Cochrane Library

Cochrane Database of Systematic Reviews

Collaborative care approaches for people with severe mental illness (Review)

Reilly S, Hobson-Merrett C, Gibbons B, Jones B, Richards D, Plappert H, Gibson J, Green M, Gask L, Huxley PJ, Druss BG, Planner CL

Reilly S, Hobson-Merrett C, Gibbons B, Jones B, Richards D, Plappert H, Gibson J, Green M, Gask L, Huxley PJ, Druss BG, Planner CL.
Collaborative care approaches for people with severe mental illness.
Cochrane Database of Systematic Reviews 2024, Issue 5. Art. No.: CD009531.
DOI: 10.1002/14651858.CD009531.pub3.

www.cochranelibrary.com

TABLE OF CONTENTS

[Intervention Review]

Collaborative care approaches for people with severe mental illness

Siobhan Reilly[1,2,3], Charley Hobson-Merrett[4,5], Bliss Gibbons[6], Ben Jones[7], Debra Richards[4], Humera Plappert[8], John Gibson[9], Maria Green[10], Linda Gask[11], Peter J Huxley[12], Benjamin G Druss[13], Claire L Planner[14]

[1]Centre for Applied Dementia Studies, Faculty of Health Studies, University of Bradford, Bradford, UK. [2]Wolfson Centre for Applied Health Research, Bradford, UK. [3]Division of Health Research, Lancaster University, Lancaster, UK. [4]Primary Care Plymouth, University of Plymouth, Plymouth, UK. [5]National Institute for Health Research Applied Research Collaboration South West Peninsula, Plymouth, UK. [6]PARTNERS2, University of Birmingham, Birmingham, UK. [7]College of Medicine and Health, University of Exeter, Exeter, UK. [8]Primary Care Clinical Sciences, University of Birmingham, Birmingham, UK. [9]The McPin Foundation, London, UK. [10]Pennine Health Care NHS Foundation Trust, Bury, UK. [11]Health Sciences Research Group, Manchester Academic Health Science Centre, University of Manchester, Manchester, UK. [12]Centre for Mental Health and Society, School of Health Sciences, Bangor University, Bangor, UK. [13]Department of Health Policy and Management, Emory University, Atlanta, USA. [14]Centre for Primary Care and Health Services Research, University of Manchester, Manchester, UK

Contact: Siobhan Reilly, s.reilly@bradford.ac.uk.

Editorial group: Cochrane Schizophrenia Group.
Publication status and date: New search for studies and content updated (no change to conclusions), published in Issue 5, 2024.

Citation: Reilly S, Hobson-Merrett C, Gibbons B, Jones B, Richards D, Plappert H, Gibson J, Green M, Gask L, Huxley PJ, Druss BG, Planner CL. Collaborative care approaches for people with severe mental illness. *Cochrane Database of Systematic Reviews* 2024, Issue 5. Art. No.: CD009531. DOI: 10.1002/14651858.CD009531.pub3.

ABSTRACT

Background

Collaborative care for severe mental illness (SMI) is a community-based intervention that promotes interdisciplinary working across primary and secondary care. Collaborative care interventions aim to improve the physical and/or mental health care of individuals with SMI. This is an update of a 2013 Cochrane review, based on new searches of the literature, which includes an additional seven studies.

Objectives

To assess the effectiveness of collaborative care approaches in comparison with standard care (or other non-collaborative care interventions) for people with diagnoses of SMI who are living in the community.

Search methods

We searched the Cochrane Schizophrenia Study-Based Register of Trials (10 February 2021). We searched the Cochrane Common Mental Disorders (CCMD) controlled trials register (all available years to 6 June 2016). Subsequent searches on Ovid MEDLINE, Embase and PsycINFO together with the Cochrane Central Register of Controlled Trials (with an overlap) were run on 17 December 2021.

Selection criteria

Randomised controlled trials (RCTs) where interventions described as 'collaborative care' were compared with 'standard care' for adults (18+ years) living in the community with a diagnosis of SMI. SMI was defined as schizophrenia, other types of schizophrenia-like psychosis or bipolar affective disorder. The primary outcomes of interest were: quality of life, mental state and psychiatric admissions at 12 months follow-up.

Data collection and analysis

Pairs of authors independently extracted data. We assessed the quality and certainty of the evidence using RoB 2 (for the primary outcomes) and GRADE. We compared treatment effects between collaborative care and standard care. We divided outcomes into short-term (up to six months), medium-term (seven to 12 months) and long-term (over 12 months).

For dichotomous data we calculated the risk ratio (RR) and for continuous data we calculated the standardised mean difference (SMD), with 95% confidence intervals (CIs). We used random-effects meta-analyses due to substantial levels of heterogeneity across trials. We created a summary of findings table using GRADEpro.

Main results

Eight RCTs (1165 participants) are included in this review. Two met the criteria for type A collaborative care (intervention comprised of the four core components). The remaining six met the criteria for type B (described as collaborative care by the trialists, but not comprised of the four core components). The composition and purpose of the interventions varied across studies. For most outcomes there was low- or very low-certainty evidence.

We found three studies that assessed the quality of life of participants at 12 months. Quality of life was measured using the SF-12 and the WHOQOL-BREF and the mean endpoint mental health component scores were reported at 12 months. Very low-certainty evidence did not show a difference in quality of life (mental health domain) between collaborative care and standard care in the medium term (at 12 months) (SMD 0.03, 95% CI -0.26 to 0.32; 3 RCTs, 227 participants). Very low-certainty evidence did not show a difference in quality of life (physical health domain) between collaborative care and standard care in the medium term (at 12 months) (SMD 0.08, 95% CI -0.18 to 0.33; 3 RCTs, 237 participants).

Furthermore, in the medium term (at 12 months) low-certainty evidence did not show a difference between collaborative care and standard care in mental state (binary) (RR 0.99, 95% CI 0.77 to 1.28; 1 RCT, 253 participants) or in the risk of being admitted to a psychiatric hospital at 12 months (RR 5.15, 95% CI 0.67 to 39.57; 1 RCT, 253 participants).

One study indicated an improvement in disability (proxy for social functioning) at 12 months in the collaborative care arm compared to usual care (RR 1.38, 95% CI 0.97 to 1.95; 1 RCT, 253 participants); we deemed this low-certainty evidence.

Personal recovery and satisfaction/experience of care outcomes were not reported in any of the included studies. The data from one study indicated that the collaborative care treatment was more expensive than standard care (mean difference (MD) international dollars (Int$) 493.00, 95% CI 345.41 to 640.59) in the short term. Another study found the collaborative care intervention to be slightly less expensive at three years.

Authors' conclusions

This review does not provide evidence to indicate that collaborative care is more effective than standard care in the medium term (at 12 months) in relation to our primary outcomes (quality of life, mental state and psychiatric admissions). The evidence would be improved by better reporting, higher-quality RCTs and the assessment of underlying mechanisms of collaborative care. We advise caution in utilising the information in this review to assess the effectiveness of collaborative care.

PLAIN LANGUAGE SUMMARY

Collaborative care approaches for people with severe mental illness

Key messages

This review does not provide evidence to indicate that collaborative care is more effective than standard care in the medium term (at 12 months) in relation to quality of life, mental state and psychiatric admissions.

No differences were shown in quality of life, mental state or admissions to a psychiatric hospital at 12 months. One study showed an improvement in disability at 12 months. Disability was used as an indirect measure of how well people function in their lives, in terms of their social roles and activities.

Most of the studies included did not meet a strict definition of collaborative care (what we called type A collaborative care) and there were large variations in the interventions delivered. Furthermore, the majority of evidence was either low- or very low-certainty.

What is severe mental illness?

Severe mental illness (SMI) refers to people with psychological problems that can be challenging to a level that impacts on their ability to engage in everyday activities. Schizophrenia, bipolar disorder and non-organic psychosis are all examples of SMIs.

What did we want to find out?

The aim of this review was to assess the effectiveness of collaborative care in comparison to standard or usual care.

What is collaborative care?

Collaborative care aims to improve both the physical and mental health of people living with long-term conditions. All definitions agree that it seeks to develop closer working relationships and better communication between primary care (general practitioners (GPs) and practice nurses) and specialist health care (such as Community Mental Health Teams, including psychiatrists and psychologists). There are different ways in which this can be achieved, making collaborative care complex. Greater joined-up working between services is expected to provide someone with a severe mental illness (SMI) with better care, based in the community, which is often a less stigmatised and stigmatising setting than hospital. It is also important because about 31% of people with SMI living in the UK are seen only in a primary care setting.

What did we do?

Electronic databases were searched in 2020 and 2021 for trials of collaborative care. The primary outcomes of interest were quality of life, mental health and admissions to hospital. We included eight studies in this review. This is an update of the original review published in 2013, which included only one study. This version is based on new searches of the literature that identified an additional seven studies.

What did we find?

No differences were shown in quality of life, mental state or admissions to a psychiatric hospital at 12 months. One study showed an improvement in disability at 12 months. Disability was used as an indirect measure of how well people function in their lives, in terms of their social roles and activities.

Although personal recovery and experience of care/satisfaction were outcomes that those with ongoing mental health problems highlighted as important, none of the included studies measured these.

What are the limitations of the evidence?

Our confidence in these findings is limited due to concerns about the certainty of the evidence. Most of the studies included did not meet a strict definition of collaborative care (what we called type A collaborative care) and there were large variations in the interventions delivered. Furthermore, the majority of evidence was either low- or very low-certainty. Further research is needed to determine whether collaborative care is good for people with a diagnosis of severe mental illness in terms of clinical outcomes or helping people feel better, as well as its cost-effectiveness. Further high-quality RCTs with a clear focus on assessing outcomes directly related to collaborative care are needed in this area, which may also benefit from mixed-methods and qualitative research to understand how collaborative care can best be delivered. None of the studies measured adverse effects of collaborative care.

The original plain language summary was written by Ben Gray and adapted by John Gibson for the updated review. Both are service user researchers.

SUMMARY OF FINDINGS

Summary of findings 1. Summary of findings table - Collaborative care compared to usual care for severe mental illness

Collaborative care compared to usual care for severe mental illness

Patient or population: severe mental illness
Setting: participants living in the community (including in independent living facilities or supported housing)
Intervention: collaborative care
Comparison: usual care

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Certainty of the evidence (GRADE)	Comments
	Risk with usual care	Risk with collaborative care				
Quality of life: average change in mental health component (proxy for binary quality of life) assessed with: SF-12/WHO-QOL-BREF follow-up: 12 months	-	SMD **0.03 SD higher** (0.26 lower to 0.32 higher)	-	227 (3 RCTs)	⊕⊖⊖⊖ Very low[a,b,c]	Very low-certainty evidence did not show a difference between collaborative care and standard care in the mental health component of quality of life at 12 months.
Mental state: clinically important change (binary) assessed with: PANSS follow-up: 12 months	512 per 1000	**507 per 1000** (394 to 655)	**RR 0.99** (0.77 to 1.28)	253 (1 RCT)	⊕⊕⊖⊖ Low[c,d]	Low-certainty evidence did not show a difference between collaborative care and standard care in mental state at 12 months.
Psychiatric hospital admissions assessed with: number of participants admitted to hospital follow-up: 12 months	12 per 1000	**60 per 1000** (8 to 460)	**RR 5.15** (0.67 to 39.57)	253 (1 RCT)	⊕⊕⊖⊖ Low[c,d,e]	Low-certainty evidence did not show a difference between collaborative care and standard care in psychiatric hospital admissions at 12 months.
Quality of life: average change in physical health component (proxy for physical health) assessed with: SF-12/WHO-QOL-BREF follow-up: 12 months	-	SMD **0.08 SD higher** (0.18 lower to 0.33 higher)	-	237 (3 RCTs)	⊕⊖⊖⊖ Very low[a,f]	Very low-certainty evidence did not show a difference between collaborative care and standard care in the physical health component of quality of life (proxy for physical health) at 12 months.

Disability (proxy for social functioning) assessed with: IDEAS follow-up: 12 months	326 per 1000	**449 per 1000** (316 to 635)	**RR 1.38** (0.97 to 1.95)	253 (1 RCT)	⊕⊕⊖⊖ Low[d,g]	Low-certainty evidence showed some evidence of a difference between collaborative care and standard care in disability (proxy for improved social functioning) at 12 months; more participants receiving collaborative care improved.
Personal recovery - not reported	-	-	-	-	-	There is no evidence regarding the effect of collaborative care on personal recovery.
Experience of care/satisfaction - not reported	-	-	-	-	-	There is no evidence regarding the effect of collaborative care on satisfaction/personal experience of care.

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

CI: confidence interval; **RR:** risk ratio; **SMD:** standardised mean difference

GRADE Working Group grades of evidence
High certainty: we are very confident that the true effect lies close to that of the estimate of the effect.
Moderate certainty: we are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low certainty: our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
Very low certainty: we have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

See interactive version of this table: https://gdt.gradepro.org/presentations/#/isof/isof_question_revman_web_430621846341005525.

[a] Downgraded two levels due to risk of bias: in two studies this outcome was rated as some concerns of risk of bias; in one study this outcome was rated as high risk of bias.
[b] There is a heterogeneity of direction of effect across studies reporting this outcome. However, due to overlapping confidence intervals and low I2 we have not downgraded the certainty of evidence in relation to inconsistency.
[c] Downgraded one level for indirectness. The interventions used in this study did not utilise Gunn's elements of collaborative care, although they were described by the study authors as collaborative care interventions.
[d] Downgraded one level for imprecision. This was based on considering a range of scenarios with varying control group rates and target relative risk reductions, the most extreme of which pertains to the assumption of a control group rate of 50% and a target relative risk reduction of 30%, which requires a total of 338 participants in order to achieve 80% power at the 5% significance level. The number of participants required for precision was not met in relation to this outcome.
[e] GRADE Handbook guidance recommends upgrading the certainty of evidence where the risk ratio exceeds 5.0. However, we have chosen not to upgrade the certainty of evidence in relation to this outcome due to the small numbers of observed events (particularly in the control arm) and therefore the substantial uncertainty of the point estimate.
[f] Downgraded one level for indirectness. The intervention used in this study did not utilise Gunn's elements of collaborative care, although it was described by the study authors as a collaborative care intervention. Additionally we are utilising quality of life (physical health domain) as a proxy for physical health.
[g] Downgraded one level for indirectness. The intervention used in this study did not utilise Gunn's elements of collaborative care, although it was described by the study authors as a collaborative care intervention. Additionally we are utilising disability as a proxy for social functioning.

BACKGROUND

Description of the condition

Defining severe mental illness

Severe mental illness (SMI) is an umbrella term commonly used to describe conditions where psychosis is often present, for example schizophrenia, schizophreniform and schizoaffective disorders, bipolar disorder and other types of psychosis.

The prevalence of SMI

In a systematic review of the prevalence of schizophrenia, median estimates were a point prevalence (measured at a particular point in time) of 4.6 per 1000 people, a period prevalence (measured over a specified period of time) of 3.3 and a lifetime prevalence (the proportion of a population that, at some point in their life, has experienced schizophrenia) of 4.0 (Saha 2005). Countries from the developing world had a lower prevalence of schizophrenia.

The World Health Organization World Mental Health Survey Initiative reported an aggregate lifetime prevalence for bipolar type I disorder of 0.6% (Merikangas 2011). In a review of 73 primary studies with data related to the prevalence of psychosis, the pooled median point and 12-month prevalence of psychosis was 3.89 and 4.03 per 1000, respectively, and the median lifetime prevalence was 7.49 per 1000 (Moreno-Küstner 2018).

The National Survey of Psychiatric Morbidity in the UK found a population prevalence of probable psychotic disorder of five per 1000 in the age group 16 to 74 years (Singleton 2001). The prevalence of SMI in England, defined as the number of people on the Quality and Outcomes Framework (QOF) SMI register, for QOF year 2011/12 was 0.8% (QOF 2012), however this is now rising closer to 1% (Whitty 2020). There are large regional variations, with recorded prevalence of 1.6 in some urban and coastal areas (Reilly 2015; Whitty 2020).

Poorer outcomes for people with SMI

People with SMI are among the most socially excluded, subject to the mutually compounding problems of impairment, discrimination, diminished social roles, unemployment and lack of social networks (Social Exclusion Unit 2004). Medical comorbidity is more common in people with SMI compared with the general population (Reilly 2015), and hospital admissions due to physical disease are higher for people with schizophrenia (Bouza 2010). Lifestyle, diet and drug side effects all contribute to poor health outcomes (Connolly 2005), including higher standardised mortality rates (Brown 2000; Brown 2010; Harris 1998; Osby 2001). Indeed, people with SMI die up to 25 years earlier than the general population (Colton 2006; Miller 2006).

Health service provision

Worldwide, spending on mental health is grossly inadequate, with wide gaps between treatments needed and those provided, especially when comparing low-income and high-income countries (Saxena 2006). There is a widespread view that mental health problems in both high-income (Blount 1998) and low-income countries could and should be tackled at the primary care level (Butler 2008; WHO 2009; WHO 2016). Treatment for SMI at the primary care level can help to reduce stigma, improve early detection and treatment, lead to cost efficiency and savings, and partly offset the limitations of mental health resources through the use of community resources. However, only 61% of countries are reported to provide this primary care (WHO 2001). In the UK, people with SMI are in contact with primary care services for a longer cumulative time than people without mental health problems (Kai 2000; Lang 1997). In fact, approximately 31% of people with SMI in the UK are seen only in the primary care setting (Reilly 2012).

Our epidemiological review of 297 randomly selected UK medical records demonstrates a number of relevant findings: (1) the biggest workload associated with this group is borne by secondary care mental health services; (2) there were high variations in care received by people with SMI; (3) when the results of this study are compared with previous evidence, where data have been collected in primary care (Reilly 2012), the information held in primary care hugely underestimates the amount of care received by most of this group and (4) there is a large imbalance in care within this group; those with SMI who are managed only in primary care receive far less intervention than most of those managed in secondary care (Reilly 2021). Furthermore, many general practitioners (GPs) feel that, in contrast to people with complex diabetes or heart failure, for example, holistic care of people with SMI is beyond their remit (Kisely 2007; Lester 2005). GPs regard themselves as involved in the monitoring and treatment of physical illness and prescribing for mental illness (Bindman 2000; Burns 2000; Kendrick 1991), with only a minority regarding themselves as involved in the monitoring and treatment of mental illness (Bindman 2000). This suggests that primary care practitioners and patients would benefit from collaborative secondary/primary mental health care.

Another recent large-scale English retrospective case-control study, using patient records from primary care linked to hospital statistics, showed that increased mortality rates observed in people with SMI may be attributable to underdiagnosis of cardiovascular disease and delays in treatment (Han 2021). There is also evidence that health prevention and promotion activities in primary care are reduced for people with SMI (Daumit 2002; Osborn 2006). Therefore, collaborative mental health care may also improve the poor physical health outcomes in SMI populations.

Policy guidance on care provision

NICE guidance in England recommends that people with an established diagnosis of schizophrenia or bipolar disorder who are managed in primary care require regular assessment of their health and social needs (NICE 2009). This should include monitoring of mental state, medication use, medication adherence, side effects, social isolation, access to services and occupational status. An individual with a diagnosis of schizophrenia or bipolar disorder should have a care plan developed jointly between primary care and secondary mental health services. Regular monitoring of physical health is also essential. With the consent of the service user, non-professional carers should be consulted at regular intervals on the needs of the service user and should also be offered an assessment of their own specific needs (NICE 2009).

Interface working and organisation of mental health care

Given that nearly all collaborative care studies have 'usual care' as the comparator group, it is important to understand what comprises usual care to contextualise the effects of collaborative care in any given study. Variation within and across countries is likely to be an important driver of differences in treatment effects across studies. In a World Health Organization (WHO) report, 42

low- and middle-income countries/territories were involved in data collection, and connections between mental health and other relevant components of the health system, as well as non-health sectors, were weak (WHO 2009). Moreover, there was minimal integration of mental health into primary health care.

Since the 1980s, multidisciplinary community mental health teams (CMHTs) have been the main vehicle for delivering co-ordinated, comprehensive, community-based mental health services in the UK (Kingdon 1989). Variation in team structures and function mean that the evidence base on the effectiveness of CMHTs is largely descriptive and relatively difficult to interpret (Burns 2004). However, CMHTs have been shown to provide better-quality care at both two and four years after referral compared with a traditional psychiatric unit (Gater 1997). Generic CMHT management also appears to be more effective than standard non-team hospital-oriented care for people with SMI, particularly in terms of patients accepting treatment and also in possibly reducing hospital admissions (Malone 2007).

CMHTs have become the backbone of mental health services over the last 30 years. Numbers have increased from 81 in 1987 (Sayce 1991) to 826 in 2006 (Centre for Public Mental Health 2006), and their core roles have been defined by the Department of Health (Department of Health 2002). There are, however, problems with CMHT staff frequently having caseloads that are too high to allow sufficient contact time to work effectively with people with SMI (Sainsbury Centre for Mental Health 1998), and problems with continuity of care across the primary, secondary and social care interface (Crawford 2004; Freeman 2002). NICE guidance for schizophrenia suggests that though “CMHTs remain the mainstay of community mental healthcare, there is surprisingly little evidence to show that they are an effective way of organising services. As such, evidence for or against the effectiveness of CMHTs in the management of schizophrenia is insufficient to make any evidence-based recommendations” (page 38, NICE 2009). NICE guidance for bipolar disorder states “There is little evidence that CMHTs have advantages or disadvantages over other means of organising care for people with bipolar disorder” (page 144, NICE 2009b). So, while there is good evidence to support intensive community services (e.g. intensive case management for severe mental illness; Dieterich 2010), there is less evidence to support large numbers of individuals who need lower-intensity care either being managed by CMHTs or being discharged back to primary care. This fits with the wider context, which indicates that the state of research on the relationship between organisational factors and outcomes of mental health treatments requires strengthening with more studies in this area (Falkenström 2018). In the US, most individuals are managed in public sector systems where psychiatrists prescribe medications, non-MDs such as social workers provide therapy and rehabilitative services, and primary care providers/GPs play a much more limited role. Other countries have varying emphasis on inpatient versus outpatient treatment, the role of PCPs versus psychiatrists and other mental health specialists, the availability of psychotropic medications, access to psychotherapy and rehabilitative treatments, and overall resources available for mental health care. In England, the Community Mental Health Framework for Adults and Older Adults provides an historic opportunity to achieve radical change in the design of community mental health care (NHS England 2019c). This will be by moving away from siloed, hard-to-reach services towards joined-up care and whole population approaches, and establishing a revitalised purpose and identity for community mental health services. It supports the development of Primary Care Networks, Integrated Care Systems (ICSs) and personalised care, including how these developments will help to improve care for people with severe mental illnesses. It is hoped that implementing this framework will break down the current barriers between: (1) mental health and physical health, (2) health, social care, voluntary, community and social enterprise organisations and local communities, and (3) primary and secondary care, to deliver integrated, personalised, place-based and well co-ordinated care.

Aim of review

As outlined above, integrated working and collaborative care may overcome some of the obstacles to optimal care provision for those with SMI diagnoses. Collaborative care for depression has a strong evidence base (Archer 2012; Druss 2005; Bauer 2009; Bower 2006; Craven 2006; Gilbody 2006; Gunn 2006). This review seeks to assess the effectiveness of collaborative care approaches in comparison to standard care for people with SMI who are living in the community.

Description of the intervention

Defining collaborative care

There is no universally agreed definition of collaborative care and variation exists in how it is operationalised. It is noted that "Interventions or organisational models similar to collaborative care are sometimes referred to as integrated care, enhanced care, or care management" (Muntingh 2016). In our original review, Reilly 2013, we reported the six unique definitions of collaborative care, cited in 13 systematic reviews of collaborative care (conducted between 2006 and 2016) for a range of mental health conditions (see Appendix 1). The description of collaborative care reported in Gunn 2006 was the most commonly cited and focuses on four 'core' elements: multi-professional work between a primary care practitioner and at least one other service, a structured management plan in the form of protocols or guidance, scheduled patient follow-ups and enhanced interprofessional communication. In all reviews, collaborative care was described as an intervention that aims to foster closer working relationships between primary care and specialist health care.

Operationalising collaborative care

Collaborative care models are often operationalised by way of a specific role, such as a case manager. In addition to prompting collaboration between services, the case manager role might involve work at the patient level according to a manual or protocol with regular follow-up periods (e.g. providing low-level psychological interventions, proactive follow-up, patient education, promotion of self-management and monitoring of clinical status, side effects and adherence, and shared decision-making with patients). In our original review we noted that even when collaborative care interventions have similar components they can differ in the way these are provided (see Appendix 2 in Reilly 2013). For example, Bauer 2001 and Baker 2019 both describe collaboration between the case manager and the patient to achieve jointly identified goals. However, Bauer 2001 does this via group patient education and Baker 2019 via a one-on-one coaching model.

Collaborative care as a complex intervention

Collaborative care meets the definition of a 'complex intervention'. It includes several interacting components, which may act independently and interdependently and within pre-existing systems for providing health care, and may create a range of possible outcomes (Craig 2008). As such, the 'active ingredient' of the intervention can be difficult to identify (Campbell 2000). For this reason, the Medical Research Council guidance recommends that the design and evaluation of complex interventions includes creation of a good theoretical understanding of how the intervention causes change (Craig 2008).

How the intervention might work

Varying definitions of collaborative care and differences in the goals, provision, complexity of interacting components and outcomes mean that explanations of mechanisms are complex. Each separate intervention might work in its own way to create the outcomes identified as important by the designers. Notwithstanding this, there is some evidence that explores how common components might lead to improved health outcomes.

Collaborative care aims to improve quality of care by ensuring that, at an individual level, both patient and case manager and, at a system level, healthcare providers *work together* to address the needs of the patient, thus improving both physical and/or mental health outcomes depending on the specific aims of the intervention. Most research has focused on integrating mental health and primary care services, to facilitate communication and joint working between health professionals (e.g. GP, psychiatrist, nurse, pharmacist, psychologist), provide the patient with care in a less stigmatised setting, promote evidence-based practice and prevent loss of contact with services. A recent feasibility study suggests that for those with a diagnosis of psychosis this integration will lead to practitioners having a better understanding of patients' needs and how to meet them, which in turn will mean that appropriate support is offered to the patient. Subsequently, this will promote behaviours that support outcomes of personal recovery, and improved mental and physical health (Baker 2019).

Evidence from the collaborative care for depression literature suggests that there may be different mechanisms of action at different levels. At the interface between patient and case manager, the focus may be on better medication management, proactive follow-up and self-management, improving health outcomes and reducing unnecessary use of health resources such as emergency admissions. At the organisational level, actions such as feedback of patient information to the GP and adherence of workforces to specific evidence-based guidelines and protocols may be key (Gask 2010).

Baker 2019 suggests that, for those with a diagnosis of psychosis, protocols that address engagement and retention, sustaining an equitable relationship, coaching, goal setting and regular review are key. These will lead to improved service user trust in the case manager (described here as 'care partner'), increased hope and self-esteem, and improved knowledge of health improvement strategies. Then, in turn, these will result in improved physical and mental health and personal recovery outcomes.

Why it is important to do this review

In view of the significantly higher mortality rate and poorer health outcomes, which are often compounded by problems with current healthcare systems, a systematic review of collaborative care approaches is required to help inform healthcare professionals and policy-makers about the provision of more effective care for people with SMI.

Since the publication of the original Cochrane review of 'Collaborative care approaches for people with severe mental illness' (Reilly 2013), there has been a substantial increase in the number of published and relevant randomised controlled trials (RCTs), as illustrated in this review, along with a refinement in defining collaborative care and working models of collaborative care. In England, the health policy landscape has changed (Mental Health Taskforce 2016): local areas will be supported to redesign and reorganise core community mental health teams to move towards a new place-based, multidisciplinary service across health and social care aligned with primary care networks (NHS England 2019a). It is now expected that all Sustainability and Transformation Partnerships (STPs)/Integrated Care Systems (ICSs) in England will receive funding to develop and begin delivering new models of integrated primary and community care for adults and older adults with severe mental illnesses (NHS England 2019b).

This review will add to the evidence base at this critical juncture in the evolution of commissioning mental health services. Despite English national guidelines recommending collaborative care for serious mental illness (Mental Health Taskforce 2016), it is still not as widely available for people with schizophrenia as it is for people with other disorders (for example, depression and diabetes). We still do not know whether collaborative care can work as an integrated intervention that can improve people's mental health, physical health and quality of life outcomes, and how these various models of collaborative care are implemented.

Patients and carers, whether family members or friends, have long been aware of the impact of severe mental illness on all aspects of the individual's life, encompassing not just their mental health, but also their physical health and overall quality of life, including their social networks and sense of isolation in the wider community. This in turn has a profound knock-on effect upon the lives of those closest to them. Lack of meaningful activities, medication side effects and general lifestyle issues all play a part in reduced quality of life and higher mortality rates. A truly patient-centred approach, focusing on individualised and holistic collaborative care, emphasises greater joined-up working between primary and secondary care, with improved communication between agencies.

OBJECTIVES

To assess the effectiveness of collaborative care in comparison with standard care (or other non-collaborative care interventions) for people with a diagnosis of severe mental illness who are living in the community.

METHODS

Criteria for considering studies for this review

Types of studies

We included all types of randomised controlled trial (RCT), including cluster-RCTs, published or unpublished.

Types of participants

We included trials where over 50% of participants fulfilled the following criteria:

- Age: adults aged 18 years or above.
- Diagnosis: severe mental illness, defined as schizophrenia or other types of schizophrenia-like psychosis (e.g. schizophreniform and schizoaffective disorders), bipolar affective disorder or other types of psychosis as defined by the trialists, irrespective of the diagnostic criteria used. Participants with substance abuse or addictive disorders were eligible for inclusion if there was a dual diagnosis of severe mental illness.
- Setting: living in the community, which could include independent living, living with family or supported housing.

Types of interventions

Experimental intervention: collaborative care

As a way of operationalising the intervention and under the guidance of Cochrane Schizophrenia, we only included interventions described as 'collaborative care' by the authors. We categorised each study as either type A or type B collaborative care (see Appendix 2).

Type A collaborative care interventions

Interventions comprise the four ‘core’ components, as defined by Gunn 2006, and are also described as ‘collaborative care’ by the trialists.

Type B collaborative care interventions

Interventions do not comprise the four ‘core’ components, but are described as 'collaborative care' by the trialists.

Comparator: standard care

We defined standard care as a community or outpatient model of care not described as 'collaborative care' by the trialists. We decided post hoc that if trial authors reported that standard care included additional 'enhancements', and these were minimal and also included as part of standard care elsewhere, we would still consider these to be standard care (see Differences between protocol and review).

Types of outcome measures

We changed the outcomes from those reported in the original review (Differences between protocol and review; Appendix 3). As this review has been funded as part of the Byng 2023 National Institute of Health Research (NIHR) grant, we were able to utilise a core outcome set for use in community-based bipolar trials to guide our choice of outcomes (Retzer 2020). We were also able to utilise an additional stakeholder consultation to select outcomes that were important to those working with and living with SMI diagnoses. This stakeholder consultation was convened to capture the wider psychosis target population in Byng 2023 and the nature of the intervention. Quality of life (QoL) was selected by the research team and Lived Experience Advisory Panels (Plappert 2021) as the most important outcome domain for stakeholders. We added this to our primary outcomes along with mental state and psychiatric hospital admissions. In response to stakeholder feedback, we also added personal recovery as an outcome and we broadened our satisfaction outcome to encompass 'experience of care'. We also included process/delivery outcomes as secondary outcomes. These changes were made before we extracted data from our included studies.

For valid scales please see Data extraction and management.

Where possible, we divided outcomes into short-term (less than six months), medium-term (seven to 12 months) and long-term (over 12 months). We endeavoured to prioritise the report of binary outcomes recording clear and clinically meaningful degrees of change ahead of continuous outcomes (e.g. global impression of much improved, or more than 50% improvement on a rating scale - as defined within the trials). For outcomes such as 'clinically important change', 'any change' and 'relapse', we used the definition used by each of the trials.

Primary outcomes

1.1 Quality of life

- Clinically important change in quality of life (as defined by individual studies) (Y/N, binary outcome) at 12 months

1.2 Mental state

- Clinically important change in mental state (as defined by individual studies) (Y/N, binary outcome) at 12 months

1.3 Psychiatric admissions

- Number of participants admitted to hospital (psychiatric admissions) at 12 months

Secondary outcomes

2.1 Quality of life (time points other than 12 months)

- Clinically important change in quality of life (as defined by individual studies) (Y/N, binary outcome)
- Clinically important change in quality of life (as defined by individual studies) (Y/N, binary outcome)
- Any change in quality of life
- Average endpoint quality of life score
- Average change in quality of life scores
- No clinically important change in specific aspects of quality of life (as defined by individual studies)
- Any change in specific aspects of quality of life
- Average endpoint in specific aspects of quality of life scores
- Average change in specific aspects of quality of life scores

2.2 Mental state

General and specific (including positive and negative symptoms of psychosis, and mood (as defined by individual studies))

- Any change in mental state
- Average endpoint mental state
- Average change in mental state

- No clinically important change in mental state (as defined by individual studies)
- Any change in specific aspects of mental state
- Average endpoint in specific aspects of mental state
- Average change in specific aspects of mental state

2.3 Psychiatric admissions

- Mean number of days in hospital for psychiatric admissions
- Length of time to readmission (psychiatric admissions)

2.4 Other hospital admissions

- Number of participants admitted to hospital (physical health admissions)
- Mean number of days in hospital for physical health admissions
- Length of time to readmission (physical health admissions)

2.5 Personal recovery

- Clinically important change in personal recovery (as defined by individual studies) (Y/N, binary outcome)
- Any change in personal recovery
- Average endpoint personal recovery score
- Average change in personal recovery scores
- No clinically important change in specific aspects of personal recovery (as defined by individual studies)
- Any change in specific aspects of personal recovery
- Average endpoint in specific aspects of personal recovery scores
- Average change in specific aspects of personal recovery scores

2.6 Physical health status (including specific measures of blood pressure, blood cholesterol, blood glucose - HbA1c, body mass index (BMI))

- Clinically important change in physical health status (as defined by individual studies)
- Any change in physical health status score
- Average endpoint physical health status score
- Average change in physical health status score

2.7 Global state

- Relapse (as defined by individual studies)
- Time to relapse
- Clinically important change in global state (as defined by individual studies)
- Any change in global state
- Average endpoint global state score
- Average change in global state score

2.8 to 2.9 Medication adherence

- Clinically important change in compliance (patient-reported)
- Any change in compliance (patient-reported)
- Clinically important change in compliance (carer-reported)
- Any change in compliance (carer-reported)

2.10 to 2.11 Social functioning

- Clinically important change in social functioning (as defined by individual studies)
- Any change in social functioning
- Average endpoint social functioning score
- Average change in social functioning scores
- Employment status
- Living tenure (number of participants homeless, in unstable housing or living independently)

2.12 Substance use (alcohol/illicit drugs/cigarettes/tobacco)

- Clinically important change in substance use (as defined by individual studies)
- Any change in substance use
- Average endpoint substance use
- Average change in substance use

2.13 Adverse effect/event(s)

- At least one adverse effect
- Incidence of specific effect (e.g. cardiovascular, metabolic, movement disorders)

2.14 Death

- Number of participants who died from suicide
- Number of participants who died from natural causes

2.15 Service use outside of mental health (i.e. primary care, emergency services, walk-in centres, social services)

- Mean number of contacts per month
- Number of participants in contact with service
- Mean number of service hours per month

2.16 to 2.17 Cost of treatment

- Direct cost of inpatient care
- Direct cost of health and social care (including the above, plus the costs of all other medical and psychiatric care, such as outpatient care and specialist service, collaborative care and community-based social services)
- Total costs, including types of costs above, plus the costs of accommodation and minus benefits, such as earnings where these are known

2.18 Experience of care/satisfaction (participant/carer/staff)

- Clinically important change in participant, carer and staff satisfaction (as defined by individual studies)
- Any change in participant, carer and staff satisfaction
- Average endpoint participant, carer and staff satisfaction score
- Average change in participant, carer and staff satisfaction score

2.19 Leaving the study early (attrition)

- For any reason
- For a specific reason

Process/delivery outcomes

- Components of collaborative care delivered
- Measures of interprofessional collaboration
- Measures of adherence to manual/algorithms/guidance
- Measures of change in management (number of contacts, referral rates, prescribing patterns and appropriateness)
- Measures of change in other health services provided

- Measures of continuity (relational, information, longitudinal)
- Measures of health care professional behaviour and knowledge (improvement in knowledge/skills, attitudes/acceptability, retention rates, absenteeism, healthcare professionals time, prescribing and management of risk factors)
- Mean percentage of case management contacts
- Mean percentage of intervention (delivered as part of collaborative care) contacts
- Mean percentage of session topics covered in training/education

Search methods for identification of studies

Electronic searches

1. Cochrane Schizophrenia Study-Based Register of Trials

On 10 February 2021, the Information Specialist searched the register using the following search strategy:

Collaborat 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 (Roberts 2021; Shokraneh 2017; Shokraneh 2021). This allows rapid and accurate searches that reduce waste in the next steps of systematic reviewing (Shokraneh 2019).

Following the methods from Cochrane (Lefebvre 2019), this register is compiled by systematic searches of major resources (the Cochrane Central Register of Controlled Trials (CENTRAL), CINAHL, ClinicalTrials.gov, Embase, ISRCTN, MEDLINE, PsycINFO, PubMed, WHO ICTRP) and their monthly updates, ProQuest Dissertations and Theses A&I and its quarterly update, handsearches, grey literature and conference proceedings (Shokraneh 2020; see Group's website). There are no language, date, document type or publication status limitations for the inclusion of records in the register.

For previous searches, please see Appendix 4.

2. Cochrane Common Mental Disorders Controlled Trials Register (CCMDCTR)

Cochrane Common Mental Disorders (CCMD) maintained a similar register of controlled trials until June 2016 (CCMDCTR). An Information Specialist with the Group searched the CCMDCTR for collaborative care studies in participants with bipolar disorder (all available years to 6 June 2016) using the following search terms:

(collab and (bipolar or mania* or manic* or hypomani* or psychos* or psychotic or postpsychotic or post-psychotic or "rapid cycling" or schizoaffective on "mixed episode"))* [all fields]

To accommodate the period when the register was out-of-date, the Information Specialist ran complementary searches on Ovid MEDLINE, Embase and PsycINFO together with CENTRAL (with an overlap) from 2014 to 6 June 2020 and a second search on 17 December 2021.

A detailed description of the CCMDCTR and the complementary database search strategies are displayed in Appendix 5.

Searching other resources

Reference searching

We checked the references of all included studies for further relevant studies. We also completed a forward citation search using Google Scholar.

Data collection and analysis

Selection of studies

1. Title/abstract screening

Pairs of authors (CP, MC, CM, SR, CHM) independently reviewed the retrieved titles and abstracts, applying the eligibility criteria. Decisions to include or exclude were recorded on an Excel spreadsheet and are summarised in Figure 1.

Figure 1. Study flow diagram

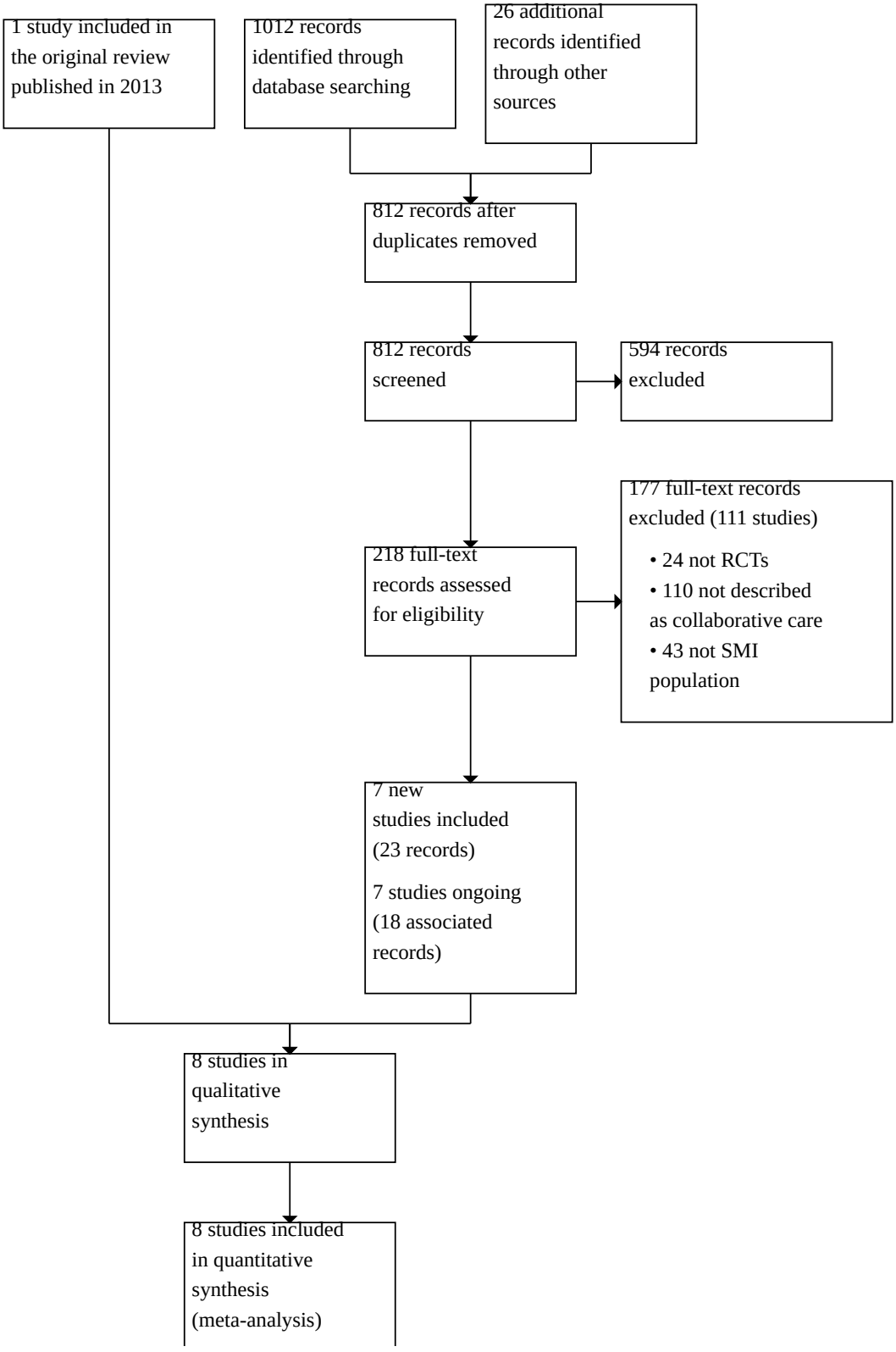
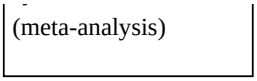


Figure 1. (Continued)



2. Full-text screening

Pairs of authors (CP, MC, CM, SR, CHM) then independently reviewed the full-text articles for studies included at the title/abstract screening stage. We maintained a log of all studies that were excluded upon review of the full text, and recorded the reason for exclusion in the Characteristics of excluded studies table.

Data extraction and management

1. Extraction

Review authors (CHM, BG, CP, MC, BJ and SR) independently extracted outcome and implementation data from the eight included studies in duplicate. We extracted the descriptions of the interventions in relation to the four 'core' components of collaborative care (see summary, Appendix 2) and constituent components identified in the original review (Reilly 2013). One author (CP) extracted descriptive information regarding the interventions being tested and fidelity assessment into TiDIER checklists (Hoffman 2014), when not published by study trialists.

2. Management

We extracted data onto a paper form and Excel spreadsheets, which we then entered into Review Manager 5 (RevMan 5).

3. Scale-derived data

We included continuous data from rating scales only if: (a) the psychometric properties of the measuring instrument have been described in a peer-reviewed journal (Marshall 2000); (b) the measuring instrument was not written or modified by one of the trialists; (c) the measuring instrument was either (i) a self-report or (ii) completed by an independent rater or relative (not the therapist).

Assessment of risk of bias in included studies

We assessed the risk of bias using the risk of bias 2.0 assessment tool for randomised trials (RoB 2) (Sterne 2019), for all trials randomised at participant level. We assessed outcomes in the van der Voort 2015 study using the RoB 2 assessment tool for cluster-randomised trials (Sterne 2019). RoB 2 assesses the risk of bias in each trial outcome independently. To balance rigour against the burden of assessment, we assessed the risk of bias only for the review's primary outcomes (quality of life, mental state, psychiatric hospital admissions) and other outcomes reported in our summary of findings table (personal recovery, experience of care/satisfaction, social functioning, physical health). This is consistent with the *Cochrane Handbook for Systematic Reviews of Interventions* Chapters 7 and 8 (Boutron 2021; Higgins 2021).

All risk of bias assessments were performed in duplicate, once by CHM and once by one of BG, DR, PH, CP and BD. Where disputes arose these were discussed and resolved by the review author team.

We assessed risk of bias from an intention-to-treat perspective in the following domains: bias arising from the randomisation process, bias due to deviations from the intended intervention, bias due to missing outcome data, bias in measurement of the outcome and bias in selection of the reported result. In addition to these domains, we assessed outcomes in the van der Voort 2015 study for risk of bias arising from the timing of identification and recruitment of participants. In each risk of bias assessment, we gave a risk of bias rating of 'low', 'some concerns' or 'high' for each individual domain and overall. An overall rating of 'low' risk of bias was only given if all domains were rated 'low'. One or more domain rated as 'some concerns' resulted in a 'some concerns' risk of bias rating overall. One or more domain rated as 'high' resulted in a 'high' risk of bias rating overall.

The impact of the risk of bias assessment on the strength of the evidence presented in this review is considered in the Discussion section of this review.

The impact of the risk of bias and other quality concerns in assessing the certainty and weight of the evidence presented in this review is discussed in the review and summarised in Summary of findings 1.

Measures of treatment effect

1. Dichotomous data

Where binary outcomes (proportions) were reported, we calculated a risk ratio (RR) using a random-effects model (Furukawa 2002), with 95% confidence intervals (CIs) for each outcome. We chose the RR over the odds ratio because the latter tends to overstate effect size when event rates are high (Sterne 2011).

2. Continuous data

2.1 Summary statistic

For continuous outcomes, we used a random-effects model to estimate standardised mean differences (SMDs) between groups. We would have preferred not to calculate SMDs, but found that studies used different measurement tools and so it was necessary to do so in order to synthesise the results.

2.2 Endpoint versus change data

Since there is no principal statistical reason why endpoint and change data should measure different effects (Sterne 2011), we used scale endpoint data as it is easier to interpret from a clinical point of view. If endpoint data had not been available, we would have used change scores.

2.3 Skewed data

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

- standard deviations (SDs) and means are reported in the paper or obtainable from the authors;
- when a scale starts from the finite number zero, the SD, when multiplied by two, is less than the mean (as otherwise the mean is unlikely to be an appropriate measure of the centre of the distribution) (Altman 1996);
- if a scale starts from a positive value, the calculation described above is modified to take the scale starting point into account. In these cases, skew is present if 2 SD > (S-S min), where S is the mean score and S min is the minimum score.

Endpoint scores on scales often have a finite start and endpoint and these rules can be applied to such values. We would have entered skewed endpoint data (from small studies of fewer than 30 participants per arm) into additional tables rather than into an analysis. Skewed data pose less of a problem if the sample size is large and, if present, we planned to enter skewed endpoint data from large trials into syntheses. When continuous data are presented on a scale that includes negative values (such as change data), it is difficult to tell whether data are skewed or not and so change data are entered into analysis.

2.4 Data synthesis

If SDs were not reported, we first tried to obtain the missing data from the study authors. If these were unavailable, Chapters 7.7.3 and 16.1.3 of the *Cochrane Handbook for Systematic Reviews of Interventions* present detailed formulae for estimating SDs from P values, t or F values, CIs, ranges or other statistics (e.g. SDs could have been calculated from standard errors (SEs) using the relationship SD = SE * square root (n)) (Higgins 2011). If these formulae were not applicable, we would have calculated the SDs according to a validated imputation method, which was based on the SDs of the other included studies (Furukawa 2002). Although some of these imputation strategies can introduce error, the alternative would be to exclude a given study's outcome(s) and thus to lose information. Nevertheless, had we identified more relevant studies, we would have examined the validity of the imputations in a sensitivity analysis (excluding imputed values).

2.5 Common measure

To facilitate comparison between trials, had multiple applicable outcomes been collected, we would have converted variables that could be reported in different metrics, such as days in hospital (mean days per year, per week or per month), to a common metric (e.g. mean days per month).

Unit of analysis issues

For repeated observations on participants in long-term studies, we assessed outcomes at different time points using separate analyses. Where possible, we presented results for several periods of follow-up (e.g. at one year and two years). We defined several different outcomes, based on different periods of follow-up, and performed separate analyses. For example, we defined time frames to reflect short-term (up to six months), medium-term (seven to 12 months) and long-term (over 12 months).

1. Cluster-randomised trials

We included one study in the review that employed cluster-randomisation. Studies increasingly employ cluster-randomisation (such as randomisation by clinician or GP practice) but the analysis and pooling of clustered data poses problems. Firstly, authors often fail to account for intracluster correlation in clustered studies, leading to a 'unit of analysis' error (Divine 1992), whereby P values are unduly low and CIs unduly narrow, increasing the risk of spurious conclusions of statistically significant efficacy or effectiveness. This causes inflated type I errors (Bland 1997; Gulliford 1999).

In order to account for the clustering inherent in data from cluster trials, the sample sizes were reduced according to the design effect to obtain effective sample sizes (ESS) as recommended in section 23.1.3 of the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2020). Specifically, for each relevant outcome at each time point, the ESS was calculated by dividing the sample size by the design effect (DE), where DE = 1 + (m – 1) * ICC, m is the cluster size and the ICC is the intracluster correlation coefficient, a measure of the degree of clustering.

In order to obtain the ESS, we had to obtain estimates of the ICC and the cluster size, m. In the study in which cluster-randomisation was employed, the ICCs were not reported. We first contacted the author and requested the ICCs for each of the included outcomes, which they were unable to provide. We therefore assumed an ICC of 0.05 for each outcome, which is reasonable in a primary care setting (Adams 2004). For each outcome at each time point, we calculated m by dividing the sample size by the number of clusters.

2. Studies with multiple treatment groups

We did not include any studies in the review with multiple treatment groups. If we had found a study that involved more than two treatment arms, we would have presented the additional treatment arms in comparisons. If data were binary, we would have simply added these and combined them within the two-by-two table. If data were continuous, we would have combined the data following the formula in section 7.7.3.8 (Combining groups) of the *Cochrane Handbook for Systematic Reviews of Interventions*. If the additional treatment arms were not relevant, we would not have reported these data (Higgins 2011).

Dealing with missing data

We contacted all trial authors of included studies to request additional information/data. Bauer 2006, Chatterjee 2011, Chwastiak 2018, Kilbourne 2012, Kilbourne 2013 and van der Voort 2015 replied and were able to provide additional information. Nevertheless, some authors could not provide all data required. Salman 2014 initially responded, however the lack of further correspondence resulted in all queries being unanswered. Mishra 2017 could not be contacted. We acknowledge that the lack of correspondence from some trial authors may be due to the demands of the COVID-19 pandemic. We documented all correspondence with trial authors.

For continuous outcomes in which SDs were not reported, and no information was available from the authors, we calculated the SDs using the SE of the mean (SEM). We have described the amount and kind of missing data related to participant attrition that was obtained from the study authors in the Characteristics of included studies table. The potential impact of the missing data on the results depends on the extent of missing data, the pooled estimate of the treatment effect and the variability of the outcomes. Variation in the degree of missing data may also be considered as a potential source of heterogeneity. We have also discussed the

impact of the missing data in the Characteristics of included studies table.

1. Overall loss of credibility

At some degree of loss to follow-up, data must lose credibility (Xia 2009). In instances where more than 50% of data is unaccounted for, we would not have reported or analysed the data. If, however, we had found a study with more than 50% of those in one arm that were lost, but the total loss was less than 50%, we would have marked such data with (*) to indicate that such a result may well be prone to bias.

2. Binary

In the case where attrition for a binary outcome was between 0% and 50% and outcomes for these participants were described, we included these data as reported. For these outcomes, the observed rate of the binary outcome for those who stay in the study - in that particular arm of the trial - was used to impute the outcome for those who did not. For primary outcomes, we undertook a sensitivity analysis to test how prone the primary outcomes are to change when data only from people who complete the study to that point are compared to the intention-to-treat (ITT) analysis using the above assumptions to impute missing data. If these data had not been clearly described, we would have presented data on a 'once-randomised-always-analyse' basis, assuming an ITT analysis.

3. Continuous

3.1 Attrition

In the case where attrition for a continuous outcome was between 0% and 50% and completer-only data were reported, we reported these.

3.2 Last observation carried forward

We anticipated that in some studies, in order to do an ITT analysis, the method of last observation carried forward (LOCF) would be employed. As with all methods of imputation to deal with missing data, LOCF introduces uncertainty about the reliability of the results. Therefore, if LOCF data had been used in the analysis, we would have indicated this in the review. Recognising that statistical analysis cannot always reliably compensate for missing data (Unnebrink 2001), we would have assessed the impact of any assumption by testing more than one method in a sensitivity analysis.

3.3 Standard deviations

Where there were missing measures of variance for continuous data but exact SE and CIs were available for group means and either P value or T value were available for differences in mean, we calculated the SD value according to the method described in Section 7.7.3 of the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2008). If SDs were not reported and could not be calculated from available data, we asked the authors to supply the data. Had there been other studies included in the review, and in the absence of data from authors, we would have used the mean SD from other studies.

Assessment of heterogeneity

1. Clinical heterogeneity

We identified issues believed to drive clinical heterogeneity, such as differences in intervention and population, and considered them in the main and sensitivity analyses for the primary outcomes.

2. Statistical

2.1 Visual inspection

Where data were available from more than one study, we inspected forest plots to assess and investigate the possibility of statistical heterogeneity.

2.2 Employing the I^2 statistic

We assessed heterogeneity between studies by considering the I^2 statistic alongside the Chi^2 P value. The I^2 statistic provides an estimate of the percentage of inconsistency thought to be due to chance (Higgins 2003). The importance of the observed value of I^2 depends on: 1) the magnitude and direction of effects and 2) the strength of evidence for heterogeneity (e.g. P value from the Chi^2 test or a CI for I^2). We interpreted an I^2 estimate greater than or equal to 50% accompanied by a statistically significant Chi^2 statistic as potentially indicative of substantial levels of heterogeneity (Deeks 2008), and explored the reasons for the heterogeneity. We also employed this approach in assessing heterogeneity in the GRADE assessment (Schünemann 2020).

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 Systematic Reviews of Interventions* (Sterne 2011). We are aware that funnel plots may be useful in investigating small-study effects but are of limited power to detect such effects when there are few studies. We planned not to use funnel plots for outcomes where there were 10 or fewer studies, or where all studies were of similar sizes. As we only included eight studies, no funnel plots were produced for this review.

Data synthesis

If possible, we would have reported the absolute differences between intervention and control groups for continuous outcomes. However, as we found different outcome measures for the same outcomes, we predominantly reported standardised mean differences. For binary outcomes, we reported relative percent differences in outcomes between the intervention and control groups. Where applicable, we synthesised the results using a random-effects model to provide a pooled estimate of effect from continuous and binary data. Although we could have assessed heterogeneity for each outcome and used a fixed-effect model when this heterogeneity was considered to be small, we opted to use random-effects models regardless, in acknowledgement of the differences in collaborative care interventions, the populations and the clinical settings across the different studies. Analyses were based on the ITT population.

Subgroup analysis and investigation of heterogeneity

For heterogeneous outcomes, we checked the data to ensure that they had been correctly extracted and entered and that there

were no unit of analysis errors. If high levels of heterogeneity are observed, meta-analysis is often not appropriate. If there is considerable variation in results, and particularly if there is inconsistency in the direction of effect, it may be misleading to quote an average value for the intervention effect. Where possible, when substantial heterogeneity was present, we explored possible reasons for this in the context of the following pre-specified characteristics of studies:

- Variations in implementation of the collaborative care intervention and healthcare systems.
- Variations in types of patients included: comparison of studies that dealt solely with people with schizophrenia, other types of schizophrenia-like psychosis, people with bipolar affective disorder or people with other types of psychosis and those that also include people with other severe mental illnesses, e.g. depression and those with a dual diagnosis.
- Quality of study: comparison of high- and low-quality studies.

If exploration of these subgroups offered no clear explanation for the heterogeneity, we would have considered other post hoc subgroups. If other characteristics of the relevant studies were identified (post hoc) as a possible cause of heterogeneity, we would have presented the subgroup analyses alongside relevant discussion.

Sensitivity analysis

We planned the following sensitivity analyses:

- Assumptions for attrition: we performed a sensitivity analysis in order to examine the robustness of the conclusions when including data according to the assumptions that were made regarding people lost to follow-up (Dealing with missing data), where we compared the findings of the primary outcome when we used our assumption compared with completer data only. Both sets of results are reported for completeness alongside appropriate discussion.

Had we found more relevant studies that had reported the required information, we would have performed further sensitivity analyses in order to examine the robustness of the conclusions of the analyses when including studies according to the following criteria:

- Randomisation: we were aiming to include trials in a sensitivity analysis if they were described in some way as to imply randomisation being performed, rather than randomisation being explicitly described. 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, then we would have employed all data from these studies.
- Types of participants: we would have explored whether studies with a higher proportion of people diagnosed with other severe mental illnesses (e.g. depression) differed substantively when compared with studies that solely included people with schizophrenia, other types of schizophrenia-like psychosis, people with bipolar affective disorder or people with other types of psychosis.

Summary of findings and assessment of the certainty of the evidence

We used the GRADE approach to interpret findings (Schünemann 2020), and we used GRADEpro GDT to export data from our review (RevMan) to create a summary of findings table. A summary of findings table provides outcome-specific information concerning the overall certainty of evidence from each included outcome 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 patient care and decision-making. The process of revising the outcomes (described in Types of outcome measures) for this review also enabled us to revise the outcomes included in the summary of findings table (Summary of findings 1). We selected the following main outcomes for inclusion in the summary of findings table:

- Quality of life
- Mental state
- Psychiatric admissions (safety outcome)
- Personal recovery
- Physical health status
- Social functioning
- Experience of care/satisfaction

If data were not available for these pre-specified outcomes but were available for ones that were similar, we presented the closest outcome to the pre-specified one in the summary of findings table, but took this into account when grading the directness of the certainty of evidence.

RESULTS

Description of studies

See Characteristics of included studies; Characteristics of excluded studies; Characteristics of ongoing studies.

Results of the search

Four separate searches of the Cochrane Schizophrenia register were carried out: February 2015 (21 references, 19 different studies), January 2019 (266 references), January 2020 (9 references, 8 studies) and December 2021 (8 references, 8 studies). Four searches were carried out by Cochrane Common Mental Disorders: September 2016 (240 references), March 2019 (357 references, 203 after duplicates removed), June 2020 (86 references, 43 references after duplicates removed) and December 2021 (25 references, 24 after duplicates removed).

We also identified linked articles of interest for included and ongoing studies through searching. A further 26 records were identified through these methods.

After removal of duplicates, we screened 812 articles and obtained 218 full-text papers for a second assessment. These were fully inspected and 177 references (111 studies) were excluded (see flow diagram in Figure 1; Characteristics of excluded studies).

Seven relevant studies are ongoing (see Characteristics of ongoing studies): Battersby 2018 is testing a comprehensive psychosocial care planning approach, building self-management capacity within a collaborative approach and providing a recovery-oriented

framework in Australia (scheduled to finish late 2020). Happell 2018 is trialling a Physical Health Nurse Consultant service for people with psychosis in Australia (findings due December 2021). Fields 2019 is testing collaborative care for people with SMI and cancer in the US (estimated finish date May 2022). Hanlon 2014 is trialling a task-sharing model of locally delivered mental health care integrated into primary health care for people with SMI in Ethiopia (study completed 2017, results were due late 2020). Nicole 2018 is piloting an interactive obesity treatment approach for people with SMI in the US (study completed June 2020, no results available at present). Aschbrenner 2019 is trialling a virtual learning collaborative to implement health promotion for people with SMI in the US (estimated completion November 2020). Byng 2023, the research study affiliated with this review, is trialling collaborative care based in GP practices for people with SMI in England (study completed March 2021, awaiting publication of results).

Included studies

Seven new studies were included in this review update (Chatterjee 2011; Chwastiak 2018; Kilbourne 2012; Kilbourne 2013; Mishra 2017; Salman 2014; van der Voort 2015), building on the one included study in the original review (Bauer 2006) (see Characteristics of included studies). Four trials were based in the US (Bauer 2006; Chatterjee 2011; Kilbourne 2012; Kilbourne 2013), one in the Netherlands (van der Voort 2015) and three in India (Chatterjee 2011; Mishra 2017; Salman 2014).

All studies required further outcome information, and we contacted all authors of these studies, obtaining additional data for four trials (Bauer 2006; Kilbourne 2012; Kilbourne 2013; van der Voort 2015).

Design and duration

A variety of different RCT designs were included in this review. Four studies were multicentre trials (Bauer 2006; Chwastiak 2018; Kilbourne 2012; Kilbourne 2013) and another was a cluster-randomised trial (van der Voort 2015). Two studies were pilot trials (Chwastiak 2018; Kilbourne 2012). Two studies were single-centre RCTs (Mishra 2017; Salman 2014).

The longest trial had a duration of 36 months (Bauer 2006), one trial reported data at 24 months (Kilbourne 2013), four studies reported data at 12 months (Chatterjee 2011; Kilbourne 2012; Kilbourne 2013; van der Voort 2015), two trials had a duration of six months (Mishra 2017; Salman 2014) and another also reported data at six months (van der Voort 2015). The shortest trial was three months (Chwastiak 2018).

Participants

Age

All eight studies focussed on adults, with one including anyone over the age of 16 (Chatterjee 2011). An upper age limit was a feature of two studies: age 60 (Chatterjee 2011) and age 70 (Chwastiak 2018). The median age of participants in the studies varied from 35.6 (10.2) (Chatterjee 2011) to 53.1 (10.6) (Kilbourne 2013). One study did not clearly report the median or mean age of participants (Mishra 2017).

Diagnosis

Studies included a variety of diagnoses on the schizophrenia and bipolar spectrum. Three studies included participants with schizophrenia and schizoaffective disorders (Chatterjee 2011; Chwastiak 2018; Salman 2014), one study bipolar disorder type 1 and 2 (Bauer 2006), three studies bipolar disorder type 1, 2 and bipolar not otherwise specified (NOS) (Kilbourne 2012; Kilbourne 2013; van der Voort 2015) and one study included people with diagnoses of schizophrenia or bipolar (Mishra 2017). Three of the studies specified that participants had to have a comorbid diagnosis of a physical health condition. One study required participants to have a comorbid diagnosis of type 2 diabetes, cardiovascular disease, haemoglobin A1c or high blood pressure (over 140/90) (Chwastiak 2018). Two studies required participants to have a comorbid diagnosis of hypertension, hyperlipidaemia, diabetes or a BMI over 25 (Kilbourne 2012; Kilbourne 2013).

Ethnicity

Five studies reported ethnicity. One study reported ethnicity by Caste, due to location (India) (Chatterjee 2011). In Bauer 2006, 23% of participants were reported as a 'minority' ethnicity. In Kilbourne 2012, 19% were reported to be 'African-American'. In Kilbourne 2013, 5.1% were reported as 'non-white'. In Chwastiak 2018, 60% were described as 'non-white'. Three studies did not report ethnicity (Mishra 2017; Salman 2014; van der Voort 2015).

Sex

All studies reported the sex of participants. In Bauer 2006, 6% were female, in Chatterjee 2011, 47%, Chwastiak 2018, 34.3%, Kilbourne 2012, 61%, Kilbourne 2013, 17%, Salman 2014, 55%, Mishra 2017, 49.5% and van der Voort 2015, 63.8%.

Setting

The majority of studies were located in secondary care outpatient services (Bauer 2006; Chwastiak 2018; Kilbourne 2012; Kilbourne 2013; Mishra 2017; Salman 2014; van der Voort 2015). One study was located in the community (Chatterjee 2011).

Study size

The eight included studies randomised a total of 1165 participants.

Interventions

We included any intervention described by the authors as 'collaborative care' (n = 8). We categorised trial interventions as type A collaborative care if they comprised the four 'core' components (multidisciplinary approach, which includes primary care, structured management plan, scheduled follow-ups and enhanced interprofessional communication) and type B collaborative care if they did not (see Appendix 2). Two of our studies met the criteria for type A (Chwastiak 2018; Kilbourne 2013).

CP extracted descriptive information regarding the interventions being tested and fidelity assessment TiDIER checklists (Hoffman 2014) for each study, as we were unable to locate checklists completed by the triallists. We describe the interventions in relation to the four 'core' components of collaborative care in summary Table 1 and Characteristics of included studies. For an overview of all constituent components of the study interventions, see Table 2.

Multidisciplinary approach

All interventions had a team that comprised a mental health professional and at least one other professional. Two interventions reported the inclusion of a primary care professional in the multidisciplinary team (Chwastiak 2018; Kilbourne 2013).

According to our definition, pharmacists are providers of primary care. However, it should be noted that where a pharmacist was included they were based in a secondary care setting (Mishra 2017; Salman 2014).

Five of the interventions had a case manager as defined by the triallists (Bauer 2006; Kilbourne 2012; Kilbourne 2013; Salman 2014; van der Voort 2015). Two of the interventions used a mental health professional case manager (Chwastiak 2018; van der Voort 2015), and two used a nurse (Bauer 2006; Kilbourne 2012). We did not systematically assess variation in the implementation of case management, for example in relation to the core tasks, intensity of involvement, breadth of services overseen and duration of involvement. However, we did note that although there were variations in the details reported, there were also some common features of case management, including care co-ordination and liaison with other providers helping to overcome fragmentation of care, patient education and patient reminders.

Structured management plan

All of the interventions had some form of a structured management plan, defined as "access to evidence-based management information. This could be in the form of guidelines or protocols. Interventions could include both pharmacological (e.g. antidepressant medication) and non-pharmacological interventions (e.g. patient screening, patient and provider education, counselling, cognitive behaviour therapy)" (Gunn 2006).

Two interventions provided evidence-based management information to providers (Kilbourne 2012; Kilbourne 2013). Four interventions included behaviour change/psycho-education or psychotherapy for participants (Chatterjee 2011; Chwastiak 2018; Kilbourne 2012; Kilbourne 2013) and five interventions included pharmacological interventions, for example facilitating treatment adherence (Chatterjee 2011; Chwastiak 2018; Mishra 2017; Salman 2014), monitoring symptoms or adverse effects (Kilbourne 2012; Kilbourne 2013; Salman 2014).

Scheduled follow-ups

All interventions reported scheduled patient follow-ups, defined as one or more scheduled telephone (Kilbourne 2012; Kilbourne 2013) or in-person follow-up appointment (Chatterjee 2011; Chwastiak 2018; Kilbourne 2012; Kilbourne 2013; Salman 2014). One study did not clearly describe follow-up but indicated that it was a feature (van der Voort 2015). The number of follow-ups scheduled varied from 4 to 27.

Enhanced interprofessional communication

Five interventions introduced mechanisms to 'facilitate communication between professionals' via interprofessional meetings (Bauer 2006; Chatterjee 2011; Chwastiak 2018; Salman 2014; van der Voort 2015).

Three studies did not include enhanced interprofessional communication as an intervention component (Kilbourne 2012; Kilbourne 2013; Mishra 2017).

Fidelity

Seven studies reported the mechanisms used to ensure the intervention was delivered as intended (Bauer 2006; Chatterjee 2011; Chwastiak 2018; Kilbourne 2012; Kilbourne 2013; Salman 2014; van der Voort 2015).

Fidelity checklist

A study-specific fidelity checklist to record the collaborative care elements delivered was reported by one study (van der Voort 2015).

Training/supervision of staff

One study reported staff training (Chwastiak 2018). One study reported supervision (Salman 2014) and four studies reported staff training alongside supervision (Bauer 2006; Chatterjee 2011; Kilbourne 2012; van der Voort 2015).

Communicating updates

One study reported regular conference calls and newsletters as a way to provide updates on treatment guidelines, to discuss difficult cases and to review access and continuity issues (Bauer 2006).

Guidelines, manuals and intervention protocols

Two studies reported using a manual (Bauer 2006; Chatterjee 2011) and two reported using a standardised set of protocols alongside a manual (Kilbourne 2012; Kilbourne 2013).

Observation and monitoring

Observation and monitoring was carried out in three studies (Bauer 2006; Kilbourne 2012; Kilbourne 2013). Specifically, Kilbourne 2013 reported observing 50% of groups and monitoring of patient and provider contacts based on the registry.

Catch-up sessions

Participants missing group sessions received catch-up sessions on the phone, to ensure the intervention was delivered as planned, in one study (Kilbourne 2012).

Comparison - usual care

Usual care, where participants continued to receive treatment as usual was the comparator in five studies (Bauer 2006; Chatterjee 2011; Chwastiak 2018; Mishra 2017; van der Voort 2015).

Three further studies reported enhanced usual care by:

- sending mailings on wellness topics (Kilbourne 2012);
- sending mailings on wellness topics, providing referrals to primary care services off-site and providing general medical and mental health providers with the same practice guidance information (Kilbourne 2013);
- providing diary cards as a medication adherence prompt (Salman 2014).

Outcome scales

Many trials used different scales in assessing treatment effects for various outcomes. We considered outcomes in relation to the impact of the intervention on the individual. Some trials had common outcomes, such as mental state, quality of life and mood. Different scales were used in assessing intervention effects for various outcomes. We conducted statistical pooling using standardised mean differences where appropriate. We show the details of scales that provided usable data in Table 3. In Table 4 we outline which predefined outcomes have no data. In the

Characteristics of included studies table under 'Outcomes' we also outline which measures were included in trials but did not provide useable data for this review.

Excluded studies

In this update, we excluded a total of 180 records (111 studies). We excluded 24 records because they were not RCTs (see Types of studies), 110 records because the intervention was not described as collaborative care by the trialists (Types of interventions) and 43 records because the sample did not meet the participant criteria for inclusion (see Types of participants). These data are presented in the Characteristics of excluded studies table and in the flow chart in Figure 1.

Risk of bias in included studies

Risk of bias was dual assessed by CHM and one other review author (BG, BD, PH or DR). Risk of bias 2.0 (RoB 2) assesses the risk of bias at outcome level. To balance rigour with the burden of assessment, we assessed the risk of bias only for the review's primary outcomes and the outcomes reported in our summary of findings table. This also meant that we assessed risk of bias in at least one outcome per study. We therefore assessed risk of bias at all time points.

1. Quality of life: Bauer 2006; Kilbourne 2012; Kilbourne 2013; Mishra 2017; Salman 2014; van der Voort 2015.
2. Mental state: Chatterjee 2011; Chwastiak 2018; Kilbourne 2012; Kilbourne 2013; van der Voort 2015.
3. Psychiatric admissions: Bauer 2006; Chatterjee 2011.
4. Social functioning: Chatterjee 2011 (binary only).

The text below describes the review authors' responses to domain signalling questions in assessing risk of bias for the primary outcomes of this review. The review authors found that in domains 1, 2, 3 and 5, risk of bias did not vary across review outcomes in the same study, with the exception of domain 5 in Chatterjee 2011. Therefore, in these domains, risk of bias is reported by study rather than by outcome.

Domain 1: Risk of bias arising from the randomisation process

All studies used a random allocation sequence. One study outlined that allocation was concealed until after participant enrolment (Chatterjee 2011). One study did not randomise participants until after recruitment (Bauer 2006). One study did not conceal allocation prior to consent from participants, recruiting researchers or those delivering the intervention. This study did conceal allocation from researchers collecting data (van der Voort 2015). The remaining studies did not provide detailed information as to allocation concealment (Chwastiak 2018; Kilbourne 2012; Kilbourne 2013; Mishra 2017; Salman 2014), however the description of job roles and task timing in three of these studies led to a judgement that allocation was probably concealed.

Most studies reported no substantial differences between arms at baseline (Bauer 2006; Chatterjee 2011; Chwastiak 2018; Kilbourne 2012; Kilbourne 2013; Salman 2014); this was the case across all outcomes assessed for risk of bias in these studies. One of these studies noted some differences at baseline in one of the recruiting sites (Chatterjee 2011), but these differences were insignificant when viewed in light of the total sample. One study noted substantial baseline differences across arms and suggested that this was because allocation was not concealed until after enrolment (van der Voort 2015). For one study, it was not possible to comment on baseline differences due to a lack of clarity regarding whether baseline data were missing or misreported as 'first follow-up' (Mishra 2017). Attempts to contact the authors to clarify this matter were unsuccessful.

Domain 1b: Risk of bias arising from the timing of identification and recruitment of individual participants in relation to timing of randomisation

We assessed one cluster-randomised study for risk of bias in this subdomain (van der Voort 2015). This study did not identify individual participants before clusters were randomised. The clinician responsible for identifying eligible patients was aware of the allocation of these clusters; therefore, knowledge of the intervention could have affected the selection of individual participants. There are substantial differences in the demographics of the two arms that suggest this may have led to differential identification of participants between the two arms.

Domain 2: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)

As these trials investigate psychosocial interventions, it is expected that participants who have provided informed consent would be aware of whether they had been assigned to the intervention or the control arm of the trial. It is also expected that the people delivering this type of intervention would be aware of the participants' assigned intervention. Therefore, the review authors did not judge participant and practitioner awareness of assignment to intervention to impact on risk of bias in this domain.

None of the studies noted any deviations from the intended intervention due to the trial context. One study reported several deviations from the intervention in the intervention arm (van der Voort 2015); it was unclear whether these were due to the trial context. Three studies reported acceptable to good fidelity to the intervention protocol: Bauer 2006, 80%; Kilbourne 2012, 79%; Kilbourne 2013, 68%. Again, it was not stated whether any deviations were due to the trial context. Four studies did not offer a quantitative measure of fidelity (Chwastiak 2018; Mishra 2017; Salman 2014).

Some studies clearly stated their analysis to be on an intention-to-treat basis (Bauer 2006; Chatterjee 2011; van der Voort 2015), which is considered appropriate for analysing the effect of assignment to the intervention. Other studies did not explicitly state whether intention-to-treat analysis was undertaken to appropriately assess assignment to intervention (Kilbourne 2012; Kilbourne 2013); as these studies do not explicitly state that those deviating from the intended intervention were removed from analysis, we have assumed that analysis was undertaken on an intention-to-treat basis.

Domain 3: Risk of bias due to missing outcome data

Studies assessed for risk of bias reported missing data at study, rather than outcome, level. Additionally, we considered missingness of all outcomes to be likely to relate to true value for all of these outcomes. Therefore, the text below considers risk of bias at study level rather than at outcome level.

In three studies, data were available for nearly all participants randomised; we defined 'nearly' as 90% of participants or clusters enrolled providing follow-up data (Bauer 2006; Chatterjee 2011; Kilbourne 2012). Four studies did not have data available for nearly all participants randomised due to high rates of participant dropout (Chwastiak 2018; Kilbourne 2013; Salman 2014; van der Voort 2015). In one study, it was not possible to accurately comment on dropout rates due to lack of clarity regarding baseline data (Mishra 2017). The authors did not respond to requests to clarify these data.

In three studies with high dropout, numbers were similar across both arms, lowering the chance of bias from missing outcome data (Chwastiak 2018; Kilbourne 2013; Salman 2014). Notwithstanding this, both a person's quality of life and mental state are likely to impact dropout, suggesting that missingness could depend on the true value of these data. In the remaining study there were higher dropouts in the intervention arm (26/71 dropouts in the intervention arm, 10/82 in the control arm), including the withdrawal of two clusters (van der Voort 2015). Inconsistency in reporting in this study makes it difficult to fully understand the dropout rates, with participants randomised to the control arm inconsistently reported as 82 or 80 participants. In this study, the withdrawal of two clusters from the intervention arm for organisational reasons suggests that the intervention may be less likely to work in these clusters; therefore, missing data in this study are likely to depend on the true value (van der Voort 2015).

Domain 4: Risk of bias in measurement of the outcome

All of the studies assessing quality of life and some of those assessing mental state did so using self-report tools (Table 3). Although this does not negatively impact the appropriateness of the measure, there is interplay between the impossibility of masking participants in psychosocial interventions and the possibility of social desirability bias when measuring outcomes. When assessing risk of bias for outcomes using self-report measures, we have considered the participant to be an outcome assessor in addition to the researcher assessor. However, we did not consider the unmasking of these participant outcome assessors to be sufficient justification to increase the bias risk. This is in part because to do so would undermine patient voice in research by suggesting that self-report measures have less scientific rigour.

Quality of life (QoL)

Studies used self-report measures to capture quality of life: two studies utilised the WHOQOL-BREF (Mishra 2017; van der Voort 2015), three the SF-12 (Kilbourne 2012; Kilbourne 2013; Salman 2014) and one the SF-36 (Bauer 2006). We consider these measures appropriate for measuring QoL. In all studies it was considered unlikely that the measurement of QoL could vary between arms. In three studies, researcher outcome assessors were blinded to intervention allocation (Bauer 2006; Salman 2014; van der Voort 2015). In one study, research assessors were not blinded to intervention allocation (Kilbourne 2013). In two studies, it was not reported whether researcher assessors were aware of the intervention received by study participants (Kilbourne 2012; Mishra 2017). As these measures are self-reported and the participants are assumed to have given informed consent, we judged the participants to be unmasked outcome assessors. As such, we judged the likelihood of social desirability bias to have probably influenced outcome assessment in all studies assessing quality of life.

Mental state

Mental state included schizophrenia symptoms, bipolar symptoms, discrete depression and mania symptoms, and overall symptoms.

Mental state in relation to schizophrenia

Two studies utilised the Positive and Negative Syndrome Scale (PANSS), which is an appropriate measure for this concept (Chatterjee 2011; Salman 2014). One study utilised the Brief Psychiatric Rating Scale (BPRS) to measure schizophrenia symptoms (Chwastiak 2018). Although the BPRS is designed to measure general psychiatric symptoms, it is commonly used to measure schizophrenia symptoms. Therefore, we judged this to be an appropriate measure. As both BPRS and PANSS are clinician-rated measures, we considered masking of assessors to be important in assessing bias. In two studies, assessors were masked to allocation (Chatterjee 2011; Salman 2014); one study did not report whether assessors were masked (Chwastiak 2018).

Mental state in relation to depression and mania symptoms

Two studies utilised a measure that captured both the mania and depression symptoms of bipolar: the Internal State Scale (ISS) (Kilbourne 2012; Kilbourne 2013). Other studies captured these symptoms separately: one study utilised the Quick Inventory of Depression Symptomatology (QIDS) to measure depression and the Altman Self-Rating Mania Scale (ASRMS) to measure mania (van der Voort 2015), and one study used the Patient Health Questionnaire-9 (PHQ-9) to measure depression symptoms (Chwastiak 2018). We considered all these measures appropriate for measuring these concepts, and that measurements were unlikely to vary across the arms of the studies. Therefore, it is unlikely that bias may have influenced the outcome measured.

In one study, researcher outcome assessors were masked to intervention allocation (van der Voort 2015). In one study, research assessors were unmasked as to intervention allocation (Kilbourne 2013). In two studies, it was not reported whether researcher assessors were aware of the intervention received by study participants (Chwastiak 2018; Kilbourne 2012).

Psychiatric admissions

In two studies, admissions were measured as the number of participants who were admitted to a psychiatric hospital, an appropriate measure (Bauer 2006; Chatterjee 2011). It is unlikely that measurement could have varied between arms. Assessors were masked to intervention allocation.

Disability (proxy measure for social functioning)

Disability was measured using the Indian Disability Evaluation and Assessment Scale (IDEAS) (Chatterjee 2011). This measure has good internal consistency and validity in schizophrenia populations (Grover 2014), and was considered appropriate by the review team. It is unlikely that measurement could have varied between arms. Researcher assessors were masked to allocation.

Domain 5: Risk of bias in selection of the reported result

Four studies did not publish a pre-specified analysis plan (Chwastiak 2018; Kilbourne 2012; Mishra 2017; Salman 2014). For two studies, this was because the study was a feasibility study with an explicit aim to test analysis options as part of the feasibility testing (Chwastiak 2018; Kilbourne 2012). We still considered the

lack of a pre-specified plan to increase the risk of bias in these studies, as the results presented may have been selected on the basis of multiple eligible outcomes and/or multiple eligible analyses of the data. In one study, this risk of bias was compounded by lack of clarity around the baseline data, which increased the likelihood of the published results being selected from multiple eligible outcomes (Mishra 2017).

Four studies did produce results in accordance with a pre-specified statistical analysis plan (Bauer 2006; Chatterjee 2011; Kilbourne 2013; van der Voort 2015), however whether these plans were published before these unblinded outcome data were available is unknown. We did not judge this to increase the risk of bias. In two studies, it was unclear in the protocol paper which outcome was intended to be the primary outcome of the study (Bauer 2006; van der Voort 2015). In one study, some outcomes were missing from the published results, including symptoms (brief symptom inventory) and severity of bipolar disorder (Clinical Global Impression for bipolar disorder) (van der Voort 2015). We judged this to increase the risk of bias in this study and its outcomes, due to the possibility of outcome data being selected from multiple eligible outcomes. In the other study, all results were still published; therefore, we did not consider this to increase the risk of bias in this study (Bauer 2006).

Disability (social functioning)

This outcome was part of a post hoc analysis undertaken in addition to the pre-specified analysis plan (Chatterjee 2011). It is likely that this analysis was undertaken to stress-test the results in light of the baseline imbalances reported in one of the study sites. As this analysis was undertaken in addition to the pre-planned analyses, we did not deem it likely that the result was likely to have been selected on the basis of multiple eligible outcome measurements or multiple eligible analyses of the data.

Overall risk of bias

Risk of bias varied considerably between included studies and across the primary outcomes of the review. In relation to quality of life, we assessed one study as having a low risk of bias (Bauer 2006), three studies as having a high risk of bias (Mishra 2017; Salman 2014; van der Voort 2015) and two studies as having some concerns (Kilbourne 2012; Kilbourne 2013). For mental state, schizophrenia symptoms, we rated one study as having a low risk of bias (Chatterjee 2011) and a second a high risk of bias (Salman 2014). Of the three studies reporting depression, mania and bipolar symptoms of mental state, we rated two as having a high risk of bias (Chwastiak 2018; van der Voort 2015) and two as having some concerns of risk of bias (Kilbourne 2012; Kilbourne 2013). We rated the two studies reporting the number of psychiatric admissions as having low risk of bias (Bauer 2006; Chatterjee 2011). We rated the one study reporting disability as having low risk of bias (Chatterjee 2011).

Effects of interventions

See: **Summary of findings 1** Summary of findings table - Collaborative care compared to usual care for severe mental illness

Collaborative care versus usual care

Eight included studies compared collaborative care with usual care (Types of interventions). The outcomes measured by each included study are described fully in Table 3. Those that were pre-specified as relevant to this review are described in the Types of outcome measures section. Each of these outcomes, which are measured in any of the eight included studies, will be discussed in turn. Sensitivity and subgroup analyses are presented concurrently where appropriate (see also, Sensitivity analysis, Subgroup analysis and investigation of heterogeneity). For those outcomes where a GRADE assessment was undertaken the assessment of the certainty of the evidence is presented alongside results.

Predefined outcomes where no data were available

There are a number of predefined outcomes for which we have no data available: personal recovery, global state, substance use (alcohol/illicit drug/cigarette/tobacco), adverse effects/events, service user outside of mental health and experience of care/satisfaction (participant, carer or staff) (Table 4). Personal recovery was not measured by any of the studies despite being highlighted as important by those with ongoing mental health problems (Retzer 2020).

The fidelity measures for which we have no data available are: measures of healthcare professional behaviour and knowledge and measures of adherence to manual/algorithms/guidance.

Primary outcomes

1.1 Quality of life: clinically important change (average endpoint in mental health component) - 12 months

See Analysis 1.1.

We found three studies that assessed quality of life of participants at 12 months (Table 3) (Kilbourne 2012; Kilbourne 2013; van der Voort 2015). Quality of life was measured using the SF-12 (Kilbourne 2012; Kilbourne 2013) and the WHOQOL-BREF (van der Voort 2015), and the mean endpoint mental health component scores were reported for 12 months. No clear difference between collaborative care and usual care was observed in the medium term (SMD 0.03, 95% CI -0.26 to 0.32; I^2 = 19%; 3 studies, 227 participants). There was very low-certainty evidence for this outcome.

1.2 Mental state: clinically important change (binary) - 12 months

See Analysis 1.2.

We found one study in which mental state was reported as a binary outcome at 12 months, where the number of participants experiencing an improvement of 20% or more on the PANSS overall score was reported (Chatterjee 2011). There was no evidence of a difference in mental state in the collaborative care arm compared to usual care (RR 0.99, 95% CI 0.77 to 1.28; 1 study, 253 participants). There was low-certainty evidence for this outcome.

1.3 Psychiatric hospital admissions: number of participants admitted to hospital - 12 months

See Analysis 1.3.

We found one study that measured participants who were admitted to a psychiatric hospital at 12 months (Chatterjee 2011). The proportion of participants who were admitted to a psychiatric hospital in the collaborative care arm was 6% and in the control arm was 1%. Whilst there is a suggestion that more psychiatric admissions were observed in the collaborative care arm compared to usual care in the medium term (RR 5.15, 95% CI 0.67 to 39.57;

1 study, 253 participants), this result is not statistically significant and there is substantial uncertainty around this estimate due to the small numbers of admissions in both arms. There was low-certainty evidence for this outcome.

Secondary outcomes

2.1 Quality of life

See Analysis 2.1.

We found six studies that assessed quality of life of participants using various scales at different time points (Table 3) (Bauer 2006; Kilbourne 2012; Kilbourne 2013; Mishra 2017; Salman 2014; van der Voort 2015). Quality of life was assessed using the SF-36, SF-12 and WHOQOL-BREF; the mean endpoint physical health and mental health component scores were reported for the up to six months, 7 to 12 months and more than 12 months follow-up periods. Clinically important change in quality of life (average endpoint in mental health component) at 12 months is a primary outcome, as is reported above (Analysis 1.1).

2.1.1 Quality of life: average endpoint in physical health - up to six months

We found five studies that assessed physical health-related quality of life up to six months (Kilbourne 2012; Kilbourne 2013; Mishra 2017; Salman 2014; van der Voort 2015). No clear difference between collaborative care and usual care was observed (SMD 0.55, 95% CI -0.24 to 1.33; I^2 = 93%; 5 studies, 406 participants). However, we observed substantial heterogeneity upon meta-analysis of these results (I^2 = 93%), alongside inconsistencies between studies in the direction of effect. The result of this meta-analysis should therefore be interpreted with caution. Subgroup analyses are explored in order to attempt to explain this heterogeneity (Analysis 4.1) (see also, Subgroup analysis and investigation of heterogeneity).

2.1.2 Quality of life: average endpoint in physical health - 12 months

We found three studies that assessed physical health-related quality of life at 12 months (Kilbourne 2012; Kilbourne 2013; van der Voort 2015). No clear difference between collaborative care and usual care was observed in the medium term (SMD 0.08, 95% CI -0.18 to 0.33; I^2 = 0%; 3 studies, 237 participants).

2.1.3 Quality of life: average endpoint in physical health - more than 12 months

The longer-term data were measured in two studies, Bauer 2006 and Kilbourne 2013, at 36 months and 24 months, respectively. No clear difference between collaborative care and usual care was observed in the long term (SMD 0.02, 95% CI -0.19 to 0.24; I^2 = 7%; 2 studies, 381 participants). There was very low-certainty evidence for this outcome.

2.1.4 Quality of life: average endpoint in mental health - up to six months

Five studies measured mean endpoint mental health-related quality of life up to six months (Kilbourne 2012; Kilbourne 2013; Mishra 2017; Salman 2014; van der Voort 2015). No clear difference between collaborative care and usual care was observed (SMD 0.71, 95% CI -0.17 to 1.59; I^2 = 94%; 5 studies, 406 participants). However, we observed substantial heterogeneity upon meta-analysis of these results (I^2 = 94%), as well as significant inconsistencies between studies in the magnitude and direction of effect. The results of the meta-analysis should therefore be interpreted with caution. Subgroup analyses are explored instead in order to attempt to explain this heterogeneity (Analysis 4.1; Analysis 4.2; Analysis 4.3; Analysis 4.4) (see also, Subgroup analysis and investigation of heterogeneity).

2.1.5 Quality of life: average endpoint in mental health - more than 12 months

The longer-term data at more than 12 months were measured in two studies, Bauer 2006 and Kilbourne 2013, at 36 months and 24 months, respectively. No clear difference between collaborative care and usual care was observed in the longer term (SMD 0.30, 95% CI -0.10 to 0.70; I^2 = 62%; 2 studies, 381 participants).

2.1.6 Quality of life: overall endpoint (WHOQOL-BREF) - six months

We found one study that measured overall quality of life using the WHOQOL-BREF at six months (van der Voort 2015). No clear difference between collaborative care and usual care was observed in the short term (SMD -0.20, 95% CI -0.61 to 0.22; 1 study, 94 participants).

2.1.7 Quality of life: overall endpoint (WHOQOL-BREF) - 12 months

We found one study that measured overall quality of life using the WHOQOL-BREF at 12 months (van der Voort 2015). No clear difference between collaborative care and usual care was observed in the medium term (SMD 0.11, 95% CI -0.31 to 0.54; 1 study, 91 participants).

2.2 Mental state

See Analysis 2.2.

2.2.1 Mental state (overall general score) - up to six months

We found one small study that measured mental state using the BPRS (Chwastiak 2018). No clear difference between collaborative care and usual care was observed in the short term (SMD -0.34, 95% CI -1.07 to 0.40; 1 study, 29 participants).

2.2.2 Mental state (general psychopathology) - up to six months

We found one study that measured mental state using the PANSS general subscale (change from baseline) (Salman 2014). No clear difference between collaborative care and usual care was observed in the short term (SMD -0.11, 95% CI -0.55 to 0.33; 1 study, 80 participants).

2.2.3 Mental state (general psychopathology) - at 12 months

We found one study that measured general symptoms in mental state using the PANSS (Chatterjee 2011). There was evidence of a difference between collaborative care compared with usual care in the medium term (SMD -0.27, 95% CI -0.53 to -0.01; 1 study, 253 participants).

2.2.4 Mental state (positive symptoms) - up to six months

We found one study that measured positive psychotic symptoms using the PANSS general subscale (change from baseline) (Salman 2014). No clear difference between collaborative care and usual care was observed in the short term (SMD -0.04, 95% CI -0.48 to 0.40; 1 study, 80 participants).

2.2.5 Mental state (positive symptoms) - at 12 months

We found one study that measured positive psychotic symptoms using the PANSS (Chatterjee 2011). No clear difference between collaborative care and usual care was observed in the short term (SMD -0.17, 95% CI -0.43 to 0.09; 1 study, 253 participants).

2.2.6 Mental state (negative symptoms) - up to six months

We found one study that measured negative psychotic symptoms using the PANSS general subscale (change from baseline) (Salman 2014). No clear difference between collaborative care and usual care was observed in the short term (SMD -0.26, 95% CI -0.70 to 0.18; 1 study, 80 participants).

2.2.7 Mental state (negative symptoms) - at 12 months

We found one study that measured negative psychotic symptoms using the PANSS (Chatterjee 2011). No clear difference between collaborative care and usual care was observed in the short term (SMD -0.08, 95% CI -0.34 to 0.18; 1 study, 253 participants).

2.2.8 Mental state (depressive symptoms) - up to six months

We found four studies that measured depressive symptoms using the Patient Health Questionnaire-9 (PHQ-9), the Internal State Scale (ISS) and the Quick Inventory for Depressive Symptomology (QIDS) up to six months (Chwastiak 2018; Kilbourne 2012; Kilbourne 2013; van der Voort 2015). No clear difference between collaborative care and usual care was observed in the short term (SMD -0.13, 95% CI -0.53 to 0.27; I^2 = 59%; 4 studies, 259 participants).

2.2.9 Mental state (depressive symptoms) - at 7 to 12 months

We found three studies that measured depressive symptoms using the Internal State Scale (ISS) and the Quick Inventory for Depressive Symptomology (QIDS) at between 7 and 12 months (Kilbourne 2012; Kilbourne 2013; van der Voort 2015). No clear difference between collaborative care and usual care was observed in the medium term (SMD -0.17, 95% CI -0.53 to 0.18; I^2 = 45%; 3 studies, 227 participants).

2.2.10 Mental state (depressive symptoms) - at 24 months

The longer-term data were provided in one study using the Internal State Scale (ISS) (Kilbourne 2013). No clear difference between collaborative care and usual care was observed in the longer term for depressive symptoms at the 24-month follow-up (SMD -0.19, 95% CI -0.64 to 0.27; 1 study, 75 participants).

2.2.11 Mental state (manic symptoms) - up to six months

We found three studies that measured manic symptoms using the Internal State Scale (ISS) and the Altman Self-Rating Mania scale (Kilbourne 2012; Kilbourne 2013; van der Voort 2015). No clear difference between collaborative care and usual care was observed in the short term (SMD -0.14, 95% CI -0.40 to 0.12; I^2 = 0%; 3 studies, 230 participants).

2.2.12 Mental state (manic symptoms) - at 7 to 12 months

We found three studies that measured manic symptoms using the Internal State Scale (ISS) and the Altman Self-Rating Mania scale (Kilbourne 2012; Kilbourne 2013; van der Voort 2015). No clear difference between collaborative care and usual care was observed in the medium term (SMD -0.08, 95% CI -0.38 to 0.22; I^2 = 22%; 3 studies, 227 participants).

2.2.13 Mental state (manic symptoms) - greater than 12 months

The longer-term data were provided in one study using the Internal State Scale (ISS) (Kilbourne 2013). The data suggested some indication that collaborative care resulted in a reduction of manic symptoms at the 24-month follow-up (SMD -0.36, 95% CI -0.82 to 0.10; 1 study, 75 participants), although this difference was not statistically significant.

2.3 Psychiatric hospital admissions: number of participants admitted to hospital - greater than 12 months

See Analysis 2.3.

Psychiatric hospital admissions: number of participants admitted to hospital (12 months) is a primary outcome, so is reported above.

2.3.1 Number of participants admitted to hospital - year two

We found one study that measured participants who were admitted to a psychiatric hospital at 24 months (Bauer 2006). Data were collected from the VA National Patient Care Database and Pharmacy Benefits Management Package. For year two, the proportion of participants hospitalised in a psychiatric hospital was statistically significantly lower in the intervention group than the standard care group: 35% compared to 47% (RR 0.75, 95% CI 0.57 to 0.99; 1 study, 306 participants).

2.3.2 Number of participants admitted to hospital - year three

We found one study that measured participants who were admitted to a psychiatric hospital at 36 months (Bauer 2006). For year three, the proportion was again lower in the collaborative care arm, but this was not statistically significant: 28% compared to 38% (RR 0.73, 95% CI 0.53 to 1.01; 306 participants).

2.4 Other hospital admissions

See Analysis 2.4.

2.4.1 Number of participants admitted to hospital - up to 12 months

We found one study that measured participants who were admitted to a hospital at 12 months (Chatterjee 2011). However, the study found no (non-psychiatric) hospital admissions in the usual care arm, and so comparative analysis between groups is not appropriate.

2.4.2 Number of participants admitted to hospital - in year two

We found one study that measured participants who were admitted to a hospital in year two (Bauer 2006). Data were collected from the VA National Patient Care Database and Pharmacy Benefits Management Package. The proportion of participants hospitalised for any reason was lower in the intervention group than the standard care group: 44% compared to 53%, although this difference was not statistically significant (RR 0.83, 95% CI 0.65 to 1.04; 1 study, 306 participants). However, it is not clear from the papers if this outcome also included psychiatric admissions.

2.4.3 Number of participants admitted to hospital - in year three

We found one study that measured participants who were admitted to hospital in year three (Bauer 2006). The data suggested that collaborative care resulted in fewer admissions to hospital than usual care in the longer term: 34% compared to 48% (RR 0.70, 95% CI 0.53 to 0.93; P = 0.01; 1 study, 306 participants). However, it is

not clear from the papers if this outcome also included psychiatric admissions.

2.5 Personal recovery

No data available.

2.6 Physical health status

See Analysis 2.6.

2.6.1 Blood pressure, mmHg systolic - up to six months

We found three studies that measured systolic blood pressure at six months (Chwastiak 2018; Kilbourne 2012; Kilbourne 2013). No clear difference between collaborative care and usual care was observed in the short term (SMD -0.15 mmHg, 95% CI -0.54 mmHg to 0.24 mmHg; I^2 = 35%; 3 studies, 165 participants).

2.6.2 Blood pressure, mmHg systolic - at 7 to 12 months

We found two studies that measured systolic blood pressure at 12 months (Kilbourne 2012; Kilbourne 2013). No clear difference between collaborative care and usual care was observed in the medium term (SMD -0.20 mmHg, 95% CI -0.54 mmHg to 0.13 mmHg; I^2 = 0%; 2 studies, 136 participants).

2.6.3 Blood pressure, mmHg systolic - 24 months

Longer-term data were provided by one study (Kilbourne 2013). No clear difference between collaborative care and usual care was observed in the longer term at the 24-month follow-up (SMD -0.22 mmHg, 95% CI -0.67 mmHg to 0.24 mmHg; 1 study, 75 participants).

2.6.4 Blood pressure, mmHg diastolic - six months

We found two studies that measured diastolic blood pressure at six months (Kilbourne 2012; Kilbourne 2013). No clear difference between collaborative care and usual care was observed in the short term (SMD -0.25 mmHg, 95% CI -0.77 mmHg to 0.27 mmHg; I^2 = 57%; 2 studies, 136 participants).

2.6.5 Blood pressure, mmHg diastolic - 7 to 12 months

We found two studies that measured diastolic blood pressure at 12 months (Kilbourne 2012; Kilbourne 2013). There was no clear difference between collaborative care and usual care in the medium term (SMD -0.29 mmHg, 95% CI -0.70 mmHg to 0.12 mmHg; I^2 = 32%; 2 studies, 136 participants).

2.6.6 Blood pressure, mmHg diastolic - 24 months

Longer-term diastolic blood pressure data were provided by one study (Kilbourne 2013). No clear difference between collaborative care and usual care was observed in the longer term at the 24-month follow-up (SMD -0.25 mmHg, 95% CI -0.70 mmHg to 0.21 mmHg; 1 study, 75 participants).

2.6.7 Body mass index (BMI) - six months

Body mass index (BMI) is a measure of body fat based on height and weight that applies to adult men and women. We found three studies that measured BMI at six months (Chwastiak 2018; Kilbourne 2012; Kilbourne 2013). No clear difference between collaborative care and usual care was observed in the short term (SMD -0.18, 95% CI -0.50 to 0.15; I^2 = 9%; 3 studies, 165 participants).

2.6.8 Body mass index (BMI) - 12 months

We found two studies that measured BMI at 12 months (Kilbourne 2012; Kilbourne 2013). The data show that there is evidence of a difference between allocated groups, indicating that the collaborative care arm had a lower BMI in the medium term compared to the control arm (SMD -0.37, 95% CI -0.71 to -0.03; I^2 = 0%; 2 studies, 136 participants).

2.6.9 Body mass index (BMI) - 24 months

Longer-term data were provided by one study (Kilbourne 2013). There was little evidence of a difference in BMI between collaborative care and usual care at the 24-month follow-up (SMD -0.35, 95% CI -0.81 to 0.11; 1 study, 75 participants).

2.6.10 Total cholesterol - six months

One study measured cholesterol at six months, 12 months and 24 months (Kilbourne 2013). There was no clear difference in cholesterol between collaborative care and usual care in the short term (SMD -0.43, 95% CI -0.90 to 0.04; 1 study, 71 participants).

2.6.11 Total cholesterol - 12 months

No clear difference in total cholesterol between collaborative care and usual care was observed in the medium term (SMD -0.19, 95% CI -0.65 to 0.28; 1 study, 71 participants).

2.6.12 Total cholesterol - 24 months

No clear difference in total cholesterol between collaborative care and usual care was observed in the longer term at the 24-month follow-up (SMD 0.07, 95% CI -0.39 to 0.52; 1 study, 75 participants).

2.6.13 Triglycerides - up to six months

One study measured triglycerides at six months (Chwastiak 2018). No clear difference between collaborative care and usual care was observed in the short term (SMD -0.35, 95% CI -1.09 to 0.38; 1 study, 29 participants).

2.6.14 High-density lipoprotein (HDL) - six months

One study measured high-density lipoprotein (HDL) at six months, 12 months and 24 months (Kilbourne 2013). This measure was multiplied by -1 prior to analysis to ensure comparability in direction with other physical health outcomes (i.e. that low values are better). No clear differences in HDL between collaborative care and usual care were observed in the short term (SMD -0.06, 95% CI -0.52 to 0.41; 1 study, 71 participants).

2.6.15 High-density lipoprotein (HDL) - 12 months

No clear difference in HDL between collaborative care and usual care was observed in the medium term (SMD 0.10, 95% CI -0.36 to 0.57; 1 study, 71 participants).

2.6.16 High-density lipoprotein (HDL) - 24 months

No clear difference between collaborative care and usual care differences were observed in the longer term (SMD -0.19, 95% CI -0.64 to 0.27; 1 study, 75 participants).

2.6.17 Low-density lipoprotein (LDL) - six months

Two studies measured LDL at six months (Chwastiak 2018; Kilbourne 2013). Kilbourne 2013 also measured LDL at 12 months and 24 months. There was little evidence that the collaborative care

group had a lower LDL in the short term (SMD -0.34, 95% CI -0.73 to 0.06; I^2 = 0%; 2 studies, 100 participants), but this result is not statistically significant.

2.5.18 Low-density lipoprotein (LDL) - 12 months

No clear difference between collaborative care and usual care was observed in the medium term (SMD -0.12, 95% CI -0.59 to 0.34; 1 study, 71 participants).

2.6.19 Low-density lipoprotein (LDL) - 24 months

No clear difference between collaborative care and usual care was observed in the longer term at the 24-month follow-up (SMD 0.00, 95% CI -0.46 to 0.45; 1 study, 75 participants).

2.6.20 HbA1c - up to six months

One study measured HbA1c at six months (Chwastiak 2018). No clear difference between collaborative care and usual care was observed in the short term (SMD -0.37, 95% CI -1.10 to 0.37; 1 study, 29 participants).

2.6.21 Waist circumference - six months

Two studies measured waist circumferences at six months and 12 months (Kilbourne 2012; Kilbourne 2013), and Kilbourne 2013 also measured waist circumferences at 24 months. No clear difference between collaborative care and usual care was observed in the short term (SMD -0.31, 95% CI -0.98 to 0.35; I^2 = 73%; 2 studies, 136 participants).

2.6.22 Waist circumference - 12 months

There was evidence of a difference between collaborative care compared to usual care in the medium term, indicating that the collaborative care group had, on average, a lower waist circumference (SMD -0.39, 95% CI -0.75 to -0.03; I^2 = 9%; 2 studies, 136 participants).

2.6.23 Waist circumference - 24 months

No clear difference between collaborative care and usual care was observed in the longer term at the 24-month follow-up (SMD -0.29, 95% CI -0.75 to 0.17; 1 study, 75 participants).

2.7 Global state

No data available.

2.8 Medication adherence (patient-reported) (DAI-10)

See Analysis 2.8.

One study measured patient-reported medication adherence using the Drug Attitude Inventory (DAI-10) at 6 and 12 months (van der Voort 2015). The DAI-10 is a series of 10 questions that are used to derive a binary response to medication adherence.

2.8.1 Medication adherence (patient-reported) - at six months

The data provide some indication that medication adherence was worse in the collaborative care arm than in the usual care arm (RR 0.83, 95% CI 0.67 to 1.04; 1 study, 94 participants), although this difference is not statistically significant.

2.8.2 Medication adherence (patient-reported) - at 12 months

No clear difference between collaborative care and usual care was observed in the medium term (RR 0.91, 95% CI 0.75 to 1.11; 1 study, 91 participants).

2.9 Medication adherence (patient-reported) (MARS)

See Analysis 2.9.

2.9.1 Medication adherence (patient-reported) - up to six months

One study measured patient-reported medication adherence using the Medication Adherence Rating Scale (MARS) at two months (Mishra 2017). There was a clear statistically significant difference showing that the collaborative care group had greater medication adherence than the usual care group (MD 1.79, 95% CI 1.56 to 2.02; 1 study, 96 participants).

2.10 to 2.11 Social functioning/disability

We found four studies assessing social function/disability of participants using various scales at different time points (Table 3) (Chatterjee 2011; Kilbourne 2012; Kilbourne 2013; van der Voort 2015). Social functioning/disability was assessed using the WHO Disability Scale (WHO-DAS), Functioning Assessment Short Test (FAST) and the Indian Disability Evaluation and Assessment Scale (IDEAS). The mean endpoint scores were reported for the up to six months, 12 months and more than 12 months follow-up period.

2.10.1 Social functioning/disability (binary) - 12 months

See Analysis 2.10.

One study also reported a post hoc analysis of disability at 12 months as a binary outcome, defined as an improvement of at least 20% on the IDEAS scale (Chatterjee 2011). This analysis provided some evidence that more participants in the intervention arm improved by this extent compared to the control arm (75/167 (48%) versus 28/86 (35%)) (RR 1.38, 95% CI 0.97 to 1.95; 1 study, 253 participants), although this result was not statistically significant. There was low-certainty evidence for this outcome.

2.11.1 Social functioning/disability - up to six months

See Analysis 2.11.

We found three studies assessing social functioning/disability up to six months (Kilbourne 2012; Kilbourne 2013; van der Voort 2015). No clear difference between collaborative care and usual care was observed in the short term (SMD -0.14, 95% CI -0.61 to 0.32; I^2 = 68%; 3 studies, 230 participants).

2.11.2 Social functioning/disability - 12 months

We found four studies assessing social functioning/disability at 12 months (Chatterjee 2011; Kilbourne 2012; Kilbourne 2013; van der Voort 2015). No clear difference between collaborative care and usual care was observed in the medium term (SMD -0.16, 95% CI -0.44 to 0.12; I^2 = 48%; 4 studies, 480 participants).

2.11.3 Social functioning/disability - 24 months

The longer-term data were provided by one study at 24 months using the WHO-DAS (Kilbourne 2013). No clear difference between collaborative care and usual care was observed in the long term (SMD -0.14, 95% CI -0.59 to 0.32; 1 study, 75 participants).

2.12 Substance use (alcohol/illicit drugs/cigarettes/tobacco)

No data available.

2.13 Adverse effect/event(s)

No data available.

2.14 Death

See Analysis 2.14.

2.14.1 Number of participants that died from suicide - 36 months

One trial reported on the number of participants that died from suicide at 36 months (Bauer 2006). However, as only one suicide was reported in the usual care arm, and none in the intervention arm, comparative analysis is not appropriate.

2.14.2 Number of participants that died from natural causes - 36 months

One trial reported on the number of participants that died from natural causes at 36 months (Bauer 2006). There was no clear difference between care groups (RR 1.48, 95% CI 0.62 to 3.53; 1 study, 330 participants).

The exact point at which participants died was not reported. We have assumed a denominator of N = 148 in the intervention group and N = 158 in the usual care group for all outcomes apart from death and leaving the study early (subtracting the numbers leaving the study early from those randomised).

2.14.3 Number of participants that died from suicide - 12 months

One trial reported on the number of participants that died from suicide at 12 months (Chatterjee 2011). However, as only one suicide was reported in each allocated group, comparative analysis is not appropriate.

2.14.4 Number of participants that died from natural causes - six months

One trial reported on the number of participants that died from natural causes at six months (Chwastiak 2018). However, as the study reported only one death in the usual care arm, and none in the intervention arm, comparative analysis is not appropriate.

2.14.5 Number of participants that died (all causes) - 36 months

One trial reported on the number of participants that died from suicide at 36 months (Kilbourne 2013). However, as no deaths were reported in the usual care arm, comparative analysis is not appropriate.

2.15 Service use outside of mental health (i.e. primary care, emergency services, walk-in centres, social services)

No data available.

2.16 to 2.17 Cost of treatment

See Analysis 2.16.

2.16.1 Cost of treatment - at 36 months

One study showed that there were no statistically significant differences in direct intervention (all-treatment) costs at three-year follow-up (Bauer 2006). Mean intervention three-year costs were USD 61,398 (95% CI USD 52,037 to 71,787) compared with USD 64,379 (95% CI USD 55,031 to 73,695) in costs for standard care (MD (USD 1000) -2.98, 95% CI -16.93 to 10.97; 1 study, 306 participants). Standard deviations were imputed from the figures reported by the study authors.

2.17 Cost of treatment (international dollars (Int$)) - up to 12 months

See Analysis 2.17.

One study showed that the collaborative care treatment was more expensive than facility-based care (MD international dollars (Int$) 493.00, 95% CI 345.41 to 640.59) (Chatterjee 2011).

2.18 Experience of care/satisfaction (participant/carer/staff)

No data available.

2.19 Attrition/leaving the study early

See Analysis 2.19.

2.19.1 Attrition/leaving the study early (lost to follow-up) - six months

We found three studies that reported attrition at six months (Chwastiak 2018; Salman 2014; van der Voort 2015). No clear difference between collaborative care and usual care was observed in the short term (RR 1.39, 95% CI 0.76 to 2.55; I^2 = 0%; 3 studies, 235 participants).

2.19.2 Attrition/leaving the study early (lost to follow-up) - 12 months

We found three studies that reported attrition at 12 months (Chatterjee 2011; Chwastiak 2018; van der Voort 2015). No clear difference between collaborative care and usual care was observed in the medium term (RR 1.11, 95% CI 0.77 to 1.58; I^2 = 0%; 3 studies, 504 participants).

2.19.3 Attrition/leaving the study early (lost to follow-up) - at 24 months

We found one study that reported attrition at 24 months (Kilbourne 2013). No clear difference between collaborative care and usual care was observed in the long term (RR 1.19, 95% CI 0.74 to 1.92; 1 study, 118 participants).

2.19.4 Attrition/leaving the study early (lost to follow-up) - at 36 months

We found one study that reported attrition at 36 months (Bauer 2006). No clear difference between collaborative care and usual care was observed in the long term (RR 1.71, 95% CI 0.77 to 3.79; 1 study, 330 participants).

Process/delivery outcomes

Components of collaborative care delivered

Two studies measured the components of collaborative care delivered (Kilbourne 2012; van der Voort 2015).

Kilbourne 2012 reported measuring the number of self-management sessions offered by the care manager and attended by patients, the number of calls made by the care manager and those completed by the patient, the number of completed registry entries on each patient, the number of focus points covered by the interventionist in the sessions and the number of follow-up contacts.

van der Voort 2015 reported that after 12 months almost 80% of patients randomised to collaborative care reported using a relapse

prevention plan, attended a psycho-education course (84%), used a Life Chart (55%), had relatives involved in treatment (86%) and received one or more sessions of Problem Solving Treatment (PST) (72%).

Measures of interprofessional collaboration

Only one study measured interprofessional collaboration (Kilbourne 2012).

Interventionist registry data indicated that the health specialist had a mean number of 1.2 (SD 1.0) and 0.3 (SD 0.6) contacts per patient with their mental health and primary care providers, respectively.

Measures of adherence to manual/algorithms/guidance

No studies measured adherence to manual/algorithms/guidance.

Measures of change in management (number of contacts, referral rates, prescribing patterns and appropriateness)

One study measured the change in the management of conditions (van der Voort 2015) and reported no difference in total number of contacts with the nurse and psychiatrist between the two treatment conditions. 'Care consumption' was measured in both groups with the Trimbos and iMTA Questionnaire for Costs Associated with Psychiatric Illness, to register elements of treatment actually delivered in each group.

One study calculated the mean number of contacts with a treating psychiatrist as 10 (95% CI 9.53 to 10.89) in the intervention group and 8 (6.98 to 9.11) in the control group (Chatterjee 2011).

Measures of change in other health services provided

One study measured service utilisation using a chart review (Kilbourne 2012) and reported no significant differences in utilisation between the Life Goals Collaborative Care (LGCC) and enhanced treatment as usual groups over the 12-month study period.

Measures of continuity (relational, information, longitudinal)

One study measured continuity and reported a critical service encounter index of 8% (quartiles 8 and 11), representing excellent continuity (Bauer 2006).

The 'critical service encounter' index is underpinned by the premise that unscheduled appointments should only be provided by a member of the team delivering the intervention. An index of 10% or less was indicative of excellent continuity (calculated as number of unscheduled appointments with a provider outside of the intervention team divided by total number of appointments). Data on each of these parameters was fed back to each site on a monthly basis via newsletter or conference call.

Measures of healthcare professional behaviour and knowledge (improvement in knowledge/skills, attitudes/acceptability, retention rates, absenteeism, healthcare professional time, prescribing and management of risk factors)

No studies measured healthcare professional behaviour.

Mean percentage of case management contacts

Four studies measured case management contacts (Chatterjee 2011; Chwastiak 2018; Kilbourne 2012; Kilbourne 2013).

Kilbourne 2012 measured the brief case management implementation by reviewing care manager contacts, and estimated the total time the care manager spent on the case management.

Kilbourne 2013 conducted a post hoc analysis to determine whether variation in health specialist-provider care management contacts might have explained changes in outcomes over time and reported that the number of care management contacts was not associated with statistically significant changes in systolic (SBP) or diastolic blood pressure (DBP).

Chwastiak 2018 measured the number of nurse care manager visits (mean 4.9) and the mean duration of treatment with the care manager (14.8 weeks; range 9 to 27).

Chatterjee 2011 reported the mean number of sessions with community health workers that were received by participants in the collaborative care group as 17.97 (SD 7.12) (95% CI 16.94 to 19.00).

Mean percentage of intervention (delivered as part of collaborative care) contacts

Chatterjee 2011 reported that 169 (90%) participants received the predefined minimally effective 12 sessions.

Kilbourne 2013 reported that the majority (68%) completed at least three of the four self-management sessions and an adequate number of follow-up contacts over the 12-month intervention phase (mean 4.6, SD 3.6). Adequate fidelity to LGCC was defined as: mean percentage of self-management sessions attended by patients is ≥ 80% (average of 4 out of 5 sessions attended), mean percentage of session topics covered in lessons is ≥ 80% and mean percentage of completed number of care management contacts to patients is ≥ 65% (mean number of 4 out of 6 required contacts).

Mean percentage of session topics covered in training/education

No studies measured session topics covered in training/education.

Other measures (not pre-specified)

Bauer 2006 reported the median monthly caseload at 47 (quartiles 41 and 48; each site was expected to manage a caseload of 45 to 50 patients). The Life Goals Program was completed within year one by 78% of the sample (quartiles 74 and 82).

Sensitivity analyses

3.1 Mental state: clinically important change (binary) - 12 months (sensitivity analysis - assumptions for attrition)

See Analysis 3.1.

In section 1.2 of this section, we report that we found one study in which mental state was reported as a binary outcome at 12 months, where the number of participants experiencing an improvement of 20% or more on the PANSS overall score was reported (Chatterjee 2011). There was no evidence of a difference in mental state in the collaborative care arm compared to usual care (RR 0.99, 95% CI 0.77 to 1.28; 253 participants). We undertook a sensitivity analysis to impute missing data according to the approach outlined in the methods section (Dealing with missing data), and the results remained robust (similar to those presented above) (RR 0.98, 95% CI 0.77 to 1.25; 282 participants).

3.2 Psychiatric hospital admissions (sensitivity analysis: assumptions for attrition)

See Analysis 3.2.

We performed a sensitivity analyses for our primary outcome of psychiatric admissions according to pre-specified assumptions for attrition (see Methods; Sensitivity analysis).

According to our protocol, the assumption is that 'events' occur at the same rate in those participants who 'leave' the study early.

In the Chatterjee 2011 study, n = 167 of 187 randomised, and n = 86 of 95 randomised participants were followed up in the collaborative care and control groups, respectively. Therefore, 20 participants were lost to follow-up in the intervention group, and nine were lost to follow-up in the control group. According to our pre-specified sensitivity analysis, this would result in an additional participant in the collaborative care arm being classified as having a psychiatric hospital admission, but no additional participants in the control arm.

In the Bauer 2006 study, n = 148 of 167 randomised, and n = 158 of 163 randomised participants were followed up in the collaborative care and control groups, respectively. Therefore, 19 patients were lost to follow-up in the intervention group and five in the standard care group. The study authors report that 52 and 41 patients in the intervention group were psychiatrically hospitalised in years two and three respectively: the corresponding numbers for the standard care group being 74 and 60. Under our assumptions for attrition, in the collaborative care group, seven patients who dropped out would have had a psychiatric admission in year two and five such patients would have had an admission in year three. The corresponding numbers for the control arm are both two.

3.2.1 Up to 12 months

In this sensitivity analysis, 6% of participants in the collaborative care arm had a psychiatric hospital admission versus 1% in the control arm. Whilst there is a suggestion that more psychiatric admissions were observed in the collaborative care arm compared to usual care in the medium term (Analysis 3.2), this result is not statistically significant and there is substantial uncertainty around this estimate due to small numbers of admissions in both allocated groups (RR 5.59, 95% CI 0.73 to 42.64; 282 participants).

3.2.2 Greater than 12 months

In this sensitivity analysis, the proportion of participants psychiatrically hospitalised was statistically significantly lower in the intervention group than the standard care group in years two and three: year two, 35% compared to 47% (RR 0.76, 95% CI 0.58 to 0.99; 330 participants); year three, 28% compared to 38% (RR 0.72, 95% CI 0.53 to 0.99; 330 participants) (Analysis 3.2).

Subgroup analyses

4.1 Quality of life, physical health at six months - subgroup analysis: quality of study

See Analysis 4.1.

We considered pre-specified subgroups defined according to the quality of the study, presenting studies with some concerns of risk of bias in one subgroup ('higher-quality studies') (Kilbourne 2012; Kilbourne 2013), and studies with a high risk of bias in the other ('lower-quality studies') (Mishra 2017; Salman 2014; van der Voort 2015). Within the higher-quality studies subgroup, no clear difference between collaborative care and usual care was observed (SMD 0.04, 95% CI -0.29 to 0.38; I^2 = 0%; 2 studies, 136 participants). Amongst the lower-quality studies, substantial heterogeneity between studies remained, including contradictory directions of effect, and there was no clear difference between collaborative care and usual care (SMD 0.89, 95% CI -0.40 to 2.18; I^2 = 96%; 3 studies, 270 participants). There was no statistically significant difference between subgroups (P = 0.21).

4.2 Quality of life, mental health at six months - subgroup analysis: quality of study

See Analysis 4.2.

We considered pre-specified subgroups defined according to the quality of the study, presenting studies with some concerns of risk of bias in one subgroup ('higher-quality studies') (Kilbourne 2012; Kilbourne 2013), and studies with a high risk of bias in the other ('lower-quality studies') (Mishra 2017; Salman 2014; van der Voort 2015). Within the higher-quality studies subgroup, no clear difference between collaborative care and usual care was observed (SMD 0.16, 95% CI -0.17 to 0.50; I^2 = 0%; 2 studies, 136 participants). Amongst the lower-quality studies, substantial heterogeneity between studies remained, including contradictory directions of effect, and no clear difference was observed between collaborative care and usual care (SMD 1.09, 95% CI -0.42 to 2.59; I^2 = 97%; 3 studies, 270 participants). There was no statistically significant difference between the subgroups (P = 0.24).

4.3 Quality of life, physical health at six months - subgroup analysis: variations in implementation of the collaborative care intervention and healthcare systems

See Analysis 4.3.

Pre-specified subgroup analysis allowed for exploration of variations in implementation of the collaborative care intervention. As such, we considered subgroups (explicitly defined post hoc) according to whether collaborative care was delivered by a pharmacist in liaison with a psychiatrist (Mishra 2017; Salman 2014) or without pharmacy (Kilbourne 2012; Kilbourne 2013; van der Voort 2015). This was because these pharmacy interventions indicated collaborative care with less complexity, narrowing the focus on improving medication management. Although substantial heterogeneity remained in the pharmacy collaborative care subgroup, the direction of the intervention effects is consistent. Within this subgroup, there is evidence of a between-group difference in favour of the collaborative care group (SMD 1.48, 95% CI 0.21 to 2.75; I^2 = 92%; 2 studies, 176 participants), although both studies were rated as high risk of bias and there is also evidence of heterogeneity between the studies. In the subgroup without pharmacy intervention, there is no evidence of a difference between collaborative care and usual care (SMD -0.09, 95% CI -0.35 to 0.18; I^2 = 0%; 3 studies, 230 participants). There is evidence of a statistically significant difference between the two subgroups (P= 0.02).

4.4 Quality of life, mental health at six months - subgroup analysis: variations in implementation of the collaborative care intervention and healthcare systems

See Analysis 4.4.

As above, our pre-specified subgroup analyses allowed for exploration of variations in implementation of the collaborative care intervention. As such, we considered subgroups (explicitly defined post hoc) according to whether collaborative care was delivered by a pharmacist in liaison with a psychiatrist (Mishra 2017; Salman 2014) or without pharmacy (Kilbourne 2012; Kilbourne 2013; van der Voort 2015). In the pharmacy collaboration subgroup, although substantial heterogeneity remained, the results of the meta-analysis are presented nonetheless because the direction of the intervention effects is consistent. Within this subgroup, there is evidence of a between-group difference in favour of the collaborative care group (SMD 1.79, 95% CI 0.36 to 3.21; I^2 = 93%; 2 studies, 176 participants); however, both studies were rated as high risk of bias and heterogeneity was observed. In the subgroup without pharmacy intervention, there is no evidence of a difference between collaborative care and usual care (SMD -0.01, 95% CI -0.33 to 0.31; I^2 = 32%; 3 studies, 230 participants). There is evidence of a statistically significant difference between subgroups (P = 0.02).

DISCUSSION

Summary of main results

Main results

The aim of this review was to assess the effectiveness of collaborative care in comparison with standard care for people with severe mental illness (SMI) who are living in the community.

Shorter-term outcomes (up to six months)

No significant shorter-term effects were observed in the included studies in favour of collaborative care for the following secondary outcomes: quality of life, mental state (general symptoms, positive symptoms, negative symptoms, depressive symptoms and manic symptoms) and number of people admitted to psychiatric hospital. No data were available for the personal recovery outcome. We noted large variations in the interventions delivered, which may have resulted in heterogeneity. In an attempt to explain this heterogeneity, we created a subgroup of studies that we categorised as 'pharmacy-based collaborative care'. The collaborative care intervention in these studies was delivered in India (a lower middle-income country). This was based on the roles and focus of the pharmacist being narrower than the 'case manager' role in the other studies. Data from this subgroup indicated that quality of life (i.e. physical health and mental health components up to six months) was better for those in the collaborative care group compared to usual care (Mishra 2017; Salman 2014). Although there was still substantial heterogeneity in this subgroup, the direction of the intervention effects is consistent in both studies; however, these were low-quality studies. In the remaining non-pharmacy studies there was no heterogeneity and a consistent null effect.

No significant short-term effects were present in favour of collaborative care for the following secondary outcomes: physical health status (systolic and diastolic blood pressure, BMI, high-density lipoprotein (HDL), low-density lipoprotein (LDL), waist circumference, total cholesterol, triglycerides), medication adherence (patient-reported - DAI-10), social functioning, costs of the intervention, mortality and attrition.

An improvement in clinician-rated medication adherence was observed in one study (Mishra 2017).

Medium-term outcomes (7 to 12 months)

No significant medium-term effects were observed in favour of collaborative care for the following primary outcomes: quality of life (mental health and physical health domains), mental state (binary), mental state (depressive symptoms), mental state (manic symptoms) or in the risk of being admitted to psychiatric hospital at 12 months.

No significant medium-term effects were observed in favour of collaborative care for the following secondary outcomes: hospital admissions for non-psychiatric conditions, physical health status (systolic blood pressure, HDL, LDL, total cholesterol, triglycerides), medication adherence (patient reported - DAI-10), social functioning, mortality and attrition.

The medium-term outcomes in this review indicate a small improvement in mental state at 12 months (general symptoms) but other mental state outcomes (described above) indicated no difference. We found one study that measured general mental state symptoms using the PANSS (Chatterjee 2011). We considered this outcome to have a low risk of bias, but also a low certainty of evidence. There was a clear difference in general symptoms in mental state between collaborative care compared with usual care in the medium term (SMD -0.27, 95% CI -0.53 to -0.01, 253 participants). However, there was no difference in positive/negative/overall symptoms; there was no difference shown in the proportion of participants who had a reduction of more than 20% in overall symptoms (85 (51%) versus 44 (51%); P = 0.89).

There was an indication from one study that there were more psychiatric admissions in the collaborative care arm compared to usual care in the medium term (12 months) (RR 5.15, 95% CI 0.67 to 39.57; 253 participants) (Chatterjee 2011). We considered this outcome to have a low risk of bias, but also a low certainty of evidence. Additionally, this result was not statistically significant and there was substantial uncertainty around this estimate due to the small numbers of admissions in both allocated groups, making it impossible to draw meaningful conclusions.

There was also a clear difference in waist circumference in the medium term, indicating that the collaborative care group had lower waist circumferences (SMD -0.39, 95% CI -0.75 to -0.03; I^2 = 9%; 2 studies, 136 participants).

Data from one study also indicated that the collaborative care treatment was more expensive than facility-based care (MD international dollars (I$) 493.00, 95% CI 345.41 to 640.59) (Chatterjee 2011). These health economic findings showed that costs in the intervention group were on average greater than those in the control group, and that about a third of these additional costs were attributable to supervision. The average greater cost for participants in the intervention group was almost INR 9500 (roughly I$ 500).

No clear difference between collaborative care and usual care was observed in the medium term for social functioning/disability (Chatterjee 2011; Kilbourne 2012; Kilbourne 2013; van der Voort 2015). However, one study reported a post hoc analysis of disability at 12 months as a binary outcome, defined as an improvement of at least 20% on the IDEAS score (Chatterjee 2011); we considered this outcome to be of low risk of bias but also low certainty of evidence. This analysis showed that more participants in the intervention

arm improved by this extent compared to the control arm (75/167 (48%) versus 28/86 (35%)) (RR 1.38, 95% CI 0.97 to 1.95; 1 study, 253 participants), although this result was not statistically significant.

Longer-term outcomes (over 12 months)

No significant longer-term effects were observed in favour of collaborative care for the following secondary outcomes: overall quality of life physical health and mental health components, mental state (general symptoms, positive symptoms, negative symptoms, depressive symptoms, manic symptoms); physical health status (systolic and diastolic blood pressure, HDL, HDL, total cholesterol, triglycerides), medication adherence (patient-reported - DAI-10), social functioning, mortality and attrition. No data were available for the personal recovery outcome.

The longer-term outcomes in the review indicate some reductions in psychiatric and other admissions. One study found that collaborative care delivered to US veterans with bipolar disorder (I or II) reduced psychiatric admissions in year two in comparison to standard care (RR 0.75, 95% CI 0.57 to 0.99; 306 participants) and in year three (RR 0.73, 95% CI 0.53 to 1.01, 306 participants) (Bauer 2006). We found this study outcome to have a low risk of bias. We carried out a sensitivity analysis to test the assumption that those participants who had withdrawn from the trial had experienced an outcome by the end of the trial (i.e. psychiatric admission) at the same rate at which those followed up experienced the outcome, by allocated group. The results of the sensitivity analysis were broadly consistent. The results show that the collaborative care intervention reduced the proportion of psychiatric admissions in year two and other admissions in year three. However, the reporting of admissions by year (rather than cumulatively) does make it more difficult to assess the mid- to long-term impact that collaborative care has on the group as a whole; it is not possible to know whether it is the same people who are hospitalised year-on-year or whether some patients never get admitted to hospital.

Longer-term data in one study captured symptoms at 24 months using the Internal State Scale (ISS) (Kilbourne 2013). The data from this study indicated that collaborative care resulted in a reduction of manic symptoms at the 24-month follow-up (SMD -0.36, 95% CI -0.82 to 0.10; 75 participants), although this difference was not statistically significant.

Overall summary of results

In summary, current limited evidence suggests that collaborative care interventions could help to improve general mental state symptoms in the medium term (Chatterjee 2011) and reduce psychiatric admissions in the longer term (Bauer 2006). However, there was only evidence from one study for both of these outcomes.

Data from one study in India indicated that collaborative care was more expensive than usual care (Chatterjee 2011). Data from one study in the US indicated that collaborative care was less expensive than usual care (Bauer 2006). However, there is uncertainty around these results due to low-certainty evidence and the variability in the interventions delivered.

Collaborative care is a complex intervention with multiple components that require a systems-level change and a different way of working. The variation in the implementation of collaborative care across included studies can be seen in Table 2 and Table 3.

There is some evidence to suggest that, in contrast to standard care, collaborative care may improve quality of life (both physical and mental health aspects) in the shorter term (at six months) when a pharmacist is an integral part of the collaborative care multidisciplinary team. However, these findings are based on the results of two low-quality studies conducted in India, a low-income country where access to mental health care may be more limited for people on standard care. This may explain why the impacts on quality of life are limited to just these studies. Assessment of the certainty of evidence suggests that caution should be applied in using the results of this review if choosing collaborative care to improve quality of life, mental state or short-term psychiatric admissions for people with a diagnosis of severe mental illness.

In relation to physical health outcomes, collaborative care was found to improve both waist circumference and BMI at 12 months (Kilbourne 2012; Kilbourne 2013). However, the Life Goals Collaborative Care intervention was designed to reduce the risk factors for cardiovascular disease, through improved control of psychiatric symptoms and increased positive health behaviours, as well as improved co-ordination of physical and mental health care. These two studies have been assessed as having some concerns in terms of risk of bias in relation to other outcomes. Overall, we advise caution in utilising the results in relation to physical health outcomes in making decisions regarding care.

Using clinician-rated medication adherence measures, collaborative care was found in one study to be more effective than standard care in increasing medication adherence in the short term (six months); this was statistically significant (Mishra 2017). This was also the case for medication adherence (although it should be noted that the desirability of medication adherence is a matter currently debated in the literature (Healy 2016; Kinderman 2014; Moncrieff 2013; Moncrieff 2015; Wunderink 2017)). These studies were categorised as type B collaborative care (comprised of one to three of the components defined by Gunn 2006). We would advise caution in using these results to inform care decisions. Furthermore, patient-reported adherence was assessed in one study and was not statistically significant (van der Voort 2015).

In the short term (12 months), collaborative care was found to be more expensive in one study and slightly cheaper in the longer term. However, again, these outcomes are each based on one study, neither of which were categorised as type A collaborative care. We would issue caution in using these cost data to decide on appropriate allocation of funds for care.

There are no data reported for our adverse effect outcome, as trial authors did not describe any outcomes as adverse effects. However, many of the outcomes we have measured could be considered a proxy measure of adverse effects, for example psychiatric admissions.

Overall, this review found that collaborative care may be associated with an advantage compared to standard care in relation to long-term psychiatric and non-psychiatric hospital admissions, medium-term waist circumference and BMI outcomes, and short-term clinician-reported medication adherence. Collaborative care was found to be more cost-effective in the long run. Patients and clinicians should be aware that the outcomes in this review are predominantly based in part on studies that were not categorised as type A collaborative care, and that the nature and purpose of the interventions included varied. Furthermore, many of these

conclusions are drawn from only one or two studies. Combined with low-certainty evidence for many of the review outcomes, we advise caution in utilising the information in this review to assess the effectiveness of collaborative care.

1. Completeness

This review included eight RCTs of collaborative care for people with SMI. This presents a limited amount of evidence and further trials are required to evaluate the effect of collaborative care on physical and mental health outcomes.

Study designs

Seven of these RCTs were individually randomised and one was a cluster-randomised study (van der Voort 2015). Cluster-randomised studies can be more susceptible to biased recruitment (Hahn 2005), and in the van der Voort 2015 study this was evidenced through an imbalance of clinical characteristics between intervention and control participants. However, the lack of masking of the researchers, clinicians and potential participants prior to randomisation likely contributed to this.

Four studies were multi-centre trials (Bauer 2006; Chwastiak 2018; Kilbourne 2012; Kilbourne 2013). Two studies were pilot trials (Chwastiak 2018; Kilbourne 2012).

1.1 Duration of follow-up

The duration of follow-up for the included studies varied significantly. One study had a three-month follow-up (Chwastiak 2018), two studies had six months (Mishra 2017; Salman 2014), three studies had 12 months (Chatterjee 2011; Kilbourne 2012; van der Voort 2015), one had 24 months (Kilbourne 2013) and one had a 36-month follow-up period (Bauer 2006). Three of these studies had a short follow-up period and, as a result, we are unable to comment on the long-term efficacy and impact of collaborative care (Chwastiak 2018; Mishra 2017; Salman 2014). We suggest that more, longer-term follow-up studies are needed, with a minimum of 12 months follow-up, ideally longer, to enable us to better understand the impact of collaborative care. This is in part because collaborative care is a re-structured way of working with service users, which takes time to take effect. People would access the intervention on a long-term basis if it was to be implemented in healthcare systems, and therefore the benefits are more likely to become apparent over a longer period of time accessing the service. Additionally, the long-term, chronic nature of severe mental illness means that changes in quality of life, recovery and mental state are likely to take longer to manifest than in other mental health populations, such as those with depression, due to problems of agency, identity and hope (Tew 2012).

1.2 Coverage of outcomes

Although individual trials have employed validated tools to measure specific outcomes, there is a particular problem with the interpretation of the scales. There appears to be confusion with authors aligning the correct scale to a specific version of a tool and some are being used interchangeably. This was particularly noted with medication adherence scales not being interpreted according to the guidance for specific versions. This presented a problem in that the data did not fall into the range of values permissible by the scale; therefore, the outcome could not be included in our analysis. Additionally, some trial authors failed to reference specific tools or versions of the tools used in their data outcomes. One of our primary outcomes, personal recovery, was not measured by any of the studies despite being highlighted as important by those with ongoing mental health problems (Retzer 2020). Furthermore, satisfaction/experience of care was not reported as an outcome in any of the included studies.

2. Applicability

Only two of our included studies included Gunn's four 'core' components of collaborative care and are therefore categorised as 'type A' collaborative care (Chwastiak 2018; Kilbourne 2013). The results of the other studies therefore lack direct applicability. The most common reason for studies not meeting the definition was the lack of involvement of primary care. This may be in part be due to studies failing to describe their intervention in sufficient detail; for example, we did not identify a published TIDieR checklist (Hoffman 2014) for any of the studies, but it also likely reflects the different contexts in which the interventions are being delivered.

2.1 Origin

The included studies were located in a variety of countries and settings across the world. Four studies were located in the US (Bauer 2006; Chwastiak 2018; Kilbourne 2012; Kilbourne 2013), three of which were within Veteran Affairs healthcare centres (Bauer 2006; Kilbourne 2012; Kilbourne 2013). Three studies were conducted in India (Chatterjee 2011; Mishra 2017; Salman 2014) and one in the Netherlands (van der Voort 2015). Future work is needed to determine generalisability across different settings in a range of countries (both high-income and low-income) with different organisational, provider and patient-level characteristics.

The seven ongoing studies are taking place in the US, Australia and England (see Ongoing studies). Usual care is very different in these countries and with most studies failing to describe what usual care consists of, we encourage researchers to provide sufficient detail.

2.2 Population

The study participants in the included studies varied greatly, with females representing between 6% (Bauer 2006) and 63.8% (van der Voort 2015) of participants. The median age of participants in studies varied between 35.6 (10.2) (Chatterjee 2011) and 53.1 (10.6) (Kilbourne 2013). Three studies included participants with schizophrenia and schizoaffective disorders (Chatterjee 2011; Chwastiak 2018; Salman 2014), one study bipolar disorder type 1 and 2 (Bauer 2006), three studies bipolar disorder type 1, 2 and bipolar not otherwise specified (NOS) (Kilbourne 2012; Kilbourne 2013; van der Voort 2015) and one study included people with diagnoses of schizophrenia or bipolar (Mishra 2017).

Three studies failed to report ethnicity (Mishra 2017; Salman 2014; van der Voort 2015) and, for those that did, the rates of minority participants varied between 5.1% and 60%. In four other studies (with the exception of Chatterjee 2011), ethnicity was reported in a very broad manner, categorising ethnicities as 'minority groups' or reporting just one ethnic minority group. This is poor practice, especially when people from ethnic minority groups often have differing risks of developing SMI compared to local populations and are often over-represented in mental health services. We recommend that future authors endeavour to collect accurate demographic data with regard to ethnicity and report this thoroughly.

The characteristics of participants were highly heterogeneous between studies and are not necessarily representative of the individuals who may be eligible for collaborative care. Seven out of eight studies reported numbers of potential participants eligible for the study who declined to participate, with the exception of Mishra 2017, which failed to report this for their participants with schizophrenia. These varied substantially, from 0% (Chwastiak 2018) to 73% (Bauer 2006) of potential participants approached. Three other studies reported reasonably low declination rates, from 24% to 27% (Chatterjee 2011; Kilbourne 2012; Kilbourne 2013), while two further studies had above-average declination rates: concerning levels, 37% (Salman 2014) and 41% (van der Voort 2015), when compared to the average rate in a recent review (Lin 2021). The high proportion of potential participants declining participation in three of the seven studies that reported this may suggest that the samples are highly selective.

Quality of the evidence

We assessed the certainty of evidence in relation to the primary outcomes of this review using the GRADE system (GRADEpro). GRADEpro prompts review authors to consider the risk of bias, inconsistency, indirectness, imprecision, publication bias and effect size to rate the certainty of evidence as very low, low, moderate or high in relation to each outcome (Schünemann 2020). The details of our assessment of risk of bias can be found in Assessment of risk of bias in included studies. Other elements of evidence certainty are discussed below. The overall results of our assessment are summarised in Summary of findings 1.

We were also able to utilise stakeholder consultation to select outcomes that were important to those working with and living with SMI diagnoses.

Inconsistency

Where substantial differences in the estimated effect of collaborative care in relation to a particular outcome are observed, this may be indicative of an issue associated with an inconsistency in the results. Statistically, this heterogeneity is estimated using the I^2 statistic. Inconsistency of evidence may be concluded when the I^2 is large and/or when there is inconsistency in the direction of effects between studies (Schünemann 2020). We did not identify any problematic inconsistency in the evidence in any of the primary outcomes assessed.

Indirectness

Caution should be used when utilising indirect evidence as the results might not be applicable to the population, intervention or outcome of interest to the review question. Evidence is considered indirect if it is gathered in relation to a different population or different intervention from the one considered in the review question. It is also considered indirect if it is gathered using an outcome that does not directly measure the concept stated in the review outcomes. Indirectness can be assessed as not serious, serious or very serious (Schünemann 2020). All of the evidence in relation to the primary outcomes is direct in that the concept of interest was directly measured and that the population of the study was those with an SMI diagnosis (either bipolar or schizophrenia). However, as noted in Table 2, most of the studies included in our review used an intervention that does not meet our 'core' definition of collaborative care (Gunn 2006). Additionally, we have used some study outcomes to indirectly measure concepts in our primary outcomes: number of psychiatric admissions as a measure of intervention safety and quality of life physical health sub-domain as a measure of physical health. Therefore, we have assessed indirectness as 'serious' in relation to all outcomes presented in the GRADE table.

Imprecision

Evidence is considered imprecise when the number of participants is low in relation to the variation in result, resulting in the inability to detect a difference that may be deemed clinically relevant with enough certainty to conclude that said difference is statistically significant. In line with GRADE guidance, we calculated the optimal information size to aid in determining for which outcomes imprecision was an issue (Schünemann 2020). Typically, larger sample sizes are required for binary outcomes. As a result, we downgraded the certainty of evidence for binary outcomes (mental state (schizophrenia symptoms) at 12 months, and psychiatric admissions at 12 months) on the basis of imprecision, but none of the continuous outcomes. Further studies of collaborative care in relation to these outcomes would improve the precision of evidence.

Other considerations

Publication bias

Publication bias is the concept that undesirable results are not disseminated, creating a skewed evidence base. Our searches did not yield any trial registrations or study protocols that would indicate research taking place that had not been published due to undesirable results. This suggests a low risk of publication bias.

Effect size

The certainty of evidence can be upgraded if there is a large effect size in relation to the outcome, but other certainty of evidence factors should be taken into account when judging effect size for binary outcomes (Schünemann 2020). We considered the effect size in relation to hospital admissions at 12 months to be very large, regardless of other factors, as the RR exceeds 5.0. The guidance suggests that this should result in an increase in the certainty of evidence. However, due to the very small number of events (particularly in the control arm) resulting in substantial uncertainty in the estimation and therefore a very large confidence interval, we have not upgraded the certainty of evidence for this outcome. We did not increase the certainty of evidence for any other outcomes on the basis of effect size.

Overall certainty of evidence

We rated the certainty of the evidence for binary schizophrenia symptoms, disability (proxy for social functioning) and psychiatric admissions at 12 months as low. For physical and mental quality of life at 12 months and for depressive and manic bipolar symptoms at 12 months, we considered the certainty of evidence to be very low. The low certainty of evidence in the included studies makes it difficult to draw strong conclusions in relation to the impact of collaborative care on the outcomes presented in our summary of findings table. This highlights the need for further good-quality studies, with particular attention to the following: participants and researchers masked prior to randomisation, participant retention strategies to reduce dropout, sufficient sample sizes, and publication of and adherence to analysis plans. The nature of the

interventions utilised in these studies, and/or the detail of the description of these interventions, makes it impossible for us to use this evidence to comment on the effectiveness of collaborative care that meets Gunn's four core elements (Gunn 2006). Further evidence, utilising studies that use interventions meeting these core elements, is required (see Table 5 for suggested future study design). 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 (Schünemann 2020). Trialists following the CONSORT recommendations will also greatly assist in the future synthesis of review data (CONSORT 2010a; CONSORT 2010b).

Potential biases in the review process

One of the potential problems with systematic reviewing is that relevant studies can be missed. The Cochrane Schizophrenia searches for studies used a specialised register, which is compiled from systematic searches of major databases, and handsearching of relevant journals and conference proceedings. The search strategy includes terms to describe schizophrenia and schizophrenia-type disorders, severe mental illness and psychosis, but not 'bipolar disorder'. To limit this potential bias in the search strategy, we supplemented the electronic searches with reference list searches and contacted experts in the field of SMI and collaborative care, who were asked to identify published and unpublished research that they were either involved in or aware of (see Searching other resources; Figure 1). The search strategy is published online. Secondly, studies that were identified by experts in the field as relevant were excluded if the intervention was not described as 'collaborative care' by the study authors in the papers or reports (for example, Byng 2004; Druss 2001; Druss 2010; Simon 2006). This approach was taken under guidance from Cochrane in an effort to reduce the degree of variation between studies. Finally, the authors of the review are researchers who are active in the development, evaluation or implementation of collaborative care for people with SMI (Druss 2001; Druss 2010; Byng 2023), but currently do not have any studies that were eligible for inclusion in this updated review.

Agreements and disagreements with other studies or reviews

This is an update of the first review of collaborative care for people with SMI (Reilly 2013). Future updates will include the ongoing studies (Characteristics of ongoing studies). We noted in our previous review the lack of other reviews in this area. Others have highlighted the same (Goodrich 2013; Planner 2016). This situation has not changed; we have not identified any other comprehensive reviews.

AUTHORS' CONCLUSIONS

Implications for practice

Collaborative care aims to provide a more patient-centred and integrated system of care. This systematic review has synthesised evidence from eight studies, spanning two decades of research on collaborative care interventions for people with severe mental illness in a range of countries and settings. The evidence was of low or very low certainty and limited data were available for our primary outcomes. Our confidence in these findings is limited due to concerns about the certainty of evidence.

Although the literature we examined in our original review showed that collaborative care aims to foster closer working relationships between primary care and specialist health care, the aims of the interventions evaluated in the trials included in this review varied substantially. None included primary care directly in their interventions.

All interventions had a team comprised of a mental health professional and at least one other professional. One study reported the inclusion of a primary care professional in the multidisciplinary team (Chwastiak 2018). According to our definition, pharmacists are providers of primary care; however, the trials where a pharmacist was included were based in a secondary care setting (Mishra 2017; Salman 2014).

This updated review, which includes eight studies (an increase of seven since the previous review), suggests that collaborative care may offer some benefit in contrast to standard care in terms of reducing psychiatric admissions in the longer term (at 24 months) and other non-psychiatric admissions (at 36 months). We found the certainty of evidence to be low. In contrast, collaborative care may slightly increase psychiatric admissions in the medium term (up to 12 months) and non-psychiatric admissions in the medium term (two years), but these were not statistically significant results, and it was not clear if this outcome also included psychiatric admissions. It is possible that collaborative care may contribute to greater admissions in particular healthcare settings because individuals may be more closely monitored or because the intervention may actively facilitate admissions when needed.

There were no data available regarding the effect of collaborative care on personal recovery.

For patients with schizophrenia, there is currently insufficient evidence to determine whether collaborative care approaches improve mental health outcomes.

Assessment of the certainty of evidence suggests that caution should be applied in using the results of this review if choosing collaborative care to improve quality of life, mental state or short-term psychiatric admissions for people with a diagnosis of severe mental illness.

Implications for policy

Healthcare policy in the UK recommends that primary care and specialist services integrate care more effectively (NICE 2009), with the compulsory introduction of integrated care systems from April 2021 (NHS England 2019b), and an overhaul of community mental health services to prevent people, especially those with severe mental illness (SMI) diagnoses, falling through gaps in care by moving away from silo-ed, criteria-led services (NHS England 2019c).

In the US, the bulk of care for individuals with severe mental illness is provided by speciality mental health providers rather than general practitioners. In collaborative care models tested in these settings, primary care providers serve in a liaison role that is analogous to the role of specialty mental health providers in primary care-based collaborative care models. Many of these interventions have primarily focused on improving medical outcomes such as cardiometabolic parameters rather than mental health outcomes (McGinty 2021).

However, this review demonstrates that these recommendations for implementing collaborative care are not currently matched by a robust, high-quality, evidence base concerning effectiveness in improving patient outcomes or cost-effectiveness. Funders should support high-quality research that investigates the effectiveness of collaborative care for people with SMI.

Policy-makers and practitioners may struggle to use the evidence in this review to inform decisions about whether or not to recommend or provide collaborative care for people with SMI. Collaborative care utilises a multi-professional approach to care, including care provided by both a primary physician and a senior mental health professional, a structured management plan, scheduled follow-ups and interprofessional communication. Case management involves a health worker taking responsibility for follow-up care to assess patient adherence to treatment, monitor progress and take action when treatments are unsuccessful. Only one of the trials in this review met this definition of collaborative care; interestingly, this was a pilot trial of participants with a physical comorbidity: patients with poorly controlled type II diabetes (Chwastiak 2018). In comparison, all 79 included studies in a Cochrane review evaluating collaborative care review for adults with depression, anxiety or both met this definition (Archer 2012). Archer showed evidence to support collaborative care (for symptoms of anxiety and depression) at six months, 12 months and 24 months: studies were conducted in the UK, US, Germany, the Netherlands, Canada, Chile, India and Puerto Rico, and almost all were from high-income countries. Others have reviewed the impact of collaborative care on depression and co-morbid chronic physical illnesses (e.g. Ekers 2013; Huang 2013). If policy-makers had even a fraction of this amount of evidence for people with severe mental illness, it is likely that it would be easier to draw clearer conclusions. Specifically, evidence from other collaborative care studies of depression, where improving outcomes for co-morbid physical health problems has been targeted, might be extrapolated to develop interventions for people with SMI, as Katon 2010's TEAMcare study was for the Chwastiak 2018 study included in this review.

Implications for research

This review has synthesised evidence from eight studies, spanning around two decades of research on collaborative care for people with severe mental illness from a range of countries and settings. The evidence was all of low or very low certainty. It is worth noting that there are seven ongoing trials that may meet the criteria for collaborative care, some of which are due to report imminently (Aschbrenner 2019; Battersby 2018; Fields 2019; Hanlon 2014; Happell 2018; Nicole 2018; Byng 2023). More of the same research with the same populations is not likely to substantially improve the evidence base around collaborative care. However, considering the literature identified by this review, we see clear opportunities to improve the evidence base, including: enhancing the quality of trial methodology and reporting, using the term 'collaborative care' consistently around an agreed definition, measuring consistent outcomes and those that matter most to people with SMI, and better understanding which people are most likely to benefit.

1. Enhancing the quality of trial methodology and reporting

Well-designed, conducted and reported RCTs are required to determine the effectiveness of collaborative care for people with serious mental illness diagnoses. A comprehensive process evaluation should be a part of the design of complex intervention trials in this area and would assist with the interpretation of trial outcomes, as would a description of the contents of the intervention, and an explanation of how and why the intervention might work (Craig 2008). Study and intervention manuals should be made available and we also recommend that authors complete a TiDIER checklist (Hoffman 2014), providing important detail on the characteristics of the intervention and assessment of fidelity.

Most services, for example community mental health teams in the UK, are organised to provide care to people living with a severe mental illness, rather than delivering condition-specific models of care. Therefore, trials of collaborative care that include participants with any type of SMI are needed, as opposed to those directed at those with a single condition.

2. Terminology

Terms such as 'collaborative care', 'shared care' and 'integrated care' are used interchangeably to describe different levels of integration between service providers providing health care across a variety of settings. More precise nomenclature would make it easier to identify relevant studies, but until that time, future reviews could seek to identify studies that evaluated interventions that were not described as collaborative care but are comprised of the following four components (as per Gunn 2006):

- A multi-professional approach to patient care. A primary care provider and at least one other health professional or paraprofessional is involved with patient care.
- A structured management plan in the form of evidence-based protocols or guidelines.
- Scheduled patient follow-ups.
- Enhanced interprofessional communication. Enhanced communication could take place through case conference, regular team meetings, case-by-case consultation and written correspondence (e.g. via email or through linked electronic records).

As stated above (Potential biases in the review process), both the search terms and the inclusion criteria excluded studies that were potentially relevant but were not described in the papers as collaborative care by the study authors (for example, Byng 2004; Druss 2001; Druss 2010; Kilbourne 2008; Simon 2006).

3. Choice and reporting of outcome measures

Trial authors should consult relevant core outcome sets (Williamson 2017) before deciding on their outcomes of interest. We are only aware of one relevant core outcome set (for those with bipolar; Retzer 2020); we consulted this when finalising our revised outcomes. Trial authors should consider carefully which outcomes their intervention is likely to affect and why. They should publish their prespecified outcomes in a study protocol prior to the reporting of results; this will enable any subsequent reviews of the literature to include an assessment of the likelihood of selective reporting bias. We need a consensus not just on what outcomes should be measured but also on how they should be measured and what constitutes clinically significant benefit, so that binary outcomes can be reported. Very few outcomes were reported in binary form. A number of outcomes, for example personal recovery and satisfaction, were not measured in any of the studies despite being deemed important.

Only published scales that have been subject to validation (internal and external) should be used. Scales that have been written or adapted by study authors need to be independently validated. This will improve the certainty of the evidence and the ability to utilise the results in making decisions for practice and policy.

4. Future reviews

The difficulties in synthesising trials of complex interventions are not new; however, the complexity of differing interventions, with different interacting components and different aims and outcomes, has made undertaking this review a challenge. It may be useful if future reviews were able to capture both quantitative and qualitative data, for example by utilising an integrated review methodology in line with the Medical Research Council guidance, placing high importance on a qualitative understanding of mechanisms of change (Craig 2008). Furthermore, depending on emerging new evidence, it would be useful for a future update to report on service-related outcomes such as cost-effectiveness. This may facilitate decisions for collaborative care service development.

5. Strengthening the evidence base

We have been limited in the conclusions that can be drawn from this review by the certainty of the evidence. Appropriate sample sizes, refinement in choice of outcomes, longer follow-up, and transparency in the randomisation and allocation process would significantly improve the evidence. Some of the ongoing trials may be more likely to resolve some of these methodological issues.

ACKNOWLEDGEMENTS

The Cochrane Schizophrenia Editorial Base, situated across the University of Melbourne, Australia, the Technical University of Munich, Germany and the University of Nottingham, UK, 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 Mahesh Jayaram and the staff at Cochrane Schizophrenia for all their practical help and advice. Thanks to Farhad Shokraneh and Sarah Dawson for running the searches and to Hui Wu for her editorial support.

Thanks to members of the PARTNERS2 live experience advisory panels who have commented on previous results of the review and contributed to the workstream on the choice of outcome measures, which informed this review.

This review presents independent research funded by the National Institute for Health Research (NIHR) under its Programme Grants for Applied Research programme (grant reference no. RP-PG-0611-20004) and supported by the NIHR Applied Research Collaboration South West Peninsula. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care.

Editorial and peer reviewer contributions

The following people conducted the editorial process for this article:

- Sign-off Editor (final editorial decision): Irene Bighelli, Technical University of Munich
- Managing Editor (selected peer reviewers, collated peer reviewer comments, provided editorial guidance to authors, edited the article): Hui Wu, Technical University of Munich
- Contact Editor (provided editorial guidance to authors): Christine Rummel-Kluge, Leipzig University
- Copy Editor (copy editing and production): Jenny Bellorini, Cochrane Central Production Service
- Information Specialist (search strategy and search results): Farhad Shokraneh, Systematic Review Consultants
- Peer reviewers* (clinical/content review, provided comments and recommended an editorial decision): Nurul Husna Salahuddin, Technical University of Munich; Nikos Christodoulou, University of Thessaly Medical School
- The previous Cochrane Schizophrenia editorial base also supported this work: Co-ordinating Editor, Clive Adams (before 2020), Managing Editor, Claire Irving (before 2020), Assistant Managing Editor, Ghazaleh Aali, University College London (before April 2021)

*Peer reviewers are members of Cochrane Schizophrenia and provided peer review comments, but they were not otherwise involved in the editorial process or decision-making for this article.

REFERENCES

References to studies included in this review

Bauer 2006 *{published and unpublished data}*

Bauer MS, Biswas K, Kilbourne A. Enhancing multiyear guideline concordance for bipolar disorder through collaborative care. *American Journal of Psychiatry* 2009;**166**(11):1244-50.

Bauer MS, Callahan AM, Jampala C, Petty F, Sajatovic M, Schaefer V, et al. Clinical practice guidelines for bipolar disorder from the Department of Veterans Affairs. *Journal of Clinical Psychiatry* 1999;**60**(1):9-21.

* Bauer MS, McBride L, Williford WO, Glick H, Kinosian B, Altshuler L, et al, Cooperative Studies Program 430 Study Team. Collaborative care for bipolar disorder: Part II. Impact on clinical outcome, function, and costs. *Psychiatric Services* 2006;**57**(7):937-45.

Bauer MS, McBride L, Williford WO, Glick H, Kinosian B, Altshuler L, et al, Cooperative Studies Program 430 Study Team. Collaborative care for bipolar disorder: part I. Intervention and implementation in a randomized effectiveness trial. *Psychiatric Services* 2006;**57**(7):927-36.

Bauer MS. Collaborative practice for bipolar disorder: a multisite controlled trial. https://docplayer.net/78847483-Psychiatric-news-philadelphia-the-21st-century-psychiatrist-volume-xxxvii-number-4-february-15-2002.html.

Bauer MS. The collaborative practice model for bipolar disorder: design and implementation in a multi-site randomized controlled trial. *Bipolar Disorders* 2001;**3**:233-44.

Pirraglia PA, Biswas K, Kilbourne AM, Bauer MS, Fenn H. A prospective study of the impact of comorbid medical disease on bipolar disorder outcomes [NCT00007761]. https://clinicaltrials.gov/study/NCT00007761.

Chatterjee 2011 *{published and unpublished data}***Std17994**

Balaji M, Chatterjee S, Koschorke M, Rangaswamy T, Chavan A, Dabholkar H, et al. The development of a lay health worker delivered collaborative community based intervention for people with schizophrenia in India. *BMC Health Services Research* 2012;**12**:42.

Chatterjee S, Leese M, Koschorke M, McCrone P, Naik S, John S, et al. Collaborative community based care for people and their families living with schizophrenia in India: protocol for a randomised controlled trial. Trials 2011;**12**:12. [CSZG: Ref22700]

Chatterjee S, Naik S, John S, Dabholkar H, Balaji M, Koschorke M, et al. Effectiveness of a community-based intervention for people with schizophrenia and their caregivers in India (COPSI): a randomised controlled trial. Lancet 2014;**383**:1385-94. [CSZG: Ref28789]

Graham Thornicroft. Email: graham.thornicroft@kcl.ac.uk July to September 2020.

ISRCTN56877013. A randomised trial comparing the clinical and cost effectiveness of usual, facility based care compared to an additional collaborative community care package for people and their families living with schizophrenia in India. https://doi.org/10.1186/ISRCTN56877013 2016.

Koschorke M, Padmavati R, Kumar S, Cohen A, Weiss HA, Chatterjee S, et al. Experiences of stigma and discrimination of people with schizophrenia in India. Social Science & Medicine 2014;**123C**:149-59. [CSZG: Ref29302]

Thornicroft G. A randomised trial comparing the clinical and cost effectiveness of usual, facility based care compared to an additional collaborative community care package for people and their families living with schizophrenia in India. http://www.controlled-trials.com 2009.

Chwastiak 2018 *{published and unpublished data}*

Chwastiak L. TEAMcare for diabetes in mental health centers. https://clinicaltrials.gov/ct2/show/NCT02011529.

Chwastiak LA, Luongo M, Russo J, Johnson L, Lowe JM, Hoffman G, et al. Use of a mental health center collaborative care team to improve diabetes care and outcomes for patients with psychosis. *Psychiatric Services (Washington, D.C.)* 2018;**69**(3):349-52.

Dr Lydia Chwastiak. Email: lchwast@uw.edu July to September.

Kilbourne 2012 *{published data only (unpublished sought but not used)}*

Amy Kilbourne. Email: amykilbo@umich.edu 7 August 2020.

Kilbourne AM, Goodrich DE, Lai Z, Clogston J, Waxmonsky J, Bauer MS. Randomized controlled pilot study of life goals collaborative care for patients with bipolar disorder and cardiovascular disease risk from community-based practices. *Psychiatric Services* 2012;**63**(12):1234-8.

Kilbourne AM, Post EP, Nossek A, Drill L, Cooley S, Bauer MS. Improving medical and psychiatric outcomes among individuals with bipolar disorder. *Psychiatric Services* 2008;**59**(7):760-8.

NCT01244854. Life goals behavioral change to improve outcomes for veterans with serious mental illness. https://clinicaltrials.gov/ct2/show/NCT01244854.

Kilbourne 2013 *{published data only (unpublished sought but not used)}*

Amy Kilbourne. Email: amykilbo@umich.edu August to September 2020.

Goodrich DE, Kilbourne AM, Lai Z, Post EP, Bowersox NW, Mezuk B, et al. Design and rationale of a randomized controlled trial to reduce cardiovascular disease risk for patients with bipolar disorder. Contemporary Clinical Trials 2012;**33**(4):666-78.

Kilbourne AM, Goodrich DE, Lai Z, Post EP, Schumacher K, Nord KM, et al. Randomized controlled trial to assess reduction of cardiovascular disease risk in patients with bipolar disorder: the self-management addressing heart risk trial (SMAHRT). *Journal of Clinical Psychiatry* 2013;**74**(7):655-62.

Kilbourne AM. Self-Management Addressing Heart Disease Risk Trial (SMAHRT). https://clinicaltrials.gov/ct2/show/study/NCT00499096 2014.

Mishra 2017 *{published data only}*

Dr Ambed Mishra. Email: ambedmishra@gmail.com July to August 2020.

Mishra A, Krishna GS, Alla S, Kurian TD, Kurian J, Ramesh M, et al. Impact of pharmacist-psychiatrist collaborative patient education on medication adherence and quality of life (QOL) of bipolar affective disorder (BPAD) patients. *Frontiers in Pharmacology* 2017;**8**:722.

Mishra A, Sai Krishna G, Sravani A, Kurian TD, Kurian J, Ramesh M, et al. Impact of pharmacist-led collaborative patient education on medication adherence and quality of life of schizophrenia patients in a tertiary care setting. *Bulletin of Faculty of Pharmacy, Cairo University* 2017;**55**(2):345-9.

Salman 2014 *{published data only (unpublished sought but not used)}***Std22430**

Saad Salman. Email: saadirph@gmail.com July to September 2020.

Salman S, Idrees J, Anees M, Arifullah M, Al Waeel M, Ismail M, et al. Collaborative care for schizophrenic patients in primary care: a double blind, randomized clinical trial of efficacy and safety. El Mednifico Journal 2014;**2**(3):240-4. [CSZG: Ref28946]

van der Voort 2015 *{published and unpublished data}*

Collaborative care for patients with a bipolar disorder: an effect study. www.trialregister.nl/trial/2483.

Nienke van der Voort. Email: t.y.g.vandervoort@saxion.nl July to September 2020.

Stegink EE, Trijntje YG, van der Voort TY, van der Hooft T, Kupka RW, Goossens PJJ, et al. The working alliance between patients with bipolar disorder and the nurse: helpful and obstructive elements during a depressive episode from the patients' perspective. *Archives of Psychiatric Nursing* 2015;**29**:290-6.

van der Voort TY, van Meijel B, Goossens PJ, Renes J, Beekman AT, Kupka RW. Collaborative care for patients with bipolar disorder: randomised controlled trial. *BMC Psychiatry* 2011;**11**:133.

van der Voort TYG, Van Meijel B, Goossens PJJ, Hoogendoorn AW, Draisma S, Beekman A, et al. Collaborative care for patients with bipolar disorder: randomised controlled trial. British Journal of Psychiatry 2015;**206**(5):393-400.

van der Voort TYG, van Meijel B, Hoogendoorn AW, Goossens PJJ, Beekman ATF, Kupka RW. Collaborative care for patients with bipolar disorder: effects on functioning and quality of life. *Journal of Affective Disorders* 2015;**179**:14-22.

van der Voort-Scholten TYG. Collaborative care for patients with bipolar disorder [PhD Thesis]. Vrije Universiteit Amsterdam, t.y.g.vandervoort@saxion.nl 2015.

References to studies excluded from this review

ACTRN12614001312639 2014 *{published data only}*

ACTRN12614001312639. The PULSAR project: a two stepped-wedge cluster randomised controlled trial to test whether training general practitioners in recovery-oriented practice using the PULSAR (Principles Unite Local Services Assisting Recovery) intervention improves personal recovery in adult consumers of mental health services in primary care settings. http://apps.who.int/trialsearch/Trial2.aspx?TrialID=ACTRN12614001312639 2014.

Ahamad 2019 *{published data only}*

Ahamad T, Ananda KN, Ghanashyam SD, Tekkalaki Bheemsain V. Effectiveness of clinical pharmacist led collaborative approach towards medication adherence in patients with schizophrenia receiving atypical antipsychotics at tertiary care hospital. *International Journal of Research in Ayurveda and Pharmacy* 2019;**10**(3):87-91.

Barnes 2007a *{published data only}*

Barnes C, Harvey R, Mitchell P, Smith M, Wilhelm K. Outcome and costs in a randomized controlled effectiveness trial of a collaborative chronic care model for bipolar disorder. *Disease Management & Health Outcomes* 2007;**15**(4):215-24.

Barnes 2007b *{published data only}*

Barnes C, Harvey R, Mitchell P, Smith M, Wilhelm K. Evaluation of an online relapse prevention program for bipolar disorder. *Disease Management & Health Outcomes* 2007;**15**(4):215-24.

Barnes 2015 *{published data only}*

Barnes C, Hadzi-Pavlovic D, Wilhelm K, Mitchell P. A web-based preventive intervention program for bipolar disorder: outcome of 12-months randomized controlled trial. *Journal of Affective Disorders* 2015;**174**:485-92.

Bauer 2019 *{published data only}*

* Bauer MS, Miller CJ, Kim B, Lew R, Stolzmann K, Sullivan J et al. Effectiveness of implementing a collaborative chronic care model for clinician teams on patient outcomes and health status in mental health: a randomized clinical trial. *JAMA Network Open* 2019;**2**(3):e190230.

Beckelman 2013 *{published data only}*

Bekelman B, Hooker S, Nowels CT, Main DS, Meek P, McBryde C, et al. Feasibility and acceptability of a collaborative care intervention to improve symptoms and quality of life in chronic heart failure: mixed methods pilot trial. *Journal of Palliative Medicine* 2014;**17**(2):145-51.

Bowden 2012 *{published data only}*

Bowden CL, Perlis RH, Thase ME, Ketter TA, Ostacher MM, Calabrese JR, et al. Aims and results of the NIMH Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD). *Neuroscience and Therapeutics* 2012;**18**(3):243-9.

Deckersbach T, Peters A, Kaur N, Corse AK, Arulpragasam AR, Sylvia L, et al. Anxiety moderates response to psychosocial treatment for depression in bipolar disorder: results from systematic treatment enhancement program for bipolar disorder. Neuropsychopharmacology 2013;**38**(1):S422-3.

Deckersbach, T, Peters, A T, Sylvia, L, Urdahl, A, Magalhães, P V, Otto, M W, Frank, E, Miklowitz, D J, Berk, M, Kinrys, G, & Nierenberg, A. The role of co-morbid anxiety in psychosocial treatment for acute bipolar depression. *Do comorbid anxiety disorders moderate the effects of psychotherapy for bipolar disorder? Results from STEP-BD.* 2014;**171**(2):178-86.

Dennehy EB, Bauer MS, Perlis RH, Kogan JN, Sachs GS. Concordance with treatment guidelines for bipolar disorder: data from the systematic treatment enhancement program for bipolar disorder. *Psychopharmacology Bulletin* 2007;**40**(3):72-84.

Miklowitz D, Stange J, Sylvia LG, Magalhaes PV, Otto M, Frank E, et al. Extreme attributions predict the course of bipolar disorder following psychotherapy treatment for depression. Bipolar Disorders 2014;**16**(Suppl 1).

Miklowitz DJ, Otto MW, Frank E, Reilly-Harrington NA, Kogan JN, Sachs GS, et al. Intensive psychosocial intervention enhances functioning in patients with bipolar depression: results from a 9-month randomized controlled trial. American Journal of Psychiatry 2007;**164**(9):1340-7.

Miklowitz, D J, Otto, M W, Frank, E, Reilly-Harrington, N A, Wisniewski, S R, Kogan, J N, Nierenberg, A A, Calabrese, J R, Marangell, L B, Gyulai, L, Araga, M, Gonzalez, J M, Shirley, E R, Thase, M E, & Sachs, G S. Psychosocial treatments for bipolar depression: A 1-year randomized trial from the systematic treatment enhancement program. *Archives of general psychiatry,* 2007;**64**(4):419-426.

Morris CD, Miklowitz DJ, Wisniewski SR, Giese AA, Thomas MR, Allen MH. Care satisfaction, hope, and life functioning among adults with bipolar disorder: data from the first 1000 participants in the Systematic Treatment Enhancement Program. *Comprehensive Psychiatry* 2005;**46**(2):98-104.

Nierenberg A, Hansen N, Peters A, Sylvia LG, Magalhaes PV, Berk M, et al. Co-morbid anxiety moderates psychosocial treatment outcome for bipolar depression. In: Proceedings of the 16th Annual Conference of the International Society for Bipolar Disorders. 2012:43-4.

Peters A, Sylvia LG, Magalhães PV, Miklowitz DJ, Frank E, Otto MW, et al. Age at onset, course of illness and response to psychotherapy in bipolar disorder: results from the Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD). *Psychological Medicine* 2014;**44**(16):3455-67.

Peters, A, Sylvia, L G, Magalhães, P V, Miklowitz, D J, Frank, E, Otto, M W, Hansen, N S, Dougherty, D D, Berk, M, Nierenberg, A A, & Deckersbach, T. Age of onset, course of illness, and response to psychotherapy for bipolar disorder: results from the Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD). *Psychological medicine* 2014;**44**(16):3455-3467.

Stange, J P, Sylvia, L G, da Silva Magalhães, P V, Miklowitz, D J, Otto, M W, Frank, E, Berk, M, Nierenberg, A A, & Deckersbach, T. Extreme attributions predict the course of bipolar depression: results from the STEP-BD randomized controlled trial of psychosocial treatment. *The Journal of clinical psychiatry* 2013;**74**(3):249-255.

Sylvia L, Salcedo S, Peters AT, Magalhaes PVS, Frank E, Miklowitz DJ, et al. Sleep disturbance moderates psychotherapy outcomes in patients with bipolar disorder: a STEP-BD report [poster P110]. In: Proceedings of the Tenth International Conference on Bipolar Disorder 2013. Chichester: John Wiley & Sons, 2013:147.

Burns 2015 *{published data only}*

Burns T. Oxford community treatment order evaluation trial. http://www.isrctn.com/ISRCTN73110773 2015.

Byng 2004 *{published data only}*

Byng R. A randomised controlled trial of service level agreements between GPs and secondary services for the care of the long term mentally ill. https://www.isrctn.com/ISRCTN14818560 2004.

Castillo 2018 *{published data only}*

Castillo EG, Shaner R, Tang L, Chung B, Jones F, Whittington Y, et al. Improving depression care for adults with serious mental illness in underresourced areas: community coalitions versus technical support. *Psychiatric Services (Washington, D.C.)* 2018;**69**(2):195-203.

D'Souza 2004 *{published data only}*

D'Souza R, Keks N. Systematic illness management skills enhancement programme – bipolar disorder (SIMSEP-BD) – shows evidence of enhanced outcomes for bipolar disorder patients. Conference of the International Society for Bipolar Disorders 2004;**6**.

Dalcin 2018 *{published data only}*

Dalcin AT, Jerome GJ, Appel LJ, Dickerson FB, Wang NY, Miller ER, et al. Need for cardiovascular risk reduction in persons with serious mental illness: design of a comprehensive intervention. *Frontiers in Psychiatry* 2018;**9**:786.

Daumit 2020 *{published data only}*

Daumit GL, Dalcin AT, Dickerson FB, Miller ER, Evins AE, Cather C, et al. Effect of a comprehensive cardiovascular risk reduction intervention in persons with serious mental illness: a randomized clinical trial. *JAMA Network Open* 2020;**3**(6):e207247.

Davidson 2005 *{published data only}*

COPES phase I randomized controlled trial of treatment for distress in heart disease patients. https://clinicaltrials.gov/ct2/show/NCT00158054 2017.

Day 2000 *{published data only}*

Day J. A multi centre study of the management of neuroleptic medication in schizophrenia. National Research Register 2000.

Dobscha 2007 *{published data only}*

Dobscha AK, Corson K, Gerrity MS. Depression treatment preferences of VA primary care patients. *Psychosomatics* 2007;**48**(6):482-8.

Donohue 2012 *{published data only}*

Donohue J, Herbeck Belnap B, Hum B, Men A, He F, Roberts M, et al. Twelve-month cost-effectiveness of telephone-delivered

collaborative care for treating depression following CABG surgery: a randomized controlled trial. *General Hospital Psychiatry* 2014;**36**:453-9.

Donohue 2014 *{published data only}*

Donohue JM, Belnap BH, Hum B, Men A, He F, Roberts MS, et al. Twelve-month cost-effectiveness of telephone-delivered collaborative care for treating depression following CABG surgery: a randomised controlled trial. *General Hospital Psychiatry* 2014;**36**:453-9.

Druss 2001 *{published data only}*

Druss BG, Rohrbaugh RM, Levinson CM, Rosenheck RA. Integrated medical care for patients with serious psychiatric illness: a randomized trial. *Archives of General Psychiatry* 2001;**58**:861-8. [MEDLINE: 11545670]

Druss 2010 *{published data only}*

Druss BG, Von Esenwein SA, Compton MT, Rask KJ, Zhao L, Parker RM. A randomized trial of medical care management for community mental health settings: the Primary Care Access, Referral, and Evaluation (PCARE) study. *American Journal of Psychiatry* 2010;**167**(2):151-9.

Duarte 2015 *{published data only}*

Duarte A, Walker J, Walker S, Richardson G, Holm Hansen C, Martin P, et al. Cost-effectiveness of integrated collaborative care for comorbid major depression in patients with cancer. *Journal of Psychosomatic Research* 2015;**79**(6):465-70.

Dwight-Johnson 2005 *{published data only}*

Dwight-Johnson M, Ell K, Lee P. Can collaborative care address the needs of low-income Latinas with comorbid depression and cancer? Results from a randomized pilot study. *Psychosomatics* 2005;**46**(3):224-32.

Dwinger 2013 *{published data only}*

Dwinger S, Dirmaier J, Herbarth L, Konig H H, Eckardt M, Kriston L, et al. Telephone-based health coaching for chronically ill patients: study protocol for a randomized controlled trial. *Trials* 2013;**14**:337.

Harter M, Dirmaier J, Dwinger S, Kriston L, Herbarth L, Siegmund-Schultze E, et al. Effectiveness of telephone-based health coaching for patients with chronic conditions: a randomised controlled trial. *PLOS One* 2016;**11**(9):e0161269.

Ell 2012 *{published data only}*

Ell K, Katon W, Lee P, Kapetanovic S, Guterman J, Xie B, et al. Depressive symptom deterioration among predominantly Hispanic diabetes patients in safety net care. *Psychosomatics* 2012;**53**(4):347-55.

Ell 2014 *{published data only}*

Ell K, Oh H, Lee PJ, Guterman J. Collaborative health literate depression care among predominantly hispanic patients with coronary heart disease in safety net care. *Psychosomatics* 2014;**55**:555-65.

Ell 2016 *{published data only}*

Ell K, Aranda M, Wu S, Oh H, Lee P, Guterman J. Promotora assisted depression care among predominately Hispanic patients with concurrent chronic illness: public care system clinical trial design. *Contemporary Clinical Trials* 2016;**46**:39-47.

EQUIP *{published data only}*

Bower P, Roberts C, O'Leary N, Callaghan P, Bee P, Fraser C, et al. A cluster randomised controlled trial and process evaluation of a training programme for mental health professionals to enhance user involvement in care planning in service users with severe mental health issues (EQUIP): study protocol for a randomised controlled trial. *Trials* 2015;**16**(1):348.

Bower P, Roberts C, O'Leary N, Callaghan P, Bee P, Fraser C, et al. A cluster randomised controlled trial and process evaluation of a training programme for mental health professionals to enhance user involvement in care planning in service users with severe mental health issues (EQUIP): study protocol for a randomised controlled trial. *Trials* 2015;**16**:348.

Camacho EM, Shields G, Lovell K, Coventry PA, Morrison AP, Davies LM. A (five-)level playing field for mental health conditions?: exploratory analysis of EQ-5D-5L-derived utility values. *Quality of Life Research* 2018;**27**(3):717-24.

Hughes-Morley A, Hann M, Fraser C, Meade O, Lovell K, Young B, et al. Does advertising patient and public involvement in a trial to potential participants improve recruitment and response rates? An embedded cluster randomised trial. *Trials* 2017;**18**(Suppl 1):200.

Hughes-Morley A, Hann M, Fraser C, Meade O, Lovell K, Young B, et al. The impact of advertising patient and public involvement on trial recruitment: embedded cluster randomised recruitment trial. *Trials* 2016;**17**(1):586.

Lovell K, Bee P, Brooks H, Cahoon P, Callaghan P, Carter LA, et al. Embedding shared decision-making in the care of patients with severe and enduring mental health problems: the EQUIP pragmatic cluster randomised trial. *PLOS One* 2018;**13**(8):e0201533.

NCT00119574. Evaluating a collaborative care model for the treatment of schizophrenia (EQUIP). https://classic.clinicaltrials.gov/ct2/show/NCT00119574 2005.

Ertem 2018 *{published data only}*

Ertem MY, Duman ZC. The effect of motivational interviews on treatment adherence and insight levels of patients with schizophrenia: a randomized controlled study. Perspectives in Psychiatric Care 2018 Jun 11 [Epub ahead of print].

Falkum 2010 *{published data only}*

Bull H, Ueland T, Lystad JU, Evensen S, Martinsen EW, Falkum E. Vocational functioning in schizophrenia spectrum disorders: does apathy matter? *Journal of Nervous and Mental Disease* 2016;**204**(8):599-605.

Falkum E, Klungsoyr O, Lystad JU, Bull HC, Evensen S, Martinsen EW, et al. Vocational rehabilitation for adults with psychotic disorders in a Scandinavian welfare society. *BMC Psychiatry* 2017;**17**(1):24.

Falkum E. Vocational rehabilitation for individuals with schizophrenia-should competitive work be the only goal. *European Archives of Psychiatry and Clinical Neuroscience* 2017;**267**(1 Suppl 1):S10-1.

Lystad JU, Falkum E, Haaland VO, Bull H, Evensen S, Bell MD, et al. Neurocognition and occupational functioning in schizophrenia spectrum disorders: the MATRICS Consensus Cognitive Battery (MCCB) and workplace assessments. *Schizophrenia Research* 2016;**170**(1):143-9.

Lystad JU, Falkum E, Haaland VO, Bull H, Evensen S, McGurk SR, et al. Cognitive remediation and occupational outcome in schizophrenia spectrum disorders: a 2 year follow-up study. *Schizophrenia Research* 2017;**185**:122-9.

NCT01139502. Cognitive behaviour therapy and cognitive training in work rehabilitation for persons with severe mental illness. http://ClinicalTrials.gov/show/NCT01139502 2010.

Ullevoldsaeter Lystad J. Cognitive remediation, cognitive behavioral therapy and occupational outcome in schizophrenia: a 2 year follow-up study. *European Archives of Psychiatry and Clinical Neuroscience* 2017;**267**(1 Suppl 1):S9-S10.

Fleehart 2015 *{published data only}*

Fleehart S, Nguyen HQ, Fan VS, Hunter C, Chen Z, Reinke LF. The effect of psychosocial behavioral therapy for patients with COPD and depression. American Journal of Respiratory and Critical Care Medicine 2015;**191**.

Fortney 2015 *{published data only}*

Fortney JC, Pyne JM, Kimbrell TA, Hudson TJ, Robinson DE, Schneider R, et al. Telemedicine-based collaborative care for posttraumatic stress disorder: a randomized clinical trial. *JAMA Psychiatry* 2015;**72**(1):58-67.

Fortney JC, Pyne JM, Kimbrell TA, Hudson TJ, Robinson DE, Schneider R, et al. Telemedicine-based collaborative care for posttraumatic stress disorder: a randomized clinical trial. *JAMA Psychiatry* 2015;**72**:58-67.

Gensichen 2006 *{published data only}*

Gensichen J, Katon W, Tang LQ, Williams JW, Kroenke K, Lin EHB, et al. IMPACT collaborative care improves depression in elderly patients in primary care in the longer term. *Evidence Based Mental Health* 2006;**9**(3):76.

Gerritsen 2014 *{published data only}*

Gerritsen DL, Smalbrugge M, Teerenstra S, Leontjevas R, Adang EM, Vernooij-Dassen MJ, et al. Act In case of Depression: the evaluation of a care program to improve the detection and treatment of depression in nursing homes. Study Protocol. *BMC Psychiatry* 2011;**11**:91. [DOI: 10.1186/1471-244X-11-91]

Goorden 2014 *{published data only}*

Goorden M, Muntingh A, van Marwijk H, Spinhoven P, Adèr H, van Balkom A, et al. Cost utility analysis of a collaborative stepped care intervention for panic and generalized anxiety disorders in primary care. *Journal of Psychosomatic Research* 2014;**77**(1):57-63.

Goorden 2015 *{published data only}*

Goorden M, Huijbregts KML, van Marwijk HWJ, Beekman ATF, van der Feltz-Cornelis CM, Hakkaart-van Roijen L. Cost-utility of collaborative care for major depressive disorder in primary care in the Netherlands. *Journal of Psychosomatic Research* 2015;**79**:316-23.

Goorden M, Vlasveld MC, Anema JR, van Mechelen W, Beekman ATF, Hoedeman R, et al. Cost-utility analysis of a collaborative care intervention for major depressive disorder in an occupational healthcare setting. *Journal of Occupational Rehabilitation* 2014;**24**(3):555-62.

Gureje 2017 *{published data only}*

Gureje O, Makanjuola V, Kola L, Yusuf B, Price L, Esan O, et al. COllaborative Shared care to IMprove Psychosis Outcome (COSIMPO): study protocol for a randomized controlled trial. *Trials* 2017;**18**(1):462.

Gureje O, Seedat S, Kola L, Appiah-Poku J, Othieno C, Harris B, et al. Partnership for mental health development in Sub-Saharan Africa (PaM-D): a collaborative initiative for research and capacity building. Epidemiology and Psychiatric Sciences 2018 Nov 27 [Epub ahead of print].

Hidalgo-Mazzei 2015 *{published data only}*

Hidalgo-Mazzei D, Mateu A, Reinares M, Undurraga J, Bonnin Cdel M, Sanchez-Moreno J, et al. Self-monitoring and psychoeducation in bipolar patients with a smart-phone application (SIMPLe) project: design, development and studies protocols. *BMC Psychiatry* 2015;**15**:52.

Hidalgo-Mazzei D, Reinares M, Murru A, Bonnin CM, Vieta E, Colom F. Signs and symptoms self-monitoring and psychoeducation in bipolar patients with a smart-phone application (SIMPLe) project. *BMC Psychiatry* 2015;**15**:52. [DOI: 10.1186/s12888-015-0437-6]

Hirayasu 2009 *{published data only}*

Hirayasu Y, Kawanishi C, Yonemoto N, Ishizuka N, Okubo Y, Sakai A, et al. A randomized controlled multicenter trial of post-suicide attempt case management for the prevention of further attempts in Japan (ACTION-J). *BMC Public Health* 2009;**9**:364.

Hogarty 1974 *{published data only}*

Hogarty GE. NIHM-PRB collaborative outpatient schizophrenic study. *Psychopharmacology Bulletin* 1974;**10**(2):54-5.

Huffman 2014 *{published data only}*

Huffman JC, Mastromauro CA, Beach SR, Celano CM, DuBois CM, Healy BC, et al. Collaborative care for depression and anxiety disorders in patients with recent cardiac events: the Management of Sadness and Anxiety in Cardiology (MOSAIC) randomized clinical trial [published correction appears in JAMA Intern Med. 2014 Aug;174(8):1419]. *JAMA Internal Medicine* 2014;**174**(6):927-35. [DOI: 10.1001/jamainternmed.2014.739]

Huffman JC, Mastromauro CA, Beach SR, Celano CM, DuBois CM, Healy BC, et al. Collaborative care for depression and anxiety disorders in patients with recent cardiac events: the Management of Sadness and Anxiety in Cardiology

(MOSAIC) randomized clinical trial. *JAMA Internal Medicine* 2014;**174**(6):927-35.

Huijbregts 2010 *{published data only}*

Huijbregts K, De Jong FJ, Martens F, Ader HJ, Van Marwijk H, Beekman ATF, et al. Effectiveness of collaborative care for depression in Dutch primary care. *Journal of Psychosomatic Research* 2010;**68**(6):74.

Huijbregts 2013 *{published data only}*

Huijbregts KM, de Jong FJ, van Marwijk HW, Beekman AT, Ader HJ, van der Feltz-Cornelis CM. A high physical symptom count reduces the effectiveness of treatment for depression, independently of chronic medical conditions. *Journal of Psychosomatic Research* 2013;**74**(3):179-85.

Iezzoni 2015 *{published data only}*

Iezzoni LI, Chang Y, Matulewicz H, Heaphy D, Warsett KS, Donelan K. Health plan enrollees with disability informing primary care practices and providers about their quality of care: a randomized trial. *Disability and Health Journal* 2018;**11**(4):537-44.

Iezzoni LI, Heaphy D, Warsett KS, Marsella SA. Description of YESHealth: a consumer-directed intervention in a randomized trial of methods to improve quality of care for persons with disability. *Disability and Health Journal* 2018;**11**(4):545-54.

NCT02390557. Persons with disabilities quality monitoring intervention. https://www.ClinicalTrials.gov.

IRCT2015060622580N1 2020 *{published data only}*

Zoladl M, Afroughi S, Nooryan K, Kharamin S, Haghgoo A, Parandvar Y. Applying collaborative care model on intensive caregiver burden and resilient family caregivers of patients with mental disorders: a randomized controlled trial. Iranian Journal of Psychiatry 2020;**15**:17-26.

Johnson 2018 *{published data only}*

Johnson S, Lamb D, Marston L, Osborn D, Mason O, Henderson C, et al. Peer-supported self-management for people discharged from a mental health crisis team: a randomised controlled trial. *Lancet* 2018;**392**(10145):409-18.

Kastner 2012 *{published data only}*

Bramesfeld A, Moock J, Kopke K, Buchtemann D, Kastner D, Radisch J, et al. Effectiveness and efficiency of assertive outreach for schizophrenia in Germany: study protocol on a pragmatic quasi-experimental controlled trial. *BMC Psychiatry* 2013;**13**:56.

Buchtemann D, Kastner D, Warnke I, Radisch J, Baumgardt J, Giersberg S, et al. Hospital utilization outcome of an assertive outreach model for schizophrenic patients - results of a quasi-experimental study. *Psychiatry research* 2016;**241**:249-55.

ISRCTN34900108. Assertive outreach for people with schizophrenia in Germany. http://www.controlled-trials.com/ISRCTN34900108/ 2012.

Kastner D, Buchtemann D, Warnke I, Radisch J, Baumgardt J, Giersberg S, et al. Clinical and functional outcome of assertive outreach for patients with schizophrenic disorder: results of a quasi-experimental controlled trial. *European Psychiatry* 2015;**30**(6):736-42.

Kendrick 2003 *{published data only}*

Kendrick A, Simons L, Mynors-Wallis L, Gray A, Lathlean J, Pickering R, et al. Trial of problem-solving by community psychiatric nurses (CPNs) for anxiety, depression and life difficulties among general practice patients. Health Technology Assessment;**9**.

Kendrick 2005 *{published data only}*

Kendrick T, Simons L, Mynors-Wallis L, Gray A, Lathlean J, Pickering R, et al. A trial of problem-solving by community mental health nurses for anxiety, depression and life difficulties among general practice patients. The CPN-GP study. *Health Technology Assessment* 2005;**9**(37):1-104, iii.

Kershaven 2003 *{published data only}*

Keshavan M. A 6-month controlled study of PsychoEducation and Collaboration Enhancement (PEACE) in early psychosis to improve treatment compliance. Stanley Foundation Research Programs 2003.

Khambaty 2015 *{published data only}*

Khambaty T, Callahan CM, Stewart JC. Depression treatment and diabetes risk: a 9-year follow-up study of the impact trial. PhD Thesis 2015.

Kikkert 2018 *{published data only}*

Kikkert M, Goudriaan A, de Waal M, Peen J, Dekker J. Effectiveness of Integrated Dual Diagnosis Treatment (IDDT) in severe mental illness outpatients with a co-occurring substance use disorder. *Journal of Substance Abuse Treatment* 2018;**95**:35-42.

Kilbourne 2009 *{published data only}*

Kilbourne AM, Biswas K, Pirraglia PA, Sajatovic M, Williford WO, Bauer MS. Is the collaborative chronic care model effective for patients withbipolar disorder and co-occurring conditions? *Journal of Affective Disorders* 2009;**112**(1-3):256-61.

Kilbourne 2012b *{published data only}*

Kilbourne AM, Neumann MS, Waxmonsky J, Bauer MS, Kim HM, Pincus HA, et al. Public-academic partnerships: evidence-based implementation: the role of sustained community-based practice and research partnerships. *Psychiatric Services* 2012;**63**(3):205-7.

Kilbourne 2013a *{published data only}*

Kilbourne AM, Li D, Lai Z, Waxmonsky J, Ketter T. Pilot randomized trial of a cross-diagnosis collaborative care program for patients with mood disorders. *Depression and Anxiety* 30;**2**:116-22.

Kilbourne 2014 *{published data only}*

Kilbourne AM, Bramlet M, Barbaresso MM, Nord KM, Goodrich DE, Lai Z, et al. SMI Life Goals: description of a randomized trial of a collaborative care model to improve outcomes for persons with serious mental illness. Contemporary Clinical Trials 2014;**39**(1):74-85. [CSZG: Ref29066]

Kilbourne AM. Life goals collaborative care to improve health outcomes in mental disorders. https://clinicaltrials.gov/ct2/show/NCT01487668.

NCT01487668. Life goals collaborative care to improve health outcomes in mental disorders. http://ClinicalTrials.gov/show/NCT01487668 2011.

Kilbourne 2015 *{published data only}*

Kilbourne AM, Goodrich DE, Nord KM, Van Poppelen C, Kyle J, Bauer MS, et al. Long-term clinical outcomes from a randomized controlled trial of two implementation strategies to promote collaborative care attendance in community practices. *Administration and Policy in Mental Health* 2015;**42**(5):642-53.

Waxmonsky J, Kilbourne AM, Goodrich DE, Nord KM, Lai Z, Laird C, et al. Enhanced fidelity to treatment for bipolar disorder: results from a randomized controlled implementation trial. Psychiatric Services 2014;**65**(1):81-90.

Kilbourne 2017 *{published data only}***Std18738**

Kilbourne A. Cross-diagnosis collaborative care program for patients with mood disorders from the U.S. national network of depression centers [abstract]. 10th International Conference on Bipolar Disorders (ICBD), 2013 June 13-16 June; Miami, Florida 2013.

Kilbourne AM, Barbaresso MM, Lai Z, Nord KM, Bramlet M, Goodrich DE, et al. Improving physical health in patients with chronic mental disorders: twelve-month results from a randomized controlled collaborative care trial. *Journal of Clinical Psychiatry* 2017;**78**(1):129-37.

NCT01244854. Life goals behavioral change to improve outcomes for veterans with serious mental illness. http://ClinicalTrials.gov/show/NCT01244854 2010.

Knight 2008 *{published data only}*

Knight MM, Houseman EA. A collaborative model for the treatment of depression in homebound elders. *Issues in Mental Health Nursing* 2008;**29**(9):974-91.

Lomax 1992 *{published data only}*

Lomax H. The Effects of Collaborative Psycho Educational Intervention on Relatives of Schizophrenic Veterans [PhD dissertation]. Saint Louis University, USA, 1992.

Mcdonough 2009 *{published data only}*

McDonough M, Thornicroft G, Barclay W, DeWet C, Kalidindi S, O'Brien T. Costs and benefits of a pilot shared care register between primary and secondary healthcare for patients with psychotic disorders. *Primary Care Mental Health* 2003;**1**(1):55-62.

McGurk *{published data only}*

McGurk SR, Mueser KT, Xie H, Welsh J, Kaiser S, Drake RE, et al. Cognitive enhancement treatment for people with mental illness who do not respond to supported employment: a randomized controlled trial. *American Journal of Psychiatry* 2015;**172**(9):852-61.

McGurk SR, Mueser KT. Cognitive remediation and supported employment: the thinking skills for work program. *Schizophrenia Bulletin* 2013;**39**:S343-4.

NCT01926613. Cognitive rehabilitation and supported employment for severe mental illness. http://clinicaltrials.gov/show/NCT01926613 2006.

Teixeira C, Mueser KT, Rogers ES, McGurk SR. Job endings and work trajectories of persons receiving supported employment and cognitive remediation. *Psychiatric Services (Washington, D.C.)* 2018;**69**(7):812-8.

Menchetti 2013 *{published data only}*

Menchetti M, Sighinolfi C, Nespeca C, Di Michele V, Peloso P, Levantesi P, et al. Italian study on a collaborative care program for primary care attenders with depressive disorders. *Journal of Psychosomatic Research* 2013;**74**:553-4.

Meyer 2014 *{published data only}*

Meyer T, Belnap BH, Herrmann-Lingen C, He F, Mazumdar S, Rollman BL. Benefits of collaborative care for post-CABG depression are not related to adjustments in antidepressant pharmacotherapy. *Journal of Psychosomatic Research* 2014;**76**(1):28-33.

Morone 2010 *{published data only}*

Morone NE, Weiner DK, Belnap BH, Karp JF, Mazumdar S, Houck PR, et al. The impact of pain and depression on recovery after coronary artery bypass grafting. *Psychosomatic Medicine* 2010;**72**(7):620-5.

NCT00137280 2005 *{published data only}*

NCT00137280. Implementing effective, collaborative care for schizophrenia. http://www.clinicaltrials.gov 2005.

NCT00919620 *{published data only}*

Chan IHH, Lai DC, Hui CLM, Li GYK, Chen EYH, Tam WWY, et al. A randomized, wait-list controlled clinical trial: the effect of life coaching on functioning in early psychosis. *Early Intervention in Psychiatry* 2012;**6**:89.

Chang WC, Hui CLM, Lai DC, Tam WWY, Tang JYM, Wong GHY, et al. Clinical and social determinants of duration of untreated psychosis: a 4-year randomized controlled-trial of specialized early intervention service for adult-onset first-episode psychosis (jcep study). *Early Intervention in Psychiatry* 2012;**6**:77.

Chen EYH, Hui CLM, Chan SKW, Chang WC, Lee EHM. 4-year outcome of a specialized early intervention treatment for adult onset psychosis (JCEP): a randomized controlled trial. *Early Intervention in Psychiatry* 2018;**12**(Suppl 1):23.

Chen EYH, Hui CLM, Tam WWY, Lam MML, Chan SKW, Li FWS, et al. A 5-year randomized controlled trial of a specialized early intervention service in first-episode psychosis. *Early Intervention in Psychiatry* 2010;**4**(Suppl 1):133.

Hui CL, Chang WC, Chan SK, Lee EH, Tam WW, Lai DC, et al. Early intervention and evaluation for adult-onset psychosis: the JCEP study rationale and design. *Early Intervention in Psychiatry* 2014;**8**(3):261-8.

Hui CLM, Chen EYH, Lam MML, Chan SKW, Li FWS, Leung KF, et al. A five-year randomised controlled trial of a specialised early intervention service in adult-onset first episode psychosis: a preliminary report. *East Asian Archives of Psychiatry* 2010;**20**:22.

Hui CLM, Lau WWY, Leung CM, Chang WC, Tang JYM, Wong GHY, et al. Clinical and social correlates of duration of untreated psychosis among adult-onset psychosis in Hong Kong Chinese: the JCEP study. *Early Intervention in Psychiatry* 2015;**9**(2):118-25.

NCT00919620. Stage-specific case management for early psychosis. https://www.ClinicalTrials.gov/ct/show/ 2009.

NCT01436331 *{published data only}*

Ruggeri M. A large pragmatic cluster randomized controlled trial of a multi-element psychosocial intervention for early psychosis GETUP PIANO. https://www.clinicaltrials.gov/study/NCT01436331 2013. [https://clinicaltrials.gov/ct2/show/NCT01436331]

NCT02440906 *{published data only}*

Evaluation of the Texas Wellness Incentives and Navigation (WIN) Project. https://clinicaltrials.gov/ct2/show/NCT02440906.

NCT02543840 *{published data only}*

Bauer MS, Miller CJ, Kim B, Lew R, Stolzmann K, Sullivan J, et al. Effectiveness of implementing a collaborative chronic care model for clinician teams on patient outcomes and health status in mental health a randomized clinical trial. *JAMA Network Open* 2019;**2**(3):e190230.

Bauer MS, Weaver K, Kim B, Miller C, Lew R, Stolzmann K, et al. The collaborative chronic care model for mental health conditions: from evidence synthesis to policy impact to scale-up and spread. *Medical Care* 2019;**57**(10 Suppl 3):S221-7.

NCT03590041 2020 *{published data only}*

Kwan BM, Dickinson LM, Glasgow RE, Sajatovic M, Gritz M, Holtrop JS, et al. Correction to: The Invested in Diabetes Study Protocol: a cluster randomized pragmatic trial comparing standardized and patient-driven diabetes shared medical appointments. Trials 2020;**21**:195.

Kwan BM, Dickinson LM, Glasgow RE, Sajatovic M, Gritz M, Holtrop JS, et al. The Invested in Diabetes Study Protocol: a cluster randomized pragmatic trial comparing standardized and patient-driven diabetes shared medical appointments. Trials 2020;**21**:65.

NCT03881657 2020 *{published data only}*

Errichetti KS, Flynn A, Gaitan E, Ramirez MM, Baker M, Xuan Z. Randomized trial of reverse colocated integrated care on persons with severe, persistent mental illness in Southern Texas. Journal of General Internal Medicine 2020;**35**:2035-42.

Sautter Errichetti K, Ramirez MM, Flynn A, Xuan Z. A reverse colocated integrated care model intervention among persons with severe persistent mental illness at the U.S.-Mexico border: a randomized controlled trial protocol. Contemporary Clinical Trials Communications 2019;**16**:100490.

NCT04324944 2021 *{published data only}*

NCT04324944. Adapting and examining collaborative decision skills training among veterans with serious mental illness. https://ClinicalTrials.gov/show/NCT04324944 2021.

NCT04600414 2020 *{published data only}*

NCT04600414. Collaborating to heal addiction and mental health in primary care. https://ClinicalTrials.gov/show/NCT04600414 2020.

NCT04601064 2021 *{published data only}*

NCT04601064. Peer supported collaborative care mental health and substance use disorder care. https://ClinicalTrials.gov/show/NCT04601064 2021.

Nordentoft 2000 *{published data only}*

Nordentoft M, Jeppesen P, Jørgensen P, Abel MB, Kassow P, Reisby N, et al. Opus-project: a randomised controlled trial of first episode psychotic patients: better compliance. *Schizophrenia Research* 2000;**41**(1):B145.

Overend 2014 *{published data only}*

Overend K, Lewis H, Bailey D, Bosanquet K, Chew-Graham C, Ekers D, et al. CASPER plus (CollAborative care in Screen-Positive ElderRs with major depressive disorder): study protocol for a randomised controlled trial. *Trials* 2014;**15**:451.

Patel 2008 *{published data only}*

Patel VH, Kirkwood BR, Pednekar S, Araya R, King M, Chisholm D, et al. Improving the outcomes of primary care attenders with common mental disorders in developing countries: a cluster randomized controlled trial of a collaborative stepped care intervention in Goa, India. *Trials* 2008;**9**:4.

Patel 2010 *{published data only}*

Patel V, Weiss HA, Chowdhary N, Naik S, Pednekar S, Chatterjee S, et al. Effectiveness of an intervention led by lay health counsellors for depressive and anxiety disorders in primary care in Goa, India (MANAS): a cluster randomised controlled trial. *Lancet* 2010;**376**(9758):2086-95.

Pereira 2011 *{published data only}*

Pereira B, Andrew G, Pednekar S, Kirkwood B R, Patel V. The integration of the treatment for common mental disorders in primary care: experiences of health care providers in the MANAS trial in Goa, India. *International Journal of Mental Health Systems* 2011;**5**(1):26.

Pin 2014 *{published data only}*

Pin N-T. Screening and treatment of depression in the community [Randomized controlled trial of a community-based early psychiatric intervention strategy to screen and manage depression in the elderly]. https://clinicaltrials.gov/study/NCT00430404 2014.

Price 2004 *{published data only}*

Price J. Collaborative care improves health outcomes in older people with depression and arthritis [IMPACT study]. *Evidence-Based Mental Health* 2004;**7**:45.

Putz 2015 *{published data only}*

Putz JW. Healthcare Integration and Hepatitis C Surveillance in A Community Mental Health Centre [Dissertation]. Ann Arbor: Indiana University, 2015.

RAISE-ETP *{published data only}*

Azorin JM, Adida M, Belzeaux R, Fakra E. A model of care for first-episode psychosis: the RAISE-ETP project. *L'Encephale* 2016;**42**(Suppl 3):S13-7.

Brown B, Alphs L, Turkoz I, Yue Y. Baseline demographics and characteristics from a paliperidone palmitate study in subjects with recent-onset schizophrenia or schizophreniform disorder. *Psychopharmacology Bulletin* 2017;**47**(3):8-16.

Brunette MF. Facilitators and barriers to implementation of coordinated specialty care in U.S. community mental health clinic. *Schizophrenia Bulletin* 2015;**41**:S304.

Cadenhead K, Addington J, Bearden C, Cannon T, Cornblatt B, Mathalon D, et al. Metabolic abnormalities prior to the onset of psychosis: another risk factor for psychosis? *Neuropsychopharmacology* 2015;**40**:S565.

Cather C, Brunette MF, Mueser KT, Babbin SF, Rosenheck R, Correll CU, et al. Impact of comprehensive treatment for first episode psychosis on substance use outcomes: a randomized controlled trial. *Psychiatry Research* 2018;**268**:303-11.

Correll CU, Robinson DG, Schooler NR, Brunette MF, Mueser KT, Rosenheck RA, et al. Cardiometabolic risk in patients with first-episode schizophrenia spectrum disorders: baseline results from the RAISE-ETP study. *JAMA Psychiatry* 2014;**71**(12):1350-63.

Fulford D, Piskulic D, Addington J, Kane JM, Schooler NR, Mueser KT. Prospective relationships between motivation and functioning in recovery after a first episode of schizophrenia. *Schizophrenia Bulletin* 2018;**44**(2):369-77.

Glynn S, Gingerich S, Meyer-Kalos P, Mueser K, Chan-Golston A, Sugar C, et al. Who participated in family work in the US raise-ETP first episode sample? *Schizophrenia Bulletin* 2018;**44**(Suppl 1):S216-7.

Glynn SM, Gingerich S, Mueser KT, Cather C, Penn D. The role of family intervention in coordinated specialty care for first episode psychosis. *Schizophrenia Bulletin* 2015;**41**:S173.

Heinssen R. Duration of untreated psychosis moderates clinical outcomes and cost-effectiveness in first episode psychosis treatment programs. *Early Intervention in Psychiatry* 2018;**12**(Suppl 1):22.

Kane J, Schooler N, Robinson D, Addington J, Kane JM. The NIMH RAISE ETP (early treatment program): initial results. *Early Intervention in Psychiatry* 2014;**8**:1.

Kane JM, Robinson DG, Schooler NR, Mueser KT, Penn DL, Rosenheck RA, et al. Comprehensive versus usual community care for first-episode psychosis: 2-year outcomes from the NIMH RAISE early treatment program. *American Journal of Psychiatry* 2016;**173**:362-72.

Kane JM. RAISE-ETP: navigate vs usual care-two year outcomes. *Schizophrenia Bulletin* 2015;**41**:S317.

Kane JM. The RAISE ETP study: initial results. *Early Intervention in Psychiatry* 2014;**8**:2.

McClellan JM. Coordinated specialty care for first-episode psychosis. *Journal of the American Academy of Child and Adolescent Psychiatry* 2018;**57**(10 Suppl):S269.

McDonell M. The impact of tobacco, alcohol, and cannabis use on treatment outcomes among patients experiencing first-episode psychosis: data from the national RA1SE Early Treatment Program (ETP) Study. *Journal of the American Academy of Child and Adolescent Psychiatry* 2018;**57**(10 Suppl):S270.

Mueser KT, Meyer-Kalos PS, Glynn SM, Lynde DW, Robinson DG, Gingerich S, et al. Implementation and fidelity assessment of the NAVIGATE treatment program for first episode psychosis in a multi-site study.. *Schizophrenia Research* 2019;**204**:271-81.

Mueser KT, Penn DL, Addington J, Brunette MF, Gingerich S, Glynn SM, et al. The NAVIGATE program for first-episode psychosis: rationale, overview, and description of psychosocial components. *Psychiatric Services* 2015;**66**(7):680-90.

Mueser KT. Description and implementation of the RAISE-ETP study psychosocial treatment model: the NAVIGATE program. *Schizophrenia Bulletin* 2015;**41**:S325-6.

Oluwoye O, Stiles B, Monroe-DeVita M, Chwastiak L, McClellan JM, Dyck D, et al. Racial-ethnic disparities in first-episode psychosis treatment outcomes from the RAISE-ETP study. *Psychiatric Services* 2018;**69**(11):1138-45.

Oluwoye O. Racial differences in first-episode psychosis treatment outcomes: data from the national RA1SE Early Treatment Program (ETP) study. *Journal of the American Academy of Child and Adolescent Psychiatry* 2018;**57**(10 Suppl):S270.

Robinson DG, Schooler NR, Correll CU, John M, Kurian BT, Marcy P, et al. Psychopharmacological treatment in the RAISE-ETP study: outcomes of a manual and computer decision support system based intervention. *American Journal of Psychiatry* 2018;**175**(2):169-79.

Robinson DG, Schooler NR, John M, Correll CU, Marcy P, Addington J, et al. Prescription practices in the treatment of first-episode schizophrenia spectrum disorders: data from the national RAISE-ETP study. *American Journal of Psychiatry* 2015;**172**(3):237-48.

Robinson DG. Randomized comparison of comprehensive versus usual community care for first-episode psychosis: the RAISE-ETP study. *Journal of the American Academy of Child and Adolescent Psychiatry* 2017;**56**(10):S340.

Rosenheck R, Mueser KT, Sint K, Lin H, Lynde DW, Glynn SM, et al. Supported employment and education in comprehensive, integrated care for first episode psychosis: effects on work, school, and disability income. *Schizophrenia Research* 2017;**182**:120-8.

Rosenheck RA, Estroff SE, Sint K, Lin H, Mueser KT, Robinson DG, et al. Incomes and outcomes: social security disability benefits in first-episode psychosis. *American Journal of Psychiatry* 2017;**174**(9):886-94.

Schooler N, Khan A, Keefe R, Marcy P, Robinson D, Kane J. Cognitive functioning in first episode psychosis: comparison of a two-year coordinated specialty care program to community care. *Neuropsychopharmacology* 2016;**41**:S593.

Schooler N, Khan A, Keefe R, Robinson D, Kane J. Cognitive functioning in first-episode psychosis:-X000B- Comparison of a 2-year coordinated specialty care program to community care. *Schizophrenia Bulletin* 2017;**43**:S24.

Schooler N. RAISE-ETP study design, site selection and implementation model. *Early Intervention in Psychiatry* 2014;**8**:1.

Schooler NR. The RAISE-ETP study design, research and implementation model. *Schizophrenia Bulletin* 2015;**41**:S332-3.

Sint K, Rosenheck R, Robinson DG, Schooler NR, Marcy P, Kane JM, et al. Accounting for group differences in study retention in a randomized trial of specialized treatment for first episode psychosis. *Schizophrenia Research* 2018;**195**:481-7.

Raube 1992 *{published data only}*

Raube K. Health and Social Support of the Elderly. RAND Corporation 1992.

Richards 2016 *{published data only}*

Richards DA, Bower P, Chew-Graham C, Gask L, Lovell K, Cape J, et al. Clinical effectiveness and cost-effectiveness of collaborative care for depression in UK primary care (CADET): a cluster randomised controlled trial. *Health Technology Assessment* 2016;**20**(14):1-192.

Richards DA, Lovell K, Gilbody S, Gask L, Torgerson D, Barkham M, et al. Collaborative care for depression in UK primary care: a randomised controlled trial. *Psychological Medicine* 2007;**38**:279-87.

Richards DA. Multi-centre randomised controlled trial of collaborative care for depression [CADET (CollAborative DEpression Trial)] [ISRCTN32829227]. http://www.pms.ac.uk/cadet/AboutCADET.aspx.

Richardson 2014 *{published data only}*

Richardson LP, Ludman E, McCauley E, Lindenbaum J, Larison C, Zhou C, et al. Collaborative care for adolescents with depression in primary care: a randomized clinical trial. *JAMA* 2014;**312**(8):809-16.

Rollman 2009 *{published data only}*

Rollman BL, Belnap BH, LeMenager MS, Mazumdar S, Schulberg HC, Reynolds CF 3rd. The Bypassing the Blues treatment protocol: stepped collaborative care for treating post-CABG depression. *Psychosomatic Medicine* 2009;**71**(2):217-30.

Rollman 2018 *{published data only}*

Rollman BL, Herbeck Belnap B, Abebe KZ, Spring MB, Rotondi AJ, Rothenberger SD, Karp JF. Effectiveness of online collaborative care for treating mood and anxiety disorders in primary care: a randomized clinical trial. *JAMA Psychiatry* 2018;**75**(1):56-64. [DOI: 10.1001/jamapsychiatry.2017.3379]

Sajatovic *{published data only}*

Sajatovic M. Treatment adherence enhancement in bipolar disorder (CAE RCT). https://www.clinicaltrials.gov/ct2/show/NCT01542008.

Sajatovic 2005a *{published data only}*

Sajatovic M, Davies M, Bauer MS, McBride L, Hays RW, Safavi R, et al. Attitudes regarding the collaborative practice model and treatment adherence among individuals with bipolar disorder. *Comprehensive Psychiatry* 2005;**46**(4):272-7.

Sajatovic 2005b *{published data only}*

Sajatovic M, Davies M, Bauer MS, McBride L, Hays R, Safav Ri, Jenkins J. Attitudes regarding the collaborative practice model and treatment adherence among individuals with bipolar disorder. Comprehensive Psychiatry 2005;**46**(4):272-7.

Sathienluckana 2018 *{published data only}*

Sathienluckana T, Unaharassamee W, Suthisisang C, Suanchang O, Suansanae T. Anticholinergic discontinuation and cognitive functions in patients with schizophrenia: a pharmacist-physician collaboration in the outpatient department. *Integrated Pharmacy Research and Practice* 2018;**7**:161-71.

TCTR20180420002. Anticholinergic discontinuation and cognitive functions in patients with schizophrenia: a pharmacist-physician collaboration in the outpatient department. http://www.clinicaltrials.in.th/index.php?tp=regtrials&menu=trialsearch&smenu=fulltext&task=search&task2=view1&2018.

Schaefert 2013 *{published data only}*

Schaefert R, Kaufmann C, Wild B, Schellberg D, Boelter R, Faber R, et al. Specific collaborative group intervention for patients with medically unexplained symptoms in general practice: a cluster randomized controlled trial. *Psychotherapy and Psychosomatics* 2013;**82**(2):106-19.

Schmidt 1998a *{published data only}*

Schmidt I, Claesson CB, Westerholm B, Nilsson LG, Svarstad BL. The impact of regular multidisciplinary team interventions on psychotropic prescribing in Swedish nursing homes. *Journal of the American Geriatrics Society* 1998;**46**(1):77-82.

Shinde 2013 *{published data only}*

Shinde S, Andrew G, Bangash O, Cohen A, Kirkwood B, Patel V. The impact of a lay counselor led collaborative care intervention for common mental disorders in public and private primary care: a qualitative evaluation nested in the MANAS trial in Goa, India. *Social Science & Medicine* 2013;**88**:48-55.

Simon 2002 *{published data only}*

Simon GE, Ludman E, Unutzer J, Bauer MS. Design and implementation of a randomized trial evaluating systematic care for bipolar disorder. *Bipolar Disorders* 2002;**4**(4):226-36.

Simon 2006 *{published data only}*

Simon G, Ludman EJ, Bauer MS, Unutzer J, Operskalski B. Long-term effectiveness and cost of a systematic care program for bipolar disorder. *Archives of General Psychiatry* 2006;**63**(5):500-8.

Smith 2003 *{published data only}*

Smith RC, Lein C, Collins C, Lyles JS, Given B, Dwamena FC, et al. Treating patients with medically unexplained symptoms in primary care. *Journal of General Internal Medicine* 2003;**18**(6):478-89.

Smith 2019 *{published data only}*

Smith, SN, Almirall D, Prenovost K, Liebrecht C, Kyle J, Eisenberg D, et al. Change in patient outcomes after augmenting a low-level implementation strategy in community practices that are slow to adopt a collaborative chronic care model: a cluster randomized implementation trial. *Medical Care* 2019;**57**(7):503-11.

Sousa 2013 *{published data only}*

Sousa SA, Corriveau D, Lee AF, Bianco LG, Sousa GM. The LORS-enabled dialogue: a collaborative intervention to promote recovery from psychotic disorders. Psychiatric Services (Washington, D.C.) 2013;**64**:58-64.

Steel *{published data only}*

Steel J, Kim K, Geller D, Brower D, Philips C, Ordos J, et al. A web-based collaborative care intervention for patients with advanced cancer [abstract]. https://www.cochranelibrary.com/central/doi/10.1002/central/CN-01062224/full.

Stewart 2014 *{published data only}*

Stewart JC, Perkins AJ, Callahan CM. Effect of collaborative care for depression on risk of cardiovascular events: data from the IMPACT randomized controlled trial. *Psychosomatic Medicine* 2014;**76**(1):29-37.

Sylvia 2013 *{published data only}*

Sylvia LG, Hay A, Ostacher MJ, Miklowitz DJ, Nierenberg AA, Thase ME, et al. Association between therapeutic alliance, care satisfaction, and pharmacological adherence in bipolar disorder. *Journal of Clinical Psychopharmacology* 2013;**33**(3):343-50.

Sylvia 2015 *{published data only}*

Sylvia LG, Shelton RC, Kemp DE, Bernstein EE, Friedman ES, Brody BD, et al. Medical burden in bipolar disorder: findings from the Clinical and Health Outcomes Initiative in Comparative Effectiveness for Bipolar Disorder study (Bipolar CHOICE). *Bipolar Disorders* 2015;**17**(2):212-23.

Tang 2010 *{published data only}*

Tang JYM, Wong GHY, Hui CLM, Lam MML, Chiu CPY, Chan SKW, et al. Early intervention for psychosis in Hong Kong -- the easy programme. *Early Intervention in Psychiatry* 2010;**4**(3):214-9.

Van der Feltz 2006 *{published data only}*

Van Der Feltz-Cornelis CM, Van Oppen P, Ader HJ, Van Dyck R. Randomised controlled trial of a collaborative care model with psychiatric consultation for persistent medically unexplained symptoms in general practice. *Psychotherapy and Psychosomatics* 2006;**75**(5):282-9.

van Orden 2009 *{published data only}*

van Orden M, Hoffman T, Haffmans J, Spinhoven P, Hoencamp E. Collaborative mental health care versus care as usual in a primary care setting: a randomized controlled trial. *Psychiatric Services* 2009;**60**(1):74-9.

Von Korff 1998 *{published data only}*

Von Korff M, Katon W, Bush T, Lin EH, Simon GE, Saunders K, et al. Treatment costs, cost offset, and cost-effectiveness of collaborative management of depression. *Psychosomatic Medicine* 1998;**60**(2):143-9.

Walker 2000 *{published data only}*

Walker EA, Katon WJ, Russo J, Von Korff M, Lin E, Simon G, et al. Predictors of outcome in a primary care depression trial. *Journal of General Internal Medicine* 2000;**15**(12):859-67.

Young 2010 *{published data only}*

Young AS, Noosha N, Cohen AN, Kessler C, McNagny K. The appropriateness of the routine medication treatment for schizophrenia. *Schizophrenia Bulletin* 2010;**36**(4):732-9.

References to ongoing studies

Aschbrenner 2019 *{published data only}*

Aschbrenner KA, Pratt SI, Bond GR, Zubkoff L, Naslund JA, Jue K, et al. A virtual learning collaborative to implement health promotion in routine mental health settings: protocol for a cluster randomized trial. *Contemporary Clinical Trials* 2019;**84**:105816.

NCT03891368. A virtual learning collborative to implement health promotion in routine mental health settings protocol. https://clinicaltrials.gov/ct2/show/NCT03891368.

Battersby 2018 *{published data only}*

Battersby M, Kidd MR, Licinio J, Aylward P, Baker A, Ratcliffe J, et al. Improving cardiovascular health and quality of life in people with severe mental illness: study protocol for a randomised controlled trial. *Trials* 2018;**19**(1):366.

Byng 2023 *{published data only}*

Baker E, Gwernan-Jones R, Britten N, Cox M, McCabe C, Retzer A, et al. Refining a model of collaborative care for people with a diagnosis of bipolar, schizophrenia or other psychoses in England: a qualitative formative evaluation. *BMC Psychiatry* 2019;**19**(1):7.

* Byng R, Creanor S, Jones B, Hosking J, Plappert H, Bevan S, et al. The effectiveness of a primary care-based collaborative care model to improve quality of life in people with severe mental illness: PARTNERS2 cluster randomised controlled trial. *British Journal of Psychiatry* 2023;**222**(6):246-56. [DOI: 10.1192/bjp.2023.28]

ISRCTN95702682. PARTNERS2: a cluster randomised control trial of a model of collaborative care for people with a diagnosis

of bipolar, schizophrenia or other psychoses. http://isrctn.com/ISRCTN95702682 2017.

Keeley T, Khan H, Pinfold V, Williamson P, Mathers J, Davies L, et al. Core outcome sets for use in effectiveness trials involving people with bipolar and schizophrenia in a community-based setting (PARTNERS2): study protocol for the development of two core outcome sets. *Trials* 2015;**16**:47. [DOI: 10.1186/s13063-015-0553-0]

PARTNERS2 writing collective. Exploring patient and public involvement (PPI) and co-production approaches in mental health research: learning from the PARTNERS2 research programme. *Research Involvement and Engagement* 2020;**6**:56.

Plappert H, Hobson-Merrett C, Gibbons B, Baker E, Bevan S, Clark M, et al. Evaluation of a primary care based collaborative care model (PARTNERS2) for people with diagnoses of schizophrenia, bipolar or other psychoses: study protocol for a cluster randomised controlled trial. *BJGP* 2021;**5**(3):BJGPO.2021.0033. [DOI: 10.1186/s13063-015-0553-0]

Reilly S, McCabe C, Marchevsky N, Green M, Davies L, Ives N, et al. Status of primary and secondary mental healthcare of people with severe mental illness: an epidemiological study from the UK PARTNERS2 programme. *British Journal of Psychiatry Open* 2021;**7**(2):e53. [DOI: 10.1192/bjo.2021.10]

Retzer A, Sayers R, Pinfold V, Gibson J, Keeley T, Taylor G, et al. Development of a core outcome set for use in community-based bipolar trials—a qualitative study and modified Delphi. *PLOS One* 2020;**15**(10):e0240518.

Fields 2019 *{published data only}*

Fields L, Callaway CA, Park ER, Nierenberg AA, Greer J, Temel JS, et al. Randomized trial protocol of bridge intervention for patients with serious mental illness and cancer. *Journal of Clinical Oncology* 2018;**36**(34 Suppl):153.

* Irwin KE, Park ER, Fields LE, Corveleyn AE, Greer JA, Perez GK, et al. Bridge: person-centered collaborative care for patients with serious mental illness and cancer. *The Oncologist* 2019;**24**:901-10.

Hanlon 2014 *{published data only}*

Hanlon C, Alem A, Medhin G, Shibre T, Ejigu DA, Negussie H, et al. Task sharing for the care of severe mental disorders in a low-income country (TaSCS): study protocol for a randomised, controlled, non-inferiority trial. *Trials* 2016;**17**(1):76.

Lund C, Alem A, Schneider M, Hanlon C, Ahrens J, Bandawe C, et al. Generating evidence to narrow the treatment gap for mental disorders in sub-Saharan Africa: rationale, overview and methods of AFFIRM. *Epidemiology and Psychiatric Sciences* 2015;**24**(3):233-40.

NCT02308956. Task sharing for the care of severe mental disorders in a low-income country (TaSCS). https://Clinicaltrials.gov/ct2/show/NCT02308956 2014.

Happell 2018 *{published data only}*

ACTRN12618000678291. Improving the cardiometabolic health of people with psychosis: The Physical Health Nurse Consultant service. http://www.anzctr.org.au/ACTRN12618000678291.aspx 2018.

Happell B, Curtis J, Banfield M, Goss J, Niyonsenga T, Watkins A, et al. Improving the cardiometabolic health of people with psychosis: a protocol for a randomised controlled trial of the Physical Health Nurse Consultant service. *Contemporary Clinical Trials* 2018;**73**:75-80.

Nicole 2018 *{published data only}*

NCT03695289. Interactive Obesity Treatment Approach (iOTA) for obesity prevention in serious mental illness. https://ClinicalTrials.gov/ 2018.

Additional references

Adams 2004

Adams G, Gulliford MC, Ukoumunne OC, Eldridge S, Chinn S, Campbell MJ. Patterns of intra-cluster correlation from primary care research to inform study design and analysis. *Journal of Clinical Epidemiology* 2004;**57**(8):785-94. [DOI: 10.1016/j.jclinepi.2003.12.013]

Altman 1996

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

Altman 1997

Altman EG, Hedeker D, Peterson JL, Davis JM. The Altman Self-Rating Mania Scale. *Biological Psychiatry* 1997;**42**:948-55.

Archer 2012

Archer J, Bower P, Gilbody S, Lovell K, Richards D, Gask L, et al. Collaborative care for depression and anxiety problems. *Cochrane Database of Systematic Reviews* 2012, Issue 10. Art. No: CD006525. [DOI: 10.1002/14651858.CD006525.pub2]

Awad 1993

Awad AG. Subjective response to neuroleptics in schizophrenia. *Schizophrenia Bulletin* 1993;**19**:609-18.

Baker 2019

Baker E, Gwernan-Jones R, Britten N, Cox M, McCabe C, Retzer A, et al. Refining a model of collaborative care for people with a diagnosis of bipolar, schizophrenia or other psychoses in England: a qualitative formative evaluation. *MBC Psychiatry* 2019;**19**(7):7. [DOI: 10.1186/s12888-018-1997-z]

Bauer 1991

Bauer MS, Crits-Christoph P, Ball WA, Dewees E, McAllister T, Alahi P, et al. Independent assessment of manic and depressive symptoms by self-rating. Scale characteristics and Implications for study mania. *Archives of General Psychiatry* 1991;**48**:807-12.

Bauer 2001

Bauer MS. The collaborative practice model for bipolar disorder: design and implementation in a multi-site randomized controlled trial. *Bipolar Disorders* 2001;**3**:233-44.

Bauer 2009

Bauer MS, Biswas K, Kilbourne A. Enhancing multiyear guideline concordance for bipolar disorder through collaborative care. *American Journal of Psychiatry* 2009;**166**(11):1244-50.

Bindman 2000

Bindman J, Johnson S, Wright S, Szmukler G, Bebbington P, Kuipers E, et al. Integration between primary and secondary services in the care of the severely mentally ill: patients and general practitioners views. *British Journal of Psychiatry* 1997;**171**:169-74.

Bland 1997

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

Blount 1998

Blount A. Introduction to integrated primary care. In: Blount A, editors(s). Integrated Primary Care: The Future of Medical and Mental Health Collaboration. New York: Norton, 1998:1-43.

Bodenheimer 2002

Bodenheimer T, Wagner EH, Grumbach K. Improving primary care for patients with chronic illness. *JAMA* 2002;**288**(14):1775-9.

Boutron 2021

Boutron I, Page MJ, Higgins JPT, Altman DG, Lundh A, Hróbjartsson A (on behalf of the Cochrane Bias Methods Group). Chapter 7: Considering bias and conflicts of interest among the included studies. In: Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 6.2 (updated February 2021). Cochrane, 2021. Available from training.cochrane.org/handbook.

Bouza 2010

Bouza C, López-Cuadrado T, Amate J M. Hospital admissions due to physical disease in people with schizophrenia: a national population-based study. *General Hospital Psychiatry* 2010;**32**(2):156-63.

Bower 2006

Bower P, Gilbody S, Richards D, Fletcher J, Sutton A. Collaborative care for depression in primary care. Making sense of a complex intervention: systematic review and meta-regression. *British Journal of Psychiatry* 2006;**189**:484-93.

Brown 2000

Brown S, Inskip H, Barraclough B. Causes of the excess mortality of schizophrenia. *British Journal of Psychiatry* 2000;**177**:212-7.

Brown 2010

Brown S, Kim M, Mitchell C, Skip H. Twenty five year mortality of a community cohort with schizophrenia. *British Journal of Psychiatry* 2010;**196**:116-21.

Burns 2000

Burns T, Greenwood N, Kendrick T, Garland C. Attitudes of general practitioners and community mental health team staff towards the locus of care for people with chronic psychotic disorders. *Primary Care Psychiatry* 2000;**6**:67-71.

Burns 2004

Burns T. Community mental health teams: a guide to current practices. Oxford University Press, 2004.

Butler 2008

Butler M, Kane RL, McAlpine D, Kathol RG, Fu SS, Hagedorn H, et al. Integration of Mental Health/ Substance Abuse and Primary Care (Prepared by the Minnesota Evidence-based Practice Center under Contract No. 290-02-0009). AHRQ Publication No. 09-E003 edition. Vol. **173**. Rockville, MD: Agency for Healthcare Research and Quality, 2008.

Campbell 2000

Campbell M, Fitzpatrick R, Haines A, Kinmonth AL, Sandercock P, Spiegelhalter D, et al. Framework for design and evaluation of complex interventions to improve health. *BMJ* 2000;**321**(7262):694-6.

CDC

Blood Pressure. https://www.cdc.gov/bloodpressure/facts.htm.

Centre for Public Mental Health 2006

Adult mental health service mapping: report Autumn 2004 and Spring 2006. Centre for Public Mental Health, Durham University 2006.

Colton 2006

Colton CW, Manderscheid RW. Congruencies in increased mortality rates, years of potential life lost, and causes of death among public mental health clients in eight states. *Preventing Chronic Disease* 2006;**3**(2):A42.

Connolly 2005

Connolly M, Kelly C. Lifestyle and physical health in schizophrenia. *Advances in Psychiatric Treatment* 2005;**11**:125-32.

CONSORT 2010a

Schulz KF, Altman DG, Moher D, for the CONSORT Group. CONSORT 2010 statement for reporting parallel group randomised trials. *BMJ* 2010;**340**:c332.

CONSORT 2010b

Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ, et al for the CONSORT Group. CONSORT 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomised trial. *BMJ* 2010;**340**:c869.

Craig 2008

Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: the new Medical Research Council guidance. *British Medical Journal* 2008;**337**:a1655. [DOI: 10.1136/bmj.a1655]

Craven 2006

Craven M, Bland R. Better practices in collaborative mental health care: an analysis of the evidence base. *Canadian Journal of Psychiatry* 2006;**51**(1):7S-72S.

Crawford 2004

Crawford M, Jonge E, Freeman G, Weaver T. Providing continuity of care for people with severe mental illness. *Social Psychiatry and Psychiatric Epidemiology* 2004;**39**:265-72.

Daumit 2002

Daumit G, Pratt L, Crum R, Powe N, Ford D. Characteristics of primary care visits for individuals with severe mentalillness in a national sample. *General Hospital Psychiatry* 2002;**24**:391–5.

Deeks 2008

Deeks JJ, Higgins JPT, Altman DG. Chapter 9: Analysing data and undertaking meta-analyses. In: Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews Version 5.0.1 (updated September 2008). The Cochrane Collaboration 2008.

Department of Health 2002

Mental health policy implementation guidance: community mental health teams. Department of Health, London 2002.

Dieterich 2010

Dieterich M, Irving CB, Park B, Marshall M. Intensive case management for severe mental illness. *Cochrane Database of Systematic Reviews* 2010, Issue 10. Art. No: CD007906. [DOI: 10.1002/14651858.CD007906]

Divine 1992

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

Druss 2005

Druss B, Esenwein S. Improving general medical care for persons with mental and addictive disorder: a systematic review. *General Hospital Psychiatry* 2005;**28**:145-53.

Egger 1997

Egger M, Davey-Smith G, Schneider M, Minder CSO. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997;**13**:629.

Ekers 2013

Ekers D, Murphy R, Archer J, Ebenezer C, Kemp D, Gilbody S. Nurse-delivered collaborative care for depression and long-term physical conditions: a systematic review and meta-analysis. *Journal of Affective Disorders* 2013;**149**(1-3):14-22.

Falkenström 2018

Falkenström F, Grant J, Holmqvist R. Review of organizational effects on the outcome of mental health treatments. *Psychotherapy Research* 2018;**28**(1):76-90. [DOI: 10.1080/10503307.2016.1158883]

Freeman 2002

Freeman G, Weaver T, Low J, Jonge E, Crawford M. Promoting continuity of care for people with severe mental illness whose needs span primary, secondary and social care: a multi-method investigation of relevant mechanisms and contexts. SDO, London 2002.

Furukawa 2002

Furukawa TA, Guyatt GH, Griffith LE. Can we individualize the 'number needed to treat'? An empirical study of summary effect measures in meta- analysis. *International Journal of Epidemiology* 2002;**31**(1):72-6.

Gask 2005

Gask L. Role of specialists in common chronic diseases. *BMJ* 2005;**19**(330):651-3.

Gask 2010

Gask L, Bower P, Lovell K, Escott D, Archer J, Gilbody S, et al. What work has to be done to implement collaborative care for depression? Process evaluation of a trial utilizing the normalisation process model. *Implementation Science* 2010;**5**(15):1-11. [DOI: 10.1186/1748-5908-5-15]

Gater 1997

Gater R, Goldberg D, Jackson G, Jennett N, Lowson K, Ratcliffe J, et al. The care of patients with chronic schizophrenia: a comparison between two services. *Psychological Medicine* 1997;**27**(6):1325-36.

Gilbody 2006

Gilbody S, Bower P, Fletcher J, Richards D, Sutton A. Collaborative care for depression: a cumulative meta-analysis and review of longer-term outcomes. *Archives of Internal Medicine* 2006;**166**:2314-21.

Glick 2003

Glick HA, McBride L, Bauer MS. A manic-depressive symptom self-report in optical scanable format. *Bipolar Disorders* 2003;**5**:366-9.

Goodrich 2013

Goodrich DE, Kilbourne AM, Nord KM, Bauer MS. Mental health collaborative care and its role in primary care settings. *Current Psychiatry Reports* 2013;**15**(8):383. [DOI: 10.1007/s11920-013-0383-2]

GRADEpro GDT [Computer program]

GRADEpro GDT. Version accessed 6 August 2016. Hamilton (ON): McMaster University (developed by Evidence Prime), 2015. Available at gradepro.org.

Grover 2014

Grover S, Shah R, Kulhara P, Malhotra R. Internal consistency & validity of Indian disability evaluation and assessment scale (IDEAS) in patients with schizophrenia. *Indian Journal of Medical Research* 2014;**140**(5):637-43. [PMID: PMC4311317]

Grundy 2018

Grundy SM, Stone NJ, Bailey AL, Beam C, Birtcher KK, Blumenthal RS, et al. ACC/AHA/AAACVPR/AAPA/ABC/ACPM/ADA/AGS/APhA/ASPC/NLA/PCNA Guideline on the management of blood cholesterol: a report of the American College of Cardiology/Foundation/American Heart Association Task Force on Clinical Practice Guidelines. *Circulation* 2018;**139**:e1082-143.

Gulliford 1999

Gulliford MC, Ukoumune OC, Chinn S. Components of variance and intraclass correlations for the design of community-based surveys and intervention studies: data from the health survey for England. *American Journal of Epidemiology* 1994;**149**:876-83.

Gunn 2006

Gunn J, Diggens J, Hegarty K, Blashki G. A systematic review of complex system interventions designed to increase recovery from depression in primary care. *BMC Health Services Research* 2006;**6**:1-11. [DOI: 10.1186/1472-6963-6-88]

Hahn 2005

Hahn S, Puffer S, Torgerson DJ, Watson J. Methodological bias in cluster randomised trials. *BMC Medical Research Methodology* 2005;**5**(10):1-8. [DOI: 10.1186/1471-2288-5-10]

Han 2021

Han L, Doran T, Holt RIG, Hewitt C, Jacobs R, Prady SL, et al. Impact of severe mental illness on healthcare use and health outcomes for people with type 2 diabetes: a longitudinal observational study in England. *British Journal of General Practice* 2021;**71**(709):e565-73. [DOI: 10.3399/BJGP.2020.0884]

Harris 1998

Harris EC, Barraclough B. Excess mortality of mental disorder. *British Journal of Psychiatry* 1998;**173**:11–53.

Healy 2016

Healy D. Psychiatric Drugs Explained . 6th edition. London: Elsevier , 2016.

Heatherton 1991

Heatherton TF, Kozlowski LT, Frecker RC, Fragerstrom KO. The Fragerstrom test for nicotine dependence: a revision of the Fagerstrom tolerance questionnaire. *British Journal of Addiction* 1991;**86**:1119-27.

Higgins 2003

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

Higgins 2008

Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews of Interventions. 1st edition. Chichester: John Wiley & Sons, 2008.

Higgins 2011

Higgins JPT, Deeks JJ (editors). Chapter 7: Selecting studies and collecting data. In: Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Intervention Version 5.0.1 (updated March 2011). The Cochrane Collaboration, 2011. Available from training.cochrane.org/handbook/archive/v5.1/.

Higgins 2020

Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 6.1 (updated September 2020). Cochrane, 2020. Available from www.training.cochrane.org/handbook.

Higgins 2021

Higgins JPT, Savović J, Page MJ, Elbers RG, Sterne JAC. Chapter 8: Assessing risk of bias in a randomized trial. In: Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 6.2 (updated February 2021). Cochrane, 2021. Available from training.cochrane.org/handbook.

Hoffman 2014

Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ* 2014;**348**:g1687. [DOI: 10.1136/bmj.g1687]

Hogan 1983

Hogan TP, Awad AG, Eastwood R. A self-report scale predictive of drug compliance in schizophrenics: reliability and discriminative validity. *Psychological Medicine* 1983;**13**:177-83.

Huang 2013

Huang Y, Wei X, Wu T, Chen R, Guo A. Collaborative care for patients with depression and diabetes mellitus: a systematic review and meta-analysis. *BMC Psychiatry* 2013;**14**(13):260.

Kai 2000

Kai J, Crosland A, Drinkwater C. Prevalence of enduring and disabling mental illness in the inner city. *British Journal of General Practice* 2000;**50**:922-4.

Katon 1995

Katon W, Von Korff M, Lin E, Walker E, Simon GE, Bush T, et al. Collaborative management to achieve treatment guidelines. Impact on depression in primary care. *JAMA* 1995;**13**:1026-31.

Katon 2001

Katon W, Von Korff M, Lin E, Simon GE. Rethinking practitioner roles in chronic illness: the specialist primary care physician and the practice nurse. *General Hospital Psychiatry* 2001;**23**:138-44.

Katon 2010

Katon WJ, Lin EHB, Von Korff M, Ciechanowski P, Ludman EJ, Young B, et al. Collaborative care for patients with depression and chronic illnesses. *New England Journal of Medicine* 2010;**363**:2611-20.

Kay 1987

Kay SR, Fiszbein A, Opler LA. The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophrenia Bulletin* 1987;**13**:261-76.

Keller 1987

Keller M, Lavori P, Friedman B. The longitudinal interval follow-up evaluation. *Archives of General Psychiatry* 1987;**44**:540-8.

Kendrick 1991

Kendrick T, Sibbald B, Burns T, Freeling P. Role of general practitioners in care of the long term mentally ill patients. *BMJ* 1991;**302**:508-10.

Kilbourne 2008

Kilbourne AM, Post EP, Nossek A, Drill L, Cooley S, Bauer MS. Improving medical and psychiatric outcomes individuals with bipolar disorder. *Psychiatric Services* 2008;**59**(7):760-8.

Kinderman 2014

Kinderman P. A Prescription for Psychiatry: Why We Need a Whole New Approach to Mental Health and Wellbeing. Hampshire: Palgrave Macmillian, 2014.

Kingdon 1989

Kingdon D. Mental health services: results of a survey of English district plans. *Psychiatric Bulletin* 1989;**13**:77-8.

Kisely 2007

Kisely S, Smith M, Lawrence D, Cox M, Campbell MA, Maaten S. Inequitable access for mentally ill patients to some medically necessary procedures. *Canadian Medical Association Journal* 2007;**176**(6):779-84.

Kroenke 2001

Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *Journal of General Internal Medicine* 2001;**16**(9):606-13.

Lang 1997

Lang F, Johnstone E, Murray D. Service provision for people with schizophrenia. Role of the general practitioner. *British Journal Psychiatry* 1997;**171**:165-8.

Lefebvre 2019

Lefebvre C, Glanville J, Briscoe S, Littlewood A, Marshall C, Metzendorf M, et al. Searching for and selecting studies. In: Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, editors(s). Cochrane Handbook for Systematic Reviews of Interventions. 2nd edition. John Wiley and Sons, 2019:67-107. [DOI: 10.1002/9781119536604.ch4]

Lester 2005

Lester H. Shared care for people with mental illness - a GP perspective. *Advances in Psychiatric Treatment* 2005;**11**:133-41.

Leverich 1998

Leverich GS, Post RM. Life Charting of affective disorders. *CNS Spectrums* 1998;**3**:21-37.

Lin 2021

Lin LY, Jochym N, Merz JF. Refusal rates and waivers of informed consent in pragmatic and comparative effectiveness RCTs: a systematic review. *Contemporary Clinical Trials* 2021;**104**:106361.

Malone 2007

Malone D, Marriott S, Newton-Howes G, Simmonds S, Tyrer P. Community mental health teams (CMHTs) for people with severe mental illnesses and disordered personality. *Cochrane Database of Systematic Reviews* 2007, Issue 3. Art. No: CD000270. [DOI: 10.1002/14651858.CD000270.pub2]

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.

McGinty 2021

McGinty EE, Presskreischer R, Breslau J, Brown JD, Domino ME, Druss BG, et al. Improving physical health among people with serious mental illness: the role of the specialty mental health sector. Psychiatric Services 2021 Jun 2 [Epub ahead of print]. [DOI: 10.1176/appi.ps.202000768] [PMID: 34074150]

Mental Health Taskforce 2016

NHS England. The Five Year Forward View for Mental Health. https://www.england.nhs.uk/publication/the-five-year-forward-view-for-mental-health/ 2016. [HTTPS:: //www.england.nhs.uk/wp-content/uploads/2016/02/ Mental-Health-Taskforce-FYFV-final.pdf]

Merikangas 2011

Merikangas KR, Jin R, He J-P, Kessler RC, Lee S, Sampson NA, et al. Prevalence and correlates of bipolar spectrum disorder in the World Mental Health Survey Initiative. *Archives of General Psychiatry* 2011;**68**(3):241-51. [DOI: 10.1001/archgenpsychiatry.2011.12]

Miller 2006

Miller BJ, Paschall CB 3rd, Svendsen DP. Mortality and medical comorbidity among patients with SMI. *Psychiatric Services* 2006;**57**(10):1482-7.

Moncrieff 2013

Moncrieff J. The Bitterest Pills. Basingstoke: Palgrave Macmillian , 2013.

Moncrieff 2015

Moncrieff J. Antipsychotic maintenance treatment: time to rethink? *PLOS Medicine* 2015;**12**(8):e1001861.

Moreno-Küstner 2018

Moreno-Küstner B, Martín C, Pastor L. Prevalence of psychotic disorders and its association with methodological issues. A systematic review and meta-analyses. *PLOS One* 2018;**13**(4):e0195687. [DOI: 10.1371/journal.pone.0195687]

Morisky 1986

Morisky DE, Green LW, Levine DM. Concurrent and predictive validity of a self-reported measure of medication adherence. *Med Care* 1986;**24**:67-74.

Muntingh 2016

Muntingh ADT, Van Der Feltz-Cornelis CM, van Marwijk HWJ, Spinhoven P, Van Balkom AJLM. Collaborative care for anxiety disorders in primary care: a systematic review and meta-analysis. *BMC Family Practice* 2016;**17**(1):62. [ISSN: 1471-2296]

Nelson 1999

Nelson CB, Lofty M. The World Health Organization's WHOQOL-Bref quality of life assessment of psychometric qualities. Results of field trial. WHO (MNH/MHP/99.7).

NHS England 2019a

NHS Long Term Plan. https://www.longtermplan.nhs.uk/ 2019. [HTTPS:: //www.longtermplan.nhs.uk]

NHS England 2019b

NHS England. NHS Mental Health Implementation Plan 2019/20 – 2023/24. https://www.longtermplan.nhs.uk/publication/nhs-mental-health-implementation-plan-2019-20-2023-24/ 2019. [HTTPS:: //www.longtermplan.nhs.uk/wp-content/uploads/2019/07/nhs-mental-health-implementation-plan-2019-20-2023-24.pdf]

NHS England 2019c

NHS England and NHS Improvement and the National Collaborating Central for Mental Health. The Community Mental Health Framework for Adults and Older Adults. In: NHS England and NHS Improvement and the National Collaborating Central for Mental Health. Vol. **1**. National Collaborating Central for Mental Health, 2019. [https://www.england.nhs.uk/wp-content/uploads/2019/09/community-mental-health-framework-for-adults-and-older-adults.pdf]

NICE 2009

Schizophrenia: core interventions in the treatment and management of schizophrenia in adults in primary and secondary care (CG82). London: National Institute for Health and Clinical Excellence. 2009.

NICE 2009b

NICE. NICE CG38. Bipolar Disorder. London: NICE, 2009.

Osborn 2006

Osborn D, Nazareth I, King M. Risk for coronary heart disease in people with severe mental illness: a cross sectional comparative study in primary care. *British Journal of Psychiatry* 2006;**188**:271-7.

Osby 2001

Ösby U, Brandt L, Correia N, Ekbom A, Sparen P. Excess mortality in bipolar and unipolar disorder in Sweden. *Archives of General Psychiatry* 2001;**58**:844-50.

Overall 1962

Overall JE, Gorham DR. The Brief Psychiatric Rating Scale. *Psychological Reports* 1962;**10**:709-812.

Peveler 1999

Peveler R, George C, Kinmonth AL, Campbell M, Thompson C. Effect of antidepressant drug counselling and information leaflets on adherence to drug treatmentin primary care: randomised controlled trial. *BMJ* 1999;**319**:612-5.

Planner 2016

Planner C. Collaborative care for severe mental illness: a mixed-methods exploration of potential to improve healthcare delivery and health outcomes. https://research.manchester.ac.uk/en/studentTheses/collaborative-care-for-severe-mental-illness-a-mixed-methods-expl 2016.

Plappert 2021

Plappert H, Hobson-Merrett C, Gibbons B, Baker E, Bevan S, Clark M, et al. Evaluation of a primary care-based collaborative care model (PARTNERS2) for people with diagnoses of schizophrenia, bipolar, or other psychoses: study protocol for a cluster randomised controlled trial. *BJGP Open* 2021;**5**(3):BJGPO.2021.0033. [DOI: 10.3399/BJGPO.2021.0033]

QOF 2012

Quality and Outcomes Framework. http://www.hscic.gov.uk/catalogue/PUB08661 (accessed 26 July 2013):The NHS Information Centre.

Reilly 2012

Reilly S, Planner C, Hann M, Reeves D, Nazareth I, Lester H. The role of primary care in service provision for people with severe mental illness: a cross sectional cohort study. *PLOS One* 2012;**7**(5):e36468. [DOI: 10.1371/journal.pone.0036468]

Reilly 2015

Reilly S, Olier I, Planner C, Doran T, Reeves D, Ashcroft DM, et al. Inequalities in physical co-morbidity: a longitudinal comparative observational study of people with severe mental illness in the UK. *BMJ Open* 2015;**5**(12):e009010. [DOI: 10.1136/bmjopen-2015-009010]

Reilly 2021

Reilly S, McCabe C, Marchevsky N, Green M, Davies L, Ives N, et al. Status of primary and secondary mental health care of people with severe mental illness: an epidemiological study from the UK PARTNERS2 study. *British Journal of Psychiatry Open* 2021;**7**(2):E53. [DOI: 10.1192/bjo.2021.10]

Retzer 2020

Retzer A, Sayers R, Pinfold V, Gibson J, Keeley T, Taylor G, et al. Development of a core outcome set for use in community-based bipolar trials - a qualitative study and modified Delphi. *PLOS One* 2020;**15**(10):e0240518.

Review Manager 5 (RevMan 5) [Computer program]

Review Manager 5 (RevMan 5). Version 5.3. Copenhagen: Nordic Cochrane Centre, The Cochrane Collaboration, 2014.

RevMan [Computer program]

Review Manager (RevMan). Version 7.2.0. The Cochrane Collaboration, 2024. Available at revman.cochrane.org.

Roberts 2021

Roberts MT, Shokraneh F, Sun Y, Groom M, Adams CE. Classification of psychotherapy interventions for people with schizophrenia: development of the Nottingham Classification of Psychotherapies. *Evidence Based Mental Health* 2021;**24**(2):62-9.

Rosa 2007

Rosa AR, Sanchez-Moreno J, Martinez-Aran A, et al. Validity and reliability of the functioning assessment short test (FAST) in bipolar disorder. *Clinical Practice and Epidemiology in Mental Health* 2007;**3**:5.

Rush 2003

Rush AJ, Trivedi MH, Ibrahim HM, Carmody TJ, Arnow B, Klein DN, et al. The 16-item Quick Inventory of Depression Symptomatology (QUIDS), clinical rating (QUIDS-C), and self report (QUIDS-SR): a psychometric evaluation in patients with chronic major depression. *Biological Psychiatry* 2003;**54**:573-83.

Saha 2005

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

Sainsbury Centre for Mental Health 1998

Keys to engagement: review of care for people with severe mental illness who are hard to engage with services. Sainsbury Centre for Mental Health, London 1998.

Saxena 2006

Saxena S, Sharan P, Garrido M, Saraceno B. World Health Organization's Mental Health Atlas 2005: implications for policy development. *World Psychiatry* 2006;**5**(3):179-84.

Sayce 1991

Sayce L, Craig TKJ, Boardman AP. The development of community mental health centres in the UK. *Social Psychiatry and Psychiatric Epidemiology* 1991;**26**:14-22.

Schooler 1979

Schooler NR, Hogarty GE, Weissman MM. Social Adjustment Scale II. In: Hargreaves WA, Attkisson CC, Sorenson JE, editors(s). Resource Materials for Community Mental Health Evaluators. Washington DC: Department of Health, Education, and Welfare, 1979:290-330.

Schünemann 2020

Schünemann HJ, Higgins JPT, Vist GE, Glasziou P, Akl EA, Skoetz N, Guyatt GH. Chapter 14: Completing 'Summary of findings' tables and grading the certainty of the evidence. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.1 (updated September 2020). Cochrane, 2020. Available from www.training.cochrane.org/handbook.

Shokraneh 2017

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

Shokraneh 2019

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

Shokraneh 2020

Shokraneh F, Adams CE. Cochrane Schizophrenia Group's Study-Based Register of Randomized Controlled Trials: Development and Content Analysis. *Schizophrenia Bulletin Open* 2020;**1**(1):sgaa061.

Shokraneh 2021

Shokraneh F, Adams CE. Classification of all pharmacological interventions tested in trials relevant to people with schizophrenia: a study-based analysis. Health Information and Libraries Journal 2021 Feb 18 [Epub ahead of print].

Singleton 2001

Singleton N, Bumpstead R, O'Brien M, Lee A, Meltzer H. Psychiatric Morbidity among Adults Living in Private Households 2000. London: TSO, 2001.

Social Exclusion Unit 2004

Social Exclusion Unit. Mental health and social exclusion: social exclusion unit report. Office of the Deputy Prime Minister, London 2004.

Sterne 2011

Sterne JAC, Egger M, Moher D. Chapter 10: Addressing reporting biases. In: Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from training.cochrane.org/handbook/archive/v5.1/.

Sterne 2019

Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 2019;**366**:l4898.

Tew 2012

Tew J, Ramon S, Slade M, Bird V, Melton J, Le Boutillier JC. Social factors and recovery from mental health difficulties: a review of the evidence. *British Journal of Social Work* 2012;**42**:443-60.

Thara 2002

Thara R. IDEAS (Indian Disability Evaluation and Assessment Scale) - a scale for measuring and quantifying disability in mental disorders. India. The Indian Psychiatry Society.

Thompsom 2000

Thompson K, Kulkarni L, Sergejew AA. Reliability and Validity of a new Medication Adherence Rating Scale (MARS) for the psychoses. *Schizophrenia Research* 2000;**42**:241-7.

Trompenaars 2005

Trompenaars FJ, Masthoff ED, Van Heck GL, Hodiamont PP, De Vries J, . Content validity, construct validity, and reliability of the WHOQOL-Bref in a population of Dutch adult psychiatric outpatients. *Quality of Life Research* 2005;**14**:151-60.

Unnebrink 2001

Unnebrink K, Windeler J. Intention-to-treat methods for dealing with missing values in clinical trials of progressively deteriorating diseases. *Statistics in Medicine* 2001;**20**:3931-46.

Ustun 2010

Ustun TB, World Health Organization. Measuring health and disability: Manual for WHO Disability Assessment Schedule WHODAS 2.0. Geneva: World Health Organization. https://www.who.int/classifications/icf/whodasii/en/index3.html 2010.

Ware 1992

Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36): I. Conceptual framework and item selection. *Medical Care* 1992;**30**:473-83.

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**(3):220-33.

Wells 2000

Wells KB, Sherbourne C, Schoenbaum M, Duan N, Meredith L, Unützer J, et al. Impact of disseminating quality improvement programmes for depression in managed primary care: a randomized controlled trial [published correction appears in JAMA. 2000;283:3204]. *JAMA* 2000;**283**(2):212-20.

Whitty 2020

Whitty C. Chief Medical Officer's Annual Report 2020. Health trends and variation in England. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/945929/Chief_Medical_Officer_s_annual_report_2020_-_health_trends_and_variation_in_England.pdf 2020.

WHO 2001

World Health Organization. Mental health: new understanding, new hope. Geneva: World Health Organization 2001.

WHO 2009

World Health Organization, Wonca. Integrating Mental Health into Primary Care: a Global Perspective. 1 edition. Geneva: WHO, 2009.

WHO 2016

WHO | mhGAP Intervention Guide for mental, neurological and substance use disorders in non-specialized health settings. https://www.who.int/publications/i/item/9789241549790.

Williamson 2017

Williamson PR, Altman DG, Bagley H, Barnes KL, Blazeby JM, Brookes ST, et al. The COMET Handbook: version 1.0. *Trials* 2017;**18**(3):280.

Wilson 1998

Wison PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. *Circulation* 1998;**97**:1837-47.

Wunderink 2017

Wunderink L. Who needs antipsychotic maintenance treatment and who does not? Our need to profile and personalize the treatment of first episode psychosis.. *Schizophrenia Research* 2017;**197**:65-6.

Xia 2009

Xia J, Adams CE, Bhagat N, Bhagat V, Bhoopathi P, Pinfold V, et al. Losing participants before the trial ends erodes credibility of findings. *Psychiatric Bulletin* 2009;**33**:254-7.

References to other published versions of this review

Reilly 2013

Reilly S, Planner C, Gask L, Hann M, Knowles S, Druss B, Lester H. Collaborative care approaches for people with severe mental illness. *Cochrane Database of Systematic Reviews* 2013, Issue 11. Art. No: CD009531. [DOI: 10.1002/14651858.CD009531.pub2]

* Indicates the major publication for the study

CHARACTERISTICS OF STUDIES

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

Bauer 2006

Study characteristics

Methods	**Allocation:** randomised **Design:** RCT, multicentre (11 outpatient VAMC clinics) **Duration:** 36-month (156-week) follow-up **Date of study:** July 1997 to December 2003 **Country:** USA **Setting:** outpatient clinics at VAMC **Recruitment method:** Potential participants were identified during acute hospitalisation for bipolar disorder and randomly assigned at discharge to either continue usual outpatient care or receive care in the intervention clinic for 3 years.

Bauer 2006 *(Continued)*

	Masking: none. "Because participants could not be blinded to the intervention, we could not guarantee blinding of the research assistants."
Participants	**Inclusion criteria:** diagnosis of bipolar disorder type I or II by criteria on the Structured Clinical Interview for Axis I DSM-IV Disorders; all psychiatric and medical comorbidities were allowed except as specified below; index episode of manic, major depressive or mixed episode, by DSM-IV criteria, requiring hospitalisation on an acute psychiatric ward; at least 2 hospitalisations on acute psychiatric wards more than 3 months apart over the prior 5 years **Exclusion criteria:** moderate to severe dementia, with a Mini-Mental State Examination score of ≤ 26; unresolved substance intoxication or withdrawal; hospitalisation on chronic or acute psychiatric wards for 6 or more months in the past year; ongoing enrolment in mental health programmes with a mobile outreach component in which clinical caregivers deliver services to the patient in the community; terminal medical illness with less than 3 years of expected longevity; unable or unwilling to give informed consent or in other ways unable to complete study requirements; participation in another concurrent experimental mental health or medical-surgical treatment protocol **Number randomised to intervention and control:** 330, 166 intervention, 164 control **Number completed study:** 306 **Age:** 46.6 years mean (SD 10.1), not reported by control vs intervention **Sex:** female 28 (9%), not reported by control vs intervention **Diagnosis:** bipolar 1 265 (87%); bipolar 2 41 (23%), not reported by control vs intervention **Ethnicity:** minority 71 (23%), not reported by control vs intervention **Any significant differences between intervention and control groups?** Participants in the intervention and usual care arms of the study did not differ in demographic or clinical characteristics except that intervention participants were somewhat older, less likely to have had a prior suicide attempt and more likely to have a diagnosis of a substance use disorder over their lifetime. Current substance disorder prevalence did not differ between groups.
Interventions	**Type of collaborative care:** B **Description of intervention:** Intervention name: collaborative care for bipolar disorder Contains 3 of 4 elements of collaborative care: 1. *A multi-professional approach to patient care:* no, there is no primary carer provider collaboration, instead the collaboration is between the nurse and the patient 2. *A structured management plan:* yes, clear protocols and algorithms for each part of the programme 3. *Scheduled patient follow-ups:* yes, there are scheduled follow-up plans with a minimum of once every 3-monthly contact 4. *Enhanced interprofessional communication:* yes, nurse manages communication with other healthcare providers and patient Other key elements of the intervention: • Psycho-education delivered to participants to encourage active self-management and monitoring of symptoms and functioning • Health promotion activities • Collaborative definition of problems • Joint goal-setting and planning • Practice guidelines for healthcare providers • Delivered in an outpatient specialist mental health clinic by a psychiatrist (0.25 FTE) and NCC (0.5 FTE)

Bauer 2006 *(Continued)*

	Description of control: Usual care; participants continued with their previous psychiatrist or were assigned one if new to VA. Clinicians who cared for participants in usual care did not care for those in collaborative care.
Outcomes	**Measures taken at:** not clearly specified. States, "The outcome battery was administered in 45 to 75 minutes every eight weeks and covered clinical and functional outcome, quality of life, non-VA clinical service use, and selected process measures"; however, some contradicting information is indicated below: **Primary outcomes:** 1) manic symptom score; 2) depressive symptom score; 3) total treatment costs **Able to use:** • Psychiatric admissions • Other hospital admissions (any reason) • Death (all causes and suicide) • Quality of life (SF-36 mental component and SF-36 physical component) (every 24 weeks) • Cost (mean intervention costs) (36 months) • Attrition (number lost to follow-up) (36 months) **Unable to use:** • Clinical outcome (Longitudinal Interval Follow-up Examination – LIFEscale) (every 8 weeks) – no mean or SD reported • Social functioning (Social Adjustment Scale II) (every 8 weeks) - no mean or SD reported • Intensity of bipolar-specific pharmacotherapy (adaptation of the National Institute of Mental Health Collaborative Study instrument) (every 24 weeks) • Patient satisfaction (Patient Satisfaction Index) - not eligible for inclusion in the review • Costs (direct all-treatment costs, psychiatric inpatient costs, inpatient costs, medical surgical inpatient costs, outpatient costs) - data not reported fully
Notes	**% lost to follow-up:** The overall protocol completion rate to week 156 was 80% and did not differ by survival analysis between intervention and usual care (respectively, 75% and 85%) or by mean retention in the protocol (123.5 ± 50.4 compared with 120.2 ± 52.0 weeks). Early terminators did not differ from completers in gender, age, homelessness, prior suicide attempts or psychosis. Ninety-six percent of all cost data points were available. Deaths did not differ (intervention, 12 deaths among 166 participants (7%); usual care, 8 deaths among 164 (5%)). There were 12 medical deaths, 4 accidents, 1 suicide (usual care participant) and 3 deaths from unknown causes. Standard deviations were imputed from the figures reported by study authors.

Chatterjee 2011

Study characteristics	
Methods	**Allocation:** stratified randomised (parallel-group); randomly assigned in a 2:1 ratio **Design:** RCT, multicentre - 3 sites in India: 4 sub-districts of Kancheepuram district, Tamil Nadu, Goa and Satara district in Maharashtra **Duration:** 12-month follow-up **Date of study:** May 2008 to December 2012

Chatterjee 2011 *(Continued)*

	Country: India **Setting:** intervention delivered within community **Recruitment method:** recruited through collaborating psychiatrists **Masking:** Outcome assessors were masked to allocation. Incidences of unmasking were recorded by researchers. If unmasking happened at the 6-month assessment, a different researcher undertook the 12-month assessment.
Participants	**Inclusion criteria:** aged 16 to 60; primary diagnosis of schizophrenia as per ICD-10 criteria; have had illness duration of at least 12 months and a moderate severity rating as rated on the Clinical Global Impression-Schizophrenia (CGI-SCH) scale; be residing within the study catchment area for the next 12 months **Exclusion criteria:** none described **Number randomised to intervention and control:** 282 (187 intervention, 95 control) **Number completed study:** 167 intervention (10 refused, 8 were not found or moved, 2 died), 86 control (6 refused, 1 not found, 2 died) **Age:** 16 to 60; intervention mean 36.2 (SD 10.2), control mean 35.6 (10.4) **Sex:** intervention 86 (46%) female, control 47 (49%) female **Diagnosis:** schizophrenia (ICD-10-DCR criteria) **Ethnicity:** reported as castes due to location Intervention: schedule caste 46 (25%), schedule tribe 4 (2%), other backward caste 45 (24%), unknown 18 (10%), none of the above 74 (40%) Control: schedule caste 20 (21%), schedule tribe 2 (2%), other backward caste 28 (29%), unknown 15 (16%), none of the above 30 (32%) **Any significant differences between intervention and control groups?** Not reported by authors, but demographics appear to be well-balanced in most cases.
Interventions	**Type of collaborative care:** B **Description of intervention:** Intervention name: community-based collaborative care + usual facility-based care Contains 3 of 4 elements of collaborative care: 1. *A multi-professional approach to patient care:* no, no primary care involvement 2. *A structured management plan:* yes, an individual treatment plan formulated in collaboration with the patient and family during the first 3 months 3. *Scheduled patient follow-ups:* yes; 6 to 8 patient visits at home in first 3 months (intensive engagement phase), 6 to 8 fortnightly sessions in the months 4 to 7 (stabilisation phase) and 6 visits between months 8 and 12 (maintenance phase) 4. *Enhanced interprofessional communication:* yes, community healthcare workers (CHWs) delivered intervention supervised by psychiatric social workers working as designated intervention co-ordinators. Psychiatrists provide clinical leadership for the community care teams, and ongoing supervision. Joint on-site visits, weekly group meetings and scheduled meetings with the psychiatrist. Other intervention components: • Structured clinical reviews by treatment team and supervision for community health workers • Psychoeducational information for participants and caregivers • Adherence management strategies

Chatterjee 2011 *(Continued)*

- Health promotion strategies to address physical health needs
- Individualised rehabilitation strategies to improve personal, social and work functioning of participants
- Specific efforts with participants and caregivers to deal with experiences of stigma and discrimination
- Linkage to self-help groups and other methods of user-led support
- Networks with community agencies to address social issues, to help with social inclusion, access to legal benefits, and employment opportunities.

The intervention is primarily delivered by the CHW. CHW have a minimum of 10 years of schooling and are trained in the intervention over a 6-week period and assessed for competence. The CHWs are co-ordinated and supervised by psychiatric social workers trained in supervision and monitoring skills. Treating psychiatrists also supervised through quarterly team reviews and regular supervision.

Maximum caseload of CHW is expected to be 25. Each participant is expected to receive 22 contacts with the CHW across 12 months.

Description of control:

Facility-based care (usual care provided by mental health providers). Varies between sites due to lack of consistency in healthcare provision in India.

Outcomes

Measures taken at: baseline, 6 and 12 months

Primary outcomes: change in symptoms, change in disabilities

Able to use:

- Change in symptoms (Positive and Negative Symptom Scale - PANSS) (baseline and 12 months)
- Change in disabilities (Indian Disability Evaluation and Assessment - IDEAS) (baseline, 6 and 12 months)
- Psychiatric admissions
- Other hospital admissions (any reason)
- Social functioning (WHO Disability Assessment Scale)
- Cost-effectiveness and cost utility (total costs in dollars) (12 months)
- Deaths from suicide

Unable to use:

- Experiences of stigma and discrimination (Discrimination and Stigma Scale - DISC) (baseline and 12 months) – not of interest
- Knowledge and attitudes about schizophrenia (Knowledge about Schizophrenia Interview - KASI) (baseline and 12 months) – not of interest
- Burden of caring (Burden Assessment Schedule – BAS) (baseline and 12 months) – not of interest
- Carer experiences of stigma and discrimination (section extracted from the Family Interview Schedule – FIS) (baseline and 12 months) – not of interest
- Willingness to disclose mental illness (scale not reported) (baseline and 12 months) – not of interest
- The caregiver summary assessment of participant adherence (same scale as participants) (baseline and 12 months)
- Adherence with antipsychotic medication using a 5-point ordinal scale, a specially designed tool developed for the study (not validated) (baseline, 6 and 12 months, if receiving medication) – not of interest
- Quality life years (quality of life - EuroQOL EQ-5D) used but not reported
- Willingness to disclose mental illness (scale not reported) - not of interest
- Experience of internalised stigma (Alienation subscale of the Internalized Stigma of Mental Illness Scale – ISMI) - not of interest

Chatterjee 2011 *(Continued)*

Notes	—

Chwastiak 2018

Study characteristics	
Methods	**Allocation:** randomised controlled pilot study. Participants were randomised in a ratio of 1:1 and randomisation was stratified based on baseline treatment with insulin or with clozapine or olanzapine. **Design:** RCT, multicentre - 2 outpatient community mental health care clinics in Seattle **Duration:** 3-month follow-up **Date of study:** November 2013 to September 2015 **Country:** USA **Setting:** outpatient clinics at CMHC in Seattle **Recruitment method:** participants invited from one of the two participating CMHCs; no specification of method **Masking:** none specified
Participants	**Inclusion criteria:** adult (18 to 70 years); enrolled to receive mental health treatment at Harborview Mental Health Services or Downtown Emergency Services Mental Health Center; a diagnosis of type 2 diabetes mellitus or cardiovascular disease; haemoglobin A1c > 8 or BP > 140/90 **Exclusion criteria:** cognitive, hearing or language impairment that would preclude a subject from providing informed consent; current suicidality, homicidality or grave disability that requires psychiatric hospitalisation; current substance abuse or dependence, as defined by Structured Clinical Interview for Diagnostic and Statistical Manual of Mental Disorders (SCID) **Number randomised to intervention and control:** 35 randomised, 18 received collaborative care, 17 care as usual **Number completed study:** 29 completed all measures **Age:** intervention mean 51.5 (10.3); control mean 51.0 (9.1) **Sex:** intervention 11 (38.9%) female; control 12 (29.4%) female **Diagnosis:** intervention 8 (44.4%); control 8 (47.1%) schizophrenia or schizoaffective disorder **Ethnicity:** intervention 10 (55.6%); control 11 (64.7%) non-white **Any significant differences between intervention and control groups?** no statistically significant baseline demographic or clinical differences between the groups
Interventions	**Type of collaborative care:** A **Description of intervention:** Intervention name: TEAMcare treatment of diabetes Contains all 4 elements of collaborative care: 1. *A multi-professional approach to patient care:* yes, participants who received the collaborative care intervention for diabetes received care from a team including a CMHC nurse care manager, a

Chwastiak 2018 *(Continued)*

	CMHC psychiatrist, an advanced practice registered nurse who provided primary care on-site and an endocrinologist consultant 2. *A structured management plan:* yes, structured needs assessments at enrolment and every 3 months thereafter to develop matched individualised treatment plans 3. *Scheduled patient follow-ups:* yes, intervention participants had a 60-min nurse care manager for comprehensive health assessment and an individualised health plan, then 30 minute visits for the support of chronic illness self-management (including medication adherence, healthy nutrition and regular physical activity) every other week for 12 weeks then monthly thereafter for up to 6 months 4. *Enhanced interprofessional communication:* yes, a treat to target approach was used for diabetes and cardiovascular risk factors through weekly systematic caseload review with team endocrinologist and psychiatrist, focussing on patients not improving as expected. Other intervention components: • All clinical visits and team meetings were conducted on-site. • Diabetes education materials were modified to address the issues unique to patients with psychosis. • Nurses used evidence-based behavioural interventions (motivational interviewing and behavioural activation) to address barriers to self-management and co-ordinated care with primary care and specialty medical providers (typically in an organisation outside of the CMHC), the CMHC clinical team and community-based agencies. • Team members received training in the TEAMcare model by the original investigators from the University of Washington. **Description of control:** Continued access to usual mental health treatment through CMHC and their usual care for diabetes
Outcomes	**Measures taken at:** baseline and 3 months **Primary outcome:** haemoglobin A1c levels **Able to use:** • Symptoms (Brief Psychiatric Rating Scale – BPRS) (baseline and 3 months) • Blood pressure (systolic BP mmHg) (baseline and 3 months) • Body mass index (BMI kg/m^2) (baseline and 3 months) • Low-density lipoprotein (LDL levels mg/dL) (baseline and 3 months) • Triglycerides (mg/dL) (baseline and 3 months) • Haemoglobin A1c levels (HbA1C %) (baseline and 3 months) • Deaths from natural causes • Attrition (number lost to follow up) (6 months) **Unable to use:** • Patient Health (Patient Health Questionnaire – PHQ-9) (baseline and 3 months) – not of interest • Smoking status (current smoker %) (baseline and 3 months) – not of interest • Nicotine dependence (Fagerstrom Nicotine Dependence Scale – FNDS) (baseline and 3 months) – not of interest
Notes	Attrition – 35 randomised, 29 patients completed study; 1 care as usual participant died of a cardiac event, 5 assumed dropouts/failed to complete all measures

Kilbourne 2012

Study characteristics	
Methods	**Allocation:** randomised in blocks of 16 to 20 stratified by age, race and diabetes diagnosis to ensure balance of characteristics **Design:** RCT, multicentre **Duration:** 12-month follow-up **Date of study:** October 2010 to August 2012 **Country:** USA **Setting:** 2 community-based mental health outpatient programmes in Southeastern Michigan **Recruitment method:** not specified **Masking:** single-blind (outcomes assessor blind to intervention)
Participants	**Inclusion criteria:** adult patients with an active diagnosis or treatment plan for bipolar disorder I, II or NOS with at least one cardiometabolic risk factor (diagnosis or indication of hypertension, hyperlipidaemia, diabetes or BMI > 25) who received care in one of 2 participating community mental health outpatient programmes. Other criteria included community-dwelling and English-speaking. **Exclusion criteria:** severely cognitively impaired or unable to give informed consent **Number randomised to intervention and control:** 68 participants enrolled and 32 randomised to the Life Goals Collaborative Care group (LGCC), 33 to enhanced treatment as usual (ETU) **Number completed study:** 65 completed both 6- and 12-month measures **Age:** intervention 47.2 ± 11.8; control 43.4 ± 13.6 **Sex:** intervention 15 (56%); control 21 (66%) female **Diagnosis:** breakdown by diagnosis not provided **Ethnicity:** intervention 7 (22%); control 5 (16%) African American **Any significant differences between intervention and control groups?** no statistically significant baseline demographic or clinical differences between groups
Interventions	**Type of collaborative care:** B **Description of intervention:** Intervention name: Life Goals Collaborative Care Contains 3 elements of collaborative care: 1. *A multi-professional approach to patient care:* no involvement with primary care 2. *A structured management plan:* yes, interventionist provided four 2-hour weekly group self-management sessions, followed by brief care management contact with patients randomised to LGCC for up to 6 months. Each group session included approximately 8 to 10 participants, and sessions were based on social cognitive theory. The sessions included active discussions by patients that were focused on their personal goals, and alignment of those goals with healthy behaviour changes and action planning to cope with current symptoms. Specific focus points covered throughout the 4 sessions included bipolar disorder and cardiovascular disease risk, stigma issues, wellness habits including diet and exercise within the context of symptom coping strategies and collaborative care management. 3. *Scheduled patient follow-ups:* yes, brief (20-minute) care management contact with patients in LGCC for up to 6 months. These were used to track symptoms and progress towards wellness goals using motivational techniques.

Kilbourne 2012 *(Continued)*

4. *Enhanced interprofessional communication:* yes, a nurse care manager served as a liaison between patients and providers regarding ongoing care and, through regular phone calls, the care manager referred urgent matters to medical and mental health providers and was involved in documenting patient progress over time, and outreach/crisis management after critical service encounters or missed appointments.

Other intervention components:

- Life Goals Collaborative Care (LGCC) is designed to reduce the risk factors for CVD, through improved control of psychiatric symptoms, increased positive health behaviours, as well as improved co-ordination of physical and mental health care.

Description of control:

Enhanced usual care. Patients receive care as usual, in addition to mailings on wellness topics over the 6-month intervention period and referral to primary care services off site.

Outcomes	**Measures taken at:** baseline, 6 months and 12 months **Primary outcomes:** cardiometabolic risk factors; waist circumference, blood pressure, BMI **Able to use:** • Quality of life (Short Form 12 - SF-12) (baseline, 6 months and 12 months) • Manic and depressive symptoms (Internal State Scale - ISS) (baseline, 6 months and 12 months) • Blood pressure (diastolic and systolic BP mmHg) (baseline, 6 months and 12 months) • Body mass index (BMI kg/m^2) (baseline, 6 months and 12 months) • Waist circumference (inches) (baseline, 6 months and 12 months) • Functioning (WHO-DAS) (baseline, 6 months and 12 months) • Attrition (number lost to follow up) (12 months) **Unable to use:** None
Notes	Study conducted in 2009

Kilbourne 2013

Study characteristics	
Methods	**Allocation:** randomised in blocks of 15 to 20 stratified by age, race and diabetes diagnosis to ensure balance of characteristics **Design:** RCT, multicentre **Duration:** 12-month and 24-month follow-up **Date of study:** May 2008 to May 2012 **Country:** USA **Setting:** 2 community-based mental health outpatient programmes in Southeastern Michigan, in a large VA healthcare system providing services to more than 158,000 veterans living in a 15-county area **Recruitment method:** Patients diagnosed with bipolar disorder and a CVD risk factor who received care between fiscal year 2008 and 2009 were identified based on a medical record review of patients.

Kilbourne 2013 *(Continued)*

	Masking: single-blind (outcomes assessor blind to intervention)
Participants	**Inclusion criteria:** adult patients with an active diagnosis or treatment plan for bipolar disorder I, II or NOS with at least one cardiometabolic risk factor (diagnosis or indication of hypertension, hyperlipidaemia, diabetes or BMI > 25) who received care in one of 2 participating community mental health outpatient programmes. Other criteria include community-dwelling and English-speaking. **Exclusion criteria:** severely cognitively impaired or unable to give informed consent **Number randomised to intervention and control:** 134 participants enrolled and 58 randomised to the Life Goals Collaborative Care group (LGCC), 60 to enhanced treatment as usual (ETU) **Number completed study:** 118 **Age:** intervention 53.1 (10.6), control 52.4 (9.2) **Sex:** intervention 10 (17.2%) female, control 10 (16.7%) female **Diagnosis:** bipolar I: intervention 20 (34.5%), control 24 (40%); bipolar II: intervention 14 (24.1%), control 12 (20%); bipolar NOS: intervention 21 (36.2%), control 24 (40%); schizoaffective: intervention 3 (5.2%), control 0 (0%) **Ethnicity:** non-white: intervention 3 (5.2%), control 3 (5.0%) **Any significant differences between intervention and control groups?** no statistically significant baseline demographic or clinical differences between the groups
Interventions	**Type of collaborative care:** A **Description of intervention:** Intervention name: Life Goals Collaborative Care 1. *A multi-professional approach to patient care:* yes, involvement with primary care professionals 2. *A structured management plan:* yes, interventionist provided four 2-hour weekly group self-management sessions, followed by brief care management contacts to patients randomised to LGCC for up to 6 months. Each group session included approximately 8 to 10 participants, and sessions were based on social cognitive theory. The sessions included active discussions by patients that were focused on their personal goals, and alignment of those goals with healthy behaviour changes and action planning to cope with current symptoms. Specific focus points covered throughout the 4 sessions included bipolar disorder and cardiovascular disease risk, stigma issues, wellness habits including diet and exercise within the context of symptom coping strategies, and collaborative care management. 3. *Scheduled patient follow-ups:* yes, brief (20-minute) care management contact with patients in LGCC for up to 6 months. These were used to track symptoms and progress towards wellness goals using motivational techniques. 4. *Enhanced interprofessional communication:* yes, a nurse care manager served as a liaison between patients and providers regarding ongoing care and, through regular phone calls, the care manager referred urgent matters to medical and mental health providers and was involved in documenting patient progress over time, and outreach/crisis management after critical service encounters or missed appointments. Other intervention components: • Brief (20-minute) care management contact with patients in LGCC for up to 6 months. These were used to track symptoms and progress towards wellness goals using motivational techniques. • Improved control of psychiatric symptoms, increased positive health behaviours, as well as improved co-ordination of physical and mental health care • Provider engagement and communication tips • Provider contacts (cues) regarding medication side effects, symptoms or urgent health concerns • Crisis management • Registry tracking

Kilbourne 2013 *(Continued)*

	• Links to community resources • Guideline dissemination to health providers of summary information on BD treatment and health issues (e.g. cardiometabolic risk monitoring) The LGCC intervention arm was implemented by a master's level-trained health specialist. The health specialist's primary roles were to: 1) lead the psychosocial educational group sessions; 2) deliver care management support; and 3) serve as an informational resource to providers by disseminating guidelines and providing information on topics specific to BD treatment and health outcomes. Following randomisation, the health specialist initiated a pre-session assessment to promote treatment engagement and participation. During this time, the health specialist assessed patient preferences for communication, motivation for health changes, availability for group participation, and principal provider contact information for emergency situations. Participants were then scheduled to attend the group self-management sessions. **Description of control:** Enhanced usual care via quarterly newsletters regarding wellness topics mailed to those in the control group. Their general medical and mental health providers received the same practice guideline information at the beginning of the study.
Outcomes	**Measures taken at:** baseline, 6 months, 12 months, 24 months **Primary outcomes:** blood pressure, lipids, functioning, non-fasting blood draw, quality of life **Able to use:** • Quality of life (Short Form 12 - SF-12) (baseline, 6 months, 12 months, 24 months) • Manic and depressive symptoms (Internal State Scale - ISS) (baseline, 6 months, 12 months, 24 months) • Blood pressure (diastolic and systolic BP mmHg) (baseline, 6 months, 12 months, 24 months) • Body mass index (BMI kg/m^2) (baseline, 6 months, 12 months, 24 months) • Total cholesterol, high-density lipoprotein and low-density lipoprotein (total cholesterol, HDL and LDL levels mg/dL) (baseline, 6 months, 12 months, 24 months) • Waist circumference (inches) (baseline, 6 months, 12 months, 24 months) • Deaths (all reasons) • Attrition (number lost to follow-up) (12 months, 24 months) • Functioning (World Health Organization Disability Assessment Scale WHO-DAS) (baseline, 6 months, 12 months, 24 months) • Estimate of heart attack risk (Framingham Risk Score) (baseline, 6 months, 12 months, 24 months) **Unable to use:** • Haemoglobin A1c levels (HbA1C %) – not reported • Estimate of heart attack risk (Framingham Risk Score) – not of interest
Notes	Self-Management Addressing Heart Risk Trial (SMAHRT), a randomised controlled effectiveness trial of an intervention (Life Goals Collaborative Care; LGCC) designed to reduce CVD risk factors and improve physical and mental health outcomes in patients with BD Conducted May 2008 to May 2012

Mishra 2017

Study characteristics	
Methods	*Results are reported in two separate papers and we have assumed these are from the same study but pertaining to two distinct subgroups (people with schizophrenia and bipolar diagnoses).*

Mishra 2017 *(Continued)*

	Allocation: simple randomisation **Design:** prospective RCT, single centre **Duration:** 6 months **Date of study:** none specified **Country:** India **Setting:** outpatient department of psychiatry in a tertiary care hospital **Recruitment method:** Patients who visited the psychiatry outpatient department, of either sex, aged 18 years, treated for schizophrenia and literate were recruited. **Masking:** none described
Participants	**Inclusion criteria:** patients who visited the psychiatry outpatient department, of either sex, over 18 and with a diagnosis of schizophrenia or bipolar and literate **Exclusion criteria:** people with comorbidities seen in other departments **Number randomised to intervention and control:** 101 enrolled **Number completed study:** 96 **Age:** not reported as means for both studies and unclearly reported as age category by sex **Sex:** intervention 25 (26%); control 22 (23%) female **Diagnosis:** diagnosed with schizophrenia (26 enrolled, 23 completed) or bipolar (75 enrolled, 73 completed) **Ethnicity:** not reported **Any significant differences between intervention and control groups?** no statistically significant baseline demographic or clinical differences between the groups *Please note – bipolar paper demographics do not add up correctly.*
Interventions	**Type of collaborative care:** B **Description of intervention:** Intervention name: pharmacist-psychiatrist collaborative care Contains 3 elements of collaborative care: 1. *A multi-professional approach to patient care:* no collaboration with primary care 2. *A structured management plan:* yes, medication review, patient and carer education sessions. These covered awareness of medication prescribed, disease, importance of adherence and how this impacts on quality of life. Motivational techniques used to encourage adherence in patients and carers. 3. *Scheduled patient follow-ups:* yes, 3 scheduled appointments 4. *Enhanced interprofessional communication:* none **Control group:** Usual care, not described
Outcomes	**Measures taken at:** 1 month, 2 months, 3 months – unclear, no baseline reported – assume 1 month means baseline **Primary outcome:** not specified

Mishra 2017 *(Continued)*

	• Medication adherence (Morisky Medication Adherence Scale - MMAS) (1 month, 2 months, 3 months) • Medication adherence (Medication Adherence Rating Scale - MARS) (1 month, 2 months, 3 months) **Unable to use:** • Quality of life (World Health Organization Quality of Life brief instrument WHOQOL-BREF) (1 month, 2 months, 3 months) - total score excluded from analysis - see notes below
Notes	The review team have made the assumption that follow-up 1 is baseline data, and follow-up 3 is at 2 months. We have also pooled the outcome data from the two papers. Pharmacists appear to be providing intervention and completing measures (bias). WHOQOL-BREF: We excluded this outcome as it was unclear how the authors derived the overall score, as they did not simply use the overall quality of life question in the WHOQOL-BREF, and we did not receive any clarification from the authors. As a result, we concluded that this outcome would not be comparable with other reported overall quality of life measures.

Salman 2014

Study characteristics	
Methods	**Allocation:** simple randomised sampling after PANSS assessment **Design:** randomised controlled trial **Duration:** 1 year and 3 months **Date of study:** not specified **Country:** India **Setting:** single site, psychiatry ward of Lady Reading Hospital Peshawar **Recruitment method:** referred by primary care providers immediately after starting antipsychotic medication for schizophrenia **Masking:** double-blind
Participants	**Inclusion criteria:** none described, can assume “diagnosis of schizophrenia and on anti-psychotic medication” **Exclusion criteria:** evidence that the patient had received an antidepressant or antipsychotic, alone or in combination, in the preceding 6 months; comorbid mania or bipolar; psychotic symptoms; eminent suicidality; substance use disorder or dependence **Number randomised to intervention and control:** 96 enrolled, 50 intervention, 46 control **Number completed study:** 80 **Age:** intervention mean 36.9 (SD 10.1), control mean 37.3 (SD 10.2) **Sex:** intervention 54.8%, control 56.1% female **Diagnosis:** schizophrenia: intervention 78.2%, control 69.5%, schizoaffective: intervention 21.7%, control 30.4% **Ethnicity:** not described

Salman 2014 *(Continued)*

	Any significant differences between intervention and control groups? There were no significant differences between the two groups with respect to age, gender, duration of illness, number of hospitalisations and number of months since the last hospitalisation. *Note - demographics numbers do not add up.*
Interventions	**Type of collaborative care:** B **Description of intervention:** Intervention name: collaborative care Contains three elements of collaborative care: 1. *Multi-professional approach to patient care:* no, psychiatrist and pharmacist care managers, liaison with clinical psychologists, does not meet criteria for a collaborative multi-professional approach 2. *A structured management plan:* yes, brief counselling on the prescribed drug, therapeutic endpoints and side effects. Participants were interviewed by care managers immediately after randomisation to assess the severity of psychopathology, identifying potential stressors and other predisposing factors. Past medication, surgical, medical and psychiatric histories were also recorded. Participants were also educated on positive, negative and general symptoms, aetiology and prognosis of schizophrenia. A detailed explanation of the role of antipsychotics was presented, including therapeutic benefits and side effects. Family members were actively engaged in this education. During visits, pharmacists followed standardised set of questions to assess drug adherence, therapeutic effects and outcomes, adverse effects and other social, psychological and medical factors. This enabled them to identify activities participants neglected during their illness and provide encouragement. 3. *Scheduled patient follow-ups:* yes, participants were scheduled for frequent follow-up every 2 weeks, via telephone call and clinic appointments. Clinic visits were scheduled on week 2, 6, 12 and 24 for psychiatric follow-ups where pharmacists would evaluate clinical progress. At week 12 necessity of treatment was determined. 4. *Enhanced interprofessional communication:* yes - clinical pharmacists met with the psychiatrist approximately daily for half an hour, 2 hours each week at least, summarising the presentation of new patients with the psychologist's assistance, as well as providing updates on the clinical progress of other subjects, and discussing it with the head clinical psychologists in the ward. Other intervention components: • Provided with diary cards as a simple medication reminder **Control group:** Enhanced usual care: participants were provided with diary cards as a medication adherence reminder
Outcomes	**Measures taken at:** baseline, 3 and 6 months **Primary outcome:** fails to report which outcomes are considered primary or secondary **Able to use:** • Symptoms (Positive and Negative Symptoms Scale – PANSS) (baseline and 6 months) • Quality of life (Short Form 12 - SF-12) (baseline and 6 months) • Attrition (number lost to follow-up) (6 months) **Unable to use:** • Medication adherence: Morisky Medication Adherence Scale (MMAS-4) (baseline and 6 months) - see notes • Medication Adherence (Medication Adherence Report Scale - MARS) (baseline and 6 months) - see notes

Salman 2014 *(Continued)*

	• Patient satisfaction with pharmacy services: (14-item 5-point Likert scale of statements, unvalidated) – (3 months)
Notes	*Note: Demographics numbers do not add up.* We excluded the MARS and the MMAS medication adherence measures as the reported results were both outside of the possible range of values that could be observed using these measures, and we did not receive any clarification from the authors.

van der Voort 2015

Study characteristics	
Methods	**Allocation:** 2-armed pragmatic cluster-randomised. Cluster-randomisation performed at the level of outpatient teams. Teams that treated at least 20 patients with bipolar disorder were asked to participate. Clusters were matched into pairs by the number of nurses in each team willing to participate in the intervention. These were then randomly assigned to either the experimental or control group by use of an internet generator, performed blind by vdV. No characteristic matching was used due to similarities in quality of care. **Design:** multi-site, cluster-randomised controlled trial **Duration:** 12 months **Date of study:** February 2011 to August 2013 – but unclear date order (2011-02-01 2013-08-01) **Country:** The Netherlands **Setting:** 16 mental health outpatient clinics **Recruitment method:** all patients seen under participating teams were invited to participate **Masking:** patients and professionals could not be blinded due to the nature of the study. However, blinding was performed in the randomisation and statistical analysis, and researchers performing the interview for the Life Chart method were also masked.
Participants	**Inclusion criteria:** diagnosed with bipolar disorder type I, II or NOS, according to DSM-IV-TR. This is assessed through medical records and confirmed by the treating psychiatrist using the Dutch language version of the Questionnaire for Bipolar Illness (QBP-NL), aged 18 to 65 years. **Exclusion criteria:** patients with severe or very severe mania or depression, with a score of 6 or 7 on the Clinical Global Impression - Bipolar Disorder scale; patients with such a stable course of illness (during the last year) that low intensity of treatment suffices (2 to 4 poly clinical visits with a psychiatrist a year); patients without sufficient command of the Dutch language to be able to fill in the questionnaires; inability or unwillingness to give informed consent **Number randomised to intervention and control:** 18 teams were randomised, 9 to intervention, 9 to control; 138 participants were randomised, 56 intervention, 82 control **Number completed study:** 72 people (88%) from both groups completed the 12-month assessment. Two teams had to drop out mid-study, meaning 38 potential participants were unable to participate in the intervention, including 15 people who had consented. 71 patients consented, 56 actually initiated the intervention. 13 discontinued the intervention and 11 were lost to follow up. 45 people in the intervention (80%) completed the 12-month assessment. **Age:** intervention mean 46.8 (9.8), control mean 44.7 (11.3) **Sex:** female: intervention 39 (70%), control 49 (60%) **Diagnosis:** bipolar type 1: intervention 39 (70%), control 49 (60%); bipolar type 2: intervention 11 (20%), control 28 (35%); bipolar NOS: intervention 2 (4%), control 4 (5%)

van der Voort 2015 *(Continued)*

	Ethnicity: not reported **Any significant differences between intervention and control groups?** Significant differences between the following baseline characteristics: patients randomised to CC reported a higher number of months with depressive symptoms during the 6 months prior to baseline than patients in the control group. Patients in CC had higher severity of depressive symptoms in the week preceding baseline. Patients randomised to CC had a lower educational level compared to control. Patients in the CC experienced more functional impairments at baseline than patients in control. Patients in control condition reported at baseline a better quality of life concerning health-related quality of life.
Interventions	**Type of collaborative care:** B **Description of intervention:** Intervention name: the Collaborative Care Programme Contains 3 elements of collaborative care: 1. *A multi-professional approach to patient care:* no collaboration with primary care 2. *A structured management plan:* yes. The patient is an active member of the CC team. One important aim is to agree on the most important problems to be worked on, the related goals and which care is needed to achieve these goals. A contract is made, in which the problems, goals, content of treatment and care, and outcomes are elaborated; monitoring and relapse prevention, by using the Life Chart Method; pharmacotherapy and somatic care, with continuous monitoring of the effects; support for developing a healthy lifestyle. 3. *Scheduled patient follow-ups:* yes, psychoeducation 6 x 2-hour sessions; problem-solving treatment x 6 sessions; pharmacotherapy and somatic care continue as appropriate. 4. *Enhanced interprofessional communication:* Collaborative Care Team consists at least of the patient (and preferably a family member or friend), the nurse and the psychiatrist. The team meets every 3 months. The primary nurse co-ordinates care and is responsible for continuity of care. The patient has an active role in his/her own treatment. If the patient agrees, then family members, friends or caregivers are invited to participate in treatment. Other intervention components: • Psychoeducation (based on the Dutch psychoeducation course, Hofman et al, 1992; Honig et al, 1997) adapted to the needs of patient and family • Problem-solving treatment (Schreuders et al, 2005/2007) • Activity scheduling, if patients have prolonged depression • Rehabilitation modules, if patients have low quality of life and minimal social participation **Control group:** Care as usual in outpatient clinics for bipolar disorder or mood disorders in general
Outcomes	**Measures taken at:** baseline, 6 and 12 months **Primary outcome:** fails to report which outcomes are considered primary or secondary, states “psychosocial functioning, course, prevalence, and severity of psychiatric symptoms and quality of life” **Able to use:** • Functioning (Functioning Assessment Short Test - FAST-NL-P) (baseline, 6 and 12 months) • Depressive symptoms (Quick Inventory for Depressive Symptomology - QIDS) (baseline, 6 and 12 months) • Mania symptoms (Altman Self-Rating Mania Scale - ASRM) (baseline, 6 and 12 months) • Average mood over last month (Life Chart Method - LCM) (baseline, 6 and 12 months) • Quality of Life (World Health Organization Quality of Life Questionnaire WHO-QOL-bref) (baseline, 6 and 12 months)

van der Voort 2015 *(Continued)*	• Attitude towards medication, adherence (Drugs Attitude Inventory - DAI-10) (baseline, 6 and 12 months) • Attrition (number lost to follow-up) (6 months and 12 months) **Unable to use**: • Current characteristics of bipolar disorder (the Questionnaire for Bipolar Illness - QBP-NL) (baseline) – not reported • Current severity of bipolar disorder (Clinical Global Impression for Bipolar Disorder – CGI-BP) (baseline, 6 and 12 months) – not reported • Fidelity; nurses in the experimental group completed fidelity checklists to register collaborative care elements delivered – not of interest • Symptoms (the Brief Symptom Inventory – BSI) (baseline, 6 and 12 months) - not reported • Assessment of needs (CANSAS-P) (baseline, 6 and 12 months) - not reported • Mastery (Sense of Mastery Scale) (baseline, 6 and 12 months) - not reported • Satisfaction with care (visual analogue scale (VAS) and qualitative interview analysed using grounded theory); VAS not reported, qualitative study not used in this study • Costs (direct and indirect) (Treatment Inventory Costs in Psychiatric patients – TIC-P) (baseline, 6 and 12 months) - not reported • Perceived burden of caregivers (Involvement Evaluation Questionnaire – IEQ) (baseline, 6 and 12 months) - not reported • Caregiver satisfaction with care (VAS) (baseline, 6 and 12 months) - not reported
Notes	Sources of monetary support GGZ Ingeest, VU University Medical Center, Dimence, AstraZeneca We "deflated" the sample sizes to account for clustering; n = 94 at 6 months and n = 91 at 12 months.

BD: bipolar disorder; BMI: body mass index; BP: blood pressure; BPD: bipolar disorder; BPRS: Brief Psychiatric Rating Scale; CC: collaborative care; CHW: community healthcare worker; CMHC: community mental health care clinic; CVD: cardiovascular disease; DSM-IV: DSM: Diagnostic and Statistical Manual, version 4; FTE: full-time equivalent; HDL: high-density lipoprotein; ICD-10 DCR: ICD-10 Diagnostic Criteria for Research; ICD-10: International Classification of Diseases, 10th revision; IQR: interquartile range; LDL: low-density lipoprotein; LGCC: Life Goals Collaborative Care; MCS: mental component score; NCC: nurse care co-ordinator; NOS: not otherwise specified; PANSS: Positive and Negative Symptom Scale; PCS: mental component score; RCT: randomised controlled trial; SD: standard deviation; SF 36: Short form 36; VAMC: Veterans Administration Medical Centre; WHO-DAS: World Health Organization Disability Assessment Scale

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

Study	Reason for exclusion
ACTRN12614001312639 2014	Allocation: randomised Intervention: not described as collaborative care Participants: diagnosed with any mental health condition
Ahamad 2019	Allocation: randomised (but does not describe the control group) Intervention: described as collaborative care Participants: diagnosed with schizophrenia, mixture of inpatients and outpatients
Barnes 2007a	Allocation: randomised Intervention: not described as collaborative care

Study	Reason for exclusion
	Participants: diagnosed with bipolar disorder
Barnes 2007b	Allocation: randomised Intervention: not described as collaborative care Participants: diagnosed with bipolar disorder
Barnes 2015	Allocation: randomised Intervention: not described as collaborative care Participants: diagnosed with bipolar disorder
Bauer 2019	Allocation: randomised Intervention: described as collaborative care Participants: diagnosed with a variety of mental health diagnoses, majority depression
Beckelman 2013	Allocation: randomised Intervention: described as collaborative care (type A, multidisciplinary (with PC), nurse follow-ups, guidelines and team meetings) Participants: diagnosed with depression
Bowden 2012	Allocation: randomised Intervention: STEP-BD trial uses collaborative care as the control, rather than the intervention Participants: diagnosed with bipolar
Burns 2015	Allocation: randomised Intervention: not described as collaborative care Participants: diagnosed with psychosis
Byng 2004	Allocation: randomised Intervention: not described as collaborative care Participants: diagnosed with chronic psychosis, bipolar disorder and chronic depression or neurotic conditions
Castillo 2018	Allocation: randomised implementation trial (community engagement and planning vs resources for services) Intervention: described as "expanded collaborative depression care" Participants: self-reported severe depression (PHQ-8 ≥ 20) at baseline or lifetime history of bipolar disorder or psychosis (41%)
D'Souza 2004	Allocation: case-control Intervention: not described as collaborative care Participants: diagnosed with bipolar
Dalcin 2018	Allocation: randomised

Study	Reason for exclusion
	Intervention: not described as collaborative care ("comprehensive CVD risk reduction intervention") Participants: diagnosed with SMI + at least one CVD risk factor
Daumit 2020	Allocation: randomised Intervention: not described as collaborative care Participants: diagnosed with schizophrenia, bipolar or major depressive disorder
Davidson 2005	Allocation: randomised Intervention: not described as collaborative care Participants: diagnosed with persistent depressive symptoms (excluding bipolar disorder and current/past psychosis)
Day 2000	Allocation: unclear Intervention: not described as collaborative care Participants: diagnosed with schizophrenia
Dobscha 2007	Allocation: randomised Intervention: not described as collaborative care (but has similar components) Participants: diagnosed with moderate to severe depression, excluding bipolar disorder or history of psychotic symptoms
Donohue 2012	Allocation: randomised Intervention: described as collaborative care Participants: diagnosed with post-CABG depression
Donohue 2014	Allocation: randomised Intervention: described as collaborative care Participants: diagnosed with depression following CABG surgery, excluding bipolar disorder
Druss 2001	Allocation: randomised Intervention: not described as collaborative care Participants: diagnosed with SMI, veterans enrolled at a VA mental health centre
Druss 2010	Allocation: randomised Intervention: not described as collaborative care Participants: diagnosed with schizophrenia, bipolar disorder, PTSD, depression and other mental illness
Duarte 2015	Allocation: randomised Intervention: described as collaborative care Participants: diagnosed with cancer and co-morbid depression

Study	Reason for exclusion
Dwight-Johnson 2005	Allocation: randomised Intervention: described as collaborative care Participants: diagnosed with co-morbid depression and cancer
Dwinger 2013	Allocation: randomised Intervention: not described as collaborative care Participants: less than half the sample had a serious mental illness and included chronic physical illnesses
Ell 2012	Allocation: randomised Intervention: described as collaborative care Participants: diagnosed with diabetes and depressive symptoms
Ell 2014	Allocation: randomised Intervention: described as collaborative care Participants: diagnosed with coronary heart disease and major depression (people taking lithium or antipsychotics excluded)
Ell 2016	Allocation: randomised Intervention: not described as collaborative care but contextually guided by the chronic care model (CCM) Participants: diagnosed with depression or anxiety
EQUIP	Allocation: randomised Intervention: not described as collaborative care Participants: diagnosed with severe mental illness such as bipolar and psychosis
Ertem 2018	Allocation: randomised Intervention: not described as collaborative care Participants: people with schizophrenia in both community and inpatient settings
Falkum 2010	Allocation: randomised Intervention: not described as collaborative care Participants: diagnosed with schizophrenia spectrum disorders
Fleehart 2015	Allocation: randomised Intervention: not described as collaborative care Participants: diagnosed with COPD and clinically significant depression
Fortney 2015	Allocation: randomised Intervention: not described as collaborative care

Study	Reason for exclusion
	Participants: diagnosed with PTSD
Gensichen 2006	Allocation: randomised Intervention: described as collaborative care Participants: diagnosed with depression (excluding SMI), elderly
Gerritsen 2014	Allocation: longitudinal controlled study Intervention: not described as collaborative care, but some elements look similar Participants: diagnosed with minor or major depression
Goorden 2014	Allocation: randomised Intervention: described as collaborative care Participants: diagnosed with anxiety and panic disorders
Goorden 2015	Allocation: randomised Intervention: described as collaborative care Participants: diagnosed with major depressive disorder (excluding patients with psychotic symptoms); some of their sample developed psychotic symptoms during the study but not formally assessed
Gureje 2017	Allocation: randomised Intervention: described as "collaborative shared care", not collaborative care Participants: diagnosed with psychosis and an inpatient at a complementary traditional health care provider in Nigeria and Ghana
Hidalgo-Mazzei 2015	Allocation: randomised Intervention: not described as collaborative care Participants: diagnosed with bipolar disorder
Hirayasu 2009	Allocation: randomised Intervention: not described as collaborative care Participants: people with history of attempted suicide
Hogarty 1974	Allocation: randomised Intervention: not described as collaborative care Participants: diagnosed with schizophrenia
Huffman 2014	Allocation: randomised Intervention: described as collaborative care Participants: diagnosed with GAD, PD, depression (SMI and psychosis excluded)
Huijbregts 2010	Allocation: unclear Intervention: described as collaborative care

Study	Reason for exclusion
	Participants: diagnosed with depression
Huijbregts 2013	Allocation: randomised Intervention: described as collaborative care Participants: diagnosed with major depressive disorder
Iezzoni 2015	Allocation: randomised Intervention: not described as collaborative care Participants: diagnosed with physical disabilities and SMI
IRCT2015060622580N1 2020	Allocation: randomised Intervention: described as collaborative care Participants: caregivers for people with 'mental disorders'
Johnson 2018	Allocation: randomised Intervention: not described as collaborative care Participants: diagnosed with a variety of mental health conditions
Kastner 2012	Allocation: not randomised Intervention: not described as collaborative care Participants: diagnosed with schizophrenia
Kendrick 2003	Allocation: randomised Intervention: not described as collaborative care Participants: diagnosed with anxiety, depression or reactions to life difficulties
Kendrick 2005	Allocation: randomised Intervention: not described as collaborative care Participants: diagnosed with depression, anxiety or life difficulties
Kershaven 2003	Allocation: not randomised Intervention: not described as collaborative care Participants: diagnosed with early psychosis
Khambaty 2015	Allocation: follow-up of IMPACT study Intervention: described as collaborative care Participants: not diagnosed with SMI
Kikkert 2018	Allocation: randomised Intervention: not described as collaborative care Participants: diagnosed with SMI and substance use disorders

Study	Reason for exclusion
	Note: however, it does appear to meet the definition of collaborative care
Kilbourne 2009	Allocation: randomised but secondary reanalysis of data Intervention: described as collaborative care Participants: diagnosed with bipolar
Kilbourne 2012b	Allocation: not randomised - adaptation of therapy Intervention: described as collaborative care Participants: not service users
Kilbourne 2013a	Allocation: randomised but secondary reanalysis of data Intervention: described as collaborative care Participants: diagnosed with mood disorders
Kilbourne 2014	Allocation: randomised Intervention: described as collaborative care Participants: diagnosed with SMI including depression
Kilbourne 2015	Allocation: randomised but an implementation study Intervention: described as collaborative care Participants: diagnosed with bipolar
Kilbourne 2017	Allocation: randomised Intervention: described as collaborative care Participants: diagnosed with mood disorders
Knight 2008	Allocation: randomised Intervention: described as collaborative care Participants: diagnosed with depression; elderly
Lomax 1992	Allocation: not randomised Intervention: not described as collaborative care Participants: relatives of veterans with schizophrenia
Mcdonough 2009	Allocation: random controls from CPA register Intervention: not described as collaborative care Participants: diagnosed with psychosis
McGurk	Allocation: randomised Intervention: not described as collaborative care Participants: diagnosed with SMI

Study	Reason for exclusion
Menchetti 2013	Allocation: randomised Intervention: described as collaborative care Participants: diagnosed with depression (people with symptoms or history of psychosis excluded)
Meyer 2014	Allocation: randomised Intervention: described as collaborative care Participants: diagnosed with depression (people with symptoms or history of psychosis excluded)
Morone 2010	Allocation: not an RCT, secondary data analysis Intervention: described as collaborative care Participants: diagnosed with depression after coronary bypass surgery
NCT00137280 2005	Allocation: not randomised Intervention: described as collaborative care Participants: diagnosed with schizophrenia, schizophreniform and schizoaffective disorder
NCT00919620	Allocation: randomised Intervention: not described as collaborative care Participants: diagnosed with psychosis
NCT01436331	Allocation: randomised Intervention: not described as collaborative care Participants: diagnosed with psychosis
NCT02440906	Allocation: randomised Intervention: not described as collaborative care Participants: diagnosed with schizophrenia, bipolar and major depression, or anxiety, depression and substance use diagnoses WITH a diagnosis of a physical health condition
NCT02543840	Allocation: implementation trial; stepped-wedge design (waiting list control); facilities were randomised into 3 different start times of 3 sites each Intervention: implementation intervention described as 'implementation support'; intervention described as collaborative chronic care Participants: diagnosed with SMI and other diagnoses
NCT03590041 2020	Allocation: randomised Intervention: not described as collaborative care Participants: diagnosed with diabetes
NCT03881657 2020	Allocation: randomised Intervention: not described as collaborative care (however, based on Wagner's model)

Study	Reason for exclusion
	Participants: diagnosed with a severe persistent mental illness (BPD, schizophrenia, depression or combination of these) + physical health conditions; the majority had depression
NCT04324944 2021	Allocation: randomised Intervention: not described as collaborative care Participants: SMI
NCT04600414 2020	Allocation: randomised Intervention: described as collaborative care Participants: diagnosed with a mental health disorder + opoid use disorder No comparator arm
NCT04601064 2021	Allocation: randomised Intervention: described as collaborative care Participants: screened positive for a mental health disorder or substance use disorder (screening tools for depression and anxiety PHQ-9 and GAD-7)
Nordentoft 2000	Allocation: randomised Intervention: not described as collaborative care Participants: diagnosed with a first episode schizophrenia spectrum disorder
Overend 2014	Allocation: randomised Intervention: described as collaborative care Participants: diagnosed with major depression
Patel 2008	Allocation: randomised Intervention: described as collaborative care Participants: diagnosed with common mental disorders
Patel 2010	Allocation: randomised Intervention: described as collaborative care Participants: diagnosed with common mental disorders
Pereira 2011	Allocation: qualitative study Intervention: described as collaborative care Participants: diagnosed with common mental disorders
Pin 2014	Allocation: randomised Intervention: described as collaborative care Participants: diagnosed with depression; bipolar and psychosis excluded
Price 2004	Allocation: not randomised

Study	Reason for exclusion
	Intervention: described as collaborative care
	Participants: diagnosed with depression, elderly
Putz 2015	Allocation: randomised
	Intervention: not described as collaborative care
	Participants: diagnosed with a variety of mental health disorders
RAISE-ETP	Allocation: randomised
	Intervention: not described as collaborative care
	Participants: diagnosed with psychosis
	Note: does appear to actually meet the definition of collaborative care
Raube 1992	Allocation: randomised
	Intervention: not described as collaborative care
	Participants: elderly people in general
Richards 2016	Allocation: randomised
	Intervention: described as collaborative care
	Participants: diagnosed with depression
Richardson 2014	Allocation: randomised
	Intervention: described as collaborative care
	Participants: diagnosed with depression in adolescence
Rollman 2009	Allocation: randomised
	Intervention: described as collaborative care
	Participants: diagnosed with depression post CABG surgery
Rollman 2018	Allocation: randomised
	Intervention: described as collaborative care
	Participants: diagnosed with anxiety or depression
Sajatovic	Allocation: not randomised
	Intervention: not described as collaborative care
	Participants: diagnosed with bipolar
Sajatovic 2005a	Allocation: not randomised
	Intervention: described as collaborative care
	Participants: diagnosed with bipolar
Sajatovic 2005b	Allocation: not randomised
	Intervention: described as collaborative care

Study	Reason for exclusion
	Participants: diagnosed with bipolar
Sathienluckana 2018	Allocation: randomised Intervention: not described as collaborative care Participants: diagnosed with schizophrenia Note: however, it might meet the definition of collaborative care
Schaefert 2013	Allocation: randomised Intervention: not described as collaborative care Participants: diagnosed with medically unexplained symptoms
Schmidt 1998a	Allocation: randomised Intervention: not described as collaborative care Participants: diagnosed with a variety of mental health conditions; 5% of participants had a diagnosis of psychosis
Shinde 2013	Allocation: not randomised Intervention: described as collaborative care Participants: diagnosed with common mental disorders
Simon 2002	Allocation: randomised Intervention: not described as collaborative care Participants: diagnosed with bipolar
Simon 2006	Allocation: randomised Intervention: not described as collaborative care Participants: diagnosed with bipolar disorder
Smith 2003	Allocation: not randomised Intervention: described as collaborative care Participants: diagnosed with medically unexplained symptoms
Smith 2019	Allocation: randomised Intervention: described as collaborative care Participants: diagnosed with mood disorders, including bipolar, but majority depressive
Sousa 2013	Allocation: randomised Intervention: not described as collaborative care but may fit remit Participants: diagnosed with psychotic disorders, both inpatient and outpatient
Steel	Allocation: randomised Intervention: described as collaborative care

Study	Reason for exclusion
	Participants: diagnosed with advanced cancer
Stewart 2014	Allocation: randomised Intervention: described as collaborative care Participants: diagnosed with depression
Sylvia 2013	Allocation: randomised Intervention: collaborative care is one small element of a complex study evaluating several different types of treatment at once, with participants combining many at once and/or at different stages over a long-term period Participants: diagnosed with bipolar
Sylvia 2015	Allocation: randomised Intervention: not described as collaborative care Participants: diagnosed with bipolar disorder
Tang 2010	Allocation: historical controls Intervention: not described as collaborative care Participants: diagnosed with early intervention psychosis
Van der Feltz 2006	Allocation: randomised cluster trial Intervention: described as collaborative care Participants: diagnosed with medically unexplained symptoms and psychiatric co-morbidity, but not serious mental illness
van Orden 2009	Allocation: randomised cluster trial Intervention: described as collaborative care Participants: diagnosed with a variety of mental disorders (only one participant diagnosed with a psychotic disorder)
Von Korff 1998	Allocation: randomised Intervention: described as collaborative care Participants: diagnosed with depression
Walker 2000	Allocation: randomised Intervention: described as collaborative care Participants: diagnosed with depression
Young 2010	Allocation: not randomised Intervention: not described as collaborative care Participants: diagnosed with schizophrenia or schizoaffective disorder

BPD: bipolar disorder; CABG: coronary artery bypass graft; COPD: chronic obstructive pulmonary disease; CPA: care programme approach; CVD: cardiovascular disease; GAD: generalised anxiety disorder; GAD-7: generalised anxiety disorder scale; PC: primary care; PD: panic disorder; PHQ/PHQ-9: Patient Health Questionnaire-9; PTSD: post-traumatic stress disorder; SMI: severe mental illness; VA: Veterans Administration

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

Aschbrenner 2019

Study name	Randomized controlled trial of a learning collaorative to implement health promotion in mental health
Methods	Setting: 48 mental health provider organisations from across the US Allocation: cluster-randomised implementation trial (sites enrolled in 3 blocks of 16 sites) Masking: single (investigator)
Participants	Diagnosis: primary DSM-V axis I diagnosis of schizophrenia, schizoaffective disorder, bipolar disorder, major depressive disorder or any other state-certified SMI diagnosis (e.g. post-traumatic stress disorder) n = 55 organisations Age: 18+
Interventions	Virtual learning collaborative comprised of an 18-month intensive training, skill building and structured implementation process focused on reinforcing fidelity to the InSHAPE model (with monthly learning sessions) InSHAPE is an evidence-based lifestyle intervention for persons with SMI consisting of a free or low cost gym membership and weekly individual meetings with a certified fitness trainer (i.e. health mentor) who provides instruction on both exercise and healthy eating, and who organises and leads group celebrations. Control: the inSHAPE intervention delivered with technical assistance (comprised of 4 scheduled conference calls and additional calls as needed) Comparison: virtual learning collaborative vs technical assistance
Outcomes	Primary outcomes: • Change in programme participation (proportion of enrolled individuals who received adequate exposure to the evidence-based practice, as defined by attending at least 50% of the InSHAPE sessions over 6 months) • Change in programme fidelity (22-item InSHAPE Fidelity Scale) • Change in participant weight (proportion of InSHAPE participants achieving clinically significant weight loss defined as ≥ 5% weight loss)
Starting date	November 2014
Contact information	kelly.aschbrenner@dartmouth.edu
Notes	This is an implementation trial of a 'virtual learning collaborative'. The inSHAPE intervention is not described as collaborative care. Estimated completion date November 2020

Battersby 2018

Study name	Improving cardiovascular health and quality of life in people with severe mental illness: a randomised trial of a 'partners in health' intervention.
Methods	Setting: southern and western Adelaide community mental health clinics (including recently discharged) Allocation: randomised controlled trial; block randomisation stratified by median age and gender Masking: participants cannot be masked; the statistician and health economist will be blinded when comparing data sets
Participants	Diagnosis: schizophrenia, schizoaffective disorder, bipolar disorder, or depressive psychosis, and at least one CVD risk factor (overweight/obesity, smoking, high blood pressure, blood lipids, glucose or diabetes) n = 358 Age: 30+
Interventions	Flinders programme: comprehensive psychosocial care planning approach, building self-management capacity within a collaborative approach and providing a recovery oriented framework Control: usual care Comparison: Flinders programme + usual care vs usual care
Outcomes	Primary outcomes: • Absolute CVD risk (general CVD risk score) • Health-related quality of life (SF-36 and SF-6D)
Starting date	Registered 31 March 2017
Contact information	malcolm.battersby@flinders.edu.au
Notes	Not currently described as 'collaborative care' by the triallist, but might meet the definition criteria. Email correspondence with PI (10 February 2020) - due to complete at end of 2020. Results available April/May 2021.

Byng 2023

Study name	PARTNERS2: A cluster randomised control trial of a model of collaborative care for people with a diagnosis of bipolar, schizophrenia or other psychoses
Methods	Setting: community, within GP practices Allocation: 1:1 cluster-randomised controlled trial Masking: data collectors, participants and clinicians masked until recruitment complete within each GP practice cluster
Participants	Diagnosis: schizophrenia, bipolar or other psychosis n = 270 Age: 18+

Byng 2023 *(Continued)*

Interventions	A specially trained secondary care mental health worker placed within the GP practice to provide collaborative care based on an individualised goal-setting and recovery model Control: usual care Comparison: PARTNERS2 vs usual care
Outcomes	Primary outcome: quality of life (MANSA V2)
Starting date	2018
Contact information	M.J.Birchwood@warwick.ac.uk; richard.byng@plymouth.ac.uk
Notes	The intervention is described as collaborative care by the trialists.

Fields 2019

Study name	Bridge: proactive psychiatry consultation and case management for patients with cancer
Methods	Setting: general hospital (US) Allocation: randomised controlled trial Masking: single (outcome assessors)
Participants	Diagnosis: schizophrenia spectrum disorder, bipolar disorder, or major depressive disorder with prior psychiatric hospitalisation + invasive breast, lung, gastrointestinal or head and neck cancer (suspected or confirmed stage I-III, or stage IV cancer that can be treated with curative intent according to the judgement of the oncologist) n = 265 Age: 18+
Interventions	Proactive psychiatry consultation (PPC) has 4 key elements- 1. Patient-centred: based on the patient's needs, the team aims to build a relationship, increase engagement and promote continuity. 2. Team-based: a psychiatrist and case manager identify goals for cancer treatment, assess psychiatric history and symptoms with a focus on the impact on cancer care, collaborate with community-based clinicians and caregivers, and address barriers to care. 3. Integrated into cancer care delivery: the psychiatry and oncology teams collaborate starting at cancer diagnosis to support patient through cancer treatment. 4. Systematic: the team monitors psychiatric and cancer-related symptoms and cancer care delivery to measure progress toward goals and rapidly adjust treatment as needed. Control: enhanced usual care (EUC) - a template email is sent to the treating oncologist informing them of the psychiatric diagnosis and available psychosocial services. Patient and caregivers are also informed of available psychosocial services. Comparison: PPC + usual care vs EUC
Outcomes	Primary outcome: • Disruptions in cancer care (the proportion of patients who experience clinically relevant disruptions in cancer care, e.g. delay to cancer diagnosis or treatment, deviation from stage-appropriate cancer treatment or interruption in planned treatment)

Fields 2019 *(Continued)*

Starting date	11 December 2017
Contact information	keirwin@partners.orgpartners.org
Notes	Estimated primary completion date: 15 May 2022; estimated study completion date 15 August 2023 Intervention described as person-centred collaborative care in the Irwin 2019 paper (but not on the trial registry) NCT03360695

Hanlon 2014

Study name	Task sharing for the care of severe mental disorders in a low-income country (TaSCS)
Methods	Setting: rural area in Ethiopia Allocation: randomised, controlled, non-inferiority trial; randomisation stratified by health centre catchment Masking: outcome assessors and investigators masked to allocation status
Participants	Diagnosis: schizophrenia, schizoaffective disorder, bipolar disorder or major depressive disorder n = 324 Age: 25+
Interventions	Task-sharing model of locally delivered mental health care integrated into primary health care (TaSCS). Primary health care based nurses and health officers trained to deliver the World Health Organization (WHO) Mental Health Gap (MhGAP) packages of mental health care supported by community-based health extension workers. Control: psychiatric nurse-led centralised model of outpatient specialist mental health care Comparison: TaSCS vs psychiatric nurse-led centralised model of outpatient specialist mental health care
Outcomes	Primary outcome: • Change in symptom severity (Brief Psychiatric Rating Scale Expanded Version, BPRS-E)
Starting date	March 2015
Contact information	charlotte.hanlon@kcl.ac.uk
Notes	The triallists do not describe this as collaborative care, but it possibly meets the definition. Email correspondence with PI in October 2020. Study completed. Results available by December 2020.

Happell 2018

Study name	Improving the cardio-metabolic health of people with psychosis
Methods	Setting: community-based mental health service in a large metropolitan city in Australia

Happell 2018 *(Continued)*

	Allocation: randomised controlled trial; block randomisation stratified by age and gender Masking: participants and outcome assessors will be masked until after baseline assessment. Team members conducting data analysis will not be involved in data collection. Treatment allocations will be masked until after data analysis.
Participants	Diagnosis: diagnosed with a DSM-V psychotic disorder n = 160 Age: 18 to 65
Interventions	Physical health nurse consultant (PHNC): will co-ordinate physical health care including supported referral to appropriate programmes/services. The PHNC will manage risk using the positive cardiometabolic health treatment framework and will work in collaboration with consumers on self-identified needs, goals and health priorities. Control: usual care Comparison: PHNC + usual care vs usual care
Outcomes	Primary outcomes: • Burden of disease risk factors • Consumer experience ('access', 'acceptability' and 'shared decision-making' dimensions of the Patient Experiences in Primary Healthcare Survey) • Quality of life (Assessment of Quality of Life - AQoL-8D) • Cost-effectiveness (assessing cost-effectiveness prevention methodology)
Starting date	Late 2018
Contact information	brenda.happell@canberra.edu.au
Notes	Not described by the triallists as 'collaborative care'. Emailed PI in October 2020; trial due to end December 2020, but in process of negotiating an extension until December 2021.

Nicole 2018

Study name	Interactive Obesity Treatment Approach (iOTA) for obesity prevention in Serious Mental Illness (iOTA-SMI)
Methods	Setting: not specified (US) Allocation: randomised controlled trial Masking: none
Participants	Diagnosis: a diagnosis of a severe and persistent mental illness n = 30 Age: 18 to 60
Interventions	An interactive obesity treatment approach (iOTA-SMI) Control: health education control receive monthly in-person health coaching visits over 16 weeks, monthly counselling on energy balance, physical activity and nutrition

Nicole 2018 *(Continued)*

	Comparison: iOTA-SMI vs health education control, a 16-week programme. They will receive an assessment of individual behaviour risks, participate in collaborative goal-setting with a health coach and use an interactive text system that will provide ongoing support and self-monitoring of behaviour change goals.
Outcomes	Primary outcome: • Change in body mass index (BMI)
Starting date	July 2018
Contact information	nicolg@wustl.edu
Notes	One aspect of the intervention is described as 'collaborative goal setting' (references other studies of collaborative care) Study due to end June 2020 Author contacted twice in October 2020 - no response

CC: collaborative care; CGI-BP: Clinical Global Impression - Bipolar disorder; CGI-SCH: Clinical Global Impression – Schizophrenia; CVD: cardiovascular disease; DSM-IV/DSM-V: Diagnostic and Statistical Manual, version 4/5; GP: general practitioner; ICD: International Classification of Diseases; SMI: severe mental illness; VAMC: Veterans Administration Medical Centre

RISK OF BIAS

Legend: Low risk of bias | High risk of bias | Some concerns

Risk of bias for analysis 1.1 Quality of life: average change in mental health component - 12 months

	Bias					
Study	**Randomisation process**	**Deviations from intended interventions**	**Missing outcome data**	**Measurement of the outcome**	**Selection of the reported results**	**Overall**
Kilbourne 2012	+	+	+	+	~	~
Kilbourne 2013	+	+	~	+	+	~
van der Voort 2015	−	~	−	+	−	−

Risk of bias for analysis 1.2 Mental state: clinically important change (binary) - 12 months

	Bias					
Study	**Randomisation process**	**Deviations from intended interventions**	**Missing outcome data**	**Measurement of the outcome**	**Selection of the reported results**	**Overall**
Chatterjee 2011	✓	✓	✓	✓	✓	✓

Risk of bias for analysis 1.3 Psychiatric hospital admissions - 12 months

	Bias					
Study	**Randomisation process**	**Deviations from intended interventions**	**Missing outcome data**	**Measurement of the outcome**	**Selection of the reported results**	**Overall**
Subgroup 1.3.1 Number of participants admitted to hospital (up to 12 months)						
Chatterjee 2011	✓	✓	✓	✓	✓	✓

DATA AND ANALYSES

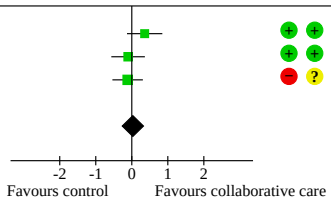
Comparison 1. Collaborative care versus usual care (primary outcomes)

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1.1 Quality of life: average change in mental health component - 12 months	3	227	Std. Mean Difference (IV, Random, 95% CI)	0.03 [-0.26, 0.32]
1.2 Mental state: clinically important change (binary) - 12 months	1	253	Risk Ratio (M-H, Random, 95% CI)	0.99 [0.77, 1.28]
1.3 Psychiatric hospital admissions - 12 months	1	253	Risk Ratio (M-H, Fixed, 95% CI)	5.15 [0.67, 39.57]
1.3.1 Number of participants admitted to hospital (up to 12 months)	1	253	Risk Ratio (M-H, Fixed, 95% CI)	5.15 [0.67, 39.57]

Analysis 1.1. Comparison 1: Collaborative care versus usual care (primary outcomes), Outcome 1: Quality of life: average change in mental health component - 12 months

Study or Subgroup	Collaborative care Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Kilbourne 2012 (1)	33.2	5.5	32	31	6.6	33	29.8%	0.36 [-0.13 , 0.85]
Kilbourne 2013 (1)	32.6	8.3	35	33.4	7.1	36	32.4%	-0.10 [-0.57 , 0.36]
van der Voort 2015 (2)	50.1	17.5	35	52.2	18.4	56	37.8%	-0.12 [-0.54 , 0.31]
Total (95% CI)			**102**			**125**	**100.0%**	**0.03 [-0.26 , 0.32]**

Heterogeneity: Tau² = 0.01; Chi² = 2.48, df = 2 (P = 0.29); I² = 19%
Test for overall effect: Z = 0.20 (P = 0.84)
Test for subgroup differences: Not applicable



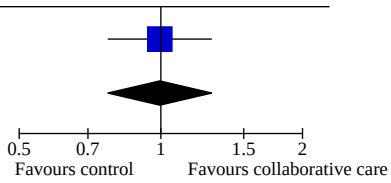
Footnotes
(1) SF-12, 12-month follow-up
(2) WHOQOL-BREF, 12-month follow-up

Risk of bias legend
(A) Bias arising from the randomization process
(B) Bias due to deviations from intended interventions
(C) Bias due to missing outcome data
(D) Bias in measurement of the outcome
(E) Bias in selection of the reported result
(F) Overall bias

Analysis 1.2. Comparison 1: Collaborative care versus usual care (primary outcomes), Outcome 2: Mental state: clinically important change (binary) - 12 months

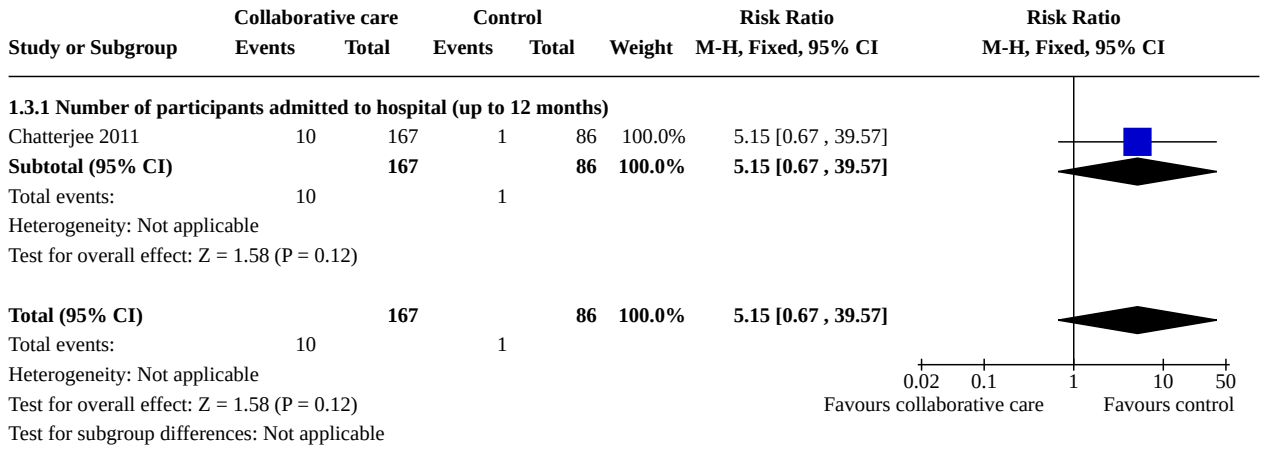
Study or Subgroup	Collaborative care Events	Total	Control Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
Chatterjee 2011 (1)	85	167	44	86	100.0%	0.99 [0.77 , 1.28]
Total (95% CI)		**167**		**86**	**100.0%**	**0.99 [0.77 , 1.28]**
Total events:	85		44			

Heterogeneity: Not applicable
Test for overall effect: Z = 0.04 (P = 0.97)
Test for subgroup differences: Not applicable



Footnotes
(1) Positive And Negative Syndrome Scale (PANSS) score

Analysis 1.3. Comparison 1: Collaborative care versus usual care (primary outcomes), Outcome 3: Psychiatric hospital admissions - 12 months



Comparison 2. Collaborative care versus usual care (secondary outcomes)

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
2.1 Quality of life	6		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
2.1.1 Quality of life: average endpoint in physical health - up to 6 months	5	406	Std. Mean Difference (IV, Random, 95% CI)	0.55 [-0.24, 1.33]
2.1.2 Quality of life: average endpoint in physical health - 12 months	3	237	Std. Mean Difference (IV, Random, 95% CI)	0.08 [-0.18, 0.33]
2.1.3 Quality of life: average endpoint in physical health - more than 12 months	2	381	Std. Mean Difference (IV, Random, 95% CI)	0.02 [-0.19, 0.24]
2.1.4 Quality of life: average endpoint in mental health - up to 6 months	5	406	Std. Mean Difference (IV, Random, 95% CI)	0.71 [-0.17, 1.59]
2.1.5 Quality of life: average endpoint in mental health component (more than 12 months)	2	381	Std. Mean Difference (IV, Random, 95% CI)	0.30 [-0.10, 0.70]
2.1.6 Quality of life: overall endpoint (WHOQOL-BREF) - 6 months	1	94	Std. Mean Difference (IV, Random, 95% CI)	-0.20 [-0.61, 0.22]
2.1.7 Quality of life: overall endpoint (WHOQOL-BREF) - 12 months	1	91	Std. Mean Difference (IV, Random, 95% CI)	0.11 [-0.31, 0.54]
2.2 Mental state	6		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
2.2.1 Mental state (overall general score) up to 6 months	1	29	Std. Mean Difference (IV, Random, 95% CI)	-0.34 [-1.07, 0.40]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
2.2.2 Mental state (general psychopathology) 6 months	1	80	Std. Mean Difference (IV, Random, 95% CI)	-0.11 [-0.55, 0.33]
2.2.3 Mental state (general psychopathology) 12 months	1	253	Std. Mean Difference (IV, Random, 95% CI)	-0.27 [-0.53, -0.01]
2.2.4 Mental state (positive symptoms) 6 months	1	80	Std. Mean Difference (IV, Random, 95% CI)	-0.04 [-0.48, 0.40]
2.2.5 Mental state (positive symptoms) 12 months	1	253	Std. Mean Difference (IV, Random, 95% CI)	-0.17 [-0.43, 0.09]
2.2.6 Mental state (negative symptoms) 6 months	1	80	Std. Mean Difference (IV, Random, 95% CI)	-0.26 [-0.70, 0.18]
2.2.7 Mental state (negative symptoms) 12 months	1	253	Std. Mean Difference (IV, Random, 95% CI)	-0.08 [-0.34, 0.18]
2.2.8 Mental state (depressive symptoms) up to 6 months	4	259	Std. Mean Difference (IV, Random, 95% CI)	-0.13 [-0.53, 0.27]
2.2.9 Mental state (depressive symptoms) 12 months	3	227	Std. Mean Difference (IV, Random, 95% CI)	-0.17 [-0.53, 0.18]
2.2.10 Mental state (depressive symptoms) more than 12 months	1	75	Std. Mean Difference (IV, Random, 95% CI)	-0.19 [-0.64, 0.27]
2.2.11 Mental state (manic symptoms) up to 6 months	3	230	Std. Mean Difference (IV, Random, 95% CI)	-0.14 [-0.40, 0.12]
2.2.12 Mental state (manic symptoms) 12 months	3	227	Std. Mean Difference (IV, Random, 95% CI)	-0.08 [-0.38, 0.22]
2.2.13 Mental state (manic symptoms) more than 12 months	1	75	Std. Mean Difference (IV, Random, 95% CI)	-0.36 [-0.82, 0.10]
2.3 Psychiatric hospital admissions: number of participants admitted to hospital (greater than 12 months)	1		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
2.3.1 Number of participants admitted to hospital (in year 2)	1	306	Risk Ratio (M-H, Fixed, 95% CI)	0.75 [0.57, 0.99]
2.3.2 Number of participants admitted to hospital (in year 3)	1	306	Risk Ratio (M-H, Fixed, 95% CI)	0.73 [0.53, 1.01]
2.4 Other hospital admissions	2		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
2.4.1 Number of participants admitted to hospital (up to 12 months)	1	253	Risk Ratio (M-H, Fixed, 95% CI)	7.77 [0.45, 134.42]
2.4.2 Number of participants admitted to hospital (in year 2)	1	306	Risk Ratio (M-H, Fixed, 95% CI)	0.83 [0.65, 1.04]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
2.4.3 Number of participants admitted to hospital (in year 3)	1	306	Risk Ratio (M-H, Fixed, 95% CI)	0.70 [0.53, 0.93]
2.5 Personal recovery	0	0	Risk Ratio (M-H, Fixed, 95% CI)	Not estimable
2.6 Physical health status	3		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
2.6.1 Blood pressure, mmHg systolic - up to 6 months	3	165	Std. Mean Difference (IV, Random, 95% CI)	-0.15 [-0.54, 0.24]
2.6.2 Blood pressure, mmHg systolic - 12 months	2	136	Std. Mean Difference (IV, Random, 95% CI)	-0.20 [-0.54, 0.13]
2.6.3 Blood pressure, mmHg systolic - 24 months	1	75	Std. Mean Difference (IV, Random, 95% CI)	-0.22 [-0.67, 0.24]
2.6.4 Blood pressure, mmHg diastolic - 6 months	2	136	Std. Mean Difference (IV, Random, 95% CI)	-0.25 [-0.77, 0.27]
2.6.5 Blood pressure, mmHg diastolic - 12 months	2	136	Std. Mean Difference (IV, Random, 95% CI)	-0.29 [-0.70, 0.12]
2.6.6 Blood pressure, mmHg diastolic - 24 months	1	75	Std. Mean Difference (IV, Random, 95% CI)	-0.25 [-0.70, 0.21]
2.6.7 Body mass index (BMI) - 6 months	3	165	Std. Mean Difference (IV, Random, 95% CI)	-0.18 [-0.50, 0.15]
2.6.8 BMI - 12 months	2	136	Std. Mean Difference (IV, Random, 95% CI)	-0.37 [-0.71, -0.03]
2.6.9 BMI - 24 months	1	75	Std. Mean Difference (IV, Random, 95% CI)	-0.35 [-0.81, 0.11]
2.6.10 Total cholesterol - 6 months	1	71	Std. Mean Difference (IV, Random, 95% CI)	-0.43 [-0.90, 0.04]
2.6.11 Total cholesterol - 12 months	1	71	Std. Mean Difference (IV, Random, 95% CI)	-0.19 [-0.65, 0.28]
2.6.12 Total cholesterol - 24 months	1	75	Std. Mean Difference (IV, Random, 95% CI)	0.07 [-0.39, 0.52]
2.6.13 Triglycerides up to 6 months	1	29	Std. Mean Difference (IV, Random, 95% CI)	-0.35 [-1.09, 0.38]
2.6.14 High-density lipoprotein (HDL) - 6 months	1	71	Std. Mean Difference (IV, Random, 95% CI)	-0.06 [-0.52, 0.41]
2.6.15 High-density lipoprotein (HDL) - 12 months	1	71	Std. Mean Difference (IV, Random, 95% CI)	0.10 [-0.36, 0.57]

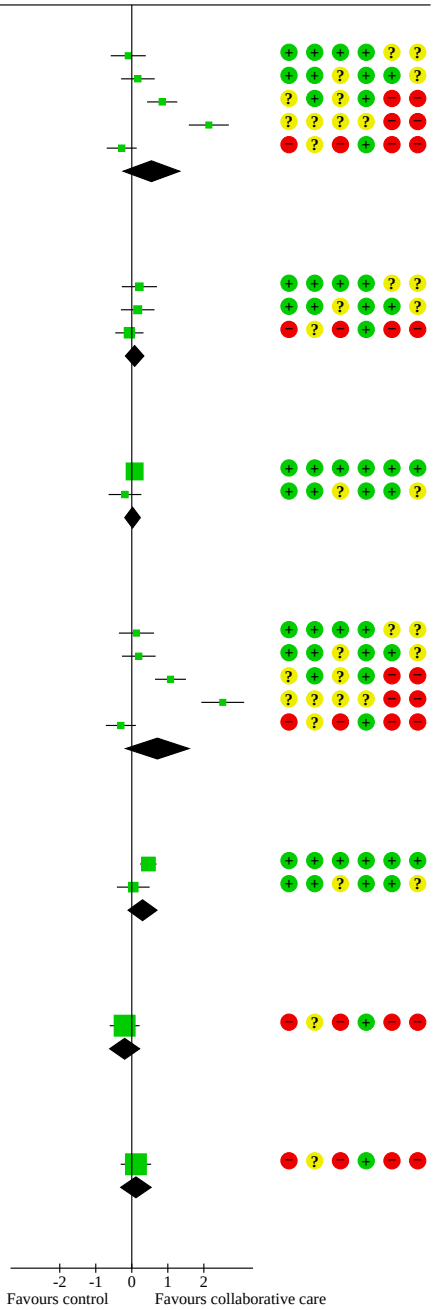
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
2.6.16 High-density lipoprotein (HDL) - 24 months	1	75	Std. Mean Difference (IV, Random, 95% CI)	-0.19 [-0.64, 0.27]
2.6.17 Low-density lipoprotein (LDL) - 6 months	2	100	Std. Mean Difference (IV, Random, 95% CI)	-0.34 [-0.73, 0.06]
2.6.18 Low-density lipoprotein (LDL) - 12 months	1	71	Std. Mean Difference (IV, Random, 95% CI)	-0.12 [-0.59, 0.34]
2.6.19 Low-density lipoprotein (LDL) - 24 months	1	75	Std. Mean Difference (IV, Random, 95% CI)	-0.00 [-0.46, 0.45]
2.6.20 HbA1c up to 6 months	1	29	Std. Mean Difference (IV, Random, 95% CI)	-0.37 [-1.10, 0.37]
2.6.21 Waist circumference - 6 months	2	136	Std. Mean Difference (IV, Random, 95% CI)	-0.31 [-0.98, 0.35]
2.6.22 Waist circumference - 12 months	2	136	Std. Mean Difference (IV, Random, 95% CI)	-0.39 [-0.75, -0.03]
2.6.23 Waist circumference - 24 months	1	75	Std. Mean Difference (IV, Random, 95% CI)	-0.29 [-0.75, 0.17]
2.7 Global state	0	0	Risk Ratio (M-H, Fixed, 95% CI)	Not estimable
2.8 Medication adherence (patient-reported) (DAI-10)	1		Risk Ratio (M-H, Random, 95% CI)	Subtotals only
2.8.1 Medication adherence (patient at 6 months)	1	94	Risk Ratio (M-H, Random, 95% CI)	0.83 [0.67, 1.04]
2.8.2 Medication adherence (patient at 12 months)	1	91	Risk Ratio (M-H, Random, 95% CI)	0.91 [0.75, 1.11]
2.9 Medication adherence (patient-reported) (MARS)	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
2.9.1 Medication adherence (up to 6 months)	1	96	Mean Difference (IV, Fixed, 95% CI)	1.79 [1.56, 2.02]
2.10 Social functioning (binary)	1	253	Risk Ratio (M-H, Random, 95% CI)	1.38 [0.97, 1.95]
2.11 Social functioning/disability	4		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
2.11.1 Social functioning/disability (up to 6 months)	3	230	Std. Mean Difference (IV, Random, 95% CI)	-0.14 [-0.61, 0.32]
2.11.2 Social functioning/disability - 12 months	4	480	Std. Mean Difference (IV, Random, 95% CI)	-0.16 [-0.44, 0.12]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
2.11.3 Social functioning/disability (more than 12 months)	1	75	Std. Mean Difference (IV, Random, 95% CI)	-0.14 [-0.59, 0.32]
2.12 Substance use (alcohol/illicit drugs/cigarettes/tobacco)	0	0	Risk Ratio (M-H, Fixed, 95% CI)	Not estimable
2.13 Adverse effect/event(s)	0	0	Risk Ratio (M-H, Fixed, 95% CI)	Not estimable
2.14 Death	4		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
2.14.1 Number of participants that died from suicide (36 months)	1	330	Risk Ratio (M-H, Fixed, 95% CI)	0.33 [0.01, 8.03]
2.14.2 Number of participants that died from natural causes (36 months)	1	330	Risk Ratio (M-H, Fixed, 95% CI)	1.48 [0.62, 3.53]
2.14.3 Number of participants that died from suicide (12 months)	1	273	Risk Ratio (M-H, Fixed, 95% CI)	0.53 [0.03, 8.30]
2.14.4 Death from natural causes (6 months)	1	35	Risk Ratio (M-H, Fixed, 95% CI)	0.32 [0.01, 7.26]
2.14.5 Any deaths (12 months)	1	118	Risk Ratio (M-H, Fixed, 95% CI)	5.17 [0.25, 105.42]
2.15 Service use outside of mental health (i.e. primary care, emergency services, walk-in centres, social services)	0	0	Risk Ratio (M-H, Fixed, 95% CI)	Not estimable
2.16 Cost of treatment	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
2.16.1 Intervention costs (at 36 months)	1	306	Mean Difference (IV, Random, 95% CI)	-2.98 [-16.93, 10.97]
2.17 Cost of treatment (international dollars)	1		Mean Difference (IV, Fixed, 95% CI)	493.00 [345.41, 640.59]
2.17.1 Total costs at 12 months	1		Mean Difference (IV, Fixed, 95% CI)	493.00 [345.41, 640.59]
2.18 Experience of care/satisfaction	0	0	Risk Ratio (M-H, Fixed, 95% CI)	Not estimable
2.19 Attrition/leaving the study early	6		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
2.19.1 Attrition/leaving the study early (lost to follow-up 6 months)	3	235	Risk Ratio (M-H, Fixed, 95% CI)	1.39 [0.76, 2.55]
2.19.2 Attrition/leaving the study early (lost to follow-up 12 months)	3	504	Risk Ratio (M-H, Fixed, 95% CI)	1.11 [0.77, 1.58]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
2.19.3 Attrition/leaving the study early (lost to follow-up 24 months)	1	118	Risk Ratio (M-H, Fixed, 95% CI)	1.19 [0.74, 1.92]
2.19.4 Attrition/leaving the study early (lost to follow-up at 36 months)	1	330	Risk Ratio (M-H, Fixed, 95% CI)	1.71 [0.77, 3.79]

Analysis 2.1. Comparison 2: Collaborative care versus usual care (secondary outcomes), Outcome 1: Quality of life

Study or Subgroup	Collaborative care Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
2.1.1 Quality of life: average endpoint in physical health - up to 6 months								
Kilbourne 2012 (1)	34.8	7	32	35.5	7.2	33	19.9%	-0.10 [-0.58 , 0.39]
Kilbourne 2013 (1)	35.8	7.8	35	34.5	7.3	36	20.0%	0.17 [-0.30 , 0.64]
Mishra 2017 (2)	56.4	16	51	45.2	8.9	45	20.3%	0.84 [0.43 , 1.26]
Salman 2014 (1)	46.7	5.7	41	35.2	4.9	39	19.5%	2.14 [1.58 , 2.69]
van der Voort 2015 (3)	53.7	17.8	37	58.6	17.1	57	20.3%	-0.28 [-0.70 , 0.14]
Subtotal (95% CI)			**196**			**210**	**100.0%**	**0.55 [-0.24 , 1.33]**
Heterogeneity: Tau² = 0.74; Chi² = 57.07, df = 4 (P < 0.00001); I² = 93%								
Test for overall effect: Z = 1.36 (P = 0.17)								
2.1.2 Quality of life: average endpoint in physical health - 12 months								
Kilbourne 2012 (4)	36	8.8	32	34.3	7.1	33	27.5%	0.21 [-0.28 , 0.70]
Kilbourne 2013 (4)	37.5	7.8	35	36.3	6.6	36	30.1%	0.16 [-0.30 , 0.63]
van der Voort 2015 (5)	56.5	18	45	57.7	17.6	56	42.4%	-0.07 [-0.46 , 0.33]
Subtotal (95% CI)			**112**			**125**	**100.0%**	**0.08 [-0.18 , 0.33]**
Heterogeneity: Tau² = 0.00; Chi² = 0.94, df = 2 (P = 0.63); I² = 0%								
Test for overall effect: Z = 0.60 (P = 0.55)								
2.1.3 Quality of life: average endpoint in physical health - more than 12 months								
Bauer 2006 (6)	43.4	6.21	148	42.9	6.41	158	78.4%	0.08 [-0.15 , 0.30]
Kilbourne 2013 (7)	34	6.6	35	35.3	7	40	21.6%	-0.19 [-0.64 , 0.27]
Subtotal (95% CI)			**183**			**198**	**100.0%**	**0.02 [-0.19 , 0.24]**
Heterogeneity: Tau² = 0.00; Chi² = 1.07, df = 1 (P = 0.30); I² = 7%								
Test for overall effect: Z = 0.19 (P = 0.85)								
2.1.4 Quality of life: average endpoint in mental health - up to 6 months								
Kilbourne 2012 (1)	32.5	7.4	32	31.5	7.9	33	20.0%	0.13 [-0.36 , 0.62]
Kilbourne 2013 (1)	34.4	6.8	35	32.9	8.4	36	20.1%	0.19 [-0.27 , 0.66]
Mishra 2017 (2)	57.4	17	51	42.3	9.3	45	20.2%	1.07 [0.64 , 1.50]
Salman 2014 (1)	68.3	9.1	41	46.3	8.1	39	19.4%	2.53 [1.93 , 3.12]
van der Voort 2015 (3)	48.2	17.3	37	53.6	17.9	57	20.3%	-0.30 [-0.72 , 0.11]
Subtotal (95% CI)			**196**			**210**	**100.0%**	**0.71 [-0.17 , 1.59]**
Heterogeneity: Tau² = 0.94; Chi² = 69.42, df = 4 (P < 0.00001); I² = 94%								
Test for overall effect: Z = 1.58 (P = 0.11)								
2.1.5 Quality of life: average endpoint in mental health component (more than 12 months)								
Bauer 2006 (6)	37.6	7.45	148	34.1	7.7	158	61.4%	0.46 [0.23 , 0.69]
Kilbourne 2013 (7)	34.9	7.5	35	34.6	7.1	40	38.6%	0.04 [-0.41 , 0.49]
Subtotal (95% CI)			**183**			**198**	**100.0%**	**0.30 [-0.10 , 0.70]**
Heterogeneity: Tau² = 0.05; Chi² = 2.63, df = 1 (P = 0.10); I² = 62%								
Test for overall effect: Z = 1.46 (P = 0.14)								
2.1.6 Quality of life: overall endpoint (WHOQOL-BREF) - 6 months								
van der Voort 2015 (3)	3.3	1	37	3.5	1	57	100.0%	-0.20 [-0.61 , 0.22]
Subtotal (95% CI)			**37**			**57**	**100.0%**	**-0.20 [-0.61 , 0.22]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.94 (P = 0.35)								
2.1.7 Quality of life: overall endpoint (WHOQOL-BREF) - 12 months								
van der Voort 2015 (5)	3.4	0.8	35	3.3	0.9	56	100.0%	0.11 [-0.31 , 0.54]
Subtotal (95% CI)			**35**			**56**	**100.0%**	**0.11 [-0.31 , 0.54]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.53 (P = 0.59)								



Footnotes

(1) SF-12, 6-month follow-up
(2) WHOQOL-BREF, 2-month follow-up
(3) WHOQOL-BREF, 6-month follow-up
(4) SF-12, 12-month follow-up
(5) WHOQOL-BREF, 12-month follow-up
(6) SF-36, 36-month follow-up
(7) SF-12, 24-month follow-up

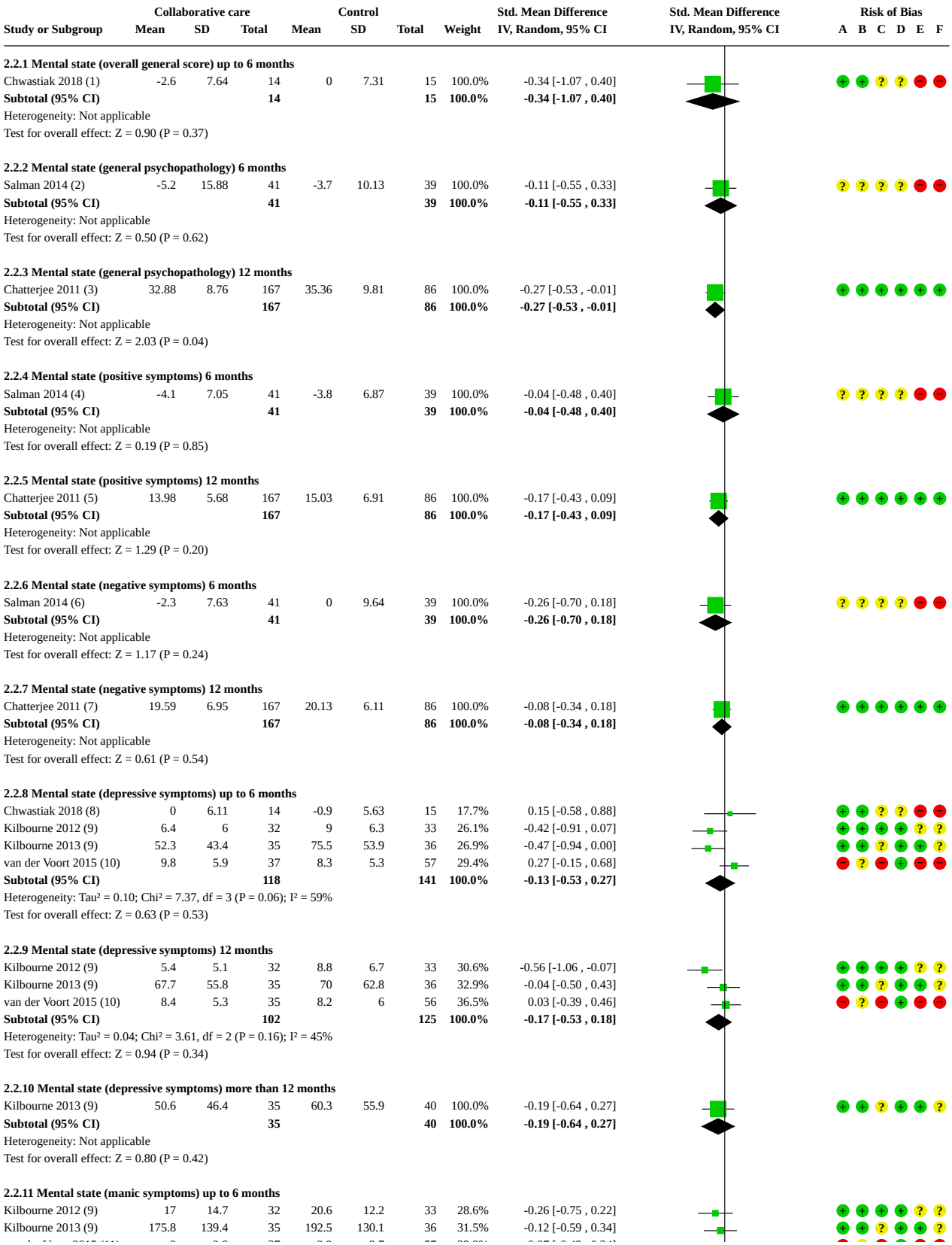
Risk of bias legend

(A) Bias arising from the randomization process
(B) Bias due to deviations from intended interventions
(C) Bias due to missing outcome data

Analysis 2.1. (Continued)

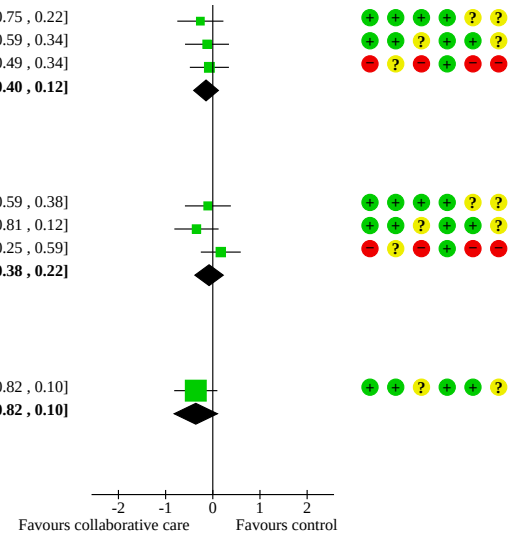
(B) Bias due to deviations from intended interventions
(C) Bias due to missing outcome data
(D) Bias in measurement of the outcome
(E) Bias in selection of the reported result
(F) Overall bias

Analysis 2.2. Comparison 2: Collaborative care versus usual care (secondary outcomes), Outcome 2: Mental state



Analysis 2.2. (Continued)

Kilbourne 2012 (9)	17	14.7	32	20.6	12.2	33	28.6%	-0.26 [-0.75 , 0.22]
Kilbourne 2013 (9)	175.8	139.4	35	192.5	130.1	36	31.5%	-0.12 [-0.59 , 0.34]
van der Voort 2015 (11)	2	2.8	37	2.2	2.7	57	39.9%	-0.07 [-0.49 , 0.34]
Subtotal (95% CI)			**104**			**126**	**100.0%**	**-0.14 [-0.40 , 0.12]**
Heterogeneity: Tau² = 0.00; Chi² = 0.35, df = 2 (P = 0.84); I² = 0%								
Test for overall effect: Z = 1.07 (P = 0.28)								
2.2.12 Mental state (manic symptoms) 12 months								
Kilbourne 2012 (9)	16.6	16	32	18	10.1	33	30.3%	-0.10 [-0.59 , 0.38]
Kilbourne 2013 (9)	153	92	35	192.5	130.1	36	32.1%	-0.35 [-0.81 , 0.12]
van der Voort 2015 (11)	1.9	2.4	35	1.5	2.3	56	37.6%	0.17 [-0.25 , 0.59]
Subtotal (95% CI)			**102**			**125**	**100.0%**	**-0.08 [-0.38 , 0.22]**
Heterogeneity: Tau² = 0.02; Chi² = 2.58, df = 2 (P = 0.28); I² = 22%								
Test for overall effect: Z = 0.51 (P = 0.61)								
2.2.13 Mental state (manic symptoms) more than 12 months								
Kilbourne 2013 (12)	148.9	120.9	35	193.9	125.9	40	100.0%	-0.36 [-0.82 , 0.10]
Subtotal (95% CI)			**35**			**40**	**100.0%**	**-0.36 [-0.82 , 0.10]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.54 (P = 0.12)								



Footnotes

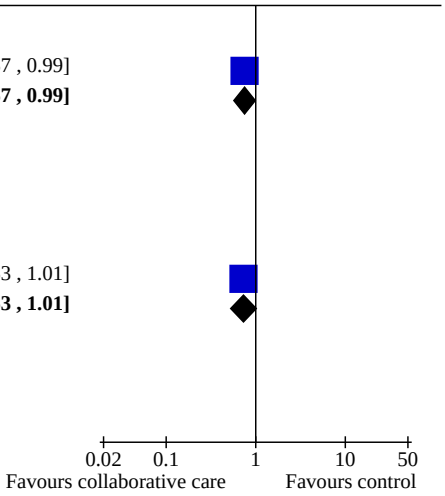
(1) Change from baseline in Brief Psychiatric Rating Scale (BPRS) at 3 months
(2) PANSS general subscale (change from baseline)
(3) PANSS general subscale
(4) PANSS positive subscale (change from baseline)
(5) PANSS positive subscale
(6) PANSS negative subscale (change from baseline)
(7) PANSS negative subscale
(8) PHQ-9 (Patient Health Questionnaire-9)
(9) Internal State Scale (ISS)
(10) Quick Inventory for Depressive Symptomatology (QIDS)
(11) Altman Self-Rating Mania scale
(12) Internal State Scale (ISS), 24 months follow-up

Risk of bias legend

(A) Bias arising from the randomization process
(B) Bias due to deviations from intended interventions
(C) Bias due to missing outcome data
(D) Bias in measurement of the outcome
(E) Bias in selection of the reported result
(F) Overall bias

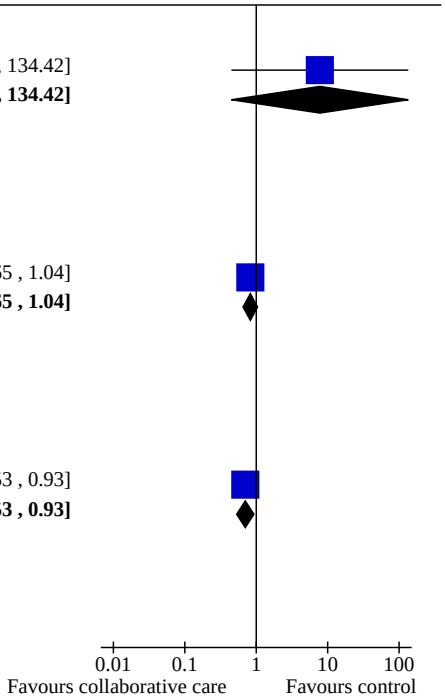
Analysis 2.3. Comparison 2: Collaborative care versus usual care (secondary outcomes), Outcome 3: Psychiatric hospital admissions: number of participants admitted to hospital (greater than 12 months)

Study or Subgroup	Collaborative care Events	Total	Control Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
2.3.1 Number of participants admitted to hospital (in year 2)						
Bauer 2006	52	148	74	158	100.0%	0.75 [0.57 , 0.99]
Subtotal (95% CI)		**148**		**158**	**100.0%**	**0.75 [0.57 , 0.99]**
Total events:	52		74			
Heterogeneity: Not applicable						
Test for overall effect: Z = 2.05 (P = 0.04)						
2.3.2 Number of participants admitted to hospital (in year 3)						
Bauer 2006	41	148	60	158	100.0%	0.73 [0.53 , 1.01]
Subtotal (95% CI)		**148**		**158**	**100.0%**	**0.73 [0.53 , 1.01]**
Total events:	41		60			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.89 (P = 0.06)						
Test for subgroup differences: Chi² = 0.02, df = 1 (P = 0.90), I² = 0%						



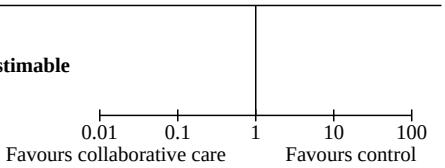
Analysis 2.4. Comparison 2: Collaborative care versus usual care (secondary outcomes), Outcome 4: Other hospital admissions

Study or Subgroup	Collaborative care Events	Total	Control Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
2.4.1 Number of participants admitted to hospital (up to 12 months)						
Chatterjee 2011	7	167	0	86	100.0%	7.77 [0.45 , 134.42]
Subtotal (95% CI)		**167**		**86**	**100.0%**	**7.77 [0.45 , 134.42]**
Total events:	7		0			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.41 (P = 0.16)						
2.4.2 Number of participants admitted to hospital (in year 2)						
Bauer 2006	65	148	84	158	100.0%	0.83 [0.65 , 1.04]
Subtotal (95% CI)		**148**		**158**	**100.0%**	**0.83 [0.65 , 1.04]**
Total events:	65		84			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.60 (P = 0.11)						
2.4.3 Number of participants admitted to hospital (in year 3)						
Bauer 2006	50	148	76	158	100.0%	0.70 [0.53 , 0.93]
Subtotal (95% CI)		**148**		**158**	**100.0%**	**0.70 [0.53 , 0.93]**
Total events:	50		76			
Heterogeneity: Not applicable						
Test for overall effect: Z = 2.49 (P = 0.01)						



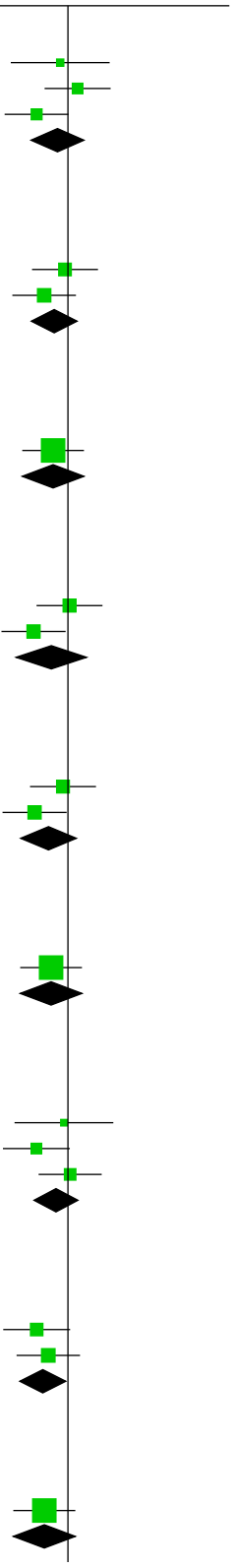
Analysis 2.5. Comparison 2: Collaborative care versus usual care (secondary outcomes), Outcome 5: Personal recovery

Study or Subgroup	Collaborative care Events	Total	Control Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Total (95% CI)		**0**		**0**		**Not estimable**
Total events:	0		0			
Heterogeneity: Not applicable						
Test for overall effect: Not applicable						
Test for subgroup differences: Not applicable						



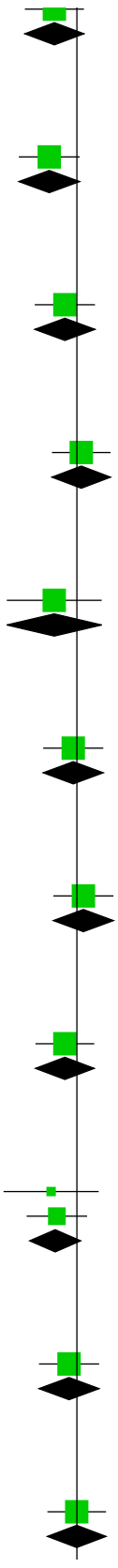
Analysis 2.6. Comparison 2: Collaborative care versus usual care (secondary outcomes), Outcome 6: Physical health status

Study or Subgroup	Collaborative care Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
2.6.1 Blood pressure, mmHg systolic - up to 6 months								
Chwastiak 2018 (1)	-1.1	25.1	14	1.6	20.85	15	22.0%	-0.11 [-0.84 , 0.61]
Kilbourne 2012	134.9	13.1	32	132.8	15.8	33	38.3%	0.14 [-0.34 , 0.63]
Kilbourne 2013	128.3	14	35	135.9	18.2	36	39.7%	-0.46 [-0.93 , 0.01]
Subtotal (95% CI)			**81**			**84**	**100.0%**	**-0.15 [-0.54 , 0.24]**
Heterogeneity: Tau² = 0.04; Chi² = 3.08, df = 2 (P = 0.21); I² = 35%								
Test for overall effect: Z = 0.77 (P = 0.44)								
2.6.2 Blood pressure, mmHg systolic - 12 months								
Kilbourne 2012	134.5	17.5	32	135.3	19	33	48.2%	-0.04 [-0.53 , 0.44]
Kilbourne 2013	127.7	17.7	35	134.2	18.9	36	51.8%	-0.35 [-0.82 , 0.12]
Subtotal (95% CI)			**67**			**69**	**100.0%**	**-0.20 [-0.54 , 0.13]**
Heterogeneity: Tau² = 0.00; Chi² = 0.80, df = 1 (P = 0.37); I² = 0%								
Test for overall effect: Z = 1.18 (P = 0.24)								
2.6.3 Blood pressure, mmHg systolic - 24 months								
Kilbourne 2013	127.2	15.4	35	130.4	13.6	40	100.0%	-0.22 [-0.67 , 0.24]
Subtotal (95% CI)			**35**			**40**	**100.0%**	**-0.22 [-0.67 , 0.24]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.94 (P = 0.35)								
2.6.4 Blood pressure, mmHg diastolic - 6 months								
Kilbourne 2012	84.9	12.3	32	84.6	13.2	33	49.4%	0.02 [-0.46 , 0.51]
Kilbourne 2013	76.3	11.7	35	82.2	11.3	36	50.6%	-0.51 [-0.98 , -0.03]
Subtotal (95% CI)			**67**			**69**	**100.0%**	**-0.25 [-0.77 , 0.27]**
Heterogeneity: Tau² = 0.08; Chi² = 2.35, df = 1 (P = 0.13); I² = 57%								
Test for overall effect: Z = 0.92 (P = 0.36)								
2.6.5 Blood pressure, mmHg diastolic - 12 months								
Kilbourne 2012	83.2	12.7	32	84.1	11.7	33	49.0%	-0.07 [-0.56 , 0.41]
Kilbourne 2013	75.3	10.6	35	80.5	10.3	36	51.0%	-0.49 [-0.96 , -0.02]
Subtotal (95% CI)			**67**			**69**	**100.0%**	**-0.29 [-0.70 , 0.12]**
Heterogeneity: Tau² = 0.03; Chi² = 1.47, df = 1 (P = 0.23); I² = 32%								
Test for overall effect: Z = 1.37 (P = 0.17)								
2.6.6 Blood pressure, mmHg diastolic - 24 months								
Kilbourne 2013	75.9	10.4	35	78.5	10.3	40	100.0%	-0.25 [-0.70 , 0.21]
Subtotal (95% CI)			**35**			**40**	**100.0%**	**-0.25 [-0.70 , 0.21]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.07 (P = 0.28)								
2.6.7 Body mass index (BMI) - 6 months								
Chwastiak 2018 (1)	-1	1.62	14	-0.9	1.66	15	18.7%	-0.06 [-0.79 , 0.67]
Kilbourne 2012	32.8	6.7	32	36.6	9.2	33	38.5%	-0.47 [-0.96 , 0.03]
Kilbourne 2013	33.3	6.9	35	33.1	5.1	36	42.7%	0.03 [-0.43 , 0.50]
Subtotal (95% CI)			**81**			**84**	**100.0%**	**-0.18 [-0.50 , 0.15]**
Heterogeneity: Tau² = 0.01; Chi² = 2.19, df = 2 (P = 0.33); I² = 9%								
Test for overall effect: Z = 1.07 (P = 0.29)								
2.6.8 BMI - 12 months								
Kilbourne 2012	32.6	5.5	32	36.5	10.4	33	47.4%	-0.46 [-0.95 , 0.03]
Kilbourne 2013	32.3	5.9	35	34	5.7	36	52.6%	-0.29 [-0.76 , 0.18]
Subtotal (95% CI)			**67**			**69**	**100.0%**	**-0.37 [-0.71 , -0.03]**
Heterogeneity: Tau² = 0.00; Chi² = 0.24, df = 1 (P = 0.62); I² = 0%								
Test for overall effect: Z = 2.14 (P = 0.03)								
2.6.9 BMI - 24 months								
Kilbourne 2013	31.3	5.8	35	33.4	6.1	40	100.0%	-0.35 [-0.81 , 0.11]
Subtotal (95% CI)			**35**			**40**	**100.0%**	**-0.35 [-0.81 , 0.11]**
Heterogeneity: Not applicable								



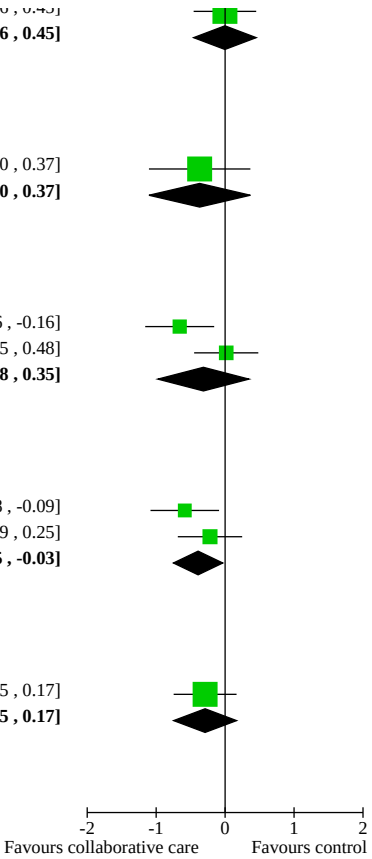
Analysis 2.6. (Continued)

Subtotal (95% CI)			**35**			**40**	**100.0%**	**-0.35 [-0.81 , 0.11]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.49 (P = 0.14)								
2.6.10 Total cholesterol - 6 months								
Kilbourne 2013	173.6	32.6	35	191.5	48.4	36	100.0%	-0.43 [-0.90 , 0.04]
Subtotal (95% CI)			**35**			**36**	**100.0%**	**-0.43 [-0.90 , 0.04]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.78 (P = 0.07)								
2.6.11 Total cholesterol - 12 months								
Kilbourne 2013	173.6	42.4	35	181.8	44.7	36	100.0%	-0.19 [-0.65 , 0.28]
Subtotal (95% CI)			**35**			**36**	**100.0%**	**-0.19 [-0.65 , 0.28]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.78 (P = 0.43)								
2.6.12 Total cholesterol - 24 months								
Kilbourne 2013	178.9	45.5	35	175.9	42.4	40	100.0%	0.07 [-0.39 , 0.52]
Subtotal (95% CI)			**35**			**40**	**100.0%**	**0.07 [-0.39 , 0.52]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.29 (P = 0.77)								
2.6.13 Triglycerides up to 6 months								
Chwastiak 2018 (1)	-33.7	103.37	14	2.9	99.49	15	100.0%	-0.35 [-1.09 , 0.38]
Subtotal (95% CI)			**14**			**15**	**100.0%**	**-0.35 [-1.09 , 0.38]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.94 (P = 0.35)								
2.6.14 High-density lipoprotein (HDL) - 6 months								
Kilbourne 2013 (2)	-36.7	13	35	-36	11.4	36	100.0%	-0.06 [-0.52 , 0.41]
Subtotal (95% CI)			**35**			**36**	**100.0%**	**-0.06 [-0.52 , 0.41]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.24 (P = 0.81)								
2.6.15 High-density lipoprotein (HDL) - 12 months								
Kilbourne 2013 (2)	-37.3	8	35	-38.3	11.3	36	100.0%	0.10 [-0.36 , 0.57]
Subtotal (95% CI)			**35**			**36**	**100.0%**	**0.10 [-0.36 , 0.57]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.42 (P = 0.67)								
2.6.16 High-density lipoprotein (HDL) - 24 months								
Kilbourne 2013 (2)	-39	12.1	35	-37.1	7.9	40	100.0%	-0.19 [-0.64 , 0.27]
Subtotal (95% CI)			**35**			**40**	**100.0%**	**-0.19 [-0.64 , 0.27]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.80 (P = 0.42)								
2.6.17 Low-density lipoprotein (LDL) - 6 months								
Chwastiak 2018 (1)	-19.4	68.44	14	7.9	64.52	15	28.8%	-0.40 [-1.14 , 0.34]
Kilbourne 2013	105.3	28.2	35	116.2	40.1	36	71.2%	-0.31 [-0.78 , 0.16]
Subtotal (95% CI)			**49**			**51**	**100.0%**	**-0.34 [-0.73 , 0.06]**
Heterogeneity: $Tau^2 = 0.00$; $Chi^2 = 0.04$, df = 1 (P = 0.84); $I^2 = 0\%$								
Test for overall effect: Z = 1.67 (P = 0.10)								
2.6.18 Low-density lipoprotein (LDL) - 12 months								
Kilbourne 2013	103.1	32.1	35	107.3	36	36	100.0%	-0.12 [-0.59 , 0.34]
Subtotal (95% CI)			**35**			**36**	**100.0%**	**-0.12 [-0.59 , 0.34]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.51 (P = 0.61)								
2.6.19 Low-density lipoprotein (LDL) - 24 months								
Kilbourne 2013	105.6	39.5	35	105.7	34.2	40	100.0%	-0.00 [-0.46 , 0.45]
Subtotal (95% CI)			**35**			**40**	**100.0%**	**-0.00 [-0.46 , 0.45]**
Heterogeneity: Not applicable								



Analysis 2.6. (Continued)

Kilbourne 2013	105.6	33.3	35	105.7	34.2	40	100.0%	-0.00 [-0.46 , 0.45]
Subtotal (95% CI)			**35**			**40**	**100.0%**	**-0.00 [-0.46 , 0.45]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.01 (P = 0.99)								
2.6.20 HbA1c up to 6 months								
Chwastiak 2018 (1)	-1.1	2.09	14	-0.4	1.58	15	100.0%	-0.37 [-1.10 , 0.37]
Subtotal (95% CI)			**14**			**15**	**100.0%**	**-0.37 [-1.10 , 0.37]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.98 (P = 0.33)								
2.6.21 Waist circumference - 6 months								
Kilbourne 2012	43.9	7.1	32	48.7	7.3	33	49.0%	-0.66 [-1.16 , -0.16]
Kilbourne 2013	43.9	6.4	35	43.8	5.8	36	51.0%	0.02 [-0.45 , 0.48]
Subtotal (95% CI)			**67**			**69**	**100.0%**	**-0.31 [-0.98 , 0.35]**
Heterogeneity: Tau² = 0.17; Chi² = 3.75, df = 1 (P = 0.05); I² = 73%								
Test for overall effect: Z = 0.93 (P = 0.35)								
2.6.22 Waist circumference - 12 months								
Kilbourne 2012	43	6.2	32	46.7	6.3	33	47.1%	-0.58 [-1.08 , -0.09]
Kilbourne 2013	43.4	5.5	35	44.7	6.2	36	52.9%	-0.22 [-0.69 , 0.25]
Subtotal (95% CI)			**67**			**69**	**100.0%**	**-0.39 [-0.75 , -0.03]**
Heterogeneity: Tau² = 0.01; Chi² = 1.10, df = 1 (P = 0.29); I² = 9%								
Test for overall effect: Z = 2.15 (P = 0.03)								
2.6.23 Waist circumference - 24 months								
Kilbourne 2013	43.2	5.3	35	44.9	6.2	40	100.0%	-0.29 [-0.75 , 0.17]
Subtotal (95% CI)			**35**			**40**	**100.0%**	**-0.29 [-0.75 , 0.17]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.25 (P = 0.21)								
Test for subgroup differences: Chi² = 8.73, df = 22 (P = 0.99), I² = 0%								



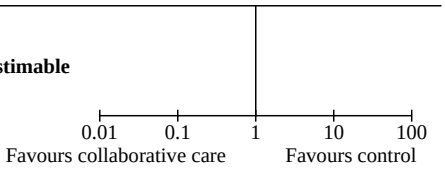
Footnotes

(1) Change from baseline reported

(2) This outcome has been multiplied by -1 in order to maintain the comparability of the direction of effect with the other outcomes in the figure

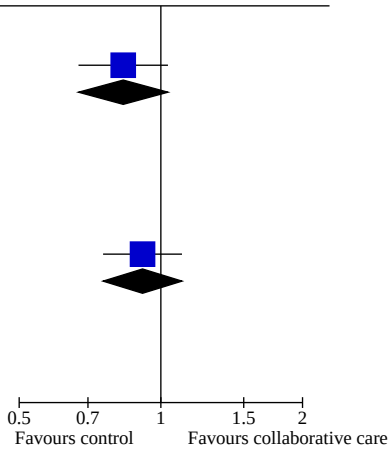
Analysis 2.7. Comparison 2: Collaborative care versus usual care (secondary outcomes), Outcome 7: Global state

Study or Subgroup	Collaborative care Events	Total	Control Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Total (95% CI)		**0**		**0**		**Not estimable**
Total events:	0		0			
Heterogeneity: Not applicable						
Test for overall effect: Not applicable						
Test for subgroup differences: Not applicable						



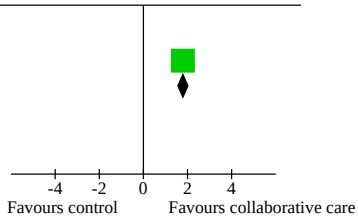
Analysis 2.8. Comparison 2: Collaborative care versus usual care (secondary outcomes), Outcome 8: Medication adherence (patient-reported) (DAI-10)

Study or Subgroup	Collaborative care Events	Total	Control Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
2.8.1 Medication adherence (patient at 6 months)						
van der Voort 2015	27	37	50	57	100.0%	0.83 [0.67 , 1.04]
Subtotal (95% CI)		**37**		**57**	**100.0%**	**0.83 [0.67 , 1.04]**
Total events:	27		50			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.65 (P = 0.10)						
2.8.2 Medication adherence (patient at 12 months)						
van der Voort 2015	28	35	49	56	100.0%	0.91 [0.75 , 1.11]
Subtotal (95% CI)		**35**		**56**	**100.0%**	**0.91 [0.75 , 1.11]**
Total events:	28		49			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.91 (P = 0.36)						
Test for subgroup differences: Chi² = 0.40, df = 1 (P = 0.53), I² = 0%						



Analysis 2.9. Comparison 2: Collaborative care versus usual care (secondary outcomes), Outcome 9: Medication adherence (patient-reported) (MARS)

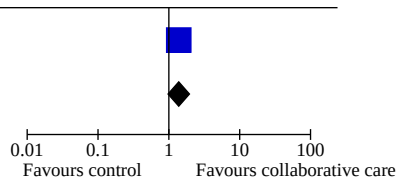
Study or Subgroup	Collaborative care Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
2.9.1 Medication adherence (up to 6 months)								
Mishra 2017 (1)	7.21	0.68	51	5.42	0.45	45	100.0%	1.79 [1.56 , 2.02]
Subtotal (95% CI)			**51**			**45**	**100.0%**	**1.79 [1.56 , 2.02]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 15.37 (P < 0.00001)								
Test for subgroup differences: Not applicable								



Footnotes
(1) 2 months

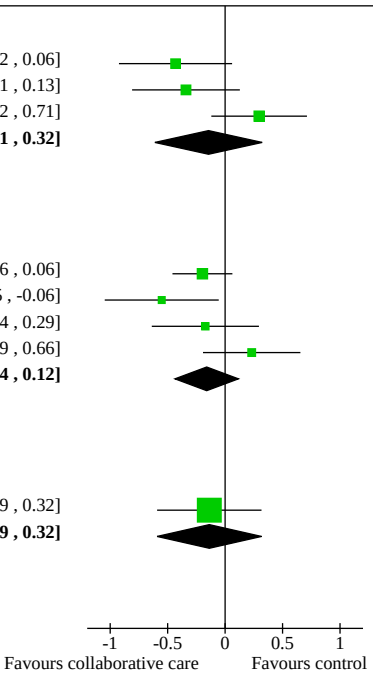
Analysis 2.10. Comparison 2: Collaborative care versus usual care (secondary outcomes), Outcome 10: Social functioning (binary)

Study or Subgroup	Collaborative care Events	Total	Control Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
Chatterjee 2011	75	167	28	86	100.0%	1.38 [0.97 , 1.95]
Total (95% CI)		**167**		**86**	**100.0%**	**1.38 [0.97 , 1.95]**
Total events:	75		28			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.81 (P = 0.07)						
Test for subgroup differences: Not applicable						



Analysis 2.11. Comparison 2: Collaborative care versus usual care (secondary outcomes), Outcome 11: Social functioning/disability

Study or Subgroup	Collaborative care Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
2.11.1 Social functioning/disability (up to 6 months)								
Kilbourne 2012 (1)	16.8	8	32	19.9	6.1	33	31.8%	-0.43 [-0.92 , 0.06]
Kilbourne 2013 (1)	15.9	8	35	19.1	10.4	36	32.9%	-0.34 [-0.81 , 0.13]
van der Voort 2015 (2)	27	16.6	37	22.6	13.4	57	35.4%	0.30 [-0.12 , 0.71]
Subtotal (95% CI)			**104**			**126**	**100.0%**	**-0.14 [-0.61 , 0.32]**
Heterogeneity: Tau² = 0.12; Chi² = 6.23, df = 2 (P = 0.04); I² = 68%								
Test for overall effect: Z = 0.61 (P = 0.54)								
2.11.2 Social functioning/disability - 12 months								
Chatterjee 2011 (3)	5.68	3.54	167	6.4	3.82	86	35.9%	-0.20 [-0.46 , 0.06]
Kilbourne 2012 (1)	15.7	11.8	32	21.2	7.5	33	19.5%	-0.55 [-1.05 , -0.06]
Kilbourne 2013 (1)	15.4	8.9	35	17	9.5	36	21.1%	-0.17 [-0.64 , 0.29]
van der Voort 2015 (2)	28.4	15.3	35	24.8	15.5	56	23.5%	0.23 [-0.19 , 0.66]
Subtotal (95% CI)			**269**			**211**	**100.0%**	**-0.16 [-0.44 , 0.12]**
Heterogeneity: Tau² = 0.04; Chi² = 5.75, df = 3 (P = 0.12); I² = 48%								
Test for overall effect: Z = 1.14 (P = 0.26)								
2.11.3 Social functioning/disability (more than 12 months)								
Kilbourne 2013 (1)	15	10.9	35	16.5	10.7	40	100.0%	-0.14 [-0.59 , 0.32]
Subtotal (95% CI)			**35**			**40**	**100.0%**	**-0.14 [-0.59 , 0.32]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.59 (P = 0.55)								

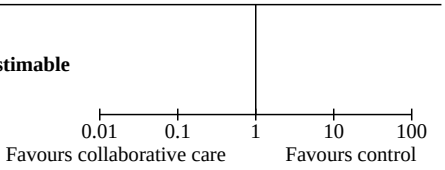


Footnotes

(1) WHO disability assessment scale
(2) Functioning Assessment Short Test (FAST)
(3) Indian Disability Evaluation and Assessment Scale

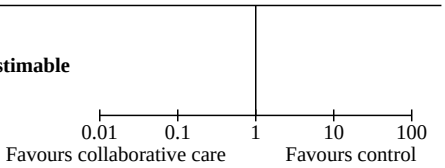
Analysis 2.12. Comparison 2: Collaborative care versus usual care (secondary outcomes), Outcome 12: Substance use (alcohol/illicit drugs/cigarettes/tobacco)

Study or Subgroup	Collaborative care Events	Total	Control Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Total (95% CI)		**0**		**0**		**Not estimable**
Total events:	0		0			
Heterogeneity: Not applicable						
Test for overall effect: Not applicable						
Test for subgroup differences: Not applicable						



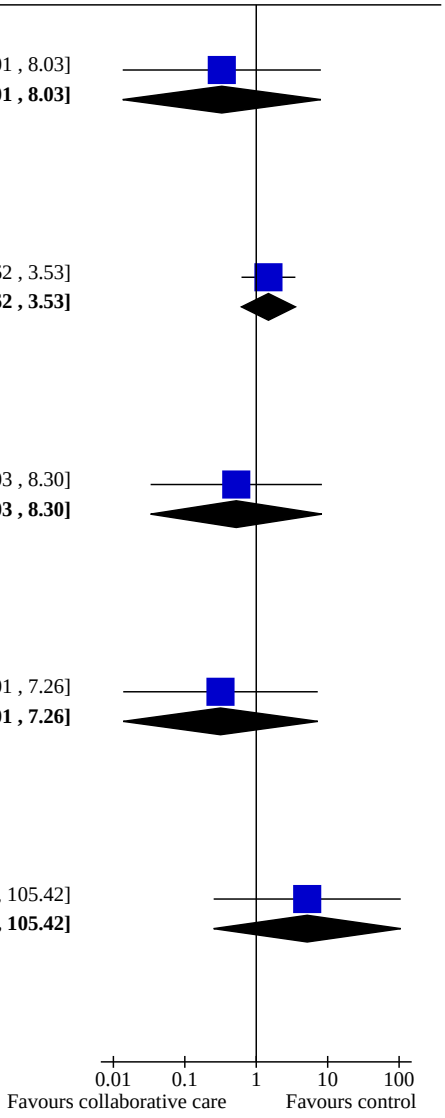
Analysis 2.13. Comparison 2: Collaborative care versus usual care (secondary outcomes), Outcome 13: Adverse effect/event(s)

Study or Subgroup	Collaborative care Events	Total	Control Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Total (95% CI)		**0**		**0**		**Not estimable**
Total events:	0		0			
Heterogeneity: Not applicable						
Test for overall effect: Not applicable						
Test for subgroup differences: Not applicable						



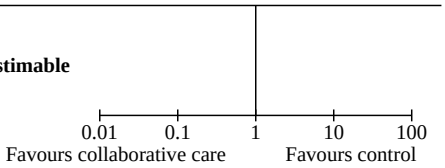
Analysis 2.14. Comparison 2: Collaborative care versus usual care (secondary outcomes), Outcome 14: Death

Study or Subgroup	Collaborative care Events	Total	Control Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
2.14.1 Number of participants that died from suicide (36 months)						
Bauer 2006	0	166	1	164	100.0%	0.33 [0.01 , 8.03]
Subtotal (95% CI)		**166**		**164**	**100.0%**	**0.33 [0.01 , 8.03]**
Total events:	0		1			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.68 (P = 0.50)						
2.14.2 Number of participants that died from natural causes (36 months)						
Bauer 2006	12	166	8	164	100.0%	1.48 [0.62 , 3.53]
Subtotal (95% CI)		**166**		**164**	**100.0%**	**1.48 [0.62 , 3.53]**
Total events:	12		8			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.89 (P = 0.37)						
2.14.3 Number of participants that died from suicide (12 months)						
Chatterjee 2011	1	179	1	94	100.0%	0.53 [0.03 , 8.30]
Subtotal (95% CI)		**179**		**94**	**100.0%**	**0.53 [0.03 , 8.30]**
Total events:	1		1			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.46 (P = 0.65)						
2.14.4 Death from natural causes (6 months)						
Chwastiak 2018	0	18	1	17	100.0%	0.32 [0.01 , 7.26]
Subtotal (95% CI)		**18**		**17**	**100.0%**	**0.32 [0.01 , 7.26]**
Total events:	0		1			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.72 (P = 0.47)						
2.14.5 Any deaths (12 months)						
Kilbourne 2013	2	58	0	60	100.0%	5.17 [0.25 , 105.42]
Subtotal (95% CI)		**58**		**60**	**100.0%**	**5.17 [0.25 , 105.42]**
Total events:	2		0			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.07 (P = 0.29)						
Test for subgroup differences: Chi² = 2.79, df = 4 (P = 0.59), I² = 0%						



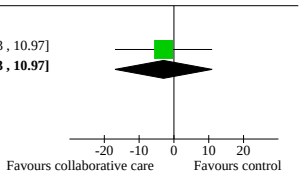
Analysis 2.15. Comparison 2: Collaborative care versus usual care (secondary outcomes), Outcome 15: Service use outside of mental health (i.e. primary care, emergency services, walk-in centres, social services)

Study or Subgroup	Collaborative care Events	Collaborative care Total	Control Events	Control Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Total (95% CI)		**0**		**0**		**Not estimable**
Total events:	0		0			
Heterogeneity: Not applicable						
Test for overall effect: Not applicable						
Test for subgroup differences: Not applicable						



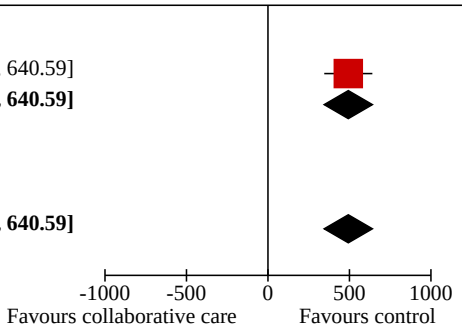
Analysis 2.16. Comparison 2: Collaborative care versus usual care (secondary outcomes), Outcome 16: Cost of treatment

Study or Subgroup	Collaborative care Mean [USD 1000]	Collaborative care SD [USD 1000]	Collaborative care Total	Control Mean [USD 1000]	Control SD [USD 1000]	Control Total	Weight	Mean Difference IV, Random, 95% CI [USD 1000]
2.16.1 Intervention costs (at 36 months)								
Bauer 2006	61.398	64.483	148	64.379	59.745	158	100.0%	-2.98 [-16.93 , 10.97]
Subtotal (95% CI)			**148**			**158**	**100.0%**	**-2.98 [-16.93 , 10.97]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.42 (P = 0.68)								



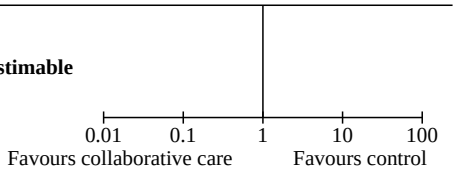
Analysis 2.17. Comparison 2: Collaborative care versus usual care (secondary outcomes), Outcome 17: Cost of treatment (international dollars)

Study or Subgroup	Mean Difference	SE	Weight	Mean Difference IV, Fixed, 95% CI
2.17.1 Total costs at 12 months				
Chatterjee 2011	493	75.3	100.0%	493.00 [345.41 , 640.59]
Subtotal (95% CI)			**100.0%**	**493.00 [345.41 , 640.59]**
Heterogeneity: Not applicable				
Test for overall effect: Z = 6.55 (P < 0.00001)				
Total (95% CI)			**100.0%**	**493.00 [345.41 , 640.59]**
Heterogeneity: Not applicable				
Test for overall effect: Z = 6.55 (P < 0.00001)				
Test for subgroup differences: Not applicable				



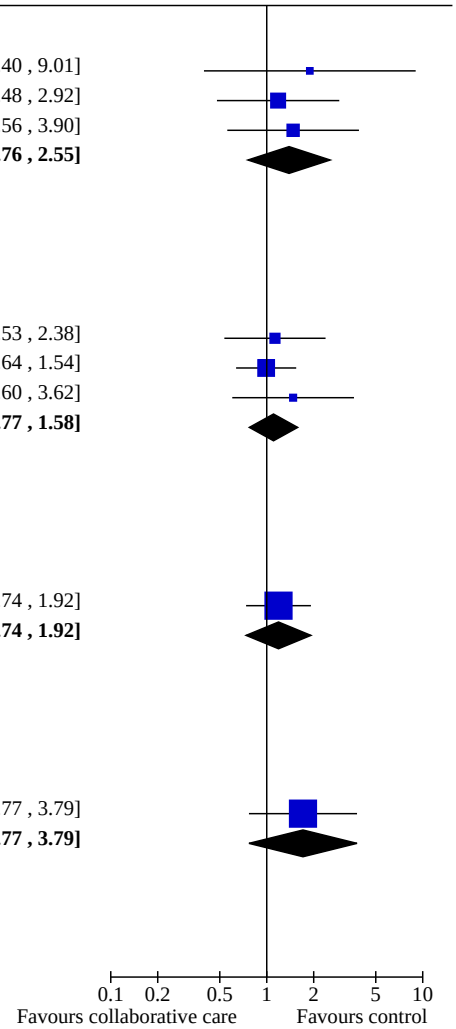
Analysis 2.18. Comparison 2: Collaborative care versus usual care (secondary outcomes), Outcome 18: Experience of care/satisfaction

Study or Subgroup	Experimental Events	Experimental Total	Control Events	Control Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Total (95% CI)		**0**		**0**		**Not estimable**
Total events:	0		0			
Heterogeneity: Not applicable						
Test for overall effect: Not applicable						
Test for subgroup differences: Not applicable						



Analysis 2.19. Comparison 2: Collaborative care versus usual care (secondary outcomes), Outcome 19: Attrition/leaving the study early

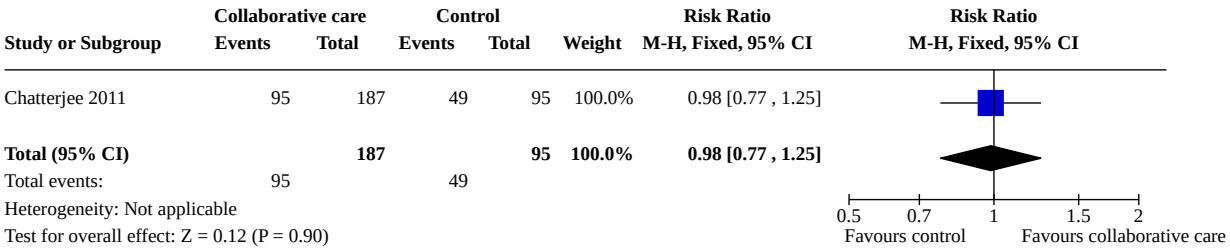
Study or Subgroup	Collaborative care Events	Collaborative care Total	Control Events	Control Total	Weight	Risk Ratio M-H, Fixed, 95% CI
2.19.1 Attrition/leaving the study early (lost to follow-up 6 months)						
Chwastiak 2018	4	18	2	17	13.7%	1.89 [0.40 , 9.01]
Salman 2014	9	50	7	46	48.6%	1.18 [0.48 , 2.92]
van der Voort 2015	7	42	7	62	37.7%	1.48 [0.56 , 3.90]
Subtotal (95% CI)		**110**		**125**	**100.0%**	**1.39 [0.76 , 2.55]**
Total events:	20		16			
Heterogeneity: Chi² = 0.29, df = 2 (P = 0.87); I² = 0%						
Test for overall effect: Z = 1.06 (P = 0.29)						
2.19.2 Attrition/leaving the study early (lost to follow-up 12 months)						
Chatterjee 2011	20	187	9	95	28.4%	1.13 [0.53 , 2.38]
Kilbourne 2013	23	58	24	60	56.2%	0.99 [0.64 , 1.54]
van der Voort 2015	8	42	8	62	15.4%	1.48 [0.60 , 3.62]
Subtotal (95% CI)		**287**		**217**	**100.0%**	**1.11 [0.77 , 1.58]**
Total events:	51		41			
Heterogeneity: Chi² = 0.63, df = 2 (P = 0.73); I² = 0%						
Test for overall effect: Z = 0.55 (P = 0.58)						
2.19.3 Attrition/leaving the study early (lost to follow-up 24 months)						
Kilbourne 2013	23	58	20	60	100.0%	1.19 [0.74 , 1.92]
Subtotal (95% CI)		**58**		**60**	**100.0%**	**1.19 [0.74 , 1.92]**
Total events:	23		20			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.71 (P = 0.48)						
2.19.4 Attrition/leaving the study early (lost to follow-up at 36 months)						
Bauer 2006	15	163	9	167	100.0%	1.71 [0.77 , 3.79]
Subtotal (95% CI)		**163**		**167**	**100.0%**	**1.71 [0.77 , 3.79]**
Total events:	15		9			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.31 (P = 0.19)						
Test for subgroup differences: Chi² = 1.17, df = 3 (P = 0.76), I² = 0%						



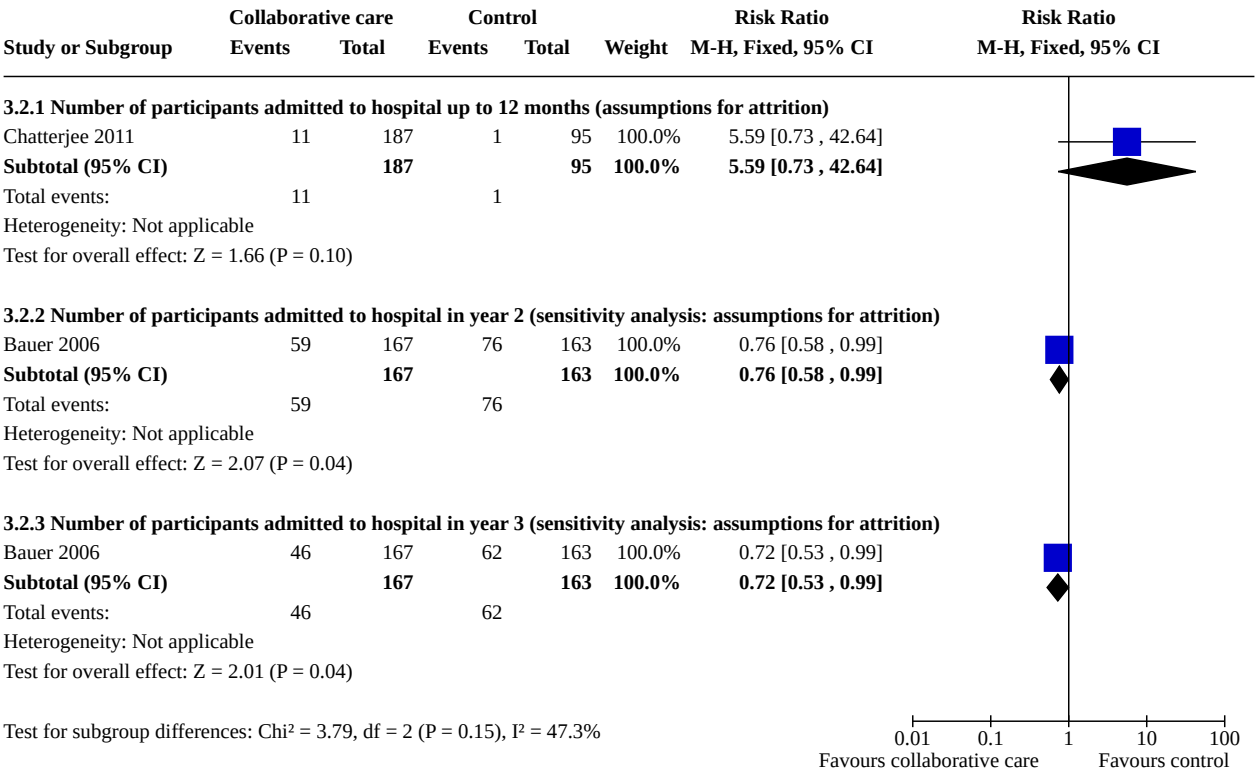
Comparison 3. Collaborative care versus usual care (sensitivity analyses)

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
3.1 Mental state: clinically important change (sensitivity analysis: assumptions for attrition)	1	282	Risk Ratio (M-H, Fixed, 95% CI)	0.98 [0.77, 1.25]
3.2 Psychiatric hospital admissions (sensitivity analysis: assumptions for attrition)	2		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
3.2.1 Number of participants admitted to hospital up to 12 months (assumptions for attrition)	1	282	Risk Ratio (M-H, Fixed, 95% CI)	5.59 [0.73, 42.64]
3.2.2 Number of participants admitted to hospital in year 2 (sensitivity analysis: assumptions for attrition)	1	330	Risk Ratio (M-H, Fixed, 95% CI)	0.76 [0.58, 0.99]
3.2.3 Number of participants admitted to hospital in year 3 (sensitivity analysis: assumptions for attrition)	1	330	Risk Ratio (M-H, Fixed, 95% CI)	0.72 [0.53, 0.99]

Analysis 3.1. Comparison 3: Collaborative care versus usual care (sensitivity analyses), Outcome 1: Mental state: clinically important change (sensitivity analysis: assumptions for attrition)



Analysis 3.2. Comparison 3: Collaborative care versus usual care (sensitivity analyses), Outcome 2: Psychiatric hospital admissions (sensitivity analysis: assumptions for attrition)



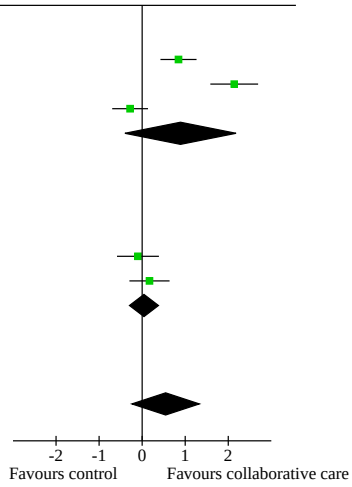
Comparison 4. Collaborative care versus usual care (subgroup analyses)

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
4.1 Quality of life, physical health at 6 months - subgroup analysis: quality of study	5	406	Std. Mean Difference (IV, Random, 95% CI)	0.55 [-0.24, 1.33]
4.1.1 Physical health at 6 months (lower-quality studies)	3	270	Std. Mean Difference (IV, Random, 95% CI)	0.89 [-0.40, 2.18]
4.1.2 Physical health at 6 months (higher-quality studies)	2	136	Std. Mean Difference (IV, Random, 95% CI)	0.04 [-0.29, 0.38]
4.2 Quality of life, mental health at 6 months - subgroup analysis: quality of study	5	406	Std. Mean Difference (IV, Random, 95% CI)	0.71 [-0.17, 1.59]
4.2.1 Mental health at 6 months (lower-quality studies)	3	270	Std. Mean Difference (IV, Random, 95% CI)	1.09 [-0.42, 2.59]
4.2.2 Mental health at 6 months (higher-quality studies)	2	136	Std. Mean Difference (IV, Random, 95% CI)	0.16 [-0.17, 0.50]
4.3 Quality of life, physical health at 6 months - subgroup analysis: variations in	5	406	Std. Mean Difference (IV, Random, 95% CI)	0.55 [-0.24, 1.33]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
implementation of the collaborative care intervention and healthcare systems				
4.3.1 Physical health at 6 months (pharmacy collaborative care)	2	176	Std. Mean Difference (IV, Random, 95% CI)	1.48 [0.21, 2.75]
4.3.2 Physical health at 6 months (no pharmacy collaborative care)	3	230	Std. Mean Difference (IV, Random, 95% CI)	-0.09 [-0.35, 0.18]
4.4 Quality of life, mental health at 6 months - subgroup analysis: variations in implementation of the collaborative care intervention and healthcare systems	5	406	Std. Mean Difference (IV, Random, 95% CI)	0.71 [-0.17, 1.59]
4.4.1 Mental health at 6 months (pharmacy collaborative care)	2	176	Std. Mean Difference (IV, Random, 95% CI)	1.79 [0.36, 3.21]
4.4.2 Mental health at 6 months (no pharmacy collaborative care)	3	230	Std. Mean Difference (IV, Random, 95% CI)	-0.01 [-0.33, 0.31]

Analysis 4.1. Comparison 4: Collaborative care versus usual care (subgroup analyses), Outcome 1: Quality of life, physical health at 6 months - subgroup analysis: quality of study

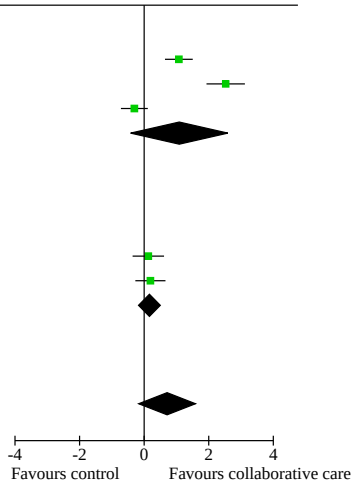
Study or Subgroup	Collaborative care Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
4.1.1 Physical health at 6 months (lower-quality studies)								
Mishra 2017 (1)	56.4	16	51	45.2	8.9	45	20.3%	0.84 [0.43 , 1.26]
Salman 2014 (2)	46.7	5.7	41	35.2	4.9	39	19.5%	2.14 [1.58 , 2.69]
van der Voort 2015 (3)	53.7	17.8	37	58.6	17.1	57	20.3%	-0.28 [-0.70 , 0.14]
Subtotal (95% CI)			**129**			**141**	**60.1%**	**0.89 [-0.40 , 2.18]**
Heterogeneity: Tau² = 1.24; Chi² = 47.61, df = 2 (P < 0.00001); I² = 96%								
Test for overall effect: Z = 1.35 (P = 0.18)								
4.1.2 Physical health at 6 months (higher-quality studies)								
Kilbourne 2012 (4)	34.8	7	32	35.5	7.2	33	19.9%	-0.10 [-0.58 , 0.39]
Kilbourne 2013 (2)	35.8	7.8	35	34.5	7.3	36	20.0%	0.17 [-0.30 , 0.64]
Subtotal (95% CI)			**67**			**69**	**39.9%**	**0.04 [-0.29 , 0.38]**
Heterogeneity: Tau² = 0.00; Chi² = 0.61, df = 1 (P = 0.44); I² = 0%								
Test for overall effect: Z = 0.25 (P = 0.81)								
Total (95% CI)			**196**			**210**	**100.0%**	**0.55 [-0.24 , 1.33]**
Heterogeneity: Tau² = 0.74; Chi² = 57.07, df = 4 (P < 0.00001); I² = 93%								
Test for overall effect: Z = 1.36 (P = 0.17)								
Test for subgroup differences: Chi² = 1.55, df = 1 (P = 0.21), I² = 35.5%								



Footnotes
(1) WHOQOL-BREF, 2 month follow-up
(2) SF-12, 6 month follow-up
(3) WHOQOL-BREF, 6 month follow-up
(4) SF-12, 6-month follow-up

Analysis 4.2. Comparison 4: Collaborative care versus usual care (subgroup analyses), Outcome 2: Quality of life, mental health at 6 months - subgroup analysis: quality of study

Study or Subgroup	Collaborative care Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
4.2.1 Mental health at 6 months (lower-quality studies)								
Mishra 2017 (1)	57.4	17	51	42.3	9.3	45	20.2%	1.07 [0.64 , 1.50]
Salman 2014 (2)	68.3	9.1	41	46.3	8.1	39	19.4%	2.53 [1.93 , 3.12]
van der Voort 2015 (3)	48.2	17.3	37	53.6	17.9	57	20.3%	-0.30 [-0.72 , 0.11]
Subtotal (95% CI)			**129**			**141**	**59.9%**	**1.09 [-0.42 , 2.59]**
Heterogeneity: Tau² = 1.71; Chi² = 60.94, df = 2 (P < 0.00001); I² = 97%								
Test for overall effect: Z = 1.41 (P = 0.16)								
4.2.2 Mental health at 6 months (higher-quality studies)								
Kilbourne 2012 (2)	32.5	7.4	32	31.5	7.9	33	20.0%	0.13 [-0.36 , 0.62]
Kilbourne 2013 (2)	34.4	6.8	35	32.9	8.4	36	20.1%	0.19 [-0.27 , 0.66]
Subtotal (95% CI)			**67**			**69**	**40.1%**	**0.16 [-0.17 , 0.50]**
Heterogeneity: Tau² = 0.00; Chi² = 0.04, df = 1 (P = 0.85); I² = 0%								
Test for overall effect: Z = 0.95 (P = 0.34)								
Total (95% CI)			**196**			**210**	**100.0%**	**0.71 [-0.17 , 1.59]**
Heterogeneity: Tau² = 0.94; Chi² = 69.42, df = 4 (P < 0.00001); I² = 94%								
Test for overall effect: Z = 1.58 (P = 0.11)								
Test for subgroup differences: Chi² = 1.38, df = 1 (P = 0.24), I² = 27.4%								

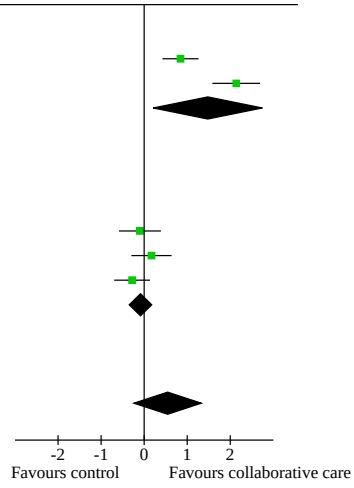


Footnotes

(1) WHOQOL-BREF, - 2 month follow-up
(2) SF-12, 6 month follow-up
(3) WHOQOL-BREF, 6 month follow-up

Analysis 4.3. Comparison 4: Collaborative care versus usual care (subgroup analyses), Outcome 3: Quality of life, physical health at 6 months - subgroup analysis: variations in implementation of the collaborative care intervention and healthcare systems

Study or Subgroup	Collaborative care Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
4.3.1 Physical health at 6 months (pharmacy collaborative care)								
Mishra 2017 (1)	56.4	16	51	45.2	8.9	45	20.3%	0.84 [0.43 , 1.26]
Salman 2014 (2)	46.7	5.7	41	35.2	4.9	39	19.5%	2.14 [1.58 , 2.69]
Subtotal (95% CI)			**92**			**84**	**39.8%**	**1.48 [0.21 , 2.75]**
Heterogeneity: Tau² = 0.78; Chi² = 13.32, df = 1 (P = 0.0003); I² = 92%								
Test for overall effect: Z = 2.28 (P = 0.02)								
4.3.2 Physical health at 6 months (no pharmacy collaborative care)								
Kilbourne 2012 (2)	34.8	7	32	35.5	7.2	33	19.9%	-0.10 [-0.58 , 0.39]
Kilbourne 2013 (2)	35.8	7.8	35	34.5	7.3	36	20.0%	0.17 [-0.30 , 0.64]
van der Voort 2015 (3)	53.7	17.8	37	58.6	17.1	57	20.3%	-0.28 [-0.70 , 0.14]
Subtotal (95% CI)			**104**			**126**	**60.2%**	**-0.09 [-0.35 , 0.18]**
Heterogeneity: Tau² = 0.00; Chi² = 2.00, df = 2 (P = 0.37); I² = 0%								
Test for overall effect: Z = 0.64 (P = 0.52)								
Total (95% CI)			**196**			**210**	**100.0%**	**0.55 [-0.24 , 1.33]**
Heterogeneity: Tau² = 0.74; Chi² = 57.07, df = 4 (P < 0.00001); I² = 93%								
Test for overall effect: Z = 1.36 (P = 0.17)								
Test for subgroup differences: Chi² = 5.60, df = 1 (P = 0.02), I² = 82.1%								

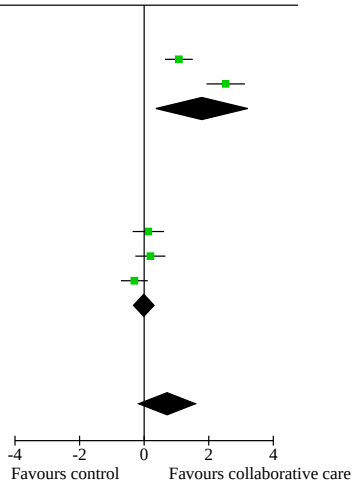


Footnotes

(1) WHOQOL-BREF, 2 month follow-up
(2) SF-12, 6 month follow-up
(3) WHOQOL-BREF, 6 month follow-up

Analysis 4.4. Comparison 4: Collaborative care versus usual care (subgroup analyses), Outcome 4: Quality of life, mental health at 6 months - subgroup analysis: variations in implementation of the collaborative care intervention and healthcare systems

Study or Subgroup	Collaborative care Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
4.4.1 Mental health at 6 months (pharmacy collaborative care)								
Mishra 2017	57.4	17	51	42.3	9.3	45	20.2%	1.07 [0.64 , 1.50]
Salman 2014	68.3	9.1	41	46.3	8.1	39	19.4%	2.53 [1.93 , 3.12]
Subtotal (95% CI)			**92**			**84**	**39.6%**	**1.79 [0.36 , 3.21]**
Heterogeneity: Tau² = 0.98; Chi² = 15.02, df = 1 (P = 0.0001); I² = 93%								
Test for overall effect: Z = 2.46 (P = 0.01)								
4.4.2 Mental health at 6 months (no pharmacy collaborative care)								
Kilbourne 2012	32.5	7.4	32	31.5	7.9	33	20.0%	0.13 [-0.36 , 0.62]
Kilbourne 2013	34.4	6.8	35	32.9	8.4	36	20.1%	0.19 [-0.27 , 0.66]
van der Voort 2015	48.2	17.3	37	53.6	17.9	57	20.3%	-0.30 [-0.72 , 0.11]
Subtotal (95% CI)			**104**			**126**	**60.4%**	**-0.01 [-0.33 , 0.31]**
Heterogeneity: Tau² = 0.03; Chi² = 2.95, df = 2 (P = 0.23); I² = 32%								
Test for overall effect: Z = 0.07 (P = 0.94)								
Total (95% CI)			**196**			**210**	**100.0%**	**0.71 [-0.17 , 1.59]**
Heterogeneity: Tau² = 0.94; Chi² = 69.42, df = 4 (P < 0.00001); I² = 94%								
Test for overall effect: Z = 1.58 (P = 0.11)								
Test for subgroup differences: Chi² = 5.85, df = 1 (P = 0.02), I² = 82.9%								



ADDITIONAL TABLES

Table 1. Collaborative care components of included studies (based on Gunn 2006 definition)

	Multi-professional approach				Structured management plan	Follow-ups	Enhanced interprofessional communication	No. components
	Primary care professional	Mental health professional	Case manager	Other				
Bauer 2006	None	Psychiatrist	Nurse	N/A		At least 1 appointment every 3 months		3
Chatterjee 2011	None	Psychiatrist Psychiatric social worker	Nurse	Community health workers	Medication management Psycho-education/health promotion	6 to 8 visits at home in months 0 to 3; 6 to 8 fortnightly visits in months 4 to 7; 6 visits in months 8 to 12	Clinical team reviews	3
Chwastiak 2018	Advanced practice registered nurse	Community Mental Health Centre (CMHC) psychiatrist CMHC nurse		Endocrinologist consultant	Health plan Motivational interviewing and behavioural activation Medication management	60-minute health assessment; 30-minute visits every other week for 12 weeks; then monthly visits for up to 6 months	Intra-clinic communication Clinical meetings Caseload review	4
Kilbourne 2012	None	Social worker interventionist	Nurse	N/A	Evidence-based guidelines Self-management support	20-minute contacts for up to 6 months		3
Kilbourne 2013	None	Mental health providers	Health specialist	General medical providers	Evidenced-based guidelines	4 x 2-hour weekly group self-management sessions and brief care management contacts for up to 6 months		3
Mishra 2017	None	Psychiatrist	None	N/A	Medication management	3 appointments	None	3

Table 1. Collaborative care components of included studies (based on Gunn 2006 definition) *(Continued)*

		Hospital pharmacist						
Salman 2014	None	Psychiatrist Psychologist Hospital pharmacist	NR	N/A	Medication management	Contact every 2 weeks, via telephone and clinic appointments. Clinic visits scheduled on weeks 2, 6, 12 and 24.	Daily meetings	3
van der Voort 2015	None	Psychiatrist Mental health nurse	Mental health nurse	N/A	Treatment plan Psychoeducation Problem-solving treatment	Psychoeducation 6 x 2-hour sessions Problem solving training x 6 sessions; other pharmacotherapy + somatic care 'continues as appropriate'	Meetings	3

CMHC: community mental health care clinic; N/A: not applicable; NR: not reported

Table 2. Collaborative care components of included studies

		Bauer 2006	Chatterjee 2011	Kilbourne 2012	Kilbourne 2013	Salman 2014	van der Voort 2015	Mishra 2017	Chwastiak 2018
Identifying patients	Provider referral		X						
	Systematic screening/medical record review		X		X				
Multi-professional approach	Psychiatrist					X	X	X	X
	Psychologist					X			
	Mental health nurse						X		X
	Medical nurse								X
	GP/family doctor								
	Pharmacist					X		X	

Table 2. Collaborative care components of included studies *(Continued)*

	Medical consultant/specialist																		X
	Community health worker			X															
	Social worker			X															
	MSW interventionist					X													
	Health specialist							X											
	Mental health provider							X											
	Primary care provider																		
	General medical provider							X											
Case manager						X		X		X						X			
Training	Staff training			X								X						X	
	Supervision			X		X						X							
Enhanced interprofessional communication	Meetings			X						X		X						X	
	Written correspondence																		
	Electronic records sharing							X											
	Caseload/clinical review			X														X	
Structured management plan	Treatment guidelines, protocol, algorithm					X		X		X									
	Medication management			X						X		X		X					
	Psychological treatment/approaches/therapy/											X						X	
Scheduled patient follow-ups		X		X		X		X		X		X		X		X		X	

Table 2. Collaborative care components of included studies *(Continued)*

Patient education	Psycho-education	X	X			X			
	Written materials						X	X	
	Mode unclear				X				
Family education		X			X	X	X		
Measurement-based care			X	X		X		X	
Tailoring (personalised health/ treatment plan/ needs assessment)		X	X			X		X	
Self-management support			X	X				X	
Community network linkages		X							

GP: general practitioner; MSW: masters degree social worker

Table 3. Outcome measures of interest from the included studies

Outcome	Studies reporting outcome	Name of measure/source	Description of validated measures used to assess outcome
Mental state			
Symptoms of schizophrenia	Chatterjee (2011) - India Salman (2014) - Pakistan	Positive And Negative Syndrome Scale (PANSS) Kay 1987	A 30-item, 7-point rating instrument, which has adapted 18 items from the Brief Psychiatric Rating Scale (BPRS) and 12 items from the Psychopathology Rating Schedule (PRS). PANSS items are rated on a 7-point scale (1 = absent, 2 = minimal, 3 = mild, 4 = moderate, 5 = moderate severe, 6 = severe, and 7 = extreme); because the absence of symptoms is equal to 1 point, the lowest possible total score on both PANSS scales is 7. The scale can be divided into three sub-scales for measuring the severity of general psychopathology, positive symptoms (PANSS-P) and negative symptoms (PANSS-N). Lower scores indicate lower symptom severity.
Symptoms of bipolar	Bauer (2006) - USA	Longitudinal Interval Follow-up Examination (LIFE) Keller 1987	LIFE is a semi-structured interview, which uses timeline follow-back methodology to provide weekly psychiatric symptom ratings (PSRs) for mania and depression based on the number of DSM-IV criteria endorsed: no or minimal symptoms (PSR 1 to 2), subthreshold symptoms (PSR 3 to 4) or episode (PSR 5 to 6). Lower scores indicate lower symptom severity.
	Van der Voort (2015) - Netherlands	Retrospective Life Chart Method (LCM) Leverich 1998	Patients were asked to rate retrospectively their average mood, in each consecutive month, over the past 6 months; scores are based on the severity of mood symptoms and the associated degree of functional impairment. The LCM consists of a scale for manic symptoms (+1 to +4) and a scale for depressive symptoms (-1 to -4); a score of 0 indicates balance or a euthymic state. Scores of ± 4 refer to syndromal episodes, whereas scores of ± 1 refer to subthreshold symptoms with only mild functional impairment. -4 represents severe depression +4 represents severe mania
	Kilbourne (2012) - USA Kilbourne (2013) - USA	The Internal State Scale (ISS) Bauer 1991 The Internal State Scale (ISS) Glick 2003	The Internal State Scale (ISS) is a simple self-report instrument for discriminating mood state and tracking both manic and depressive symptoms in bipolar disorder (Bauer et al 2000). The ISS is a 15-item self-report instrument using the visual analogue line scale format. Each item is a statement followed by a 100 mm line with anchor points at 0 and 100. The 0 anchor point is 'Not at all, rarely' and the 100 anchor point is 'Very much so, much of the time'. Items for each of the subscales are then summed to provide the subscale score (Bauer et al 2000). Lower scores indicate lower symptom severity. Converted visual analogue scale-based scoring to 10-point Likert scoring.
Symptoms of psychosis, anxiety and depression	Chwastiak (2018) - USA	Brief Psychiatric Rating Scale (BPRS) Overall 1962	The BPRS is a rating scale developed to characterise psychopathology. The scale was originally developed with 16 items, and updated in 1965 to the standard 18-item version. The BPRS is widely used to assess the effectiveness of treat-

Table 3. Outcome measures of interest from the included studies *(Continued)*

			ment. A clinician rates each item on a scale ranging from 1 (not present) to 7 (extremely severe). The BPRS can also yield an overall score, with scores ranging from 0 to 126, with higher scores indicating more severe (positive, negative and affective) symptoms of psychosis. Lower scores indicate lower symptom severity.
Symptoms of depression	Chwastiak (2018) - USA	Patient Health Questionnaire-9 (PHQ-9) Kroenke 2001	PHQ-9 is a self-administered patient questionnaire version of the PRIME-MD diagnostic instrument for common mental disorders. The PHQ-9 is the depression module, which scores each of the nine DSM-IV criteria as "0" (not at all) to "3" (nearly every day). The scores from each of the 9 criteria are then totalled. Depression severity: 0 to 4 none, 5 to 9 mild, 10 to 14 moderate, 15 to 19 moderately severe, 20 to 27 severe. Lower score indicates lower symptom severity.
	Van der Voort (2015) - Netherlands	Quick Inventory of Depressive Symptomatology (QIDS) scale Rush 2003	A 16-item instrument for depressive symptom severity derived from the 30-item Inventory of Depressive Symptomatology (IDS). It assesses the 9 DSM-IV diagnostic symptom domains and is available in clinician rating and self-report. The scores for three domains (sleep, appetite/weight and restlessness/agitation) are based upon the maximum score (most pathological) of 2 or more questions. Each of the remaining domains are rated by a single item. All domains are scored from 0 to 3, with higher scores reflecting greater psychopathology. Total QIDS scores range from 0 to 27, with scores of 5 or lower indicative of no depression, scores from 6 to 10 indicating mild depression, 11 to 15 indicating moderate depression, 16 to 20 reflecting severe depression and total scores greater than 21 indicating very severe depression. Higher scores indicate greater psychopathology.
Symptoms of mania	Van der Voort (2015) - Netherlands	The Altman Self-Rating Mania (ASRM) scale Altman 1997	The Altman Self-Rating Mania Scale is a short, 5-item self-assessment questionnaire for assessing the presence and severity of manic or hypomanic symptoms. Each item on the measure (elevated mood, increased self-esteem, decreased sleep, pressured sleep and psychomotor agitation) is rated on a 5-point scale (i.e. 0 to 4) with the response categories having different anchors depending on the item. A score of 6 or higher indicates a high probability of a manic or hypomanic condition. A score of 5 or lower is less likely to be associated with significant symptoms of mania. Higher scores indicate greater symptom severity.
Physical health status			
	Kilbourne (2012) - USA	Systolic/diastolic blood pressure CDC BMI	< 120/< 80 mmHg = normal 120 to 129/< 80 mmHg = elevated 130 to 139 mmHg or 80 to 89 mmHg = hypertension, stage 1 ≥ 140 mmHg or ≥ 90 mmHg = hypertension, stage 2

Table 3. Outcome measures of interest from the included studies *(Continued)*

			kg/m^2: healthy range 18 to 24 kg/m^2, > 25 = overweight, > 30 = obesity
	Kilbourne (2013) - USA	Systolic/diastolic blood pressure Total cholesterol Grundy 2018 BMI Waist circumference Haemoglobin A1c (HbA1c) Framingham Risk Score Wilson 1998	mmHg Desirable total cholesterol level = < 200 mg/dL kg/m^2: healthy range 18 to 24 kg/m^2, > 25 = overweight, > 30 = obesity cm/inches The Framingham Risk Score is an algorithm calculation of an individual's risk of developing or dying from coronary heart disease within the next 10 years. Individuals receive a point score based on categorical values of age, total cholesterol, high-density lipoprotein cholesterol, blood pressure, smoking and diabetes. Scores are gender-specific. < 10% = low risk, 10% to 20% = intermediate risk, > 20% = high risk.
	Chwastiak (2018) - USA	HbA1c Systolic blood pressure Total cholesterol BMI Smoking Fagerström Test for Nicotine Dependence (FTND) Heatherton 1991	mmHg mg/dL kg/m^2: healthy range 18 to 24 kg/m^2, > 25 = overweight, > 30 = obesity The Fagerström Test for Nicotine Dependence is a standard instrument for assessing the intensity of physical addiction to nicotine. It contains 6 items that evaluate the quantity of cigarette consumption, the compulsion to use, and dependence. Yes/no items are scored from 0 to 1 and multiple-choice items are scored from 0 to 3. The items are summed to yield a total score of 0 to 10. The higher the total Fagerström score, the more intense is the patient's physical dependence on nicotine.
Quality of life	Bauer (2006) - USA	Medical Outcomes Study 36- item Short Form Health Survey (SF-36) Ware 1992	The SF-36 is a multi-purpose, short-form health survey with 36 questions. It measures 8 health concepts: 1) physical functioning; 2) role limitations because of physical health problems; 3) bodily pain; 4) social functioning; 5) general mental health (psychological distress and psychological well-being); 6) role limitations because of emotional problems; 7) vitality (energy/fatigue); and 8) general health perceptions. Summary measures of physical and mental health, PCS and MCS, are calculated from the 8 scales using algorithms, which are strictly controlled by a private company. Scores are calibrated so that 50 is the average score or norm. Individual respondent's scale scores below 45, or a group mean scale score below 47, would suggest health status to be below the average range. Higher scores indicate better quality of life.
	Kilbourne (2012) - USA Kilbourne (2013) - USA	Short Form Health Survey (SF-12) Ware 1996	The 12-item Short Form Health Survey is a shortened version of its predecessor, the SF-36, using the same 8 domains: 1) physical functioning; 2) role limitations because of physical health problems; 3) bodily pain; 4) social functioning; 5) general mental health (psychological distress and psychological well-being); 6) role limitations because of emotional problems; 7) vitality (energy/fatigue); and 8) general health perceptions. It

Table 3. Outcome measures of interest from the included studies *(Continued)*

	Salman (2014) - Pakistan		includes two composite scores for physical (PCS) and mental health (MHS) (range 0 to 100). However, it is recommended that users base their interpretations on norm-based scores (mean = 50, SD = 10) rather than 0 to 100 scores. Individual respondent's scale scores below 45, or a group mean scale score below 47, would suggest health status to be below the average range. Higher scores indicate better physical health.
	Mishra (2017) - India Van der Voort (2015) - Netherlands	WHOQOL-BREF Nelson 1999 Trompenaars 2005 Dutch population	The WHOQOL-BREF is a self- administered, short form quality of life assessment, abbreviated from the WHOQOL-100. It contains 26 questions and is based on a 4 domain structure (physical health, psychological, social relationships and environment), plus one question for overall quality of life and one question for general health. Domain scores are scaled in a positive direction (i.e. higher scores denote higher quality of life). The mean score of items within each domain is used to calculate the domain score. Mean scores are then multiplied by 4 in order to make domain scores comparable with the scores used in the WHOQOL-100. The first transformation method converts scores to range between 4 and 20, comparable with the WHOQOL-100. A second transformation method converts domain scores to a 0 to 100 scale. Higher scores indicate higher quality of life.
Functioning	Bauer (2006) - USA	Social Adjustment Scale II SAS Schooler 1979	SAS contains 42 items that assess role performance in the past 2 weeks across 6 domains: work/school role, social/leisure time, family outside the home, primary relationship, parental role and family unit. Each item is rated on a 5-point scale. An overall adjustment score is obtained by summing the scores of all the items and dividing by the number of items actually answered. Lower scores indicate poorer functioning. This measure could not be included due to the data not being reported.
	Van der Voort (2015) - Netherlands	Functioning Assessment Short Test (FAST) Rosa 2007	FAST is a short instrument, patient-rated, and scores are rated on a 4-point Likert scale. It comprises 24 items, and covers 6 areas of functioning: autonomy, occupational functioning, cognitive functioning, financial issues, interpersonal relationships and leisure time. Higher scores indicate greater impairment in functioning.
Disability Assessment Scale	Kilbourne (2012) - USA Kilbourne (2013)- USA	World Health Organization Disability Assessment Schedule 2.0 (WHO-DAS) Ustun 2010	WHO-DAS is a self-administered, 36-item questionnaire. It assesses disability across 6 domains (cognition, mobility, self-care, getting along, life activities and participation). The individual rates how much difficulty he or she has had in specific areas of functioning during the past 30 days. There are two ways of scoring the questionnaire. Simple - the scores assigned to each of the items are on a 0 to 4 scale, 0 representing no difficulty, 4 representing extreme difficulty. These scores can then be summed. Or via a complex method of 'item response theory', which uses a computer to determine the summary score by differentially weighting the items and the levels of severity. Kilbourne et al have used the 12-item brief assessment form, which allows for calculation of an overall functioning score, ex-

Table 3. Outcome measures of interest from the included studies *(Continued)*

			plaining 81% variance of the 36-item version. Scores for each question are scored on a 0 to 4 scale as for the 36-item version. Higher score indicates higher disability.
Disability Assessment Scale	Chatterjee (2011) - India	Indian Disability Evaluation and Assessment Scale (IDEAS) Thara 2002	IDEAS is best suited for the purpose of measuring and certifying disability. It has 4 items: self care, interpersonal activities (social relationships), communication and understanding, and work. Each item is scored between 0 and 4, i.e. from no to profound disability; adding scores on 4 items gives the 'total disability score'. Global disability score is calculated by adding the 'total disability score' and MI2Y score (months in 2 years - a score ranging between 1 and 4, depending on the number of months in the last 2 years the patient exhibited symptoms). Global disability score of 0 (i.e. 0%) corresponds to 'no disability', a score between 1 and 7 (i.e. 40% corresponds to moderate to profound disability. A higher score indicates a greater disability.
Medication adherence	Mishra (2017) - India Salman (2014) - Pakistan	Medication Adherence Rating Scale (MARS) Thompsom 2000	MARS describes an individual's medication adherence in 3 dimensions: medication adherence behaviour, attitude toward taking medication, and negative side effects and attitudes to psychotropic medication. It is a 10-item self-report questionnaire developed after combining the Medication Adherence Questionnaire and the Drug Attitude Inventory. Each question has a yes or no response. A response consistent with non-adherence is coded as 0, whereas a response consistent with adherence is coded as 1. For questions 1 to 6 and 9 to 10, a no response is indicative of adherence and is coded as 1, while for questions 7 and 8, a yes response is indicative of adherence and is coded as 1. Total scores on the MARS may range between 0 and 10. Higher scores indicate better medication adherence. This measure could not be included for the Salman study due to the data not falling into the range of values permissible by the scale.
	Salman (2014) - Pakistan	Morisky Medication Adherence Scale (MMAS-4) Morisky 1986	A structured 4-item self-reported adherence measure. The MMAS-4 is used mainly as a screening test in the clinical setting. This 4-item version (MMAS-4) requires a dichotomous response of yes or no, and includes elements of forgetfulness and symptom severity. Scores are high to low, with yes = 0 and no = 1. Total scores range between 0 and 4. This measure could not be included due to the data not falling into the range of values permissible by the scale.
	Van der Voort (2015) - USA	DAI-10 Awad 1993 Hogan 1983	The DAI-10 is a shortened version of the DAI-30. The 10-item questionnaire requires a true or false response. A patient who is fully adherent to prescribed medication would answer true to 6 items and false to 4 items. Each positive answer is given a score of plus one, and each negative answer is given a score of minus one. The total score for each patient is calculated as the sum of the positive scores, minus the negative scores. A positive total score indicates a positive subjective response (adherent) and a negative total score indicates a negative subjective response (non-adherent).

BMI: body mass index; DSM-IV: Diagnostic and Statistical Manual, version 4; MCS: mental component score; PCS: mental component score

Table 4. Outcomes prespecified and data available

Primary outcomes pre-specified in review protocol	Data available (and section reported in results)
1.1 Quality of life	
Clinically important change in quality of life (as defined by individual studies) (Y/N, binary outcome) at 12 months	
	1.1 Quality of life: clinically important change (average endpoint in mental health component) - 12 months
1.2 Mental state	
Clinically important change in mental state (as defined by individual studies) (Y/N, binary outcome) at 12 months	1.2 Mental state: clinically important change (binary) - 12 months
1.3 Psychiatric admissions	
Number of participants admitted to hospital (psychiatric admissions) at 12 months	1.3 Psychiatric hospital admissions: number of participants admitted to hospital (12 months)
Secondary outcomes pre-specified in review protocol	
2.1 Quality of life	
Clinically important change in quality of life (as defined by individual studies) (Y/N, binary outcome) (time points other than 12 months)	
Clinically important change in quality of life at 12 months (as defined by individual studies) (Y/N, binary outcome)	
Any change in quality of life	
Average endpoint quality of life score	2.1.6 Quality of life: overall endpoint (WHOQOL-BREF) - 6 months 2.1.7 Quality of life: overall endpoint (WHOQOL-BREF) - 12 months
Average change in quality of life scores	
No clinically important change in specific aspects of quality of life (as defined by individual studies)	
Any change in specific aspects of quality of life	
Average endpoint in specific aspects of quality of life scores	2.1.1 Quality of life: average endpoint in physical health - up to 6 months 2.1.2 Quality of life: average endpoint in physical health - 12 months 2.1.3 Quality of life: average endpoint in physical health - more than 12 months 2.1.4 Quality of life: average endpoint in mental health - up to 6 months

Table 4. Outcomes prespecified and data available *(Continued)*

	2.1.5 Quality of life: average endpoint in mental health - more than 12 months
Average change in specific aspects of quality of life scores	
2.2 Mental state **General and specific (including positive and negative symptoms of psychosis, and mood (as defined by individual studies)**	
Any change in mental state Average endpoint mental state Average change in mental state No clinically important change in mental state (as defined by individual studies) Any change in specific aspects of mental state Average endpoint in specific aspects of mental state Average change in specific aspects of mental state	2.2.1 Mental state (overall general score) up to 6 months 2.2.2 Mental state (general psychopathology) up to 6 months 2.2.3 Mental state (general psychopathology) at 12 months 2.2.4 Mental state (positive symptoms) up to 6 months 2.2.5 Mental state (positive symptoms) at 12 months 2.2.6 Mental state (negative symptoms) up to 6 months 2.2.7 Mental state (negative symptoms) at 12 months 2.2.8 Mental state (depressive symptoms) up to 6 months 2.2.9 Mental state (depressive symptoms) at 7 to 12 months 2.2.10 Mental state (depressive symptoms) at 24 months 2.2.11 Mental state (manic symptoms) up to 6 months 2.2.12 Mental state (manic symptoms) at 7 to 12 months 2.2.13 Mental state (manic symptoms) greater than 12 months
2.3 Psychiatric admissions	
Number of participants admitted to hospital (psychiatric admissions) greater than 12 months	2.3.1 Number of participants admitted to hospital (year 2)
	2.3.2 Number of participants admitted to hospital (year 3)
Mean number of days in hospital for psychiatric admissions	
Length of time to readmission (psychiatric admissions)	
2.4 Other hospital admissions	
Number of participants admitted to hospital (physical health admission)	2.4.1 Number of participants admitted to hospital (up to 12 months) 2.4.2 Number of participants admitted to hospital (in year 2) 2.4.3 Number of participants admitted to hospital (in year 3)

Table 4. Outcomes prespecified and data available *(Continued)*

Mean number of days in hospital for physical health admissions	
Length of time to readmission (physical health admissions)	
2.5 Personal recovery	2.5 No data available
Clinically important change in personal recovery (as defined by individual studies) (Y/N, binary outcome)	
Any change in personal recovery	
Average endpoint personal recovery score	
Average change in personal recovery scores	
No clinically important change in specific aspects of personal recovery (as defined by individual studies)	
Any change in specific aspects of personal recovery	
Average endpoint in specific aspects of personal recovery scores	
Average change in specific aspects of personal recovery scores	
2.6 Physical health status (including specific measures of blood pressure, blood cholesterol, blood glucose - HbA1c, body mass index (BMI)	
Clinically important change in physical health status (as defined by individual studies) Any change in physical health status score Average endpoint physical health status score Average change in physical health status score	2.6.1 Blood pressure, mmHg systolic - up to 6 months 2.6.2 Blood pressure, mmHg systolic - at 7 to 12 months 2.6.3 Blood pressure, mmHg systolic - 24 months 2.6.4 Blood pressure, mmHg diastolic - 6 months 2.6.5 Blood pressure, mmHg diastolic - 7 to 12 months 2.6.6 Blood pressure, mmHg diastolic - 24 months 2.6.7 Body mass index (BMI) - 6 months 2.6.8 Body mass index (BMI) - 12 months 2.6.9 Body mass index (BMI) - 24 months 2.6.10 Total cholesterol - 6 months 2.6.11 Total cholesterol - 12 months 2.6.12 Total cholesterol - 24 months 2.6.13 Triglycerides - up to 6 months 2.6.14 High-density lipoprotein (HDL) - 6 months 2.6.15 High-density lipoprotein (HDL) - 12 months 2.6.16 High-density lipoprotein (HDL) - 24 months 2.6.17 Low-density lipoprotein (LDL) - 6 months

Table 4. Outcomes prespecified and data available *(Continued)*

	2.6.18 Low-density lipoprotein (LDL) - 12 months 2.6.19 Low-density lipoprotein (LDL) - 24 months 2.6.20 HbA1c - up to 6 months 2.6.21 Waist circumference - 6 months 2.6.22 Waist circumference - 12 months 2.6.23 Waist circumference - 24 months
2.7 Global state	2.7 No data available
Relapse (as defined by individual studies)	
Time to relapse	
Clinically important change in global state (as defined by individual studies)	
Any change in global state	
Average endpoint global state score	
Average change in global state score	
2.8 - 2.9 Medication adherence	
Clinically important change in compliance (patient-reported) Any change in compliance (patient-reported) Clinically important change in compliance (carer-reported) Any change in compliance (carer-reported)	2.8 Medication adherence (patient-reported) (DAI-10) 2.8.1 Medication adherence (patient-reported) at 6 months 2.8.2 Medication adherence (patient-reported)at 12 months 2.9.1 Medication adherence (patient-reported) - up to 6 months
2.10 - 2.11 Social functioning	
Clinically important change in social functioning (as defined by individual studies) Any change in social functioning Average endpoint social functioning score Average change in social functioning scores	2.10.1 Social functioning/disability (binary) - 12 months 2.11.1 Social functioning/disability - up to 6 months 2.11.2 Social functioning/disability - 12 months 2.11.3 Social functioning/disability - 24 months
Employment status	
Living tenure (number of participants homeless, in unstable housing or living independently)	
2.12 Substance use (alcohol/illicit drugs/cigarettes/tobacco)	2.12 No data available
Clinically important change in substance use (as defined by individual studies)	

Table 4. Outcomes prespecified and data available *(Continued)*

Any change in substance use	
Average endpoint substance use	
Average change in substance use	
2.13 Adverse effect/event(s)	2.13 No data available
At least one adverse effect	
Incidence of specific effect (e.g. cardiovascular, metabolic, movement disorders)	
2.14 Death	
Number of participants who died from suicide	2.14.1 Number of participants that died from suicide (36 months) 2.12.3 Number of participants that died from suicide (12 months)
Number of participants who died from natural causes	2.14.2 Number of participants that died from natural causes (36 months) 2.14.4 Number of participants that died from natural causes (6 months)
	2.14.5 Number of participants that died (all causes) (12 months)
2.15 Service use outside of mental health (i.e. primary care, emergency services, walk-in centres, social services)	2.15 No data available
Mean number of contacts per month	
Number of participants in contact with service	
Mean number of service hours per month	
2.16 - 2.17 Cost of treatment	
Direct cost of inpatient care	
Direct cost of health and social care (including the above, plus the costs of all other medical and psychiatric care, such as: outpatient care and specialist service, collaborative care and community-based social services)	
Total costs, including types of costs above, plus the costs of accommodation and minus benefits, such as earnings where these are known	2.16.1 Cost of treatment (USD 1000) - at 36 months 2.17 Cost of treatment (international dollars (Int$)) up to 12 months
2.18 Experience of care/satisfaction (participant/carer/staff)	2.18 No data available
Clinically important change in experience of care/participant, carer and staff satisfaction (as defined by individual studies)	

Table 4. Outcomes prespecified and data available *(Continued)*

Any change in experience of care/participant, carer and staff satisfaction	
Average endpoint experience of care/participant, carer and staff satisfaction score	
Average change in experience of care/participant, carer and staff satisfaction score	
2.19 Leaving the study early (attrition)	
For any reason	2.19.1 Attrition/leaving the study early (lost to follow-up 6 months) 2.19.2 Attrition/leaving the study early (lost to follow-up 12 months) 2.19.3 Attrition/leaving the study early (lost to follow-up at 24 months) 2.19.4 Attrition/leaving the study early (lost to follow-up at 36 months)
For specific reason	

Table 5. Future study design

Study design	Study design	Recommendation
Setting		Primary care and community mental healthcare services
Method	Duration	Minimum of 12-month follow-up
	Allocation	Randomised (cluster or individual)
	Blinding	Blinding of outcome assessors Blinding of statisticians Allocation concealment
	Outcomes	• Psychiatric admissions • Intervention costs • Quality of life • Mental state • Social functioning • Personal recovery • Adverse effects, e.g. acute mental health episodes (defined as number of crises in time period) • Process outcomes • Experience of care/satisfaction • Fidelity Core outcome set for bipolar I and II (Retzer 2020)
	Retention	Utilise participant retention strategies

Table 5. Future study design *(Continued)*

	Analysis	Produce and report analysis plans prior to research being conducted
Participants	Diagnosis	Severe mental illness
	Age	18+ (or younger if in receipt of adult services)
	Gender	All
	N	300+
Intervention		Collaborative care according to Gunn 2006 description to include: • Multidisciplinary working (which includes primary care) • Enhanced communication between providers (e.g. via case manager, multi-disciplinary team meetings) • Scheduled and proactive follow-up • Delivery of evidence-based treatment according to algorithms/protocols

APPENDICES

Appendix 1. Previous definitions of collaborative care

Authors	Conditions under review	Definition of collaborative care
Druss 2005	Mental and addictive disorders	*"This approach, based on Wagner's Chronic Care Model (Bodenheimer 2002), uses a multidisciplinary team including both mental health and primary care providers to ensure co-ordination and follow-up with care (Katon 1995). Regardless of whether services are collocated, the key element of these collaborative care approaches is that they involve functionally integrated care teams." (pg 150)*
Bower 2006	Depression	*"A multifaceted organisational intervention, which could include a number of components:* *(a) the introduction of a new role (case manager) into primary care, to assist in the management of patients with depression through structured and systematic delivery of interventions;* *(b) the introduction of mechanisms to foster closer liaison between primary care clinicians and mental health specialists (including case managers) around individual patient care;* *(c) the introduction of mechanisms to collect and share information on the progress of individual patients." (pg 485)*
Craven 2006	A range of mental health disorders, including depression and severe mental illness	*"Collaborative care involves providers from different specialties, disciplines or sectors working together to offer complementary services and mutual support, to ensure that individuals receive the most appropriate service from the most appropriate provider in the most suitable location, as quickly as necessary, and with a minimum of obstacles. Collaboration can involve better communication, closer personal contacts, sharing of clinical care, joint educational programs and/or joint program and system planning." (pg 9)*

(Continued)

Gilbody 2006	Depression	*"involving a structured approach to care based on chronic disease management principles and a greater role for nonmedical specialists such as nurse practitioners working in conjunction with the primary care physician and a mental health specialist (*Katon 2001*). Collaborative care captures a range of interventions of varying intensity, ranging from simple telephone interventions to encourage compliance with medication (*Peveler 1999*) to more complex interventions that involve intensive follow-up and incorporate a form of structured psychosocial intervention.(*Wells 2000*)" (pg 2314-5)*
Gunn 2006	Depression	*"1. A multi-professional approach to patient care. This required that a general practitioner (GP) or family physician and at least one other health professional (e.g. nurse, psychologist, psychiatrist, pharmacist) were involved with patient care.* *2. A structured management plan. In line with introducing an organised approach to patient care 'systems' trials were required to offer practitioners access to evidence based management information. This could be in the form of guidelines or protocols. Interventions could include both pharmacological (e.g. antidepressant medication) and non-pharmacological interventions (e.g. patient screening, patient and provider education, counselling, cognitive behaviour therapy).* *3. Scheduled patient follow-ups. A 'systems' approach required interventions to have an organised approach to patient follow-up. We defined this as one or more scheduled telephone or in-person follow-up appointments to provide specific interventions, facilitate treatment adherence, or monitor symptoms or adverse effects.* *4. Enhanced inter-professional communication. This required that the intervention introduced mechanisms to facilitate communication between professionals caring for the depressed person. This included team meetings, case conferences, individual consultation/supervision, shared medical records, patient-specific written or verbal feedback between care-givers and was sometimes referred to as 'collaborative care' in the publications." (pg 2)*
Archer 2012	Depression and anxiety	*"For the purposes of this review, collaborative care is defined as a multifaceted intervention which involves 3 distinct professionals working collaboratively within the primary care setting. One professional works as a case manager, one as a primary care practitioner and the other as the mental health specialist."* (Katon 2001) *(pg 3).* *"The specific roles each of these professionals are detailed below:* *-Primary care practitioner: will provide the initial recognition, diagnosis and treatment.* *-Case manager: will provide medication management and psychological intervention, proactively follow-up patients, assess adherence to treatment and monitor progress and feedback to the primary care physician.* *-Mental health specialist: will provide support/consultation to either the case manager or the primary care physician. This role maybe played by others other than a medically qualified professional i.e. nurse specialists (*Gask 2005*)." (pg 3)*

Appendix 2. Defining type A and B collaborative care

Type A collaborative care

Type A interventions are described as collaborative care by the trialists and are comprised of the four 'core' components as outlined below:

a) A multi-professional approach to patient care. A primary care provider and at least one other health professional or paraprofessional is involved with patient care.

A primary care professional could be a General Practitioner (or family doctor), practice nurse, pharmacist, dentist, optician or any other generalist providing medical health care within the community.

Collaborative care interventions aim to integrate generalist primary care with specialist health services so in addition to primary care there should be another non-primary care professional involved in the patient's care, for example a social worker, community psychiatric nurse, psychiatrist or occupational therapist.

b) A structured management plan in the form of evidence-based protocols or guidelines.

Any form of guideline, protocol or algorithm can be defined as a structured management plan, for example, a study protocol and/or manual, treatment algorithm (for medication and/or therapy), treatment management plan, care plan or stepped care plan.

The aim of collaborative care is to develop plans which are evidence based.

c) Scheduled patient follow-ups.

Follow-ups may be for the purpose of monitoring clinical status, side effects or medication adherence. A protocol may have been developed to manualise the process of follow-up, for example, frequency, purpose and format for contact.

d) Enhanced inter-professional communication.

Enhanced communication could take place through case conference, regular team meetings, case by case consultation, written correspondence, e.g. via email or through linked electronic records.

Any method/approach used to ensure regular communication between the people involved in caring for the patient takes place can be defined as 'enhanced'.

Type B collaborative care

Type B interventions are described as collaborative care by the trialists but are **not** comprised of the four 'core' components.

Appendix 3. Outcomes reported in our previously published Cochrane review

We have changed the outcomes from those reported in the original review (shown below). We outline this in the Types of outcome measures section and explain the reason for the change in Differences between protocol and review.

Types of outcome measures

Primary outcomes

1. Psychiatric admissions

1.1 Number of participants admitted to hospital

Secondary outcomes

2. Hospital admissions

2.1 Mean number of days in hospital for psychiatric admissions
2.2 Length of time to readmission (psychiatric admissions)
2.3 Number of participants admitted to hospital (physical health problem)
2.4 Mean number of days in hospital for physical health admissions
2.5 Length of time to readmission (physical health admissions)

3. Mental state

3.1 Clinically important change in general mental state symptoms (as defined by individual studies)
3.2 Any change in general mental state
3.3 Average endpoint general mental state score
3.4 Average change in general mental state scores
3.5 Clinically important change in specific symptoms, including positive and negative symptoms of psychosis, and mood (as defined by individual studies)
3.6 Any change in specific symptoms
3.7 Average endpoint specific symptoms score
3.8 Average change in specific symptoms scores

4. Physical health status (including specific measures of blood pressure, blood cholesterol, blood glucose- HbA1c, body mass index (BMI)

4.1 Clinically important change in physical health status (as defined by individual studies)
4.2 Any change in physical health status score
4.3 Average endpoint physical health status score
4.4 Average change in physical health status score

5. Global state

5.1 Relapse (as defined by individual studies)
5.2 Time to relapse
5.3 Clinically important change in global state (as defined by individual studies)
5.4 Any change in global state
5.5 Average endpoint global state score
5.6 Average change in global state score

7. Social functioning

7.1 Clinically important change in social functioning (as defined by individual studies)
7.2 Any change in social functioning
7.3 Average endpoint social functioning score
7.4 Average change in social functioning scores
7.5 Employment status
7.6 Living tenure (number of participants homeless, in unstable housing or living independently)

8. Alcohol use

8.1 Clinically important change in alcohol use (as defined by individual studies)
8.2 Any change in alcohol use
8.3 Average endpoint alcohol use
8.4 Average change in alcohol use

9. Illicit drug use

9.1 Clinically important change in illicit drug use (as defined by individual studies)
9.2 Any change in illicit drug use
9.3 Average endpoint in illicit drug use
9.4 Average change in illicit drug use

10. Cigarettes/tobacco smoked

10.1 Clinically important change in cigarettes/tobacco smoked (as defined by individual studies)
10.2 Any change in average number of cigarettes smoked (or rolling tobacco)
10.3 Average endpoint number of cigarettes smoked (or rolling tobacco)
10.4 Average change in number of cigarettes smoked (or rolling tobacco)

11. Death

11.1 Number of participants who died from suicide
11.2 Number of participants who died from natural causes

12. Quality of life

12.1 Clinically important change in quality of life (as defined by individual studies)
12.2 Any change in quality of life
12.3 Average endpoint quality of life score
12.4 Average change in quality of life scores
12.5 No clinically important change in specific aspects of quality of life (as defined by individual studies)
12.6 Any change in specific aspects of quality of life
12.7 Average endpoint in specific aspects of quality of life scores
12.8 Average change in specific aspects of quality of life scores

13. Service use outside of mental health (i.e. primary care, emergency services, walk-in centres, social services)

13.1 Mean number of contacts per month
13.2 Number of participants in contact with service
13.3 Mean number of services' hours per month

14. Cost of treatment

14.1 Direct cost of inpatient care
14.2 Direct cost of health and social care (including the above, plus the costs of all other medical and psychiatric care, such as: outpatient care and specialist service, collaborative care and community-based social services)
14.3 Total costs, including types of costs above, plus the costs of accommodation and minus benefits, such as earnings where these are known

15. Satisfaction (participant and carer)

15.1 Clinically important change in participant and carer satisfaction (as defined by individual studies)
15.2 Any change in participant and carer satisfaction
15.3 Average endpoint participant and carer satisfaction score
15.4 Average change in participant and carer satisfaction score

16. Staff satisfaction

16.1 Change in staff satisfaction (as defined by individual studies)
16.2 Average endpoint staff satisfaction score
16.3 Average change in staff satisfaction score

17. Attrition

17.1 Leaving the study early (lost to follow-up)
17.2 Leaving the study for a specific reason (as defined by individual studies)

Appendix 4. Searches by Cochrane Schizophrenia

1. Cochrane Schizophrenia Register of Trials

We searched the Cochrane Schizophrenia register using the terms:

(*collaborative care* OR *collab* in title, abstract, indexing terms of REFERENCE or interventions of STUDY)

This register is compiled by systematic searches of major databases, handsearches and searches of conference proceedings (see Group module). We recognised that using this register alone may have limited identifying trials of bipolar disorder and other types of psychosis. We supplemented the electronic searches with reference list searches and contacted experts in the field of collaborative care (see: Searching other resources).

2. Searching other resources

2.1 Reference searching

We examined the reference lists of all included studies for additional trials.

2.1 Author contact

We contacted the authors of significant papers identified from trials and review articles found in the search and asked for their knowledge of other studies, published or unpublished, relevant to the review. We also contacted other experts in the field for similar information.

Appendix 5. Searches by Cochrane Common Mental Disorders

1. Cochrane Common Mental Disorders Controlled Trials Register (CCMDCTR)

Cochrane Common Mental Disorders maintained a specialised register of randomised controlled trials, the CCMDCTR, until June 2016. This register contains over 40,000 reference records (reports of RCTs) for anxiety disorders, depression, bipolar disorder, eating disorders, self-harm and other mental disorders within the scope of this group. The CCMDCTR is a partially study-based register with > 50% of reference records tagged to c12,500 individually PICO-coded study records. Reports of trials for inclusion in the register were collated from (weekly) generic searches of MEDLINE (1950-), Embase (1974-) and PsycINFO (1967-), quarterly searches of the Cochrane Central Register of Controlled Trials (CENTRAL) and review-specific searches of additional databases. Reports of trials were also sourced from international trial registries, drug companies, the handsearching of key journals, conference proceedings and other (non-Cochrane) systematic reviews and meta-analyses. Details of CCMD's core search strategies (used to identify RCTs) can be found on the Group's website with an example of the core MEDLINE search displayed below.

CCMD’s core Ovid MEDLINE search strategy used to inform the Group’s specialised register:

A weekly search alert based on condition + RCT filter only

1. *[MeSH Headings]:*
eating disorders/ or anorexia nervosa/ or binge-eating disorder/ or bulimia nervosa/ or female athlete triad syndrome/ or pica/ or hyperphagia/ or bulimia/ or self-injurious behavior/ or self mutilation/ or suicide/ or suicidal ideation/ or suicide, attempted/ or mood disorders/ or affective disorders, psychotic/ or bipolar disorder/ or cyclothymic disorder/ or depressive disorder/ or depression, postpartum/ or depressive disorder, major/ or depressive disorder, treatment-resistant/ or dysthymic disorder/ or seasonal affective disorder/ or neurotic disorders/ or depression/ or adjustment disorders/ or exp antidepressive agents/ or anxiety disorders/ or agoraphobia/ or neurocirculatory asthenia/ or obsessive-compulsive disorder/ or obsessive hoarding/ or panic disorder/ or phobic disorders/ or stress disorders, traumatic/ or combat disorders/ or stress disorders, post-traumatic/ or stress disorders, traumatic, acute/ or anxiety/ or anxiety, castration/ or koro/ or anxiety, separation/ or panic/ or exp anti-anxiety agents/ or somatoform disorders/ or body dysmorphic disorders/ or conversion disorder/ or hypochondriasis/ or neurasthenia/ or hysteria/ or munchausen syndrome by proxy/ or munchausen syndrome/ or fatigue syndrome, chronic/ or obsessive behavior/ or compulsive behavior/ or behavior, addictive/ or impulse control disorders/ or firesetting behavior/ or gambling/ or trichotillomania/ or stress, psychological/ or burnout, professional/ or sexual dysfunctions, psychological/ or vaginismus/ or Anhedonia/ or Affective Symptoms/ or *Mental Disorders/

2. *[Title/ Author Keywords]:*
(eating disorder* or anorexia nervosa or bulimi* or binge eat* or (self adj (injur* or mutilat*)) or suicide* or suicidal or parasuicid* or mood disorder* or affective disorder* or bipolar i or bipolar ii or (bipolar and (affective or disorder*)) or mania or manic or cyclothymic* or depression or depressive or dysthymi* or neurotic or neurosis or adjustment disorder* or antidepress* or anxiety disorder* or agoraphobia or obsess* or compulsi* or panic or phobi* or ptsd or posttrauma* or post trauma* or combat or somatoform or somati#ation or medical* unexplained or body dysmorphi* or conversion disorder or hypochondria* or neurastheni* or hysteria or munchausen or chronic fatigue* or gambling or trichotillomania or vaginismus or anhedoni* or affective symptoms or mental disorder* or mental health).ti,kf.

3. *[RCT filter]:*
(controlled clinical trial.pt. or randomized controlled trial.pt. or (randomi#ed or randomi#ation).ab,ti. or randomly.ab. or (random* adj3 (administ* or allocat* or assign* or class* or control* or determine* or divide* or distribut* or expose* or fashion or number* or place* or recruit* or subsitut* or treat*)).ab. or placebo*.ab,ti. or drug therapy.fs. or trial.ab,ti. or groups.ab. or (control* adj3 (trial* or study or studies)).ab,ti. or ((singl* or doubl* or tripl* or trebl*) adj3 (blind* or mask* or dummy*)).mp. or clinical trial, phase ii/ or clinical trial, phase iii/ or clinical trial, phase iv/ or randomized controlled trial/ or pragmatic clinical trial/ or (quasi adj (experimental or random*)).ti,ab. or ((waitlist* or wait* list* or treatment as usual or TAU) adj3 (control or group)).ab.)

4. (1 and 2 and 3)

Records were screened for reports of RCTs within the scope of the Cochrane Common Mental Disorders Group. Secondary reports of RCTs were tagged to the appropriate study record. Similar weekly search alerts were also conducted on OVID EMBASE and PsycINFO, using relevant subject headings (controlled vocabularies) and search syntax, appropriate to each resource.

For this review, the CCMDCTR studies and references register was cross-searched using the following terms:

(collab and (bipolar or mania* or manic* or hypomani* or psychos* or psychotic or postpsychotic or post-psychotic or "rapid cycling" or schizoaffective on "mixed episode")) [all fields]*

****

2. Additional searches run by Cochrane Common Mental Disorders

As the CCMDCTR was only current to 6 June 2016, the Information Specialist ran complementary searches in the following databases:

Ovid databases (2014 to 2 June 2020): PsycINFO, Embase, MEDLINE(R) and Epub Ahead of Print, In-Process & Other Non-Indexed Citations and Daily

Search strategy:

---

1 BIPOLAR DISORDER/ or BIPOLAR DEPRESSION/ or BIPOLAR I DISORDER/ or BIPOLAR II DISORDER/ or BIPOLAR MANIA/ or CYCLOTHYMIA/ or "MIXED MANIA AND DEPRESSION"/ or RAPID CYCLING BIPOLAR DISORDER/

2 AFFECTIVE PSYCHOSIS/ or CYCLOTHYMIC PERSONALITY/ or MANIA/ or HYPOMANIA/

3 (bipolar adj3 (affective or depress* or disorder* or episode* or mood or psychosis or spectrum or state or states)).ti,ab,id,kf,kw.

4 (affective psycho* or mania or manic or hypermani* or hypomani* or rapid cycling).ti,ab,id,kf,kw.

5 or/1-4

6 (RCT or "at random" or (random* adj3 (administ* or allocat* or assign* or class* or cluster or crossover or cross-over or control* or determine* or divide* or division or distribut* or expose* or fashion or number* or place* or pragmatic or quasi or recruit* or split or subsitut* or treat*))).ti,ab,id,kf,kw.

7 (randomi#ed or randomi#ation or randomi#ing).ti,ab,id,kf,hw.

8 randomized controlled trial.pt,sh.

9 randomization.sh.

10 treatment effectiveness evaluation.sh.

11 controlled clinical trial.pt,sh.

12 (control* and (trial or study or group?) and (waitlist* or wait* list* or ((treatment or care) adj2 usual))).ti,ab,id,kf,kw.

13 ((single or double or triple or treble) adj2 (blind* or mask* or dummy)).ti,ab,id,kf,kw.

14 double blind procedure/

15 placebo.sh,ti. or (placebo adj3 (control or group?)).ti,ab,id,kf,kw.

16 or/6-15

17 5 and 16 (17352)

18 (2014* or 2015* or 2016* or 2017* or 2018* or 2019*).yr,dc,dd,dp,dt,em,ep,ez.

19 (17 and 18)

20 collaborat*.ti,ab,id,kf,kw. or collaborative care.hw.

21 (19 and 20)

22 remove duplicates from 21

****

Cochrane Central Register of Controlled Trials

Issue 6 of 12, March 2020

#1 MeSH descriptor: [Bipolar and Related Disorders] explode all trees

#2 (bipolar NEAR (affective or depress* or disorder* or episode* or mood or psychosis or spectrum or state or states)):ti,ab,kw

#3 ((affective next psycho*) or mania or manic or hypermani* or hypomani* or "rapid cycling"):ti,ab,kw

#4 cyclothymi*:ti,ab,kw

#5 (#1 or #2 or #3 or #4)

#6 collaborat*:ti,ab,kw

#7 (#5 and #6)

Limit to Trials

****

WHAT'S NEW

Date	Event	Description
7 May 2024	New citation required but conclusions have not changed	Seven additional studies included in the review.

Date	Event	Description
7 May 2024	New search has been performed	Searches updated.

HISTORY

Protocol first published: Issue 1, 2012
Review first published: Issue 11, 2013

CONTRIBUTIONS OF AUTHORS

SR: developed and led the writing of the protocol, original review and updated review, screened titles and abstracts, screened full-text articles, extracted data and wrote the final review.

CP: drafted and edited the protocol, contacted experts in the field and study authors, screened titles and abstracts, screened full-text articles, extracted data and wrote the final report.

BG: screened titles and abstracts, extracted data and rated risk of bias, completed the PRISMA flow diagram, contributed to the final version of the review.

CHM: screened titles and abstracts, extracted data, led the quality assessment: risk of bias and summary of findings. Contributed to the final report: risk of bias methods, risk of bias results, quality of evidence assessment and discussion, recommendations for practice.

DR: extracted data, rated risk of bias, completed the outcome measures of interest table, contributed to the final version of the review.

BJ: led the statistical elements of the study, including data extraction, manipulation, insertion into RevMan and analysis. Assisted in the interpretation of results and contributed to the writing of the final report.

JG: contributed to the lay summary and final version of the review.

HK: contributed to the outcome measures of interest table and the final version of the review.

LG: provided a clinical perspective, provided general advice on the review, secured funding for the review, agreed on the final version of the protocol and review.

BD: agreed on the final version of the protocol, provided a clinical perspective, provided general advice on the review, rated risk of bias, commented on and approved the final review.

PH: provided a social care perspective, provided general advice on the review, rated risk of bias, agreed on the final version of the review.

A number of the authors of this review were also involved in the core outcome set development (HP, BG, PH, SR, LG, JG).

DECLARATIONS OF INTEREST

SR was one of the researchers involved in one of the ongoing studies: PARTNERS2, NIHR-funded study.

CHM was one of the researchers involved in one of the ongoing studies: PARTNERS2, NIHR-funded study.

BG was one of the researchers involved in one of the ongoing studies: PARTNERS2, NIHR-funded study.

BJ was one of the researchers involved in one of the ongoing studies: PARTNERS2, NIHR-funded study.

DR was one of the researchers involved in one of the ongoing studies: PARTNERS2, NIHR-funded study.

HP was one of the researchers involved in one of the ongoing studies: PARTNERS2, NIHR-funded study.

JG was one of the researchers involved in one of the ongoing studies: PARTNERS2, NIHR-funded study.

MG was one of the researchers involved in one of the ongoing studies: PARTNERS2, NIHR-funded study.

LG was co-author on the 'Collaborative care for depression and anxiety problems in primary care' review (Archer 2012), and is one of the researchers involved in one of the ongoing studies: PARTNERS2, NIHR-funded study.

PH was one of the researchers involved in one of the ongoing studies: PARTNERS2, NIHR-funded study.

BD has no known conflicts of interest.

CP was one of the researchers involved in one of the ongoing studies: PARTNERS2, NIHR-funded study.

Those authors involved in the PARTNERS2 study will not be involved in the extraction or ratings of the risk of bias for that study.

SOURCES OF SUPPORT

Internal sources

- University of Bradford, UK

  Employs Siobhan Reilly as a Professor of Applied Dementia Research
- Lancaster University, UK

  Employed Siobhan Reilly as a senior lecturer
- National Institute for Health and Care Research (NIHR), UK

  Provided funding for Cochrane Schizophrenia
- Bangor University, UK

  Employs Peter Huxley as a Professor of Mental Health Research
- University of Manchester, UK

  Employs Dr Claire Planner as a Research Associate at the Centre for Primary Care and Health Services Research
- University of Birmingham, UK

  Employs Humera Plappert as Programme Manager

External sources

- National Institute for Health Research (NIHR) RP-PG-0611-20004, UK

  This review presents independent research funded by the National Institute for Health Research (NIHR) under its Programme Grants for Applied Research programme (grant reference no. RP-PG-0611-20004). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care.

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

Standard care

We defined standard care as a community or outpatient model of care not described as 'collaborative care' by the trialists. We made the post hoc decision that if study authors mentioned additional 'enhancements' to standard care, and these were minimal, then these could be included in the standard care comparison.

Reason for change

The reason for including studies in which the comparator included minimal enhancements to standard care was to ensure that 'standard care' reflected what might feasibly be delivered in healthcare settings.

Outcomes

We changed the outcomes from those reported in the original review (Appendix 3), before we extracted data from our included studies.

Reason for change

As this review has been funded as part of the Byng 2023 National Institute of Health Research (NIHR) grant, we were able to utilise a core outcome set for use in community-based bipolar trials to guide our choice of outcomes (Retzer 2020). We were also able to utilise an additional stakeholder consultation to select outcomes and measures. This was convened to capture the wider psychosis target population in Byng 2023 and the nature of the intervention. Quality of life (QoL) was selected as the most important outcome domain by all stakeholders. We added this as a primary outcome along with mental state and psychiatric hospital admissions.

The changes to the naming of outcomes are to maintain consistency with Cochrane Schizophrenia's classification of outcomes; the renaming and addition of outcomes does not alter the types of outcome we were/are interested in. On the contrary, the additional outcomes are clinically very relevant to people with severe mental illness.

Methods for handling unit of analysis issues in cluster-randomised controlled trials

In the original review, we stated that we would have assumed an intracluster correlation coefficient (ICC) of 0.1 if it was not possible to obtain the estimate from published papers or the authors. We have amended this assumption to 0.05.

We have also clarified the methodology used to account for clustering in meta-analyses.

Reason for change

A value of 0.05 is a more appropriate 'rule of thumb' estimate, particularly for trials in a primary care setting (Adams 2004).

Data synthesis methods

We have clarified our intention to use a random-effects model for meta-analysis, rather than fixed-effect, instead of assessing heterogeneity to determine which approach to undertake.

Reason for change

It was agreed that random-effects meta-analysis would be the appropriate default choice of analysis method in recognition of the differences between collaborative care interventions, the populations and the clinical settings across different studies.

Subgroup analysis and investigation of heterogeneity

We removed the pre-specification of undertaking a subgroup analysis on the basis of leaving the study early, but assuming all missing participants experienced a negative event.

Reason for change

It was agreed that this was not in fact a subgroup analysis, but rather an additional sensitivity analysis. It was agreed that the existing pre-specified sensitivity analysis to impute missing binary outcome data was sufficient, in line with what was actually done in the original review.

Measures of treatment effect - skewed data

We have softened the criteria for inclusion of skewed data, and amended the number of participants required to override issues with skew from a total of 200 to 30 per arm.

Reason for change

A sample size of 30 is the 'rule of thumb' for invocation of the central limit theorem, and so we believed that this was a more appropriate criterion than a total sample size of 200.

INDEX TERMS

Medical Subject Headings (MeSH)

Bias; Bipolar Disorder [therapy]; Community Mental Health Services; *Mental Disorders [therapy]; Patient Care Team; *Quality of Life; *Randomized Controlled Trials as Topic; *Schizophrenia [therapy]

MeSH check words

Adult; Humans