difference Fixed, 95% CI	Weight	Mean Difference Fixed, 95% CI
1.19.1 Assertiveness							
Cook 2012b	220	3.8 (0.8)	227	3.8 (0.7)		100%	0.07[-0.06,0.2]
Subtotal ***	**220**		**227**			**100%**	**0.07[-0.06,0.2]**
Heterogeneity: Not applicable							
Test for overall effect: Z=1.05(P=0.29)							



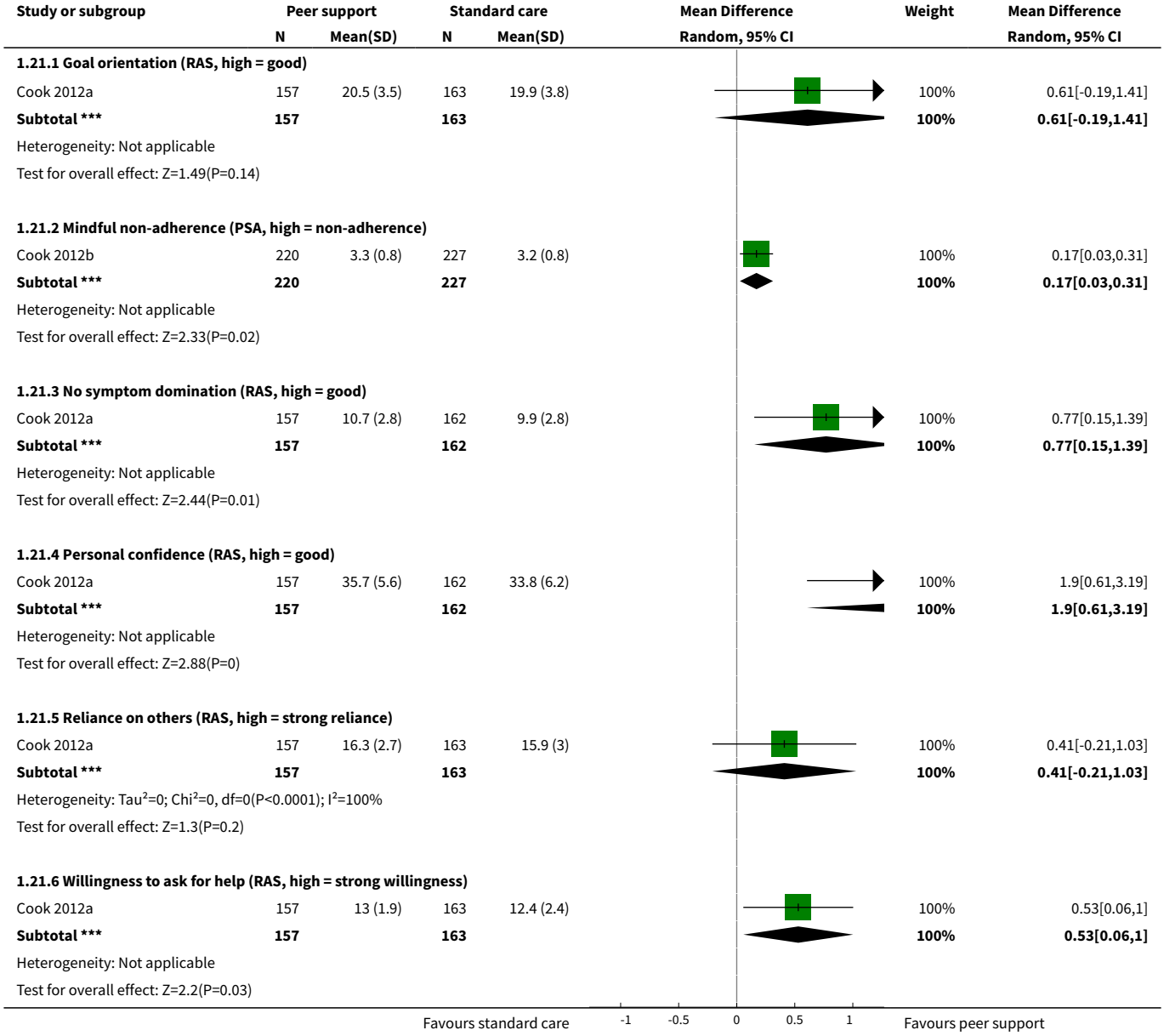
Analysis 1.20. Comparison 1 Peer support + standard care versus standard care alone, Outcome 20 Behaviour: 4c. Specific: various behaviours – mean endpoint score (various subscales) – medium term.

Study or subgroup	Peer support N	Peer support Mean(SD)	Standard care N	Standard care Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
1.20.1 Goal orientation (RAS, high = good)							
Cook 2012a	171	20.4 (3.7)	172	19.6 (3.9)		100%	0.72[-0.09,1.53]
Subtotal ***	**171**		**172**			**100%**	**0.72[-0.09,1.53]**
Heterogeneity: Not applicable							
Test for overall effect: Z=1.75(P=0.08)							
1.20.2 Healthy eating (IMSM, high = good)							
Goldberg 2013	28	2.8 (1)	29	2.4 (1.1)		100%	0.4[-0.15,0.95]
Subtotal ***	**28**		**29**			**100%**	**0.4[-0.15,0.95]**
Heterogeneity: Not applicable							
Test for overall effect: Z=1.44(P=0.15)							
1.20.3 Internal locus of control for health (Multidimensional Health Locus of Control Scale (MHLC), high = greater control)							
Goldberg 2013	28	28.9 (4)	29	25.3 (5.9)		100%	3.6[0.99,6.21]
Subtotal ***	**28**		**29**			**100%**	**3.6[0.99,6.21]**
Heterogeneity: Tau^2=0; Chi^2=0, df=0(P<0.0001); I^2=100%							
Test for overall effect: Z=2.7(P=0.01)							
1.20.4 Mindful non-adherence (PSA, high = non-adherence)							
Cook 2012b	224	3.3 (0.7)	232	3.2 (0.7)		100%	0.09[-0.05,0.23]
Subtotal ***	**224**		**232**			**100%**	**0.09[-0.05,0.23]**
Heterogeneity: Not applicable							
Test for overall effect: Z=1.3(P=0.19)							
1.20.5 No symptom domination (RAS, high = good)							
Cook 2012a	170	10.4 (2.9)	172	10.1 (2.8)		100%	0.29[-0.31,0.89]
Subtotal ***	**170**		**172**			**100%**	**0.29[-0.31,0.89]**
Heterogeneity: Not applicable							
Test for overall effect: Z=0.95(P=0.34)							
1.20.6 Personal confidence (RAS, high = good)							
Cook 2012a	171	35.2 (5.9)	172	33.6 (6.3)		100%	1.59[0.3,2.88]
Subtotal ***	**171**		**172**			**100%**	**1.59[0.3,2.88]**
Heterogeneity: Not applicable							
Test for overall effect: Z=2.42(P=0.02)							
1.20.7 Reliance on others (RAS, high = strong reliance)							
Cook 2012a	171	16.2 (2.6)	172	15.4 (3.3)		100%	0.8[0.17,1.43]
Subtotal ***	**171**		**172**			**100%**	**0.8[0.17,1.43]**
Heterogeneity: Not applicable							
Test for overall effect: Z=2.5(P=0.01)							
1.20.8 Self-management behaviours (SMS, high = good)							
Goldberg 2013	28	2.9 (1.4)	29	2.3 (1.3)		100%	0.6[-0.1,1.3]
Subtotal ***	**28**		**29**			**100%**	**0.6[-0.1,1.3]**
Heterogeneity: Not applicable							
Test for overall effect: Z=1.68(P=0.09)							





Analysis 1.21. Comparison 1 Peer support + standard care versus standard care alone, Outcome 21 Behaviour: 4d. Specific: various behaviours – mean endpoint score (various subscales) – long term.



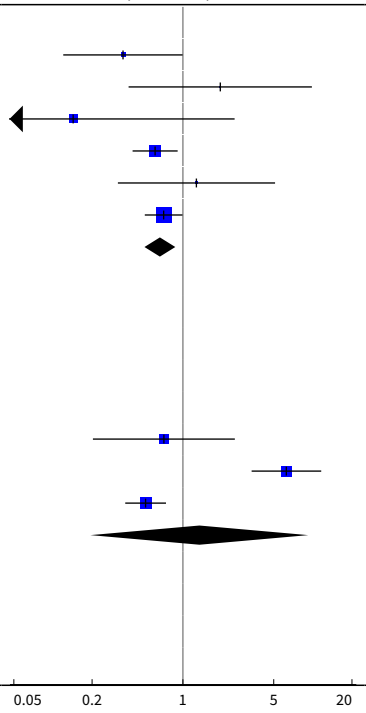
Analysis 1.22. Comparison 1 Peer support + standard care versus standard care alone, Outcome 22 Behaviour: 5. Specific: alcohol or drug use (various subscales) (skewed data).

Behaviour: 5. Specific: alcohol or drug use (various subscales) (skewed data)

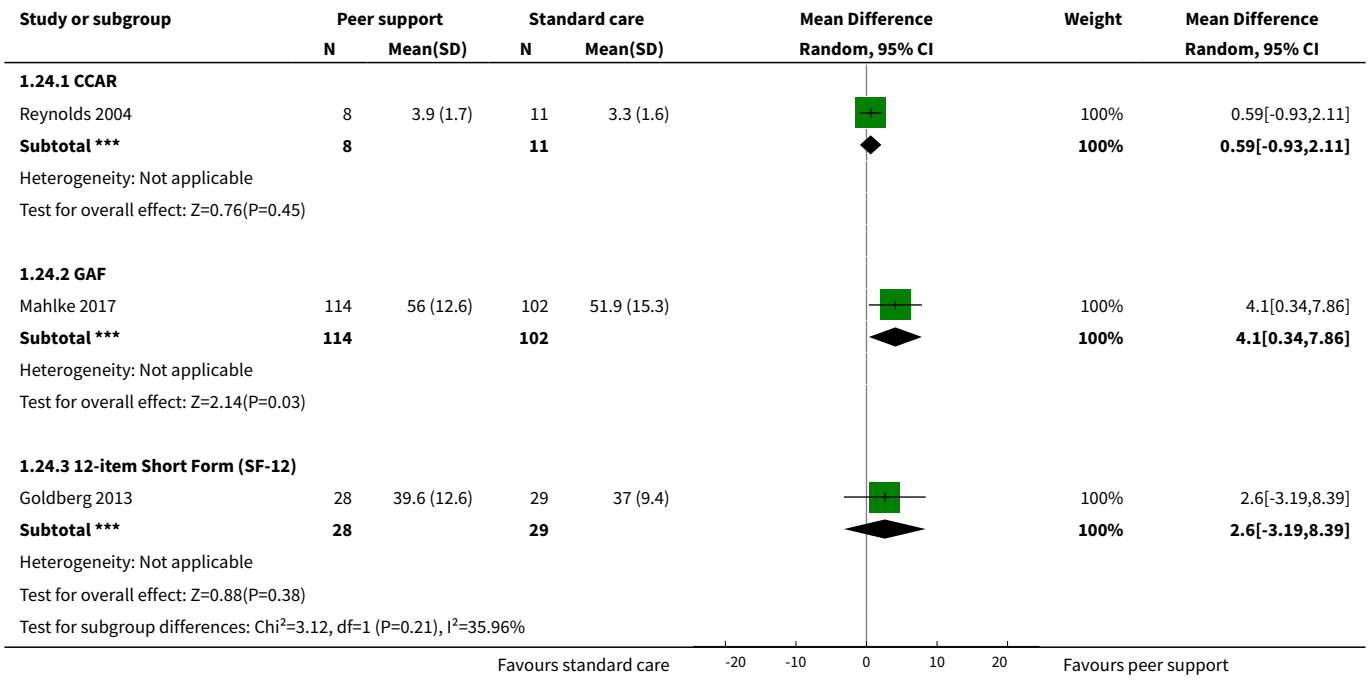
Study	Intervention	Mean	SD	N
Alcohol/drug use (BASIS-24, high = strong) – medium term				
Eisen 2012	Peer support	0.51	0.62	74
Eisen 2012	Standard care	0.56	0.83	84
Alcohol use (Addiction Severity Index (ASI), high = strong) – medium term				
Rowe 2007	Peer support	0.10	0.18	41
Rowe 2007	Standard care	0.10	0.13	27
Alcohol use (ASI, high = strong) – long term				
Rowe 2007	Peer support	0.07	0.13	40
Rowe 2007	Standard care	0.11	0.16	29
Drug use (ASI, high = strong) – medium term				
Rowe 2007	Peer support	0.04	0.06	41
Rowe 2007	Standard care	0.07	0.09	27
Drug use (ASI, high = strong) – long term				
Rowe 2007	Peer support	0.04	0.05	40
Rowe 2007	Standard care	0.04	0.07	29

Analysis 1.23. Comparison 1 Peer support + standard care versus standard care alone, Outcome 23 Leaving the study early – for any reason.

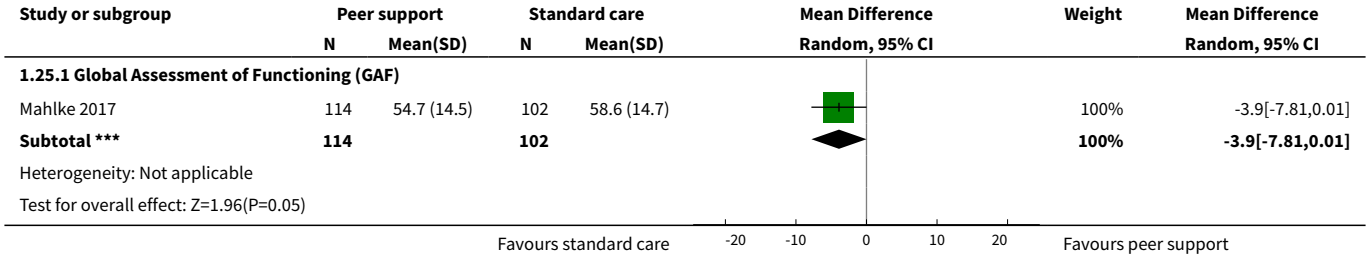
Study or subgroup	Peer support n/N	Standard care n/N	Risk Ratio M-H, Random, 95% CI	Weight	Risk Ratio M-H, Random, 95% CI
1.23.1 Medium term					
Druss 2010	4/41	11/39		6.45%	0.35[0.12,1]
Goldberg 2013	4/32	2/31		2.79%	1.94[0.38,9.83]
Kelly 2014	0/12	3/12		0.91%	0.14[0.01,2.5]
Mahlke 2017	28/114	41/102		37.2%	0.61[0.41,0.91]
Reynolds 2004	3/11	3/14		3.78%	1.27[0.32,5.12]
Van Gestel-Timmermans 2012	42/168	58/165		48.87%	0.71[0.51,0.99]
Subtotal (95% CI)	**378**	**363**		**100%**	**0.66[0.51,0.87]**
Total events: 81 (Peer support), 118 (Standard care)					
Heterogeneity: Tau^2=0.01; Chi^2=5.42, df=5(P=0.37); I^2=7.72%					
Test for overall effect: Z=2.93(P=0)					
1.23.2 Long term					
Castelein 2008	4/56	5/50		30.79%	0.71[0.2,2.51]
Cook 2012b	68/276	11/279		34.2%	6.25[3.38,11.56]
Mahlke 2017	30/114	52/102		35%	0.52[0.36,0.74]
Subtotal (95% CI)	**446**	**431**		**100%**	**1.34[0.19,9.22]**
Total events: 102 (Peer support), 68 (Standard care)					
Heterogeneity: Tau^2=2.73; Chi^2=53.42, df=2(P<0.0001); I^2=96.26%					
Test for overall effect: Z=0.3(P=0.77)					
Test for subgroup differences: Chi^2=0.5, df=1 (P=0.48), I^2=0%					



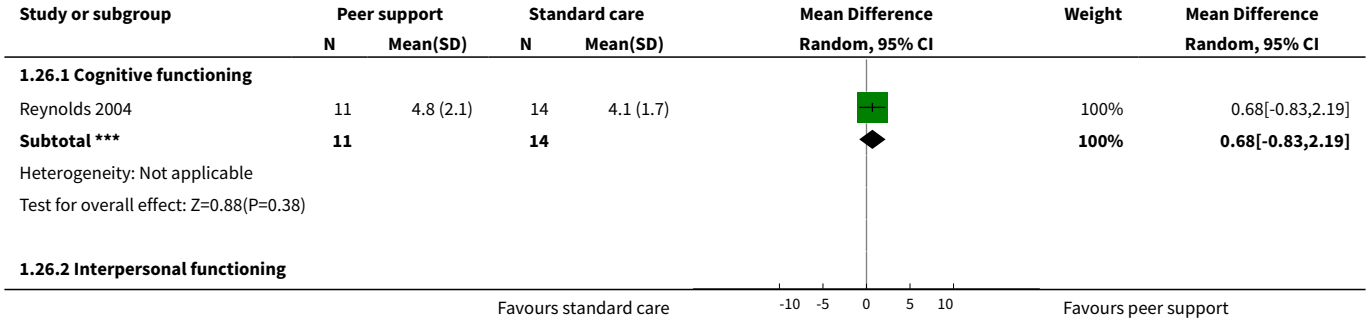
Analysis 1.24. Comparison 1 Peer support + standard care versus standard care alone, Outcome 24 Functioning: 1a. General: mean total endpoint score (various scales, high = good) – medium term.

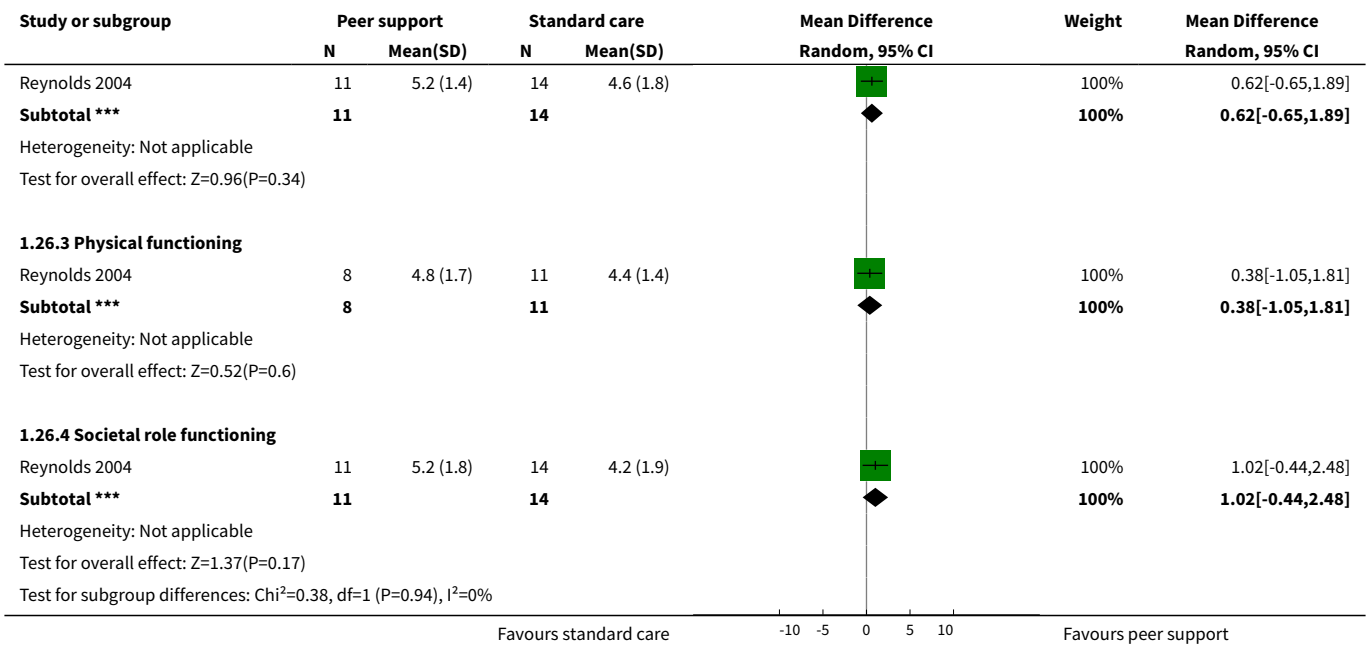


Analysis 1.25. Comparison 1 Peer support + standard care versus standard care alone, Outcome 25 Functioning: 1b. General: mean total endpoint score (various scales, high = good) – long term.

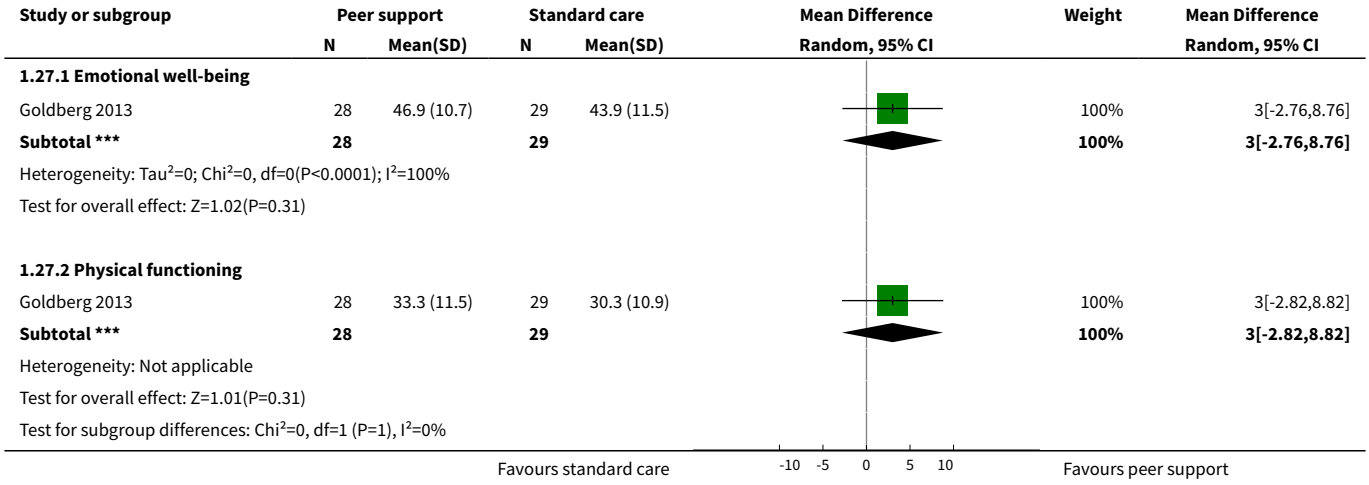


Analysis 1.26. Comparison 1 Peer support + standard care versus standard care alone, Outcome 26 Functioning: 2a. Specific: various aspects – mean endpoint score (CCAR subscales, high = good) – medium term.





Analysis 1.27. Comparison 1 Peer support + standard care versus standard care alone, Outcome 27 Functioning: 2b. Specific: various aspects – mean endpoint score (SF-12 subscales, high = good) – medium term.



Analysis 1.28. Comparison 1 Peer support + standard care versus standard care alone, Outcome 28 Functioning: 3. Specific: daily living – mean endpoint score (CCAR, high = good) – medium term (skewed data).

Functioning: 3. Specific: daily living – mean endpoint score (CCAR, high = good) – medium term (skewed data)

Study	Intervention	Mean	SD	N
Reynolds 2004	peer support	5.11	1.62	11
Reynolds 2004	standard care	3.6	1.9	14

Analysis 1.29. Comparison 1 Peer support + standard care versus standard care alone, Outcome 29 Functioning: 4. Specific: self-management – mean endpoint score (IMSM, high = good) (skewed data).

Functioning: 4. Specific: self-management – mean endpoint score (IMSM, high = good) (skewed data)

Study	Heading 1	Heading 2	Heading 3	Heading 4	Heading 5
			IMSM		
Goldberg 2013	Peer support	2.9	1.2	28	

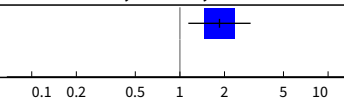
Analysis 1.30. Comparison 1 Peer support + standard care versus standard care alone, Outcome 30 Functioning: 5. Specific: contact with justice system – criminal justice charges (skewed data).

Functioning: 5. Specific: contact with justice system – criminal justice charges (skewed data)

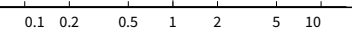
Study	Intervention	Mean	SD	N
Felony (counts of criminal justice charges, high = more criminal charges) -- medium term				
Rowe 2007	peer support	0.19	0.46	73
Rowe 2007	standard care	0.10	0.49	41
Felony (counts of criminal justice charges, high = more criminal charges) – long term				
Rowe 2007	peer support	0.10	0.30	73
Rowe 2007	standard care	0.02	0.16	41
Infraction (counts of criminal justice charges, high = more criminal charges) – medium term				
Rowe 2007	peer support	0.08	0.28	73
Rowe 2007	standard care	0.15	0.48	41
Infraction (counts of criminal justice charges, high = more criminal charges) – long term				
Rowe 2007	peer support	0.05	0.23	73
Rowe 2007	standard care	0.00	0.00	41
Misdemeanour (counts of criminal justice charges, high = more criminal charges) – medium term				
Rowe 2007	peer support	0.89	1.50	73
Rowe 2007	standard care	0.46	1.03	41
Misdemeanour (counts of criminal justice charges, high = more criminal charges) – long term				
Rowe 2007	peer support	0.53	1.30	73
Rowe 2007	standard care	0.27	0.63	41
Total charges (counts of criminal justice charges, high = more criminal charges) – medium term				
Rowe 2007	peer support	1.18	1.87	73
Rowe 2007	standard care	0.76	1.50	41
Total charges (counts of criminal justice charges, high = more criminal charges) – long term				
Rowe 2007	peer support	0.75	1.71	73
Rowe 2007	standard care	0.32	0.76	41
Violation (counts of criminal justice charges, high = more criminal charges) -- medium term				
Rowe 2007	peer support	0.01	0.12	73
Rowe 2007	standard care	0.05	0.22	41
Violation (counts of criminal justice charges, high = more criminal charges) – long term				
Rowe 2007	peer support	0.07	0.30	73
Rowe 2007	standard care	0.02	0.16	41

Analysis 1.31. Comparison 1 Peer support + standard care versus standard care alone, Outcome 31 Peer outcomes: 1a. Impact on the participant and peer supporter: improved peer contact – mean endpoint score (Personal Network Questionnaire (PNQ), high = good) – long term.

Study or subgroup	Peer support n/N	Standard care n/N	Risk Ratio M-H, Random, 95% CI	Weight	Risk Ratio M-H, Random, 95% CI
Castelein 2008	31/56	15/50		100%	1.85[1.14,3]



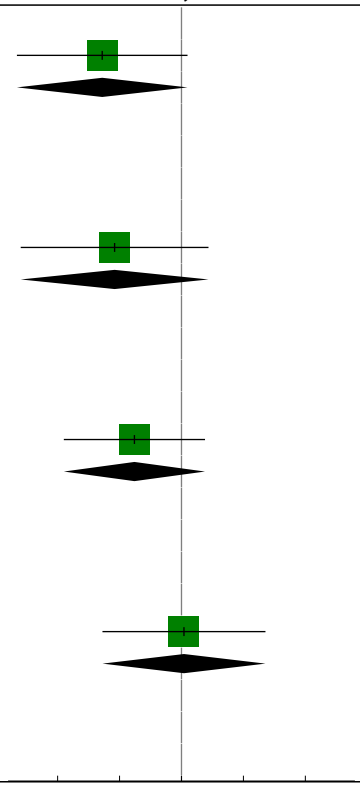
Study or subgroup	Peer support n/N	Standard care n/N	Risk Ratio M-H, Random, 95% CI	Weight	Risk Ratio M-H, Random, 95% CI
Total (95% CI)	**56**	**50**		**100%**	**1.85[1.14,3]**
Total events: 31 (Peer support), 15 (Standard care)					
Heterogeneity: Not applicable					
Test for overall effect: Z=2.48(P=0.01)					



Favours Standard care | Favours Peer Support

Analysis 1.32. Comparison 1 Peer support + standard care versus standard care alone, Outcome 32 Peer outcomes: 1b. Impact on participant and peer supporter: negative aspects – mean endpoint score (BLR subscales, high = true) – medium term.

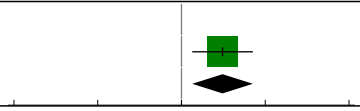
Study or subgroup	Peer support N	Peer support Mean(SD)	Standard care N	Standard care Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
1.32.1 Negative empathy							
Sells 2008	54	3.5 (0.9)	51	3.9 (0.9)		100%	-0.32[-0.66,0.02]
Subtotal ***	**54**		**51**			**100%**	**-0.32[-0.66,0.02]**
Heterogeneity: Not applicable							
Test for overall effect: Z=1.82(P=0.07)							
1.32.2 Negative regard							
Sells 2008	54	2.7 (1)	51	2.9 (1)		100%	-0.27[-0.65,0.11]
Subtotal ***	**54**		**51**			**100%**	**-0.27[-0.65,0.11]**
Heterogeneity: Not applicable							
Test for overall effect: Z=1.4(P=0.16)							
1.32.3 Negative overall relationship							
Sells 2008	54	3.3 (0.8)	51	3.5 (0.7)		100%	-0.19[-0.48,0.1]
Subtotal ***	**54**		**51**			**100%**	**-0.19[-0.48,0.1]**
Heterogeneity: Tau^2=0; Chi^2=0, df=0(P<0.0001); I^2=100%							
Test for overall effect: Z=1.31(P=0.19)							
1.32.4 Negative unconditionality							
Sells 2008	54	3.8 (0.9)	51	3.8 (0.8)		100%	0.01[-0.32,0.34]
Subtotal ***	**54**		**51**			**100%**	**0.01[-0.32,0.34]**
Heterogeneity: Not applicable							
Test for overall effect: Z=0.06(P=0.95)							
Test for subgroup differences: Chi^2=2.13, df=1 (P=0.54), I^2=0%							



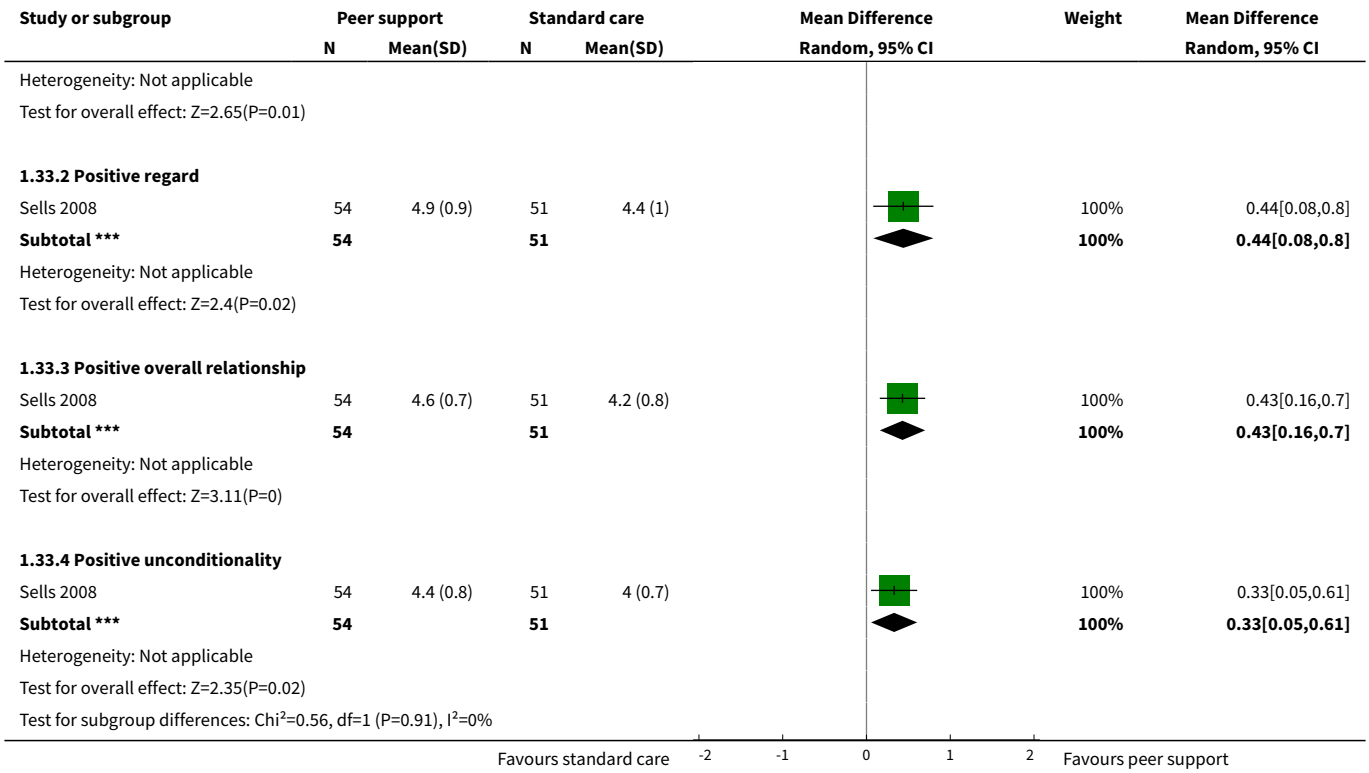
Favours peer support | Favours standard care

Analysis 1.33. Comparison 1 Peer support + standard care versus standard care alone, Outcome 33 Peer outcomes: 1c. Impact on participant and peer supporter: positive aspects – mean endpoint score (Barrett-Lennard Relationship Inventory (BLRI) subscales, high = true) – medium term.

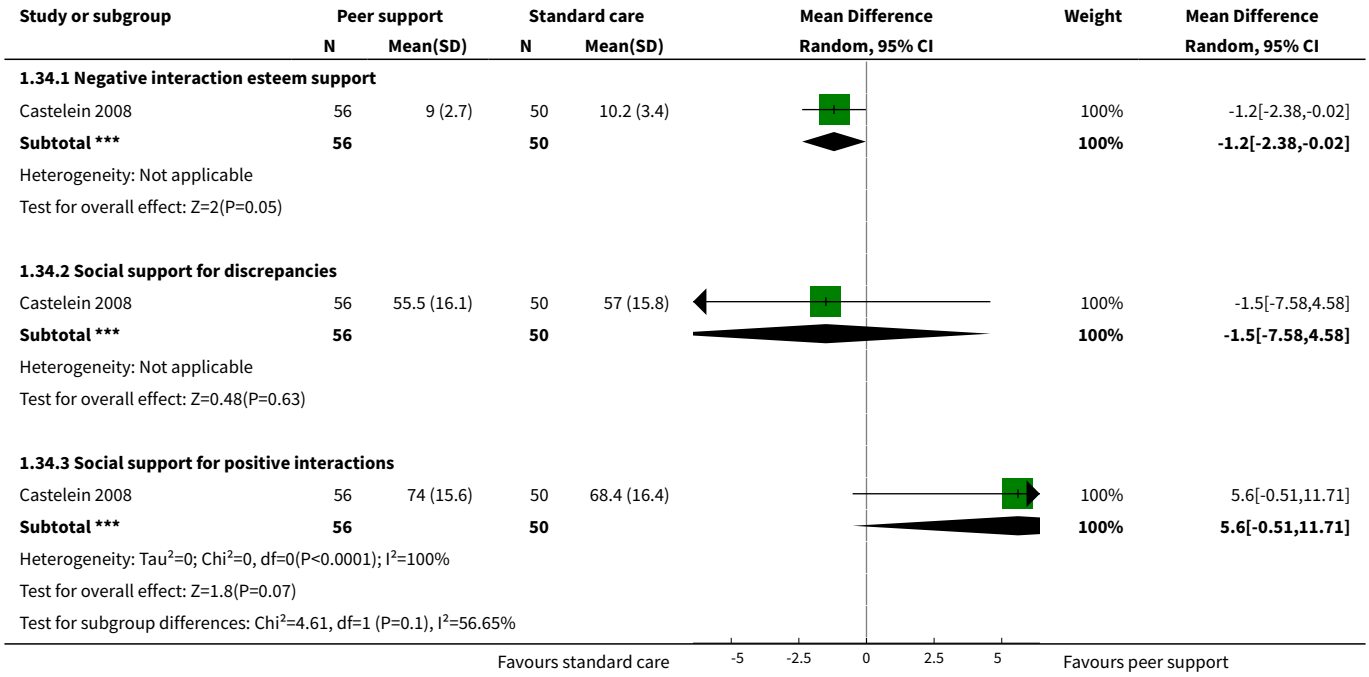
Study or subgroup	Peer support N	Peer support Mean(SD)	Standard care N	Standard care Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
1.33.1 Positive empathy							
Sells 2008	54	4.7 (0.9)	51	4.2 (1)		100%	0.49[0.13,0.85]
Subtotal ***	**54**		**51**			**100%**	**0.49[0.13,0.85]**



Favours standard care | Favours peer support

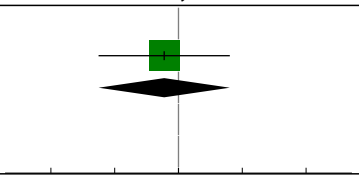


Analysis 1.34. Comparison 1 Peer support + standard care versus standard care alone, Outcome 34 Peer outcomes: 1d. Impact on participant and peer supporter: various aspects – mean endpoint score (Social Support List (SSL) subscales, high = increased need for support) – long term.



Analysis 1.35. Comparison 1 Peer support + standard care versus standard care alone, Outcome 35 Peer outcomes: 1e. Impact on participant and peer supporter: social support – mean endpoint score (Medical Outcomes Study Social Support Survey (MOSSSS), high = good).

Study or subgroup	Peer support N	Peer support Mean(SD)	Standard care N	Standard care Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
1.35.1 Medium term							
Eisen 2012	74	63.2 (16.6)	84	64.3 (16.3)		100%	-1.12[-6.26,4.02]
Subtotal ***	**74**		**84**			**100%**	**-1.12[-6.26,4.02]**
Heterogeneity: Not applicable							
Test for overall effect: Z=0.43(P=0.67)							



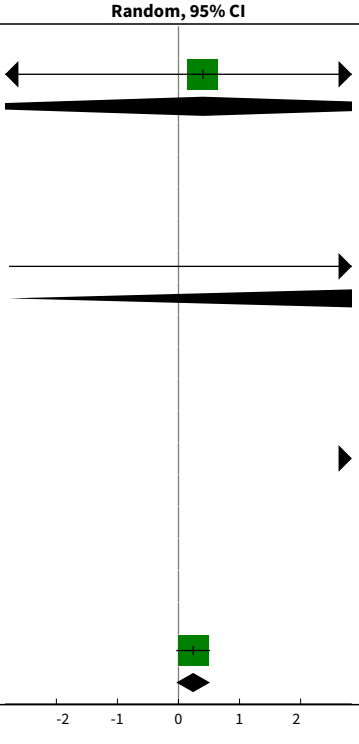
Analysis 1.36. Comparison 1 Peer support + standard care versus standard care alone, Outcome 36 Peer outcomes: 1f. Impact on participant and peer supporter: accessing social support (IMSM, high = greater amount of support obtained) – medium term (skewed data).

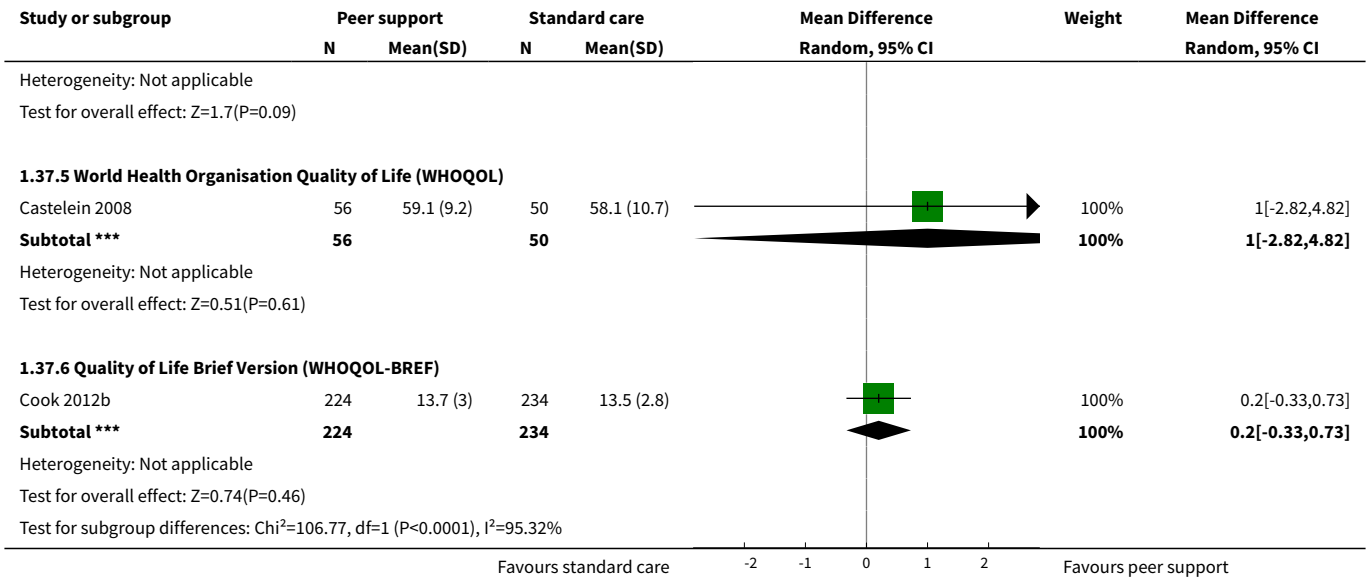
Peer outcomes: 1f. Impact on participant and peer supporter: accessing social support (IMSM, high = greater amount of support obtained) – medium term (skewed data)

Study	Intervention	Mean	SD	N
Goldberg 2013	peer support	2.5	1.3	28
Goldberg 2013	standard care	2.5	1.3	29

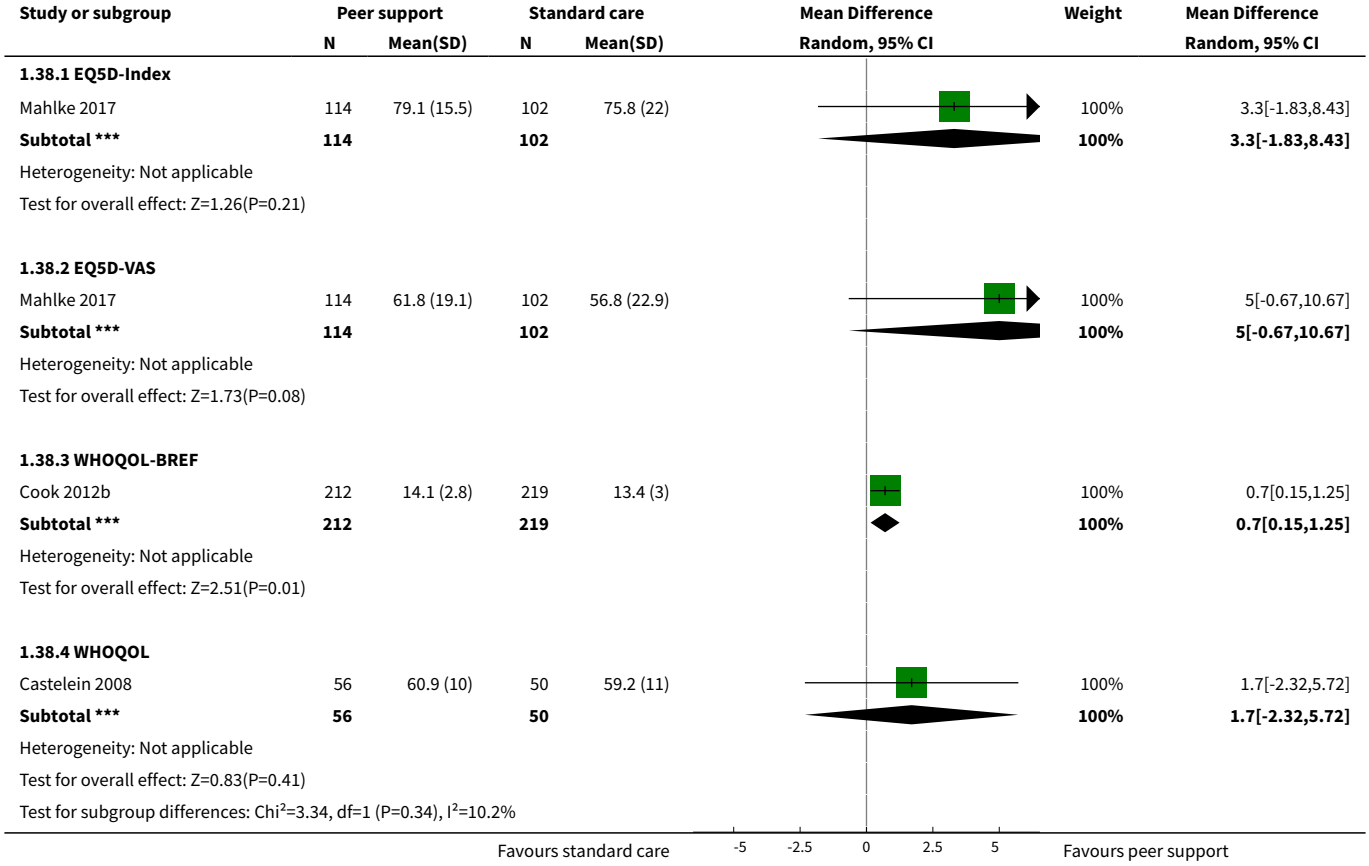
Analysis 1.37. Comparison 1 Peer support + standard care versus standard care alone, Outcome 37 Peer outcomes: 2a. Quality of life for participant and peer supporter: overall – mean total endpoint (various scales, high = good) – medium term.

Study or subgroup	Peer support N	Peer support Mean(SD)	Standard care N	Standard care Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
1.37.1 EuroQol: Five Dimensions (EQ5D)-Index							
Mahlke 2017	114	75.4 (16.5)	102	75 (20)		100%	0.4[-4.52,5.32]
Subtotal ***	**114**		**102**			**100%**	**0.4[-4.52,5.32]**
Heterogeneity: Tau^2=0; Chi^2=0, df=0(P<0.0001); I^2=100%							
Test for overall effect: Z=0.16(P=0.87)							
1.37.2 EuroQol: Five Dimensions-Visual Analogue Scale (EQ5D-VAS)							
Mahlke 2017	114	59.7 (22.7)	102	56.5 (22.1)		100%	3.2[-2.77,9.17]
Subtotal ***	**114**		**102**			**100%**	**3.2[-2.77,9.17]**
Heterogeneity: Not applicable							
Test for overall effect: Z=1.05(P=0.29)							
1.37.3 General Quality of Life Inventory (GQOLI-74)							
Qian 2015	50	216.3 (19)	50	175.9 (20)		100%	40.34[32.7,47.98]
Subtotal ***	**50**		**50**			**100%**	**40.34[32.7,47.98]**
Heterogeneity: Not applicable							
Test for overall effect: Z=10.35(P<0.0001)							
1.37.4 Manchester Short Assessment of Quality of Life (MSAQOL)							
Van Gestel-Timmermans 2012	111	4.6 (1)	97	4.4 (1.1)		100%	0.24[-0.04,0.52]
Subtotal ***	**111**		**97**			**100%**	**0.24[-0.04,0.52]**



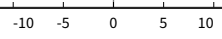


Analysis 1.38. Comparison 1 Peer support + standard care versus standard care alone, Outcome 38 Peer outcomes: 2b. Quality of life for participant and peer supporter: overall – mean total endpoint (various scales, high = good) – long term.



Analysis 1.39. Comparison 1 Peer support + standard care versus standard care alone, Outcome 39 Peer outcomes: 3a. Quality of life for participant and peer supporter: specific aspects – mean endpoint score (GQOLI-74 subscales, high = good) – medium term.

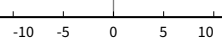
Study or subgroup	Peer support N	Peer support Mean(SD)	Standard care N	Standard care Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
1.39.1 Mental health							
Qian 2015	50	55 (8.9)	50	38 (9.5)		100%	16.95[13.34,20.56]
Subtotal ***	**50**		**50**			**100%**	**16.95[13.34,20.56]**
Heterogeneity: Not applicable							
Test for overall effect: Z=9.2(P<0.0001)							
1.39.2 Physical quality of life							
Qian 2015	50	51.4 (10.8)	50	50 (8.1)		100%	1.43[-2.31,5.17]
Subtotal ***	**50**		**50**			**100%**	**1.43[-2.31,5.17]**
Heterogeneity: Not applicable							
Test for overall effect: Z=0.75(P=0.45)							
1.39.3 Physical health							
Qian 2015	50	55.6 (9)	50	40.6 (10.3)		100%	15.08[11.29,18.87]
Subtotal ***	**50**		**50**			**100%**	**15.08[11.29,18.87]**
Heterogeneity: Tau^2=0; Chi^2=0, df=0(P<0.0001); I^2=100%							
Test for overall effect: Z=7.81(P<0.0001)							
1.39.4 Social function							
Qian 2015	50	56 (7.3)	50	40.2 (9)		100%	15.87[12.66,19.08]
Subtotal ***	**50**		**50**			**100%**	**15.87[12.66,19.08]**
Heterogeneity: Not applicable							
Test for overall effect: Z=9.68(P<0.0001)							
Test for subgroup differences: Chi^2=45.44, df=1 (P<0.0001), I^2=93.4%							



Favours standard care | Favours peer support

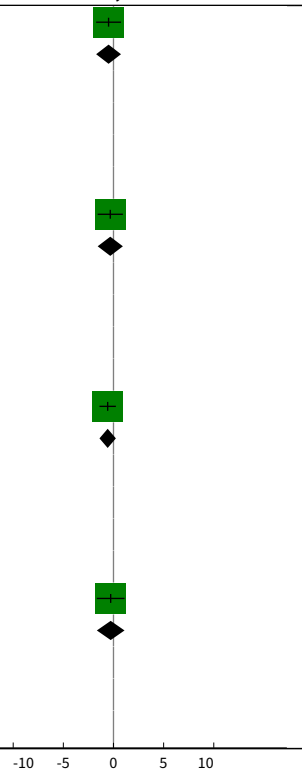
Analysis 1.40. Comparison 1 Peer support + standard care versus standard care alone, Outcome 40 Peer outcomes: 3b. Quality of life for participant and peer supporter: specific aspects – mean endpoint score (QOLI-BREF subscales, high = good) – medium term.

Study or subgroup	Peer support N	Peer support Mean(SD)	Standard care N	Standard care Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
1.40.1 Amount of time spent with others							
Reynolds 2004	8	4.4 (1.5)	11	4.4 (1.3)		100%	0.04[-1.24,1.32]
Subtotal ***	**8**		**11**			**100%**	**0.04[-1.24,1.32]**
Heterogeneity: Not applicable							
Test for overall effect: Z=0.06(P=0.95)							
1.40.2 General life satisfaction							
Reynolds 2004	8	4.6 (1.6)	11	4.6 (0.8)		100%	-0.04[-1.25,1.17]
Subtotal ***	**8**		**11**			**100%**	**-0.04[-1.25,1.17]**
Heterogeneity: Not applicable							
Test for overall effect: Z=0.06(P=0.95)							
1.40.3 Life in general							



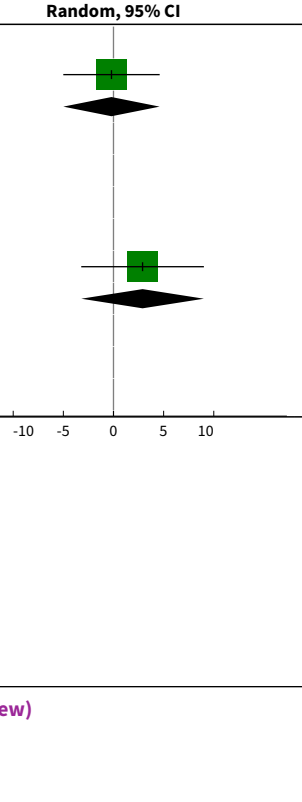
Favours standard care | Favours peer support

Study or subgroup	Peer support N	Peer support Mean(SD)	Standard care N	Standard care Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
Reynolds 2004	8	4.1 (1.5)	11	4.6 (1.1)		100%	-0.49[-1.73,0.75]
Subtotal ***	**8**		**11**			**100%**	**-0.49[-1.73,0.75]**
Heterogeneity: Not applicable							
Test for overall effect: Z=0.77(P=0.44)							
1.40.4 Living arrangements							
Reynolds 2004	8	4.8 (1.4)	11	5.1 (1.4)		100%	-0.32[-1.58,0.94]
Subtotal ***	**8**		**11**			**100%**	**-0.32[-1.58,0.94]**
Heterogeneity: Not applicable							
Test for overall effect: Z=0.5(P=0.62)							
1.40.5 Privacy							
Reynolds 2004	8	5.2 (1)	11	5.8 (0.8)		100%	-0.58[-1.4,0.24]
Subtotal ***	**8**		**11**			**100%**	**-0.58[-1.4,0.24]**
Heterogeneity: Not applicable							
Test for overall effect: Z=1.39(P=0.16)							
1.40.6 Relax							
Reynolds 2004	8	4.2 (1.7)	11	4.5 (1.2)		100%	-0.28[-1.66,1.1]
Subtotal ***	**8**		**11**			**100%**	**-0.28[-1.66,1.1]**
Heterogeneity: Not applicable							
Test for overall effect: Z=0.4(P=0.69)							
Test for subgroup differences: Chi^2=0.97, df=1 (P=0.97), I^2=0%							



Analysis 1.41. Comparison 1 Peer support + standard care versus standard care alone, Outcome 41 Peer outcomes: 3c. Quality of life for participant and peer supporter: specific aspects – mean endpoint score (36-item Short Form (SF-36) subscales, high = good) – medium term.

Study or subgroup	Peer support N	Peer support Mean(SD)	Standard care N	Standard care Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
1.41.1 Mental health							
Druss 2010	41	36.8 (10)	39	37 (11.8)		100%	-0.2[-5,4.6]
Subtotal ***	**41**		**39**			**100%**	**-0.2[-5,4.6]**
Heterogeneity: Not applicable							
Test for overall effect: Z=0.08(P=0.93)							
1.41.2 Physical health							
Druss 2010	41	42.9 (14.2)	39	40 (13.7)		100%	2.9[-3.21,9.01]
Subtotal ***	**41**		**39**			**100%**	**2.9[-3.21,9.01]**
Heterogeneity: Tau^2=0; Chi^2=0, df=0(P<0.0001); I^2=100%							
Test for overall effect: Z=0.93(P=0.35)							
Test for subgroup differences: Chi^2=0.61, df=1 (P=0.43), I^2=0%							

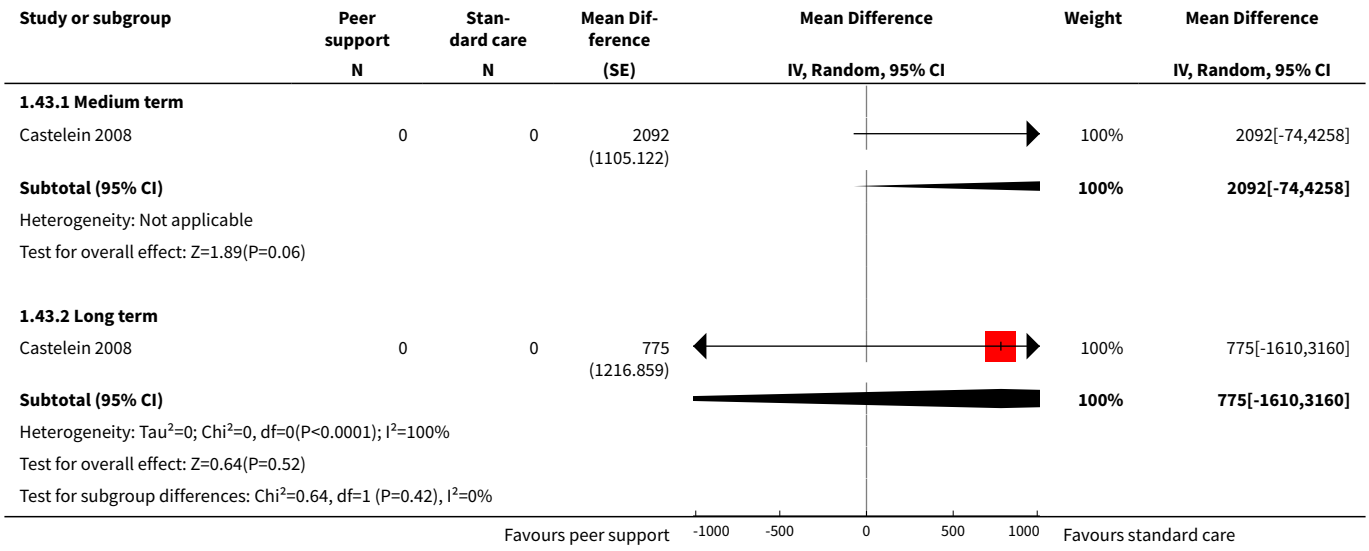


Analysis 1.42. Comparison 1 Peer support + standard care versus standard care alone, Outcome 42 Peer outcomes: 3d. Quality of life for participant and peer supporter: specific aspects – mean endpoint score (QOL-BREF subscale, high = good) – medium term (skewed data).

Peer outcomes: 3d. Quality of life for participant and peer supporter: specific aspects – mean endpoint score (QOL-BREF subscale, high = good) – medium term (skewed data)

Study	Intervention	Mean	SD	N
Reynolds 2004	Peer support	3.56	2.01	8
Reynolds 2004	Standard care	4.1	1.52	11

Analysis 1.43. Comparison 1 Peer support + standard care versus standard care alone, Outcome 43 Economic cost: 1. Direct and indirect costs (Euro): total cost (high = poor).



Analysis 1.44. Comparison 1 Peer support + standard care versus standard care alone, Outcome 44 Economic outcomes: 2. Direct costs (Euro): for minimally guided peer support (high = poor) – long term (skewed data).

Economic outcomes: 2. Direct costs (Euro): for minimally guided peer support (high = poor) – long term (skewed data)

Study	Intervention	Mean	SD	N
Castelein 2008	Peer support+ standard care	250	97	56
Castelein 2008	Standard care	0	0	50

Analysis 1.45. Comparison 1 Peer support + standard care versus standard care alone, Outcome 45 Economic outcomes: 3a. Indirect cost of care (Euro): for inpatient and semi-inpatient care (high = poor) – long term (skewed data).

Economic outcomes: 3a. Indirect cost of care (Euro): for inpatient and semi-inpatient care (high = poor) – long term (skewed data)

Study	Intervention	Mean	SD	N
		Hospital admission		
Castelein 2008	Peer support+ standard care	1712	5314	56
Castelein 2008	Standard care	1471	5741	50
		Day care		
Castelein 2008	Peer support+ standard care	767	2377	56
Castelein 2008	Standard care	687	2166	50

Economic outcomes: 3a. Indirect cost of care (Euro): for inpatient and semi-inpatient care (high = poor) – long term (skewed data)

Study	Intervention	Mean	SD	N
Sheltered living				
Castelein 2008	Peer support+ standard care	820	2984	56
Castelein 2008	Standard care	230	1624	50

Analysis 1.46. Comparison 1 Peer support + standard care versus standard care alone, Outcome 46 Economic outcomes: 3b. Indirect cost of care (Euro): for outpatient and community care (high = poor) – long term (skewed data).

Economic outcomes: 3b. Indirect cost of care (Euro): for outpatient and community care (high = poor) – long term (skewed data)

Study	Intervention	Mean	SD	N
Psychiatrist				
Castelein 2008	Peer support+ standard care	255	348	56
Castelein 2008	Standard care	164	218	50
Psychologist				
Castelein 2008	Peer support+ standard care	153	359	56
Castelein 2008	Standard care	81	208	50
Social psychiatric nurse				
Castelein 2008	Peer support+ standard care	249	558	56
Castelein 2008	Standard care	203	409	50
Social worker				
Castelein 2008	Peer support+ standard care	0	0	56
Castelein 2008	Standard care	54	210	50
Crisis intervention				
Castelein 2008	Peer support+ standard care	23	77	56
Castelein 2008	Standard care	13	51	50
Psychiatric home care				
Castelein 2008	Peer support+ standard care	249	1069	56
Castelein 2008	Standard care	242	996	50
Consultation clinic for alcohol and drug addiction				
Castelein 2008	Peer support+ standard care	16	122	56
Castelein 2008	Standard care	9	64	50
Other outpatient care				
Castelein 2008	Peer support+ standard care	23	96	56
Castelein 2008	Standard care	89	405	50

Analysis 1.47. Comparison 1 Peer support + standard care versus standard care alone, Outcome 47 Economic outcomes: 3c. Indirect cost of care (Euro): for general healthcare (high = poor) – long term (skewed data).

Economic outcomes: 3c. Indirect cost of care (Euro): for general healthcare (high = poor) – long term (skewed data)

Study	Intervention	Mean	SD	N
General practitioner				
Castelein 2008	Peer support+ standard care	18	46	56
Castelein 2008	Standard care	29	90	50
Alternative health care				
Castelein 2008	Peer support+ standard care	13	86	56
Castelein 2008	Standard care	2	13	50
Emergency care				
Castelein 2008	Peer support+ standard care	0	0	56
Castelein 2008	Standard care	6	28	50
Other general health care				
Castelein 2008	Peer support+ standard care	8	57	56

Economic outcomes: 3c. Indirect cost of care (Euro): for general healthcare (high = poor) – long term (skewed data)

Study	Intervention	Mean	SD	N
Castelein 2008	Standard care	5	31	50

Analysis 1.48. Comparison 1 Peer support + standard care versus standard care alone, Outcome 48 Economic outcomes: 3d. Indirect costs (Euro): of day activity institutions (high = poor) – long term (skewed data).

Economic outcomes: 3d. Indirect costs (Euro): of day activity institutions (high = poor) – long term (skewed data)

Study	Intervention	Mean	SD	N
Day activity centre				
Castelein 2008	Peer support+ standard care	83	217	56
Castelein 2008	Standard care	137	399	50
Drop-in centre				
Castelein 2008	Peer support+ standard care	79	321	56
Castelein 2008	Standard care	145	493	50
Recreation/activity centre				
Castelein 2008	Peer support+ standard care	6	42	56
Castelein 2008	Standard care	32	132	50
Other institutions				
Castelein 2008	Peer support+ standard care	29	173	56
Castelein 2008	Standard care	43	165	50

Analysis 1.49. Comparison 1 Peer support + standard care versus standard care alone, Outcome 49 Economic outcomes: 3e. Indirect cost (Euro): of medication (high = poor) – long term (skewed data).

Economic outcomes: 3e. Indirect cost (Euro): of medication (high = poor) – long term (skewed data)

Study	Intervention	Mean	SD	N
Prescribed				
Castelein 2008	Peer support+ standard care	503	553	56
Castelein 2008	Standard care	504	460	50
Non-prescribed				
Castelein 2008	Peer support+ standard care	13	54	56
Castelein 2008	Standard care	6	32	50

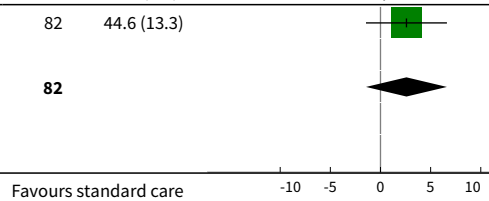
Comparison 2. Peer support plus standard care versus clinician-led support plus standard care

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Global state: 1. General health – mean total endpoint score (Veterans RAND 12-Item Health Survey (VR-12), high = good) – medium term	1	156	Mean Difference (IV, Random, 95% CI)	2.59 [-1.45, 6.63]
2 Mental state: 1a. Specific: various aspects – mean endpoint score (various scales, high = good) – medium term	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
2.1 Hope (State Hope Scale (SHS))	1	156	Mean Difference (IV, Random, 95% CI)	-0.59 [-1.80, 0.62]
2.2 Recovery (Recovery Assessment Scale (RAS))	1	156	Mean Difference (IV, Random, 95% CI)	-0.5 [-7.13, 6.13]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
2.3 Empowerment (Rogers Empowerment Scale (RES))	1	156	Mean Difference (IV, Random, 95% CI)	-0.65 [-2.95, 1.65]
3 Mental state: 1b. Specific: various aspects – mean endpoint score (Patient Activation Scale (PAS) subscales, high = good) – medium term	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
3.1 Activation (patient)	1	156	Mean Difference (IV, Random, 95% CI)	0.30 [-1.64, 2.24]
4 Mental state: 1c. Specific: various aspects – mean endpoint score (BASIS subscales, high = poor) – medium term (skewed data)			Other data	No numeric data
4.1 Self-harm			Other data	No numeric data
4.2 Emotional liability			Other data	No numeric data
4.4 Psychotic symptoms			Other data	No numeric data
4.5 Interpersonal relationship			Other data	No numeric data
4.6 Depression			Other data	No numeric data
4.17 Psychotic symptoms			Other data	No numeric data
5 Behaviour: 1. Specific: drug/alcohol use – mean endpoint score (BASIS subscale, high = poor) – medium term (skewed data)			Other data	No numeric data
5.3 Alcohol/drug use			Other data	No numeric data
6 Peer outcomes: 1. Impact on the service user and peer supporter: social support – mean endpoint score (MOSSSS, high = good) – medium term	1	156	Mean Difference (IV, Random, 95% CI)	4.97 [-0.62, 10.56]

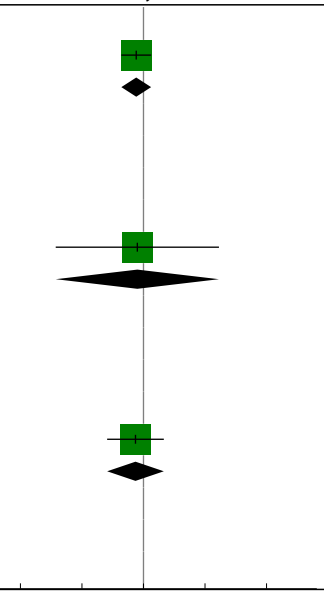
Analysis 2.1. Comparison 2 Peer support plus standard care versus clinician-led support plus standard care, Outcome 1 Global state: 1. General health – mean total endpoint score (Veterans RAND 12-Item Health Survey (VR-12), high = good) – medium term.

Study or subgroup	Peer support N	Peer support Mean(SD)	Standard care N	Standard care Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
Eisen 2012	74	47.2 (12.4)	82	44.6 (13.3)		100%	2.59[-1.45,6.63]
Total ***	**74**		**82**			**100%**	**2.59[-1.45,6.63]**
Heterogeneity: Not applicable							
Test for overall effect: Z=1.26(P=0.21)							



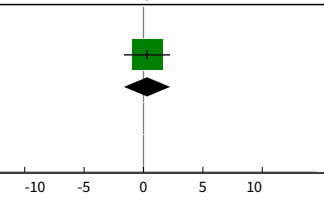
Analysis 2.2. Comparison 2 Peer support plus standard care versus clinician-led support plus standard care, Outcome 2 Mental state: 1a. Specific: various aspects – mean endpoint score (various scales, high = good) – medium term.

Study or subgroup	Peer support		Other psychological interventions		Mean Difference	Weight	Mean Difference
	N	Mean(SD)	N	Mean(SD)	Random, 95% CI		Random, 95% CI
2.2.1 Hope (State Hope Scale (SHS))							
Eisen 2012	74	23.5 (3.7)	82	24.1 (4)		100%	-0.59[-1.8,0.62]
Subtotal ***	**74**		**82**			**100%**	**-0.59[-1.8,0.62]**
Heterogeneity: Not applicable							
Test for overall effect: Z=0.96(P=0.34)							
2.2.2 Recovery (Recovery Assessment Scale (RAS))							
Eisen 2012	74	164.2 (20.1)	82	164.7 (22.2)		100%	-0.5[-7.13,6.13]
Subtotal ***	**74**		**82**			**100%**	**-0.5[-7.13,6.13]**
Heterogeneity: Not applicable							
Test for overall effect: Z=0.15(P=0.88)							
2.2.3 Empowerment (Rogers Empowerment Scale (RES))							
Eisen 2012	74	80.2 (6.7)	82	80.9 (8)		100%	-0.65[-2.95,1.65]
Subtotal ***	**74**		**82**			**100%**	**-0.65[-2.95,1.65]**
Heterogeneity: Not applicable							
Test for overall effect: Z=0.55(P=0.58)							
Test for subgroup differences: Chi^2=0, df=1 (P=1), I^2=0%							



Analysis 2.3. Comparison 2 Peer support plus standard care versus clinician-led support plus standard care, Outcome 3 Mental state: 1b. Specific: various aspects – mean endpoint score (Patient Activation Scale (PAS) subscales, high = good) – medium term.

Study or subgroup	Peer support		Other psychological interventions		Mean Difference	Weight	Mean Difference
	N	Mean(SD)	N	Mean(SD)	Random, 95% CI		Random, 95% CI
2.3.1 Activation (patient)							
Eisen 2012	74	29.1 (5.9)	82	28.8 (6.5)		100%	0.3[-1.64,2.24]
Subtotal ***	**74**		**82**			**100%**	**0.3[-1.64,2.24]**
Heterogeneity: Not applicable							
Test for overall effect: Z=0.3(P=0.76)							



Analysis 2.4. Comparison 2 Peer support plus standard care versus clinician-led support plus standard care, Outcome 4 Mental state: 1c. Specific: various aspects – mean endpoint score (BASIS subscales, high = poor) – medium term (skewed data).

Mental state: 1c. Specific: various aspects – mean endpoint score (BASIS subscales, high = poor) – medium term (skewed data)

Study	Intervention	Mean	SD	N
		Self-harm		
Eisen 2012	peer support	0.18	0.50	74
Eisen 2012	other psychological interventions	0.22	0.57	82

Mental state: 1c. Specific: various aspects – mean endpoint score (BASIS subscales, high = poor) – medium term (skewed data)

Study	Intervention	Mean	SD	N
		Emotional liability		
Eisen 2012	peer support	1.32	1.06	74
Eisen 2012	other psychological interventions	1.64	0.97	82
		Psychotic symptoms		
Eisen 2012	peer support	0.58	0.87	74
Eisen 2012	other psychological interventions	0.84	0.96	82
		Interpersonal relationship		
Eisen 2012	peer support	1.28	0.76	74
Eisen 2012	other psychological interventions	1.5	0.82	82
		Depression		
Eisen 2012	peer support	1.3	0.9	74
Eisen 2012	other psychological interventions	1.38	0.95	82
		Psychotic symptoms		
Eisen 2012	peer support	0.58	0.87	74
Eisen 2012	standard care	0.84	0.96	82

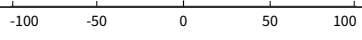
Analysis 2.5. Comparison 2 Peer support plus standard care versus clinician-led support plus standard care, Outcome 5 Behaviour: 1. Specific: drug/alcohol use – mean endpoint score (BASIS subscale, high = poor) – medium term (skewed data).

Behaviour: 1. Specific: drug/alcohol use – mean endpoint score (BASIS subscale, high = poor) – medium term (skewed data)

Study	Intervention	Mean	SD	N
		Alcohol/drug use		
Eisen 2012	peer support	0.51	0.62	74
Eisen 2012	other psychological interventions	0.70	0.89	82

Analysis 2.6. Comparison 2 Peer support plus standard care versus clinician-led support plus standard care, Outcome 6 Peer outcomes: 1. Impact on the service user and peer supporter: social support – mean endpoint score (MOSSSS, high = good) – medium term.

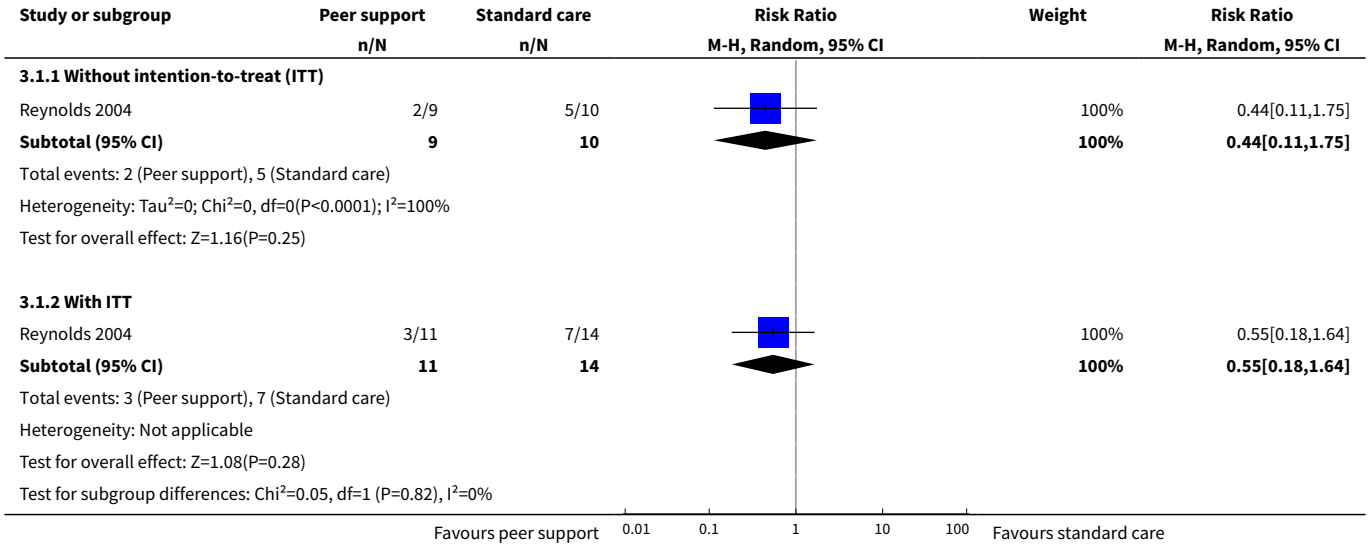
Study or subgroup	Peer support		Other psychological interventions		Mean Difference	Weight	Mean Difference
	N	Mean(SD)	N	Mean(SD)	Random, 95% CI		Random, 95% CI
Eisen 2012	74	63.2 (16.6)	82	58.2 (19)		100%	4.97[-0.62,10.56]
Total ***	**74**		**82**			**100%**	**4.97[-0.62,10.56]**
Heterogeneity: Not applicable							
Test for overall effect: Z=1.74(P=0.08)							



Comparison 3. Sensitivity analysis (assumptions for lost binary data): peer support + standard care versus standard care

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Service use: 1. Hospital admission – medium term	1		Risk Ratio (M-H, Random, 95% CI)	Subtotals only
1.1 Without intention-to-treat (ITT)	1	19	Risk Ratio (M-H, Random, 95% CI)	0.44 [0.11, 1.75]
1.2 With ITT	1	25	Risk Ratio (M-H, Random, 95% CI)	0.55 [0.18, 1.64]

Analysis 3.1. Comparison 3 Sensitivity analysis (assumptions for lost binary data): peer support + standard care versus standard care, Outcome 1 Service use: 1. Hospital admission – medium term.



ADDITIONAL TABLES

Table 1. Details of peer-support intervention in each included study

Study ID	Peer-support intervention		
	Treatment duration	Who delivered/led the intervention	Element of peer support
Castelein 2008	8 months	People with schizophrenia or related psychotic disorder	Guided peer support group; participants decided the topic of each session; each session had the same structure discussing daily life experiences in pairs; it is to provide peer-to-peer interaction.
Cook 2012a	8 weeks	Peer instructors	Peer-led, mental illness education intervention called Building Recovery of Individual Dreams and Goals through Education

Table 1. Details of peer-support intervention in each included study *(Continued)*

			and Support (BRIDGES). Classes were delivered interactive, and included group discussion, illustrative anecdotes and structured exercises designed to apply information to everyday situations. Course topics included recovery principles and stages, strategies for building interpersonal and community support systems, brain biology and psychiatric medications, diagnoses and related symptom complexes, traditional and non-traditional treatments and relapse prevention and coping skills.
Cook 2012b	8 weeks	Peer instructors	Peer-led illness self-management intervention called Wellness Recovery Action Planning (WRAP). Course work included lectures, group discussions, personal examples from the lives of the educators and participants, individual and group exercises, and voluntary homework assignments. Session 1: introduction of key concepts of WRAP; session 2 and 3: development of personalised wellness strategies; session 4: introduction of a daily maintenance plan to use every day to stay emotionally and physically healthy; session 5: educating of early warning signs; session 6 and 7: creation of a crisis plan specifying signs of impending crisis, names of individuals willing to help, and types of assistance preferred; session 8: post crisis support.
Druss 2010	6 sessions	Peer specialists	6 group sessions led by peer specialists, the following topics were discussed: overview of self-management; exercise and physical activity; pain and fatigue management; healthy eating on a limited budget; medication management; finding and working with a regular doctor.
Eisen 2012	12 weeks	Peer facilitators	Peer facilitators used written recovery material such as the Spanior Recovery Workbook available from the Boston University. Peer leaders also shared their personal experiences as veterans with mental illness.
Van Gestel-Timmermans 2012			
Goldberg 2013	13 weeks	People with mental illness	Living well group; the first 3 sessions of the living well intervention focus on the basic strategies of self-management; the remaining weekly sessions focus on training in specific disease management techniques and skills.
Kelly 2014	6 months	People with mental illness	Manualised intervention. Navigators encouraged development of self-management of healthcare through a series of psychoeducation and behavioural strategies.
Mahlke 2017	6 months	People with mental illness	1-to-1 peer support in addition to standard care. Peer supporters contacted patients within the first week after randomisation and then established 1-to-1 meetings. The minimum number of meetings required to build a supporting relationship and be effective for the patient, based on the experiences in delivering support by the peers themselves.
Qian 2015	5 weeks	People with mental illness	Peer support and psychoeducation.
Reynolds 2004	5 months	People with mental illness	The transitional discharge model; this peer support provided friendship, understanding and encouragement for the discharged patient.

Table 1. Details of peer-support intervention in each included study *(Continued)*

Rowe 2007	4 months	People with mental illness	Citizenship intervention plus valued-roles projects. Consist of classes with topics related to social participation and community integration (citizenship classes), followed by projects designed to foster participants' acquisition of valued social roles (valued-roles projects).
Sells 2008	12 months	Peer providers	Peer-based group; use past experiences with recovery as a tool for understanding, role modelling and hope building for others.
Van Gestel-Timmermans 2012	12 weeks	People with mental illness	Each session had the same structure and was organised around a specific, recovery-related theme, such as the meaning of recovery to participants, personal experiences of recovery, personal desires for the future, making choices, goal setting, participation in society, roles in daily life, personal values, how to get social support, abilities and personal resources, and empowerment and assertiveness.

APPENDICES

Appendix 1. Previous search terms

1. Cochrane Schizophrenia Group's Trials Register

The Trials Search Co-ordinator searched the Cochrane Schizophrenia Group's Trials Register applying the following search strategy based on the terms recommended by Doull 2005:

peer*:ti or "self help":ti.or (social NEXT (support* or network* advis* or advice* or counsel*)):ti or peer*:ab or "self help":ab or (social NEXT (support* or network* advis* or advice* or counsel*)): ab

The Cochrane Schizophrenia Group's Trials Register was compiled by systematic searches of major databases and their monthly updates, handsearches and conference proceedings (see the Cochrane Schizophrenia Group Module).

WHAT'S NEW

Date	Event	Description
28 May 2019	Amended	Due to copyright issues and requests for payment to reproduce previously published data we have removed, at the request of the authors of this scale, all information connected with the Morisky Medication Adherence Scales. The removal of these data does not materially affect the results of the review as the relevant data from this scale were skewed and were presented as 'other data'.

CONTRIBUTIONS OF AUTHORS

WTC: project initiation, primary review author, protocol and review writing, abstract inspection, full report inspection, data extraction.

AC: primary review author, helped with writing the protocol, checking screening results, advice on report writing.

SZ: revised the format of all tables, helped write the review.

SL: primary review author, helped with writing the protocol, screening search results, full report inspection and review, data extraction.

DECLARATIONS OF INTEREST

WTC: none.

SL: none.

AC: none.

SZ: none.

SOURCES OF SUPPORT

Internal sources

- University of Huddersfield, UK.

  Employs review author Steve Lui
- The Nethersole School of Nursing, The Chinese University of Hong Kong, China.

  Employs review author Wai Tong Chien
- De Montfort University, Leicester, UK.

  Employs review author Andrew Clifton

External sources

- No sources of support supplied

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

Objectives

We reworded the objectives to clarify that the comparator interventions were interventions not delivered by peers.

Previous objective text: To assess the effects of peer-support interventions for people with schizophrenia or schizophrenia-like disorders in the community, compared to standard care and other psychosocial interventions.

Inclusion criteria

In the protocol, we stated that majority of participants should be within the adult age range and be diagnosed with schizophrenia, schizophrenia-like disorders, bipolar disorder or serious affective disorders, preferably as defined by National Institute of Mental Health (NIMH) criteria (NIMH 1987). Moreover, we indicated that if a trial included participants with a range of serious mental illnesses we would have included it only if the majority had schizophrenia.

In the review, we decided to change the inclusion criteria to reflect the circumstances of clinical practice which means peer support is usually delivered to populations with mixed diagnosis and consequently this reflects what researchers have been trialling thus far. We included studies with schizophrenia or schizophrenia-like disorders at least 20% of the participants. Where a paper did not report the proportion of various diagnoses, we included such paper but conducted sensitivity analysis to test whether the paper influences the pooled results. Besides, we also changed our objectives to keep consistent with our inclusion criteria.

Outcomes

We also make some amendment on the of outcomes that planned to be included in the 'Summary of findings' table in our protocol. We added relapse to the 'Summary of findings' table as it is a primary outcomes in our protocol, therefore should also be one main outcome in the 'Summary of findings' table. We also changed "adverse events – suicide or all-cause mortality" to "adverse events – all cause", and added in 'sub-groups' of outcomes to the peer outcomes: quality of life and satisfaction with care for service user and peer supporter in line with standard Cocharane Schizophrenia's template outcomes..

INDEX TERMS

Medical Subject Headings (MeSH)

*Peer Group; *Schizophrenic Psychology; *Social Support; Quality of Life; Recurrence; Schizophrenia [*therapy]

MeSH check words

Humans