National Institute for Health and Care Excellence

Draft for consultation

Depression in adults

[A] Service delivery
Models and settings for delivery of services

NICE guideline CG90 (update)

Evidence reviews underpinning recommendations 1.15.7 to 1.15.14 in the NICE guideline

November 2021

Draft for consultation

These evidence reviews were developed by the National Guideline Alliance which is a part of the Royal College of Obstetricians and Gynaecologists

Disclaimer

The recommendations in this guideline represent the view of NICE, arrived at after careful consideration of the evidence available. When exercising their judgement, professionals are expected to take this guideline fully into account, alongside the individual needs, preferences and values of their patients or service users. The recommendations in this guideline are not mandatory and the guideline does not override the responsibility of healthcare professionals to make decisions appropriate to the circumstances of the individual patient, in consultation with the patient and/or their carer or guardian.

Local commissioners and/or providers have a responsibility to enable the guideline to be applied when individual health professionals and their patients or service users wish to use it. They should do so in the context of local and national priorities for funding and developing services, and in light of their duties to have due regard to the need to eliminate unlawful discrimination, to advance equality of opportunity and to reduce health inequalities. Nothing in this guideline should be interpreted in a way that would be inconsistent with compliance with those duties.

NICE guidelines cover health and care in England. Decisions on how they apply in other UK countries are made by ministers in the Welsh Government, Scottish Government, and Northern Ireland Executive. All NICE guidance is subject to regular review and may be updated or withdrawn.

ISBN:

Contents

Service delivery

This evidence report contains 2 reviews relating to service delivery

- Review question 1.1 For adults with depression, what are the relative benefits and harms associated with different models for the coordination and delivery of services?
- Review question 1.2 For adults with depression, what are the relative benefits and harms associated with different settings for the delivery of care?

Models of care

Review question

For adults with depression, what are the relative benefits and harms associated with different models for the coordination and delivery of services?

Introduction

To improve the treatment of adult depression, there has been a growing interest in the development of systems of care, with some influences from chronic disease management programmes seen in physical healthcare. Different systems of care have been developed and evaluated to see which may improve access to and efficacy of treatment and the efficiency and cost-efficiency of services. Models widely adopted in the UK include the stepped-care model, often associated with the Improving Access to Psychological Therapies (IAPT) programme. This seeks to offer people their least burdensome, most effective therapy first, usually a low intensity therapy (such as guided self-help) where appropriate, and then have their progress reviewed in conjunction with a therapist at regular intervals, with the option to step-up to higher intensity treatment, or step-across to another treatment of the same intensity, depending on progress. Alternatively, people can start on a higher intensity treatment where appropriate and step across or step down, depending on progress. Another model widely used is collaborative care, where a case manager or key worker is in regular contact with the person with depression to help coordinate their care, often involving liaison with the person's GP, specialists such as psychiatrists, and other psychological therapists if required. They may also support additional needs such employment. There may be overlap between these models of care where, for example, collaborative care may also include stepped care, and there are a number of other models including medication management, the attached professional (where a mental health professional has direct responsibility for the care of a person), and shared care, which may be delivered separately, or may be delivered within a broader place-based or community-based model of care.

The aim of this review is to identify benefits associated with different models of care for adults with depression.

Summary of the protocol

See Table 1 for a summary of the Population, Intervention, Comparison and Outcome (PICO) characteristics of this review.

Table 1: Summary of the protocol (PICO table)

• Population	• Adults with a diagnosis of depression according to DSM, ICD or similar criteria, or depressive symptoms as indicated by baseline depression scores on validated scales (and including those with subthreshold [just below threshold] depressive symptoms) For studies on relapse prevention: • Adults whose depression has responded to treatment (in full or partial remission) according to DSM, ICD or similar criteria, or indicated by below clinical threshold depression symptom scores on validated scales

• Intervention	Models for the coordination and delivery of services, including: • Collaborative care (simple and complex) • Stepped care • Medication management • Attached professional model • Care co-ordination • Integrated care pathways (including primary care liaison or shared care) • Measurement-based care
• Comparison	• Treatment as usual • Waitlist • Any other service delivery model
• Outcomes	**Critical** ○ Depression symptomatology (mean endpoint score or change in depression score from baseline) at 6 and 12 months ○ Response (usually defined as at least 50% improvement from the baseline score on a depression scale) at 6 and 12 months ○ Remission (usually defined as a score below clinical threshold on a depression scale) at 6 and 12 months ○ Relapse (number of people who returned to a depressive episode whilst in remission) at 6 and 12 months **Important** • Antidepressant use at 6 and 12 months • Discontinuation (due to any reason) at 6 and 12 months

DSM: Diagnostic and Statistical Manual of Mental Disorders; ICD: International Classification of Diseases.

For further details see the review protocol in appendix A.

Methods and process

This evidence review was developed using the methods and process described in Developing NICE guidelines: the manual. Methods specific to this review question are described in the review protocol in appendix A.

Declarations of interest were recorded according to NICE's 2014 conflicts of interest policy until 31 March 2018. From 1 April 2018, declarations of interest were recorded according to NICE's 2018 conflicts of interest policy. Those interests declared until April 2018 were reclassified according to NICE's 2018 conflicts of interest policy (see Register of Interests).

Clinical evidence

Included studies

56 randomised controlled trials (RCTs) were identified for inclusion in this review and the model of care described was identified.

For this review, a coding system for classifying the complexity and type of service delivery model was developed by the committee specifically for the purpose of this guideline. The service delivery model was rated on this 17-item coding system to generate an overall rating between 0-20 (see Figure 1). Service delivery models scoring at least 6 were considered a collaborative care intervention. Collaborative care interventions were further sub-divided into simple collaborative care (score of 6-

12) and complex collaborative care (score ≥13). Service delivery models scoring below 6 were classified as an alternative service delivery model (e.g. care coordination) or a stand-alone psychological intervention (e.g. self-help with support).

Figure 1: Coding system for service delivery models (Collaborative Care Component Score Method)

Item	Score
1. Active and integrated case recognition/identification* (Systematic identification- from a clinical database or screened positive for depression)	0 1
2. Collaborative assessment and plan included (Collaborative assessment with the patient)	0 1
3. Case Management (Case manager present- can include pharmacist for medication management)	0 1
4. Active liaison with primary care and other services (System set up for structured liaison/ regular meetings)	0 1
5. Case Manager has MH background (A prior mental health background, not just training in mental health)	0 1
6. Supervision provided for case manager	0 1
7. Senior MH professional consultation/involvement (Broad definition- just need to be available)	0 1
8. Psychoeducation delivered	0 1
9. Algorithm(s) used to determine care*	0 1
10. Integration with physical health care where necessary	0 1
11. Social/psychosocial interventions provided	0 1
12. Case manager delivers intervention	0 1
13. Medication management provided	0 1
14. Routine outcome monitoring (Scheduled, using a tool)	0 1
15. Psychological interventions provided None Low intensity High intensity	 0 1 2
16. Duration of programme contact ≤6 mths 7-12mths 1year plus	 0 1 2
17. Number of sessions (F-t-F and Telephone) ≤6 sessions 6 – 12 sessions 13 + sessions	 0 1 2
Total (maximum 20)	
***Including stepped care** **Rating** <5 – not collaborative care 6-12 – simple collaborative care 13+ – complex collaborative care	

39 RCTS were categorised as collaborative care (Aragones 2012; Araya 2003; Berghofer 2012; Bjorkelund 2018; Bosanquet 2017; Bruce 2004; Buszewicz 2016; Capoccia 2004; Chen 2015; Curth 2020; Dobscha 2006; Ell 2007; Finley 2003; Fortney 2007; Gensichen 2009; Gilbody 2017/Lewis 2017; Harter 2018; Holzel 2018; Huang 2018; Huijbregts 2013; Jarjoura 2004; Jeong 2013; Katon 1999; Katzelnick 2000; Landis 2007; Ludman 2007; Morriss 2016; Ng 2020; Oladeji 2015; Richards 2013/2016; Simon 2004 (CM); Simon 2004 (CM + psych); Simon 2006; Smit 2006; Swindle 2003; Unutzer 2002/Arean 2005; Wells 2000; Yeung 2010; Yeung 2016.

Of the 39 RCTs categorised as collaborative care, 6 were categorised as complex collaborative care (score ≥13) (Fortney 2007; Holzel 2018; Huijbregts 2013; Morris

2016; Simon 2004 CM+psych; Unutzer 2002/Arean 2005) and the remaining 33 RCTs were categorised as simple collaborative care (score of 6 to 12).

1 RCT was categorised as collaborative care for relapse prevention (Katon 2001).

5 RCTs were categorised as stepped care (Adewuya 2019; Callahan 1994; Gureje 2019; Knapstad 2020; Van Der Weele 2012).

1 RCT was categorised as stepped care for relapse prevention (Apil 2012).

5 RCTs were categorised as medication management (Akerblad 2003; Aljumah 2015; Rickles 2005; Rubio-Valera 2013a; Sirey 20105).

2 RCTs were categorised as care coordination (McMahon 2007; Salisbury 2016).

1 RCT was categorised as attached professional model (Bedoya 2014).

1 RCT was categorised as shared care (Banerjee 1996).

1 RCT was categorised as measurement-based care (Guo 2015).

The included studies are summarised in Table 2 to Table 10.

Planned subgroup analyses were outlined in the full review protocol (see appendix A) to include (where possible) for all reviews, the influence of the following subgroups: chronic depression; depression with coexisting personality disorder; psychotic depression; older adults; BME populations; men. For the collaborative care review, planned subgroup analyses included the following which were informed by the collaborative care component score method (in Figure 1): type of collaborative care; stepped care component; case manager background; psychological interventions delivered as part of the model of care; number of contacts/sessions/follow-up visits provided as part of the intervention. The committee were also interested in post-hoc subgroup analyses comparing outcomes by baseline severity. Subgroup analysis was considered for all critical outcomes with at least 2 studies in each subgroup. Subgroup analysis was only possible for the collaborative care dataset, where subgroup analyses were possible for older adults, BME groups, baseline severity, and the different collaborative care components outlined above.

See the literature search strategy in appendix B and study selection flow chart in appendix C.

Excluded studies

Studies not included in this review with reasons for their exclusions are provided in appendix K.

Summary of clinical studies included in the evidence review

Comparison 1. Collaborative care (simple or complex) versus standard care/enhanced standard care

Collaborative care is defined as a multi-professional approach to care for people with depression, involving a structured management plan, scheduled follow-ups and enhanced inter-professional communication. Collaborative care may also include elements of other models, such as stepped care, psychoeducation, psychological interventions or medication management.

Summaries of the studies included for the comparison of collaborative care versus standard care or enhanced standard care are presented in Table 2.

Subgroup analysis of the collaborative care dataset was possible for:

- Older adults (mean age ≥ 60 years) versus younger adults (mean age <60 years) for the following outcomes: depression symptomatology at 6 months; depression symptomatology at 12 months; response at 6 months; response at 12 months; remission at 6 months; remission at 12 months
- BME groups, comparing studies where less than 50% of the population were from a BME group with studies where 50-100% of the population were from a BME group, for the following outcome: remission at 6 months
- Baseline severity, comparing studies where the mean depression scale score indicated less severe depression (corresponding to the traditional categories of mild and subthreshold) with more severe depression (corresponding to the traditional categories of moderate and severe depression), for the following outcomes: depression symptomatology at 6 months; depression symptomatology at 12 months; response at 6 months; response at 12 months; remission at 6 months; remission at 12 months
- Type of collaborative care, simple versus complex, for the following outcomes: depression symptomatology at 6 months; depression symptomatology at 12 months; response at 6 months; response at 12 months; remission at 6 months; remission at 12 months
- Stepped care component, comparing interventions that included a stepped care component, interventions that included only a medication algorithm, and interventions with no stepped care component or algorithm, for the following outcomes: depression symptomatology at 6 months; depression symptomatology at 12 months; response at 6 months; response at 12 months; remission at 6 months; remission at 12 months
- Case manager background, comparing studies where the case manager had a prior mental health background and studies where the case manager did not have a prior mental health background, for the following outcomes: depression symptomatology at 6 months; depression symptomatology at 12 months
- Inclusion of psychological interventions, comparing studies where psychological interventions were delivered as part of the model of care with studies where psychological interventions were not part of the service delivery model, for the following outcomes: depression symptomatology at 6 months; depression symptomatology at 12 months; response at 6 months; response at 12 months; remission at 6 months; remission at 12 months
- Number of contacts provided as part of the intervention, comparing less than 13 contacts with 13 or more contacts, for the following outcomes: depression symptomatology at 6 months; depression symptomatology at 12 months; response at 6 months; response at 12 months; remission at 6 months; remission at 12 months

Table 2: Summary of included studies for Comparison 1: Collaborative care (simple or complex) versus standard care/enhanced standard care.

Study	Population	Intervention	Comparison	Comments
Aragones 2012 RCT Spain	N=360 Baseline severity: More severe	Simple collaborative care	Standard care	Duration of programme contact (in months): NR

Study	Population	Intervention	Comparison	Comments
	Mean age (years): 47.6 Sex (% female): 79 Ethnicity (% BME): NR	Collaborative care component score: 9		Outcomes: • Depression symptomatology at 6 months • Depression symptomatology at 12 months • Response at 6 months • Response at 12 months • Remission at 6 months • Remission at 12 months • Antidepressant use at 6 months • Antidepressant use at 12 months • Discontinuation at 6 months • Discontinuation at 12 months
Araya 2003 RCT Chile	N=240 Baseline severity: More severe Mean age (years): 42.6 Sex (% female): 100 Ethnicity (% BME): NR	Simple collaborative care Collaborative care component score: 7	Standard care	Duration of programme contact (in months): 3 Outcomes: • Response at 6 months • Remission at 6 months • Antidepressant use at 6 months • Discontinuation at 6 months
Berghofer 2012 RCT Germany	N=63 Baseline severity: More severe Mean age (years): 49.7 Sex (% female): 73 Ethnicity (% BME): NR	Simple collaborative care Collaborative care component score: 10	Standard care	Duration of programme contact (in months): 6 Outcomes: • Response at 6 months • Response at 12 months
Bjorkelund 2018	N=385	Simple collaborative care	Standard care	Duration of programme

Study	Population	Intervention	Comparison	Comments
RCT Sweden	Baseline severity: Less severe Mean age (years): 41.2 Sex (% female): 71 Ethnicity (% BME): NR	Collaborative care component score: 9		contact (in months): 3 Outcomes: • Remission at 6 months • Antidepressant use at 6 months • Discontinuation at 6 months
Bosanquet 2017 RCT UK	N=485 Baseline severity: Less severe Mean age (years): 72.2 Sex (% female): 62 Ethnicity (% BME): 2	Simple collaborative care Collaborative care component score: 8	Enhanced standard care	Duration of programme contact (in months): 2 Outcomes: • Depression symptomatology at 12 months • Antidepressant use at 12 months • Discontinuation at 12 months
Bruce 2004 RCT US	N=598 Baseline severity: More severe Mean age (years): NR (>60) Sex (% female): 72 Ethnicity (% BME): 28	Simple collaborative care Collaborative care component score: 11.5	Enhanced standard care	Duration of programme contact (in months): 12 Outcomes: • Depression symptomatology at 12 months • Response at 12 months • Remission at 12 months • Antidepressant use at 12 months • Discontinuation at 12 months
Buszewicz 2016 RCT UK	N=558 Baseline severity: More severe Mean age (years): 48.4	Simple collaborative care Collaborative care component score: 11	Standard care	Duration of programme contact (in months): 24 Outcomes: • Depression symptomatology at 6 months

Study	Population	Intervention	Comparison	Comments
	Sex (% female): 75 Ethnicity (% BME): 12			• Depression symptomatology at 12 months • Discontinuation at 6 months • Discontinuation at 12 months
Capoccia 2004 RCT US	N=74 Baseline severity: NR Mean age (years): 38.7 Sex (% female): 77 Ethnicity (% BME): 22	Simple collaborative care Collaborative care component score: 8	Standard care	Duration of programme contact (in months): 12 Outcomes: • Antidepressant use at 12 months • Discontinuation at 12 months
Chen 2015 RCT China	N=326 Baseline severity: More severe Mean age (years): NR (>60) Sex (% female): 63 Ethnicity (% BME): NR	Simple collaborative care Collaborative care component score: 12	Enhanced standard care	Duration of programme contact (in months): 4 Outcomes: • Depression symptomatology at 6 months • Depression symptomatology at 12 months • Response at 6 months • Response at 12 months • Remission at 6 months • Remission at 12 months • Discontinuation at 6 months • Discontinuation at 12 months
Curth 2020 RCT Denmark	N=325 Baseline severity: Less severe Mean age (years): 39	Simple collaborative care Collaborative care component score: 11	Standard care	Duration of programme contact (in months): 4 Outcomes: • Depression symptomatology at 6 months

Study	Population	Intervention	Comparison	Comments
	Sex (% female): 67 Ethnicity (% BME): NR			• Discontinuation at 6 months
Dobscha 2006 RCT US	N=375 Baseline severity: Less severe Mean age (years): 56.8 Sex (% female): 7 Ethnicity (% BME): 3	Simple collaborative care Collaborative care component score: 9	Enhanced standard care	Duration of programme contact (in months): 12 Outcomes: • Antidepressant use at 12 months • Discontinuation at 12 months
Ell 2007 RCT US	N=311 Baseline severity: NR Mean age (years): NR (>60) Sex (% female): 72 Ethnicity (% BME): 27	Simple collaborative care Collaborative care component score: 10.5	Enhanced standard care	Duration of programme contact (in months): 12 Outcomes: • Response at 12 months • Remission at 12 months • Antidepressant use at 12 months • Discontinuation at 12 months
Finley 2003 RCT US	N=125 Baseline severity: NR Mean age (years): 54.3 Sex (% female): 85 Ethnicity (% BME): NR	Simple collaborative care Collaborative care component score: 6.5	Standard care	Duration of programme contact (in months): 6 Outcomes: • Antidepressant use at 6 months • Discontinuation at 6 months
Fortney 2007 RCT US	N=395 Baseline severity: More severe Mean age (years): 59.2	Complex collaborative care Collaborative care component score: 13	Enhanced standard care	Duration of programme contact (in months): 12 Outcomes:

Study	Population	Intervention	Comparison	Comments
	Sex (% female): 8 Ethnicity (% BME): 25			• Antidepressant use at 12 months • Discontinuation at 12 months
Gensichen 2009 RCT Germany	N=626 Baseline severity: More severe Mean age (years): 51.1 Sex (% female): 76 Ethnicity (% BME): NR	Simple collaborative care Collaborative care component score: 7	Standard care	Duration of programme contact (in months): 12 Outcomes: • Depression symptomatology at 12 months • Response at 12 months • Remission at 12 months • Antidepressant use at 12 months • Discontinuation at 12 months
Gilbody 2017/Lewis 2017 RCT UK	N=705 Baseline severity: Less severe Mean age (years): 77.3 Sex (% female): 58 Ethnicity (% BME): 1	Simple collaborative care Collaborative care component score: 10	Standard care	Duration of programme contact (in months): 2 Outcomes: • Depression symptomatology at 12 months • Antidepressant use at 12 months • Discontinuation at 12 months
Harter 2018 RCT Germany	N=779 Baseline severity: Less severe Mean age (years): 42.9 Sex (% female): 73 Ethnicity (% BME): NR	Simple collaborative care Collaborative care component score: 11	Standard care	Duration of programme contact (in months): NR Outcomes: • Depression symptomatology at 6 months • Depression symptomatology at 12 months • Response at 12 months • Remission at 12 months

Study	Population	Intervention	Comparison	Comments
				• Discontinuation at 6 months • Discontinuation at 12 months
Holzel 2018 RCT Germany	N=248 Baseline severity: Less severe Mean age (years): 71.4 Sex (% female): 77 Ethnicity (% BME): NR	Complex collaborative care Collaborative care component score: 14	Standard care	Duration of programme contact (in months): 12 Outcomes: • Depression symptomatology at 12 months • Response at 12 months • Remission at 12 months
Huang 2018 RCT China	N=280 Baseline severity: More severe Mean age (years): 47.4 Sex (% female): 85 Ethnicity (% BME): 100	Simple collaborative care Collaborative care component score: 10	Standard care	Duration of programme contact (in months): 6 Outcomes: • Depression symptomatology at 6 months • Discontinuation at 6 months
Huijbregts 2013 RCT Netherlands	N=150 Baseline severity: Less severe Mean age (years): 48.7 Sex (% female): 73 Ethnicity (% BME): 29	Complex collaborative care Collaborative care component score: 13	Standard care	Duration of programme contact (in months): 12 Outcomes: • Response at 6 months • Response at 12 months • Remission at 6 months • Remission at 12 months • Discontinuation at 6 months • Discontinuation at 12 months
Jarjoura 2004 RCT	N=61 Baseline severity: More severe	Simple collaborative care	Enhanced standard care	Duration of programme contact (in months): NR

Study	Population	Intervention	Comparison	Comments
US	Mean age (years): 45.5 Sex (% female): 69 Ethnicity (% BME): NR	Collaborative care component score: 6		Outcomes: • Antidepressant use at 12 months
Jeong 2013 RCT Korea	N=57 Baseline severity: More severe Mean age (years): NR (>60) Sex (% female): 58 Ethnicity (% BME): NR	Simple collaborative care Collaborative care component score: 7	Standard care	Duration of programme contact (in months): 6 Outcomes: • Remission at 6 months • Antidepressant use at 6 months • Discontinuation at 6 months
Katon 1999 RCT US	N=228 Baseline severity: NR Mean age (years): 47 Sex (% female): 75 Ethnicity (% BME): 20	Simple collaborative care Collaborative care component score: 6	Standard care	Duration of programme contact (in months): 3 Outcomes: • Remission at 6 months • Antidepressant use at 6 months
Katzelnick 2000 RCT US	N=407 Baseline severity: More severe Mean age (years): 45.5 Sex (% female): 77 Ethnicity (% BME): 21	Simple collaborative care Collaborative care component score: 9	Standard care	Duration of programme contact (in months): 7 Outcomes: • Response at 12 months • Remission at 12 months • Discontinuation at 12 months
Landis 2007 RCT	N=45 Baseline severity: More severe	Simple collaborative care	Enhanced standard care	Duration of programme contact (in months): 6

Study	Population	Intervention	Comparison	Comments
US	Mean age (years): 39.7 Sex (% female): 96 Ethnicity (% BME): 28	Collaborative care component score: 9		Outcome: • Depression symptomatology at 6 months
Ludman 2007 RCT US	N=52 Baseline severity: NR Mean age (years): 50.3 Sex (% female): 69 Ethnicity (% BME): 13	Simple collaborative care Collaborative care component score: 9	Standard care	Duration of programme contact (in months): NR Outcomes: • Remission at 12 months • Antidepressant use at 12 months • Discontinuation at 12 months
Morris 2016 RCT UK	N=187 Baseline severity: More severe Mean age (years): 46.5 Sex (% female): 61 Ethnicity (% BME): NR	Complex collaborative care Collaborative care component score: 14	Standard care	Duration of programme contact (in months): 12 Outcomes: • Depression symptomatology at 12 months • Response at 12 months • Remission at 12 months • Discontinuation at 12 months
Ng 2020 RCT Singapore	N=274 Baseline severity: Less severe Mean age (years): 73.5 Sex (% female): 56 Ethnicity (% BME): NR	Simple collaborative care Collaborative care component score: 9	Standard care	Duration of programme contact (in months): 6 Outcomes: • Depression symptomatology at 6 months • Depression symptomatology at 12 months • Response at 6 months • Response at 12 months

Study	Population	Intervention	Comparison	Comments
				• Remission at 6 months • Remission at 12 months • Discontinuation at 6 months • Discontinuation at 12 months
Oladeji 2015 RCT Nigeria	N=234 Baseline severity: Less severe Mean age (years): 43.2 Sex (% female): 80 Ethnicity (% BME): NR	Simple collaborative care Collaborative care component score: 12	Enhanced standard care	Duration of programme contact (in months): 6 Outcomes: • Depression symptomatology at 6 months • Discontinuation at 6 months
Richards 2013/2016 RCT UK	N=581 Baseline severity: More severe Mean age (years): 44.8 Sex (% female): 72 Ethnicity (% BME): 15	Simple collaborative care Collaborative care component score: 12	Standard care	Duration of programme contact (in months): 3 Outcomes: • Depression symptomatology at 12 months • Response at 12 months • Remission at 12 months • Antidepressant use at 12 months • Discontinuation at 12 months
Simon 2004 (CM) RCT US	N=402 Baseline severity: Less severe Mean age (years): 44.5 Sex (% female): 75 Ethnicity (% BME): 20	Simple collaborative care Collaborative care component score: 9	Standard care	Duration of programme contact (in months): 5 Outcomes: • Antidepressant use at 6 months • Discontinuation at 6 months

Study	Population	Intervention	Comparison	Comments
Simon 2004 (CM + psych) RCT US	N=393 Baseline severity: Less severe Mean age (years): 44.4 Sex (% female): 76 Ethnicity (% BME): 23	Complex collaborative care Collaborative care component score:13	Standard care	Duration of programme contact (in months): 5 Outcomes: • Antidepressant use at 6 months • Discontinuation at 6 months
Simon 2006 RCT US	N=207 Baseline severity: Less severe Mean age (years): 43 Sex (% female): 65 Ethnicity (% BME): 11	Simple collaborative care Collaborative care component score: 9	Standard care	Duration of programme contact (in months): 3 Outcomes: • Antidepressant use at 6 months • Discontinuation at 6 months
Smit 2006 RCT Netherlands	N=267 Baseline severity: Less severe Mean age (years): 42.8 Sex (% female): 64 Ethnicity (% BME): NR	Simple collaborative care Collaborative care component score: 9.5	Standard care	Duration of programme contact (in months): 6 Outcomes: • Remission at 6 months • Antidepressant use at 6 months • Discontinuation at 6 months
Swindle 2003 RCT US	N=268 Baseline severity: Less severe Mean age (years): 56.3 Sex (% female): 3 Ethnicity (% BME): 15	Simple collaborative care Collaborative care component score: 8	Enhanced standard care	Duration of programme contact (in months): 2 Outcomes: • Depression symptomatology at 12 months • Discontinuation at 12 months

Study	Population	Intervention	Comparison	Comments
Unutzer 2002/Arean 2005 RCT US	N=1901 Baseline severity: NR Mean age (years): 71.2 Sex (% female): 65 Ethnicity (% BME): 23	Complex collaborative care Collaborative care component score: 14.5	Standard care	Duration of programme contact (in months): 12 Outcomes: • Antidepressant use at 6 months • Antidepressant use at 12 months • Discontinuation at 6 months • Discontinuation at 12 months
Wells 2000 RCT US	N=1356 Baseline severity: NR Mean age (years): 43.7 Sex (% female): 71 Ethnicity (% BME): 43	Simple collaborative care Collaborative care component score: 8.5	Standard care	Duration of programme contact (in months): 12 Outcomes: • Remission at 6 months • Remission at 12 months • Discontinuation at 6 months • Discontinuation at 12 months
Yeung 2010 RCT US	N=100 Baseline severity: More severe Mean age (years): 49 Sex (% female): 68 Ethnicity (% BME): 100	Simple collaborative care Collaborative care component score: 8	Standard care	Duration of programme contact (in months): 6 Outcomes: • Response at 6 months • Remission at 6 months
Yeung 2016 RCT US	N=190 Baseline severity: More severe Mean age (years): 50 Sex (% female): 63	Simple collaborative care Collaborative care component score: 8	Enhanced standard care	Duration of programme contact (in months): 6 Outcomes: • Response at 6 months • Remission at 6 months

Study	Population	Intervention	Comparison	Comments
	Ethnicity (% BME): 100			

BME: black minority ethnic; N: number; NR: not reported; RCT: randomised controlled trial

There were no statistically significant subgroup differences between older and younger adults for the comparison collaborative care versus standard care or enhanced standard care on: depression symptomatology at 6 months (Test for subgroup differences: Chi² = 0.74, df = 1, p = 0.39); depression symptomatology at 12 months (Test for subgroup differences: Chi² = 1.01, df = 1, p = 0.32); response at 6 months (Test for subgroup differences: Chi² = 1.34, df = 1, p = 0.25); response at 12 months (Test for subgroup differences: Chi² = 1.34, df = 1, p = 0.25); remission at 6 months (Test for subgroup differences: Chi² = 1.20, df = 1, p = 0.27); remission at 12 months (Test for subgroup differences: Chi² = 0.52, df = 1, p = 0.47). Although there was a consistent trend for larger benefits for older adults, for example, for younger adults the effect estimate for collaborative care versus standard care/enhanced standard care on depression symptomatology at 12 months was SMD -0.25 [-0.33, -0.17] (K=7; N=2865) relative to older adults where the effect estimate was SMD -0.47 [-0.88, -0.05] (K=6; N=2543).

There was no statistically significant subgroup differences between studies with a predominantly white population and studies where the majority of participants were from BME groups for the comparison collaborative care versus standard care or enhanced standard care on: remission at 6 months (Test for subgroup differences: Chi² = 0.79, df = 1, p = 0.38).

There was a statistically significant subgroup difference between studies where the mean depression scale score indicated less severe depression and studies where participants had more severe depression, for the comparison collaborative care versus standard care or enhanced standard care, on remission at 6 months (Test for subgroup differences: Chi² = 8.54, df = 1, p = 0.003). Larger benefits were observed for more severe depression populations (RR 2.31 [1.59, 3.36]; K=6; N=1273), relative to less severe depression (RR 1.21 [0.97, 1.51]; K=4; N=1076). However, this pattern was not consistent across outcomes, and subgroup differences were not statistically significant for: depression symptomatology at 6 months (Test for subgroup differences: Chi² = 0.07, df = 1, p = 0.79); depression symptomatology at 12 months (Test for subgroup differences: Chi² = 0.47, df = 1, p = 0.49); response at 6 months (Test for subgroup differences: Chi² = 0.49, df = 1, p = 0.49); response at 12 months (Test for subgroup differences: Chi² = 1.05, df = 1, p = 0.31); remission at 12 months (Test for subgroup differences: Chi² = 0.32, df = 1, p = 0.57).

There were no statistically significant subgroup differences between simple and complex collaborative care for the comparison collaborative care versus standard care or enhanced standard care on: depression symptomatology at 12 months (Test for subgroup differences: Chi² = 0.69, df = 1, p = 0.41); response at 12 months (Test for subgroup differences: Chi² = 0.17, df = 1, p = 0.68); remission at 12 months (Test for subgroup differences: Chi² = 2.79, df = 1, p = 0.09).

There were no statistically significant subgroup differences between interventions that included a stepped care component, interventions that included only a medication algorithm, and interventions with no stepped care component or algorithm for the comparison collaborative care versus standard care or enhanced standard care on: depression symptomatology at 6 months (Test for subgroup differences: Chi² = 2.33, df = 2, p = 0.31); depression symptomatology at 12 months (Test for subgroup differences: Chi² = 5.44, df = 2, p = 0.07); response at 6 months (Test for

subgroup differences: Chi² = 2.07, df = 2, p = 0.36); response at 12 months (Test for subgroup differences: Chi² = 3.96, df = 2, p = 0.14); remission at 6 months (Test for subgroup differences: Chi² = 4.02, df = 2, p = 0.13); remission at 12 months (Test for subgroup differences: Chi² = 4.30, df = 2, p = 0.12). Although there was a consistent trend for larger benefits for interventions that included a stepped care component, for example, for interventions that included a stepped care component the effect estimate for collaborative care versus standard care/enhanced standard care on depression symptomatology at 12 months was SMD -0.61 [-1.10, -0.11] (K=5; N=1717) relative to interventions that included a medication algorithm-only where the effect estimate was SMD -0.10 [-0.23, 0.03] (K=3; N=1081), or no stepped care component where the effect estimate was SMD -0.25 [-0.39, -0.12] (K=5; N=2610).

There were no statistically significant subgroup differences between interventions where the case manager had a prior mental health background and interventions where the case manager did not have a prior mental health background, for the comparison collaborative care versus standard care or enhanced standard care on: depression symptomatology at 6 months (Test for subgroup differences: Chi² = 0.18, df = 1, p = 0.67); depression symptomatology at 12 months (Test for subgroup differences: Chi² = 1.02, df = 1, p = 0.31).

There were no statistically significant subgroup differences between studies where psychological interventions were delivered as part of the model of care and studies where psychological interventions were not part of the service delivery model, for the comparison collaborative care versus standard care or enhanced standard care on: depression symptomatology at 6 months (Test for subgroup differences: Chi² = 0.00, df = 1, p = 0.98); depression symptomatology at 12 months (Test for subgroup differences: Chi² = 0.01, df = 1, p = 0.91); response at 6 months (Test for subgroup differences: Chi² = 0.01, df = 1, p = 0.94); response at 12 months (Test for subgroup differences: Chi² = 0.14, df = 1, p = 0.71); remission at 6 months (Test for subgroup differences: Chi² = 1.12, df = 1, p = 0.29); remission at 12 months (Test for subgroup differences: Chi² = 0.09, df = 1, p = 0.76).

There was a statistically significant subgroup difference between interventions with fewer than 13 contacts and interventions with 13 or more contacts, for the comparison collaborative care versus standard care or enhanced standard care, on remission at 12 months (Test for subgroup differences: Chi² = 4.23, df = 1, p = 0.04). Interventions with 13+ contacts showed larger benefits (RR 1.97 [1.33, 2.91]; K=8; N=3188) than interventions with <13 contacts (RR 1.25 [1.06, 1.48]; K=6; N=3067). Although heterogeneity remained fairly high within (as well as between) subgroups, with I^2 values of 79% for interventions with 13+ contacts and 56% for interventions with <13 contacts. There was a trend for larger benefits associated with more contacts across other outcomes, although subgroup differences were not statistically significant for: depression symptomatology at 6 months (Test for subgroup differences: Chi² = 0.35, df = 1, p = 0.55); depression symptomatology at 12 months (Test for subgroup differences: Chi² = 1.13, df = 1, p = 0.29); response at 6 months (Test for subgroup differences: Chi² = 0.02, df = 1, p = 0.88); response at 12 months (Test for subgroup differences: Chi² = 0.41, df = 1, p = 0.52); remission at 6 months (Test for subgroup differences: Chi² = 0.84, df = 1, p = 0.36).

Comparison 2. Collaborative care versus standard care for relapse prevention

Collaborative care can also be used for those in full or partial remission from depression, particularly those at higher risk of relapse, as a strategy to keep well.

A summary of the study included for the comparison of collaborative care versus standard care for relapse prevention is presented in Table 3.

Table 3: Summary of included studies for Comparison 2: Collaborative care versus standard care for relapse prevention

Study	Population	Intervention	Comparison	Comments
Katon 2001 RCT US	N=386 Baseline severity: Recovered but at high risk of relapse (<4 DSM-IV MDD symptoms and a history of ≥3 episodes of MDD or dysthymia or 4 residual depressive symptoms but mean SCL-20 depression score < 1.0 and a history of MDD/dysthymia) Mean age (years): 46 Sex (% female): 74 Ethnicity (% BME): 10	Simple collaborative care Collaborative care component score: 9	Standard care	Duration of programme contact (in months): 12 Outcomes: • Relapse at 12 months • Antidepressant use at 6 months • Antidepressant use at 12 months • Discontinuation at 12 months

BME: black minority ethnic; DSM: diagnostic statistical manual; MDD: major depressive disorder; N: number; NR: not reported; RCT: randomised controlled trial; SCL-20: symptom checklist

Comparison 3. Stepped care versus standard care/enhanced standard care

Stepped care provides the most effective yet least burdensome treatment for people with depression first, but if a person does not benefit from an initial intervention they are 'stepped up' to a more complex intervention. Typically, stepped care starts by providing a low intensity intervention, but in patient-specific stepped care a higher intensity intervention may be commenced if, for example, a person is very ill or suicidal and a low intensity intervention would not be appropriate.

Summaries of the studies included for the comparison of stepped care versus standard care or enhanced standard care are presented in Table 4.

Table 4: Summary of included studies for Comparison 3: Stepped care versus standard care/enhanced standard care

Study	Population	Intervention	Comparison	Comments
Adewuya 2019 RCT	N=907	Stepped care	Enhanced standard care	Duration of programme contact (in months): NR

Study	Population	Intervention	Comparison	Comments
Nigeria	Baseline severity: More severe Mean age (years): 34.3 Sex (% female): 53 Ethnicity (% BME): NR	Step 1: Psychoeducation; Step 2: Problem solving or amitriptyline (if contraindicated, fluoxetine) monotherapy Step 3: Combination from step 2 Step 4: Support and supervision from mental health team		Outcomes: • Remission at 6 months • Remission at 12 months • Discontinuation at 6 months • Discontinuation at 12 months
Callahan 1994 RCT US	N=175 Baseline severity: More severe Mean age (years): 65.3 Sex (% female): 76 Ethnicity (% BME): 51	Stepped care Step 1: Nortriptyline or desipramine Step 2: Fluoxetine Step 3: Psychiatry consultation	Standard care	Duration of programme contact (in months): 3 Outcomes: • Remission at 6 months • Antidepressant use at 6 months • Discontinuation at 6 months
Gureje 2019 RCT Nigeria	N=1178 Baseline severity: Less severe Mean age (years): 47.3 Sex (% female): 83 Ethnicity (% BME): NR	Stepped care Step 1: Psychological intervention (BA & problem solving) for mild, combined psychological intervention and amitriptyline for moderate and severe Step 2: Additional therapy sessions or psychological intervention + AD Step 3: Cases discussed with a psychiatrist	Enhanced standard care	Duration of programme contact (in months): NR Outcomes: • Depression symptomatology at 6 months • Depression symptomatology at 12 months • Remission at 12 months • Discontinuation at 6 months • Discontinuation at 12 months
Knapstad 2020 RCT Norway	N=774 Baseline severity: Less severe	Stepped care Norwegian version of IAPT - low-intensity (guided self-help, psychoeducation	Standard care	Duration of programme contact (in months): NR Outcomes:

Study	Population	Intervention	Comparison	Comments
	Mean age (years): 34.8 Sex (% female): 67 Ethnicity (% BME): NR	al courses) and high-intensity (individual treatment)		• Depression symptomatology at 6 months • Discontinuation at 6 months
Van Der Weele 2012 RCT Netherlands	N=239 Baseline severity: Less severe Mean age (years): NR (median 80) Sex (% female): NR Ethnicity (% BME): NR	Stepped care Step 1: Individual counselling concerning treatment needs and motivation Step 2: Coping with depression course Step 3: Referral back to GP	Standard care	Duration of programme contact (in months): NR Outcomes: • Depression symptomatology at 6 months • Depression symptomatology at 12 months • Response at 6 months • Response at 12 months • Discontinuation at 6 months • Discontinuation at 12 months

AD: antidepressant; BA: behavioural activation; BME: black minority ethnic; IAPT: improving access to psychological therapies service; N: number; NR: not reported; RCT: randomised controlled trial

Comparison 4. Stepped care versus standard care for relapse prevention

Stepped care can also be used for those in full or partial remission from depression, as a strategy to keep well.

A summary of the study included for the comparison of stepped care versus standard care for relapse prevention is presented in Table 5.

Table 5: Summary of included studies for Comparison 4: Stepped care versus standard care for relapse prevention

Study	Population	Intervention	Comparison	Comments
Apil 2012 RCT Netherlands	N=136 Baseline severity: Less severe Mean age (years): 65.6 Sex (% female): 72	Stepped care relapse prevention programme Step 1: Watchful waiting for 6 weeks (no intervention offered)	Standard care	Duration of programme contact (in months): 12 Outcomes: • Relapse at 12 months • Antidepressant use at 12 months

Study	Population	Intervention	Comparison	Comments
	Ethnicity (% BME): NR	Step 2: Cognitive bibliotherapy for 6 weeks Step 3: Individual coping with depression course (12x weekly sessions of 45 mins) Step 4: Indicated treatment (referred to a physician/psychotherapist and treatment could consist of any intervention considered necessary)		• Discontinuation at 12 months

BME: black minority ethnic; N: number; NR: not reported; RCT: randomised controlled trial

Comparison 5. Pure medication management versus standard care

Medication management can be a component of a broader service delivery model (for example, as part of collaborative care) or as a stand-alone intervention (pure medication management). Medication management is an intervention to ensure medication taken for depression has the greatest opportunity to be effective, by working with people to increase understanding of their medication, promote adherence, ensure adequate therapeutic levels are obtained, and allow people to discuss their medicine use and so reduce unnecessary discontinuation of medication due to lack of benefits or side effects.

Summaries of the studies included for the comparison of pure medication management versus standard care are presented in Table 6.

Table 6: Summary of included studies for Comparison 5: Pure medication management versus standard care

Study	Population	Intervention	Comparison	Comments
Akerblad 2003 RCT Sweden	N=665 Baseline severity: More severe Mean age (years): 48.5 Sex (% female): 72 Ethnicity (% BME): NR	Pure medication management Therapeutic drug monitoring (TDM). All patients were treated with sertraline. Plasma levels of sertraline and desmethylsertraline were determined at weeks 4 and 12 and reported	Standard care	Duration of programme contact (in months): 6 Outcomes: • Antidepressant use at 6 months • Discontinuation at 6 months

Study	Population	Intervention	Comparison	Comments
		back to the GP for continued discussion with the patients. Intervention included monitoring for side effects		
Aljumah 2015 RCT Saudi Arabia	N=239 Baseline severity: More severe Mean age (years): NR Sex (% female): 55 Ethnicity (% BME): NR	Pure medication management Pharmacist intervention involving assessing patients' beliefs and knowledge about antidepressants and distribution of a decision aid to patients	Standard care	Duration of programme contact (in months): 3 Outcomes: • Depression symptomatology at 6 months • Antidepressant use at 6 months • Discontinuation at 6 months
Rickles 2005 RCT US	N=63 Baseline severity: Less severe Mean age (years): 38 Sex (% female): 84 Ethnicity (% BME): 8	Pure medication management Pharmacist-guided education and monitoring (PGEM) included assessing patient's antidepressant knowledge and beliefs, adverse effects and other concerns, treatment goals, and how the medication was being used, reviewing of current adherence, and any new adverse effects and concerns	Standard care	Duration of programme contact (in months): 3 Outcomes: • Antidepressant use at 6 months • Discontinuation at 6 months
Rubio-Valera 2013a RCT Spain	N=179 Baseline severity: Less severe Mean age (years): 46.6 Sex (% female): 75	Pure medication management Community pharmacist intervention included provision of an educational intervention aimed at	Standard care	Duration of programme contact (in months): 6 Outcomes: • Depression symptomatology at 6 months

Study	Population	Intervention	Comparison	Comments
	Ethnicity (% BME): NR	improving patients' knowledge of antidepressants and awareness of the importance of adherence, and monitoring of patient progress (improvement, appearance of side effects, or queries)		• Antidepressant use at 6 months • Discontinuation at 6 months
Sirey 2010 RCT US	N=70 Baseline severity: More severe Mean age (years): 76 Sex (% female): 77 Ethnicity (% BME): 29	Pure medication management Treatment Initiation and Participation (TIP) programme, included reviewing symptoms and antidepressant therapy regimen and conducting a barriers assessment, defining personal treatment goal, provision of education about depression and antidepressants, discussing barriers to adherence, creating an adherence strategy, and encouraging the patient to talk directly with the primary care physician about treatment	Standard care	Duration of programme contact (in months): 2 Outcomes: • Response at 6 months • Discontinuation at 6 months

BME: black minority ethnic; N: number; NR: not reported; RCT: randomised controlled trial

Comparison 6. Care coordination versus standard care/enhanced standard care

Care coordination can be a component of a broader service delivery model (for example, as part of collaborative care) or as a stand-alone intervention. Care coordination (also known as case management) is a system where an individual healthcare professional takes responsibility for the coordination of the care of a person with depression, but is not necessarily directly involved in the provision of any intervention; it may also involve the coordination of follow-up.

Summaries of the studies included for the comparison of care coordination versus standard care or enhanced standard care are presented in Table 7.

Table 7: Summary of included studies for Comparison 6: Care coordination versus standard care/enhanced standard care

Study	Population	Intervention	Comparison	Comments
McMahon 2007 RCT UK	N=62 Baseline severity: More severe Mean age (years): NR Sex (% female): NR Ethnicity (% BME): NR	Care coordination Case management from graduate primary care mental health workers + TAU from GP. Minimal supportive counselling provided and could recommend increase in antidepressant dosage to GP	Enhanced standard care	Duration of programme contact (in months): 4 Outcomes: • Depression symptomatology at 6 months • Discontinuation at 6 months
Salisbury 2016 RCT UK	N=609 Baseline severity: More severe Mean age (years): 49.6 Sex (% female): 68 Ethnicity (% BME): 3	Care coordination Telephone calls with health adviser, includes information signposting, access to computerized CBT (CCBT) and support in use of CCBT, minimal supportive counselling and could recommend increase in antidepressant dosage to GP	Standard care	Duration of programme contact (in months): 10 Outcomes: • Depression symptomatology at 12 months • Remission at 12 months • Discontinuation at 12 months

BME: black minority ethnic; N: number; NR: not reported; RCT: randomised controlled trial; TAU: treatment as usual

Comparison 7. Attached professional model versus enhanced standard care

In this model a mental health professional has direct responsibility for the care of a person (usually in primary care) focusing on the primary treatment of the depression. The coordination of care remains with the GP/primary care team. Contact with the attached professional is usually limited to treatment and involves little or no follow-up beyond that determined by the specific intervention offered (for example, booster sessions in CBT).

A summary of the study included for the comparison of attached professional model versus enhanced standard care is presented in Table 8.

Table 8: Summary of included studies for Comparison 7: Attached professional model versus enhanced standard care

Study	Population	Intervention	Comparison	Comments
Bedoya 2014 RCT US	N=120 Baseline severity: More severe Mean age (years): 42.4 Sex (% female): 69 Ethnicity (% BME): 100	Attached professional model Culturally focused psychiatric (CFP) consultation service. Study clinicians (psychologists or psychiatrists) provided a psychiatric assessment, psychoeducation, cognitive-behavioural tools, and tailored treatment recommendations ; primary care providers were provided a consultation summary	Enhanced standard care	Duration of programme contact (in months): 0.5 Outcomes: • Depression symptomatology at 6 months • Discontinuation at 6 months

BME: black minority ethnic; N: number; NR: not reported; RCT: randomised controlled trial

Comparison 8. Shared care versus standard care

Shared care is the involvement of a multidisciplinary team who work together to plan and deliver individualised care for people with depression. The team will usually include involvement from both primary care and specialist services.

A summary of the study included for the comparison of shared care versus standard care is presented in Table 9.

Table 9: Summary of included studies for Comparison 8: Shared care versus standard care

Study	Population	Intervention	Comparison	Comments
Banerjee 1996 RCT UK	N=69 Baseline severity: More severe Mean age (years): 80.7 Sex (% female): 83 Ethnicity (% BME): NR	Shared care Individual package of care formulated by the community psychogeriatric team in their catchment area and implemented by a researcher working as a member of that team. Each case was presented at	Standard care	Duration of programme contact (in months): 6 Outcomes: • Depression symptomatology at 6 months • Remission at 6 months • Antidepressant use at 6 months

Study	Population	Intervention	Comparison	Comments
		a multidisciplinary team meeting which included CPNs, OTs, senior and junior medical staff, a social worker, and a psychologist. A management plan was formulated by the team for each person on an individual basis and could include any combination of antidepressants, psychological interventions and social interventions. A psychiatrist acted as each person's keyworker		• Discontinuation at 6 months

BME: black minority ethnic; CPN: community psychiatric nurse; N: number; NR: not reported; OT: occupational therapist; RCT: randomised controlled trial

Comparison 9. Measurement-based care versus standard care

Measurement-based care is similar to stepped care with defined levels of treatment but progression to different steps or alternative treatments is guided by the use of a predefined algorithm that utilises objective measures of efficacy.

A summary of the study included for the comparison of measurement-based care versus standard care is presented in Table 10.

Table 10: Summary of included studies for Comparison 9: Measurement-based care versus standard care

Study	Population	Intervention	Comparison	Comments
Guo 2015 RCT China	N=120 Baseline severity: More severe Mean age (years): 41.1 Sex (% female): 64 Ethnicity (% BME): NR	Measurement-based care Guideline- and rating scale-based decisions. The treating psychiatrists made treatment decisions about starting dosages, dose adjustments and medication changes of paroxetine (20–60mg/day) or	Standard care	Duration of programme contact (in months): 3 Outcomes: • Depression symptomatology at 6 months • Response at 6 months • Remission at 6 months • Discontinuation at 6 months

Study	Population	Intervention	Comparison	Comments
		mirtazapine (15–45mg/day), on the basis of ratings on QIDS-SR and the Frequency, Intensity, and Burden of Side Effects Rating scale		

BME: black minority ethnic; N: number; NR: not reported; QIDS-SR: quick inventory of depressive symptomatology-self report; RCT: randomised controlled trial; SR: self-report

See the full evidence tables in appendix D and the forest plots in appendix E.

Quality assessment of clinical outcomes included in the evidence review

See the clinical evidence profiles in appendix F.

Economic evidence

Included studies

A single economic search was undertaken for all topics included in the scope of this guideline. See the literature search strategy in appendix B and economic study selection flow chart in appendix G. Details on the hierarchy of inclusion criteria for economic studies are provided in supplement 1 (methods supplement).

The systematic search of the literature identified 12 studies (in 13 publications) on the cost effectiveness of different models for the coordination and delivery of services for adults with depression.

There were 3 UK studies that assessed simple collaborative care (Bosanquet 2017; Green 2014; Lewis 2017) and 1 UK study that assessed complex collaborative care (Morriss 2016). Following the hierarchy of inclusion criteria regarding country settings, 1 Dutch (Goorden 2015) and 1 German (Grochtdreis 2019) studies assessing the cost effectiveness of complex collaborative care were also included in the review. In addition, the search identified 1 US study assessing the cost effectiveness of simple collaborative care in relapse prevention (Simon 2002) and given that the study focused on a different population that was not covered by UK studies or other studies ranking higher on the hierarchy of inclusion criteria, this study was also included in the review.

One UK study assessed the cost effectiveness of stepped care (Mukuria 2013). Following the hierarchy of inclusion criteria regarding country settings, 2 Dutch (van der Weele 2012, Meeuwissen 2019) and 1 Canadian economic study (Health Quality Ontario 2019) were also included in the economic review of stepped care.

No UK studies on the cost effectiveness of medication management for adults with depression were identified. Following the hierarchy of inclusion criteria regarding country settings, 1 Spanish study (Rubio-Valera 2013) was included in the review.

No UK studies on the cost effectiveness of shared care for adults with depression were identified. Following the hierarchy of inclusion criteria regarding country settings, 1 US study (Wiley-Exley 2009) was included in the review.

No studies assessing the cost effectiveness of care coordination, the attached professional model, or measurement-based care for adults with depression were identified.

Economic evidence tables are provided in appendix H. Economic evidence profiles are shown in appendix I.

Excluded studies

A list of excluded economic and utility studies, with reasons for exclusion, is provided in supplement 3 - Economic evidence included & excluded studies.

Summary of studies included in the economic evidence review

Simple collaborative care

Bosanquet 2017 performed a cost-utility analysis alongside a RCT (Bosanquet 2017; N=485; at 18 months n=344; cost data available for n=447) that compared simple collaborative care in addition to usual primary care versus primary care alone for older adults who screened positive for major depression in the UK. The perspective of the analysis was the NHS and personal social services (PSS). Healthcare costs consisted exclusively of intervention and primary care costs. National unit costs were used. The outcome measure was the QALY estimated based on SF-6D ratings (UK tariff). The duration of the analysis was 18 months.

Simple collaborative care was found to be more effective and more costly than usual (primary) care alone, with an ICER of £28,765/QALY (uplifted to 2020 prices). The probability of simple collaborative care being cost-effective at the NICE lower (£20,000/QALY) and upper (£30,000/QALY) cost effectiveness threshold was 0.39 and 0.55, respectively. When only participants who engaged with 5 or more sessions of collaborative care were included in the analysis, the ICER fell at £10,922/QALY (in 2020 prices). The study is directly applicable to the UK context but is characterised by potentially serious limitations, mainly the inclusion of intervention and primary care costs only.

Green 2014 conducted a cost-utility analysis alongside a RCT (Richards 2013; N=581, efficacy data available for n=466; resource use data available for n=447) that compared simple collaborative care in addition to usual primary care versus primary care alone for adults with depression in the UK. The perspective of the analysis was the NHS and personal social services (PSS); a broader perspective that included informal care costs and service user expenses was considered in a sensitivity analysis. Healthcare costs consisted of intervention costs, staff time (such as GP, mental health nurse, mental health worker, psychiatrist, psychologist), other outpatient and inpatient care, day care, walk-in-centre, and A&E. National unit costs were used. The outcome measure was the QALY estimated based on EQ-5D ratings (UK tariff); QALY estimates based on the SF-6D (UK tariff) were used in sensitivity analysis. The duration of the analysis was 12 months.

Simple collaborative care was found to be more effective and more costly than usual (primary) care alone, with an Incremental Cost Effectiveness Ratio (ICER) of £16,361/QALY (in 2020 prices). The probability of simple collaborative care being cost-effective at the NICE lower (£20,000/QALY) and upper (£30,000/QALY) cost

effectiveness threshold was 0.58 and 0.65, respectively. Results were robust to multiple imputation of missing data, use of SF-6D utility values, and use of alternative collaborative care costs. The study is directly applicable to the UK context and is characterised by minor limitations.

Lewis 2017 also conducted a cost-utility analysis alongside a RCT (Gilbody 2017; N=705, complete data for economic analysis n=448) that compared simple collaborative care in addition to usual primary care versus primary care alone for older adults who screened positive for subthreshold depression in the UK. The perspective of the analysis was the NHS and PSS. Healthcare costs consisted exclusively of intervention and primary care costs. National unit costs were used. The outcome measure was the QALY estimated based on EQ-5D ratings (UK tariff). The duration of the analysis was 12 months.

Simple collaborative care was found to be more effective and more costly than usual (primary) care alone, with an ICER of £10,653/QALY (in 2020 prices). The probability of simple collaborative care being cost-effective at the NICE lower (£20,000/QALY) and upper (£30,000/QALY) cost effectiveness threshold was 0.92 and 0.97, respectively. Accounting for the true observed case manager contact rate (rather than the expected contact rate that was used in the base-case analysis), the ICER fell at £3,681/QALY (in 2020 prices). The study is directly applicable to the UK context but is characterised by potentially serious limitations, mainly the high attrition that was markedly greater in the collaborative care arm, and the consideration of intervention and primary care costs only.

Simple collaborative care for relapse prevention

Simon 2002 assessed the cost effectiveness of simple collaborative care versus usual care alongside a RCT (Katon 2001; N=386, 82% completed all follow-up assessments and 98% remained enrolled throughout the follow-up period) that compared simple collaborative care with treatment as usual for adults with a history of either recurrent major depression or dysthymia that had recovered from a depressive episode following antidepressant treatment in primary care in the US. The study, which adopted a 3rd party payer perspective, considered costs of medication, staff time, as well as costs of any inpatient and outpatient services for mental health or general medical care; local prices were used. The outcome measure was the number of depression-free days, defined as days with a Hopkins Symptoms Checklist (HSCL) depression score ≤ 0.5; days with a HSCL score above 0.5 but < 2 were considered as being 50% depression free. The time horizon of the analysis was 12 months.

Simple collaborative care was found to be more effective and more costly than usual care, with an ICER of $1 per depression-free day (95%CI -$134 to $344, 1998 US$), which translates to £1.2 per depression-free day in 2020 prices. The study is only partially applicable to the NICE decision-making context as it was conducted in the US and does not use the QALY as the outcome measure, which requires judgement on whether the additional benefit is worth the extra cost. It is also characterised by potentially serious limitations, resulting mainly from the fact that analyses of clinical data included only those completing all blinded follow-up assessments; cost analyses included only those remaining enrolled throughout the follow-up period. However, participation in follow-up interviews was significantly greater in the intervention group than in usual care, introducing a possibility of bias.

Complex collaborative care

Morriss 2016 assessed the cost-utility of complex collaborative care versus usual secondary mental health care in the UK. The economic analysis was carried out alongside a RCT (Morriss 2016; N=187; 84% completed at 6 months, 72% at 12 months and 59% at 18 months). Complex collaborating care comprised secondary outpatient specialist depression services offering tailored integrated pharmacological and psychological (CBT, MBCT and compassion focused therapy, as appropriate) treatment within a collaborative care approach for 12-15 months. The analysis adopted a NHS and PSS perspective. Healthcare costs consisted of intervention costs, primary care (GP surgery and home attendances), inpatient and outpatient (psychiatric or other) care, other staff time (practice - district - community psychiatric nurse, psychotherapist), A&E attendances, and medication. National unit costs were used. The outcome measure was the QALY estimated based on EQ-5D ratings (UK tariff). The duration of the analysis was 18 months.

Complex collaborative care was more effective and more costly than usual secondary mental health care, with an ICER of £47,690/QALY (in 2020 prices). Controlling for baseline differences and cluster effects, the probability of complex collaborative care being cost-effective exceeded 50% at a cost effectiveness threshold of £45,500/QALY, which is well above the NICE cost effectiveness threshold of £30,000/QALY. The study is directly applicable to the UK context and is characterised by minor limitations.

Goorden 2015 assessed the cost effectiveness of complex collaborative care versus treatment as usual in a Dutch primary care setting. The study, which was conducted alongside a RCT (Huijbregts 2013), adopted a healthcare perspective, with productivity losses being reported separately. Healthcare costs consisted of intervention costs (care manager), other staff time (such as GP, mental health care professional, psychologist/psychiatrist, social worker, occupational therapist), self-help groups, day care, psychiatric inpatient care and medication. National unit costs were used. The outcome measure was the QALY estimated based on EQ-5D ratings (Dutch tariff). The time horizon was 12 months.

Complex collaborative care was found to be more effective and more costly than treatment as usual, with an ICER of €53,717/QALY in 2013 prices (£54,087 in 2020 prices), and a probability of being cost-effective of 0.20 and 0.70 at a cost effectiveness threshold of £20,100 and £80,500/QALY, respectively. The study is partially applicable to the UK context and is characterised by potentially serious limitations, mainly by the fact that, although the RCT included 150 participants, 93 identified by screening and 47 by GP referral, the cost-utility analysis was based only on the 93 participants that were identified by screening.

Grochtdreis 2019 assessed the cost effectiveness of complex collaborative care versus treatment as usual for adults aged ≥ 60 years with moderate depressive symptoms in Germany. The study was undertaken alongside a cluster RCT (Hölzel 2018; N=246 from 71 clusters) and adopted a healthcare perspective, with informal care costs being reported separately. Healthcare costs consisted of outpatient physician and non-physician services (e.g. physiotherapy, occupational therapy, massage), inpatient care, rehabilitation, formal nursing care (professional nurse or housekeeper), informal nursing care (family or friends), medication and medical devices. National unit costs were used. The outcome measure was the number of depression-free days (DFDs), determined by a PHQ-9 score <5. QALYs were also used as a secondary outcome, estimated based on EQ-5D ratings (UK tariff). The time horizon was 12 months.

Complex collaborative care was found to be more effective and more costly than treatment as usual, with an ICER of €26.07/DFD or €55,800/QALY in 2013 prices (£26/DFD or £56,184/QALY in 2020 prices), and a probability of being cost-effective of 0.95 at a cost-effetiveness threshold of £204/DFD and 0.45 at a cost-effectiveness threshold of £50,400/QALY. The study is partially applicable to the UK context and is characterised by minor limitations.

Stepped care

Mukuria 2013 assessed the cost-utility of stepped care for people with depression or anxiety in the UK, as reflected in the Improving Access to Psychological Therapies (IAPT) service, in addition to treatment as usual, versus treatment as usual alone; the latter comprised GP care, primary care counselling and referral to secondary mental health services. The study was conducted alongside a prospective cohort study with matched sites (N=403), and more than 95% of the study sample included people with a primary diagnosis of depression. The analysis adopted a NHS and social services perspective; productivity losses were assessed separately. Healthcare costs consisted of intervention (staff time, training, equipment, facilities and overheads), other mental healthcare (psychiatrist, psychologist, community psychiatric nurse, etc.), primary and secondary care, and social care; medication costs were not considered. Unit costs were based on IAPT data and national sources. The outcome measures of the analysis were the proportion of people with a reliable and clinically significant (RCS) improvement on the PHQ-9 and the QALY based on SF-6D ratings (UK tariff); QALYs estimated based on predicted EQ-5D ratings (UK tariff), estimated from SF-6D using an empirical mapping function, were used in sensitivity analysis. The duration of the analysis was 8 months.

IAPT added to treatment as usual was more costly and more effective than treatment as usual alone, with ICERs of £11,234 per additional participant with RCS improvement, £35,106/QALY using the SF-6D and £20,059/QALY using predicted EQ-5D scores (figures uplifted to 2020 prices). The probability of IAPT being cost-effective using SF-6D QALYs was less than 0.40 at a cost effectiveness threshold of £30,000/QALY; using QALYs estimated based on predicted EQ-5D ratings the probability of IAPT being cost-effective was 0.38 and 0.53 at cost effectiveness thresholds of £20,000 and £30,000/QALY, respectively. Using national unit costs instead of IAPT financial data resulted in an ICER of £4,522 per additional participant achieving RCS improvement and £14,132/QALY using SF-6D ratings (2020 prices). It is noted that NICE recommends use of EQ-5D for the estimation of QALYs in adults.

The study is directly applicable to the UK context and is characterised by potentially serious limitations such as its short time horizon, its study design, the sensitivity of results to unit costs of IAPT, the low response rate at recruitment (403 out of 3,391, 11.9%); and the fact that the IAPT service was assessed over the first 2 years of establishment, therefore costs associated with learning effects were likely.

Meeuwissen 2019 assessed the cost-utility of stepped care versus treatment as usual for adults with mild, moderate or severe major depression in the Netherlands. The study employed decision-economic modelling and adopted a healthcare perspective. Efficacy data were taken from a literature review, resource use data were based on published literature and national unit costs were likely used. Healthcare costs consisted of health professional time (GP, psychologist, psychiatrist, etc.), antidepressants, telephone consultation, self-help book or information leaflet, group therapy, crisis intervention, inpatient care, day care, homecare, and other out-patient care. The outcome measure of the analysis was the

QALY, following transformation of the effect size into a utility increment. The time horizon of the analysis was 5 years.

Stepped care was found to dominate treatment as usual in adults with mild depression; it was more effective and costlier in adults with moderate/severe depression, with an ICER of €3,166/QALY (in 2017 prices) or £3,159/QALY (in 2020 prices). The probability of stepped care being dominant was 0.67 in adults with mild depression and 0.33 in adults with moderate/severe depression. The probability of stepped care being cost-effective at a cost-effectiveness threshold of approximately £20,000/QALY was more than 0.95 in both populations.

The study is partially applicable to the UK NHS context, as it was conducted in the Netherlands and the method of estimation of QALYs was not the one recommended by NICE, and is characterised by minor limitations.

Van der Weele 2012 assessed the cost-utility of stepped care versus treatment as usual for adults aged ≥ 75 years with depressive symptoms in the Netherlands. The study was undertaken alongside a cluster RCT (van der Weele 2012; N=239; completers n=194) and adopted a healthcare perspective, with service user and informal care costs being reported separately. Healthcare costs consisted of intervention costs (individual consultation, course sessions, course instructors, room rental, refreshments, course materials), staff time (psychiatrist, psychologist, GP, physiotherapist), medication, hospitalisation (psychiatric & general), hospital day care, specialist care, paramedical care, service user costs (time & travel) and informal care. National unit costs were used. The outcome measures were the MADRS change score, and the QALY based on EQ-5D and SF-6D ratings (UK tariff). The time horizon was 12 months.

Stepped care was found to be dominated by treatment as usual in adults aged 75-79 years, when QALYs were derived by EQ-5D ratings, and to dominate treatment as usual in adults aged ≥80 years. The study is partially applicable to the UK NHS context, as it was conducted in the Netherlands, and is characterised by potentially serious limitations, mainly because there was no estimation of the uncertainty in the cost effectiveness results.

Health Quality Ontario 2019 assessed the cost-utility of stepped care for people with mild to moderate depression in Canada based on decision-economic modelling. Two separate analyses were conducted: one analysis compared stepped care comprising computerised CBT (cCBT) with support followed by individual or group CBT with treatment as usual; the other analysis assessed stepped care comprising cCBT without support followed by cCBT with support versus individual CBT, group CBT and treatment as usual in people who are likely to drop out of treatment. The perspective of the analysis was that of healthcare and long term care. Efficacy data were taken from a systematic literature review, resource use data were based on published literature and expert opinion and national unit costs were used. Costs consisted of intervention costs (health professional time, training and supervision, equipment), assessment, medication, follow-up care with GP, and psychiatrist time. The outcome measure of the analysis was the QALY; utility data were derived from a literature review; various scales were used for the quality of life ratings. The time horizon was lifetime for the first analysis and 1 year for the second analysis (the one on adults with mild to moderate depression at risk of dropping out).

Stepped care was found to dominate treatment as usual in adults with mild to moderate depression (first analysis); results were robust to change in efficacy, dropout rates, utilities, medication costs, time horizon. The probability of stepped care where cCBT was followed by individual CBT was 0.60 at a cost effectiveness

threshold of about £30,000/QALY. Regarding adults with mild to moderate depression at risk of dropping out, stepped care was the most cost-effective option assessed: it was more effective and costlier than treatment as usual, with an ICER of Can$19,454/QALY (in 2018 prices) or £11,666/QALY (in 2019 prices). Individual and group CBT were less cost-effective than stepped care at a cost-effectiveness threshold of about £30,000/QALY, as their ICERs versus stepped care reached or exceeded £40,000/QALY. The probability of stepped care being cost-effective among individual CBT, group CBT and treatment as usual was 0.48 at this threshold.

The study is partially applicable to the UK NHS context, as it was conducted in Canada and the method of estimation of QALYs was not the one recommended by NICE, and is characterised by minor limitations.

Medication management

Rubio-Valera 2013 conducted an economic evaluation of medication management versus treatment as usual for adults with depression treated in primary care. The study was undertaken alongside a RCT (Rubio-Valera 2013, N=179; 71% completed at 6 months; n=151 received intervention as allocated). The study adopted a healthcare and a societal perspective; costs included intervention, publicly funded healthcare services (GP, nurse, psychologist, psychiatrist, other specialists, social worker, hospital emergency visits, hospital stay, diagnostic tests, medication), privately funded healthcare services (psychiatrist, psychologist, medical specialist, GP), and absenteeism from paid labour. Regional unit prices were used. The study used 3 outcome measures: adherence to antidepressant treatment measured using electronic pharmacy records; remission of depressive symptoms defined as a reduction in the Patient Health Questionnaire 9-item (PHQ-9) of at least 50%; and the QALY based on EQ-5D ratings and the Spanish tariff. The time horizon of the analysis was 6 months.

Under the healthcare perspective, medication management was more expensive than treatment is usual. It was also more effective in terms of adherence to antidepressant treatment and the QALYs gained. The respective ICERs were €962 per extra adherent service user and €3,592/QALY (2009 prices; translating into figures of £935 per extra adherent service user and £3,495/QALY in 2020 prices). However, when remission was used as an outcome, medication management was dominated by treatment as usual, as it was more expensive and less effective. The probability of medication management being cost-effective was 0.71 and 0.76 for WTP £5,800/adherent service user and £29,000/QALY, respectively (2020 prices). Using remission as an outcome, the maximum probability of medication management being cost-effective was only 0.46, irrespective of the cost effectiveness threshold used. Results were robust to different scenarios such as a per protocol or complete case analysis, use of different diagnostic criteria for depression, changes in intervention costs or different methodology used for estimating indirect costs. The study is partially applicable to the UK decision-making context, as it was conducted in Spain. The findings of the study are inconsistent across the outcome measures used (i.e. the study appears to be cost-effective using the QALY, but cost-ineffective using remission as measure of outcome). The study was characterised by potentially serious limitations, mainly its contradictory results, its short time horizon and the use of regional unit costs.

Shared care

Wiley-Exley 2009 evaluated the cost effectiveness of integrated (shared) care compared with primary care with a referral system to specialist care for older adults with depression in the US. The study, which was conducted alongside a RCT

(N=840), analysed 4 different combinations of populations and settings: people major and minor depression (full sample) in the Veteran Affairs (VA) setting (n=365), full sample outside VA (n=475); people with major depression within VA (n=214), and people with major depression outside VA (n=302). The analysis adopted a healthcare and service users' and carers' perspective and included intervention costs, outpatient and inpatient care, nursing home, rehabilitation, emergency room, medication, service users' and caregivers' time and travel costs. National unit costs were used. The study included various measures of outcome, such as the CES-D score; the number of depression-free days derived from CES-D; the number of QALYs estimated based on depression-free days, using utility weights of health=1, depression=0.59; the number of QALYs estimated based on SF-36, using preferences for matched vignettes created following cluster analysis of SF-12 mental and physical component scores, elicited by US service users with depression using SG. Only results for the latter are reported here (full results of the study are provided in the study's evidence table in appendix H). The time horizon of the analysis was 6 months.

Integrated care was found to dominate usual primary care in the full sample (major and minor depression), VA setting. It was more costly and more effective than usual primary care regarding the full sample outside VA setting and major depression sample in the VA setting, with ICERs of £91,674/QALY and £56,799/QALY, respectively (2020 prices). It was less effective and less costly than usual primary care in the major depression sample, outside the VA setting, with an ICER of £76,861/QALY (saving per QALY lost).

The probability of integrated care being cost-effective was more than 0.70 for any cost effectiveness threshold only in the full sample and VA setting. The probability of integrated care being cost-effective was low at levels of willingness to pay that corresponded to NICE cost effectiveness thresholds. The study is partially applicable to the UK as it was conducted in the US, and is characterised by potentially serious limitations, including the short time horizon and the contradictory results across sub-analyses.

Economic model

No economic modelling was undertaken for this review because the committee agreed that other topics were higher priorities for economic evaluation.

Evidence statements

Clinical evidence statements

Comparison 1. Collaborative care (simple or complex) versus standard care/enhanced standard care

Critical outcomes

Depression symptomatology

- Very low quality evidence from 9 RCTs (N=2791) shows a statistically significant but not clinically important benefit of collaborative care, relative to standard care or enhanced standard care, on depression symptomatology at 6 months for adults with depression.
- Very low quality evidence from 13 RCTs (N=5408) shows a statistically significant but not clinically important benefit of collaborative care, relative to

standard care or enhanced standard care, on depression symptomatology at 12 months for adults with depression.

Response

- Low quality evidence from 8 RCTs (N=1703) shows a clinically important and statistically significant benefit of collaborative care, relative to standard care or enhanced standard care, on the rate of response at 6 months for adults with depression.
- Low quality evidence from 13 RCTs (N=4910) shows a clinically important and statistically significant benefit of collaborative care, relative to standard care or enhanced standard care, on the rate of response at 12 months for adults with depression.

Remission

- Low quality evidence from 12 RCTs (N=3933) shows a clinically important and statistically significant benefit of collaborative care, relative to standard care or enhanced standard care, on the rate of remission at 6 months for adults with depression.
- Very low quality evidence from 14 RCTs (N=6255) shows a clinically important and statistically significant benefit of collaborative care, relative to standard care or enhanced standard care, on the rate of remission at 12 months for adults with depression.

Important outcomes

Antidepressant use

- Very low quality evidence from 11 RCTs (N=4022) shows neither a clinically important nor statistically significant effect of collaborative care, relative to standard care or enhanced standard care, on antidepressant use at 6 months for adults with depression.
- Very low quality evidence from 13 RCTs (N=5666) shows a statistically significant but not clinically important benefit of collaborative care, relative to standard care or enhanced standard care, on antidepressant use at 12 months for adults with depression.

Discontinuation

- Low quality evidence from 19 RCTs (N=8305) shows neither a clinically important nor statistically significant effect of collaborative care, relative to standard care or enhanced standard care, on discontinuation at 6 months for adults with depression.
- Moderate quality evidence from 22 RCTs (N=10,916) shows neither a clinically important nor statistically significant effect of collaborative care, relative to standard care or enhanced standard care, on discontinuation at 12 months for adults with depression

Subgroup analysis 1a: Simple versus complex collaborative care

- Subgroup analysis of collaborative care compared to standard care or enhanced standard care for adults with depression, shows no statistically significant difference between simple and complex collaborative care, on any of the outcomes for which sub-analysis was possible: depression

symptomatology at 12 months; response at 12 months; remission at 12 months.

Subgroup analysis 1b: Older adults

- Subgroup analysis of collaborative care compared to standard care or enhanced standard care for adults with depression, shows no statistically significant difference between older adults and younger adults, on any of the outcomes for which sub-analysis was possible: depression symptomatology at 6 months; depression symptomatology at 12 months; response at 6 months; response at 12 months; remission at 6 months; remission at 12 months. Although there was a consistent trend for larger benefits for older adults.

Subgroup analysis 1c: BME groups

- Subgroup analysis of collaborative care compared to standard care or enhanced standard care for adults with depression, shows no statistically significant difference between studies with a predominantly white population and studies where the majority of participants were from BME groups, on the one outcome for which sub-analysis was possible: remission at 6 months.

Subgroup analysis 1d: Stepped care component

- Subgroup analysis of collaborative care compared to standard care or enhanced standard care for adults with depression, shows no statistically significant difference between interventions that included a stepped care component, interventions that included only a medication algorithm, and interventions with no stepped care component or algorithm, on any of the outcomes for which sub-analysis was possible: depression symptomatology at 6 months; depression symptomatology at 12 months; response at 6 months; response at 12 months; remission at 6 months; remission at 12 months. Although there was a consistent trend for larger benefits for interventions that included a stepped care component.

Subgroup analysis 1e: Case manager background

- Subgroup analysis of collaborative care compared to standard care or enhanced standard care for adults with depression, shows no statistically significant difference between interventions where the case manager had a prior mental health background and interventions where the case manager did not have a prior mental health background, on any of the outcomes for which sub-analysis was possible: depression symptomatology at 6 months; depression symptomatology at 12 months.

Subgroup analysis 1f: Psychological intervention

- Subgroup analysis of collaborative care compared to standard care or enhanced standard care for adults with depression, shows no statistically significant difference between studies where psychological interventions were delivered as part of the model of care and studies where psychological interventions were not part of the service delivery model, on any of the outcomes for which sub-analysis was possible: depression symptomatology at 6 months; depression symptomatology at 12 months; response at 6 months; response at 12 months; remission at 6 months; remission at 12 months.

Subgroup analysis 1g: Number of contacts

- Subgroup analysis of collaborative care compared to standard care or enhanced standard care for adults with depression, showed a statistically significant subgroup difference between interventions with fewer than 13 contacts and interventions with 13 or more contacts on the rate of remission at 12 months, with larger benefits associated with 13+ contacts. Although heterogeneity remained fairly high within (as well as between) subgroups. There was a trend for larger benefits associated with more contacts across other outcomes, although subgroup differences were not statistically significant for: depression symptomatology at 6 months; depression symptomatology at 12 months; response at 6 months; response at 12 months; remission at 6 months.

Subgroup analysis 1h: Baseline severity

- Subgroup analysis of collaborative care compared to standard care or enhanced standard care for adults with depression, showed a statistically significant subgroup difference between studies where the mean depression scale score indicated less severe depression and studies where participants had more severe depression on the rate of remission at 6 months, with larger benefits associated with more severe depression. However, this pattern was not consistent across outcomes, and subgroup differences were not statistically significant for: depression symptomatology at 6 months; depression symptomatology at 12 months; response at 6 months; response at 12 months; remission at 12 months.

Comparison 2: Collaborative care versus standard care for relapse prevention

Critical outcomes

Relapse

- Very low quality evidence from 1 RCT (N=386) shows neither a clinically important nor statistically significant effect of collaborative care, relative to standard care, on the rate of relapse for adults with remitted depression.

Important outcomes

Antidepressant use

- Low quality evidence from 1 RCT (N=386) shows a statistically significant but not clinically important benefit of collaborative care, relative to standard care, on antidepressant use at 6 months for adults with remitted depression.
- Low quality evidence from 1 RCT (N=386) shows a clinically important and statistically significant benefit of collaborative care, relative to standard care, on antidepressant use at 12 months for adults with remitted depression.

Discontinuation

- Low quality evidence from 1 RCT (N=386) shows a clinically important and statistically significant benefit of collaborative care, relative to standard care, on discontinuation at 12 months for adults with remitted depression.

Comparison 3: Stepped care versus standard care/enhanced standard care

Critical outcomes

Depression symptomatology

- Very low quality evidence from 2 RCTs (N=1614) shows a statistically significant but not clinically important benefit of stepped care, relative to standard care or enhanced standard care, on depression symptomatology endpoint score at 6 months for adults with depression.
- Very low quality evidence from 2 RCTs (N=826) shows a clinically important and statistically significant benefit of stepped care, relative to standard care, on depression symptomatology change score at 6 months for adults with depression.
- Moderate quality evidence from 1 RCT (N=998) shows neither a clinically important nor statistically significant effect of stepped care, relative to enhanced standard care, on depression symptomatology endpoint score at 12 months for adults with depression.
- Low quality evidence from 1 RCT (N=194) shows neither a clinically important nor statistically significant effect of stepped care, relative to standard care, on depression symptomatology change score at 12 months for adults with depression.

Response

- Very low quality evidence from 1 RCT (N=239) shows a clinically important but not statistically significant benefit of standard care, relative to stepped care, on the rate of response at 6 months for adults with depression.
- Low quality evidence from 1 RCT (N=239) shows a clinically important but not statistically significant benefit of standard care, relative to stepped care, on the rate of response at 12 months for adults with depression.

Remission

- Low quality evidence from 2 RCTs (N=1082) shows a clinically important and statistically significant benefit of stepped care, relative to standard care or enhanced standard care, on the rate of remission at 6 months for adults with depression.
- Very low quality evidence from 2 RCTs (N=2085) shows a clinically important but not statistically significant benefit of stepped care, relative to enhanced standard care, on the rate of remission at 12 months for adults with depression.

Important outcomes

Antidepressant use

- Moderate quality evidence from 1 RCT (N=175) shows a clinically important and statistically significant benefit of stepped care, relative to standard care, on antidepressant use at 6 months for adults with depression.

Discontinuation

- Low quality evidence from 5 RCTs (N=3180) shows a clinically important and statistically significant benefit of stepped care, relative to standard care or

enhanced standard care, on discontinuation at 6 months for adults with depression.

- Moderate quality evidence from 3 RCTs (N=2324) shows a clinically important and statistically significant benefit of stepped care, relative to standard care or enhanced standard care, on discontinuation at 12 months for adults with depression.

Comparison 4: Stepped care versus standard care for relapse prevention

Critical outcomes

Relapse

- Low quality evidence from 1 RCT (N=135) shows a clinically important but not statistically significant benefit of standard care, relative to stepped care, on the rate of relapse at 12 months in adults with remitted depression.

Important outcomes

Antidepressant use

- Very low quality evidence from 1 RCT (N=94) shows neither a clinically important nor statistically significant effect of stepped care, relative to standard care, on antidepressant use at 12 months for adults with remitted depression.

Discontinuation

- Low quality evidence from 1 RCT (N=74) shows neither a clinically important nor statistically significant effect of stepped care, relative to standard care, on discontinuation at 12 months for adults with remitted depression.

Comparison 5: Pure medication management versus standard care

Critical outcomes

Depression symptomatology

- High quality evidence from 2 RCTs (N=399) shows neither a clinically important nor statistically significant benefit of pure medication management, relative to standard care, on depression symptomatology at 6 months for adults with depression.

Response

- Moderate quality evidence from 1 RCT (N=70) shows a clinically important but not statistically significant benefit of pure medication management, relative to standard care, on the rate of response at 6 months for adults with depression.

Important outcomes

Antidepressant use

- Low quality evidence from 3 RCTs (N=904) shows a clinically important and statistically significant benefit of pure medication management, relative to standard care, on antidepressant use at 6 months for adults with depression.

Discontinuation

- Moderate quality evidence from 5 RCTs (N=1216) shows neither a clinically important nor statistically significant benefit of pure medication management, relative to standard care, on discontinuation at 6 months for adults with depression.

Comparison 6: Care coordination versus standard care/enhanced standard care

Critical outcomes

Depression symptomatology

- Very low quality evidence from 1 RCT (N=62) shows neither a clinically important nor statistically significant benefit of care coordination, relative to enhanced standard care, on depression symptomatology at 6 months for adults with depression.
- Moderate quality evidence from 1 RCT (N=516) shows neither a clinically important nor statistically significant benefit of care coordination, relative to standard care, on depression symptomatology at 12 months for adults with depression.

Remission

- Low quality evidence from 1 RCT (N=609) shows neither a clinically important nor statistically significant benefit of care coordination, relative to standard care, on the rate of remission at 12 months for adults with depression.

Important outcomes

Discontinuation

- Very low quality evidence from 1 RCT (N=62) shows neither a clinically important nor statistically significant effect of care coordination, relative to enhanced standard care, on discontinuation at 6 months for adults with depression.
- Low quality evidence from 1 RCT (N=609) shows a clinically important but not statistically significant benefit of standard care, relative to care coordination, on discontinuation at 12 months for adults with depression.

Comparison 7: Attached professional model versus enhanced standard care

Critical outcomes

Depression symptomatology

- Very low quality evidence from 1 RCT (N=118) shows neither a clinically important nor statistically significant benefit of attached professional model care, relative to enhanced standard care, on depression symptomatology at 6 months for adults with depression.

Important outcomes

Discontinuation

- Very low quality evidence from 1 RCT (N=120) shows a clinically important but not statistically significant benefit of attached professional model care, relative

to enhanced standard care, on discontinuation at 6 months for adults with depression.

Comparison 8: Shared care versus standard care

Critical outcomes

Depression symptomatology

- High quality evidence from 1 RCT (N=69) shows a clinically important and statistically significant benefit of shared care, relative to standard care, on depression symptomatology at 6 months for adults with depression.

Remission

- Moderate quality evidence from 1 RCT (N=69) shows a clinically important and statistically significant benefit of shared care, relative to standard care, on the rate of remission at 6 months for adults with depression.

Important outcomes

Antidepressant use

- High quality evidence from 1 RCT (N=69) shows a clinically important and statistically significant benefit of shared care, relative to standard care, on antidepressant use at 6 months for adults with depression.

Discontinuation

- Low quality evidence from 1 RCT (N=69) shows neither a clinically important nor statistically significant effect of shared care, relative to standard care, on discontinuation at 6 months for adults with depression.

Comparison 9: Measurement-based care versus standard care

Critical outcomes

Depression symptomatology

- Moderate quality evidence from 1 RCT (N=81) shows a clinically important and statistically significant benefit of measurement-based care, relative to standard care, on depression symptomatology at 6 months for adults with depression.

Response

- Low quality evidence from 1 RCT (N=120) shows a clinically important and statistically significant benefit of measurement-based care, relative to standard care, on the rate of response at 6 months for adults with depression.

Remission

- Moderate quality evidence from 1 RCT (N=120) shows a clinically important and statistically significant benefit of measurement-based care, relative to standard care, on remission at 6 months for adults with depression.

Important outcomes

Discontinuation

- Very low quality evidence from 1 RCT (N=120) shows a clinically important but not statistically significant benefit of measurement-based care, relative to standard care, on discontinuation at 6 months for adults with depression.

Economic evidence statements

Collaborative care

- Evidence from 3 UK economic evaluations conducted alongside RCTs (N = 1,771; complete data for economic analysis n=1341) suggest that simple collaborative care is possibly a cost-effective model for delivering services to adults or older adults with depression. This evidence is directly applicable to the UK context and is coming from one study with minor and two studies with potentially serious methodological limitations.
- Evidence from 1 US study conducted alongside a RCT (N=386) suggests that simple collaborative care aiming at relapse prevention may be cost-effective in adults with depression that is in remission. This evidence is partially applicable to the NICE decision-making context as it comes from a US study and is not using the QALY as the outcome measure. The study is characterised by potentially serious methodological limitations.
- Evidence from 1 UK study conducted alongside a RCT (N=187) suggests that complex collaborative care is not cost-effective compared with usual secondary mental health care for adults with depression. This evidence is directly applicable to the UK context and is characterised by minor limitations.
- Evidence from 1 Dutch study and 1 German study conducted alongside RCTs (N=396) suggest that complex collaborative care is unlikely to be cost-effective compared with treatment as usual in adults with depression in primary care. This evidence is partially applicable to the NICE decision-making context as the studies were conducted outside the UK and, in the Dutch study, utility values were based on EQ-5D ratings using the Dutch tariff. One study is characterised by potentially serious limitations and the other study by minor limitations.

Stepped care

- Evidence from 1 UK study conducted alongside a cohort study with matched sites (N=403), and 3 non-UK studies (2 Dutch and 1 Canadian) based on decision-analytic economic modelling suggests that stepped care might be cost-effective for adults with depression in primary care, although results were inconsistent within and across studies. This evidence is directly applicable (UK study) and partially applicable (Dutch and Canadian studies) to the NICE decision-making context. The UK study is characterised by potentially serious limitations; of the 3 non-UK studies, 1 is characterised by potentially serious limitations and 2 are characterised by minor limitations.

Medication management

- Evidence from 1 Spanish study conducted alongside a RCT (N=179) suggests that medication management may be cost-effective for adults with depression. This evidence is partially applicable to the NICE decision-making context as it was conducted outside the UK and is characterised by potentially serious limitations.

Care co-ordination

- No evidence on the cost effectiveness of care co-ordination for adults with depression is available.

Attached professional model

- No evidence on the cost effectiveness of the attached professional model for adults with depression is available.

Shared care

- Evidence from 1 US study conducted alongside a multi-site pragmatic RCT (N=840) is inconclusive regarding the cost effectiveness of shared care compared with usual primary care that includes a referral system to specialist care. The evidence is partially applicable to the NICE decision making context (US study, QALYs based on SF-36 using preferences for matched vignettes created following cluster analysis of SF-12 mental and physical component scores, elicited by US service users with depression using SG) and is characterised by potentially serious limitations.

Measurement-based care

- No evidence on the cost effectiveness of measurement-based care for adults with depression is available.

The committee's discussion of the evidence

Interpreting the evidence

The outcomes that matter most

The aim of this review was to determine if different models of service delivery improved outcomes for people with depression so the committee identified depression symptomatology and response, remission and relapse to be the critical outcomes for this question. Antidepressant use and discontinuation were identified as important outcomes. For all outcomes, time points of 6 and 12 months were used, to ensure comparability across interventions.

Evidence was available for all outcomes and time points of interest for the collaborative care dataset (comparison 1), but for all other comparisons data were only available for some of the outcomes. A number of different care models did not have available data on the outcomes of remission and response. Therefore when considering the evidence the committee placed the greatest emphasis on depression symptomatology and antidepressant use, as these provided the best point of comparison across different interventions.

The quality of the evidence

The committee noted that most outcomes for most of the comparisons had been assessed in GRADE as either low or very low quality. Most outcomes were downgraded due to risk of bias (common reasons for downgrading included a lack of blinding of participants and intervention administrators, and non-blind or unclear blinding of outcome assessment, and significant baseline differences between groups) and imprecision. The committee also noted the absence of evidence identified for head-to-head comparisons of different service delivery models.

Benefits and harms

The committee considered that effective service delivery models would enhance clinical outcomes by improved engagement with effective interventions and thereby improve outcomes in terms of depression symptomatology and response, remission and relapse.

For collaborative care, the committee noted that there was evidence from a number of UK and international trials for clinical benefits associated with the use of collaborative care compared to standard care or enhanced standard care, with higher rates of response and remission at both 6 and 12 months. However, the committee noted that the heterogeneity was very high, and effect sizes for depression symptomatology were small compared to first-line acute treatments. Based on these factors, the committee made a 'consider' rather than 'offer' recommendation and identified groups where collaborative care may confer significant added value, for example, those with significant physical health problems or who are socially isolated.

Older adults were also identified as a group that may particularly benefit from collaborative care. Subgroup analysis comparing outcomes for older (mean age ≥ 60 years) and younger (mean age <60 years) adults did not identify statistically significant subgroup differences. However, there was a consistent trend for larger benefits of collaborative care for older adults. Considered together with the committee knowledge and experience of difficulties with engagement in older adults particularly for those with physical health problems, and evidence for the cost-effectiveness of collaborative care in older people, the committee agreed to also recommend collaborative care for this group.

The committee defined the components of collaborative care that are important, based both on their expertise and experience and on the results of sub-analyses of the collaborative care dataset. Subgroup analyses examined the impact of complex (relative to simple) collaborative care, case manager background, use of a psychological intervention or stepped care algorithm, and the number of contacts provided as part of the intervention. No significant subgroup differences or consistent pattern in results were observed for analyses comparing outcomes for complex versus simple collaborative care, or case manager with mental health background versus case manager without a mental health background.

The inclusion of a stepped care algorithm showed a trend for larger effect sizes compared to no stepped care algorithm. There were no significant subgroup differences for the inclusion of psychological interventions, however, the committee agreed based on their knowledge and experience that collaborative care should include delivery of psychological and psychosocial interventions within a structured protocol. This was also reinforced by evidence for the benefits of stepped care interventions (that were not integrated into collaborative care models) relative to standard care on depression symptomatology, the rate of remission and antidepressant use at 6 months. The committee considered adding a separate recommendation for stepped care but agreed that the key principles were covered by existing recommendations, and this also had the advantage of being integrated into a care pathway that emphasises patient choice.

Subgroup analysis comparing the outcomes of collaborative care between interventions with fewer than 13 contacts and interventions with 13 contacts or more contacts, showed a trend for larger effects sizes with more contacts and this difference was statistically significant for remission at 12 months. However, the committee did not consider this evidence sufficiently compelling to specify the number of contacts that a collaborative care intervention should include.

The committee were aware of the importance of medication adherence, in particular, for people with severe and chronic depressive symptoms and noted that although the evidence for pure medication management was limited and did not show significant benefits on clinical outcomes, there were benefits on antidepressant use at 6 months. Based on this limited evidence, the committee agreed not to make any recommendations about the use of medication management as an independent service delivery model. For people with depression who may have specific difficulties with the uptake of, or engagement with, treatment the committee agreed that medication adherence would be more effectively promoted through the delivery of care in a collaborative, multidisciplinary manner and that included medication management as a component within a collaborative care model.

The committee acknowledged that for more severe depression or chronic depression with multiple complicating problems or significant coexisting conditions there was no direct evidence to guide the development of recommendations. The committee were, however, aware of the very significant difficulties that people with severe, chronic and complex depression face and the burden of suffering this represents for families and carers. Such high levels of need are best met by specialist services within specialist secondary care. The committee therefore drew on their expert knowledge and experience of specialist services and used informal consensus to develop a series of recommendations on who might benefit from specialist services, how these services should be co-ordinated and what the nature of the co-ordination of the services should involve. The committee were of the view that the development of a comprehensive multidisciplinary care plan will allow more timely, appropriate, and individualised planning and delivery of care to people with more severe or more chronic depression with multiple complicating problems or significant coexisting conditions, and that these benefits should offset (fully or partially) the costs associated with development of the care plan. In contrast, lack of a detailed care plan may lead to sub-optimal, less clinically and cost-effective care pathways and inappropriate treatments, ultimately leading to sub-optimal outcomes for the person and higher healthcare costs.

Cost effectiveness and resource use

The committee agreed that, overall, the published economic evidence indicated that simple collaborative care is potentially a cost-effective model for delivering services to adults with depression, including older adults. This is because out of the 3 UK cost-utility studies included in the review, 2 found simple collaborative care cost-effective when added to usual primary care compared with usual primary care alone at the NICE lower cost-effectiveness threshold of £20,000/QALY. The third study reported an ICER for simple collaborative care between the NICE lower and upper (£30,000/QALY) cost-effectiveness thresholds. The two studies that found simple collaborative care cost-effective were also somewhat larger than the one that found it cost-ineffective at the lower NICE cost-effectiveness threshold. The committee also noted that, among the 3 studies, there was one with minor methodological limitations (the other two were characterised by potentially serious limitations), and this found simple collaborative care to be cost-effective. In contrast, the only UK study on more resource-intensive complex collaborative care included in the review suggested this is unlikely to be cost-effective compared with usual secondary mental health care, as its ICER was well above the NICE upper cost-effectiveness threshold of £30,000/QALY. Therefore, the committee decided to recommend collaborative care with the characteristics of the less resource-intensive, simple collaborative care, as defined in this review, for organising the delivery of care and treatment of people with depression.

The committee noted, based on the evidence, that stepped care might also be cost effective for adults with depression.

The committee noted that no UK economic evidence was available and non-UK evidence did not provide any substantial support for the cost effectiveness of medication management as an independent service delivery model for adults with depression. They also noted that non-UK economic evidence on shared care was inconclusive.

The committee acknowledged that referring people with more severe depression or chronic depressive symptoms and multiple complicating problems (such as unemployment, poor housing or financial problems) or significant coexisting conditions to specialist mental health services, if they have not benefitted from treatment or if they have impaired functioning, is likely to incur additional costs compared with no referral. However, they agreed that the number of people affected would be small and any additional costs were likely to be offset by cost-savings resulting from more appropriate care for this population following referral (compared with treatment in primary care settings), leading to improved outcomes and reduction in the need for potentially costly care further down the care pathway.

Recommendations supported by this evidence review

This evidence review supports recommendations 1.15.7 to 1.15.10 in the NICE guideline.

References

Adewuya 2019

Adewuya AO, Ola BA, Coker O, Atilola O, Fasawe A, Ajomale T. A stepped care intervention for non-specialist health workers' management of depression in the Mental Health in Primary Care (MeHPriC) project, Lagos, Nigeria: a cluster randomised controlled trial. General Hospital Psychiatry. 2019 Sep 1;60:76-82.

Åkerblad 2003

Åkerblad AC, Bengtsson F, Ekselius L, von Knorring L. Effects of an educational compliance enhancement programme and therapeutic drug monitoring on treatment adherence in depressed patients managed by general practitioners. International Clinical Psychopharmacology. 2003 Nov 1;18(6):347-54.

Aljumah 2015

Aljumah K, Hassali MA. Impact of pharmacist intervention on adherence and measurable patient outcomes among depressed patients: a randomised controlled study. BMC Psychiatry. 2015 Dec;15(1):219.

Apil 2012

Apil SR, Hoencamp E, Judith Haffmans PM, Spinhoven P. A stepped care relapse prevention program for depression in older people: a randomized controlled trial. International Journal of Geriatric Psychiatry. 2012 Jun;27(6):583-91.

Aragonès 2012

Aragonès E, Piñol JL, Caballero A, López-Cortacans G, Casaus P, Hernández JM, Badia W, Folch S. Effectiveness of a multi-component programme for managing depression in primary care: a cluster randomized trial. The INDI project. Journal of Affective Disorders. 2012 Dec 15;142(1-3):297-305.

Araya 2003

Araya R, Rojas G, Fritsch R, Gaete J, Rojas M, Simon G, Peters TJ. Treating depression in primary care in low-income women in Santiago, Chile: a randomised controlled trial. The Lancet. 2003 Mar 22;361(9362):995-1000.

Banerjee 1996

Banerjee S, Shamash K, Macdonald AJ, Mann AH. Randomised controlled trial of effect of intervention by psychogeriatric team on depression in frail elderly people at home. BMJ. 1996 Oct 26;313(7064):1058-61.

Bedoya 2014

Bedoya CA, Traeger L, Trinh NH, Chang TE, Brill CD, Hails K, Hagan PN, Flaherty K, Yeung A. Impact of a culturally focused psychiatric consultation on depressive symptoms among Latinos in primary care. Psychiatric Services. 2014 Oct;65(10):1256-62.

Berghöfer 2012

Berghöfer A, Hartwich A, Bauer M, Unützer J, Willich SN, Pfennig A. Efficacy of a systematic depression management program in high utilizers of primary care: a randomized trial. BMC Health Services Research. 2012 Dec;12(1):298.

Björkelund 2018

Björkelund C, Svenningsson I, Hange D, Udo C, Petersson EL, Ariai N, Nejati S, Wessman C, Wikberg C, André M, Wallin L. Clinical effectiveness of care managers in collaborative care for patients with depression in Swedish primary health care: a pragmatic cluster randomized controlled trial. BMC Family Practice. 2018 Dec;19(1):28.

Bosanquet 2017

Bosanquet K, Adamson J, Atherton K, Bailey D, Baxter C, Beresford-Dent J, et al. CollAborative care for Screen-Positive ElderRs with major depression (CASPER plus): a multicentred randomised controlled trial of clinical effectiveness and cost-effectiveness. Health Technol Assess 2017;21(67).

Bruce 2004

Bruce ML, Ten Have TR, Reynolds III CF, Katz II, Schulberg HC, Mulsant BH, Brown GK, McAvay GJ, Pearson JL, Alexopoulos GS. Reducing suicidal ideation and depressive symptoms in depressed older primary care patients: a randomized controlled trial. JAMA. 2004 Mar 3;291(9):1081-91.

Buszewicz 2016

Buszewicz M, Griffin M, McMahon EM, Walters K, King M. Practice nurse-led proactive care for chronic depression in primary care: a randomised controlled trial. The British Journal of Psychiatry. 2016 Apr;208(4):374-80.

Callahan 1994

Callahan CM, Hendrie HC, Dittus RS, Brater DC, Hui SL, Tierney WM. Improving treatment of late life depression in primary care: a randomized clinical trial. Journal of the American Geriatrics Society. 1994 Aug;42(8):839-46.

Capoccia 2004

Capoccia KL, Boudreau DM, Blough DK, Ellsworth AJ, Clark DR, Stevens NG, Katon WJ, Sullivan SD. Randomized trial of pharmacist interventions to improve depression care and outcomes in primary care. American Journal of Health System Pharmacy. 2004 Feb 15;61(4):364-72.

Chen 2015

Chen S, Conwell Y, He J, Lu N, Wu J. Depression care management for adults older than 60 years in primary care clinics in urban China: a cluster-randomised trial. The Lancet Psychiatry. 2015 Apr 1;2(4):332-9.

Curth 2020

Curth NK, Brinck-Claussen UØ, Hjorthøj C, Davidsen AS, Mikkelsen JH, Lau ME, Lundsteen M, Csillag C, Christensen KS, Jakobsen M, Bojesen AB. Collaborative care for depression and anxiety disorders: results and lessons learned from the Danish cluster-randomized Collabri trials. BMC Family Practice. 2020 Dec;21(1):1-5.

Dobscha 2006

Dobscha SK, Corson K, Hickam DH, Perrin NA, Kraemer DF, Gerrity MS. Depression decision support in primary care: a cluster randomized trial. Annals of Internal Medicine. 2006 Oct 3;145(7):477-87.

Ell 2007

Ell K, Unützer J, Aranda M, Gibbs NE, Lee PJ, Xie B. Managing depression in home health care: a randomized clinical trial. Home Health Care Services Quarterly. 2007 Jul 12;26(3):81-104.

Finley 2003

Finley PR, Rens HR, Pont JT, Gess SL, Louie C, Bull SA, Lee JY, Bero LA. Impact of a collaborative care model on depression in a primary care setting: a randomized controlled trial. Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy. 2003 Sep;23(9):1175-85.

Fortney 2007

Fortney JC, Pyne JM, Edlund MJ, Williams DK, Robinson DE, Mittal D, Henderson KL. A randomized trial of telemedicine-based collaborative care for depression. Journal of General Internal Medicine. 2007 Aug 1;22(8):1086-93.

Gensichen 2009

Gensichen J, von Korff M, Peitz M, Muth C, Beyer M, Güthlin C, Torge M, Petersen JJ, Rosemann T, König J, Gerlach FM. Case management for depression by health care assistants in small primary care practices: a cluster randomized trial. Annals of Internal Medicine. 2009 Sep 15;151(6):369-78.

Gilbody 2017 / Lewis 2017

Gilbody S, Lewis H, Adamson J, Atherton K, Bailey D, Birtwistle J, Bosanquet K, Clare E, Delgadillo J, Ekers D, Foster D. Effect of collaborative care vs usual care on

depressive symptoms in older adults with subthreshold depression: the CASPER randomized clinical trial. JAMA. 2017 Feb 21;317(7):728-37.

Lewis H, Adamson J, Atherton K, Bailey D, Birtwistle J, Bosanquet K, Clare E, Delgadillo J, Ekers D, Foster D, Gabe R. CollAborative care and active surveillance for Screen-Positive EldeRs with subthreshold depression (CASPER): a multicentred randomised controlled trial of clinical effectiveness and cost-effectiveness. Health Technology Assessment (Winchester, England). 2017 Feb;21(8):1.

Guo 2015

Guo T, Xiang YT, Xiao L, Hu CQ, Chiu HF, Ungvari GS, Correll CU, Lai KY, Feng L, Geng Y, Feng Y. Measurement-based care versus standard care for major depression: a randomized controlled trial with blind raters. American Journal of Psychiatry. 2015 Aug 28;172(10):1004-13.

Gureje 2019

Gureje O, Oladeji BD, Montgomery AA, Bello T, Kola L, Ojagbemi A, Chisholm D, Araya R. Effect of a stepped-care intervention delivered by lay health workers on major depressive disorder among primary care patients in Nigeria (STEPCARE): a cluster-randomised controlled trial. The Lancet Global Health. 2019 Jul 1;7(7):e951-60.

Härter 2018

Härter M, Watzke B, Daubmann A, Wegscheider K, König HH, Brettschneider C, Liebherz S, Heddaeus D, Steinmann M. Guideline-based stepped and collaborative care for patients with depression in a cluster-randomised trial. Scientific Reports. 2018;8.

Health Quality Ontario 2019

Health Quality Ontario (2019). Internet-Delivered Cognitive Behavioural Therapy for Major Depression and Anxiety Disorders: A Health Technology Assessment. Ont Health Technol Assess Ser, 19(6):1-199.

Hölzel 2018 / Grochtdreis 2019

Hölzel LP, Bjerregaard F, Bleich C, Boczor S, Härter M, König HH, Kloppe T, Niebling W, Scherer M, Tinsel I, Hüll M. Coordinated Treatment of Depression in Elderly People in Primary Care: A Cluster-Randomized, Controlled Study (GermanIMPACT). Deutsches Ärzteblatt International. 2018 Nov;115(44):741.

Grochtdreis, T., Brettschneider, C., Bjerregaard, F., Bleich, C., Boczor, S., Harter, M. et al. (2019). Cost-effectiveness analysis of collaborative treatment of late-life depression in primary care (GermanIMPACT). European Psychiatry, 57, 10-18.

Huang 2018

Huang HC, Liu SI, Hwang LC, Sun FJ, Tjung JJ, Huang CR, Li TC, Huang YP, Yeung A. The effectiveness of Culturally Sensitive Collaborative Treatment of depressed Chinese in family medicine clinics: A randomized controlled trial. General Hospital Psychiatry. 2018 Jan 1;50:96-103.

Huijbregts 2013 / Goorden 2015

Huijbregts KM, de Jong FJ, van Marwijk HW, Beekman AT, Adèr HJ, Hakkaart-van Roijen L, Unützer J, van der Feltz-Cornelis CM. A target-driven collaborative care

model for Major Depressive Disorder is effective in primary care in the Netherlands. A randomized clinical trial from the depression initiative. Journal of Affective Disorders. 2013 Apr 25;146(3):328-37.

Goorden M, Huijbregts KM, van Marwijk HW, Beekman AT, van der Feltz-Cornelis CM, Hakkaart-van Roijen L (2015) Cost-utility of collaborative care for major depressive disorder in primary care in the Netherlands. Journal of Psychosomatic Research, 79(4), 316-23.

Jarjoura 2004

Jarjoura D, Polen A, Baum E, Kropp D, Hetrick S, Rutecki G. Effectiveness of screening and treatment for depression in ambulatory indigent patients. Journal of General Internal Medicine. 2004 Jan;19(1):78-84.

Jeong 2013

Jeong H, Yim HW, Jo SJ, Nam B, Kwon SM, Choi JY, Yang SK. The effects of care management on depression treatment in a psychiatric clinic: a randomized controlled trial. International Journal Of Geriatric Psychiatry. 2013 Oct;28(10):1023-30.

Katon 1999

Katon W, Von Korff M, Lin E, Simon G, Walker E, Unützer J, Bush T, Russo J, Ludman E. Stepped collaborative care for primary care patients with persistent symptoms of depression: a randomized trial. Archives of General Psychiatry. 1999 Dec 1;56(12):1109-15.

Katon 2001 / Simon 2002

Katon W, Rutter C, Ludman EJ, Von Korff M, Lin E, Simon G, Bush T, Walker E, Unützer J. A randomized trial of relapse prevention of depression in primary care. Archives of General Psychiatry. 2001 Mar 1;58(3):241-7.

Simon GE, Von Korff M, Ludman EJ, Katon WJ, Rutter C, Unutzer J, Lin EH, Bush T, Walker E (2002) Cost-effectiveness of a program to prevent depression relapse in primary care. Medical care 40: 941-950.

Katzelnick 2000

Katzelnick DJ, Simon GE, Pearson SD, Manning WG, Helstad CP, Henk HJ, Cole SM, Lin EH, Taylor LH, Kobak KA. Randomized trial of a depression management program in high utilizers of medical care. Archives of Family Medicine. 2000 Apr 1;9(4):345.

Knapstad 2020

Knapstad M, Lervik LV, Sæther SM, Aarø LE, Smith OR. Effectiveness of prompt mental health care, the Norwegian version of improving access to psychological therapies: a randomized controlled trial. Psychotherapy and Psychosomatics. 2020;89(2):90-105.

Landis 2007

Landis SE, Gaynes BN, Morrissey JP, Vinson N, Ellis AR, Domino ME. Generalist care managers for the treatment of depressed medicaid patients in North Carolina: a pilot study. BMC Family Practice. 2007 Dec;8(1):7.

Ludman 2007

Ludman EJ, Simon GE, Grothaus LC, Luce C, Markley DK, Schaefer J. A pilot study of telephone care management and structured disease self-management groups for chronic depression. Psychiatric Services. 2007 Aug;58(8):1065-72.

McMahon 2007

McMahon L, Foran KM, Forrest SD, Taylor ML, Ingram G, Rajwal M, Cornwall PL, McAllister-Williams RH. Graduate mental health worker case management of depression in UK primary care: a pilot study. Br J Gen Pract. 2007 Nov 1;57(544):880-5.

Meeuwissen 2019

Meeuwissen JAC, Feenstra TL, Smit F, Blankers M, Spijker J, Bockting CLH et al. (2019). The cost-utility of stepped-care algorithms according to depression guideline recommendations - Results of a state-transition model analysis. Journal of Affective Disorders, 242, 244-254.

Morriss 2016

Morriss R, Garland A, Nixon N, Guo B, James M, Kaylor-Hughes C, Moore R, Ramana R, Sampson C, Sweeney T, Dalgleish T. Efficacy and cost-effectiveness of a specialist depression service versus usual specialist mental health care to manage persistent depression: a randomised controlled trial. The Lancet Psychiatry. 2016 Sep 1;3(9):821-31.

Mukuria 2013

Mukuria C, Brazier J, Barkham M, Connell J, Hardy G, Hutten R, Saxon D, Dent-Brown K, Parry G (2013) Cost-effectiveness of an improving access to psychological therapies service. British Journal of Psychiatry 202: 220-227.

Ng 2020

Ng TP, Nyunt MS, Feng L, Kumar R, Fones CS, Ko SM. Collaborative care for primary care treatment of late-life depression in Singapore: Randomized controlled trial. International Journal of Geriatric Psychiatry. 2020 Oct;35(10):1171-80.

Oladeji 2015

Oladeji BD, Kola L, Abiona T, Montgomery AA, Araya R, Gureje O. A pilot randomized controlled trial of a stepped care intervention package for depression in primary care in Nigeria. BMC Psychiatry. 2015 Dec;15(1):96.

Richards 2013 / Green 2014

Richards DA, Hill JJ, Gask L, Lovell K, Chew-Graham C, Bower P, Cape J, Pilling S, Araya R, Kessler D, Bland JM. Clinical effectiveness of collaborative care for depression in UK primary care (CADET): cluster randomised controlled trial. Br Med J. 2013 Aug 19;347:f4913.

Richards DA, Bower P, Chew-Graham C, Gask L, Lovell K, Cape J, Pilling S, Araya R, Kessler D, Barkham M, Bland JM. 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).

Green C, Richards DA, Hill JJ, Gask L, Lovell K, Chew-Graham C, et al. (2014) Cost-effectiveness of collaborative care for depression in UK primary care: Economic evaluation of a randomised controlled trial (CADET). PLoS ONE 9(8): e104225.

Rickles 2005

Rickles NM, Svarstad BL, Statz-Paynter JL, Taylor LV, Kobak KA. Pharmacist telemonitoring of antidepressant use: effects on pharmacist–patient collaboration. Journal of the American Pharmacists Association. 2005 May 1;45(3):344-53.

Rubio-Valera 2013

Rubio-Valera M, Pujol MM, Fernández A, Peñarrubia-María MT, Travé P, del Hoyo YL, Serrano-Blanco A. Evaluation of a pharmacist intervention on patients initiating pharmacological treatment for depression: a randomized controlled superiority trial. European Neuropsychopharmacology. 2013 Sep 1;23(9):1057-66.

Rubio-Valera M, Bosmans J, Fernandez A, Penarrubia-Maria M, March M, Trave P, Bellon JA, Serrano-Blanco A (2013) Cost-Effectiveness of a Community Pharmacist Intervention in Patients with Depression: A Randomized Controlled Trial (PRODEFAR Study). PLoS ONE 8(8): e70588.

Salisbury 2016

Salisbury C, O'Cathain A, Edwards L, Thomas C, Gaunt D, Hollinghurst S, Nicholl J, Large S, Yardley L, Lewis G, Foster A. Effectiveness of an integrated telehealth service for patients with depression: a pragmatic randomised controlled trial of a complex intervention. The Lancet Psychiatry. 2016 Jun 1;3(6):515-25.

Simon 2004 (CM)

Simon GE, Ludman EJ, Tutty S, Operskalski B, Von Korff M. Telephone psychotherapy and telephone care management for primary care patients starting antidepressant treatment: a randomized controlled trial. JAMA. 2004 Aug 25;292(8):935-42.

Simon 2004 (CM+psych)

Simon GE, Ludman EJ, Tutty S, Operskalski B, Von Korff M. Telephone psychotherapy and telephone care management for primary care patients starting antidepressant treatment: a randomized controlled trial. JAMA. 2004 Aug 25;292(8):935-42.

Simon 2006

Simon GE, Ludman EJ, Operskalski BH. Randomized trial of a telephone care management program for outpatients starting antidepressant treatment. Psychiatric Services. 2006 Oct;57(10):1441-5.

Sirey 2010

Sirey JA, Bruce ML, Kales HC. Improving antidepressant adherence and depression outcomes in primary care: the treatment initiation and participation (TIP) program. The American Journal of Geriatric Psychiatry. 2010 Jun 1;18(6):554-62.

Smit 2006

Smit A, Kluiter H, Conradi HJ, Van der Meer K, Tiemens BG, Jenner JA, Van Os TW, Ormel J. Short-term effects of enhanced treatment for depression in primary care: results from a randomized controlled trial. Psychological Medicine. 2006 Jan;36(1):15-26.

Swindle 2003

Swindle RW, Rao JK, Helmy A, Plue L, Zhou XH, Eckert GJ, Weinberger M. Integrating clinical nurse specialists into the treatment of primary care patients with depression. The International Journal of Psychiatry in Medicine. 2003 Mar;33(1):17-37.

Unützer 2002

Unützer J, Katon W, Callahan CM, Williams Jr JW, Hunkeler E, Harpole L, Hoffing M, Della Penna RD, Noël PH, Lin EH, Areán PA. Collaborative care management of late-life depression in the primary care setting: a randomized controlled trial. JAMA. 2002 Dec 11;288(22):2836-45.

Areán PA, Ayalon L, Hunkeler E, Lin EH, Tang L, Harpole L, Hendrie H, Williams Jr JW, Unützer J. Improving depression care for older, minority patients in primary care. Medical Care. 2005 Apr 1:381-90.

van der Weele 2012

van der Weele GM, de Waal MW, van den Hout WB, de Craen AJ, Spinhoven P, Stijnen T, Assendelft WJ, van der Mast RC, Gussekloo J. Effects of a stepped-care intervention programme among older subjects who screened positive for depressive symptoms in general practice: the PROMODE randomised controlled trial. Age and Ageing. 2012 Mar 16;41(4):482-8.

Wells 2000

Wells KB, Sherbourne C, Schoenbaum M, Duan N, Meredith L, Unützer J, Miranda J, Carney MF, Rubenstein LV. Impact of disseminating quality improvement programs for depression in managed primary care: a randomized controlled trial. JAMA. 2000 Jan 12;283(2):212-20.

Wiley-Exley 2009

Wiley-Exley E, Domino ME, Maxwell J, Levkoff SE. Cost-effectiveness of integrated care for elderly depressed patients in the PRISM-E study. Journal of Mental Health Policy and Economics 2009; 12: 205-213+217.

Yeung 2010

Yeung A, Shyu I, Fisher L, Wu S, Yang H, Fava M. Culturally sensitive collaborative treatment for depressed Chinese Americans in primary care. American Journal Of Public Health. 2010 Dec;100(12):2397-402.

Yeung 2016

Yeung A, Martinson MA, Baer L, Chen J, Clain A, Williams A, Chang TE, Trinh NH, Alpert JE, Fava M. The Effectiveness of Telepsychiatry-Based Culturally Sensitive Collaborative Treatment for Depressed Chinese American Immigrants: A Randomized Controlled Trial. The Journal of Clinical Psychiatry. 2016 Aug;77(8):e996-1002.

Settings of care

Review question

For adults with depression, what are the relative benefits and harms associated with different settings for the delivery of care?

Introduction

Care for adults with depression can be provided in a variety of different settings, ranging from care in people's own homes, primary care and day hospitals, through to inpatient care or tertiary settings, and the setting in which care is delivered may have a bearing on the outcomes for individuals, and the effectiveness of the interventions.

The aim of this review is to identify if there is a setting which delivers optimal results for people with depression, and if there is anything about the general management of care that should be done differently when delivered in different settings.

Summary of the protocol

Please see Table 1 for a summary of the Population, Intervention, Comparison and Outcome (PICO) characteristics of this review.

Table 11: Summary of the protocol (PICO table)

Population	• Adults with a diagnosis of depression according to DSM, ICD or similar criteria, or depressive symptoms as indicated by baseline depression scores on validated scales (and including those with subthreshold [just below threshold] depressive symptoms)
Intervention	Settings for the delivery of care, which may include: • Primary care • Crisis resolution and home treatment teams • Inpatient setting • Acute psychiatric day hospital care • Non-acute day hospital care and recovery centres • Specialist tertiary affective disorders settings • Community mental health teams • Residential services
Comparison	• Any other setting for the delivery of care
Outcomes	**Critical:** • Depression symptomatology (mean endpoint score or change in depression score from baseline) • Remission (usually defined as a score below clinical threshold on a depression scale) • Response (usually defined as at least 50% improvement from the baseline score on a depression scale) • Relapse (number of people who returned to a depressive episode whilst in remission) **Important:** • Service utilisation/resource use (e.g. antidepressant use) • Psychological functioning • Social functioning

	• Satisfaction • Carer distress

DSM: diagnostic and statistical manual of mental disorders; ICD: international classification of diseases

For further details see the review protocol in appendix A.

Methods and process

This evidence review was developed using the methods and process described in Developing NICE guidelines: the manual. Methods specific to this review question are described in the review protocol in appendix A.

Declarations of interest were recorded according to NICE's 2014 conflicts of interest policy until 31 March 2018. From 1 April 2018, declarations of interest were recorded according to NICE's 2018 conflicts of interest policy. Those interests declared until April 2018 were reclassified according to NICE's 2018 conflicts of interest policy (see Register of Interests).

Clinical evidence

Included studies

No randomised controlled trial (RCT) evidence was identified that specifically addressed the following settings: primary care, and inpatient care, therefore, as specified in the full protocol (see Appendix A), indirect evidence was considered in the form of sub-analyses of the NMA dataset (Evidence report B: Treatment of a new episode of depression).

Comparison 1. Primary care versus secondary care

As outlined above, no RCT evidence was identified that specifically addressed this comparison. Therefore the committee considered indirect evidence in the form of sub-analyses of the NMA dataset (Evidence report B: Treatment of a new episode of depression). Primary versus secondary care differences were examined for critical outcomes that had more than 2 studies in each subgroup.

Subgroup analysis of primary care versus secondary care was possible for 5 comparisons in the NMA dataset, and all 5 comparisons were for adults with more severe depression (corresponding to the categories of moderate and severe depression):

- Comparison 1a. Cognitive and cognitive behavioural therapies individual + antidepressant versus antidepressant, with 2 RCTs included for primary care (Naeem 2011; Scott 1997) and 4 RCTs included for secondary care (Ashouri 2013; Hautzinger 1996; Hollon 1992; Zu 2014). Primary care versus secondary care subgroup analysis was possible for the depression symptoms endpoint outcome only.
- Comparison 1b. Selective serotonin reuptake inhibitors (SSRIs) versus placebo, with 5 RCTs included for primary care (Bjerkenstedt 2005; Doogan 1994; Lepola 2003; Lopez-Rodriguez 2004; Wade 2002) and 78 RCTs included for secondary care (29060 07 001; Andreoli 2002/ Dubini 1997/ Massana 1998_study 1 [1 study reported across 3 papers]; Baune 2018; Binnemann 2008; Bose 2008; Burke 2002; Byerley 1988; Claghorn 1992a; Claghorn 1992b; Clayton 2006_study 1; Clayton 2006_study 2; CL3-20098-022; CL3-20098-023; CL3-20098-024; Detke 2004; Dube 2010; Dunbar 1993; Eli Lilly

HMAT-A; Emsley 2018; Fabre 1992; Fabre 1995a; Fava 1998a; Fava 2005; FDA 245 (EMD 68 843-010); FDA 246 (SB 659746-003); Forest Laboratories 2000; Forest Research Institute 2005; Golden 2002_448; Golden 2002_449; Goldstein 2002; Goldstein 2004; Gual 2003; Higuchi 2009; Hirayasu 2011a; Hirayasu 2011b; Jefferson 2000; Kasper 2012; Katz 2004; Keller 2006_Study 062; Kramer 1998; Kranzler 2006_Group A; Lam 2016; Macias-Cortes 2015; Mathews 2015; Mendels 1999; Miller 1989a; Montgomery 1992; Mundt 2012; MY-1042/BRL-029060/CPMS-251; MY-1042/BRL-029060/1 (PAR 128); Nemeroff 2007; Nierenberg 2007; NKD20006 (NCT00048204); Nyth 1992; Olie 1997; PAR 01 001 (GSK & FDA); Perahia 2006; Peselow 1989a; Peselow 1989b; Rapaport 2009; Ratti 2011_study 096; Ravindran 1995; Reimherr 1990; Rickels 1992; Rudolph 1999; SER 315 (FDA); Sheehan 2009b; Smith 1992; Stark 1985; Study 62b (FDA); Study F1J-MC-HMAQ- Study Group B; Tollefson 1993/1995 [1 study reported across 2 papers]; Valle-Cabrera 2018; VEN XR 367 (FDA); Wang 2014c; WELL AK1A4006; Wernicke 1987; Wernicke 1988). Primary care versus secondary care subgroup analyses were possible for the depression symptoms endpoint, depression symptoms change score, and response outcomes.

- Comparison 1c. SSRIs versus tricyclic antidepressants (TCAs), with 10 RCTs included for primary care (Christiansen 1996; Freed 1999; Hutchinson 1992; Kyle 1998; Moon 1994; Moon 1996; PAR 29060/281; PAR MDUK 032; Rosenberg 1994; Serrano-Blanco 2006) and 47 RCTs included for secondary care (29060 07 001; 29060/299; Akhondzadeh 2003; Arminem 1992; Beasley 1993b; Bersani 1994; Bhargava 2012; Bremner 1984; Byerley 1988; Chiu 1996; Cohn 1984b; Cohn 1990b; Danish University Antidepressant Group 1986; Danish University Antidepressant Group 1990; De Ronchi 1998; Demyttenaere 1998; Deuschle 2003; Fabre 1991; Fabre 1992; Fawcett 1989; Feighner 1993; Forlenza 2001; Geretsegger 1995; GSK_29060/103; Hashemi 2012; Keegan 1991; Laakmann 1988; Laakmann 1991; Levine 1989; Marchesi 1998; MDF/29060/III/070/88/MC; Miura 2000; Moller 1993; Moller 1998; Mulsant 1999; Navarro 2001; Ontiveros Sanchez 1998; Peselow 1989a; Peselow 1989b; Peters 1990; Preskorn 1991; Reimherr 1990; Ropert 1989; SER 315 (FDA); Staner 1995; Stark 1985; Suleman 1997). Primary care versus secondary care subgroup analyses were possible for the depression symptoms endpoint, depression symptoms change score, remission and response outcomes.
- Comparison 1d. TCAs versus placebo, with 6 RCTs included for primary care (Barge-Schaapveld 2002; Blashki 1971; Lecrubier 1997; Mynors-Wallis 1995; Philipp 1999; Schweizer 1998) and 30 RCTs included for secondary care (29060 07 001; Amsterdam 1986; Bakish 1992b; Bremner 1995; Byerley 1988; Cassano 1986; Elkin 1989/Imber 1990 [1 study reported across 2 papers]; Escobar 1980; Fabre 1992; Feiger 1996; Feighner 1982; Feighner 1989b; Fontaine 1994; Goldberg 1980; Kusalic 1993; McCallum 1975; MIR 003-020 (FDA); Peselow 1989a; Peselow 1989b; Reimherr 1990; Rickels 1982e; Rickels 1991; Rickels 1995_Study 006-1; Rickels 1995_Study 006-2; Schweizer 1994; SER 315 (FDA); Silverstone 1994; Smith 1990; Stark 1985; White 1984). Primary care versus secondary care subgroup analyses were possible for the depression symptoms endpoint, depression symptoms change score, and response outcomes.
- Comparison 1e. Serotonin–norepinephrine reuptake inhibitors (SNRIs) versus SSRIs, with 2 RCTs included for primary care (Montgomery 2004; Tylee 1997) and 29 RCTs included for secondary care (Allard 2004; Alves 1999; Bielski 2004; Clerc 1994; Costa 1998; DeNayer 2002; Detke 2004; Diaz-Martinez 1998; Dierick 1996; Eli Lilly HMAT-A; Goldstein 2002; Goldstein 2004; Hao

2014; Higuchi 2009; Hwang 2004; Jiang 2017; Khan 2007; Kornaat 2000; Mehtonen 2000; Nemeroff 2007; Nierenberg 2007; Perahia 2006; Rickels 2000; Rudolph 1999; Sheehan 2009b; Shelton 2006; Sir 2005; Study F1J-MC-HMAQ-Study Group B; Tzanakaki 2000). Primary care versus secondary care subgroup analyses were possible for the remission and response outcomes.

Comparison 2. Crisis resolution team care versus standard care (for adults with non-psychotic severe mental illness)

No RCT evidence was identified that specifically addressed this comparison for adults with depression. The committee therefore agreed to consider a wider evidence base including non-psychotic severe mental illness. A systematic review (Murphy 2015; updated version of Joy 2003 used in 2009 guideline) was identified that examined crisis intervention for people with severe mental illness. This Cochrane review was used as a source of studies with inclusion criteria into this review of over 50% of the population having a non-psychotic disorder. Of the 8 RCTs included in Murphy 2015, 1 of these studies met the >50% non-psychotic disorder inclusion criterion (Johnson 2005).

Comparison 3. Inpatient versus outpatient settings

No randomised controlled trial (RCT) evidence was identified that specifically addressed this comparison. Therefore the committee considered indirect evidence in the form of sub-analyses of the NMA dataset (Evidence report B: Treatment of a new episode of depression). Differences between inpatient and outpatient settings were examined for critical outcomes that had more than 2 studies in each subgroup.

Subgroup analysis of inpatient versus outpatient settings was possible for 6 comparisons in the NMA dataset, and all 6 comparisons were for adults with more severe depression (corresponding to the categories of moderate and severe depression):

- Comparison 3a. Selective serotonin reuptake inhibitors (SSRIs) versus placebo, with 3 RCTs included for inpatient settings (29060 07 001; Katz 2004; Sheehan 2009b) and 74 RCTs included for outpatient settings (Baune 2018; Binnemann 2008; Bjerkenstedt 2005; Blumenthal 2007/Hoffman 2011 [1 study reported across 2 papers]; Bose 2008; Burke 2002; Byerley 1988; Claghorn 1992a; Claghorn 1992b; Clayton 2006_study 1; Clayton 2006_study 2; Detke 2004; Doogan 1994; Dube 2010; Dunbar 1993; Eli Lilly HMAT-A; Emsley 2018; Fabre 1992; Fava 1998a; Fava 2005; FDA 245 (EMD 68 843-010); Forest Laboratories 2000; Forest Research Institute 2005; Golden 2002_448; Golden 2002_449; Goldstein 2002; Goldstein 2004; Gual 2003; Hirayasu 2011a; Hirayasu 2011b; Hunter 2010_study 1; Hunter 2011; Jefferson 2000; Keller 2006_Study 062; Komulainen 2018; Kramer 1998; Kranzler 2006_Group A; Lam 2016; Lepola 2003; Macias-Cortes 2015; Mathews 2015; Mendels 1999; Miller 1989a; Mundt 2012; MY-1042/BRL-029060/CPMS-251; MY-1045/BRL-029060/1 (PAR 128); Nemeroff 2007; Nierenberg 2007; NKD20006 (NCT00048204); Olie 1997; PAR 01 001 (GSK & FDA); Perahia 2006; Peselow 1989a; Peselow 1989b; Rapaport 2009; Ratti 2011_study 096; Ravindran 1995; Reimherr 1990; Rickels 1992; Roose 2004; Rudolph 1999; SER 315 (FDA); Smith 1992; Stark 1985; Study 62b (FDA); Study F1J-MC-HMAQ – Study Group B; Tollefson 1993/1995 [1 study reported across 2 papers]; Valle-Cabrera 2018; VEN XR 367 (FDA); Wade 2002; Wang 2014c; WELL AK1A4006; Wernicke 1987; Wernicke 1988). Inpatient versus outpatient subgroup analysis was possible for the depression symptoms change score and response outcomes.

- Comparison 3b. SSRIs versus tricyclic antidepressants (TCAs), with 11 RCTs included for inpatient settings (29060/299; 29060 07 001; Arminen 1992; Danish University Antidepressant Group 1986; Danish University Antidepressant Group 1990; Deushle 2003; Geretsegger 1995; Laakmann 1991; Moller 1993; Moller 1998; Staner 1995), and 40 RCTs included for outpatient settings (Akhondzadeh 2003; Beasley 1993b; Bersani 1994; Bhargava 2012; Bremner 1984; Byerley 1988; Christiansen 1996; Cohn 1984b; Cohn 1990b; De Ronchi 1998; Demyttenaere 1998; Fabre 1991; Fabre 1992; Fawcett 1989; Feighner 1993; Forlenza 2001; Freed 1999; Hashemi 2012; Hutchinson 1992; Kyle 1998; Laakmann 1988; Marchesi 1998; MDF/29060/III/070/88/MC; Moller 2000; Moon 1994; Moon 1996; Ontiveros Sanchez 1998; PAR 29060/281; PAR MDUK 032; Peselow 1989a; Peselow 1989b; Peters 1990; Preskorn 1991; Reimherr 1990; Ropert 1989; Rosenberg 1994; SER 315 (FDA); Serrano-Blanco 2006; Stark 1985; Suleman 1997). Inpatient versus outpatient subgroup analysis was possible for the depression symptoms endpoint, depression symptoms change score, remission, and response outcomes.
- Comparison 3c. Serotonin–norepinephrine reuptake inhibitors (SNRIs) versus placebo, with 2 RCTs included for inpatient settings (Guelfi 1995; Sheehan 2009b), and 26 RCTs included for outpatient settings (Brannan 2005; Cutler 2009; Detke 2002a; Detke 2002b; Detke 2004; Eli Lilly HMAT-A; Goldstein 2002; Goldstein 2004; Hewett 2009; Hewett 2010; Higuchi 2016; Khan 1998; Levin 2013; Mendels 1993; Nemeroff 2007; Nierenberg 2007; Perahia 2006; Raskin 2007; Robinson 2014; Rudolph 1999; Schweizer 1994; Study F1J-MC-HMAQ-Study Group B; Thase 1997; VEN 600A-303 (FDA); VEN 600A-313 (FDA); VEN XR 367 (FDA)). Inpatient versus outpatient subgroup analysis was possible for the depression symptoms endpoint, depression symptoms change score, and remission outcomes.
- Comparison 3d. SNRIs versus SSRIs, with 4 RCTs included for inpatient settings (Clerc 1994; Hwang 2004; Sheehan 2009b; Tzanakaki 2000), and 32 RCTs included for outpatient settings (Allard 2004; Alves 1999; Bielski 2004; Casabona 2004; Chang 2015; Costa 1998; DeNayer 2002; Detke 2004; Diaz-Martinez 1998; Dierick 1996; Eli Lilly HMAT-A; Goldstein 2002; Goldstein 2004; Hackett 1996; Heller 2009; Jiang 2017; Khan 2007; Kornaat 2000; Mehtonen 2000; Montgomery 2004; Mowla 2016; Nemeroff 2007; Nierenberg 2007; Perahia 2006; Rickels 2000; Rudolph 1999; Shelton 2006; Sir 2005; Study F1J-MC-HMAQ-Study Group B; Tylee 1997; VEN XR 367 (FDA); Wade 2007) . Inpatient versus outpatient subgroup analysis was possible for the depression symptoms endpoint, depression symptoms change score, remission, and response outcomes.
- Comparison 3e. Mirtazapine versus TCAs, with 2 RCTs included for inpatient settings (Richou 1995; Zivkov 1995), and 4 RCTs included for outpatient settings (Bremner 1995; MIR 003-020 (FDA); MIR 003-021 (FDA); Smith 1990). Inpatient versus outpatient subgroup analysis was only possible for the response outcome.
- Comparison 3f. Acupuncture + antidepressants versus antidepressants, with 2 RCTs included for inpatient settings (Wang 2014a; Zhang 2007a), and 2 RCTs included for outpatient settings (Qu 2013; Zhao 2019a). Inpatient versus outpatient subgroup analysis was only possible for the depression symptoms change score outcome.

Comparison 4. Acute psychiatric day hospital care versus inpatient care (for adults with depression and non-psychotic severe mental illness)

Acute psychiatric day hospitals are units that provide diagnostic and treatment services for acutely ill individuals who would otherwise be treated in traditional psychiatric inpatient units. 1 RCT (Dinger 2014) was identified that specifically addressed acute psychiatric day hospital care for adults with depression. The committee therefore agreed to consider a wider evidence base including non-psychotic severe mental illness. A systematic review (Marshall 2011) was identified that compared day hospital to inpatient care for people with acute psychiatric disorders. This Cochrane review was used as a source of studies with inclusion criteria into this review of over 50% of the population having a non-psychotic disorder.

Of the 10 RCTs included in Marshall 2011, 5 of these studies met the >50% non-psychotic disorder inclusion criterion (Creed 1990; Creed 1997; Dick 1985; Kallert 2007; Schene 1993).

Comparison 5. Non-acute day hospital care versus outpatient care (for adults with depression and non-psychotic severe mental illness)

No RCT evidence was identified that specifically addressed this setting for adults with depression. The committee therefore agreed to consider a wider evidence base including non-psychotic severe mental illness. A systematic review (Marshall 2001) was identified that examined the use of day hospitals as an alternative to outpatient care for people with psychiatric disorders. This Cochrane review was used as a source of studies with inclusion criteria into this review of over 50% of the population having a non-psychotic disorder.

Of the 8 studies included in Marshall 2001, 3 of these studies met the >50% non-psychotic disorder inclusion criterion (Dick 1991; Glick 1986; Tyrer 1979).

Comparison 6. Community mental health teams versus standard care (for adults with non-psychotic severe mental illness)

No RCT evidence was identified that specifically addressed this setting for adults with depression. The committee therefore agreed to consider a wider evidence base including non-psychotic severe mental illness. A systematic review (Malone 2007) was identified that examined community mental health teams (CMHTs) for people with severe mental illnesses and disordered personality. This Cochrane review was used as a source of studies with inclusion criteria into this review of over 50% of the population having a non-psychotic disorder.

Of the 3 studies included in Malone 2007, 1 of these studies met the >50% non-psychotic disorder inclusion criterion (Merson 1992).

See the literature search strategy in appendix B and study selection flow chart in appendix C.

Excluded studies

Studies not included in this review with reasons for their exclusions are provided in appendix K.

Summary of clinical studies included in the evidence review

Comparison 1. Primary care versus secondary care

Summaries of the studies included in the primary care versus secondary care subgroup analysis of the cognitive and cognitive behavioural therapies individual + antidepressant versus antidepressant comparison are presented in Table 2.

There were no significant subgroup differences between primary care and secondary care for the comparison cognitive and cognitive behavioural therapies individual + antidepressant versus antidepressant on: depression symptoms endpoint (Test for subgroup differences: $Chi^2 = 0.27$, df = 1, p = 0.60).

Table 12: Summary of included studies for primary care versus secondary care subgroup analysis for comparison 1a Cognitive and cognitive behavioural therapies individual + antidepressant versus antidepressant

Study	Population	Intervention	Comparison	Comments
Primary care (K=2, N=82)				
Naeem 2011 RCT Pakistan	Primary care N=34 Baseline severity: More severe Mean age (years): 33.0 Sex (% female): 74 Ethnicity (% BME): NR	CBT individual (9 weekly or fortnightly sessions) + SSRI (paroxetine or fluoxetine 20mg/day)	SSRI (paroxetine or fluoxetine 20mg/day)	Treatment duration (weeks): 12 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint
Scott 1997 RCT UK	Primary care N=48 Baseline severity: More severe Mean age (years): 41.0 Sex (% female): 67 Ethnicity (% BME): NR	CBT individual (6x weekly 30-min sessions) + any antidepressant	Any antidepressant	Treatment duration (weeks): 7 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint
Secondary care (K=4, N=311)				
Ashouri 2013 RCT Iran	Secondary care N=33 Baseline severity: More severe Mean age (years): 32.5 Sex (% female): 61 Ethnicity (% BME): NR	Third-wave cognitive therapy individual or CBT individual (number of sessions not reported) + any antidepressant	Any antidepressant	Treatment duration (weeks): NR Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint

Study	Population	Intervention	Comparison	Comments
Hautzinger 1996a RCT Germany	Secondary care N=76 Baseline severity: More severe Mean age (years): NR Sex (% female): NR Ethnicity (% BME): NR	CBT individual (24x 50-60 min sessions) + amitriptyline 150mg/day	Amitriptyline 150mg/day	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint
Hollon 1992a RCT US	Secondary care N=82 Baseline severity: More severe Mean age (years): 32.6 Sex (% female): 80 Ethnicity (% BME): 9	CBT individual (maximum 20x 50-min weekly or fortnightly sessions) + imipramine 75-450mg/day	Imipramine 75-450mg/day	Treatment duration (weeks): 12 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint
Zu 2014b RCT China	Secondary care N=120 Baseline severity: More severe Mean age (years): 38.3 Sex (% female): 49 Ethnicity (% BME): NR	CBT individual (20x 1-hour sessions) + any SSRI (dose NR, within recommended therapeutic dose ranges)	Any SSRI (dose NR, within recommended therapeutic dose ranges)	Treatment duration (weeks): 24 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint

a Three-armed trial but where possible the demographics reported here are for only the two relevant arms.
b Four-armed trial but where possible the demographics reported here are for only the two relevant arms.
BME: black, minority, ethnic; CBT: cognitive behavioural therapy; K: number of studies; mg: milligrams; N: number of participants; NR: not reported; RCT: randomised controlled trial; SSRI: selective serotonin reuptake inhibitor.

Summaries of the studies included in the primary care versus secondary care subgroup analysis of the selective serotonin reuptake inhibitors (SSRIs) versus placebo comparison are presented in Table 3.

There were no significant subgroup differences between primary care and secondary care for the comparison SSRIs versus placebo on: depression symptoms endpoint (Test for subgroup differences: $Chi^2 = 0.01$, df = 1, p = 0.91); depression symptoms change score (Test for subgroup differences: $Chi^2 = 0.26$, df = 1, p = 0.61); response (Test for subgroup differences: $Chi^2 = 1.75$, df = 1, p = 0.19).

Table 13: Summary of included studies for primary care versus secondary care subgroup analysis for comparison 1b Selective serotonin reuptake inhibitors (SSRIs) versus placebo

Study	Population	Intervention	Comparison	Comments
Primary care (K=5, N=1,184)				
Bjerkenstedt 2005 RCT Sweden	Primary care N=115 Baseline severity: More severe Mean age (years): 50.9 Sex (% female): 79 Ethnicity (% BME): 0	Fluoxetine 20mg/day	Placebo	Treatment duration (weeks): 4 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score • Response
Doogan 1994 RCT UK	Primary care N=200 Baseline severity: More severe Mean age (years): 45.7 Sex (% female): 68 Ethnicity (% BME): NR	Sertraline 50-100mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Response
Lepola 2003 RCT Belgium, Canada, Finland, France, Norway, Sweden, Switzerland & UK	Primary care N=469 Baseline severity: More severe Mean age (years): 43.3 Sex (% female): 72 Ethnicity (% BME): NR	Escitalopram 10-20mg/day or citalopram 20-40mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Response
Lopez-Rodriguez 2004 RCT South America	Primary care N=20 Baseline severity: More severe Mean age (years): 31.9 Sex (% female): NR Ethnicity (% BME): NR	Fluoxetine 20mg/day	Placebo	Treatment duration (weeks): 9 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint

Study	Population	Intervention	Comparison	Comments
Wade 2002 RCT Canada, Estonia, France, Netherlands & UK	Primary care N=380 Baseline severity: More severe Mean age (years): 40.5 Sex (% female): 76 Ethnicity (% BME): 3	Escitalopram 10mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score • Response
Secondary care (K=78, N=18,070)				
29060 07 001a RCT US	Secondary care N=25 Baseline severity: More severe Mean age (years): 42.5 Sex (% female): 56 Ethnicity (% BME): NR	Paroxetine 10-60mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score
Andreoli 2002/ Dubini 1997/ Massana 1998_study 1 RCT Brazil, France, Ireland, Italy, Poland, and UK	Secondary care N=255 Baseline severity: More severe Mean age (years): 42.2 Sex (% female): 60 Ethnicity (% BME): NR	Fluoxetine 20-40mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score • Response
Baune 2018 RCT Estonia, Finland, Germany, & Lithuania	Secondary care N=104 Baseline severity: More severe Mean age (years): 45.7 Sex (% female): 64 Ethnicity (% BME): 2	Paroxetine 20mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score

Study	Population	Intervention	Comparison	Comments
Binnemann 2008 RCT US, Serbia and Montenegro, & the Russian Federation	Secondary care N=82 Baseline severity: More severe Mean age (years): 49 Sex (% female): 39 Ethnicity (% BME): NR	Sertraline 100mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score • Response
Bose 2008 RCT US	Secondary care N=267 Baseline severity: More severe Mean age (years): 68.3 Sex (% female): 59 Ethnicity (% BME): 11	Escitalopram 10-20mg/day	Placebo	Treatment duration (weeks): 12 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score • Response
Burke 2002 RCT US	Secondary care N=491 Baseline severity: More severe Mean age (years): 40.1 Sex (% female): 65 Ethnicity (% BME): NR	Escitalopram 10mg/day or 20mg/day, or citalopram 40mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score • Response
Byerley 1988a RCT US	Secondary care N=61 Baseline severity: More severe Mean age (years): 38.3 Sex (% female): 68 Ethnicity (% BME): NR	Fluoxetine 40-80mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint • Response
Claghorn 1992a RCT US	Secondary care N=59 Baseline severity: More severe	Paroxetine 10-50mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus

Study	Population	Intervention	Comparison	Comments
	Mean age (years): NR Sex (% female): NR Ethnicity (% BME): NR			secondary care subgroup analysis): • Depression symptoms change score
Claghorn 1992b RCT US	Secondary care N=72 Baseline severity: More severe Mean age (years): 35 Sex (% female): 32 Ethnicity (% BME): NR	Paroxetine 10-50mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score • Response
Clayton 2006_study 1 RCT US	Secondary care N=283 Baseline severity: More severe Mean age (years): 35 Sex (% female): 61 Ethnicity (% BME): 35	Escitalopram 10-20mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score • Response
Clayton 2006_study 2 RCT US	Secondary care N=286 Baseline severity: More severe Mean age (years): 36.5 Sex (% female): 56 Ethnicity (% BME): 27	Escitalopram 10-20mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score • Response
CL3-20098-022 RCT Europe	Secondary care N=286 Baseline severity: More severe Mean age (years): 43 Sex (% female): 67 Ethnicity (% BME): NR	Fluoxetine 20mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint • Response

Study	Population	Intervention	Comparison	Comments
CL3-20098-023 RCT Cross-continental	Secondary care N=275 Baseline severity: More severe Mean age (years): 41.1 Sex (% female): 75 Ethnicity (% BME): NR	Paroxetine 20mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint
CL3-20098-024 RCT Cross-continental	Secondary care N=306 Baseline severity: More severe Mean age (years): 41.5 Sex (% female): 73 Ethnicity (% BME): NR	Fluoxetine 20mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint • Response
Detke 2004b RCT US	Secondary care N=179 Baseline severity: More severe Mean age (years): 42.9 Sex (% female): 71 Ethnicity (% BME): 0	Paroxetine 20mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score • Response
Dube 2010 RCT India, US, Mexico & Romania	Secondary care N=200 Baseline severity: More severe Mean age (years): 36.5 Sex (% female): 44 Ethnicity (% BME): NR	Escitalopram 10-20mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score • Response
Dunbar 1993 RCT US	Secondary care N=341 Baseline severity: More severe	Paroxetine 10-50mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus

Study	Population	Intervention	Comparison	Comments
	Mean age (years): 41 Sex (% female): 51 Ethnicity (% BME): NR			secondary care subgroup analysis): • Response
Eli Lilly HMAT-Aa RCT US	Secondary care N=179 Baseline severity: More severe Mean age (years): NR Sex (% female): NR Ethnicity (% BME): NR	Paroxetine 20mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score • Response
Emsley 2018 RCT Bulgaria, Estonia, Finland, France, Republic of Korea, Malaysia, Mexico, Poland, Romania, & Slovakia	Secondary care N=206 Baseline severity: More severe Mean age (years): 70.6 Sex (% female): 75 Ethnicity (% BME): NR	Escitalopram 10mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score • Response
Fabre 1992a RCT US	Secondary care N=80 Baseline severity: More severe Mean age (years): 35.8 Sex (% female): 59 Ethnicity (% BME): NR	Paroxetine 10-50mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score
Fabre 1995a RCT US	Secondary care N=369 Baseline severity: More severe Mean age (years): 37.6 Sex (% female): 53	Sertraline 50mg/day, 100mg/day, or 200mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care

Study	Population	Intervention	Comparison	Comments
	Ethnicity (% BME): 9			subgroup analysis): • Depression symptoms change score • Response
Fava 1998a RCT US	Secondary care N=128 Baseline severity: More severe Mean age (years): 41.3 Sex (% female): 51 Ethnicity (% BME): NR	Paroxetine 20-50mg/day or fluoxetine 20-80mg/day	Placebo	Treatment duration (weeks): 12 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score • Response
Fava 2005 RCT US	Secondary care N=90 Baseline severity: More severe Mean age (years): 37.2 Sex (% female): 59 Ethnicity (% BME): NR	Fluoxetine 20mg/day	Placebo	Treatment duration (weeks): 12 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score
FDA 245 (EMD 68 843-010) RCT US	Secondary care N=191 Baseline severity: More severe Mean age (years): NR Sex (% female): NR Ethnicity (% BME): NR	Fluoxetine 20mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score
FDA 246 (SB 659746-003) RCT US	Secondary care N=246 Baseline severity: More severe Mean age (years): NR	Citalopram 20mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care

Study	Population	Intervention	Comparison	Comments
	Sex (% female): NR Ethnicity (% BME): NR			subgroup analysis): • Depression symptoms endpoint
Forest Laboratories 2000 RCT US	Secondary care N=386 Baseline severity: More severe Mean age (years): 42 Sex (% female): 52 Ethnicity (% BME): NR	Escitalopram 10-20mg/day or citalopram 20-40mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score • Response
Forest Research Institute 2005 RCT US	Secondary care N=409 Baseline severity: More severe Mean age (years): 40 Sex (% female): 56 Ethnicity (% BME): NR	Escitalopram 10-20mg/day or sertraline 50-200mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score • Response
Golden 2002_448 RCT US	Secondary care N=315 Baseline severity: More severe Mean age (years): 39 Sex (% female):NR Ethnicity (% BME): NR	Paroxetine 20-62.5mg/day	Placebo	Treatment duration (weeks): 12 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score
Golden 2002_449 RCT US	Secondary care N=330 Baseline severity: More severe Mean age (years): 41.2 Sex (% female): NR	Paroxetine 20-62.5mg/day	Placebo	Treatment duration (weeks): 12 Outcomes (for primary versus secondary care subgroup analysis):

Study	Population	Intervention	Comparison	Comments
	Ethnicity (% BME): NR			• Depression symptoms change score
Goldstein 2002a RCT US	Secondary care N=103 Baseline severity: More severe Mean age (years): 40.9 Sex (% female): 65 Ethnicity (% BME): 21	Fluoxetine 20mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Response
Goldstein 2004a RCT US	Secondary care N=176 Baseline severity: More severe Mean age (years): 40 Sex (% female): 64 Ethnicity (% BME): 22	Paroxetine 20mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Response
Gual 2003 RCT Spain	Secondary care N=83 Baseline severity: More severe Mean age (years): 46.7 Sex (% female): 47 Ethnicity (% BME): NR	Sertraline 50-150mg/day	Placebo	Treatment duration (weeks): 24 Outcomes (for primary versus secondary care subgroup analysis): • Response
Higuchi 2009a RCT Japan	Secondary care N=294 Baseline severity: More severe Mean age (years): 38.3 Sex (% female): NR Ethnicity (% BME): NR	Paroxetine 20-40mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score • Response
Hirayasu 2011a RCT Japan	Secondary care N=310 Baseline severity: More severe Mean age (years): 34.6	Escitalopram 10mg/day or 20mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis):

Study	Population	Intervention	Comparison	Comments
	Sex (% female): NR Ethnicity (% BME): NR			• Depression symptoms endpoint • Response
Hirayasu 2011b RCT Japan	Secondary care N=485 Baseline severity: More severe Mean age (years): 36.2 Sex (% female): NR Ethnicity (% BME): NR	Escitalopram 10mg/day or 20mg/day, or paroxetine 20-40mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint • Response
Jefferson 2000 RCT US	Secondary care N=415 Baseline severity: More severe Mean age (years): 39.9 Sex (% female): NR Ethnicity (% BME): NR	Paroxetine 25mg/day, or citalopram 20mg/day or 40mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score • Response
Kasper 2012 RCT Russia & Austria	Secondary care N=211 Baseline severity: More severe Mean age (years): 41.9 Sex (% female): 71 Ethnicity (% BME): 0	Escitalopram 20mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score • Response
Katz 2004 RCT US	Secondary care N=53 Baseline severity: More severe Mean age (years): NR Sex (% female): NR Ethnicity (% BME): NR	Paroxetine 20-60mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Response

Study	Population	Intervention	Comparison	Comments
Keller 2006_Study 062 RCT Cross-continental	Secondary care N=325 Baseline severity: More severe Mean age (years):41 Sex (% female): 67 Ethnicity (% BME): 43	Paroxetine 20mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score
Kramer 1998 RCT US	Secondary care N=142 Baseline severity: More severe Mean age (years): NR Sex (% female): NR Ethnicity (% BME): NR	Paroxetine 20mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Response
Kranzler 2006_Group A RCT US	Secondary care N=189 Baseline severity: More severe Mean age (years): 42.9 Sex (% female): 35 Ethnicity (% BME): 10	Sertraline 50-200mg/day	Placebo	Treatment duration (weeks): 10 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score • Response
Lam 2016b RCT Canada	Secondary care N=61 Baseline severity: More severe Mean age (years): 36.8 Sex (% female): 72 Ethnicity (% BME): NR	Fluoxetine 20mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score • Response
Macias-Cortes 2015 RCT Mexico	Secondary care N=89 Baseline severity: More severe Mean age (years): 49	Fluoxetine 20mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care

Study	Population	Intervention	Comparison	Comments
	Sex (% female): 100 Ethnicity (% BME): 100			subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score • Response
Mathews 2015 RCT US	Secondary care N=579 Baseline severity: More severe Mean age (years): 42.3 Sex (% female): 57 Ethnicity (% BME): 32	Citalopram 40mg/day	Placebo	Treatment duration (weeks): 10 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score • Response
Mendels 1999 RCT US	Secondary care N=180 Baseline severity: More severe Mean age (years): 43 Sex (% female): 33 Ethnicity (% BME): 13	Citalopram 20-80mg/day	Placebo	Treatment duration (weeks): 4 Outcomes (for primary versus secondary care subgroup analysis): • Response
Miller 1989a RCT UK	Secondary care N=47 Baseline severity: More severe Mean age (years): 42.5 Sex (% female): 68 Ethnicity (% BME): NR	Paroxetine 30mg/day	Placebo	Treatment duration (weeks): 4 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score
Montgomery 1992 RCT UK	Secondary care N=199 Baseline severity: More severe Mean age (years): 44	Citalopram 20mg/day or 40mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care

Study	Population	Intervention	Comparison	Comments
	Sex (% female): 69 Ethnicity (% BME): NR			subgroup analysis): • Depression symptoms change score
Mundt 2012 RCT US	Secondary care N=165 Baseline severity: More severe Mean age (years): 37.8 Sex (% female): 63 Ethnicity (% BME): 24	Sertraline 50-100mg/day	Placebo	Treatment duration (weeks): 4 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score • Response
MY-1042/BRL-029060/CPMS-251 RCT US	Secondary care N=254 Baseline severity: More severe Mean age (years): 41.9 Sex (% female): NR Ethnicity (% BME): NR	Paroxetine 20-50mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score • Response
MY-1042/BRL-029060/1 (PAR 128) RCT US	Secondary care N=848 Baseline severity: More severe Mean age (years): 41.8 Sex (% female): NR Ethnicity (% BME): NR	Paroxetine 20-50mg/day or fluoxetine 20-80mg/day	Placebo	Treatment duration (weeks): 12 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score • Response
Nemeroff 2007a RCT US	Secondary care N=206 Baseline severity: More severe Mean age (years): 39.1 Sex (% female): 61	Fluoxetine 20-60mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis):

Study	Population	Intervention	Comparison	Comments
	Ethnicity (% BME): 7			• Response
Nierenberg 2007a RCT US	Secondary care N=411 Baseline severity: More severe Mean age (years): 43 Sex (% female): 66 Ethnicity (% BME): 21	Escitalopram 10mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score • Response
NKD20006 (NCT00048204) RCT US	Secondary care N=250 Baseline severity: More severe Mean age (years): 38 Sex (% female): 60 Ethnicity (% BME): NR	Paroxetine 20mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score • Response
Nyth 1992 RCT Denmark, Norway & Sweden	Secondary care N=149 Baseline severity: More severe Mean age (years): 76.7 Sex (% female): 69 Ethnicity (% BME): NR	Citalopram 10-30mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score • Response
Olie 1997 RCT France	Secondary care N=258 Baseline severity: More severe Mean age (years): 43.8 Sex (% female): 63 Ethnicity (% BME): 1	Sertraline 50-200mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Response

Study	Population	Intervention	Comparison	Comments
PAR 01 001 (GSK & FDA) RCT US	Secondary care N=50 Baseline severity: More severe Mean age (years): 43.1 Sex (% female): 35 Ethnicity (% BME): NR	Paroxetine 10-50mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score • Response
Perahia 2006b RCT Bulgaria, Croatia, Hungary, Poland, Romania, Russia, & Slovakia	Secondary care N=196 Baseline severity: More severe Mean age (years): 45.2 Sex (% female): 68 Ethnicity (% BME): 0	Paroxetine 20mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Response
Peselow 1989aa RCT US	Secondary care N=73 Baseline severity: More severe Mean age (years): 43.2 Sex (% female): 38 Ethnicity (% BME): NR	Paroxetine 10-50mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Response
Peselow 1989ba RCT US	Secondary care N=82 Baseline severity: More severe Mean age (years): NR Sex (% female): NR Ethnicity (% BME): NR	Paroxetine 20-50mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Response
Rapaport 2009 RCT US	Secondary care N=357 Baseline severity: More severe Mean age (years): 67.5 Sex (% female): 62 Ethnicity (% BME): 17	Paroxetine 25mg/day	Placebo	Treatment duration (weeks): 10 Outcomes (for primary versus secondary care

Study	Population	Intervention	Comparison	Comments
				subgroup analysis): • Depression symptoms change score • Response
Ratti 2011_study 096 RCT 11 countries in Europe and Latin America	Secondary care N=236 Baseline severity: More severe Mean age (years): 44 Sex (% female): 72 Ethnicity (% BME): NR	Paroxetine 20-30mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Response
Ravindran 1995 RCT Canada	Secondary care N=66 Baseline severity: More severe Mean age (years): 38.9 Sex (% female): 62 Ethnicity (% BME): NR	Sertraline 50-200mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Response
Reimherr 1990a RCT US & Canada	Secondary care N=299 Baseline severity: More severe Mean age (years): 39.6 Sex (% female): 53 Ethnicity (% BME): 8	Sertraline 20-200mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score • Response
Rickels 1992 RCT US	Secondary care N=111 Baseline severity: More severe Mean age (years): 44.7 Sex (% female): 48 Ethnicity (% BME): NR	Paroxetine (dose NR)	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Response

Study	Population	Intervention	Comparison	Comments
Rudolph 1999a RCT US	Secondary care N=201 Baseline severity: More severe Mean age (years): 40 Sex (% female): 66 Ethnicity (% BME): NR	Fluoxetine 20-60mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint • Response
SER 315 (FDA)a RCT Europe	Secondary care N=165 Baseline severity: More severe Mean age (years): 42.0 Sex (% female): 72 Ethnicity (% BME): NR	Sertraline 50-200mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score
Sheehan 2009ba RCT US	Secondary care N=194 Baseline severity: More severe Mean age (years): 38.8 Sex (% female): 66 Ethnicity (% BME): NR	Fluoxetine 60-80mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score • Response
Smith 1992 RCT US	Secondary care N=77 Baseline severity: More severe Mean age (years): 44.8 Sex (% female): 50 Ethnicity (% BME): NR	Paroxetine 10-50mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Response
Stark 1985a RCT US	Secondary care N=354 Baseline severity: More severe	Fluoxetine 60-80mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus

Study	Population	Intervention	Comparison	Comments
	Mean age (years): 40.5 Sex (% female): 68 Ethnicity (% BME): NR			secondary care subgroup analysis): • Depression symptoms change score • Response
Study 62b (FDA) RCT Country NR	Secondary care N=356 Baseline severity: More severe Mean age (years): 40 Sex (% female): 57 Ethnicity (% BME): NR	Fluoxetine 20mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score
Study F1J-MC-HMAQ- Study Group Ba RCT US	Secondary care N=112 Baseline severity: More severe Mean age (years): 40.8 Sex (% female): NR Ethnicity (% BME): NR	Fluoxetine 20mg/day	Placebo	Treatment duration (weeks): 10 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score • Response
Tollefson 1993/1995 RCT US	Secondary care N=671 Baseline severity: More severe Mean age (years): 67.7 Sex (% female): 55 Ethnicity (% BME): 6	Fluoxetine maximum 20mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score • Response
Valle-Cabrera 2018 RCT Cuba	Secondary care N=77 Baseline severity: More severe Mean age (years): 45.2	Sertraline 50-200mg/day	Placebo	Treatment duration (weeks): 10 Outcomes (for primary versus secondary care

Study	Population	Intervention	Comparison	Comments
	Sex (% female): 92 Ethnicity (% BME): NR			subgroup analysis): • Response
VEN XR 367 (FDA)b RCT Europe	Secondary care N=164 Baseline severity: More severe Mean age (years): NR Sex (% female): 61 Ethnicity (% BME): NR	Paroxetine 20mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score
Wang 2014c RCT Canada, China, Finland, South Korea, Malaysia, Mexico, The Philippines, South Africa, & Spain	Secondary care N=314 Baseline severity: More severe Mean age (years): 40 Sex (% female): 71 Ethnicity (% BME): 46	Escitalopram 10-20mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Response
WELL AK1A4006 RCT US	Secondary care N=309 Baseline severity: More severe Mean age (years): 37.9 Sex (% female): NR Ethnicity (% BME): NR	Fluoxetine 20-60mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score • Response
Wernicke 1987 RCT US	Secondary care N=356 Baseline severity: More severe Mean age (years): 39.8 Sex (% female): 57 Ethnicity (% BME): NR	Fluoxetine 20mg/day, 40mg/day, or 60mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score • Response

Study	Population	Intervention	Comparison	Comments
Wernicke 1988 RCT US	Secondary care N=267 Baseline severity: More severe Mean age (years): NR Sex (% female): NR Ethnicity (% BME): NR	Fluoxetine 20mg/day or 40mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score • Response

a Three-armed trial but where possible the demographics reported here are for only the two relevant arms.
b Four-armed trial but where possible the demographics reported here are for only the two relevant arms.
BME: black, minority, ethnic; K: number of studies; mg: milligrams; N: number of participants; NR: not reported; RCT: randomised controlled trial.

Summaries of the studies included in the primary care versus secondary care subgroup analysis of the SSRIs versus tricyclic antidepressants (TCAs) comparison are presented in Table 4.

There were no significant subgroup differences between primary care and secondary care for the comparison SSRIs versus TCAs on: depression symptoms endpoint (Test for subgroup differences: $Chi^2 = 0.09$, df = 1, p = 0.76); depression symptoms change score (Test for subgroup differences: $Chi^2 = 1.46$, df = 1, p = 0.23); remission (Test for subgroup differences: $Chi^2 = 2.19$, df = 1, p = 0.14); response (Test for subgroup differences: $Chi^2 = 2.22$, df = 1, p = 0.14).

Table 14: Summary of included studies for primary care versus secondary care subgroup analysis for comparison 1c SSRIs versus tricyclic antidepressants (TCAs)

Study	Population	Intervention	Comparison	Comments
Primary care (K=10, N=2,014)				
Christiansen 1996 RCT Denmark	Primary care N=144 Baseline severity: More severe Mean age (years): NR Sex (% female): NR Ethnicity (% BME): NR	Paroxetine 20-40mg/day	Amitriptyline 75-150mg/day	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint • Response
Freed 1999 RCT Australia	Primary care N=375 Baseline severity: More severe Mean age (years): 48	Paroxetine 20mg/day	Amitriptyline 75mg/day	Treatment duration (weeks): 9 Outcomes (for primary versus secondary care

Study	Population	Intervention	Comparison	Comments
	Sex (% female): 65 Ethnicity (% BME): NR			subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score
Hutchinson 1992 RCT UK	Primary care N=90 Baseline severity: More severe Mean age (years): 71.8 Sex (% female): 77 Ethnicity (% BME): NR	Paroxetine 30mg/day	Amitriptyline 100mg/day	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Remission • Response
Kyle 1998 RCT UK	Primary care N=365 Baseline severity: More severe Mean age (years): 73.8 Sex (% female): 73 Ethnicity (% BME): NR	Citalopram 20-40mg/day	Amitriptyline 50-100mg/day	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Remission
Moon 1994 RCT UK	Primary care N=106 Baseline severity: More severe Mean age (years): 43.7 Sex (% female): 52 Ethnicity (% BME): NR	Sertraline 50-150mg/day	Clomipramine 50-150mg/day	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Response
Moon 1996 RCT UK	Primary care N=138 Baseline severity: More severe Mean age (years): 43.7 Sex (% female): 71 Ethnicity (% BME): NR	Paroxetine 20-30mg/day	Lofepramine 70-210mg/day	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Remission • Response

Study	Population	Intervention	Comparison	Comments
PAR 29060/281 RCT Europe	Primary care N=162 Baseline severity: More severe Mean age (years): 38.8 Sex (% female): 77 Ethnicity (% BME): NR	Paroxetine 30mg/day	Amitriptyline 75-150mg/day	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint
PAR MDUK 032 RCT Country NR	Primary care N=59 Baseline severity: More severe Mean age (years): 44.4 Sex (% female): NR Ethnicity (% BME): NR	Paroxetine 20-30mg/day	Amitriptyline 100-150mg/day	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint
Rosenberg 1994 RCT Denmark, Norway, Sweden & Finland	Primary care N=472 Baseline severity: More severe Mean age (years): 47.6 Sex (% female): 69 Ethnicity (% BME): NR	Citalopram 10-30mg/day or 20-60mg/day	Imipramine 50-150mg/day	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Response
Serrano-Blanco 2006 RCT Spain	Primary care N=103 Baseline severity: More severe Mean age (years): 43.5 Sex (% female): 73 Ethnicity (% BME): NR	Fluoxetine 10-40mg/day	Imipramine 25-125mg/day	Treatment duration (weeks): 24 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score
Secondary care (K=47, N=5,482)				

Study	Population	Intervention	Comparison	Comments
29060 07 001a RCT US	Secondary care N=26 Baseline severity: More severe Mean age (years): 42.3 Sex (% female): 65 Ethnicity (% BME): NR	Paroxetine 10-60mg/day	Amitriptyline (dose NR)	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score
29060/299 RCT Europe	Secondary care N=217 Baseline severity: More severe Mean age (years): 40.4 Sex (% female): NR Ethnicity (% BME): NR	Paroxetine 20-50mg/day	Amitriptyline 100-250mg/day	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score
Akhondzadeh 2003 RCT Iran	Secondary care N=48 Baseline severity: More severe Mean age (years): 35.8 Sex (% female): 40 Ethnicity (% BME): NR	Fluoxetine 60mg/day	Nortriptyline 150mg/day	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score
Arminem 1992 RCT Finland	Secondary care N=57 Baseline severity: More severe Mean age (years): NR Sex (% female): 54 Ethnicity (% BME): NR	Paroxetine 20-40mg/day	Imipramine 100-200mg/day	Treatment duration (weeks): 12 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint
Beasley 1993b RCT US	Secondary care N=136 Baseline severity: More severe Mean age (years): 44.8 Sex (% female): 70	Fluoxetine 40-80mg/day	Amitriptyline 150-300mg/day	Treatment duration (weeks): 5 Outcomes (for primary versus secondary care

Study	Population	Intervention	Comparison	Comments
	Ethnicity (% BME): 4			subgroup analysis): • Depression symptoms change score • Remission • Response
Bersani 1994 RCT Italy	Secondary care N=68 Baseline severity: More severe Mean age (years): 47.1 Sex (% female): 63 Ethnicity (% BME): NR	Sertraline 50-150mg/day	Amitriptyline 50-150mg/day	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score
Bhargava 2012 RCT India	Secondary care N=60 Baseline severity: More severe Mean age (years): 36.2 Sex (% female): 52 Ethnicity (% BME): NR	Sertraline 50-150mg/day	Imipramine 75-150mg/day	Treatment duration (weeks): 12 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score
Bremner 1984 RCT US	Secondary care N=40 Baseline severity: More severe Mean age (years): 42.6 Sex (% female): 51 Ethnicity (% BME): NR	Fluoxetine 60-80mg/day	Imipramine 125-300mg/day	Treatment duration (weeks): 5 Outcomes (for primary versus secondary care subgroup analysis): • Response
Byerley 1988a RCT US	Secondary care N=66 Baseline severity: More severe Mean age (years): 39.1	Fluoxetine 40-80mg/day	Imipramine 150-300mg/day	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care

Study	Population	Intervention	Comparison	Comments
	Sex (% female): 68 Ethnicity (% BME): NR			subgroup analysis): • Depression symptoms endpoint • Response
Chiu 1996 RCT Taiwan	Secondary care N=40 Baseline severity: More severe Mean age (years): 45.7 Sex (% female): 63 Ethnicity (% BME): NR	Paroxetine 20-30mg/day	Imipramine 125-150mg/day	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score • Response
Cohn 1984b RCT US	Secondary care N=66 Baseline severity: More severe Mean age (years): 42 Sex (% female): NR Ethnicity (% BME): NR	Fluoxetine (dose NR)	Imipramine (dose NR)	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint
Cohn 1990b RCT US	Secondary care N=241 Baseline severity: More severe Mean age (years): 70.3 Sex (% female): 49 Ethnicity (% BME): NR	Sertraline 50-200mg/day	Amitriptyline 50-150mg/day	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score • Response
Danish University Antidepressant Group 1986 RCT Denmark	Secondary care N=114 Baseline severity: More severe Mean age (years): NR Sex (% female): 70	Citalopram 40mg/day	Clomipramine 150mg/day	Treatment duration (weeks): 5 **Outcomes (for primary versus secondary care subgroup analysis):**

Study	Population	Intervention	Comparison	Comments
	Ethnicity (% BME): NR			• Remission
Danish University Antidepressant Group 1990 RCT Denmark	Secondary care N=120 Baseline severity: More severe Mean age (years): NR Sex (% female): 66 Ethnicity (% BME): NR	Paroxetine 30mg/day	Clomipramine 150mg/day	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Remission
De Ronchi 1998 RCT Italy	Secondary care N=65 Baseline severity: More severe Mean age (years): 68.9 Sex (% female): 72 Ethnicity (% BME): NR	Fluoxetine 20mg/day	Amitriptyline 50-100mg/day	Treatment duration (weeks): 10 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score • Response
Demyttenaere 1998 RCT Belgium	Secondary care N=66 Baseline severity: More severe Mean age (years): 41.7 Sex (% female): 55 Ethnicity (% BME): NR	Fluoxetine 20mg/day	Amitriptyline 50mg/day	Treatment duration (weeks): 9 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score • Response
Deuschle 2003 RCT Germany	Secondary care N=126 Baseline severity: More severe Mean age (years): 54.1 Sex (% female): 67 Ethnicity (% BME): NR	Paroxetine 40mg/day	Amitriptyline 150mg/day	Treatment duration (weeks): 5 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint

Study	Population	Intervention	Comparison	Comments
				• Depression symptoms change score
Fabre 1991 RCT US	Secondary care N=205 Baseline severity: More severe Mean age (years): 37 Sex (% female): 57 Ethnicity (% BME): NR	Fluoxetine 40mg/day	Nortriptyline 100mg/day	Treatment duration (weeks): 5 Outcomes (for primary versus secondary care subgroup analysis): • Response
Fabre 1992a RCT US	Secondary care N=80 Baseline severity: More severe Mean age (years): 35.4 Sex (% female): 61 Ethnicity (% BME): NR	Paroxetine 10-50mg/day	Imipramine 65-275mg/day	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score
Fawcett 1989 RCT US	Secondary care N=40 Baseline severity: More severe Mean age (years): 42.2 Sex (% female): 65 Ethnicity (% BME): NR	Fluoxetine 20-60mg/day	Amitriptyline 50-200mg/day	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score • Remission • Response
Feighner 1993a RCT US	Secondary care N=477 Baseline severity: More severe Mean age (years): 40.4 Sex (% female): 53 Ethnicity (% BME): NR	Paroxetine 10-50mg/day	Imipramine 65-275mg/day	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Remission

Study	Population	Intervention	Comparison	Comments
Forlenza 2001 RCT Brazil	Secondary care N=55 Baseline severity: More severe Mean age (years): 68.5 Sex (% female): 69 Ethnicity (% BME): NR	Sertraline 50mg/day	Imipramine 150mg/day	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score • Remission • Response
Geretsegger 1995 RCT Austria & Germany	Secondary care N=91 Baseline severity: More severe Mean age (years): 71.2 Sex (% female): 86 Ethnicity (% BME): NR	Paroxetine 20-30mg/day	Amitriptyline 100-150mg/day	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Remission • Response
GSK_29060/103 RCT UK	Secondary care N=106 Baseline severity: More severe Mean age (years): 75.3 Sex (% female): 74 Ethnicity (% BME): NR	Paroxetine 20-30mg/day	Lofepramine 70-210mg/day	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score • Response
Hashemi 2012 RCT Iran	Secondary care N=120 Baseline severity: More severe Mean age (years): 34.8 Sex (% female): 53 Ethnicity (% BME): NR	Fluoxetine maximum 60mg/day	Nortriptyline maximum 150mg/day	Treatment duration (weeks): 26 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint

Study	Population	Intervention	Comparison	Comments
				• Depression symptoms change score
Keegan 1991 RCT Canada	Secondary care N=42 Baseline severity: More severe Mean age (years): 43.8 Sex (% female): NR Ethnicity (% BME): NR	Fluoxetine 20-80mg/day	Amitriptyline 100-250mg/day	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Remission • Response
Laakmann 1988 RCT Germany	Secondary care N=128 Baseline severity: More severe Mean age (years): NR Sex (% female): 72 Ethnicity (% BME): NR	Fluoxetine 20-60mg/day	Amitriptyline 50-150mg/day	Treatment duration (weeks): 5 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint • Response
Laakmann 1991 RCT Germany	Secondary care N=174 Baseline severity: More severe Mean age (years): NR Sex (% female): NR Ethnicity (% BME): NR	Fluoxetine (dose NR)	Amitriptyline 100-200mg/day	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint
Levine 1989 RCT UK	Secondary care N=60 Baseline severity: More severe Mean age (years): 45.8 Sex (% female): 70 Ethnicity (% BME): NR	Fluoxetine 40-60mg/day	Imipramine 75-150mg/day	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Remission
Marchesi 1998 RCT Italy	Secondary care N=142 Baseline severity: More severe Mean age (years): 43.6	Fluoxetine 20mg/day	Amitriptyline 75-225mg/day	Treatment duration (weeks): 10 Outcomes (for primary versus secondary care

Study	Population	Intervention	Comparison	Comments
	Sex (% female): 74 Ethnicity (% BME): NR			subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score • Response
MDF/29060/III/070/88/MC RCT Europe	Secondary care N=62 Baseline severity: More severe Mean age (years): 73 Sex (% female): NR Ethnicity (% BME): NR	Paroxetine 20-30mg/day	Clomipramine 60-75mg/day	Treatment duration (weeks): 5 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score • Remission • Response
Miura 2000 RCT Japan	Secondary care N=228 Baseline severity: More severe Mean age (years): 46.5 Sex (% female): NR Ethnicity (% BME): NR	Paroxetine 20-40mg/day	Amitriptyline 50-150mg/day	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score
Moller 1993 RCT Germany and Hungary	Secondary care N=223 Baseline severity: More severe Mean age (years): 47.1 Sex (% female): NR Ethnicity (% BME): NR	Paroxetine 30-50mg/day	Amitriptyline 150-250mg/day	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score • Remission • Response

Study	Population	Intervention	Comparison	Comments
Moller 1998 RCT Germany, Hungary, & Czech Republic	Secondary care N=160 Baseline severity: More severe Mean age (years): 48.6 Sex (% female): 70 Ethnicity (% BME): NR	Sertraline 50-150mg/day	Amitriptyline 75-150mg/day	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score • Response
Mulsant 1999 RCT US	Secondary care N=80 Baseline severity: More severe Mean age (years): 65 Sex (% female): 74 Ethnicity (% BME): 15	Paroxetine 20-30mg/day	Nortriptyline (Mean dose 51.4mg/day)	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score • Remission
Navarro 2001 RCT Spain	Secondary care N=58 Baseline severity: More severe Mean age (years): 70.7 Sex (% female): 64 Ethnicity (% BME): NR	Citalopram 30-40mg/day	Nortriptyline 50-100mg/day	Treatment duration (weeks): 12 Outcomes (for primary versus secondary care subgroup analysis): • Remission
Ontiveros Sanchez 1998 RCT South America	Secondary care N=42 Baseline severity: More severe Mean age (years): 37.6 Sex (% female): 53 Ethnicity (% BME): NR	Fluoxetine 20mg/day	Amitriptyline 150-250mg/day	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint • Response

Study	Population	Intervention	Comparison	Comments
Peselow 1989aa RCT US	Secondary care N=66 Baseline severity: More severe Mean age (years): 45.9 Sex (% female): 35 Ethnicity (% BME): NR	Paroxetine 10-50mg/day	Imipramine 65-275mg/day	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Response
Peselow 1989ba RCT US	Secondary care N=80 Baseline severity: More severe Mean age (years): NR Sex (% female): NR Ethnicity (% BME): NR	Paroxetine 20-50mg/day	Imipramine 65-275mg/day	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Response
Peters 1990 RCT Germany	Secondary care N=102 Baseline severity: More severe Mean age (years): 44.5 Sex (% female): 63 Ethnicity (% BME): NR	Fluoxetine 20mg/day	Amitriptyline 100mg/day	Treatment duration (weeks): 5 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint • Response
Preskorn 1991 RCT US	Secondary care N=61 Baseline severity: More severe Mean age (years): NR Sex (% female): NR Ethnicity (% BME): 2	Fluoxetine 20-60mg/day	Amitriptyline 50-200mg/day	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score
Reimherr 1990a RCT US & Canada	Secondary care N=298 Baseline severity: More severe Mean age (years): 38.4 Sex (% female): 55	Sertraline 20-200mg/day	Amitriptyline 50-150mg/day	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care

Study	Population	Intervention	Comparison	Comments
	Ethnicity (% BME): 10			subgroup analysis): • Depression symptoms change score • Response
Ropert 1989 RCT France	Secondary care N=143 Baseline severity: More severe Mean age (years): 43.8 Sex (% female): 64 Ethnicity (% BME): NR	Fluoxetine	Clomipramine	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score
SER 315 (FDA)a RCT Europe	Secondary care N=162 Baseline severity: More severe Mean age (years): 42.4 Sex (% female): 69 Ethnicity (% BME): NR	Sertraline 50-200mg/day	Amitriptyline 50-200mg/day	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score
Staner 1995 RCT Belgium	Secondary care N=40 Baseline severity: More severe Mean age (years): 42.1 Sex (% female): 83 Ethnicity (% BME): NR	Paroxetine 30mg/day	Amitriptyline 150mg/day	Treatment duration (weeks): 5 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score • Response
Stark 1985a RCT US	Secondary care N=371 Baseline severity: More severe Mean age (years): 41.0	Fluoxetine 60-80mg/day	Imipramine 100-300mg/day	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care

Study	Population	Intervention	Comparison	Comments
	Sex (% female): 69 Ethnicity (% BME): NR			subgroup analysis): • Depression symptoms change score • Response
Suleman 1997 RCT Zimbabwe	Secondary care N=30 Baseline severity: More severe Mean age (years): NR Sex (% female): NR Ethnicity (% BME): NR	Fluoxetine 20mg/day	Amitriptyline 100mg/day	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score

a Three-armed trial but where possible the demographics reported here are for only the two relevant arms.
BME: black, minority, ethnic; K: number of studies; mg: milligrams; N: number of participants; NR: not reported; RCT: randomised controlled trial.

Summaries of the studies included in the primary care versus secondary care subgroup analysis of the TCAs versus placebo comparison are presented in Table 5.

There were no significant subgroup differences between primary care and secondary care for the comparison TCAs versus placebo on: depression symptoms endpoint (Test for subgroup differences: $Chi^2 = 0.49$, df = 1, p = 0.49); depression symptoms change score (Test for subgroup differences: $Chi^2 = 0.32$, df = 1, p = 0.57); response (Test for subgroup differences: $Chi^2 = 2.87$, df = 1, p = 0.09).

Table 15: Summary of included studies for primary care versus secondary care subgroup analysis for comparison 1d TCAs versus placebo

Study	Population	Intervention	Comparison	Comments
Primary care (K=6, N=597)				
Barge-Schaapveld 2002 RCT Netherlands	Primary care N=63 Baseline severity: More severe Mean age (years): 43.4 Sex (% female): 73 Ethnicity (% BME): NR	Imipramine 200mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint

Study	Population	Intervention	Comparison	Comments
Blashki 1971 RCT Australia	Primary care N=45 Baseline severity: More severe Mean age (years): 36.7 Sex (% female): 100 Ethnicity (% BME): NR	Amitriptyline 75mg/day or 150mg/day	Placebo	Treatment duration (weeks): 4 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score
Lecrubier 1997a RCT France, Italy & UK	Primary care N=151 Baseline severity: More severe Mean age (years): 40.6 Sex (% female): 66 Ethnicity (% BME): NR	Imipramine 75-150mg/day	Placebo	Treatment duration (weeks): 13 Outcomes (for primary versus secondary care subgroup analysis): • Response
Mynors-Wallis 1995a RCT UK	Primary care N=61 Baseline severity: More severe Mean age (years): 37.1 Sex (% female): 74 Ethnicity (% BME): 5	Amitriptyline maximum 150mg/day	Placebo	Treatment duration (weeks): 12 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score
Philipp 1999 RCT Germany	Primary care N=157 Baseline severity: More severe Mean age (years): 46.5 Sex (% female): 75 Ethnicity (% BME): NR	Imipramine 100mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score • Response

Study	Population	Intervention	Comparison	Comments
Schweizer 1998 RCT US	Primary care N=120 Baseline severity: More severe Mean age (years): NR Sex (% female): NR Ethnicity (% BME): NR	Imipramine 50-150mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Response
Secondary care (K=30, N=3,444)				
29060 07 001a RCT US	Secondary care N=25 Baseline severity: More severe Mean age (years): 44.8 Sex (% female): 52 Ethnicity (% BME): NR	Amitriptyline (dose NR)	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score
Amsterdam 1986 RCT US	Secondary care N=109 Baseline severity: More severe Mean age (years): 41 Sex (% female): 33 Ethnicity (% BME): NR	Amitriptyline 200-600mg/day	Placebo	Treatment duration (weeks): 4 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score • Response
Bakish 1992b RCT Canada	Secondary care N=115 Baseline severity: More severe Mean age (years): 43 Sex (% female): 43 Ethnicity (% BME): NR	Amitriptyline 150mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Response
Bremner 1995a RCT US	Secondary care N=100 Baseline severity: More severe	Amitriptyline 40-280mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care

Study	Population	Intervention	Comparison	Comments
	Mean age (years): 38.0 Sex (% female): 72 Ethnicity (% BME): NR			subgroup analysis): • Response
Byerley 1988a RCT US	Secondary care N=63 Baseline severity: More severe Mean age (years): 38.5 Sex (% female): 61 Ethnicity (% BME): NR	Imipramine 150-300mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint • Response
Cassano 1986 RCT US, Canada, UK, & France	Secondary care N=314 Baseline severity: More severe Mean age (years): 41.7 Sex (% female): 62 Ethnicity (% BME): NR	Imipramine 50-300mg/day	Placebo	Treatment duration (weeks): 4 Outcomes (for primary versus secondary care subgroup analysis): • Response
Elkin 1989/Imber 1990b RCT US	Secondary care N=125 Baseline severity: More severe Mean age (years): 35 Sex (% female): 70 Ethnicity (% BME): 11	Imipramine (mean final dose 185mg/day)	Placebo	Treatment duration (weeks): 16 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score
Escobar 1980a RCT Colombia	Secondary care N=27 Baseline severity: More severe Mean age (years): 46.1 Sex (% female): 59 Ethnicity (% BME): NR	Imipramine 100-300mg/day	Placebo	Treatment duration (weeks): 4 Outcomes (for primary versus secondary care subgroup analysis): • Response

Study	Population	Intervention	Comparison	Comments
Fabre 1992a RCT US	Secondary care N=80 Baseline severity: More severe Mean age (years): 35.5 Sex (% female): 66 Ethnicity (% BME): NR	Imipramine 65-275mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score
Feiger 1996 RCT US	Secondary care N=81 Baseline severity: More severe Mean age (years): 39.7 Sex (% female): 56 Ethnicity (% BME): 11	Imipramine 50-300mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Response
Feighner 1982 RCT US	Secondary care N=139 Baseline severity: More severe Mean age (years): NR Sex (% female): 71 Ethnicity (% BME): NR	Lofepramine 105-280mg/day or Imipramine 75-200mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint • Response
Feighner 1989b RCT US	Secondary care N=30 Baseline severity: More severe Mean age (years): 44 Sex (% female): 50 Ethnicity (% BME): NR	Imipramine 50-250mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Response
Fontaine 1994 RCT Canada	Secondary care N=90 Baseline severity: More severe Mean age (years): 43.1 Sex (% female): 58	Imipramine 50-250mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Response

Study	Population	Intervention	Comparison	Comments
	Ethnicity (% BME): NR			
Goldberg 1980a RCT US	Secondary care N=122 Baseline severity: More severe Mean age (years): 36.1 Sex (% female): 74 Ethnicity (% BME): NR	Amitriptyline 75-200mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Response
Kusalic 1993 RCT Canada	Secondary care N=28 Baseline severity: More severe Mean age (years): NR Sex (% female): NR Ethnicity (% BME): NR	Amitriptyline (mean final dose 109.93mg/day)	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Response
McCallum 1975 RCT US	Secondary care N=24 Baseline severity: More severe Mean age (years): 41.5 Sex (% female): 83 Ethnicity (% BME): NR	Amitriptyline 150mg/day	Placebo	Treatment duration (weeks): 3 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score
MIR 003-020 (FDA)a RCT US	Secondary care N=86 Baseline severity: More severe Mean age (years): 44.0 Sex (% female): 55 Ethnicity (% BME): NR	Amitriptyline 40-280mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score • Response

Study	Population	Intervention	Comparison	Comments
Peselow 1989aa RCT US	Secondary care N=71 Baseline severity: More severe Mean age (years): 44.7 Sex (% female): 35 Ethnicity (% BME): NR	Imipramine 65-275mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Response
Peselow 1989ba RCT US	Secondary care N=82 Baseline severity: More severe Mean age (years): NR Sex (% female): NR Ethnicity (% BME): NR	Imipramine 65-275mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Response
Reimherr 1990a RCT US & Canada	Secondary care N=299 Baseline severity: More severe Mean age (years): 39.0 Sex (% female): 54 Ethnicity (% BME): 9	Amitriptyline 50-150mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score • Response
Rickels 1982e RCT US	Secondary care N=97 Baseline severity: More severe Mean age (years): NR Sex (% female): NR Ethnicity (% BME): NR	Imipramine 150-200mg/day	Placebo	Treatment duration (weeks): 4 Outcomes (for primary versus secondary care subgroup analysis): • Response
Rickels 1991 RCT US	Secondary care N=131 Baseline severity: More severe Mean age (years): NR Sex (% female): NR Ethnicity (% BME): NR	Imipramine minimum 150mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Response

Study	Population	Intervention	Comparison	Comments
Rickels 1995_Study 006-1 RCT US	Secondary care N=77 Baseline severity: More severe Mean age (years): NR Sex (% female): NR Ethnicity (% BME): NR	Imipramine 100-300mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Response
Rickels 1995_Study 006-2 RCT US	Secondary care N=80 Baseline severity: More severe Mean age (years): NR Sex (% female): NR Ethnicity (% BME): NR	Imipramine 100-300mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Response
Schweizer 1994a RCT US	Secondary care N=151 Baseline severity: More severe Mean age (years): 42.5 Sex (% female): 64 Ethnicity (% BME): NR	Imipramine 75-225mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score • Response
SER 315 (FDA)a RCT Europe	Secondary care N=157 Baseline severity: More severe Mean age (years): 43.5 Sex (% female): 75 Ethnicity (% BME): NR	Amitriptyline 50-200mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score
Silverstone 1994 RCT UK	Secondary care N=166 Baseline severity: More severe Mean age (years): NR Sex (% female): NR	Imipramine 150mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care

Study	Population	Intervention	Comparison	Comments
	Ethnicity (% BME): NR			subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score • Response
Smith 1990a RCT US	Secondary care N=100 Baseline severity: More severe Mean age (years): NR Sex (% female): NR Ethnicity (% BME): NR	Amitriptyline 80-280mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Response
Stark 1985a RCT US	Secondary care N=355 Baseline severity: More severe Mean age (years): 41.5 Sex (% female): 68 Ethnicity (% BME): NR	Imipramine 100-300mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score • Response
White 1984 RCT US	Secondary care N=120 Baseline severity: More severe Mean age (years): 37.1 Sex (% female): 48 Ethnicity (% BME): NR	Nortriptyline 75-150mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Depression symptoms change score

a Three-armed trial but where possible the demographics reported here are for only the two relevant arms.
BME: black, minority, ethnic; K: number of studies; mg: milligrams; N: number of participants; NR: not reported; RCT: randomised controlled trial.

Summaries of the studies included in the primary care versus secondary care subgroup analysis of the serotonin–norepinephrine reuptake inhibitors (SNRIs) versus SSRIs comparison are presented in Table 6.

There were no significant subgroup differences between primary care and secondary care for the comparison SNRIs versus SSRIs on: remission (Test for subgroup

differences: Chi² = 1.55, df = 1, p = 0.21); response (Test for subgroup differences: Chi² = 0.62, df = 1, p = 0.43).

Table 16: Summary of included studies for primary care versus secondary care subgroup analysis for comparison 1e Serotonin–norepinephrine reuptake inhibitors (SNRIs) versus SSRIs

Study	Population	Intervention	Comparison	Comments
Primary care (K=2, N=634)				
Montgomery 2004 RCT Denmark, Finland, France, Germany, Ireland, Spain, & Switzerland	Primary care N=293 Baseline severity: More severe Mean age (years): 48 Sex (% female): 71 Ethnicity (% BME): NR	Venlafaxine 75-150mg/day	Escitalopram 10-20mg/day	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Remission • Response
Tylee 1997 RCT UK	Primary care N=341 Baseline severity: More severe Mean age (years): 44.5 Sex (% female): 71 Ethnicity (% BME): NR	Venlafaxine 75mg/day	Fluoxetine 20mg/day	Treatment duration (weeks): 12 Outcomes (for primary versus secondary care subgroup analysis): • Remission • Response
Secondary care (K=29, N=5,484)				
Allard 2004 RCT Sweden & Denmark	Secondary care N=151 Baseline severity: More severe Mean age (years): 73 Sex (% female): 80 Ethnicity (% BME): NR	Venlafaxine 75-150mg/day	Citalopram 10-20mg/day	Treatment duration (weeks): 22 Outcomes (for primary versus secondary care subgroup analysis): • Remission • Response
Alves 1999 RCT Portugal	Secondary care N=87 Baseline severity: More severe Mean age (years): 43.7 Sex (% female): 92 Ethnicity (% BME): NR	Venlafaxine 75-150mg/day	Fluoxetine 20-40mg/day	Treatment duration (weeks): 12 Outcomes (for primary versus secondary care subgroup analysis): • Remission • Response

Study	Population	Intervention	Comparison	Comments
Bielski 2004 RCT US	Secondary care N=202 Baseline severity: More severe Mean age (years): 37.4 Sex (% female): 58 Ethnicity (% BME): 25	Venlafaxine 225mg/day	Escitalopram 20mg/day	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Remission • Response
Clerc 1994 RCT France & Belgium	Secondary care N=68 Baseline severity: More severe Mean age (years): 51.3 Sex (% female): 68 Ethnicity (% BME): NR	Venlafaxine 200mg/day	Fluoxetine 40mg/day	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Response
Costa 1998 RCT Argentina, Brazil, Chile, Colombia, Uruguay, & Venezuela	Secondary care N=382 Baseline severity: More severe Mean age (years): 40.2 Sex (% female): 79 Ethnicity (% BME): NR	Venlafaxine 75-150mg/day	Fluoxetine 20-40mg/day	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Remission • Response
DeNayer 2002 RCT Belgium	Secondary care N=146 Baseline severity: More severe Mean age (years): 42.8 Sex (% female): 68 Ethnicity (% BME): NR	Venlafaxine 75-150mg/day	Fluoxetine 20-40mg/day	Treatment duration (weeks): 12 Outcomes (for primary versus secondary care subgroup analysis): • Remission • Response
Detke 2004a RCT US	Secondary care N=274 Baseline severity: More severe Mean age (years): 43.3 Sex (% female): 72 Ethnicity (% BME): 0	Duloxetine 80mg/day or 120mg/day	Paroxetine 20mg/day	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Remission • Response

Study	Population	Intervention	Comparison	Comments
Diaz-Martinez 1998 RCT Mexico	Secondary care N=145 Baseline severity: More severe Mean age (years): NR Sex (% female): 72 Ethnicity (% BME): NR	Venlafaxine 75-150mg/day	Fluoxetine 20-40mg/day	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Response
Dierick 1996 RCT Belgium, Italy, Switzerland & France	Secondary care N=314 Baseline severity: More severe Mean age (years): 43.4 Sex (% female): 65 Ethnicity (% BME): NR	Venlafaxine 75-150mg/day	Fluoxetine 20mg/day	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Response
Eli Lilly HMAT-Aa RCT US	Secondary care N=173 Baseline severity: More severe Mean age (years): NR Sex (% female): NR Ethnicity (% BME): NR	Duloxetine 80mg/day	Paroxetine 20mg/day	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Remission • Response
Goldstein 2002a RCT US	Secondary care N=103 Baseline severity: More severe Mean age (years): 41.5 Sex (% female): 61 Ethnicity (% BME): 17	Duloxetine 40-120mg/day	Fluoxetine 20mg/day	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Remission • Response
Goldstein 2004a RCT US	Secondary care N=178 Baseline severity: More severe Mean age (years): 40.5 Sex (% female): 63 Ethnicity (% BME): 21	Duloxetine 80mg/day	Paroxetine 20mg/day	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Remission • Response

Study	Population	Intervention	Comparison	Comments
Hao 2014 RCT China	Secondary care N=281 Baseline severity: More severe Mean age (years): 38.5 Sex (% female): 59 Ethnicity (% BME): NR	Duloxetine 60mg/day	Paroxetine 20mg/day	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Remission • Response
Higuchi 2009a RCT Japan	Secondary care N=223 Baseline severity: More severe Mean age (years): 38.3 Sex (% female): NR Ethnicity (% BME): NR	Duloxetine 60mg/day	Paroxetine 20-40mg/day	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Remission • Response
Hwang 2004 RCT Taiwan	Secondary care N=105 Baseline severity: More severe Mean age (years): 65.1 Sex (% female): 58 Ethnicity (% BME): NR	Venlafaxine 75-150mg/day	Paroxetine 20-40mg/day	Treatment duration (weeks): 4 Outcomes (for primary versus secondary care subgroup analysis): • Response
Jiang 2017 RCT China	Secondary care N=26 Baseline severity: More severe Mean age (years): 45.5 Sex (% female): 73 Ethnicity (% BME): NR	Duloxetine (mean final dose 60mg/day)	Escitalopram (mean final dose 13.13mg/day)	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Response
Khan 2007 RCT US	Secondary care N=278 Baseline severity: More severe Mean age (years): 42.4 Sex (% female): 61 Ethnicity (% BME): 20	Duloxetine 60mg/day	Escitalopram 10-20mg/day	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Remission • Response

Study	Population	Intervention	Comparison	Comments
Kornaat 2000 RCT Country NR	Secondary care N=156 Baseline severity: More severe Mean age (years): NR Sex (% female): 64 Ethnicity (% BME): NR	Venlafaxine 75-225mg/day	Fluoxetine 20-40mg/day	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Remission • Response
Mehtonen 2000 RCT Finland	Secondary care N=147 Baseline severity: More severe Mean age (years): 42.6 Sex (% female): 66 Ethnicity (% BME): NR	Venlafaxine 75-150mg/day	Sertraline 50-100mg/day	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Remission • Response
Nemeroff 2007a RCT US	Secondary care N=206 Baseline severity: More severe Mean age (years): 39 Sex (% female): 65 Ethnicity (% BME): 11	Venlafaxine 75-225mg/day	Fluoxetine 20-60mg/day	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Remission • Response
Nierenberg 2007a RCT US	Secondary care N=547 Baseline severity: More severe Mean age (years): 42.2 Sex (% female): 66 Ethnicity (% BME): 24	Duloxetine 60mg/day	Escitalopram 10mg/day	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Remission • Response
Perahia 2006a RCT Bulgaria, Croatia, Hungary, Poland, Romania, Russia, & Slovakia	Secondary care N=293 Baseline severity: More severe Mean age (years): 45.4 Sex (% female): 71 Ethnicity (% BME): 0	Duloxetine 80mg/day or 120mg/day	Paroxetine 20mg/day	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Remission • Response

Study	Population	Intervention	Comparison	Comments
Rickels 2000 RCT Country NR	Secondary care N=51 Baseline severity: More severe Mean age (years): 37.4 Sex (% female): 75 Ethnicity (% BME): NR	Venlafaxine 150-225mg/day	Fluoxetine 20-40mg/day	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Remission
Rudolph 1999a RCT US	Secondary care N=203 Baseline severity: More severe Mean age (years): 40 Sex (% female): 72 Ethnicity (% BME): NR	Venlafaxine 75-225mg/day	Fluoxetine 20-60mg/day	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Remission • Response
Sheehan 2009ba RCT US	Secondary care N=194 Baseline severity: More severe Mean age (years): 39.7 Sex (% female): 59 Ethnicity (% BME): NR	Venlafaxine 225-375mg/day	Fluoxetine 60-80mg/day	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Remission • Response
Shelton 2006 RCT US	Secondary care N=160 Baseline severity: More severe Mean age (years): 39.3 Sex (% female): 53 Ethnicity (% BME): 17	Venlafaxine 75-225mg/day	Sertraline 50-150mg/day	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Remission • Response
Sir 2005 RCT Australia & Turkey	Secondary care N=163 Baseline severity: More severe Mean age (years): 37 Sex (% female): 69 Ethnicity (% BME): 2	Venlafaxine 75-225mg/day	Sertraline 50-150mg/day	Treatment duration (weeks): 8 Outcomes (for primary versus secondary care subgroup analysis): • Remission • Response

Study	Population	Intervention	Comparison	Comments
Study F1J-MC-HMAQ- Study Group Ba RCT US	Secondary care N=119 Baseline severity: More severe Mean age (years): 39.8 Sex (% female): NR Ethnicity (% BME): NR	Duloxetine 40-120mg/day	Fluoxetine 20mg/day	Treatment duration (weeks): 10 Outcomes (for primary versus secondary care subgroup analysis): • Remission • Response
Tzanakaki 2000 RCT Greece & Italy	Secondary care N=109 Baseline severity: More severe Mean age (years): 48 Sex (% female): 79 Ethnicity (% BME): NR	Venlafaxine 225mg/day	Fluoxetine 60mg/day	Treatment duration (weeks): 6 Outcomes (for primary versus secondary care subgroup analysis): • Remission • Response

a Three-armed trial but where possible the demographics reported here are for only the two relevant arms.
BME: black, minority, ethnic; K: number of studies; mg: milligrams; N: number of participants; NR: not reported; RCT: randomised controlled trial.

Comparison 2. Crisis resolution team care versus standard care (for adults with non-psychotic severe mental illness)

Summary of the study included in the crisis resolution team care and standard care comparison is presented in Table 7.

Table 17: Summary of included studies for comparison 2 Crisis resolution versus standard care

Study	Population	Intervention	Comparison	Comments
Johnson 2005 RCT UK	N=260 Non-psychotic severe mental illness Diagnosis: 25% schizophrenia or schizoaffective disorder; 10% bipolar affective disorder; 7% other psychosis; 30% unipolar depression; 13% personality disorder; 4% other non-psychotic disorder; 5% substance misuse only (data only reported for	Crisis resolution team augmented existing acute services and aimed to assess all patients and manage them at home if feasible. Staff were available 24 hours but on call from home after 10pm	Standard care included care from the inpatient unit, crisis houses, and community mental health teams	Outcomes assessed at 8 weeks and 6 months after crisis Outcomes: • Symptom severity (BPRS) 8 weeks after crisis • Admission as inpatient 6 months after crisis • Bed days in hospital 6 months after crisis • Patient satisfaction

Study	Population	Intervention	Comparison	Comments
	123/135 of experimental group so percentages do not add up to 100%) Mean age (years): 37.9 Sex (% female): 49 Ethnicity (% BME): 22			(CSQ-8) 8 weeks after crisis • Quality of life (MANSA) 8 weeks after crisis • Social functioning (LSP) 8 weeks after crisis • Social functioning (LSP) 6 months after crisis

BME: black, minority, ethnic; BPRS: brief psychiatric rating scale; CSQ-8: client satisfaction questionnaire - 8 item version; LSP: life skills profile; MANSA: Manchester short assessment of quality of life; N: number of participants; RCT: randomised controlled trial

Comparison 3. Inpatient versus outpatient settings

Summaries of the studies included in the inpatient versus outpatient subgroup analysis of the selective serotonin reuptake inhibitors (SSRIs) versus placebo comparison are presented in Table 8Table 18.

There were no significant subgroup differences between inpatient and outpatient settings for the comparison SSRIs versus placebo on: depression symptoms change score (Test for subgroup differences: Chi² = 2.47, df = 1, p = 0.12); response (Test for subgroup differences: Chi² = 0.11, df = 1, p = 0.74).

Table 18: Summary of included studies for inpatient versus outpatient subgroup analysis for comparison 3a Selective serotonin reuptake inhibitors (SSRIs) versus placebo

Study	Population	Intervention	Comparison	Comments
Inpatient setting (K=3, N=272)				
29060 07 001a RCT US	Inpatient N=25 Baseline severity: More severe Mean age (years): 42.5 Sex (% female): 56 Ethnicity (% BME): NR	Paroxetine 10-60mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score
Katz 2004 RCT US	Inpatient N=53 Baseline severity: More severe Mean age (years): NR	Paroxetine 20-60mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient

Study	Population	Intervention	Comparison	Comments
	Sex (% female): NR Ethnicity (% BME): NR			subgroup analysis): • Response
Sheehan 2009ba RCT US	Inpatient N=194 Baseline severity: More severe Mean age (years): 38.8 Sex (% female): 66 Ethnicity (% BME): NR	Fluoxetine 60-80mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score • Response
Outpatient setting (K=74, N=16,736)				
Baune 2018 RCT Estonia, Finland, Germany, & Lithuania	Outpatient N=104 Baseline severity: More severe Mean age (years): 45.7 Sex (% female): 64 Ethnicity (% BME): 2	Paroxetine 20mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score
Binnemann 2008 RCT US, Serbia and Montenegro, & the Russian Federation	Outpatient N=82 Baseline severity: More severe Mean age (years): 49 Sex (% female): 39 Ethnicity (% BME): NR	Sertraline 100mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score • Response
Bjerkenstedt 2005 RCT Sweden	Outpatient N=115 Baseline severity: More severe Mean age (years): 50.9 Sex (% female): 79 Ethnicity (% BME): 0	Fluoxetine 20mg/day	Placebo	Treatment duration (weeks): 4 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score • Response

Study	Population	Intervention	Comparison	Comments
Blumenthal 2007/Hoffman 2011b RCT US	Outpatient N=98 Baseline severity: More severe Mean age (years): 52 Sex (% female): 77 Ethnicity (% BME): 33	Sertraline 50-200mg/day	Placebo	Treatment duration (weeks): 16 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score
Bose 2008 RCT US	Outpatient N=267 Baseline severity: More severe Mean age (years): 68.3 Sex (% female): 59 Ethnicity (% BME): 11	Escitalopram 10-20mg/day	Placebo	Treatment duration (weeks): 12 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score • Response
Burke 2002 RCT US	Outpatient N=491 Baseline severity: More severe Mean age (years): 40.1 Sex (% female): 65 Ethnicity (% BME): NR	Escitalopram 10mg/day or 20mg/day, or citalopram 40mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score • Response
Byerley 1988a RCT US	Outpatient N=61 Baseline severity: More severe Mean age (years): 38.2 Sex (% female): 68 Ethnicity (% BME): NR	Fluoxetine 40-80mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient subgroup analysis): • Response
Claghorn 1992a RCT US	Outpatient N=59 Baseline severity: More severe Mean age (years): NR	Paroxetine 10-50mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient

Study	Population	Intervention	Comparison	Comments
	Sex (% female): NR Ethnicity (% BME): NR			subgroup analysis): • Depression symptoms change score
Claghorn 1992b RCT US	Outpatient N=72 Baseline severity: More severe Mean age (years): 35 Sex (% female): 32 Ethnicity (% BME): NR	Paroxetine 10-50mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score • Response
Clayton 2006_study 1 RCT US	Outpatient N=283 Baseline severity: More severe Mean age (years): 35 Sex (% female): 61 Ethnicity (% BME): 35	Escitalopram 10-20mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score • Response
Clayton 2006_study 2 RCT US	Outpatient N=286 Baseline severity: More severe Mean age (years): 36.5 Sex (% female): 56 Ethnicity (% BME): 27	Escitalopram 10-20mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score • Response
Detke 2004a RCT US	Outpatient N=179 Baseline severity: More severe Mean age (years): 42.9 Sex (% female): 71 Ethnicity (% BME): 0	Paroxetine 20mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score

Study	Population	Intervention	Comparison	Comments
				• Response
Doogan 1994 RCT UK	Outpatient N=200 Baseline severity: More severe Mean age (years): 45.7 Sex (% female): 68 Ethnicity (% BME): NR	Sertraline 50-100mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient subgroup analysis): • Response
Dube 2010 RCT India, US, Mexico & Romania	Outpatient N=200 Baseline severity: More severe Mean age (years): 36.5 Sex (% female): 44 Ethnicity (% BME): NR	Escitalopram 10-20mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score • Response
Dunbar 1993 RCT US	Outpatient N=341 Baseline severity: More severe Mean age (years): 41 Sex (% female): 51 Ethnicity (% BME): NR	Paroxetine 10-50mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient subgroup analysis): • Response
Eli Lilly HMAT-Aa RCT US	Outpatient N=179 Baseline severity: More severe Mean age (years): NR Sex (% female): NR Ethnicity (% BME): NR	Paroxetine 20mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score • Response
Emsley 2018 RCT Bulgaria, Estonia, Finland, France, Republic of Korea, Malaysia, Mexico, Poland,	Outpatient N=206 Baseline severity: More severe Mean age (years): 70.6 Sex (% female):	Escitalopram 10mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient

Study	Population	Intervention	Comparison	Comments
Romania, & Slovakia	75 Ethnicity (% BME): NR			subgroup analysis): • Depression symptoms change score • Response
Fabre 1992a RCT US	Outpatient N=80 Baseline severity: More severe Mean age (years): 35.8 Sex (% female): 59 Ethnicity (% BME): NR	Paroxetine 10-50mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score
Fava 1998a RCT US	Outpatient N=128 Baseline severity: More severe Mean age (years): 41.3 Sex (% female): 51 Ethnicity (% BME): NR	Paroxetine 20-50mg/day or fluoxetine 20-80mg/day	Placebo	Treatment duration (weeks): 12 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score • Response
Fava 2005 RCT US	Outpatient N=90 Baseline severity: More severe Mean age (years): 37.2 Sex (% female): 59 Ethnicity (% BME): NR	Fluoxetine 20mg/day	Placebo	Treatment duration (weeks): 12 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score
FDA 245 (EMD 68 843-010) RCT US	Outpatient N=191 Baseline severity: More severe Mean age (years): NR Sex (% female): NR Ethnicity (% BME): NR	Fluoxetine 20mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score

Study	Population	Intervention	Comparison	Comments
Forest Laboratories 2000 RCT US	Outpatient N=386 Baseline severity: More severe Mean age (years): 42 Sex (% female): 52 Ethnicity (% BME): NR	Escitalopram 10-20mg/day or citalopram 20-40mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score • Response
Forest Research Institute 2005 RCT US	Outpatient N=409 Baseline severity: More severe Mean age (years): 40 Sex (% female): 56 Ethnicity (% BME): NR	Escitalopram 10-20mg/day or sertraline 50-200mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score • Response
Golden 2002_448 RCT US	Outpatient N=315 Baseline severity: More severe Mean age (years): 39 Sex (% female): NR Ethnicity (% BME): NR	Paroxetine 20-62.5mg/day	Placebo	Treatment duration (weeks): 12 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score
Golden 2002_449 RCT US	Outpatient N=330 Baseline severity: More severe Mean age (years): 41.2 Sex (% female): NR Ethnicity (% BME): NR	Paroxetine 20-62.5mg/day	Placebo	Treatment duration (weeks): 12 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score
Goldstein 2002a RCT US	Outpatient N=103 Baseline severity: More severe Mean age (years): 40.9	Fluoxetine 20mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient

Study	Population	Intervention	Comparison	Comments
	Sex (% female): 65 Ethnicity (% BME): 21			subgroup analysis): • Response
Goldstein 2004a RCT US	Outpatient N=176 Baseline severity: More severe Mean age (years): 40 Sex (% female): 64 Ethnicity (% BME): 22	Paroxetine 20mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Response
Gual 2003 RCT Spain	Outpatient N=83 Baseline severity: More severe Mean age (years): 46.7 Sex (% female): 47 Ethnicity (% BME): NR	Sertraline 50-150mg/day	Placebo	Treatment duration (weeks): 24 Outcomes (for inpatient versus outpatient subgroup analysis): • Response
Hirayasu 2011a RCT Japan	Outpatient N=310 34.6 Sex (% female): NR Ethnicity (% BME): NR	Escitalopram 10mg/day or 20mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Response
Hirayasu 2011b RCT Japan	Outpatient N=485 Baseline severity: More severe Mean age (years): 36.2 Sex (% female): NR Ethnicity (% BME): NR	Escitalopram 10mg/day or 20mg/day, or paroxetine 20-40mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Response
Hunter 2010_study 1 RCT US	Outpatient N=28 Baseline severity: More severe Mean age (years): 42.4 Sex (% female): 68	Fluoxetine 20mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Response

Study	Population	Intervention	Comparison	Comments
	Ethnicity (% BME): NR			
Hunter 2011 RCT US	Outpatient N=24 Baseline severity: More severe Mean age (years): 40.4 Sex (% female): 65 Ethnicity (% BME): NR	Fluoxetine 20mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score • Response
Jefferson 2000 RCT US	Outpatient N=415 Baseline severity: More severe Mean age (years): 39.9 Sex (% female): NR Ethnicity (% BME): NR	Paroxetine 25mg/day, or citalopram 20mg/day or 40mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score • Response
Keller 2006_Study 062 RCT Cross-continental	Outpatient N=325 Baseline severity: More severe Mean age (years): 41 Sex (% female): 67 Ethnicity (% BME): 43	Paroxetine 20mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score
Komulainen 2018 RCT Finland	Outpatient N=37 Baseline severity: More severe Mean age (years): median 25.1 Sex (% female): 44 Ethnicity (% BME): NR	Escitalopram 10mg/day	Placebo	Treatment duration (weeks): 1 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score
Kramer 1998 RCT US	Outpatient N=142 Baseline severity: More severe	Paroxetine 20mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for inpatient versus

Study	Population	Intervention	Comparison	Comments
	Mean age (years): NR Sex (% female): NR Ethnicity (% BME): NR			outpatient subgroup analysis): • Response
Kranzler 2006_Group A RCT US	Outpatient N=189 Baseline severity: More severe Mean age (years): 42.9 Sex (% female): 35 Ethnicity (% BME): 10	Sertraline 50-200mg/day	Placebo	Treatment duration (weeks): 10 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score • Response
Lam 2016b RCT Canada	Outpatient N=61 Baseline severity: More severe Mean age (years): 36.8 Sex (% female): 72 Ethnicity (% BME): NR	Fluoxetine 20mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score • Response
Lepola 2003 RCT Belgium, Canada, Finland, France, Norway, Sweden, Switzerland & UK	Outpatient N=469 Baseline severity: More severe Mean age (years): 43.3 Sex (% female): 72 Ethnicity (% BME): NR	Escitalopram 10-20mg/day or citalopram 20-40mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Response
Macias-Cortes 2015 RCT Mexico	Outpatient N=89 Baseline severity: More severe Mean age (years): 49 Sex (% female): 100 Ethnicity (% BME): 100	Fluoxetine 20mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score • Response

Study	Population	Intervention	Comparison	Comments
Mathews 2015 RCT US	Outpatient N=579 Baseline severity: More severe Mean age (years): 42.3 Sex (% female): 57 Ethnicity (% BME): 32	Citalopram 40mg/day	Placebo	Treatment duration (weeks): 10 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score • Response
Mendels 1999 RCT US	Outpatient N=180 Baseline severity: More severe Mean age (years): 43 Sex (% female): 33 Ethnicity (% BME): 13	Citalopram 20-80mg/day	Placebo	Treatment duration (weeks): 4 Outcomes (for inpatient versus outpatient subgroup analysis): • Response
Miller 1989a RCT UK	Outpatient N=47 Baseline severity: More severe Mean age (years): 42.5 Sex (% female): 68 Ethnicity (% BME): NR	Paroxetine 30mg/day	Placebo	Treatment duration (weeks): 4 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score
Mundt 2012 RCT US	Outpatient N=165 Baseline severity: More severe Mean age (years): 37.8 Sex (% female): 63 Ethnicity (% BME): 24	Sertraline 50-100mg/day	Placebo	Treatment duration (weeks): 4 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score • Response
MY-1042/BRL-029060/CPMS-251 RCT US	Outpatient N=254 Baseline severity: More severe Mean age (years): 41.9	Paroxetine 20-50mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient

Study	Population	Intervention	Comparison	Comments
	Sex (% female): NR Ethnicity (% BME): NR			subgroup analysis): • Depression symptoms change score • Response
MY-1045/BRL-029060/1 (PAR 128) RCT US	Outpatient N=848 Baseline severity: More severe Mean age (years): 41.8 Sex (% female): NR Ethnicity (% BME): NR	Paroxetine 20-50mg/day or fluoxetine 20-80mg/day	Placebo	Treatment duration (weeks): 12 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score • Response
Nemeroff 2007a RCT US	Outpatient N=206 Baseline severity: More severe Mean age (years): 39.1 Sex (% female): 61 Ethnicity (% BME): 10	Fluoxetine 20-60mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient subgroup analysis): • Response
Nierenberg 2007a RCT US	Outpatient N=411 Baseline severity: More severe Mean age (years): 43 Sex (% female): 66 Ethnicity (% BME): 21	Escitalopram 10mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score • Response
NKD20006 (NCT00048204) RCT US	Outpatient N=250 Baseline severity: More severe Mean age (years): 38 Sex (% female): 60 Ethnicity (% BME): NR	Paroxetine 20mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score • Response

Study	Population	Intervention	Comparison	Comments
Olie 1997 RCT France	Outpatient N=258 Baseline severity: More severe Mean age (years): 43.8 Sex (% female): 63 Ethnicity (% BME): 1	Sertraline 50-200mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient subgroup analysis): • Response
PAR 01 001 (GSK & FDA) RCT US	Outpatient N=50 Baseline severity: More severe Mean age (years): 43.1 Sex (% female): 35 Ethnicity (% BME): NR	Paroxetine 10-50mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score • Response
Perahia 2006a RCT Bulgaria, Croatia, Hungary, Poland, Romania, Russia, & Slovakia	Outpatient N=196 Baseline severity: More severe Mean age (years): 68.4 Sex (% female): 68 Ethnicity (% BME): 0	Paroxetine 20mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Response
Peselow 1989aa RCT US	Outpatient N=73 Baseline severity: More severe Mean age (years): 46.1 Sex (% female): 38 Ethnicity (% BME): NR	Paroxetine 10-50mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient subgroup analysis): • Response
Peselow 1989ba RCT US	Outpatient N=82 Baseline severity: More severe Mean age (years): NR Sex (% female): NR Ethnicity (% BME): NR	Paroxetine 10-50mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient subgroup analysis): • Response

Study	Population	Intervention	Comparison	Comments
Rapaport 2009 RCT US	Outpatient N=357 Baseline severity: More severe Mean age (years): 67.5 Sex (% female): 62 Ethnicity (% BME): 17	Paroxetine 25mg/day	Placebo	Treatment duration (weeks): 10 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score • Response
Ratti 2011_study 096 RCT 11 countries in Europe and Latin America	Outpatient N=236 Baseline severity: More severe Mean age (years): 44 Sex (% female): 72 Ethnicity (% BME): NR	Paroxetine 20-30mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Response
Ravindran 1995 RCT Canada	Outpatient N=66 Baseline severity: More severe Mean age (years): 38.9 Sex (% female): 62 Ethnicity (% BME): NR	Sertraline 50-200mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Response
Reimherr 1990 RCT US & Canada	Outpatient N=299 Baseline severity: More severe Mean age (years): 39.6 Sex (% female): 53 Ethnicity (% BME): 8	Sertraline 20-200mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score • Response
Rickels 1992 RCT US	Outpatient N=111 Baseline severity: More severe Mean age (years): 44.7 Sex (% female): 48	Paroxetine (dose NR)	Placebo	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient

Study	Population	Intervention	Comparison	Comments
	Ethnicity (% BME): NR			subgroup analysis): • Response
Roose 2004 RCT US	Outpatient N=178 Baseline severity: More severe Mean age (years): 79.6 Sex (% female): 58 Ethnicity (% BME): NR	Citalopram 20-40mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Response
Rudolph 1999a RCT US	Outpatient N=200 Baseline severity: More severe Mean age (years): 40 Sex (% female): 66 Ethnicity (% BME): NR	Fluoxetine 20-60mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Response
SER 315 (FDA)a RCT Europe	Outpatient N=165 Baseline severity: More severe Mean age (years): 42.0 Sex (% female): 72 Ethnicity (% BME): NR	Sertraline 50-200mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score
Smith 1992 RCT US	Outpatient N=77 Baseline severity: More severe Mean age (years): 44.8 Sex (% female): 50 Ethnicity (% BME): NR	Paroxetine 10-50mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient subgroup analysis): • Response
Stark 1985a RCT US	Outpatient N=354 Baseline severity: More severe Mean age (years): 40.5 Sex (% female): 68	Fluoxetine 60-80mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient

Study	Population	Intervention	Comparison	Comments
	Ethnicity (% BME): NR			subgroup analysis): • Depression symptoms change score • Response
Study 62b (FDA) RCT Country NR	Outpatient N=356 Baseline severity: More severe Mean age (years): 40 Sex (% female): 57 Ethnicity (% BME): NR	Fluoxetine 20mg/day, 40mg/day, or 60mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score
Study F1J-MC-HMAQ – Study Group Ba RCT US	Outpatient N=112 Baseline severity: More severe Mean age (years): 40.8 Sex (% female): NR Ethnicity (% BME): NR	Fluoxetine 20mg/day	Placebo	Treatment duration (weeks): 10 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score • Response
Tollefson 1993/1995 RCT US	Outpatient N=671 Baseline severity: More severe Mean age (years): 67.7 Sex (% female): 55 Ethnicity (% BME): 6	Fluoxetine maximum 20mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score • Response
Valle-Cabrera 2018 RCT Cuba	Outpatient N=77 Baseline severity: More severe Mean age (years): 45.2 Sex (% female): 92 Ethnicity (% BME): NR	Sertraline 50-200mg/day	Placebo	Treatment duration (weeks): 10 Outcomes (for inpatient versus outpatient subgroup analysis): • Response

Study	Population	Intervention	Comparison	Comments
VEN XR 367 (FDA)a RCT Europe	Outpatient N=164 Baseline severity: More severe Mean age (years): NR Sex (% female): 61 Ethnicity (% BME): NR	Paroxetine 20mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score
Wade 2002 RCT Canada, Estonia, France, Netherlands & UK	Outpatient N=380 Baseline severity: More severe Mean age (years): 40.5 Sex (% female): 76 Ethnicity (% BME): 3	Escitalopram 10mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score • Response
Wang 2014c RCT Canada, China, Finland, South Korea, Malaysia, Mexico, The Philippines, South Africa, & Spain	Outpatient N=314 Baseline severity: More severe Mean age (years): 40 Sex (% female): 71 Ethnicity (% BME): 46	Escitalopram 10-20mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Response
WELL AK1A4006 RCT US	Outpatient N=309 Baseline severity: More severe Mean age (years): 37.9 Sex (% female): NR Ethnicity (% BME): NR	Fluoxetine 20-60mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score • Response
Wernicke 1987 RCT US	Outpatient N=356 Baseline severity: More severe Mean age (years): 39.8	Fluoxetine 20mg/day, 40mg/day, or 60mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient

Study	Population	Intervention	Comparison	Comments
	Sex (% female): 57 Ethnicity (% BME): NR			subgroup analysis): • Depression symptoms change score • Response
Wernicke 1988 RCT US	Outpatient N=267 Baseline severity: More severe Mean age (years): NR Sex (% female): NR Ethnicity (% BME): NR	Fluoxetine 20mg/day or 40mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score • Response

a Three-armed trial but where possible the demographics reported here are for only the two relevant arms.
b Four-armed trial but where possible the demographics reported here are for only the two relevant arms
BME: black, minority, ethnic; mg: milligrams; N: number of participants; NR: not reported; RCT: randomised controlled trial.

Summaries of the studies included in the inpatient versus outpatient subgroup analysis of the SSRIs versus tricyclic antidepressants (TCAs) comparison are presented in Table 9.

There were no significant subgroup differences between inpatient and outpatient settings for the comparison SSRIs versus TCAs on: depression symptoms endpoint (Test for subgroup differences: $Chi^2 = 1.08$, df = 1, p = 0.30); remission (Test for subgroup differences: $Chi^2 = 2.11$, df = 1, p = 0.15); response (Test for subgroup differences: $Chi^2 = 1.03$, df = 1, p = 0.31). There was a statistically significant subgroup difference between inpatient and outpatient settings for depression change score (Test for subgroup differences: $Chi^2 = 7.03$, df = 1, p = 0.008). In inpatient settings TCAs showed a small benefit over SSRIs (SMD 0.27 [0.08, 0.47]), whereas in outpatient settings SSRIs showed a small benefit over TCAs (SMD -0.05 [-0.19, 0.09]), however, in both inpatient and outpatient settings the difference between TCAs and SSRIs was non-significant.

Table 19: Summary of included studies for inpatient versus outpatient subgroup analysis for comparison 3b SSRIs versus tricyclic antidepressants (TCAs)

Study	Population	Intervention	Comparison	Comments
Inpatient setting (K=11, N=1,347)				
29060/299 RCT Europe	Inpatient N=217 Baseline severity: More severe Mean age (years): 40.4 Sex (% female): NR	Paroxetine 20-50mg/day	Amitriptyline 100-250mg/day	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient

Study	Population	Intervention	Comparison	Comments
	Ethnicity (% BME): NR			subgroup analysis): • Depression symptoms change score
29060 07 001a RCT US	Inpatient N=26 Baseline severity: More severe Mean age (years): 42.3 Sex (% female): 65 Ethnicity (% BME): NR	Paroxetine 10-60mg/day	Amitriptyline (dose NR)	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score
Arminen 1992 RCT Finland	Inpatient N=57 Baseline severity: More severe Mean age (years): NR Sex (% female): 54 Ethnicity (% BME): NR	Paroxetine 20-40mg/day	Imipramine 100-200mg/day	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms endpoint
Danish University Antidepressant Group 1986 RCT Denmark	Inpatient N=114 Baseline severity: More severe Mean age (years): NR Sex (% female): 70 Ethnicity (% BME): NR	Citalopram 40mg/day	Clomipramine 150mg/day	Treatment duration (weeks): 5 Outcomes (for inpatient versus outpatient subgroup analysis): • Remission
Danish University Antidepressant Group 1990 RCT Denmark	Inpatient N=120 Baseline severity: More severe Mean age (years): NR Sex (% female): 66 Ethnicity (% BME): NR	Paroxetine 30mg/day	Clomipramine 150mg/day	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient subgroup analysis): • Remission
Deushle 2003 RCT Germany	Inpatient N=126 Baseline severity: More severe	Paroxetine 40mg/day	Amitriptyline 150mg/day	Treatment duration (weeks): 5 Outcomes (for inpatient versus outpatient

Study	Population	Intervention	Comparison	Comments
	Mean age (years): 54.1 Sex (% female): 67 Ethnicity (% BME): NR			subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score
Geretsegger 1995 RCT Austria & Germany	Inpatient N=91 Baseline severity: More severe Mean age (years): 71.2 Sex (% female): 86 Ethnicity (% BME): NR	Paroxetine 20-30mg/day	Amitriptyline 100-150mg/day	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient subgroup analysis): • Remission • Response
Laakmann 1991 RCT Germany	Inpatient N=174 Baseline severity: More severe Mean age (years): NR Sex (% female): NR Ethnicity (% BME): NR	Fluoxetine (dose NR)	Amitriptyline 100-200mg/day	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms endpoint
Moller 1993 RCT Germany & Hungary	Inpatient N=222 Baseline severity: More severe Mean age (years): 47.1 Sex (% female): NR Ethnicity (% BME): NR	Paroxetine 30-50mg/day	Amitriptyline 150-250mg/day	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score • Remission • Response
Moller 1998 RCT Germany, Hungary, & Czech Republic	Inpatient N=160 Baseline severity: More severe Mean age (years): 48.6	Sertraline 50-150mg/day	Amitriptyline 75-150mg/day	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient

Study	Population	Intervention	Comparison	Comments
	Sex (% female): 70 Ethnicity (% BME): NR			subgroup analysis): • Depression symptoms change score • Response
Staner 1995 RCT Belgium	Inpatient N=40 Baseline severity: More severe Mean age (years): 42.1 Sex (% female): 83 Ethnicity (% BME): NR	Paroxetine 30mg/day	Amitriptyline 150mg/day	Treatment duration (weeks): 5 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score • Response
Outpatient setting (K=40, N=5,774)				
Akhondzadeh 2003 RCT Iran	Outpatient N=48 Baseline severity: More severe Mean age (years): 35.8 Sex (% female): 40 Ethnicity (% BME): NR	Fluoxetine 60mg/day	Nortriptyline 150mg/day	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score
Beasley 1993b RCT US	Outpatient N=136 Baseline severity: More severe Mean age (years): 44.8 Sex (% female): 70 Ethnicity (% BME): 4	Fluoxetine 40-80mg/day	Amitriptyline 150-300mg/day	Treatment duration (weeks): 5 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score • Remission • Response
Bersani 1994 RCT Italy	Outpatient N=68 Baseline severity: More severe Mean age (years): 47.1	Sertraline 50-150mg/day	Amitriptyline 50-150mg/day	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient

Study	Population	Intervention	Comparison	Comments
	Sex (% female): 63 Ethnicity (% BME): NR			subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score
Bhargava 2012 RCT India	Outpatient N=60 Baseline severity: More severe Mean age (years): 36.2 Sex (% female): 52 Ethnicity (% BME): NR	Sertraline 50-150mg/day	Imipramine 75-150mg/day	Treatment duration (weeks): 12 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score
Bremner 1984 RCT US	Outpatient N=40 Baseline severity: More severe Mean age (years): 42.6 Sex (% female): 51 Ethnicity (% BME): NR	Fluoxetine 60-80mg/day	Imipramine 125-300mg/day	Treatment duration (weeks): 5 Outcomes (for inpatient versus outpatient subgroup analysis): • Response
Byerley 1988a RCT US	Outpatient N=66 Baseline severity: More severe Mean age (years): 39.3 Sex (% female): 68 Ethnicity (% BME): NR	Fluoxetine 40-80mg/day	Imipramine 150-300mg/day	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms endpoint • Response
Christiansen 1996 RCT Denmark	Outpatient N=144 Baseline severity: More severe Mean age (years): NR Sex (% female): NR	Paroxetine 20-40mg/day	Amitriptyline 75-150mg/day	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient

Study	Population	Intervention	Comparison	Comments
	Ethnicity (% BME): NR			subgroup analysis): • Depression symptoms endpoint • Response
Cohn 1984b RCT US	Outpatient N=66 Baseline severity: More severe Mean age (years): 42 Sex (% female): NR Ethnicity (% BME): NR	Fluoxetine (dose NR)	Imipramine (dose NR)	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms endpoint
Cohn 1990b RCT US	Outpatient N=241 Baseline severity: More severe Mean age (years): 70.3 Sex (% female): 49 Ethnicity (% BME): NR	Sertraline 50-200mg/day	Amitriptyline 50-150mg/day	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score • Response
De Ronchi 1998 RCT Italy	Outpatient N=65 Baseline severity: More severe Mean age (years): 68.9 Sex (% female): 72 Ethnicity (% BME): NR	Fluoxetine 20mg/day	Amitriptyline 50-100mg/day	Treatment duration (weeks): 10 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score • Response
Demyttenaere 1998 RCT Belgium	Outpatient N=66 Baseline severity: More severe Mean age (years): 41.7 Sex (% female): 55	Fluoxetine 20mg/day	Amitriptyline 50mg/day	Treatment duration (weeks): 9 Outcomes (for inpatient versus outpatient

Study	Population	Intervention	Comparison	Comments
	Ethnicity (% BME): NR			subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score • Response
Fabre 1991 RCT US	Outpatient N=205 Baseline severity: More severe Mean age (years): 37 Sex (% female): 57 Ethnicity (% BME): NR	Fluoxetine 40mg/day	Nortriptyline 100mg/day	Treatment duration (weeks): 5 Outcomes (for inpatient versus outpatient subgroup analysis): • Response
Fabre 1992a RCT US	Outpatient N=80 Baseline severity: More severe Mean age (years): 35.4 Sex (% female): 61 Ethnicity (% BME): NR	Paroxetine 10-50mg/day	Imipramine 65-275mg/day	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score
Fawcett 1989 RCT US	Outpatient N=40 Baseline severity: More severe Mean age (years): 42.2 Sex (% female): 65 Ethnicity (% BME): NR	Fluoxetine 20-60mg/day	Amitriptyline 50-200mg/day	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score • Remission • Response
Feighner 1993a RCT US	Outpatient N=477 Baseline severity: More severe Mean age (years): 40.1	Paroxetine 10-50mg/day	Imipramine 65-275mg/day	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient

Study	Population	Intervention	Comparison	Comments
	Sex (% female): 53 Ethnicity (% BME): NR			subgroup analysis): • Remission
Forlenza 2001 RCT Brazil	Outpatient N=55 Baseline severity: More severe Mean age (years): 68.5 Sex (% female): 69 Ethnicity (% BME): NR	Sertraline 50mg/day	Imipramine 150mg/day	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score • Remission • Response
Freed 1999 RCT Australia	Outpatient N=375 Baseline severity: More severe Mean age (years): 48 Sex (% female): 65 Ethnicity (% BME): NR	Paroxetine 20mg/day	Amitriptyline 75mg/day	Treatment duration (weeks): 9 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score
Hashemi 2012 RCT Iran	Outpatient N=120 Baseline severity: More severe Mean age (years): 34.8 Sex (% female): 53 Ethnicity (% BME): NR	Fluoxetine maximum 60mg/day	Nortriptyline maximum 150mg/day	Treatment duration (weeks): 26 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score
Hutchinson 1992 RCT UK	Outpatient N=90 Baseline severity: More severe	Paroxetine 30mg/day	Amitriptyline 100mg/day	Treatment duration (weeks): 26 Outcomes (for inpatient versus

Study	Population	Intervention	Comparison	Comments
	Mean age (years): 71.8 Sex (% female): 77 Ethnicity (% BME): NR			outpatient subgroup analysis): • Remission • Response
Kyle 1998 RCT UK	Outpatient N=365 Baseline severity: More severe Mean age (years): 73.8 Sex (% female): 73 Ethnicity (% BME): NR	Citalopram 20-40mg/day	Amitriptyline 50-100mg/day	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Remission
Laakmann 1988 RCT Germany	Outpatient N=128 Baseline severity: More severe Mean age (years): NR Sex (% female): 72 Ethnicity (% BME): NR	Fluoxetine 20-60mg/day	Amitriptyline 50-150mg/day	Treatment duration (weeks): 5 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms endpoint • Response
Marchesi 1998 RCT Italy	Outpatient N=142 Baseline severity: More severe Mean age (years): 43.6 Sex (% female): 74 Ethnicity (% BME): NR	Fluoxetine 20mg/day	Amitriptyline 75-225mg/day	Treatment duration (weeks): 10 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score • Response
MDF/29060/III/070/88/MC RCT Europe	Outpatient N=62 Baseline severity: More severe Mean age (years): 73 Sex (% female): NR Ethnicity (% BME): NR	Paroxetine 20-30mg/day	Clomipramine 60-75mg/day	Treatment duration (weeks): 5 Outcomes (for inpatient versus outpatient

Study	Population	Intervention	Comparison	Comments
				subgroup analysis): • Depression symptoms change score • Remission • Response
Moller 2000 RCT Germany	Outpatient N=240 Baseline severity: More severe Mean age (years): 47.9 Sex (% female): 67 Ethnicity (% BME): NR	Sertraline 50-100mg/day	Amitriptyline 75-150mg/day	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score • Response
Moon 1994 RCT UK	Outpatient N=106 Baseline severity: More severe Mean age (years): 43.7 Sex (% female): 52 Ethnicity (% BME): NR	Sertraline 50-150mg/day	Clomipramine 50-150mg/day	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient subgroup analysis): • Response
Moon 1996 RCT UK	Outpatient N=138 Baseline severity: More severe Mean age (years): 43.7 Sex (% female): 71 Ethnicity (% BME): NR	Paroxetine 20-30mg/day	Lofepramine 70-210mg/day	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient subgroup analysis): • Remission • Response
Ontiveros Sanchez 1998 RCT South America	Outpatient N=42 Baseline severity: More severe Mean age (years): 37.6 Sex (% female): 53 Ethnicity (% BME): NR	Fluoxetine 20mg/day	Amitriptyline 150-250mg/day	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms endpoint • Response

Study	Population	Intervention	Comparison	Comments
PAR 29060/281 RCT Europe	Outpatient N=162 Baseline severity: More severe Mean age (years): 38.8 Sex (% female): 77 Ethnicity (% BME): NR	Paroxetine 30mg/day	Amitriptyline 75-150mg/day	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms endpoint
PAR MDUK 032 RCT Country NR	Outpatient N=59 Baseline severity: More severe Mean age (years): 44.4 Sex (% female): NR Ethnicity (% BME): NR	Paroxetine 20-30mg/day	Amitriptyline 100-150mg/day	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms endpoint
Peselow 1989aa RCT US	Outpatient N=66 Baseline severity: More severe Mean age (years): 45.9 Sex (% female): 35 Ethnicity (% BME): NR	Paroxetine 10-50mg/day	Imipramine 65-275mg/day	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient subgroup analysis): • Response
Peselow 1989ba RCT US	Outpatient N=80 Baseline severity: More severe Mean age (years): NR Sex (% female): NR Ethnicity (% BME): NR	Paroxetine 20-50mg/day	Imipramine 65-275mg/day	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient subgroup analysis): • Response
Peters 1990 RCT Germany	Outpatient N=102 Baseline severity: More severe Mean age (years): 44.5 Sex (% female): 63 Ethnicity (% BME): NR	Fluoxetine 20mg/day	Amitriptyline 100mg/day	Treatment duration (weeks): 5 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms endpoint

Study	Population	Intervention	Comparison	Comments
				• Response
Preskorn 1991 RCT US	Outpatient N=61 Baseline severity: More severe Mean age (years): NR Sex (% female): NR Ethnicity (% BME): 2	Fluoxetine 20-60mg/day	Amitriptyline 50-200mg/day	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score
Reimherr 1990a RCT US & Canada	Outpatient N=298 Baseline severity: More severe Mean age (years): 38.4 Sex (% female): 55 Ethnicity (% BME): 10	Sertraline 20-200mg/day	Amitriptyline 50-150mg/day	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score • Response
Ropert 1989 RCT France	Outpatient N=143 Baseline severity: More severe Mean age (years): 43.8 Sex (% female): 64 Ethnicity (% BME): NR	Fluoxetine 20mg/day	Clomipramine 75mg/day	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score
Rosenberg 1994 RCT Denmark, Norway, Sweden & Finland	Outpatient N=472 Baseline severity: More severe Mean age (years): 47.6 Sex (% female): 69 Ethnicity (% BME): NR	Citalopram 10-30mg/day or 20-60mg/day	Imipramine 50-150mg/day	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient subgroup analysis): • Response
SER 315 (FDA)a RCT Europe	Outpatient N=162 Baseline severity: More severe	Sertraline 50-200mg/day	Amitriptyline 50-200mg/day	Treatment duration (weeks): 8 Outcomes (for inpatient versus

Study	Population	Intervention	Comparison	Comments
	Mean age (years): 42.4 Sex (% female): 69 Ethnicity (% BME): NR			outpatient subgroup analysis): • Depression symptoms change score
Serrano-Blanco 2006 RCT Spain	Outpatient N=103 Baseline severity: More severe Mean age (years): 43.5 Sex (% female): 73 Ethnicity (% BME): NR	Fluoxetine 10-40mg/day	Imipramine 25-125mg/day	Treatment duration (weeks): 24 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score
Stark 1985a RCT US	Outpatient N=371 Baseline severity: More severe Mean age (years): 41.0 Sex (% female): 69 Ethnicity (% BME): NR	Fluoxetine 60-80mg/day	Imipramine 100-300mg/day	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score • Response
Suleman 1997 RCT Zimbabwe	Outpatient N=30 Baseline severity: More severe Mean age (years): NR Sex (% female): NR Ethnicity (% BME): NR	Fluoxetine 20mg/day	Amitriptyline 100mg/day	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score

a Three-armed trial but where possible the demographics reported here are for only the two relevant arms.
BME: black, minority, ethnic; mg: milligrams; N: number of participants; NR: not reported; RCT: randomised controlled trial.

Summaries of the studies included in the inpatient versus outpatient subgroup analysis of the serotonin–norepinephrine reuptake inhibitors (SNRIs) versus placebo comparison are presented in Table 10.

There were no significant subgroup differences between inpatient and outpatient settings for the comparison SNRIs versus placebo on: depression symptoms endpoint (Test for subgroup differences: Chi² = 0.03, df = 1, p = 0.87); depression symptoms change score (Test for subgroup differences: Chi² = 3.12, df = 1, p = 0.08); remission (Test for subgroup differences: Chi² = 0.25, df = 1, p = 0.62).

Table 20: Summary of included studies for inpatient versus outpatient subgroup analysis for comparison 3c Serotonin–norepinephrine reuptake inhibitors (SNRIs) versus placebo

Study	Population	Intervention	Comparison	Comments
Inpatient setting (K=2, N=283)				
Guelfi 1995 RCT France	Inpatient N=93 Baseline severity: More severe Mean age (years): 56 Sex (% female): 85 Ethnicity (% BME): NR	Venlafaxine 150-375mg/day	Placebo	Treatment duration (weeks): 4 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score • Remission
Sheehan 2009ba RCT US	Inpatient N=190 Baseline severity: More severe Mean age (years): 40.8 Sex (% female): 56 Ethnicity (% BME): NR	Venlafaxine 225-375mg/day	Placebo	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score • Remission
Outpatient setting (K=26, N=6,784)				
Brannan 2005 RCT US	Outpatient N=282 Baseline severity: More severe Mean age (years): 40.6	Duloxetine 60mg/day	Placebo 2 capsules/day	Treatment duration (weeks): 7 Outcomes (for inpatient versus outpatient

Study	Population	Intervention	Comparison	Comments
	Sex (% female): 65 Ethnicity (% BME): 20			subgroup analysis): • Depression symptoms change score • Remission
Cutler 2009 RCT US	Outpatient N=308 Baseline severity: More severe Mean age (years): 41.3 Sex (% female): 63 Ethnicity (% BME): 28	Duloxetine 60mg/day	Placebo	Treatment duration (weeks): 6 Outcome (for inpatient versus outpatient subgroup analysis): • Remission
Detke 2002a RCT US	Outpatient N=267 Baseline severity: More severe Mean age (years): 41 Sex (% female): 69 Ethnicity (% BME): 22	Duloxetine 60mg/day	Placebo 3 capsules/day	Treatment duration (weeks): 9 Outcome (for inpatient versus outpatient subgroup analysis): • Remission
Detke 2002b RCT US	Outpatient N=245 Baseline severity: More severe Mean age (years): 42.4 Sex (% female): 67 Ethnicity (% BME): 14	Duloxetine 40-60mg/day	Placebo 2-3 capsules/day	Treatment duration (weeks): 9 Outcome (for inpatient versus outpatient subgroup analysis): • Remission
Detke 2004a RCT US	Outpatient N=281 Baseline severity: More severe Mean age (years): 43.8 Sex (% female): 74 Ethnicity (% BME): 0	Duloxetine 80mg/day or 120mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score • Remission

Study	Population	Intervention	Comparison	Comments
Eli Lilly HMAT-Aa RCT US	Outpatient N=174 Baseline severity: More severe Mean age (years): NR Sex (% female): NR Ethnicity (% BME): NR	Duloxetine 80mg/day or 120mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score • Remission
Goldstein 2002a RCT US	Outpatient N=140 Baseline severity: More severe Mean age (years): 41.9 Sex (% female): 66 Ethnicity (% BME): 15	Duloxetine 40-120mg/day	Placebo	Treatment duration (weeks): 8 Outcome (for inpatient versus outpatient subgroup analysis): • Remission
Goldstein 2004a RCT US	Outpatient N=180 Baseline severity: More severe Mean age (years): 40.5 Sex (% female): 63 Ethnicity (% BME): 16	Duloxetine 80mg/day	Placebo	Treatment duration (weeks): 8 Outcome (for inpatient versus outpatient subgroup analysis): • Remission
Hewett 2009 RCT Country NR	Outpatient N=384 Baseline severity: More severe Mean age (years): 42.2 Sex (% female): 70 Ethnicity (% BME): 3	Venlafaxine 75-150mg/day	Placebo	Treatment duration (weeks): 8 Outcome (for inpatient versus outpatient subgroup analysis): • Remission
Hewett 2010 RCT Country NR	Outpatient N=385 Baseline severity: More severe Mean age (years): 44.3 Sex (% female): 68 Ethnicity (% BME): 5	Venlafaxine 75-150mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient

Study	Population	Intervention	Comparison	Comments
				subgroup analysis): • Depression symptoms change score • Remission
Higuchi 2016 RCT Japan	Outpatient N=538 Baseline severity: More severe Mean age (years): 38.4 Sex (% female): NR Ethnicity (% BME): 100	Venlafaxine 75mg/day or 75-225mg/day	Placebo	Treatment duration (weeks): 8 Outcome (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score
Khan 1998 RCT US	Outpatient N=403 Baseline severity: More severe Mean age (years): 41.7 Sex (% female): 63 Ethnicity (% BME): NR	Venlafaxine 75mg/day, 150mg/day or 200mg/day	Placebo	Treatment duration (weeks): 12 Outcome (for inpatient versus outpatient subgroup analysis): • Depression symptoms endpoint
Levin 2013 RCT US	Outpatient N=103 Baseline severity: More severe Mean age (years): 35.1 Sex (% female): 26 Ethnicity (% BME): 54	Venlafaxine maximum 375mg/day	Placebo	Treatment duration (weeks): 12 Outcome (for inpatient versus outpatient subgroup analysis): • Remission
Mendels 1993 RCT US	Outpatient N=157 Baseline severity: More severe Mean age (years): 38.5 Sex (% female): 65 Ethnicity (% BME): NR	Venlafaxine 150-200mg/day	Placebo	Treatment duration (weeks): 6 Outcome (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score

Study	Population	Intervention	Comparison	Comments
Nemeroff 2007a RCT US	Outpatient N=204 Baseline severity: More severe Mean age (years): 40.2 Sex (% female): 59 Ethnicity (% BME): 10	Venlafaxine 75-225mg/day	Placebo	Treatment duration (weeks): 6 Outcome (for inpatient versus outpatient subgroup analysis): • Remission
Nierenberg 2007a RCT US	Outpatient N=410 Baseline severity: More severe Mean age (years): 41.6 Sex (% female): 63 Ethnicity (% BME): 22	Duloxetine 60mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score • Remission
Perahia 2006a RCT Bulgaria, Croatia, Hungary, Poland, Romania, Russia, & Slovakia	Outpatient N=295 Baseline severity: More severe Mean age (years): 45 Sex (% female): 69 Ethnicity (% BME): 0	Duloxetine 80mg/day or 120mg/day	Placebo	Treatment duration (weeks): 8 Outcome (for inpatient versus outpatient subgroup analysis): • Remission
Raskin 2007 RCT US	Outpatient N=311 Baseline severity: More severe Mean age (years): 72.8 Sex (% female): 59 Ethnicity (% BME): 22	Duloxetine 60mg/day	Placebo	Treatment duration (weeks): 8 Outcome (for inpatient versus outpatient subgroup analysis): • Remission
Robinson 2014 RCT France, Mexico, Puerto Rico, & US	Outpatient N=370 Baseline severity: More severe Mean age (years): 72.9 Sex (% female): 63 Ethnicity (% BME): 22	Duloxetine 60mg/day	Placebo	Treatment duration (weeks): 12 Outcomes (for inpatient versus outpatient

Study	Population	Intervention	Comparison	Comments
				subgroup analysis): • Depression symptoms change score • Remission
Rudolph 1999a RCT US	Outpatient N=192 Baseline severity: More severe Mean age (years): 40 Sex (% female): 71 Ethnicity (% BME): NR	Venlafaxine 75-225mg/day	Placebo	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms endpoint • Remission
Schweizer 1994a RCT US	Outpatient N=151 Baseline severity: More severe Mean age (years): 41.5 Sex (% female): 69 Ethnicity (% BME): NR	Venlafaxine 75-225mg/day	Placebo	Treatment duration (weeks): 6 Outcome (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score
Study F1J-MC-HMAQ-Study Group Ba RCT US	Outpatient N=157 Baseline severity: More severe Mean age (years): 40.6 Sex (% female): NR Ethnicity (% BME): NR	Duloxetine 40-120mg/day	Placebo	Treatment duration (weeks): 10 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score • Remission
Thase 1997 RCT US	Outpatient N=197 Baseline severity: More severe Mean age (years): 41 Sex (% female): 61 Ethnicity (% BME): NR	Venlafaxine 75-225mg/day	Placebo 1-3 capsules/day	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms endpoint

Study	Population	Intervention	Comparison	Comments
				• Remission
VEN 600A-303 (FDA) RCT US	Outpatient N=165 Baseline severity: More severe Mean age (years): 38.5 Sex (% female): 69 Ethnicity (% BME): NR	Venlafaxine 150-225mg/day	Placebo	Treatment duration (weeks): 6 Outcome (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score
VEN 600A-313 (FDA) RCT US	Outpatient N=237 Baseline severity: More severe Mean age (years): 38.4 Sex (% female): 67 Ethnicity (% BME): NR	Venlafaxine 75mg/day or 200mg/day	Placebo	Treatment duration (weeks): 6 Outcome (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score
VEN XR 367 (FDA)a RCT Europe	Outpatient N=248 Baseline severity: More severe Mean age (years): NR Sex (% female): 66 Ethnicity (% BME): NR	Venlafaxine 75mg/day or 150mg/day	Placebo	Treatment duration (weeks): 8 Outcome (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score

a Three-armed trial but where possible the demographics reported here are for only the two relevant arms.
BME: black, minority, ethnic; mg: milligrams; N: number of participants; NR: not reported; RCT: randomised controlled trial.

Summaries of the studies included in the inpatient versus outpatient subgroup analysis of the SNRIs versus SSRIs comparison are presented in Table 11.

There were no significant subgroup differences between inpatient and outpatient settings for the comparison SNRIs versus SSRIs on: depression symptoms endpoint (Test for subgroup differences: $Chi^2 = 2.03$, df = 1, p = 0.15); remission (Test for subgroup differences: $Chi^2 = 1.08$, df = 1, p = 0.30); response (Test for subgroup differences: $Chi^2 = 0.49$, df = 1, p = 0.48). There was a statistically significant subgroup difference between inpatient and outpatient settings for depression change score (Test for subgroup differences: $Chi^2 = 8.03$, df = 1, p = 0.005). SNRIs showed a benefit over SSRIs in both settings, although this effect was larger in inpatient settings (SMD -0.48 [-0.73, -0.23]) relative to outpatient settings (SMD -0.09 [-0.19, 0.01]), however, this was a difference in magnitude rather than direction and even in inpatient settings the difference was not clinically important.

Table 21: Summary of included studies for inpatient versus outpatient subgroup analysis for comparison 3d SNRIs versus SSRIs

Study	Population	Intervention	Comparison	Comments
Inpatient setting (K=4, N=476)				
Clerc 1994 RCT France & Belgium	Inpatient N=68 Baseline severity: More severe Mean age (years): 51.3 Sex (% female): 68 Ethnicity (% BME): NR	Venlafaxine 200mg/day	Fluoxetine 40mg/day	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score • Response
Hwang 2004 RCT Taiwan	Inpatient N=105 Baseline severity: More severe Mean age (years): 65.1 Sex (% female): 58 Ethnicity (% BME): NR	Venlafaxine 75-150mg/day	Paroxetine 20-40mg/day	Treatment duration (weeks): 4 Outcome (for inpatient versus outpatient subgroup analysis): • Response
Sheehan 2009ba RCT US	Inpatient N=194 Baseline severity: More severe Mean age (years): 39.7 Sex (% female): 59 Ethnicity (% BME): NR	Venlafaxine 225-375mg/day	Fluoxetine 60-80mg/day	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score • Remission • Response
Tzanakaki 2000 RCT Greece & Italy	Inpatient N=109 Baseline severity: More severe Mean age (years): 48 Sex (% female): 79	Venlafaxine 225mg/day	Fluoxetine 60mg/day	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient

Study	Population	Intervention	Comparison	Comments
	Ethnicity (% BME): NR			subgroup analysis): • Remission • Response
Outpatient setting (K=32, N=6,238)				
Allard 2004 RCT Sweden & Denmark	Outpatient N=151 Baseline severity: More severe Mean age (years): 73 Sex (% female): 80 Ethnicity (% BME): NR	Venlafaxine 75-150mg/day	Citalopram 10-20mg/day	Treatment duration (weeks): 22 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score • Remission • Response
Alves 1999 RCT Portugal	Outpatient N=87 Baseline severity: More severe Mean age (years): 43.7 Sex (% female): 92 Ethnicity (% BME): NR	Venlafaxine 75-150mg/day	Fluoxetine 20-40mg/day	Treatment duration (weeks): 12 Outcomes (for inpatient versus outpatient subgroup analysis): • Remission • Response
Bielski 2004 RCT US	Outpatient N=202 Baseline severity: More severe Mean age (years): 37.4 Sex (% female): 58 Ethnicity (% BME): 25	Venlafaxine 225mg/day	Escitalopram 20mg/day	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score • Remission • Response
Casabona 2004 RCT Country NR	Outpatient N=114 Baseline severity: More severe Mean age (years): NR	Venlafaxine 75mg/day	Paroxetine 20mg/day	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient

Study	Population	Intervention	Comparison	Comments
	Sex (% female): 77 Ethnicity (% BME): NR			subgroup analysis): • Depression symptoms endpoint • Remission • Response
Chang 2015 RCT Taiwan	Outpatient N=112 Baseline severity: More severe Mean age (years): 39.7 Sex (% female): 73 Ethnicity (% BME): NR	Venlafaxine 75-225mg/day	Fluoxetine 20-80mg/day	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score
Costa 1998 RCT Argentina, Brazil, Chile, Colombia, Uruguay, & Venezuela	Outpatient N=382 Baseline severity: More severe Mean age (years): 40.2 Sex (% female): 79 Ethnicity (% BME): NR	Venlafaxine 75-150mg/day	Fluoxetine 20-40mg/day	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score • Remission • Response
DeNayer 2002 RCT Belgium	Outpatient N=146 Baseline severity: More severe Mean age (years): 42.8 Sex (% female): 68 Ethnicity (% BME): NR	Venlafaxine 75-150mg/day	Fluoxetine 20-40mg/day	Treatment duration (weeks): 12 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score • Remission • Response

Study	Population	Intervention	Comparison	Comments
Detke 2004a RCT US	Outpatient N=274 Baseline severity: More severe Mean age (years): 43.3 Sex (% female): 72 Ethnicity (% BME): 0	Duloxetine 80mg/day or 120mg/day	Paroxetine 20mg/day	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score • Remission • Response
Diaz-Martinez 1998 RCT Mexico	Outpatient N=145 Baseline severity: More severe Mean age (years): NR Sex (% female): 72 Ethnicity (% BME): NR	Venlafaxine 75-150mg/day	Fluoxetine 20-40mg/day	Treatment duration (weeks): 8 Outcome (for inpatient versus outpatient subgroup analysis): • Response
Dierick 1996 RCT Belgium, Italy, Switzerland & France	Outpatient N=314 Baseline severity: More severe Mean age (years): 43.4 Sex (% female): 65 Ethnicity (% BME): NR	Venlafaxine 75-150mg/day	Fluoxetine 20mg/day	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score • Response
Eli Lilly HMAT-Aa RCT US	Outpatient N=173 Baseline severity: More severe Mean age (years): NR Sex (% female): NR Ethnicity (% BME): NR	Duloxetine 80mg/day	Paroxetine 20mg/day	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score • Remission • Response

Study	Population	Intervention	Comparison	Comments
Goldstein 2002a RCT US	Outpatient N=103 Baseline severity: More severe Mean age (years): 41.5 Sex (% female): 61 Ethnicity (% BME): 17	Duloxetine 40-120mg/day	Fluoxetine 20mg/day	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Remission • Response
Goldstein 2004a RCT US	Outpatient N=178 Baseline severity: More severe Mean age (years): 40.5 Sex (% female): 63 Ethnicity (% BME): 21	Duloxetine 80mg/day	Paroxetine 20mg/day	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Remission • Response
Hackett 1996 RCT Europe	Outpatient N=241 Baseline severity: More severe Mean age (years): NR Sex (% female): NR Ethnicity (% BME): NR	Venlafaxine 150mg/day	Paroxetine (dose NR)	Treatment duration (weeks): 8 Outcome (for inpatient versus outpatient subgroup analysis): • Depression symptoms endpoint
Heller 2009 RCT US	Outpatient N=29 Baseline severity: More severe Mean age (years): 31.9 Sex (% female): 55 Ethnicity (% BME): NR	Venlafaxine 75-300mg/day	Fluoxetine 20-80mg/day	Treatment duration (weeks): 26 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score
Jiang 2017 RCT China	Outpatient N=26 Baseline severity: More severe Mean age (years): 45.5	Duloxetine (dose NR)	Escitalopram (dose NR)	Treatment duration (weeks): 8 Outcome (for inpatient versus outpatient

Study	Population	Intervention	Comparison	Comments
	Sex (% female): 73 Ethnicity (% BME): NR			subgroup analysis): • Response
Khan 2007 RCT US	Outpatient N=278 Baseline severity: More severe Mean age (years): 42.4 Sex (% female): 61 Ethnicity (% BME): 20	Duloxetine 60mg/day	Escitalopram 10-20mg/day	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score • Remission • Response
Kornaat 2000 RCT Country NR	Outpatient N=156 Baseline severity: More severe Mean age (years): NR Sex (% female): 64 Ethnicity (% BME): NR	Venlafaxine 75-225mg/day	Fluoxetine 20-40mg/day	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Remission • Response
Mehtonen 2000 RCT Finland	Outpatient N=147 Baseline severity: More severe Mean age (years): 42.6 Sex (% female): 66 Ethnicity (% BME): NR	Venlafaxine 75-150mg/day	Sertraline 50-100mg/day	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Remission • Response
Montgomery 2004 RCT Denmark, Finland, France, Germany, Ireland, Spain, & Switzerland	Outpatient N=293 Baseline severity: More severe Mean age (years): 48 Sex (% female): 71 Ethnicity (% BME): NR	Venlafaxine 75-150mg/day	Escitalopram 10-20mg/day	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Remission • Response
Mowla 2016 RCT Iran	Outpatient N=63 Baseline severity: More severe	Duloxetine 20-60mg/day	Sertraline 50-200mg/day	Treatment duration (weeks): 6 Outcomes (for inpatient versus

Study	Population	Intervention	Comparison	Comments
	Mean age (years): 41.2 Sex (% female): 60 Ethnicity (% BME): NR			outpatient subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score
Nemeroff 2007a RCT US	Outpatient N=206 Baseline severity: More severe Mean age (years): 39 Sex (% female): 65 Ethnicity (% BME): 11	Venlafaxine 75-225mg/day	Fluoxetine 20-60mg/day	Treatment duration (weeks): 6 Outcomes (for inpatient versus outpatient subgroup analysis): • Remission • Response
Nierenberg 2007a RCT US	Outpatient N=547 Baseline severity: More severe Mean age (years): 42.2 Sex (% female): 66 Ethnicity (% BME): 24	Duloxetine 60mg/day	Escitalopram 10mg/day	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score • Remission • Response
Perahia 2006a RCT Bulgaria, Croatia, Hungary, Poland, Romania, Russia, & Slovakia	Outpatient N=293 Baseline severity: More severe Mean age (years): 45.4 Sex (% female): 71 Ethnicity (% BME): 0	Duloxetine 80mg/day or 120mg/day	Paroxetine 20mg/day	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Remission • Response
Rickels 2000 RCT Country NR	Outpatient N=51 Baseline severity: More severe Mean age (years): 37.4 Sex (% female): 75 Ethnicity (% BME): NR	Venlafaxine 150-225mg/day	Fluoxetine 20-40mg/day	Treatment duration (weeks): 6 Outcome (for inpatient versus outpatient subgroup analysis): • Remission

Study	Population	Intervention	Comparison	Comments
Rudolph 1999a RCT US	Outpatient N=203 Baseline severity: More severe Mean age (years): 40 Sex (% female): 72 Ethnicity (% BME): NR	Venlafaxine 75-225mg/day	Fluoxetine 20-60mg/day	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms endpoint • Remission • Response
Shelton 2006 RCT US	Outpatient N=160 Baseline severity: More severe Mean age (years): 39.3 Sex (% female): 53 Ethnicity (% BME): 17	Venlafaxine 75-225mg/day	Sertraline 50-150mg/day	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms endpoint • Depression symptoms change score • Remission • Response
Sir 2005 RCT Australia & Turkey	Outpatient N=163 Baseline severity: More severe Mean age (years): 37 Sex (% female): 69 Ethnicity (% BME): 2	Venlafaxine 75-225mg/day	Sertraline 50-150mg/day	Treatment duration (weeks): 8 Outcomes (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score • Remission • Response
Study F1J-MC-HMAQ-Study Group Ba RCT US	Outpatient N=119 Baseline severity: More severe Mean age (years): 39.8 Sex (% female): NR Ethnicity (% BME): NR	Duloxetine 40-120mg/day	Fluoxetine 20mg/day	Treatment duration (weeks): 10 Outcomes (for inpatient versus outpatient

Study	Population	Intervention	Comparison	Comments
				subgroup analysis): • Depression symptoms change score • Remission • Response
Tylee 1997 RCT UK	Outpatient N=341 Baseline severity: More severe Mean age (years): 44.5 Sex (% female): 71 Ethnicity (% BME): NR	Venlafaxine 75mg/day	Fluoxetine 20mg/day	Treatment duration (weeks): 12 Outcomes (for inpatient versus outpatient subgroup analysis): • Remission • Response
VEN XR 367 (FDA)a RCT Europe	Outpatient N=246 Baseline severity: More severe Mean age (years): NR Sex (% female): 59 Ethnicity (% BME): NR	Venlafaxine 75mg/day or 150mg/day	Paroxetine 20mg/day	Treatment duration (weeks): 8 Outcome (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score
Wade 2007 RCT Belgium, Canada, the Czech Republic, France, Germany, Italy, Spain, Sweden & UK	Outpatient N=295 Baseline severity: More severe Mean age (years): 43.9 Sex (% female): 72 Ethnicity (% BME): 4	Duloxetine 60mg/day	Escitalopram 20mg/day	Treatment duration (weeks): 24 Outcomes (for inpatient versus outpatient subgroup analysis): • Remission • Response

a Three-armed trial but where possible the demographics reported here are for only the two relevant arms.
BME: black, minority, ethnic; mg: milligrams; N: number of participants; NR: not reported; RCT: randomised controlled trial.

Summaries of the studies included in the inpatient versus outpatient subgroup analysis of the mirtazapine versus TCAs comparison are presented in Table 12Table 18.

There was not a significant subgroup difference between inpatient and outpatient settings for the comparison mirtazapine versus TCAs on response (Test for subgroup differences: $Chi^2 = 0.19$, df = 1, p = 0.66).

Table 22: Summary of included studies for inpatient versus outpatient subgroup analysis for comparison 3e Mirtazapine versus TCAs

Study	Population	Intervention	Comparison	Comments
Inpatient setting (K=2, N=425)				
Richou 1995 RCT France	Inpatient N=174 Baseline severity: More severe Mean age (years): 50.7 Sex (% female): 67 Ethnicity (% BME): NR	Mirtazapine 20-80mg/day	Clomipramine 50-200mg/day	Treatment duration (weeks): 6 Outcome (for inpatient versus outpatient subgroup analysis): • Response
Zivkov 1995 RCT Former Yugoslavia	Inpatient N=251 Baseline severity: More severe Mean age (years): 46.9 Sex (% female): 78 Ethnicity (% BME): NR	Mirtazapine 20-60mg/day	Amitriptyline 75-225mg/day	Treatment duration (weeks): 6 Outcome (for inpatient versus outpatient subgroup analysis): • Response
Outpatient setting (K=4, N=387)				
Bremner 1995a RCT US	Outpatient N=100 Baseline severity: More severe Mean age (years): 39.0 Sex (% female): 67 Ethnicity (% BME): NR	Mirtazapine 5-35mg/day	Amitriptyline 40-280mg/day	Treatment duration (weeks): 6 Outcome (for inpatient versus outpatient subgroup analysis): • Response
MIR 003-020 (FDA)a RCT US	Outpatient N=87 Baseline severity: More severe Mean age (years): 43.5 Sex (% female): 45 Ethnicity (% BME): NR	Mirtazapine 5-35mg/day	Amitriptyline 40-280mg/day	Treatment duration (weeks): 6 Outcome (for inpatient versus outpatient subgroup analysis): • Response
MIR 003-021 (FDA)a RCT US	Outpatient N=100 Baseline severity: More severe Mean age (years): 44.5 Sex (% female): 55	Mirtazapine 5-35mg/day	Amitriptyline 40-280mg/day	Treatment duration (weeks): 6 Outcome (for inpatient versus outpatient

Study	Population	Intervention	Comparison	Comments
	Ethnicity (% BME): NR			subgroup analysis): • Response
Smith 1990a RCT US	Outpatient N=100 Baseline severity: More severe Mean age (years): NR Sex (% female): NR Ethnicity (% BME): NR	Mirtazapine 10-35mg/day	Amitriptyline 80-280mg/day	Treatment duration (weeks): 6 Outcome (for inpatient versus outpatient subgroup analysis): • Response

a Three-armed trial but where possible the demographics reported here are for only the two relevant arms.
BME: black, minority, ethnic; mg: milligrams; N: number of participants; NR: not reported; RCT: randomised controlled trial.

Summaries of the studies included in the inpatient versus outpatient subgroup analysis of the acupuncture + antidepressants versus antidepressants comparison are presented in Table 13Table 18.

There was not a significant subgroup difference between inpatient and outpatient settings for the comparison acupuncture + antidepressants versus antidepressants on depression symptoms change score (Test for subgroup differences: $Chi^2 = 1.18$, df = 1, p = 0.28).

Table 23: Summary of included studies for inpatient versus outpatient subgroup analysis for comparison 3f Acupuncture + antidepressants versus antidepressants

Study	Population	Intervention	Comparison	Comments
Inpatient setting (K=2, N=119)				
Wang 2014a RCT China	Inpatient N=77 Baseline severity: More severe Mean age (years): NR Sex (% female): 72 Ethnicity (% BME): NR	Traditional acupuncture (30 sessions) + any SSRI (dose NR)	Any SSRI (dose NR)	Treatment duration (weeks): 6 Outcome (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score
Zhang 2007a RCT China	Inpatient N=42 Baseline severity: More severe Mean age (years): 36.8 Sex (% female): 50 Ethnicity (% BME): NR	Electroacupuncture (36x 30-min sessions) + paroxetine 10-40mg/day	Paroxetine 10-40mg/day	Treatment duration (weeks): 6 Outcome (for inpatient versus outpatient

Study	Population	Intervention	Comparison	Comments
				subgroup analysis): • Depression symptoms change score
Outpatient setting (K=2, N=637)				
Qu 2013 RCT China	Outpatient N=160 Baseline severity: More severe Mean age (years): 33.3 Sex (% female): 59 Ethnicity (% BME): NR	Traditional acupuncture or electroacupuncture (18 sessions) + paroxetine 20-40mg/day	Paroxetine 20-40mg/day	Treatment duration (weeks): 6 Outcome (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score
Zhao 2019a RCT China	Outpatient N=477 Baseline severity: More severe Mean age (years): 41.5 Sex (% female): 65 Ethnicity (% BME): NR	Traditional acupuncture or electroacupuncture (18x 30-min sessions) + any SSRI (most commonly paroxetine 20mg/day)	Any SSRI (most commonly paroxetine 20mg/day)	Treatment duration (weeks): 6 Outcome (for inpatient versus outpatient subgroup analysis): • Depression symptoms change score

BME: black, minority, ethnic; mg: milligrams; N: number of participants; NR: not reported; RCT: randomised controlled trial; SSRI: selective serotonin reuptake inhibitor.

Comparison 4. Acute psychiatric day hospital care versus inpatient care (for adults with depression and non-psychotic severe mental illness)

Table 24: Summary of included studies for comparison 4 acute psychiatric day hospital versus inpatient care

Study	Population	Intervention	Comparison	Comments
Creed 1990 RCT UK	N=102 Non-psychotic severe mental illness Diagnosis: 27% schizophrenia; 20% depression; 9% mania; 27% neurotic disorder; 9% personality disorder; 8% addiction/organic disorder Mean age (years): 42.5 Sex (% female):	Acute day hospital care. Teaching hospital serving small socially deprived inner city area. Day hospital designed to take acute admissions because of few beds (8 nurses, 3 OTs)	Inpatient care (routine inpatient)	Duration of follow-up: 12 months Outcomes: • Duration of index admission • Readmission at 12 months post-admission • Social functioning response at 12 months post-admission

Study	Population	Intervention	Comparison	Comments
	51 Ethnicity (% BME): NR			
Creed 1997 RCT UK	N=187 Non-psychotic severe mental illness Diagnosis: 43% schizophrenia; 34% depression; 23% neurosis Mean age (years): 38.0 Sex (% female): 43 Ethnicity (% BME): 18	Acute day hospital care. Teaching hospital serving small socially deprived inner city area. Day hospital designed to take acute admissions because of few beds (CPN out of hours).	Inpatient care (routine inpatient)	Duration of follow-up: 12 months Outcomes: • Psychiatric symptom severity at 3 months post-admission • Psychiatric symptom severity at 12 months post-admission • Duration of index admission • Readmission at 12 months post-admission • Carer distress at 3 months post-admission • Carer distress at 12 months post-admission
Dick 1985 RCT UK	N=91 Non-psychotic severe mental illness Diagnosis: Neurosis (56% depressive neurosis), personality disorder, or adjustment reaction Mean age (years): ~35 Sex (% female): 68 Ethnicity (% BME): NR	Acute day hospital care. 2 trained staff + OT, patient/staff ratio: 12.5:1, individual counselling, groups, activities and medication	Inpatient care. Mixed sex and female wards	Duration of follow-up: 12 months Outcomes: • Readmission at 4 months post-admission • Emergency contacts at 4 months post-admission • Outpatient contact at 4 months post-admission • Satisfaction at 4 months post-admission
Dinger 2014 RCT Germany	N=44 Depression Diagnosis: 97.7% had a major depressive	Acute day hospital care. Therapeutic staff were the same for both treatment arms. Both	Inpatient care. Therapeutic staff were the same for both treatment arms. Both groups received	Duration of follow-up: 3 months Outcomes:

Study	Population	Intervention	Comparison	Comments
	episode, 2.3% had primary dysthymia Mean age (years): 35.1 Sex (% female): 50 Ethnicity (% BME): NR	groups received equal amounts of psychotherapeutic interventions. Day-clinic patients attended therapy on 5 weekdays from 8 a.m. to 4 p.m. (8 weeks of treatment)	equal amounts of psychotherapeutic interventions. Inpatients were free to leave the unit outside of night hours and therapy sessions and spent 6 weekends at home (8 weeks of treatment)	• Depression symptomatology at 3 months post-admission • Remission at 3 months post-admission • Response at 3 months post-admission
Kallert 2007 RCT Germany, UK, Poland, Slovakia and Czech Republic	N=1117 Non-psychotic severe mental illness Diagnosis: 27% schizophrenia, schizotypal, delusional, and other non-mood psychotic disorders (ICD-10 F20-F29); 41% mood [affective] disorders (ICD-10 F30-F39); 22% anxiety, dissociative, stress-related, somatoform and other nonpsychotic mental disorders (ICD-10 F40-F49); 9% disorders of adult personality and behaviour (ICD-10 F60-F69) Mean age (years): ~38 Sex (% female): 56 Ethnicity (% BME): NR	Acute day hospital care. Provided between 15 and 35 places, mean staff hours per week per treatment place ranged from 8.8 to 16.0. Staff patient ratios not reported	Inpatient care (routine inpatient)	Duration of follow-up: 14 months Outcomes: • Psychiatric symptom severity at 2 months post-admission • Psychiatric symptom severity at 14 months post-admission • Duration of index admission • Quality of life at 2 months post-admission • Quality of life at 14 months post-admission • Social functioning at 2 months post-admission • Social functioning at 14 months post-admission • Satisfaction at 2 months post-admission
Schene 1993 RCT Netherlands	N=222 Non-psychotic severe mental illness Diagnosis: 21% psychosis; 38% mood disorders; 24% anxiety disorders; 10%	Acute day hospital care. Provided 24 places. For each day treatment patient, a 0.08 full-time equivalent social	Inpatient care. Open inpatient ward with 20 beds. For each inpatient, a 0.40 full-time equivalent psychiatric nurse was available	Duration of follow-up: 13 months Outcomes: • Remission at 13 months post-admission

Study	Population	Intervention	Comparison	Comments
	eating disorders; 8% other Mean age (years): 31.9 Sex (% female): 58 Ethnicity (% BME): NR	psychiatric nurse was available		• Duration of index admission • Social functioning response at 13 months post-admission

BME: black, minority, ethnic; CPN: community psychiatric nurse; ICD: International Classification of Diseases; N: number of participants; NR: not reported; OT: occupational therapist; RCT: randomised controlled trial

Comparison 5. Non-acute day hospital care versus outpatient care (for adults with depression and non-psychotic severe mental illness)

Table 25: Summary of included studies for comparison 5 non-acute day hospital versus outpatient care

Study	Population	Intervention	Comparison	Comments
Dick 1991 RCT UK	N=96 Depression Diagnosis: 92% DSM-III major depressive disorder; 8% dysthymic disorder Mean age (years): NR Sex (% female): 75 Ethnicity (% BME): NR	Non-acute day hospital care. Places for up to 40 patients. Treatment is eclectic, with a focus on time structuring and socialisation, and a problem-orientated supportive/behavioural rather than a psychodynamic approach. Staffing comprises three sessions per week of consultant time, three sessions per week of support medical time, three full-time trained nurses, and one full-time occupational therapist. Mean duration of day treatment was 10.7 weeks	Outpatient care. Patients allocated to continued outpatient treatment were seen approximately monthly and given advice on relaxation, anxiety management, and alternative approaches to time structuring and handling relationships	Duration of follow-up: 6 months Outcomes: • Admission as an inpatient 6 months post-admission • Satisfaction at 6 months post-admission
Glick 1986 RCT US	N=79 Non-psychotic severe mental illness	Non-acute day hospital care. Transitional day care following	Outpatient care. Outpatient follow-up post-inpatient admission	Duration of follow-up: 12 months Outcomes:

Study	Population	Intervention	Comparison	Comments
	Diagnosis: 47% schizophrenia; 53% major affective disorder Mean age (years): 35 Sex (% female): 63 Ethnicity (% BME): NR	inpatient admission (about 15 hours/week and limited to 6-12 weeks) involving milieu, family, supportive & group therapy, medication, care management, recreation & dance therapy, and discharge planning	involving 6-12 weeks in outpatient group therapy (90 mins/week), medication management and 24 hour crisis intervention	• Psychiatric symptom severity at 6 months post-admission • Psychiatric symptom severity at 12 months post-admission • Admission as an inpatient 12 months post-admission • Social functioning at 6 months post-admission • Social functioning at 12 months post-admission • Global functioning at 6 months post-admission • Global functioning at 12 months post-admission
Tyrer 1979 RCT UK	N=106 Non-psychotic severe mental illness Diagnosis: Neurotic disorder (severe enough for day hospital treatment) Mean age (years): NR Sex (% female): NR Ethnicity (% BME): NR	Non-acute day hospital care. Two different types of day hospital: one specialising in neurotic disorders (well-staffed with psychotherapeutic orientation) and the other a standard day hospital (psychiatrists, nurses, occupational & art therapists)	Outpatient care (routine outpatient)	Duration of follow-up: 24 months Outcomes: • Psychiatric symptom severity at 4 months post-admission • Psychiatric symptom severity at 8 months post-admission • Admission as an inpatient 8 months post-admission • Social functioning at 4 months post-admission • Social functioning at 8

Study	Population	Intervention	Comparison	Comments
				months post-admission • Satisfaction at 4 months post-admission

BME: black, minority, ethnic; DSM Diagnostic and Statistical Manual of Mental Disorders; N: number of participants; NR: not reported; RCT: randomised controlled trial

Comparison 6. Community mental health teams versus standard care (for adults with non-psychotic severe mental illness)

Table 26: Summary of included studies for comparison 6 community mental health teams versus standard care

Study	Population	Intervention	Comparison	Comments
Merson 1992 RCT UK	N=100 Non-psychotic severe mental illness Diagnosis: 38% ICD-10 schizophrenia and related disorders; 32% mood disorder; 25% neurotic and stress-related disorders; 4% substance misuse; 1% personality disorder only Mean age (years): NR (median 32) Sex (% female): 60 Ethnicity (% BME): 32	Community mental health team (CMHT). Early intervention from a multidisciplinary community-based team, open referral, in-home assessments, collaboration maintained with already involved agencies, clinical decisions by team consensus	Standard care included conventional hospital-based psychiatric services, usually outpatient clinic assessments with occasional home visits	Duration of follow-up: 3 months Outcomes: • Psychiatric symptom severity at 3 months post-entry • Admission as an inpatient 3 months post-entry • Admission as an inpatient for >10 days at 3 months post-entry • Satisfaction (number of participants satisfied with their treatment) at 3-months post-entry • Satisfaction (service satisfaction score) at 3-months post-entry

BME: black, minority, ethnic; ICD: International Classification of Diseases; N: number of participants; NR: not reported; RCT: randomised controlled trial

See the full evidence tables in appendix D and the forest plots in appendix E.

Quality assessment of clinical outcomes included in the evidence review

See the clinical evidence profiles in appendix F.

Economic evidence

Included studies

A single economic search was undertaken for all topics included in the scope of this guideline but no economic studies were identified which were applicable to this review question. See the literature search strategy in appendix B and economic study selection flow chart in appendix G.

Excluded studies

A list of excluded economic and utility studies, with reasons for exclusion, is provided in supplement 3 - Health economic included & excluded studies.

Economic model

No economic modelling was undertaken for this review because the committee agreed that other topics were higher priorities for economic evaluation.

Evidence statements

Clinical evidence statements

Comparison 1. Primary care versus secondary care

Primary care versus secondary care subgroup analysis for Comparison 1a Cognitive and cognitive behavioural therapies individual + antidepressant versus antidepressant

Critical outcomes

Depression symptomatology

- Subgroup analysis of primary care and secondary care, for the comparison of combined individual CBT and antidepressant versus antidepressant-only, shows no statistically significant subgroup difference in depression symptomatology at endpoint for adults receiving first-line treatment for depression.

Primary care versus secondary care subgroup analysis for Comparison 1b. Selective serotonin reuptake inhibitors (SSRIs) versus placebo

Critical outcomes

Depression symptomatology

- Subgroup analysis of primary care and secondary care, for the comparison of SSRIs versus placebo, shows no statistically significant subgroup difference in depression symptomatology at endpoint, or change from baseline to endpoint, for adults receiving first-line treatment for depression.

Response

- Subgroup analysis of primary care and secondary care, for the comparison of SSRIs versus placebo, shows no statistically significant subgroup difference in the rate of response for adults receiving first-line treatment for depression.

Primary care versus secondary care subgroup analysis for Comparison 1c. SSRIs versus tricyclic antidepressants (TCAs)

Critical outcomes

Depression symptomatology

- Subgroup analysis of primary care and secondary care, for the comparison of SSRIs versus TCAs, shows no statistically significant subgroup difference in depression symptomatology at endpoint, or change from baseline to endpoint, for adults receiving first-line treatment for depression.

Remission

- Subgroup analysis of primary care and secondary care, for the comparison of SSRIs versus TCAs, shows no statistically significant subgroup difference in the rate of remission for adults receiving first-line treatment for depression.

Response

- Subgroup analysis of primary care and secondary care, for the comparison of SSRIs versus TCAs, shows no statistically significant subgroup difference in the rate of response for adults receiving first-line treatment for depression.

Primary care versus secondary care subgroup analysis for Comparison 1d. TCAs versus placebo

Critical outcomes

Depression symptomatology

- Subgroup analysis of primary care and secondary care, for the comparison of TCAs versus placebo, shows no statistically significant subgroup difference in depression symptomatology at endpoint, or change from baseline to endpoint, for adults receiving first-line treatment for depression.

Response

- Subgroup analysis of primary care and secondary care, for the comparison of TCAs versus placebo, shows no statistically significant subgroup difference in the rate of response for adults receiving first-line treatment for depression.

Primary care versus secondary care subgroup analysis for Comparison 1e. Serotonin–norepinephrine reuptake inhibitors (SNRIs) versus SSRIs

Critical outcomes

Remission

- Subgroup analysis of primary care and secondary care, for the comparison of SNRIs versus SSRIs, shows no statistically significant subgroup difference in the rate of remission for adults receiving first-line treatment for depression.

Response

- Subgroup analysis of primary care and secondary care, for the comparison of SNRIs versus SSRIs, shows no statistically significant subgroup difference in the rate of response for adults receiving first-line treatment for depression.

Comparison 2. Crisis resolution team care versus standard care (for adults with non-psychotic severe mental illness)

Critical outcomes

Psychiatric symptom severity

- Very low quality evidence from 1 RCT (N=211) shows a statistically significant but not clinically important benefit of crisis resolution team care relative to standard care on psychiatric symptom severity 8 weeks after crisis, for adults with non-psychotic severe mental illness.

Important outcomes

Service utilisation

- Very low quality evidence from 1 RCT (N=258) shows a clinically important and statistically significant benefit of crisis resolution team care relative to standard care on the rate of inpatient admission 6 months after crisis, for adults with non-psychotic severe mental illness.
- Very low quality evidence from 1 RCT (N=257) shows a statistically significant but not clinically important benefit of crisis resolution team care relative to standard care on the number of bed days in hospital 6 months after crisis, for adults with non-psychotic severe mental illness.

Psychological functioning

- Very low quality evidence from 1 RCT (N=217) shows neither a clinically important nor statistically significant difference between crisis resolution team care and standard care on quality of life 8 weeks after crisis, for adults with non-psychotic severe mental illness.

Social functioning

- Very low quality evidence from 1 RCT (N=255-257) shows neither a clinically important nor statistically significant difference between crisis resolution team care and standard care on social functioning at 8 weeks or 6 months after crisis, for adults with non-psychotic severe mental illness.

Satisfaction

- Very low quality evidence from 1 RCT (N=226) shows neither a clinically important nor statistically significant difference between crisis resolution team care relative and standard care on patient satisfaction ratings 8 weeks after crisis, for adults with non-psychotic severe mental illness.

Comparison 3. Inpatient versus outpatient settings

Inpatient versus outpatient subgroup analysis for Comparison 3a Selective serotonin reuptake inhibitors (SSRIs) versus placebo

Critical Outcomes

Depression symptomatology

- Subgroup analysis of inpatient and outpatient settings, for the comparison of SSRIs versus placebo, shows no statistically significant subgroup difference in

depression symptomatology change score for adults receiving first-line treatment for depression.

Response

- Subgroup analysis of inpatient and outpatient settings, for the comparison of SSRIs versus placebo, shows no statistically significant subgroup difference in the rate of response for adults receiving first-line treatment for depression.

Inpatient versus outpatient subgroup analysis for Comparison 3b SSRIs versus Tricyclic Antidepressants (TCAs)

Critical Outcomes

Depression symptomatology

- Subgroup analysis of inpatient and outpatient settings, for the comparison of SSRIs versus TCAs, shows no statistically significant subgroup difference in depression symptomatology at endpoint for adults receiving first-line treatment for depression.
- Subgroup analysis of inpatient and outpatient settings, for the comparison of SSRIs versus TCAs, shows a statistically significant subgroup difference in depression symptomatology change score for adults receiving first-line treatment for depression. In inpatient settings TCAs show a small benefit over SSRIs, and in outpatient settings SSRIs show a small benefit over TCAs, however, in both inpatient and outpatient settings the difference between TCAs and SSRIs is non-significant.

Remission

- Subgroup analysis of inpatient and outpatient settings, for the comparison of SSRIs versus TCAs, shows no statistically significant subgroup difference in the rate of remission for adults receiving first-line treatment for depression.

Response

- Subgroup analysis of inpatient and outpatient settings, for the comparison of SSRIs versus TCAs, shows no statistically significant subgroup difference in the rate of response for adults receiving first-line treatment for depression.

Inpatient versus outpatient subgroup analysis for Comparison 3c Serotonin–norepinephrine reuptake inhibitors (SNRIs) versus placebo

Critical Outcomes

Depression symptomatology

- Subgroup analysis of inpatient and outpatient settings, for the comparison of SNRIs versus placebo, shows no statistically significant subgroup difference in depression symptomatology at endpoint for adults receiving first-line treatment for depression.
- Subgroup analysis of inpatient and outpatient settings, for the comparison of SNRIs versus placebo, shows no statistically significant subgroup difference in depression symptomatology change scores for adults receiving first-line treatment for depression.

Remission

- Subgroup analysis of inpatient and outpatient settings, for the comparison of SNRIs versus placebo, shows no statistically significant subgroup difference in the rate of remission for adults receiving first-line treatment for depression.

Inpatient versus outpatient subgroup analysis for Comparison 3d SNRIs versus SSRIs

Critical Outcomes

Depression symptomatology

- Subgroup analysis of inpatient and outpatient settings, for the comparison of SNRIs versus SSRIs, shows no statistically significant subgroup difference in depression symptomatology at endpoint for adults receiving first-line treatment for depression.
- Subgroup analysis of inpatient and outpatient settings, for the comparison of SNRIs versus SSRIs, shows a statistically significant subgroup difference in depression symptomatology change score for adults receiving first-line treatment for depression. In both inpatient and outpatient settings SNRIs show a benefit over SSRIs however this effect is larger in inpatient relative to outpatient settings, although this is a difference in magnitude rather than direction and even in inpatient settings the difference is not clinically important.

Remission

- Subgroup analysis of inpatient and outpatient settings, for the comparison of SNRIs versus SSRIs, shows no statistically significant subgroup difference in the rate of remission for adults receiving first-line treatment for depression.

Response

- Subgroup analysis of inpatient and outpatient settings, for the comparison of SNRIs versus SSRIs, shows no statistically significant subgroup difference in the rate of response for adults receiving first-line treatment for depression.

Inpatient versus outpatient subgroup analysis for Comparison 3e Mirtazapine versus TCAs

Critical Outcomes

Response

- Subgroup analysis of inpatient and outpatient settings, for the comparison of mirtazapine versus TCAs, shows no statistically significant subgroup difference in the rate of response for adults receiving first-line treatment for depression.

Inpatient versus outpatient subgroup analysis for Comparison 3f Acupuncture + antidepressants versus antidepressants

Critical Outcomes

Depression symptomatology

- Subgroup analysis of inpatient and outpatient settings, for the comparison of combined acupuncture and antidepressant versus antidepressants-only, shows

no statistically significant subgroup difference in depression symptomatology change score for adults receiving first-line treatment for depression.

Comparison 4. Acute psychiatric day hospital care versus inpatient care (for adults with depression and non-psychotic severe mental illness)

Critical outcomes

Psychiatric symptom severity

- Very low quality evidence from 2-3 RCTs (N=1249-1281) shows neither clinically important nor statistically significant differences between acute day hospital care compared to inpatient care on psychiatric symptom severity at 2-3 months or 12-14 months post-admission, for adults with depression or non-psychotic severe mental illness.

Remission

- Very low quality evidence from 2 RCTs (N=151) shows neither clinically important nor statistically significant effects differences between acute day hospital care compared to inpatient care on the rate of remission at 3 or 13 months post-admission, for adults with depression or non-psychotic severe mental illness.

Response

- Very low quality evidence from 1 RCT (N=44) including only adults with depression shows a clinically important but not statistically significant benefit of inpatient care relative to acute day hospital care on the rate of response at 3 months post-admission.

Important outcomes

Service utilisation

- Very low quality evidence from 4 RCTs (N=1535) shows a clinically important and statistically significant benefit of inpatient care, relative to acute day hospital care, on the duration of index admission for adults with depression or non-psychotic severe mental illness.
- Very low quality evidence from 3 RCTs (N=372) shows a clinically important but not statistically significant benefit of acute day hospital care relative to inpatient care on readmission at 4 months or 12 months post-admission, for adults with depression or non-psychotic severe mental illness.
- Very low quality evidence from 1 RCT (N=83) shows clinically important but not statistically significant benefits of inpatient care relative to acute day hospital care on the number of emergency contacts and the number of outpatient contacts, for adults with non-psychotic severe mental illness.

Psychological functioning

- Very low quality evidence from 1 RCT (N= 1117) shows neither clinically important nor statistically significant differences between acute day hospital care compared to inpatient care on quality of life at 2 or 14 months post-admission, for adults with non-psychotic severe mental illness.

Social functioning

- Very low quality evidence from 1 RCT (N= 1117) shows a statistically significant but not clinically important benefit of acute day hospital care relative to inpatient care on social functioning impairment at 2 and 14 months post-admission, for adults with non-psychotic severe mental illness.
- Very low quality evidence from 2 RCTs (N=181) shows a clinically important but not statistically significant benefit of acute day hospital care relative to inpatient care on the number of people achieving significant improvement in social functioning at 12-13 months post-admission, for adults with non-psychotic severe mental illness.

Satisfaction

- Very low quality evidence from 1 RCT (N= 83) shows a clinically important and statistically significant benefit of acute day hospital care relative to inpatient care in the number of people who are satisfied or very satisfied with their treatment, for adults with non-psychotic severe mental illness.
- Very low quality evidence from 1 RCT (N=1117) shows neither clinically important nor statistically significant differences between acute day hospital care compared to inpatient care on patient satisfaction ratings at 2 months post-admission, for adults with non-psychotic severe mental illness.

Carer distress

- Very low quality evidence from 1 RCT (N=55-77) shows neither clinically important nor statistically significant differences between acute day hospital care compared to inpatient care on carer distress at 3 or 12 months post-admission, for adults with non-psychotic severe mental illness.

Comparison 5. Non-acute day hospital care versus outpatient care (for adults with depression and non-psychotic severe mental illness)

Critical outcomes

Psychiatric symptom severity

- Low to very low quality evidence from 2 RCTs (N=139-144) shows neither clinically important nor statistically significant differences between non-acute day hospital care compared to outpatient care on psychiatric symptom severity at 4-6 months and 8-12 months post-admission, for adults with non-psychotic severe mental illness.

Important outcomes

Service utilisation

- Very low quality evidence from 3 RCTs (N=281) shows a clinically important but not statistically significant benefit of outpatient care relative to non-acute day hospital care on the number of people admitted as an inpatient at 6-12 months post-admission, for adults with non-psychotic severe mental illness.

Social functioning

- Very low quality evidence from 2 RCTs (N=141) shows neither clinically important nor statistically significant differences between non-acute day

hospital care compared to outpatient care on social functioning at 4-6 or 8-12 months post-admission, for adults with non-psychotic severe mental illness.

- Very low quality evidence from 1 RCT (N=51-52) shows neither clinically important nor statistically significant differences between non-acute day hospital care compared to outpatient care on global functioning at 6 and 12 months post-admission, for adults with non-psychotic severe mental illness.

Satisfaction

- Very low quality evidence from 2 RCTs (N=198) shows neither clinically important nor statistically significant differences between non-acute day hospital care compared to outpatient care on the number of people satisfied or very satisfied with their treatment at 4-6 months post-admission, for adults with non-psychotic severe mental illness.

Comparison 6. Community mental health teams versus standard care (for adults with non-psychotic severe mental illness)

Critical outcomes

Psychiatric symptom severity

- Low quality evidence from 1 RCT (N=100) shows neither a clinically important nor statistically significant difference between community mental health team care compared to standard care on psychiatric symptom severity at 3 months post-entry, for adults with non-psychotic severe mental illness.

Important outcomes

Service utilisation

- Very low quality evidence from 1 RCT (N=100) shows a clinically important but not statistically significant benefit of community mental health team care relative to standard care on the number of people admitted to inpatient care, and a clinically important and statistically significant benefit on the number of people admitted to inpatient care for longer than 10 days, for adults with non-psychotic severe mental illness.

Satisfaction

- Very low quality evidence from 1 RCT (N=87) shows clinically important and statistically significant benefits of community mental health team care, relative to standard care, on both continuous and dichotomous measures of satisfaction for adults with non-psychotic severe mental illness.

Economic evidence statements

No economic evidence was identified which was applicable to this review question.

The committee's discussion of the evidence

Interpreting the evidence

The outcomes that matter most

The aim of this review was to determine if different settings for the delivery of care improved outcomes for people with depression so the committee identified depression symptomatology, response, remission and relapse to be the critical outcomes for this question. If the evidence specific to depression was limited, it was pre-defined in the protocol that the inclusion criteria would be expanded to include those with non-psychotic severe mental illness, and for these populations psychiatric symptom severity was a critical outcome. Service utilisation and resource use were identified as important outcomes, as a measure of uptake and persistence with treatment. Psychological functioning, social functioning, satisfaction, and carer distress were also considered important outcomes, in order to assess the broader impact of setting on the person with depression and their family or carer.

For all comparisons there was evidence for at least one critical outcome – most commonly symptom severity – and at least one important outcome. Carer distress was rarely reported and this outcome was only available for comparison 4.

The quality of the evidence

The committee noted that all outcomes had been assessed as either very low or low in GRADE. Most outcomes were downgraded due to imprecision and/or risk of bias. A number of the comparisons also included people with non-psychotic severe mental illness, and so were not specific to the population of people with depression, and these comparisons were downgraded again due to indirectness.

Benefits and harms

The comparisons included in this review included a number of different settings such as the primary care setting (where people are living in their own home and are cared for by their GPs), and a number of different secondary care or specialist services, where care is provided to people in their own homes, as outpatients, or as inpatients.

During the protocol development, the committee had noted that the best evidence to examine the benefits and harms associated with settings would require randomised controlled trials (RCTs) that randomised the same population to different settings for the delivery of care. However, trials of interventions delivered in certain settings will recruit populations considered to be relevant to that setting. Evidence is particularly limited where the comparison includes inpatient care, as the large majority of people with depression are never admitted to hospital. The committee therefore agreed to consider a wider evidence base for settings where there was limited direct RCT evidence by including evidence on the care of people with severe, non-psychotic mental illness as well as or instead of, those with depression. The committee also agreed that where specific RCT evidence was limited for particular comparisons, indirect evidence in the form of subgroup analyses of the NMA dataset (Evidence report B: Treatment of a new episode of depression) may be informative.

For crisis resolution team care, no RCT evidence was identified that specifically addressed this setting for adults with depression, and only 1 RCT was identified that included people with severe non-psychotic mental illness. The evidence showed a small but statistically significant benefit of crisis resolution team care (relative to standard care) on psychiatric symptom severity, and benefits in terms of service utilisation (on the number of people admitted as an inpatient, and bed days in

hospital). Based on their experience, the committee recognised the potential benefits that crisis resolution team care may bring to adults with severe depression (particularly those at significant risk of harming themselves through suicide attempts or self-neglect) in providing an alternative to inpatient treatment and thus potentially avoiding the stigma and costs associated with hospital admission. They also recognised that crisis resolution and home treatment team care may have an important role in supporting people at home after an inpatient stay and so facilitate an early discharge, reducing the likelihood of a readmission to hospital. The committee therefore included in their recommendations some guidance on the type of people with depression who should be seen by crisis resolution teams, and what that care should involve. However, given the limited and indirect evidence base, the committee agreed that a 'consider' rather than 'offer' recommendation was appropriate.

There was no specific RCT evidence for inpatient settings. Therefore the committee considered indirect evidence in the form of subgroup analyses of the NMA dataset (acute treatment of depressive episodes). Differences between delivery in inpatient and outpatient settings were explored for depression symptomatology, remission, and response for all treatment comparisons with at least 2 studies in each subgroup (SSRIs versus placebo; SSRIs versus TCAs; SNRIs versus placebo; SNRIs versus SSRIs; mirtazapine versus TCAs; acupuncture + antidepressant versus antidepressant). Most subgroup differences were non-significant. There was, however, a statistically significant subgroup difference between inpatient and outpatient settings for depression change score for the SSRIs versus TCAs comparison, with TCAs showing a small benefit over SSRIs in inpatient settings and SSRIs showing a small benefit over TCAs in outpatient settings, however, the difference between TCAs and SSRIs was non-significant in both inpatient and outpatient settings. There was also a statistically significant subgroup difference between inpatient and outpatient settings for depression change score for the SNRIs versus SSRIs comparison, however, this was a difference in magnitude rather than direction with a benefit of SNRIs relative to SSRIs shown in both inpatient and outpatient settings but larger effects shown in inpatient settings. Despite the lack of evidence for clear clinical benefits associated with inpatient care, the committee drew on their clinical knowledge and expertise, and recognised that inpatient care may be necessary for people with more severe depression who could not be adequately supported by a crisis resolution and home treatment team, particularly if they were socially isolated, and so they made a recommendation to this effect.

For primary care compared to secondary care, no RCT evidence was identified that specifically addressed this setting. Therefore the committee considered indirect evidence in the form of subgroup analyses of the NMA dataset (acute treatment of depressive episodes). For all valid treatment comparisons (at least 2 studies per subgroup), subgroup analyses compared whether different outcomes were associated with delivery of treatment in primary compared to secondary care. For all comparisons (combined individual CBT and antidepressant versus antidepressant-only; SSRIs versus placebo; SSRIs versus TCAs; TCAs versus placebo; SNRIs versus SSRIs) there was no good evidence to show any difference between delivery in primary care or secondary care on depression symptomatology, response, or remission. Based on this evidence and their knowledge and experience, the committee agreed that there was no need to add a recommendation that specified whether interventions should be delivered in primary or secondary care, except where there were safety concerns for certain pharmacological interventions but this was captured in the specific treatment recommendations.

For all other comparisons, very few RCTs were identified that included only adults with depression (only 2 RCTs across 2 separate comparisons of non-acute day hospital versus outpatient care, and acute psychiatric day hospital versus inpatient

care), and a wider evidence base including those with non-psychotic severe mental illness was considered. For acute psychiatric day hospital care (relative to inpatient care), non-acute day hospital care (relative to outpatient care), and community mental health team care (relative to standard care) no significant (both clinically important and statistically significant) differences were shown for the critical outcomes of psychiatric symptom severity, remission or response. No eligible evidence was identified for specialist tertiary affective disorders settings or residential settings. On the basis of the limited evidence base, the committee agreed that there were no grounds (including their clinical knowledge and experience) on which to base a recommendation that care for people with depression should be delivered in these specific settings.

The committee raised the importance of equity of access to interventions in inpatient care that is equivalent to those available in community settings. They therefore recommended that the full range of psychological interventions available in community settings should also be available in inpatient settings. They also recognised that the intensity and/or duration of these interventions may need to be altered commensurate with the level of severity and need in inpatient settings.

Cost effectiveness and resource use

No evidence on the cost-effectiveness of different settings for the delivery of care for adults with depression was identified and no further economic analysis was undertaken. The committee considered the costs associated with crisis resolution and home treatment and estimated that these are higher than routine primary care but significantly lower than inpatient care. The committee expressed the opinion that, compared with routine primary care, crisis resolution treatment is often more appropriate for people with more severe depression who are at significant risk of suicide, harm to self or to others, self-neglect or complications in response to their treatment, leading to better outcomes and reduced need for more costly inpatient care.

The committee took into account the high costs associated with inpatient care, and decided to recommend inpatient treatment only for people with more severe depression who cannot be adequately supported by a crisis resolution and home treatment team.

Considering the benefits and costs of crisis resolution and home treatment teams (CRHT teams) relative to other care settings, the committee expressed the opinion that CRHT comprises an effective and likely cost-effective model of care for people with depression who would benefit from early discharge from hospital after a period of inpatient care.

The committee took into account the cost effectiveness of psychological treatments in the acute treatment of people with depression based on the results of the economic analysis undertaken for this guideline (Evidence report B: Treatment of a new episode of depression), and expressed the view that the full range of such treatments should also be available in inpatient settings, to allow provision of clinically and cost-effective care in populations treated in such settings. The committee acknowledged the fact that increasing the intensity and duration of psychological interventions for people with depression in inpatient settings has resource implications, but expressed the view that the benefits of more intensive treatment in this group would outweigh the additional intervention costs. Moreover, if improved outcomes result in earlier discharge, then cost-savings may outweigh the intervention costs of more intensive psychological treatment.

Recommendations supported by this evidence review

This evidence review supports recommendations 1.15.11 to 1.15.14 in the NICE guideline.

References

This reference list does not include studies analysed as a sub-set of the NMA data for comparisons 1 and 3. Please see evidence review B.

Creed 1990

Creed F, Black D, Anthony P, Osborn M, Thomas P, Tomenson B. Randomised controlled trial of day patient versus inpatient psychiatric treatment. Br Med J. 1990 Apr 21;300(6731):1033-7.

Creed 1997

Creed F, Mbaya P, Lancashire S, Tomenson B, Williams B, Holme S. Cost effectiveness of day and inpatient psychiatric treatment: results of a randomised controlled trial. Br Med J. 1997 May 10;314(7091):1381.

Dick 1985

Dick P, Cameron L, Cohen D, Barlow M, Ince AN. Day and full time psychiatric treatment: a controlled comparison. The British Journal of Psychiatry. 1985 Sep 1;147(3):246-9.

Dick 1991

Dick PH, Sweeney ML, Crombie IK. Controlled comparison of day-patient and out-patient treatment for persistent anxiety and depression. The BritishJjournal of Psychiatry. 1991 Jan 1;158(1):24-7.

Dinger 2014

Dinger U, Klipsch O, Köhling J, Ehrenthal JC, Nikendei C, Herzog W, Schauenburg H. Day-clinic and inpatient psychotherapy for depression (DIP-D): a randomized controlled pilot study in routine clinical care. Psychotherapy andPpsychosomatics. 2014 Apr 17;83(3):194-5.

Glick 1986

Glick ID, Fleming L, DeChillo N, Meyerkopf N, Jackson C, Muscara D, Good-Ellis M. A controlled study of transitional day care for non-chronically ill patients. American Journal of Psychiatry. 1986 Dec 1;143(12):1551-6.

Johnson 2005

Johnson S, Nolan F, Pilling S, Sandor A, Hoult J, McKenzie N, White IR, Thompson M, Bebbington P. Randomised controlled trial of acute mental health care by a crisis resolution team: the north Islington crisis study. Br Med J. 2005 Sep 15;331(7517):599.

Kallert 2007

Kallert TW, Priebe S, McCabe R, Kiejna A, Rymaszewska J, Nawka P, Ocvár L, Raboch J, Stárková-Kalisová L, Koch R, Schutzwohl M. Are Day Hospitals Effective for Acutely III Psychiatric Patients? A European Multicenter Randomized Controlled Trial. Journal of Clinical Psychiatry. 2007 Feb 15;68(2):278-87.

Merson 1992

Merson S, Tyrer P, Lack S, Birkett P, Lynch S, Onyett S, Johnson T. Early intervention in psychiatric emergencies: a controlled clinical trial. The Lancet. 1992 May 30;339(8805):1311-4.

Schene 1993

Schene AH, Wijngaarden BV, Poelijoe NW, Gersons BP. The Utrecht comparative study on psychiatric day treatment and inpatient treatment. Acta Psychiatrica Scandinavica. 1993 Jun 1;87(6):427-36.

Tyrer 1979

Tyrer PJ, Remington M. Controlled comparison of day-hospital and outpatient treatment for neurotic disorders. The Lancet. 1979 May 12;313(8124):1014-6.

Appendices

Appendix A – Review protocols

Review protocol for review question 1.1: For adults with depression, what are the relative benefits and harms associated with different models for the coordination and delivery of services?

Table 27: Review protocol for different models of care

Field (based on PRISMA-P)	Content
Review question	For adults with depression, what are the relative benefits and harms associated with different models for the coordination and delivery of services?
Type of review question	Intervention review
Objective of the review	To identify the optimal model of delivery of services for adults with an acute episode of depression, or adults whose depression has responded fully or partially to treatment.
Population	• Adults with a diagnosis of depression according to DSM, ICD or similar criteria, or depressive symptoms as indicated by baseline depression scores on validated scales (and including those with subthreshold [just below threshold] depressive symptoms) For studies on relapse prevention: • Adults whose depression has responded to treatment (in full or partial remission) according to DSM, ICD or similar criteria, or indicated by below clinical threshold depression symptom scores on validated scales If some, but not all, of a study's participants are eligible for the review, for instance, mixed anxiety and depression diagnoses, then we will include a study if at least 80% of its participants are eligible for this review
Exclude	• Trials of women with antenatal or postnatal depression • Trials of children and young people (mean age under 18 years) • Trials of people with learning disabilities • Trials of adults in contact with the criminal justice system (not solely as a result of being a witness or victim)

Field (based on PRISMA-P)	Content
	• Trials that specifically recruit participants with a physical health condition in addition to depression (e.g. depression in people with diabetes)
Intervention	Models for the coordination and delivery of services: • Collaborative care (simple and complex) • Stepped care • Medication management • Attached professional model • Care coordination • Integrated care pathways (including primary care liaison or shared care) • Measurement-based care
Comparison	• Treatment as usual • Waitlist • Any other service delivery model
Outcomes and prioritisation	**Critical outcomes:** • Depression symptomatology (mean endpoint score or change in depression score from baseline) • Response (usually defined as at least 50% improvement from the baseline score on a depression scale) • Remission (usually defined as a score below clinical threshold on a depression scale) • Relapse (number of people who returned to a depressive episode whilst in remission) The following depression scales will be included in the following hierarchy: • MADRS • HAMD • QIDS • PHQ • CGI (for dichotomous outcomes only) • CES-D • BDI • HADS-D (depression subscale)

Field (based on PRISMA-P)	Content
	Important outcomes: • Antidepressant use • Discontinuation due to any reason Outcomes will be assessed at 6 months and 12 months.
Study design	• RCTs • Systematic reviews of RCTs
Include unpublished data?	Conference abstracts, dissertations and unpublished data will not be included unless the data can be extracted from elsewhere (for instance, from the previous guideline)
Restriction by date?	All relevant studies from existing reviews from the 2009 guideline and from previous searches (pre-2016) will be carried forward. No restriction on date for the updated search, studies published between database inception and the date the searches are run will be sought.
Minimum sample size	• Minimum sample size N = 10 in each arm • Studies with <50% completion data (drop out of >50%) will be excluded
Study setting	Primary, secondary, tertiary and social care settings. Non-English-language papers will be excluded (unless data can be obtained from an existing review).
Review strategy	**Coding Strategy** For this review, a coding system for classifying the complexity and type of service delivery model has been developed specifically for the purpose of this guideline. The service delivery model described in each study will be rated on this 17-item coding system which will generate an overall rating between 0-20 (see Table 1). Service delivery models which score above 6 will be considered a collaborative care intervention; those scoring 13+ will be coded as complex collaborative care and those scoring 6-12 will be coded as simple collaborative care. Service delivery models that score below 6 will be classified as an alternative service delivery model (e.g. care coordination) or a stand-alone psychological intervention (e.g. self-help with support). **Data Extraction (selection and coding)** Citations from each search will be downloaded into EndNote and duplicates removed. Titles and abstracts of identified studies will be screened by two reviewers for inclusion against criteria, until a good inter-rater reliability has been observed (percentage agreement =>90%). Initially 10% of references will be double-

Field (based on PRISMA-P)	Content
	screened. If inter-rater agreement is good then the remaining references will be screened by one reviewer. All primary-level studies included after the first scan of citations will be acquired in full and re-evaluated for eligibility at the time they are being entered into a study database (standardised template created in Microsoft Excel). At least 10% of data extraction will be double-coded. Discrepancies or difficulties with coding will be resolved through discussion between reviewers or the opinion of a third reviewer will be sought. **Data Analysis** A meta-analysis using a random-effects model will be conducted to combine results from similar studies. An intention to treat (ITT) approach will be taken where possible. Risk of bias will be assessed at the study level using the Cochrane risk of bias tool. This assessment includes: adequacy of randomisation (sufficient description of randomisation method, allocation concealment and any baseline difference between groups); blinding (of participants, intervention administrators and outcome assessors); attrition ('at risk of attrition bias' defined as a dropout of more than 20% and completer analysis used, or a difference of >20% between the groups); selective reporting bias (is the protocol registered, are all outcomes reported); other bias (for instance, conflict of interest in funding). Risk of bias will also be assessed at the outcome level using GRADE. For heterogeneity, outcomes will be downgraded once if I^2>50%, twice if I^2 >80%. For imprecision, outcomes will be downgraded using rules of thumb. If the 95% CI is imprecise i.e. crosses the line of no effect and the threshold for clinical benefit/harm, 0.8 or 1.25 (dichotomous) or -0.5 or 0.5 SMD (for continuous), the outcome will be downgraded. Outcomes will be downgraded one or two levels depending on how many lines it crosses. If the 95% CI is not imprecise, we will consider whether the criterion for Optimal Information Size is met (for dichotomous outcomes, 300 events; for continuous outcomes, 400 participants), if not we will downgrade one level. **Coding system for service delivery models** **Collaborative Care Component Score Method**

Field (based on PRISMA-P)	Content

Item	Score
1. Active and integrated case recognition/identification* (Systematic identification- from a clinical database or screened positive for depression)	0 1
2. Collaborative assessment and plan included (Collaborative assessment with the patient)	0 1
3. Case Management (Case manager present- can include pharmacist for medication management)	0 1
4. Active liaison with primary care and other services (System set up for structured liaison/ regular meetings)	0 1
5. Case Manager has MH background (A prior mental health background, not just training in mental health)	0 1
6. Supervision provided for case manager	0 1
7. Senior MH professional consultation/involvement (Broad definition- just need to be available)	0 1
8. Psychoeducation delivered	0 1
9. Algorithm(s) used to determine care*	0 1
10. Integration with physical health care where necessary	0 1
11. Social/psychosocial interventions provided	0 1
12. Case manager delivers intervention	0 1
13. Medication management provided	0 1
14. Routine outcome monitoring (Scheduled, using a tool)	0 1
15. Psychological interventions provided None Low intensity	 0 1

Field (based on PRISMA-P)	Content
	<table><tr><td>High intensity</td><td>2</td></tr><tr><td>16. Duration of programme contact ≤6 months 7-12months 1year plus</td><td> 0 1 2</td></tr><tr><td>17. Number of sessions (F-t-F and Telephone) ≤6 sessions 6 – 12 sessions 13 + sessions</td><td> 0 1 2</td></tr><tr><td>Total (maximum 20)</td><td></td></tr><tr><td colspan="2">*Including stepped care Rating <5 – not collaborative care 6-12 – simple collaborative care 13+ – complex collaborative care</td></tr></table>
Heterogeneity (sensitivity analysis and subgroups)	Where possible, the influence of the following subgroups will be considered: For the review of collaborative care only: • Type of collaborative care (simple vs complex) • Stepped care component included in collaborative care intervention • Case manager background • Psychological interventions delivered as part of the model of care • Number of contacts/sessions/follow-up visits provided as part of intervention (less than 13 sessions, 13+ sessions) For all reviews: • Chronic depression • Depression with coexisting personality disorder • Psychotic depression • Older adults • BME populations

Field (based on PRISMA-P)	Content
	• Men
Data management (software)	Endnote was used to sift through the references identified by the search, Excel was used for data extraction Pairwise meta-analyses and production of forest plots was done using Cochrane Review Manager (RevMan5). 'GRADEpro' was used to assess the quality of evidence for each outcome.
Notes	The committee identified one good quality systematic review of RCTs (Coventry et al., 2014) which reviewed collaborative care interventions. The review was used as a source to identify any additional eligible studies *Coventry PA, Hudson JL, Kontopantelis E, Archer J, Richards DA, et al. (2014) Characteristics of Effective Collaborative Care for Treatment of Depression: A Systematic Review and Meta-Regression of 74 Randomised Controlled Trials. PLoS ONE 9(9): e108114.* Separate reviews (if applicable) will be conducted for service delivery models which were aimed at: • Treating an episode of depression • Preventing relapse of a future episode of depression
Information sources – databases and dates	Database(s): Embase 1974 to Present, Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R) 1946 to Present; Cochrane Library; WEB OF SCIENCE
Identify if an update	Update of CG90 (2009)
Author contacts	For details please see the guideline in development web site.
Highlight if amendment to previous protocol	For details please see section 4.5 of Developing NICE guidelines: the manual 2014
Search strategy – for one database	For details please see appendix B.
Data collection process – forms/duplicate	A standardised evidence table format will be used, and published as appendix D (clinical evidence tables) or H (economic evidence tables).
Data items – define all variables to be collected	For details please see evidence tables in appendix D (clinical evidence tables) or H (economic evidence tables).
Methods for assessing bias at outcome/study level	Standard study checklists were used to critically appraise individual studies. For details please see section 6.2 of Developing NICE guidelines: the manual 2014.

Field (based on PRISMA-P)	Content
	The risk of bias across all available evidence was evaluated for each outcome using an adaptation of the 'Grading of Recommendations Assessment, Development and Evaluation (GRADE) toolbox' developed by the international GRADE working group http://www.gradeworkinggroup.org/.
Criteria for quantitative synthesis	For details please see section 6.4 of Developing NICE guidelines: the manual 2014
Methods for quantitative analysis – combining studies and exploring (in)consistency	For details please see the methods chapter.
Meta-bias assessment – publication bias, selective reporting bias	For details please see section 6.2 of Developing NICE guidelines: the manual 2014.
Confidence in cumulative evidence	For details please see sections 6.4 and 9.1 of Developing NICE guidelines: the manual 2014
Rationale/context – what is known	For details please see the introduction to the evidence review.
Describe contributions of authors and guarantor	A multidisciplinary committee developed the evidence review. The committee was convened by the National Guideline Alliance (NGA) and chaired by Dr Navneet Kapur in line with section 3 of Developing NICE guidelines: the manual 2014. Staff from the NGA undertook systematic literature searches, appraised the evidence, conducted meta-analysis and cost effectiveness analysis where appropriate, and drafted the guideline in collaboration with the committee. For details please see the methods chapter.
Sources of funding/support	The NGA is funded by NICE and hosted by the Royal College of Obstetricians and Gynaecologists.
Name of sponsor	The NGA is funded by NICE and hosted by the Royal College of Obstetricians and Gynaecologists.
Roles of sponsor	NICE funds NGA to develop guidelines for those working in the NHS, public health and social care in England
PROSPERO registration number	CRD42019151323

BDI: Beck Depression Inventory; BME: black, minority, ethnic; CDSR: Cochrane Database of Systematic Reviews; CENTRAL: Cochrane Central Register of Controlled Trials; CES-D: Centre of Epidemiology Studies – Depression; CGI: Clinical Global Impressions; CI: confidence interval; DARE: Database of Abstracts of Reviews of Effects; DSM: Diagnostic and Statistical Manual of Mental Disorders; GRADE: Grading of Recommendations Assessment, Development and Evaluation; HADS-D: Hospital Anxiety and Depression Scale (-Depression); HAMD: Hamilton Depression Rating Scale ; ICD: International Statistical Classification of Diseases;ITT: intention to treat; MADRS: Montgomery–Åsberg Depression Rating Scale; N: number; NGA: National Guideline Alliance; NHS: National health service; NICE: National Institute for Health and Care Excellence; PHQ: Patient Health Questionnaire; QIDS: Quick Inventory of Depressive Symptomatology; RCT: randomised controlled trial; RoB: risk of bias; SMD: standardised mean difference;

Review protocol for review question 1.2 For adults with depression, what are the relative benefits and harms associated with different settings for the delivery of care?

Table 28: Review protocol for different settings for the delivery of care

Field (based on PRISMA-P)	Content
Review question	For adults with depression, what are the relative benefits and harms associated with different settings for the delivery of care?
Type of review question	Intervention review
Objective of the review	To identify the optimal settings for the delivery of care for adults with depression
Population	• Adults with a diagnosis of depression according to DSM, ICD or similar criteria, or depressive symptoms as indicated by baseline depression scores on validated scales (and including those with subthreshold [just below threshold] depressive symptoms) • If the evidence specific to depression is limited then the inclusion criteria may be expanded to include those with non-psychotic severe mental illness. • If some, but not all, of a study's participants are eligible for the review, then we will include a study if the majority (at least 51%) of its participants are eligible for this review.
Exclude	• Trials of women with antenatal or postnatal depression • Trials of children and young people (mean age under 18 years) • Trials of people with learning disabilities • Trials of adults in contact with the criminal justice system (not solely as a result of being a witness or victim) • Trials that specifically recruit participants with a physical health condition in addition to depression (e.g. depression in people with diabetes)
Intervention	Settings for the delivery of care, which may include: • Primary care • Crisis resolution and home treatment teams • Inpatient setting • Acute psychiatric day hospital care • Non-acute day hospital care and recovery centres

Field (based on PRISMA-P)	Content
	• Specialist tertiary affective disorders settings • Community Mental Health Teams • Residential services •
Comparison	Any other setting for the delivery of care
Outcomes and prioritisation	Critical outcomes: • Depression symptomatology (mean endpoint score or change in depression score from baseline) • Response (usually defined as at least 50% improvement from the baseline score on a depression scale) • Remission (usually defined as a score below clinical threshold on a depression scale) • Relapse (number of people who returned to a depressive episode whilst in remission) **Important outcomes:** • Service utilisation/resource use (e.g. antidepressant use) • Psychological functioning • Social functioning • Satisfaction • Carer distress Outcomes will be assessed at endpoint and follow-up.
Study design	Only published full-text papers of the following types of studies: systematic reviews of RCTs; RCTs If no RCT evidence is identified that specifically addresses the following settings: primary care, and inpatient care, then indirect evidence will be considered in the form of sub-analyses of the NMA dataset (first-line treatment of depressive episodes)
Include unpublished data?	Conference abstracts, dissertations and unpublished data will not be included unless the data can be extracted from elsewhere (for instance, from the previous guideline)
Restriction by date?	All relevant studies from existing reviews from the 2009 guideline and from previous searches (pre-2016) will be carried forward. No restriction on date for the updated search, studies published between database inception and the date the searches are run will be sought.
Minimum sample size	• Minimum sample size N = 10 in each arm • Studies with <50% completion data (drop out of >50%) will be excluded

Field (based on PRISMA-P)	Content
Study setting	Primary, secondary, tertiary and social care settings. Non-English-language papers will be excluded (unless data can be obtained from an existing review).
Review strategy	Data Extraction (selection and coding) Citations from each search will be downloaded into EndNote and duplicates removed. Titles and abstracts of identified studies will be screened by two reviewers for inclusion against criteria, until a good inter-rater reliability has been observed (percentage agreement =>90%). Initially 10% of references will be double-screened. If inter-rater agreement is good then the remaining references will be screened by one reviewer. All primary-level studies included after the first scan of citations will be acquired in full and re-evaluated for eligibility at the time they are being entered into a study database (standardised template created in Microsoft Excel). At least 10% of data extraction will be double-coded. Discrepancies or difficulties with coding will be resolved through discussion between reviewers or the opinion of a third reviewer will be sought. Data Analysis A meta-analysis using a random-effects model will be conducted to combine results from similar studies. An intention to treat (ITT) approach will be taken where possible. Risk of bias will be assessed at the study level using the Cochrane risk of bias tool. This assessment includes: adequacy of randomisation (sufficient description of randomisation method, allocation concealment and any baseline difference between groups); blinding (of participants, intervention administrators and outcome assessors); attrition ('at risk of attrition bias' defined as a dropout of more than 20% and completer analysis used, or a difference of >20% between the groups); selective reporting bias (is the protocol registered, are all outcomes reported); other bias (for instance, conflict of interest in funding). Risk of bias will also be assessed at the outcome level using GRADE. For heterogeneity, outcomes will be downgraded once if I2>50%, twice if I2 >80%. For imprecision, outcomes will be downgraded using rules of thumb. If the 95% CI is imprecise i.e. crosses the line of no effect and the threshold for clinical benefit/harm, 0.8 or 1.25 (dichotomous) or -0.5 or 0.5 SMD (for continuous), the outcome will be downgraded. Outcomes will be downgraded one or two levels depending on how many lines it crosses. If the 95% CI is not imprecise, we will consider whether the criterion for Optimal Information Size is met (for dichotomous outcomes, 300 events; for continuous outcomes, 400 participants), if not we will downgrade one level
Heterogeneity (sensitivity analysis and subgroups)	Where possible, the influence of the following subgroups will be considered: • Chronic depression • Depression with coexisting personality disorder • Psychotic depression

Field (based on PRISMA-P)	Content
	• Older adults
Data management (software)	STAR was used to sift through the references identified by the search, and for data extraction Pairwise meta-analyses and production of forest plots was done using Cochrane Review Manager (RevMan5). 'GRADEpro' was used to assess the quality of evidence for each outcome.
Information sources – databases and dates	Database(s): Embase 1974 to Present, Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R) 1946 to Present; Cochrane Library; WEB OF SCIENCE
Identify if an update	Update of CG90 (2009)
Author contacts	For details please see the guideline in development web site.
Highlight if amendment to previous protocol	For details please see section 4.5 of Developing NICE guidelines: the manual 2014
Search strategy – for one database	For details please see appendix B.
Data collection process – forms/duplicate	A standardised evidence table format will be used, and published as appendix D (clinical evidence tables) or H (economic evidence tables).
Data items – define all variables to be collected	For details please see evidence tables in appendix D (clinical evidence tables) or H (economic evidence tables).
Methods for assessing bias at outcome/study level	Standard study checklists were used to critically appraise individual studies. For details please see section 6.2 of Developing NICE guidelines: the manual 2014. The risk of bias across all available evidence was evaluated for each outcome using an adaptation of the 'Grading of Recommendations Assessment, Development and Evaluation (GRADE) toolbox' developed by the international GRADE working group http://www.gradeworkinggroup.org/.
Criteria for quantitative synthesis	For details please see section 6.4 of Developing NICE guidelines: the manual 2014
Methods for quantitative analysis – combining studies and exploring (in)consistency	For details please see the methods chapter.
Meta-bias assessment – publication bias, selective reporting bias	For details please see section 6.2 of Developing NICE guidelines: the manual 2014.
Confidence in cumulative evidence	For details please see sections 6.4 and 9.1 of Developing NICE guidelines: the manual 2014

Field (based on PRISMA-P)	Content
Rationale/context – what is known	For details please see the introduction to the evidence review.
Describe contributions of authors and guarantor	A multidisciplinary committee developed the evidence review. The committee was convened by the National Guideline Alliance (NGA) and chaired by Dr Navneet Kapur in line with section 3 of Developing NICE guidelines: the manual 2014. Staff from the NGA undertook systematic literature searches, appraised the evidence, conducted meta-analysis and cost effectiveness analysis where appropriate, and drafted the guideline in collaboration with the committee. For details please see the methods chapter.
Sources of funding/support	The NGA is funded by NICE and hosted by the Royal College of Obstetricians and Gynaecologists.
Name of sponsor	The NGA is funded by NICE and hosted by the Royal College of Obstetricians and Gynaecologists.
Roles of sponsor	NICE funds NGA to develop guidelines for those working in the NHS, public health and social care in England
PROSPERO registration number	Not applicable

CDSR: Cochrane Database of Systematic Reviews; CENTRAL: Cochrane Central Register of Controlled Trials; CI: confidence interval; DARE: Database of Abstracts of Reviews of Effects; DSM: Diagnostic and Statistical Manual of Mental Disorders; GRADE: Grading of Recommendations Assessment, Development and Evaluation; ICD: International Statistical Classification of Diseases;ITT: intention to treat; N: number; NGA: National Guideline Alliance; NHS: National health service; NICE: National Institute for Health and Care Excellence; NMA: network meta-analysis; RCT: randomised controlled trial; RoB: risk of bias; SMD: standardised mean difference;

Appendix B – Literature search strategies

Literature search strategies for review question 1.1: For adults with depression, what are the relative benefits and harms associated with different models for the coordination and delivery of services?

Clinical search

Database(s): Embase 1974 to 2019 March 04, Ovid MEDLINE(R) and Epub Ahead of Print, In-Process & Other Non-Indexed Citations and Daily 1946 to March 04, 2019, PsycINFO 1806 to February Week 4 2019

Date of search: 05/03/2019

Search updated: 02/03/2021

#	Searches
1	(depression/ or agitated depression/ or atypical depression/ or depressive psychosis/ or dysphoria/ or dysthymia/ or endogenous depression/ or involutional depression/ or late life depression/ or major depression/ or masked depression/ or melancholia/ or "mixed anxiety and depression"/ or "mixed depression and dementia"/ or premenstrual dysphoric disorder/ or reactive depression/ or recurrent brief depression/ or seasonal affective disorder/ or treatment resistant depression/) use oemezd
2	(Depression/ or exp Depressive Disorder/ or Adjustment Disorders/ or Affective Disorders, Psychotic/ or Factitious Disorders/ or Premenstrual Dysphoric Disorder/) use ppez
3	("depression (emotion)"/ or exp major depression/ or affective disorders/ or atypical depression/ or premenstrual dysphoric disorder/ or seasonal affective disorder/) use psyh
4	(depress* or dysphori* or dysthym* or melanchol* or seasonal affective disorder* or ((affective or mood) adj disorder*)).tw.
5	or/1-4
6	Case Management/
7	(collaboration or teamwork*).tw.
8	Intersectoral Collaboration/
9	collaboration/ use psyh
10	collaborative care team/ use oemezd
11	integrated health care system/ use oemezd
12	Delivery of Health Care, Integrated/ use ppez
13	(interdisciplinary treatment approach/ or integrated services/) use psyh
14	(Community-Institutional Relations/ or Hospital-Patient Relations/ or Hospital-Physician Relations/ or Interdepartmental Relations/ or Interinstitutional Relations/ or exp Interprofessional Relations/) use ppez
15	public relations/ use oemezd
16	(multidisciplinary care team* or MDT*1).tw.
17	patient care planning/ use oemezd
18	(Patient-Centered Care/ or exp Patient Care Planning/) use ppez
19	((collaborat* or coordinat* or co ordinat* or integrat* or shared or stepped or systematic) adj2 (care or effort* or health* or interven* or liais* or manag* or model* or pathway* or service* or work*)).tw.
20	(case manag* or disease manag* or enhanced care or managed care or multi-component or multicomponent).tw.
21	(care manag* or chronic care* or complex intervention* or cooperative behav* or co-operative behav* or joint working or interprofessional or inter-professional or interdisciplinary or inter-disciplinary or multidisciplin* or mulit-disciplin* or multiprofession* or multi-profession* or transdisciplin* or trans-disciplin* or multifacet* or multi-facet* or multiple intervention* or multi-intervention* or organi?ational intervention* or interpersonal relation* or inter-personal relation* or interinstitutional relation* or inter-insitutional relation* or consultation liais* or algorithm*).tw.
22	((drug* or medication* or therap* or treatment*) adj (guideline* or protocol* or manag* or model or adherence or complian* or concordance)).tw.
23	(patient care team or patient care management or patient care planning or managed care program* or (healthcare adj3 delivery) or (continuity adj3 care) or (measur* adj2 care) or professional-patient relations or interprofessional relations or inter-professional relations).tw.
24	or/6-23
25	5 and 24
26	Letter/ use ppez
27	letter.pt. or letter/ use oemezd
28	note.pt.
29	editorial.pt.

#	Searches
30	Editorial/ use ppez
31	News/ use ppez
32	exp Historical Article/ use ppez
33	Anecdotes as Topic/ use ppez
34	Comment/ use ppez
35	Case Report/
36	case study/ use oemezd
37	(letter or comment*).ti.
38	or/26-37
39	randomized controlled trial/
40	random*.ti,ab.
41	39 or 40
42	38 not 41
43	(animals/ not humans/) use ppez
44	(animal/ not human/) use oemezd
45	nonhuman/ use oemezd
46	exp animals/ use psyh
47	"primates (nonhuman)"/ use psyh
48	exp Animals, Laboratory/ use ppez
49	exp Animal Experimentation/ use ppez
50	exp animal experiment/ use oemezd
51	exp experimental animal/ use oemezd
52	exp Models, Animal/ use ppez
53	animal model/ use oemezd
54	animal models/ use psyh
55	animal research/ use psyh
56	exp Rodentia/ use ppez
57	exp rodent/ use oemezd
58	exp rodents/ use psyh
59	(rat or rats or mouse or mice).ti.
60	or/42-59
61	25 not 60
62	limit 61 to english language
63	clinical Trials as topic.sh. or (controlled clinical trial or pragmatic clinical trial or randomized controlled trial).pt. or (placebo or randomi?ed or randomly).ab. or trial.ti.
64	63 use ppez
65	(controlled clinical trial or pragmatic clinical trial or randomized controlled trial).pt. or drug therapy.fs. or (groups or placebo or randomi?ed or randomly or trial).ab.
66	65 use ppez
67	crossover procedure/ or double blind procedure/ or randomized controlled trial/ or single blind procedure/ or (assign* or allocat* or crossover* or cross over* or ((doubl* or singl*) adj blind*) or factorial* or placebo* or random* or volunteer*).ti,ab.
68	67 use oemezd
69	clinical trials/ or (placebo or randomi?ed or randomly).ab. or trial.ti.
70	69 use psyh
71	64 or 66
72	68 or 70 or 71
73	Meta-Analysis/
74	Meta-Analysis as Topic/
75	systematic review/
76	meta-analysis/
77	(meta analy* or metanaly* or metaanaly*).ti,ab.
78	((systematic or evidence) adj2 (review* or overview*)).ti,ab.
79	((systematic* or evidence*) adj2 (review* or overview*)).ti,ab.
80	(reference list* or bibliograph* or hand search* or manual search* or relevant journals).ab.
81	(search strategy or search criteria or systematic search or study selection or data extraction).ab.
82	(search* adj4 literature).ab.
83	(medline or pubmed or cochrane or embase or psychlit or psyclit or psychinfo or psycinfo or cinahl or science citation index or bids or cancerlit).ab.
84	cochrane.jw.
85	((pool* or combined) adj2 (data or trials or studies or results)).ab.
86	(or/73-75,77,79-84) use ppez
87	(or/75-78,80-85) use oemezd
88	(or/73,77,79-84) use psyh
89	or/86-88
90	72 or 89
91	62 and 90

The Cochrane Library: Cochrane Database of Systematic Reviews, Issue 3 of 12, March 2019; Cochrane Central Register of Controlled Trials, Issue 3 of 12, March 2019

Date of search: 05/03/2019

Search updated: 04/03/2021

ID	Search
#1	MeSH descriptor: [Depression] this term only
#2	MeSH descriptor: [Depressive Disorder, Major] explode all trees
#3	MeSH descriptor: [Adjustment Disorders] this term only
#4	MeSH descriptor: [Affective Disorders, Psychotic] this term only
#5	MeSH descriptor: [Factitious Disorders] this term only
#6	MeSH descriptor: [Premenstrual Dysphoric Disorder] this term only
#7	(depress* or dysphori* or dysthym* or melanchol* or seasonal affective disorder* or ((affective or mood) next disorder*))
#8	{or #1-#7}
#9	MeSH descriptor: [Case Management] this term only
#10	(collaboration or teamwork*):ti,ab
#11	MeSH descriptor: [Delivery of Health Care, Integrated] this term only
#12	MeSH descriptor: [Community-Institutional Relations] this term only
#13	MeSH descriptor: [Hospital-Patient Relations] this term only
#14	MeSH descriptor: [Hospital-Physician Relations] this term only
#15	MeSH descriptor: [Interdepartmental Relations] this term only
#16	MeSH descriptor: [Interdepartmental Relations] this term only
#17	MeSH descriptor: [Interprofessional Relations] explode all trees
#18	(multidisciplinary care team* or MDT or MDTs):ti,ab
#19	MeSH descriptor: [Patient-Centered Care] this term only
#20	MeSH descriptor: [Patient Care Planning] explode all trees
#21	((collaborat* or coordinat* or “co ordinat*” or integrat* or shared or stepped or systematic) near/2 (care or effort* or health* or interven* or liais* or manag* or model* or pathway* or service* or work*)):ti,ab
#22	(“case manag*” or “disease manag*” or “enhanced care” or “manag* care” or “multi component” or multicomponent):ti,ab
#23	(“care manag*” or “chronic care*” or “complex intervention*” or “cooperative behav*” or “co operative behav*” or “joint working” or interprofessional or "inter professional" or interdisciplinary or “inter disciplinary” or multidisciplin* or “mulit disciplin*” or multiprofession* or “multi profession*” or transdisciplin* or “trans disciplin*” or multifacet* or “multi facet*” or “multiple intervention*” or “multi intervention*” or “organi?ational intervention*” or “interpersonal relation*” or “inter personal relation*” or “interinstitutional relation*” or ”inter insitutional relation*” or “consultation liais*” or algorithm*):ti,ab
#24	((drug* or medication* or therap* or treatment*) NEXT (guideline* or protocol* or manag* or model* or adherence or complian* or concordance)):ti,ab
#25	(“patient care team*” or “patient care manag*” or “patient care plan*” or “managed care program*” or (healthcare near/3 delivery) or (continuity near/3 care) or (measur* near/2 care) or “professional-patient relations” or “interprofessional relations” or “inter professional relations”):ti,ab
#26	{or #9-#25}
#27	#8 and #26 in Cochrane Reviews, Cochrane Protocols, Trials

Health Economics search

Database(s): Embase 1974 to 2019 Week 08, Ovid MEDLINE(R) and Epub Ahead of Print, In-Process & Other Non-Indexed Citations and Daily 1946 to February 26, 2019, PsycINFO 1806 to February Week 1 2019

Searched: 27/02/2019

Search updated: 02/03/2021

#	Searches
1	(depression/ or agitated depression/ or atypical depression/ or depressive psychosis/ or dysphoria/ or dysthymia/ or endogenous depression/ or involutional depression/ or late life depression/ or major depression/ or masked depression/ or melancholia/ or "mixed anxiety and depression"/ or "mixed depression and dementia"/ or premenstrual dysphoric disorder/ or reactive depression/ or recurrent brief depression/ or seasonal affective disorder/ or treatment resistant depression/) use oemezd

#	Searches
2	((Depression/ or exp Depressive Disorder/ or Adjustment Disorders/ or Affective Disorders, Psychotic/ or Factitious Disorders/ or Premenstrual Dysphoric Disorder/) use ppez
3	("depression (emotion)"/ or exp major depression/ or affective disorders/ or atypical depression/ or premenstrual dysphoric disorder/ or seasonal affective disorder/) use psyh
4	(depress* or dysphori* or dysthym* or melanchol* or seasonal affective disorder* or ((affective or mood) adj disorder*)).tw.
5	or/1-4
6	Letter/ use ppez
7	letter.pt. or letter/ use oemezd
8	note.pt.
9	editorial.pt.
10	Editorial/ use ppez
11	News/ use ppez
12	exp Historical Article/ use ppez
13	Anecdotes as Topic/ use ppez
14	Comment/ use ppez
15	Case Report/
16	case study/ use oemezd
17	(letter or comment*).ti.
18	or/6-17
19	randomized controlled trial/
20	random*.ti,ab.
21	19 or 20
22	18 not 21
23	(animals/ not humans/) use ppez
24	(animal/ not human/) use oemezd
25	nonhuman/ use oemezd
26	exp animals/ use psyh
27	"primates (nonhuman)"/ use psyh
28	exp Animals, Laboratory/ use ppez
29	exp Animal Experimentation/ use ppez
30	exp animal experiment/ use oemezd
31	exp experimental animal/ use oemezd
32	exp Models, Animal/ use ppez
33	animal model/ use oemezd
34	animal models/ use psyh
35	animal research/ use psyh
36	exp Rodentia/ use ppez
37	exp rodent/ use oemezd
38	exp rodents/ use psyh
39	(rat or rats or mouse or mice).ti.
40	or/22-39
41	5 not 40
42	Economics/
43	Value of life/
44	exp "Costs and Cost Analysis"/
45	exp Economics, Hospital/
46	exp Economics, Medical/
47	Economics, Nursing/
48	Economics, Pharmaceutical/
49	exp "Fees and Charges"/
50	exp Budgets/
51	(or/42-50) use ppez
52	health economics/
53	exp economic evaluation/
54	exp health care cost/
55	exp fee/
56	budget/
57	funding/
58	(or/52-57) use oemezd
59	exp economics/
60	exp "costs and cost analysis"/
61	cost containment/
62	money/
63	resource allocation/
64	(or/59-63) use psyh
65	budget*.ti,ab.

#	Searches
66	cost*.ti.
67	(economic* or pharmaco?economic*).ti.
68	(price* or pricing*).ti,ab.
69	(cost* adj2 (effective* or utilit* or benefit* or minimi* or unit* or estimat* or variable*)).ab.
70	(financ* or fee or fees).ti,ab.
71	(value adj2 (money or monetary)).ti,ab.
72	or/65-70
73	51 or 58 or 64 or 72
74	Quality-Adjusted Life Years/ use ppez
75	Sickness Impact Profile/
76	quality adjusted life year/ use oemezd
77	"quality of life index"/ use oemezd
78	(quality adjusted or quality adjusted life year*).tw.
79	(qaly* or qal or qald* or qale* or qtime* or qwb* or daly).tw.
80	(illness state* or health state*).tw.
81	(hui or hui2 or hui3).tw.
82	(multiattibute* or multi attribute*).tw.
83	(utilit* adj3 (score*1 or valu* or health* or cost* or measur* or disease* or mean or gain or gains or index*)).tw.
84	utilities.tw.
85	(eq-5d* or eq5d* or eq-5* or eq5* or euroqual* or euro qual* or euroqual 5d* or euro qual 5d* or euro qol* or euroqol*or euro quol* or euroquol* or euro quol5d* or euroquol5d* or eur qol* or eurqol* or eur qol5d* or eurqol5d* or eur?qul* or eur?qul5d* or euro* quality of life or european qol).tw.
86	(euro* adj3 (5 d* or 5d* or 5 dimension* or 5dimension* or 5 domain* or 5domain*)).tw.
87	(sf36 or sf 36 or sf thirty six or sf thirtysix).tw.
88	(time trade off*1 or time tradeoff*1 or tto or timetradeoff*1).tw.
89	Quality of Life/ and ((quality of life or qol) adj (score*1 or measure*1)).tw.
90	Quality of Life/ and ec.fs.
91	Quality of Life/ and (health adj3 status).tw.
92	(quality of life or qol).tw. and Cost-Benefit Analysis/ use ppez
93	(quality of life or qol).tw. and cost benefit analysis/ use oemezd
94	(quality of life or qol).tw. and "costs and cost analysis"/ use psyh
95	((qol or hrqol or quality of life).tw. or *quality of life/) and ((qol or hrqol* or quality of life) adj2 (increas* or decreas* or improv* or declin* or reduc* or high* or low* or effect or effects or worse or score or scores or change*1 or impact*1 or impacted or deteriorat*)).ab.
96	Cost-Benefit Analysis/ use ppez and cost-effectiveness ratio*.tw. and (cost-effectiveness ratio* and (perspective* or life expectanc*)).tw.
97	cost benefit analysis/ use oemezd and cost-effectiveness ratio*.tw. and (cost-effectiveness ratio* and (perspective* or life expectanc*)).tw.
98	"costs and cost analysis"/ use psyh and cost-effectiveness ratio*.tw. and (cost-effectiveness ratio* and (perspective* or life expectanc*)).tw.
99	*quality of life/ and (quality of life or qol).ti.
100	quality of life/ and ((quality of life or qol) adj3 (improv* or chang*)).tw.
101	quality of life/ and health-related quality of life.tw.
102	Models, Economic/ use ppez
103	economic model/ use oemezd
104	or/74-101
105	73 or 104
106	41 and 105
107	limit 106 to english language
108	limit 107 to yr="2016 -Current"

Database(s): NIHR Centre for Reviews and Dissemination: Health Technology Assessment Database (HTA)

Searched: 26/02/2019

#	Searches
#1	MESH DESCRIPTOR: depressive disorder EXPLODE ALL TREES
#2	((depres* or dysphori* or dysthymi* or melancholi* or seasonal affective disorder* or affective disorder* or mood disorder*))
#3	#1 or #2 IN HTA FROM 2016 TO 2019

Database(s): CINAHL Plus (Cumulative Index to Nursing and Allied Health Literature) 1937-current, EBSCO Host

Searched: 26/02/2019

Search updated: 02/03/2021

#	Query	Limiters/Expanders
S31	S4 AND S30	Limiters - Publication Year: 2016-2019; Exclude MEDLINE records; Language: English Search modes - Boolean/Phrase
S30	S10 OR S29	Search modes - Boolean/Phrase
S29	S11 OR S12 OR S13 OR S14 OR S15 OR S16 OR S17 OR S18 OR S19 OR S20 OR S21 OR S22 OR S23 OR S24 OR S25 OR S26 OR S27 OR S28	Limiters - Exclude MEDLINE records; Language: English Search modes - Boolean/Phrase
S28	(MH "Quality of Life") AND TX (health-related quality of life)	Search modes - Boolean/Phrase
S27	(MH "Quality of Life") AND TI (quality of life or qol)	Search modes - Boolean/Phrase
S26	AB ((qol or hrqol or quality of life) AND ((qol or hrqol* or quality of life) N2 (increas* or decreas* or improv* or declin* or reduc* or high* or low* or effect or effects or worse or score or scores or change*1 or impact*1 or impacted or deteriorat*)))	Search modes - Boolean/Phrase
S25	(MH "Cost Benefit Analysis") AND TX ((quality of life or qol) or (cost-effectiveness ratio* and (perspective* or life expectanc*))	Search modes - Boolean/Phrase
S24	(MH "Quality of Life") TX (health N3 status)	Search modes - Boolean/Phrase
S23	(MH "Quality of Life") AND TX ((quality of life or qol) N (score*1 or measure*1))	Search modes - Boolean/Phrase
S22	TX (time trade off*1 or time tradeoff*1 or tto or timetradeoff*1)	Search modes - Boolean/Phrase
S21	TX (sf36 or sf 36 or sf thirty six or sf thirtysix)	Search modes - Boolean/Phrase
S20	TX (euro* N3 (5 d* or 5d* or 5 dimension* or 5dimension* or 5 domain* or 5domain*))	Search modes - Boolean/Phrase
S19	TX (eq-5d* or eq5d* or eq-5* or eq5* or euroqual* or euro qual* or euroqual 5d* or euro qual 5d* or euro qol* or euroqol*or euro quol* or euroquol* or euro quol5d* or euroquol5d* or eur qol* or eurqol* or eur qol5d* or eurqol5d* or eur?qul* or eur?qul5d* or euro* quality of life or european qol)	Search modes - Boolean/Phrase
S18	TI utilities	Search modes - Boolean/Phrase
S17	TX (utilit* N3 (score*1 or valu* or health* or cost* or measur* or disease* or mean or gain or gains or index*))	Search modes - Boolean/Phrase
S16	TX (multiattibute* or multi attribute*)	Search modes - Boolean/Phrase
S15	TX (hui or hui2 or hui3)	Search modes - Boolean/Phrase
S14	TX (illness state* or health state*)	Search modes - Boolean/Phrase
S13	TX (quality adjusted or quality adjusted life year*or qaly* or qal or qald* or qale* or qtime* or qwb* or daly)	Search modes - Boolean/Phrase
S12	(MH "Sickness Impact Profile")	Search modes - Boolean/Phrase
S11	(MH "Quality-Adjusted Life Years")	Search modes - Boolean/Phrase
S10	S5 OR S6 OR S7 OR S8 OR S9	Limiters - Exclude MEDLINE records; Language: English Search modes - Boolean/Phrase
S9	TX (value N2 (money or monetary))	Search modes - Boolean/Phrase
S8	TX (cost* N2 (effective* or utilit* or benefit* or minimi* or unit* or estimat* or variable*))	Search modes - Boolean/Phrase
S7	TI cost* or economic* or pharmaco?economic*	Search modes - Boolean/Phrase
S6	TX budget* or fee or fees or finance* or price* or pricing	Search modes - Boolean/Phrase
S5	(MH "Fees and Charges+") OR (MH "Costs and Cost Analysis+") OR (MH "Economics") OR (MH "Economic Value of Life") OR (MH "Economics, Pharmaceutical") OR (MH "Economic Aspects of Illness") OR (MH "Resource Allocation+")	Search modes - Boolean/Phrase
S4	S1 OR S2 OR S3	Limiters - Exclude MEDLINE records; Language: English Search modes - Boolean/Phrase
S3	TX (depress* or dysphori* or dysthym* or melanchol* or seasonal affective disorder)	Search modes - Boolean/Phrase
S2	(MH "Adjustment Disorders+") OR (MH "Factitious Disorders") OR (MH "Affective Disorders, Psychotic")	Search modes - Boolean/Phrase
S1	(MH "Depression+") OR (MH "Premenstrual Dysphoric Disorder") OR (MH "Seasonal Affective Disorder")	Search modes - Boolean/Phrase

Literature search strategies for review question 1.2 For adults with depression, what are the relative benefits and harms associated with different settings for the delivery of care?

Clinical search

Database(s): Embase 1974 to 2019 March 13, Ovid MEDLINE(R) and Epub Ahead of Print, In-Process & Other Non-Indexed Citations and Daily 1946 to March 13, 2019, PsycINFO 1806 to March Week 1 2019

Searched: 14/03/2019

Search updated: 03/03/2021

#	Searches
1	(depression/ or agitated depression/ or atypical depression/ or depressive psychosis/ or dysthymia/ or endogenous depression/ or involutional depression/ or late life depression/ or major depression/ or masked depression/ or melancholia/ or "mixed anxiety and depression"/ or reactive depression/ or recurrent brief depression/ or treatment resistant depression/) use oemezd
2	(Depression/ or Depressive Disorder/ or Depressive Disorder, Major/ or Depressive Disorder, Treatment-Resistant/ or Disorders, Psychotic/ or Dysthymic Disorder/) use ppez
3	("depression (emotion)"/ or exp major depression/ or affective disorders/ or atypical depression/) use psyh
4	(depress* or dysthym* or melancholic* or ((affective or mood) adj disorder*)).tw.
5	((severe or serious or persistent or major or critical or clinical or acute) adj2 (anxiety* or (mental adj2 (disorder* or health or illness* or ill-health)) or (obsessive adj2 disorder*) or OCD or panic attack* or panic disorder* or phobi* or personality disorder* or psychiatric disorder* or psychiatric illness* or psychiatric ill-health*)).tw.
6	or/1-5
7	exp Primary Health Care/
8	Physicians, Family/
9	Family Practice/
10	General Practice/
11	General Practitioners/
12	Primary Care Nursing/
13	Family Nursing/
14	Mental Health Services/
15	Community Mental Health Services/
16	Community Health Nursing/
17	exp Community Health Centers/
18	Home Care Services/ or Home Care Services, Hospital-Based/ or Home Care Agencies/ or Home Health Nursing/ or exp Home Nursing/
19	Crisis Intervention/
20	Emergency Services, Psychiatric/
21	Psychiatric Department, Hospital/ or Psychiatric Hospitals/
22	Residential Facilities/
23	Hospitalization/
24	Ambulatory Care/ or Ambulatory Care Facilities/ or Outpatients Clinics, Hosptial/
25	Day Care, Medical/
26	Adult Day Care Centers/
27	Assisted Living Facilities/
28	Psychiatric Rehabilitation/ or Mental Health Recovery/
29	Tertiary Care Centers/
30	(or/7-29) use ppez
31	exp primary health care/
32	general practitioner/
33	community care/ or community health nursing/ or community psychiatric nursing/

#	Searches
34	home care/ or home mental health care/ or visiting nurse service/
35	crisis intervention/
36	psychiatric emergency service/
37	mental health center/ or mental health service/ or mental hospital/ or psychiatric department/ or psychiatric intensive care unit/
38	residential care/ or residential home/
39	ambulatory care/ or ambulatory care nursing/ or outpatient care/ or outpatient department/
40	adult day care/
41	rehabilitation center/ or mental health recovery/
42	tertiary care center/
43	(or/31-42) use oemezd
44	primary health care/
45	family medicine/ or family physicians/ or general practitioners/
46	community mental health/ or community mental health centers/ or community mental health services/ or community psychiatry/ or community psychology/
47	home care/ or home visiting programs/ or homebound/
48	crisis intervention services/ or suicide prevention centers/
49	psychiatric units/ or psychiatric hospitals/ or exp psychiatric hospitalization/
50	exp hospitalization/
51	exp residential care/ or residential home/ or exp residential care institutions/
52	psychiatric clinics/ or outpatient treatment/ or partial hospitalization/
53	adult day care/ or day care centers/
54	deinstitutionalization/ or rehabilitation centers/
55	(or/44-54) use psyh
56	(primary adj2 (care or health*)).tw.
57	((general or family) adj (practice* or practitioner*)).tw.
58	(GP or GPs).tw.
59	((family or community or practice*) adj (centre* or center*1 or clinic* or doctor* or health* or medic* or nurs* or physician* or service* or setting* or team*)).tw.
60	(communit* adj2 (care or centre* or center*1 or facilit* or hospital* or service* or setting* or team* or unit*)).tw.
61	(home adj2 (based or care or service* or setting* or team*)).tw.
62	((crisis or emergency) adj2 (centre* or center*1 or department* or facilit* or service* or setting* or team* or unit*)).tw.
63	((acute or inpatient* or mental health or psychiatric) adj2 (care or centre* or center*1 or department* or facilit* or hospital* or institution* or service* or setting* or team* or unit*)).tw.
64	((assisted living or housing or residential) adj2 (care or centre* or center*1 or facilit* or home* or hospital* or institution* or service* or setting* or support* or team* or unit*)).tw.
65	(((day or drop-in) adj2 (centre* or center*1 or care* or hospital* or unit*)) or community mental health cent* or CMHC).tw.
66	((rehabilitat* or recovery) adj2 (centre* or center*1 or facilit* or hospital* or service* or setting* or team* or unit*)).tw.
67	((specialist or tertiary) adj2 (care or centre* or center*1 or facilit* or hospital or service* or setting* or team* or unit*)).tw.
68	or/56-67
69	30 or 43 or 55 or 68
70	5 and 69
71	limit 70 to english language
72	Letter/ use ppez
73	letter.pt. or letter/ use oemezd
74	note.pt.
75	editorial.pt.
76	Editorial/ use ppez
77	News/ use ppez
78	exp Historical Article/ use ppez
79	Anecdotes as Topic/ use ppez

#	Searches
80	Comment/ use ppez
81	Case Report/
82	case study/ use oemezd
83	(letter or comment*).ti.
84	or/72-83
85	randomized controlled trial/
86	random*.ti,ab.
87	85 or 86
88	84 not 87
89	(animals/ not humans/) use ppez
90	(animal/ not human/) use oemezd
91	nonhuman/ use oemezd
92	exp animals/ use psyh
93	"primates (nonhuman)"/ use psyh
94	exp Animals, Laboratory/ use ppez
95	exp Animal Experimentation/ use ppez
96	exp animal experiment/ use oemezd
97	exp experimental animal/ use oemezd
98	exp Models, Animal/ use ppez
99	animal model/ use oemezd
100	animal models/ use psyh
101	animal research/ use psyh
102	exp Rodentia/ use ppez
103	exp rodent/ use oemezd
104	exp rodents/ use psyh
105	(rat or rats or mouse or mice).ti.
106	or/88-105
107	71 not 106
108	clinical Trials as topic.sh. or (controlled clinical trial or pragmatic clinical trial or randomized controlled trial).pt. or (placebo or randomi?ed or randomly).ab. or trial.ti.
109	108 use ppez
110	(controlled clinical trial or pragmatic clinical trial or randomized controlled trial).pt. or drug therapy.fs. or (groups or placebo or randomi?ed or randomly or trial).ab.
111	110 use ppez
112	crossover procedure/ or double blind procedure/ or randomized controlled trial/ or single blind procedure/ or (assign* or allocat* or crossover* or cross over* or ((doubl* or singl*) adj blind*) or factorial* or placebo* or random* or volunteer*).ti,ab.
113	112 use oemezd
114	clinical trials/ or (placebo or randomi?ed or randomly).ab. or trial.ti.
115	114 use psyh
116	109 or 111
117	113 or 115 or 116
118	Meta-Analysis/
119	exp Meta-Analysis as Topic/
120	systematic review/
121	meta-analysis/
122	(meta analy* or metanaly* or metaanaly*).ti,ab.
123	((systematic or evidence) adj2 (review* or overview*)).ti,ab.
124	((systematic* or evidence*) adj2 (review* or overview*)).ti,ab.
125	(reference list* or bibliograph* or hand search* or manual search* or relevant journals).ab.
126	(search strategy or search criteria or systematic search or study selection or data extraction).ab.
127	(search* adj4 literature).ab.

#	Searches
128	(medline or pubmed or cochrane or embase or psychlit or psyclit or psychinfo or psycinfo or cinahl or science citation index or bids or cancerlit).ab.
129	cochrane.jw.
130	((pool* or combined) adj2 (data or trials or studies or results)).ab.
131	(or/118-120,122,124-129) use ppez
132	(or/120-123,125-130) use oemezd
133	(or/118,122,124-129) use psyh
134	or/131-133
135	117 or 134
136	107 and 135

The Cochrane Library: Cochrane Database of Systematic Reviews, Issue 3 of 12, March 2019; Cochrane Central Register of Controlled Trials, Issue 3 of 12, March 2019

Searched: 14/03/2019

Search updated: 04/03/2021

ID	Search
#1	MeSH descriptor: [Depression] this term only
#2	MeSH descriptor: [Depressive Disorder] this term only
#3	MeSH descriptor: [Depressive Disorder, Major] this term only
#4	MeSH descriptor: [Depressive Disorder, Treatment-Resistant] this term only
#5	MeSH descriptor: [Affective Disorders, Psychotic] this term only
#6	MeSH descriptor: [Dysthymic Disorder] this term only
#7	(depress* or dysphori* or dysthym* or melanchol* or ((affective or mood) next disorder*)):ti,ab
#8	((sever* or serious* or resist* or persist* or major or endur* or chronic or acute or complex) next/2 anxiet* or (mental next/2 (disorder* or health or illness* or ill-health)) or (obsessive next/2 disorder*) or OCD or "panic attack*" or "panic disorder*" or phobi* or "personality disorder*" or "psychiatric disorder*" or "psychiatric illness*" or "psychiatric ill-health*"):ti,ab
#9	{or #1-#8}
#10	MeSH descriptor: [Primary Health Care] explode all trees
#11	MeSH descriptor: [Physicians, Family] this term only
#12	MeSH descriptor: [Family Practice] this term only
#13	MeSH descriptor: [General Practice] this term only
#14	MeSH descriptor: [General Practitioners] this term only
#15	MeSH descriptor: [Primary Care Nursing] this term only
#16	MeSH descriptor: [Family Nursing] this term only
#17	MeSH descriptor: [Mental Health Services] this term only
#18	MeSH descriptor: [Community Mental Health Services] this term only
#19	MeSH descriptor: [Community Health Nursing] this term only
#20	MeSH descriptor: [Community Health Centers] explode all trees
#21	MeSH descriptor: [Home Care Services] this term only
#22	MeSH descriptor: [Home Care Services, Hospital-Based] this term only
#23	MeSH descriptor: [Home Care Agencies] this term only
#24	MeSH descriptor: [Home Health Nursing] this term only
#25	MeSH descriptor: [Home Nursing] explode all trees
#26	MeSH descriptor: [Crisis Intervention] this term only
#27	MeSH descriptor: [Emergency Services, Psychiatric] this term only
#28	MeSH descriptor: [Psychiatric Department, Hospital] this term only
#29	MeSH descriptor: [Hospitals, Psychiatric] this term only
#30	MeSH descriptor: [Residential Facilities] this term only
#31	MeSH descriptor: [Hospitalization] this term only
#32	MeSH descriptor: [Ambulatory Care] this term only
#33	MeSH descriptor: [Ambulatory Care Facilities] this term only

ID	Search
#34	MeSH descriptor: [Outpatient Clinics, Hospital] this term only
#35	MeSH descriptor: [Day Care, Medical] this term only
#36	MeSH descriptor: [Adult Day Care Centers] this term only
#37	MeSH descriptor: [Assisted Living Facilities] this term only
#38	MeSH descriptor: [Psychiatric Rehabilitation] this term only
#39	MeSH descriptor: [Mental Health Recovery] this term only
#40	MeSH descriptor: [Tertiary Care Centers] this term only
#41	(primary next (care or health*)):ti,ab
#42	((general or family) next (practice* or practitioner*)):ti,ab
#43	(GP or GPs):ti,ab
#44	((family or community or practice*) next (centre* or center or centers or clinic* or doctor* or health* or medic* or nurs* or physician* or service* or setting* or team*)):ti,ab
#45	(communit* next/2 (care or centre* or center or centers or facilit* or hospital* or service* or setting* or team* or unit*)):ti,ab
#46	(home next (based or care or service* or setting* or team*)):ti,ab
#47	((crisis or emergency) near (centre* or center or centers or department* or facilit* or service* or setting* or team* or unit*)):ti,ab
#48	((acute or inpatient* or "mental health" or psychiatric) next (care or centre* or center or centers or department* or facilit* or hospital* or institution* or service* or setting* or team* or unit*)):ti,ab
#49	("assisted living" or ((residential or housing) next (care or centre* or center or centers or facilit* or home* or hospital* or institution* or service* or support or setting* or team* or unit*))):ti,ab
#50	(((day or drop-in) near (centre* or center or centers or care* or hospital* or unit*)) or "community mental health cent*" or CMHC):ti,ab
#51	((rehabilitat* or recovery) next (centre* or center or centers or facilit* or hospital* or service* or setting* or team* or unit*)):ti,ab
#52	((specialist or tertiary) near (care or centre* or center or centers or facilit* or hospital or service* or setting* or team* or unit*)):ti,ab
#53	{or #10-#52}
#54	#9 and #53 in Cochrane Reviews, Cochrane Protocols, Trials

Health Economics search

Database(s): Embase 1974 to 2019 Week 08, Ovid MEDLINE(R) and Epub Ahead of Print, In-Process & Other Non-Indexed Citations and Daily 1946 to February 26, 2019, PsycINFO 1806 to February Week 1 2019

Date of initial search: 27/02/12019

Search updated: 02/03/2021

#	Searches
1	(depression/ or agitated depression/ or atypical depression/ or depressive psychosis/ or dysphoria/ or dysthymia/ or endogenous depression/ or involutional depression/ or late life depression/ or major depression/ or masked depression/ or melancholia/ or "mixed anxiety and depression"/ or "mixed depression and dementia"/ or premenstrual dysphoric disorder/ or reactive depression/ or recurrent brief depression/ or seasonal affective disorder/ or treatment resistant depression/) use oemezd
2	((Depression/ or exp Depressive Disorder/ or Adjustment Disorders/ or Affective Disorders, Psychotic/ or Factitious Disorders/ or Premenstrual Dysphoric Disorder/) use ppez
3	("depression (emotion)"/ or exp major depression/ or affective disorders/ or atypical depression/ or premenstrual dysphoric disorder/ or seasonal affective disorder/) use psyh
4	(depress* or dysphori* or dysthym* or melanchol* or seasonal affective disorder* or ((affective or mood) adj disorder*)).tw.
5	or/1-4
6	Letter/ use ppez
7	letter.pt. or letter/ use oemezd

#	Searches
8	note.pt.
9	editorial.pt.
10	Editorial/ use ppez
11	News/ use ppez
12	exp Historical Article/ use ppez
13	Anecdotes as Topic/ use ppez
14	Comment/ use ppez
15	Case Report/
16	case study/ use oemezd
17	(letter or comment*).ti.
18	or/6-17
19	randomized controlled trial/
20	random*.ti,ab.
21	19 or 20
22	18 not 21
23	(animals/ not humans/) use ppez
24	(animal/ not human/) use oemezd
25	nonhuman/ use oemezd
26	exp animals/ use psyh
27	"primates (nonhuman)"/ use psyh
28	exp Animals, Laboratory/ use ppez
29	exp Animal Experimentation/ use ppez
30	exp animal experiment/ use oemezd
31	exp experimental animal/ use oemezd
32	exp Models, Animal/ use ppez
33	animal model/ use oemezd
34	animal models/ use psyh
35	animal research/ use psyh
36	exp Rodentia/ use ppez
37	exp rodent/ use oemezd
38	exp rodents/ use psyh
39	(rat or rats or mouse or mice).ti.
40	or/22-39
41	5 not 40

#	Searches
42	Economics/
43	Value of life/
44	exp "Costs and Cost Analysis"/
45	exp Economics, Hospital/
46	exp Economics, Medical/
47	Economics, Nursing/
48	Economics, Pharmaceutical/
49	exp "Fees and Charges"/
50	exp Budgets/
51	(or/42-50) use ppez
52	health economics/
53	exp economic evaluation/
54	exp health care cost/
55	exp fee/
56	budget/
57	funding/
58	(or/52-57) use oemezd
59	exp economics/
60	exp "costs and cost analysis"/
61	cost containment/
62	money/
63	resource allocation/
64	(or/59-63) use psyh
65	budget*.ti,ab.
66	cost*.ti.
67	(economic* or pharmaco?economic*).ti.
68	(price* or pricing*).ti,ab.
69	(cost* adj2 (effective* or utilit* or benefit* or minimi* or unit* or estimat* or variable*)).ab.
70	(financ* or fee or fees).ti,ab.
71	(value adj2 (money or monetary)).ti,ab.
72	or/65-70
73	51 or 58 or 64 or 72
74	Quality-Adjusted Life Years/ use ppez
75	Sickness Impact Profile/

#	Searches
76	quality adjusted life year/ use oemezd
77	"quality of life index"/ use oemezd
78	(quality adjusted or quality adjusted life year*).tw.
79	(qaly* or qal or qald* or qale* or qtime* or qwb* or daly).tw.
80	(illness state* or health state*).tw.
81	(hui or hui2 or hui3).tw.
82	(multiattibute* or multi attribute*).tw.
83	(utilit* adj3 (score*1 or valu* or health* or cost* or measur* or disease* or mean or gain or gains or index*)).tw.
84	utilities.tw.
85	(eq-5d* or eq5d* or eq-5* or eq5* or euroqual* or euro qual* or euroqual 5d* or euro qual 5d* or euro qol* or euroqol*or euro quol* or euroquol* or euro quol5d* or euroquol5d* or eur qol* or eurqol* or eur qol5d* or eurqol5d* or eur?qul* or eur?qul5d* or euro* quality of life or european qol).tw.
86	(euro* adj3 (5 d* or 5d* or 5 dimension* or 5dimension* or 5 domain* or 5domain*)).tw.
87	(sf36 or sf 36 or sf thirty six or sf thirtysix).tw.
88	(time trade off*1 or time tradeoff*1 or tto or timetradeoff*1).tw.
89	Quality of Life/ and ((quality of life or qol) adj (score*1 or measure*1)).tw.
90	Quality of Life/ and ec.fs.
91	Quality of Life/ and (health adj3 status).tw.
92	(quality of life or qol).tw. and Cost-Benefit Analysis/ use ppez
93	(quality of life or qol).tw. and cost benefit analysis/ use oemezd
94	(quality of life or qol).tw. and "costs and cost analysis"/ use psyh
95	((qol or hrqol or quality of life).tw. or *quality of life/) and ((qol or hrqol* or quality of life) adj2 (increas* or decreas* or improv* or declin* or reduc* or high* or low* or effect or effects or worse or score or scores or change*1 or impact*1 or impacted or deteriorat*)).ab.
96	Cost-Benefit Analysis/ use ppez and cost-effectiveness ratio*.tw. and (cost-effectiveness ratio* and (perspective* or life expectanc*)).tw.
97	cost benefit analysis/ use oemezd and cost-effectiveness ratio*.tw. and (cost-effectiveness ratio* and (perspective* or life expectanc*)).tw.
98	"costs and cost analysis"/ use psyh and cost-effectiveness ratio*.tw. and (cost-effectiveness ratio* and (perspective* or life expectanc*)).tw.
99	*quality of life/ and (quality of life or qol).ti.
100	quality of life/ and ((quality of life or qol) adj3 (improv* or chang*)).tw.
101	quality of life/ and health-related quality of life.tw.
102	Models, Economic/ use ppez
103	economic model/ use oemezd
104	or/74-101
105	73 or 104
106	41 and 105

#	Searches
107	limit 106 to english language
108	limit 107 to yr="2016 -Current"

Database(s): NIHR Centre for Reviews and Dissemination: Health Technology Assessment Database (HTA)

Searched: 26/02/2019

#	Searches
#1	MESH DESCRIPTOR: depressive disorder EXPLODE ALL TREES
#2	((depres* or dysphori* or dysthymi* or melancholi* or seasonal affective disorder* or affective disorder* or mood disorder*))
#3	#1 or #2 IN HTA FROM 2016 TO 2019

Database(s): CINAHL Plus (Cumulative Index to Nursing and Allied Health Literature) 1937-current, EBSCO Host

Date of initial search: 26/02/2019

Search updated: 02/03/2021

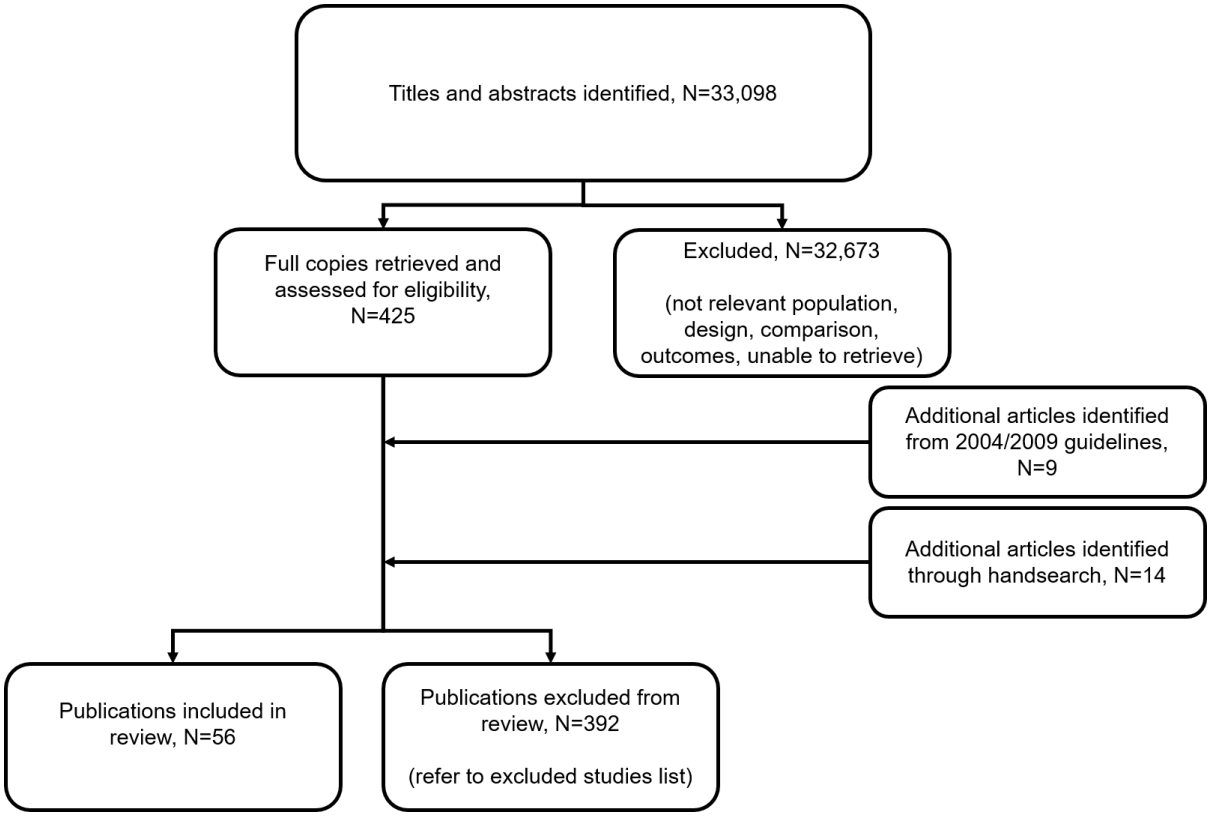
#	Query	Limiters/Expanders
S31	S4 AND S30	Limiters - Publication Year: 2016-2019; Exclude MEDLINE records; Language: English Search modes - Boolean/Phrase
S30	S10 OR S29	Search modes - Boolean/Phrase
S29	S11 OR S12 OR S13 OR S14 OR S15 OR S16 OR S17 OR S18 OR S19 OR S20 OR S21 OR S22 OR S23 OR S24 OR S25 OR S26 OR S27 OR S28	Limiters - Exclude MEDLINE records; Language: English Search modes - Boolean/Phrase
S28	(MH "Quality of Life") AND TX (health-related quality of life)	Search modes - Boolean/Phrase
S27	(MH "Quality of Life") AND TI (quality of life or qol)	Search modes - Boolean/Phrase
S26	AB ((qol or hrqol or quality of life) AND ((qol or hrqol* or quality of life) N2 (increas* or decreas* or improv* or declin* or reduc* or high* or low* or effect or effects or worse or score or scores or change*1 or impact*1 or impacted or deteriorat*)))	Search modes - Boolean/Phrase
S25	(MH "Cost Benefit Analysis") AND TX ((quality of life or qol) or (cost-effectiveness ratio* and (perspective* or life expectanc*))	Search modes - Boolean/Phrase
S24	(MH "Quality of Life") TX (health N3 status)	Search modes - Boolean/Phrase
S23	(MH "Quality of Life") AND TX ((quality of life or qol) N (score*1 or measure*1))	Search modes - Boolean/Phrase
S22	TX (time trade off*1 or time tradeoff*1 or tto or timetradeoff*1)	Search modes - Boolean/Phrase
S21	TX (sf36 or sf 36 or sf thirty six or sf thirtysix)	Search modes - Boolean/Phrase
S20	TX (euro* N3 (5 d* or 5d* or 5 dimension* or 5dimension* or 5 domain* or 5domain*))	Search modes - Boolean/Phrase
S19	TX (eq-5d* or eq5d* or eq-5* or eq5* or euroqual* or euro qual* or euroqual 5d* or euro qual 5d* or euro qol* or euroqol*or euro quol* or euroquol* or euro quol5d* or euroquol5d* or eur qol* or eurqol* or eur qol5d* or eurqol5d* or eur?qul* or eur?qul5d* or euro* quality of life or european qol)	Search modes - Boolean/Phrase
S18	TI utilities	Search modes - Boolean/Phrase

#	Query	Limiters/Expanders
S17	TX (utilit* N3 (score*1 or valu* or health* or cost* or measur* or disease* or mean or gain or gains or index*))	Search modes - Boolean/Phrase
S16	TX (multiattibute* or multi attribute*)	Search modes - Boolean/Phrase
S15	TX (hui or hui2 or hui3)	Search modes - Boolean/Phrase
S14	TX (illness state* or health state*)	Search modes - Boolean/Phrase
S13	TX (quality adjusted or quality adjusted life year*or qaly* or qal or qald* or qale* or qtime* or qwb* or daly)	Search modes - Boolean/Phrase
S12	(MH "Sickness Impact Profile")	Search modes - Boolean/Phrase
S11	(MH "Quality-Adjusted Life Years")	Search modes - Boolean/Phrase
S10	S5 OR S6 OR S7 OR S8 OR S9	Limiters - Exclude MEDLINE records; Language: English Search modes - Boolean/Phrase
S9	TX (value N2 (money or monetary))	Search modes - Boolean/Phrase
S8	TX (cost* N2 (effective* or utilit* or benefit* or minimi* or unit* or estimat* or variable*))	Search modes - Boolean/Phrase
S7	TI cost* or economic* or pharmaco?economic*	Search modes - Boolean/Phrase
S6	TX budget* or fee or fees or finance* or price* or pricing	Search modes - Boolean/Phrase
S5	(MH "Fees and Charges+") OR (MH "Costs and Cost Analysis+") OR (MH "Economics") OR (MH "Economic Value of Life") OR (MH "Economics, Pharmaceutical") OR (MH "Economic Aspects of Illness") OR (MH "Resource Allocation+")	Search modes - Boolean/Phrase
S4	S1 OR S2 OR S3	Limiters - Exclude MEDLINE records; Language: English Search modes - Boolean/Phrase
S3	TX (depress* or dysphori* or dysthym* or melanchol* or seasonal affective disorder)	Search modes - Boolean/Phrase
S2	(MH "Adjustment Disorders+") OR (MH "Factitious Disorders") OR (MH "Affective Disorders, Psychotic")	Search modes - Boolean/Phrase
S1	(MH "Depression+") OR (MH "Premenstrual Dysphoric Disorder") OR (MH "Seasonal Affective Disorder")	Search modes - Boolean/Phrase

Appendix C – Clinical evidence study selection

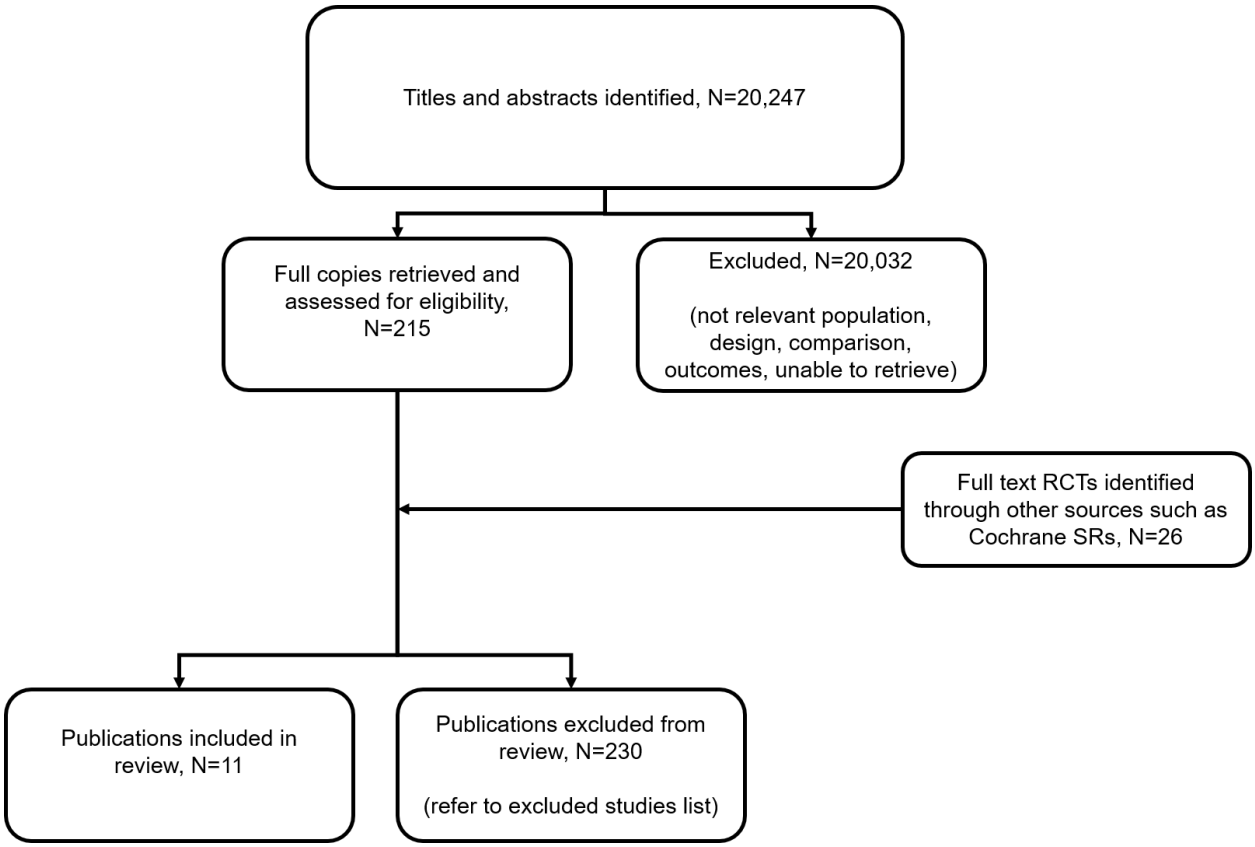
Clinical study selection review question 1.1 For adults with depression, what are the relative benefits and harms associated with different models for the coordination and delivery of services?

Figure 2: Study selection flow chart



Clinical study selection review question 1.2 For adults with depression, what are the relative benefits and harms associated with different settings for the delivery of care?

Figure 3: Study selection flow chart (does not include studies analysed as a sub-set of the NMA data for comparisons 1 and 3)



Appendix D – Clinical evidence tables

Clinical evidence tables for review question 1.1 For adults with depression, what are the relative benefits and harms associated with different models for the coordination and delivery of services?

Please refer to the clinical evidence tables in supplement A1 – Clinical evidence tables for review 1.1

Clinical evidence tables for review question 1.2 For adults with depression, what are the relative benefits and harms associated with different settings for the delivery of care?

Please refer to the clinical evidence tables in supplement A2 – Clinical evidence tables for review 1.2

Appendix E – Forest plots

Forest plots for review question 1.1 For adults with depression, what are the relative benefits and harms associated with different models for the coordination and delivery of services?

Comparison 1: Collaborative care versus standard care/enhanced standard care

Critical outcomes

Figure 4: Depression symptomatology at 6 months

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
1.1.1 Simple collaborative care								
Aragones 2012	7.97	7.22	177	10.46	7.22	134	11.6%	-0.34 [-0.57, -0.12]
Buszewicz 2016	25.8	12.7	201	28.8	13.8	167	11.7%	-0.23 [-0.43, -0.02]
Chen 2015	8	3	164	14.1	4.9	162	11.4%	-1.50 [-1.75, -1.25]
Curth 2020	13.3	5.47	272	19.2	12.88	53	11.0%	-0.82 [-1.12, -0.52]
Harter 2018	11.21	5.92	569	11.9	5.37	168	11.9%	-0.12 [-0.29, 0.05]
Huang 2018	8.8	7.1	139	10.4	7	141	11.5%	-0.23 [-0.46, 0.01]
Landis 2007	14.4	6.4	23	12.4	7.8	22	8.5%	0.28 [-0.31, 0.86]
Ng 2020	6.6	6.17	99	7.8	6.07	99	11.2%	-0.20 [-0.47, 0.08]
Oladeji 2015	4.1	4.4	137	5.5	5.2	64	11.1%	-0.30 [-0.60, -0.00]
Subtotal (95% CI)			**1781**			**1010**	**100.0%**	**-0.40 [-0.71, -0.09]**

Heterogeneity: Tau² = 0.20; Chi² = 107.05, df = 8 (P < 0.00001); I² = 93%
Test for overall effect: Z = 2.54 (P = 0.01)

Test for subgroup differences: Not applicable

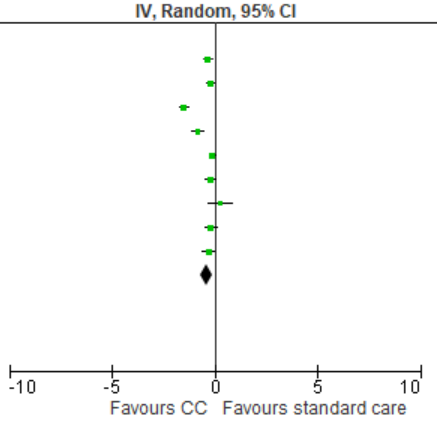


Figure 5: Depression symptomatology at 12 months

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
1.2.1 Simple collaborative care								
Aragones 2012	7.15	7.11	172	8.78	6.99	130	7.6%	-0.23 [-0.46, -0.00]
Bosanquet 2017	10.51	3.280735	249	10.74	3.056782	236	8.0%	-0.07 [-0.25, 0.11]
Bruce 2004	9.77	7.28	320	10.35	6.78	278	8.1%	-0.08 [-0.24, 0.08]
Buszewicz 2016	25.2	12.8	201	27.9	13.6	166	7.8%	-0.20 [-0.41, 0.00]
Chen 2015	6.1	2.6	164	12.6	5.2	162	7.5%	-1.58 [-1.83, -1.33]
Gensichen 2009	10.72	5.43	267	12.13	5.6	288	8.0%	-0.26 [-0.42, -0.09]
Gilbody 2017/Lewis 2017	6.01	2.767891	344	7.26	2.568878	361	8.1%	-0.47 [-0.62, -0.32]
Harter 2018	10.33	6.03	569	12.12	5.53	168	8.0%	-0.30 [-0.47, -0.13]
Ng 2020	7.2	7.06	91	6.9	7.02	90	7.1%	0.04 [-0.25, 0.33]
Richards 2013/2016	10	7.1	235	11.7	6.8	263	8.0%	-0.24 [-0.42, -0.07]
Swindle 2003	17.9	10.7	113	19.9	10.9	106	7.3%	-0.18 [-0.45, 0.08]
Subtotal (95% CI)			**2725**			**2248**	**85.5%**	**-0.32 [-0.53, -0.11]**

Heterogeneity: Tau² = 0.11; Chi² = 127.56, df = 10 (P < 0.00001); I² = 92%
Test for overall effect: Z = 3.04 (P = 0.002)

1.2.2 Complex collaborative care								
Holzel 2018	8.13	2.15044	139	9.38	1.558056	109	7.4%	-0.65 [-0.91, -0.39]
Morriss 2016	14.8	7.9	93	17.2	7.3	94	7.1%	-0.31 [-0.60, -0.03]
Subtotal (95% CI)			**232**			**203**	**14.5%**	**-0.49 [-0.82, -0.16]**

Heterogeneity: Tau² = 0.04; Chi² = 2.92, df = 1 (P = 0.09); I² = 66%
Test for overall effect: Z = 2.91 (P = 0.004)

Total (95% CI)			**2957**			**2451**	**100.0%**	**-0.35 [-0.53, -0.16]**

Heterogeneity: Tau² = 0.11; Chi² = 134.39, df = 12 (P < 0.00001); I² = 91%
Test for overall effect: Z = 3.62 (P = 0.0003)
Test for subgroup differences: Chi² = 0.69, df = 1 (P = 0.41), I² = 0%

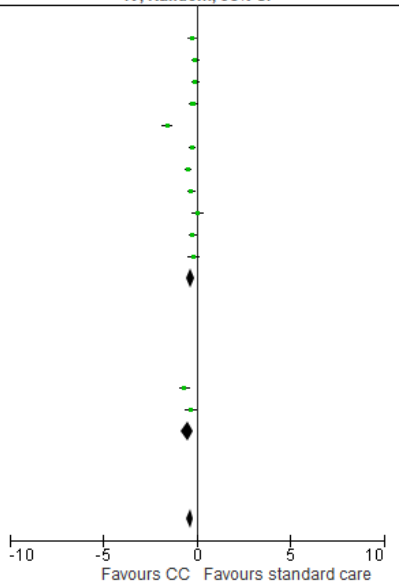


Figure 6: Response at 6 months

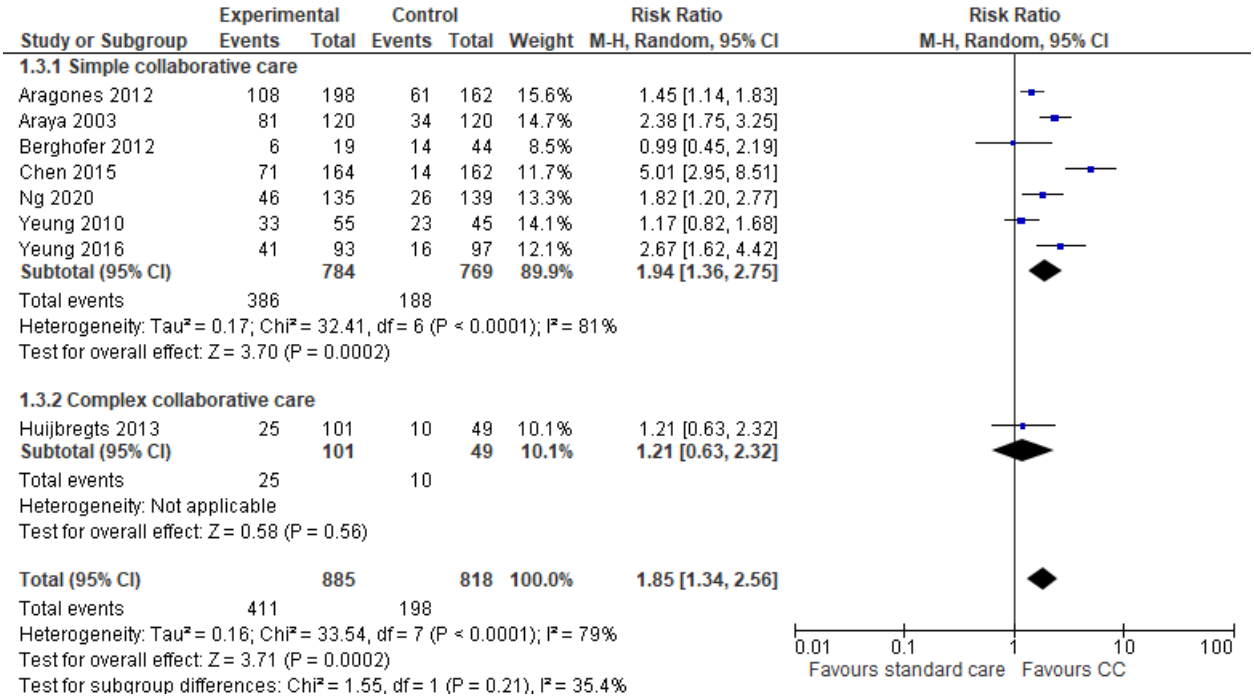


Figure 7: Response at 12 months

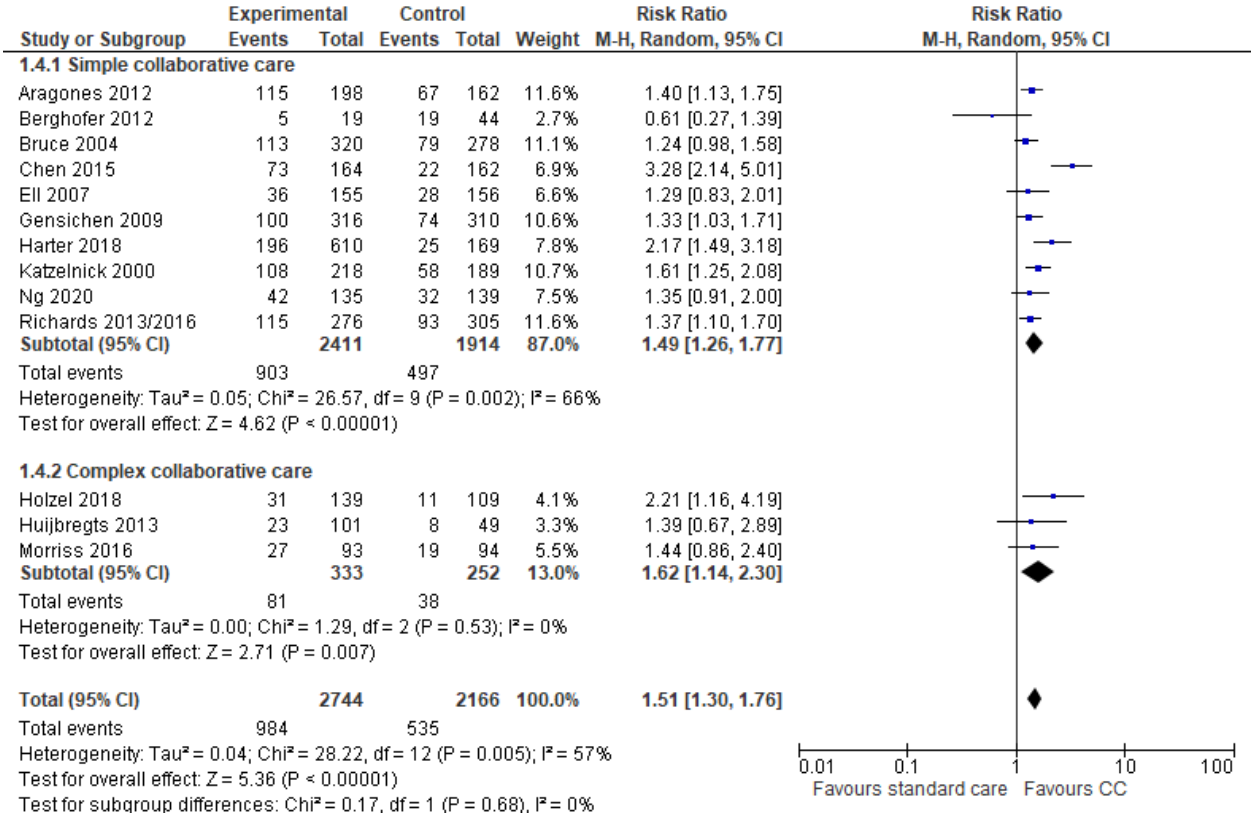


Figure 8: Remission at 6 months

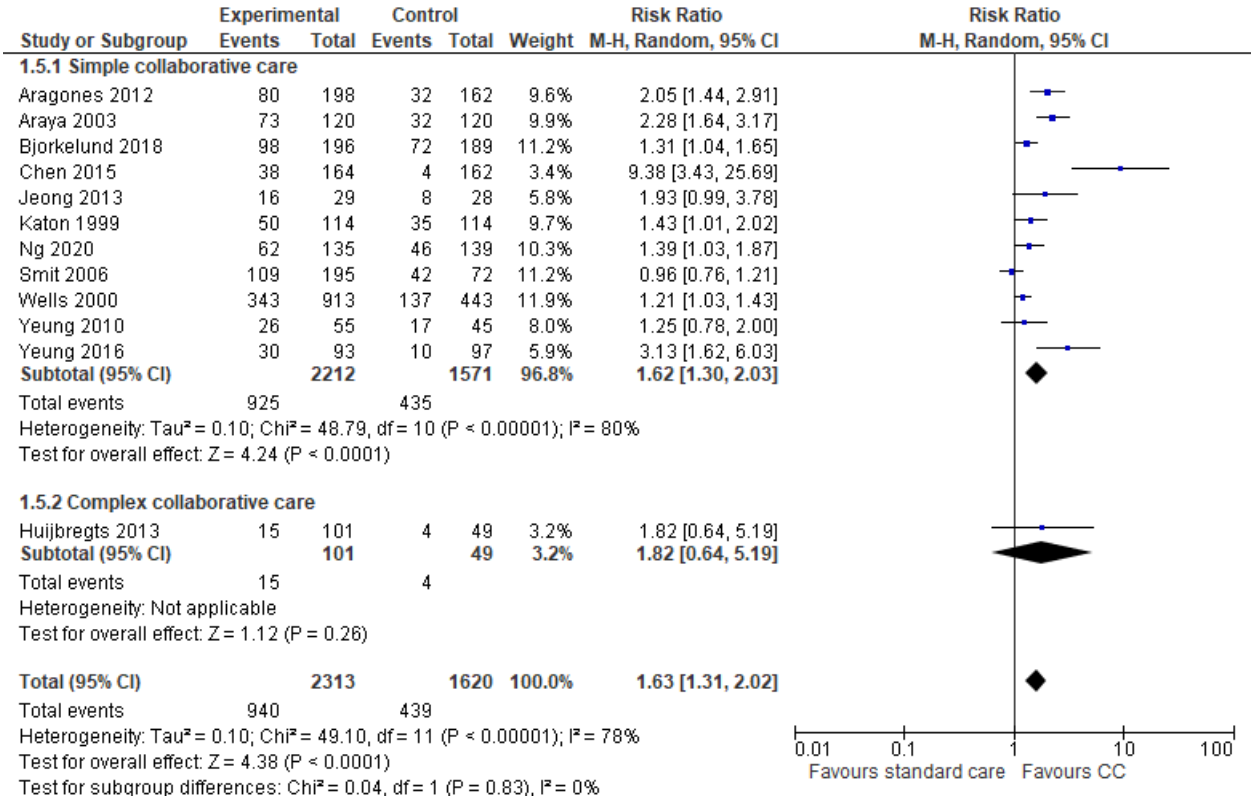
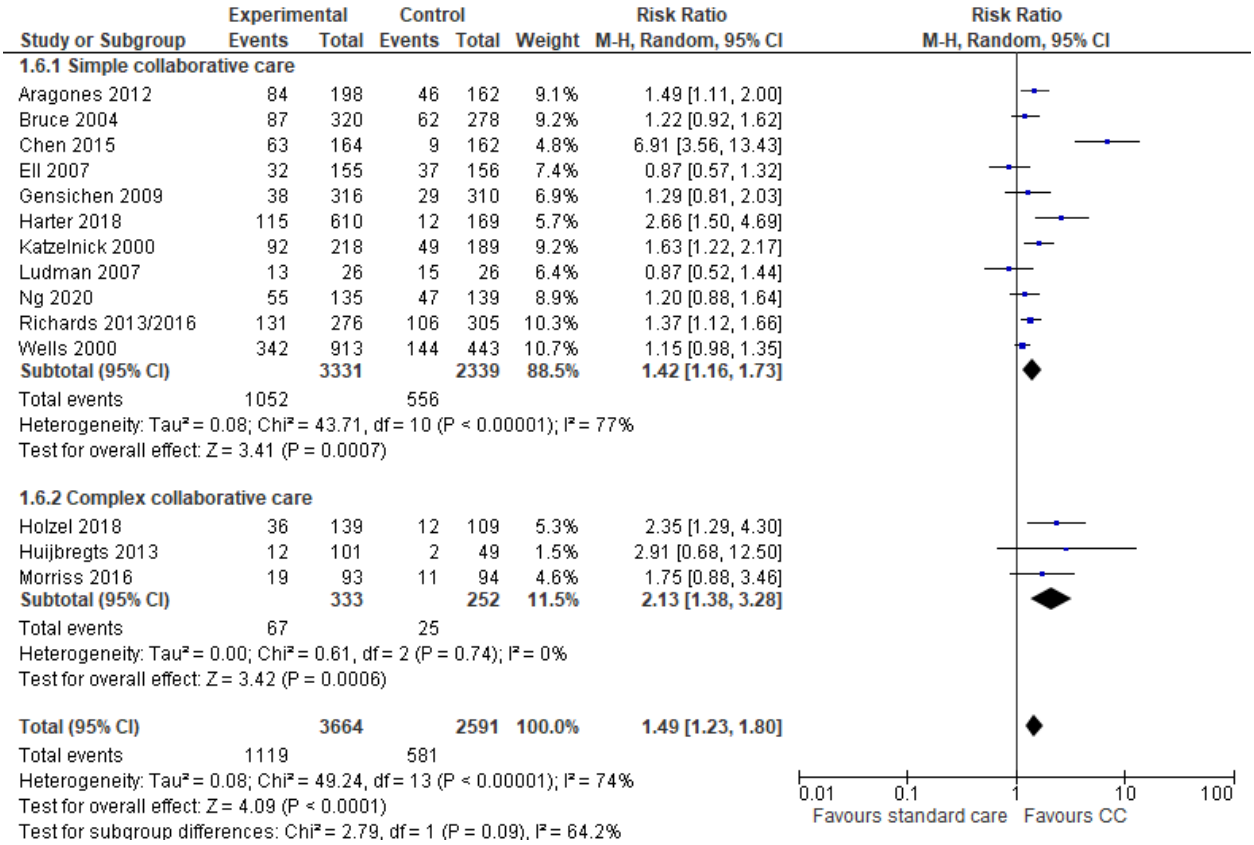


Figure 9: Remission at 12 months



Important outcomes

Figure 10: Antidepressant use at 6 months

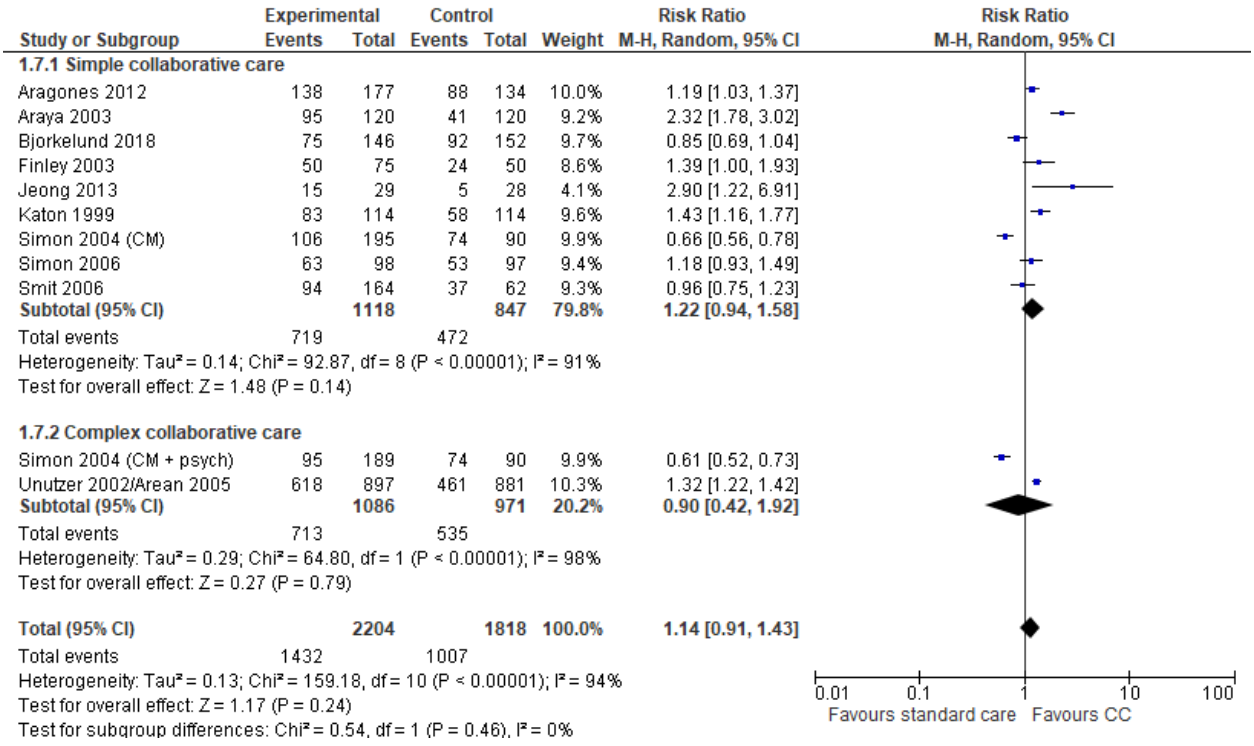


Figure 11: Antidepressant use at 12 months

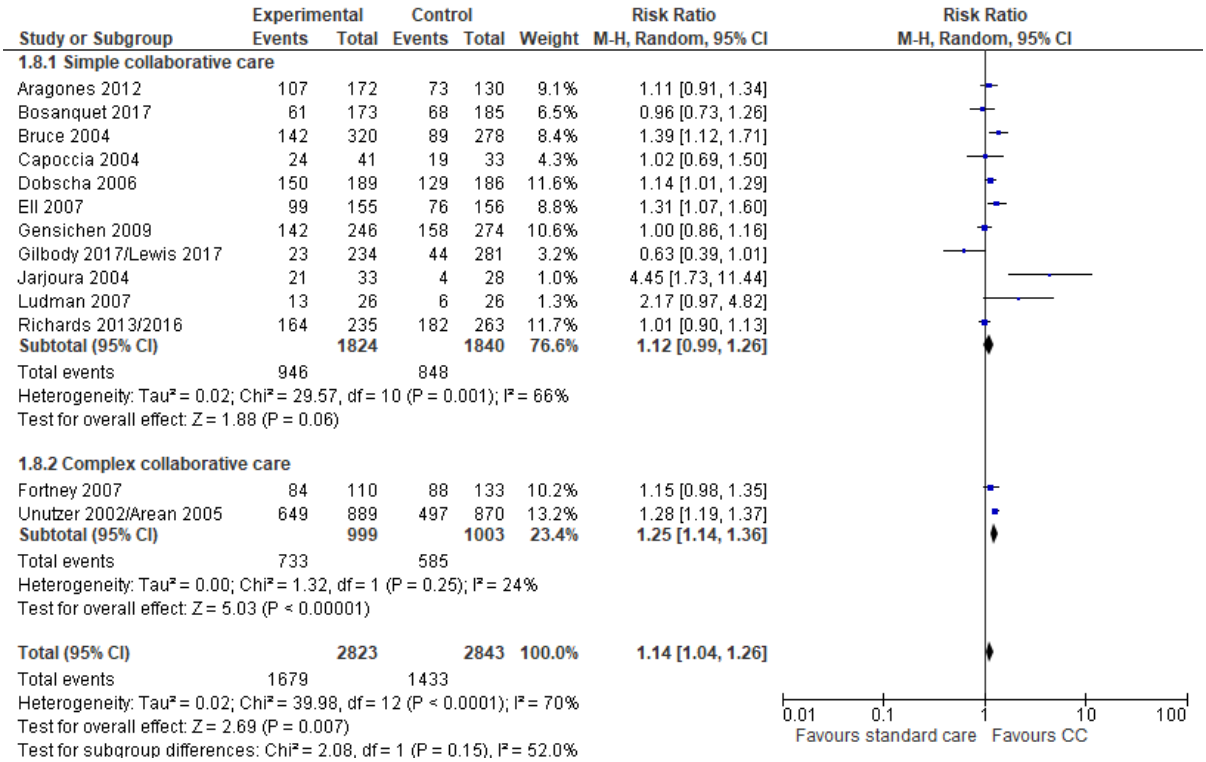


Figure 12: Discontinuation at 6 months

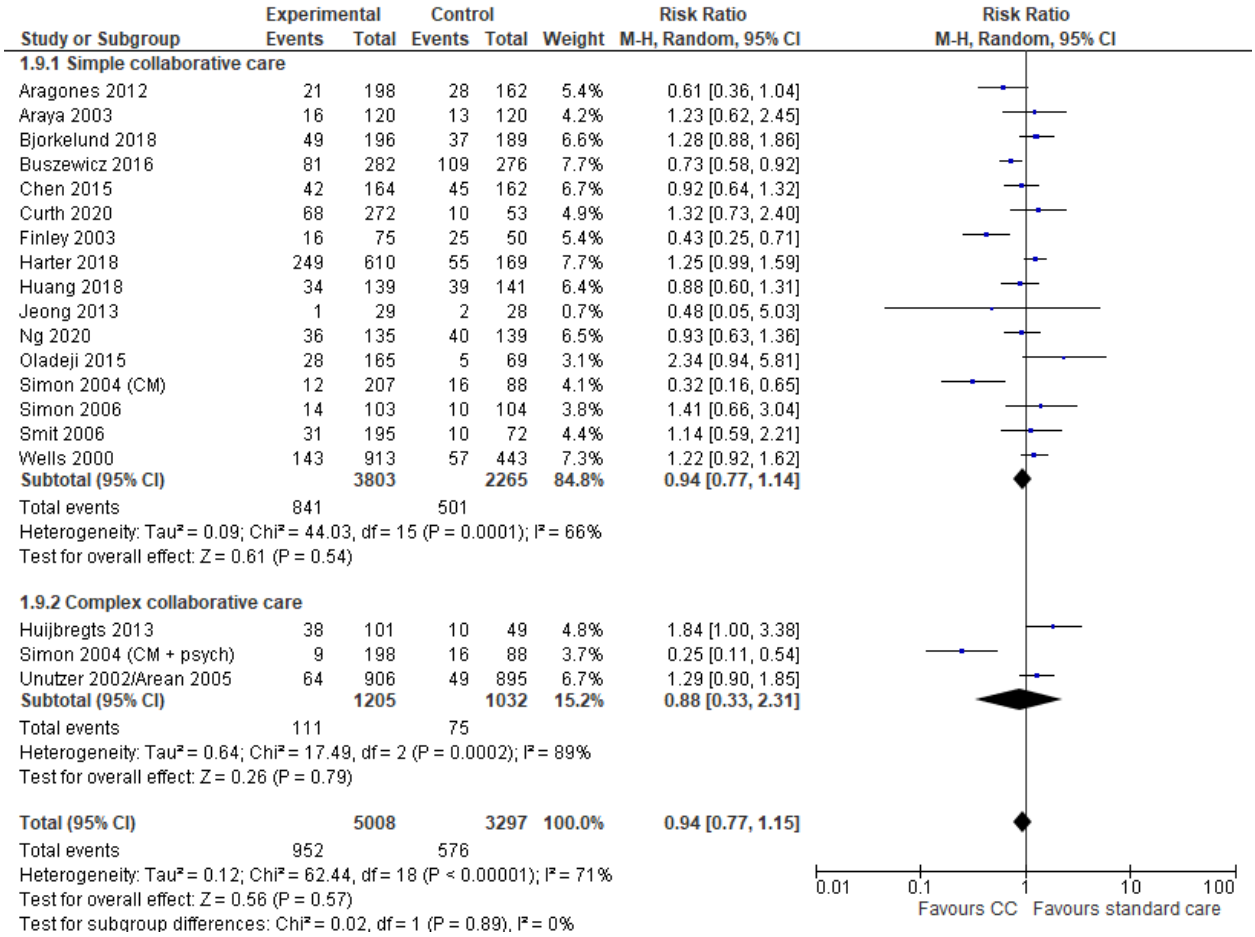
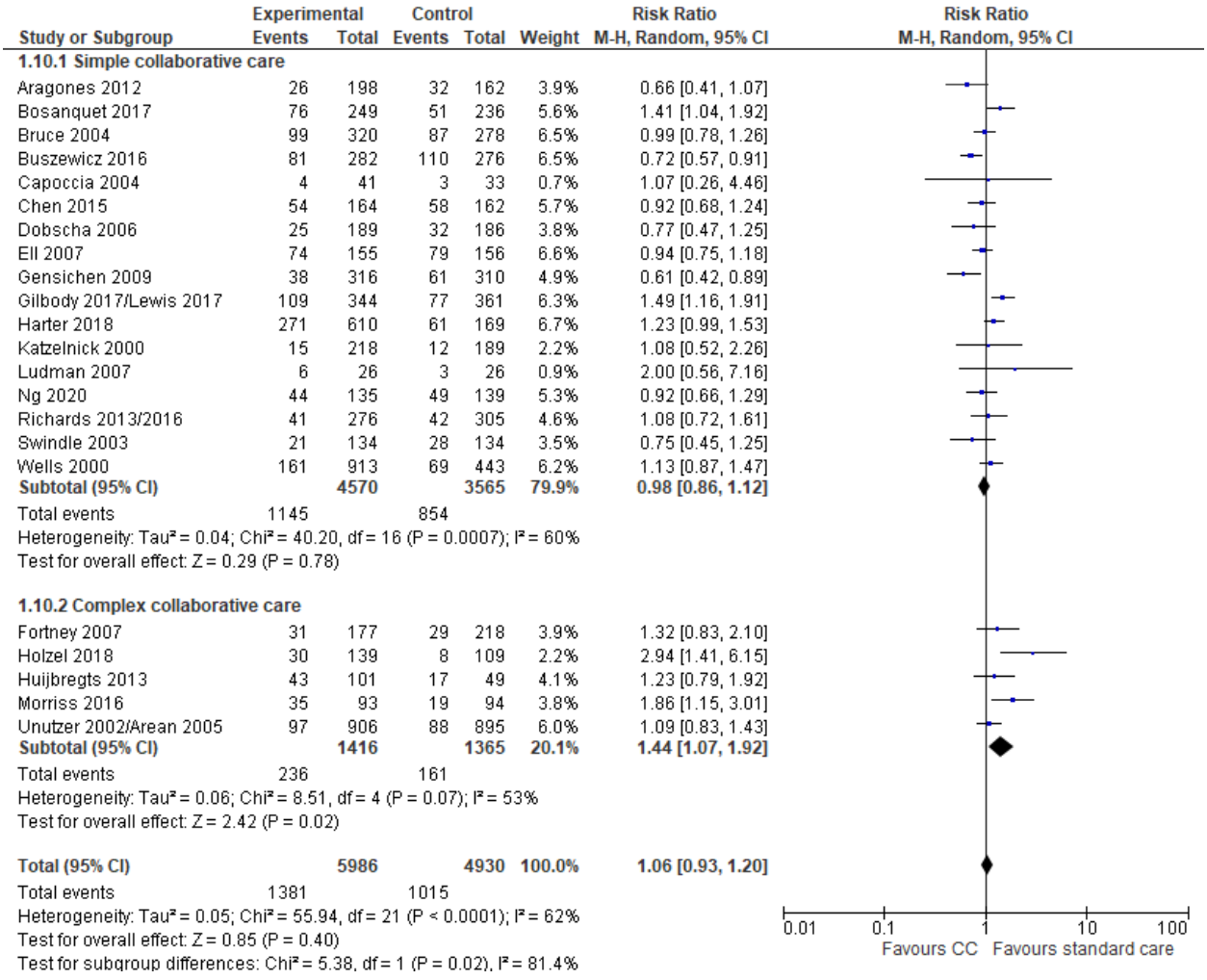


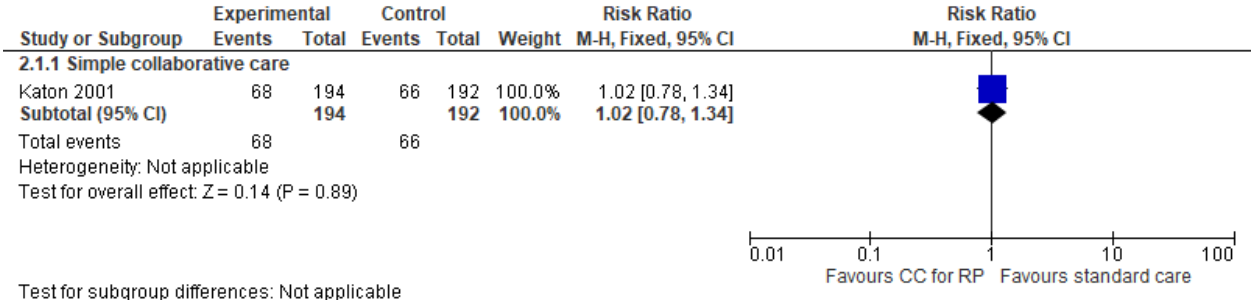
Figure 13: Discontinuation at 12 months



Comparison 2: Collaborative care versus standard care for relapse prevention

Critical outcomes

Figure 14: Relapse at 12 months



Important outcomes

Figure 15: Antidepressant use at 6 months

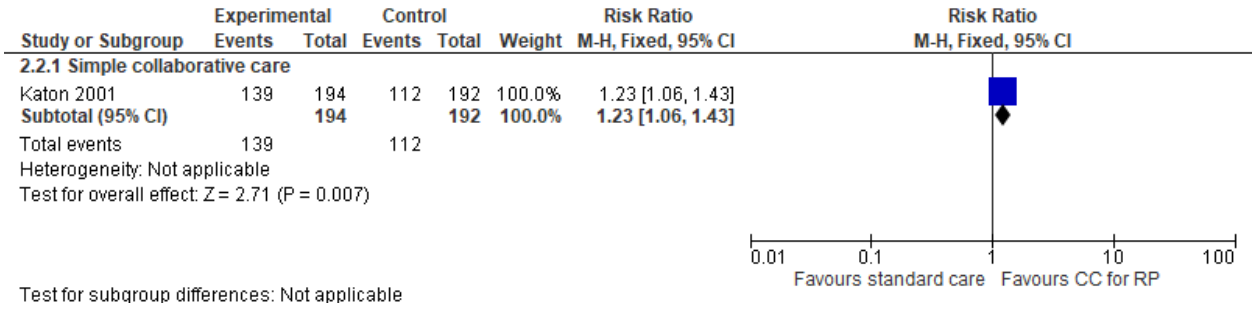


Figure 16: Antidepressant use at 12 months

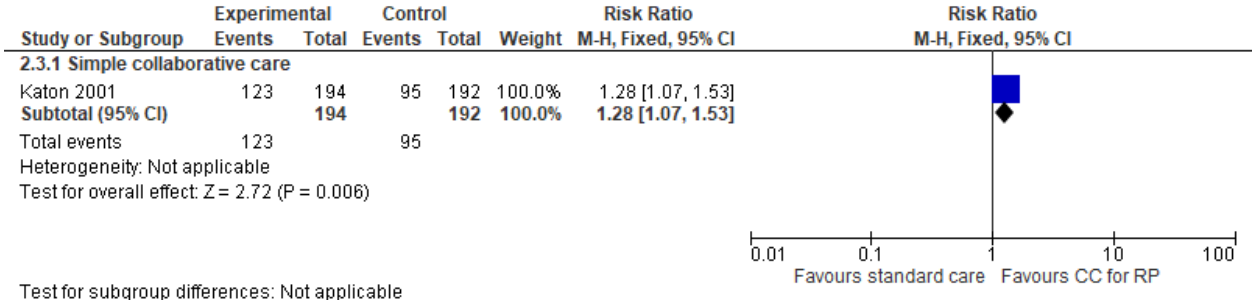
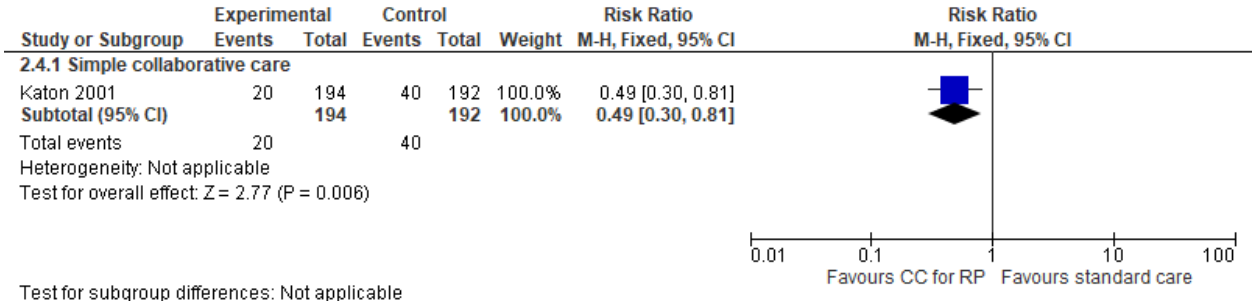


Figure 17: Discontinuation at 12 months



Comparison 3: Stepped care versus standard care/enhanced standard care

Critical outcomes

Figure 18: Depression symptomatology endpoint score at 6 months

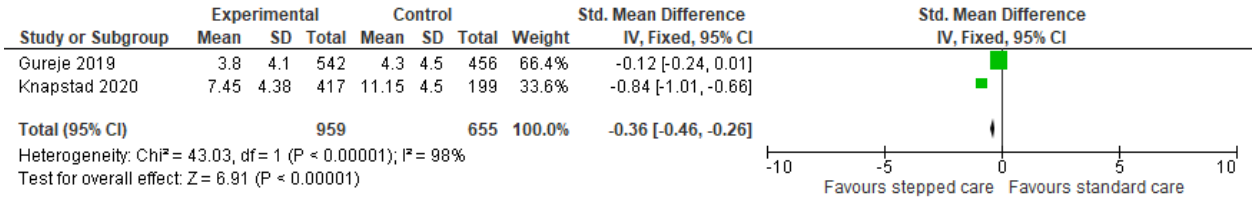


Figure 19: Depression symptomatology change score at 6 months

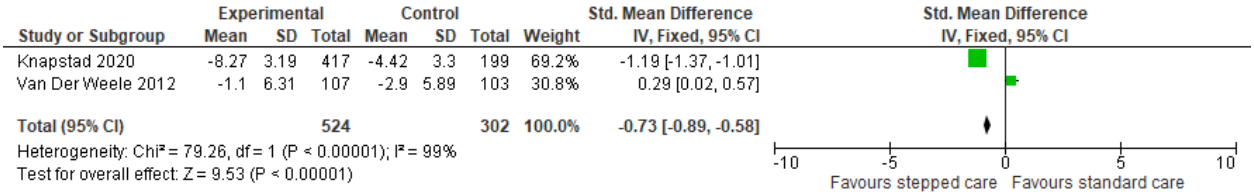


Figure 20: Depression symptomatology endpoint score at 12 months

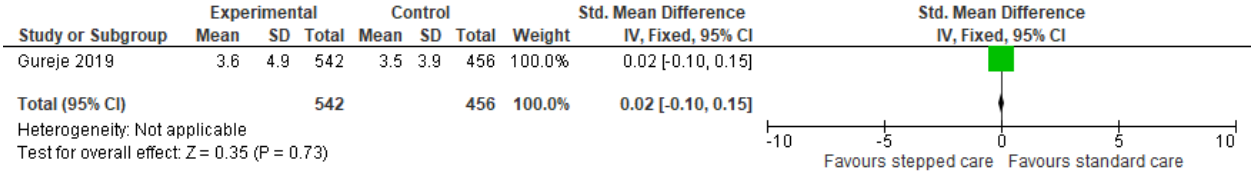


Figure 21: Depression symptomatology change score at 12 months

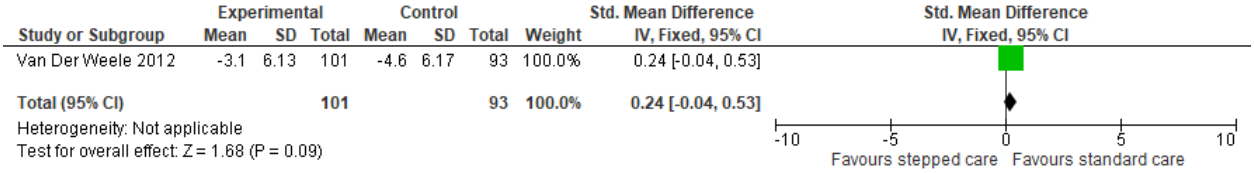


Figure 22: Response at 6 months

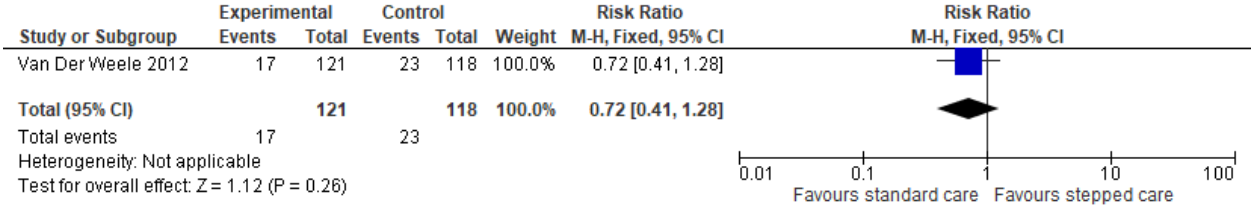


Figure 23: Response at 12 months

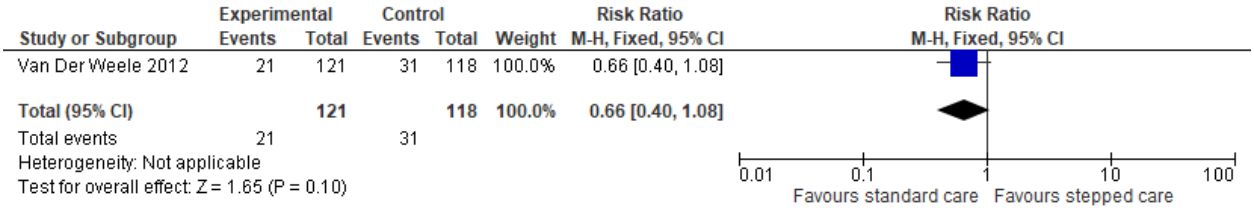


Figure 24: Remission at 6 months

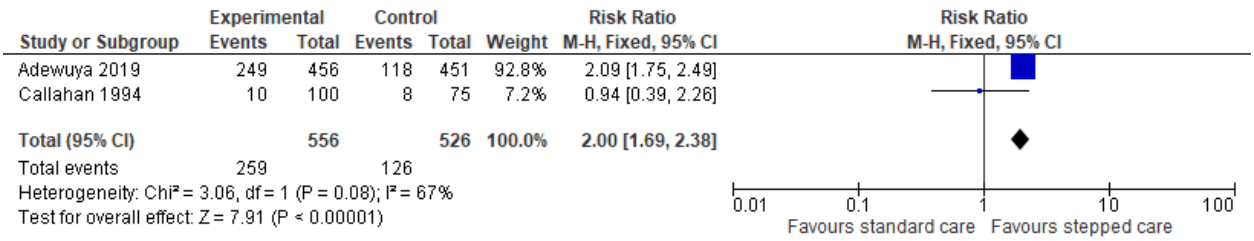
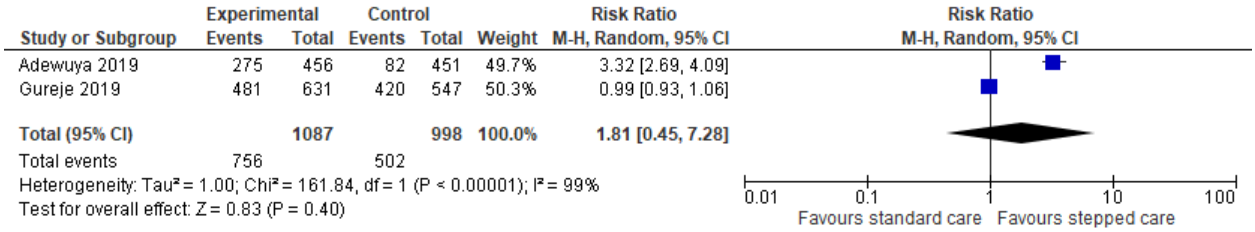


Figure 25: Remission at 12 months



Important outcomes

Figure 26: Antidepressant use at 6 months

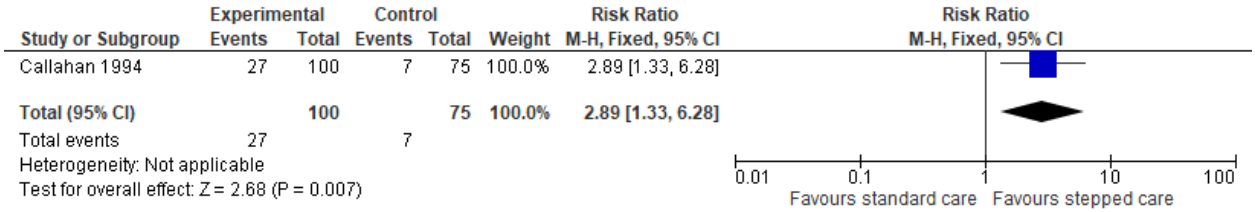


Figure 27: Discontinuation at 6 months

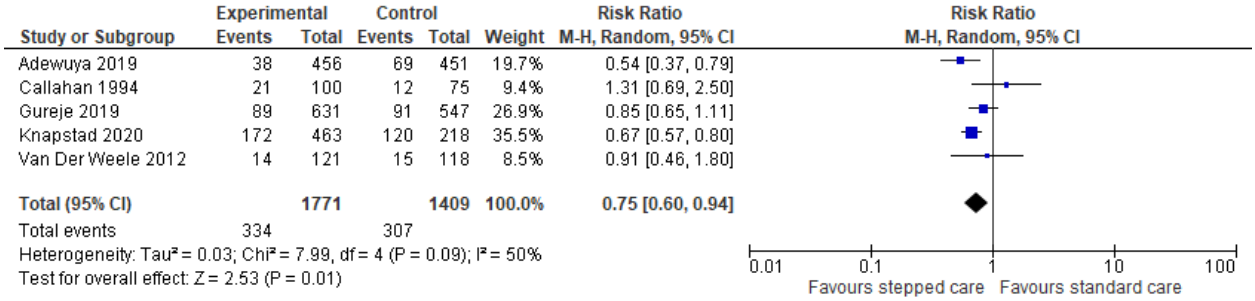
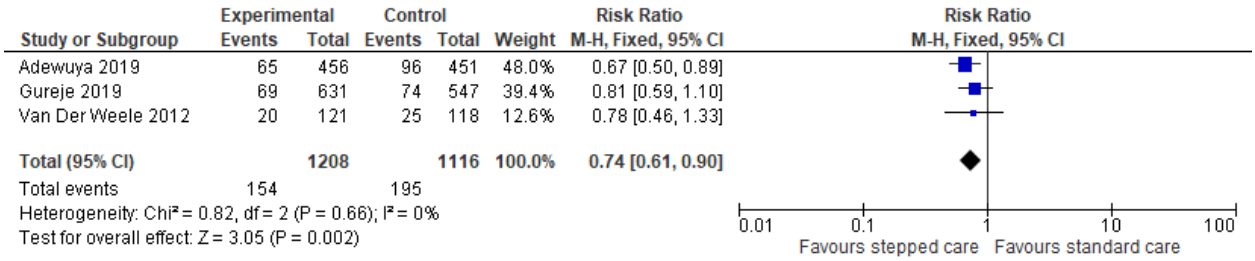


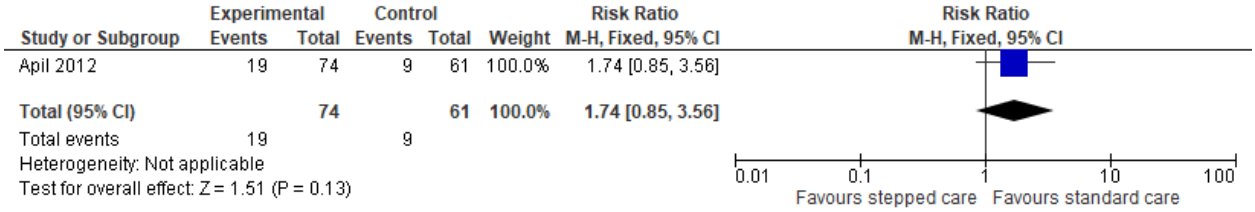
Figure 28: Discontinuation at 12 months



Comparison 4: Stepped care versus standard care for relapse prevention

Critical outcomes

Figure 29: Relapse at 12 months



Important outcomes

Figure 30: Antidepressant use at 12 months

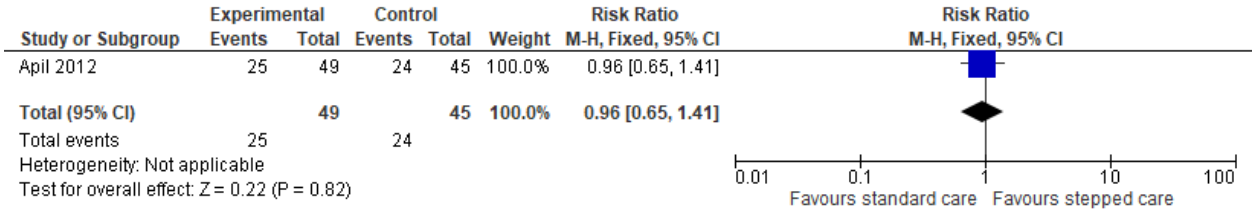
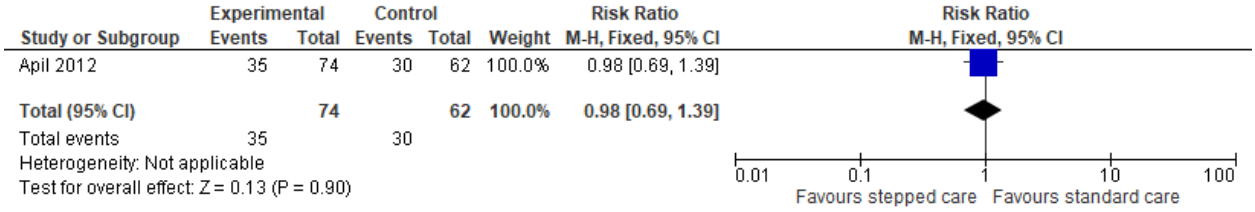


Figure 31: Discontinuation at 12 months



Comparison 5: Pure medication management versus standard care

Critical outcomes

Figure 32: Depression symptomatology at 6 months

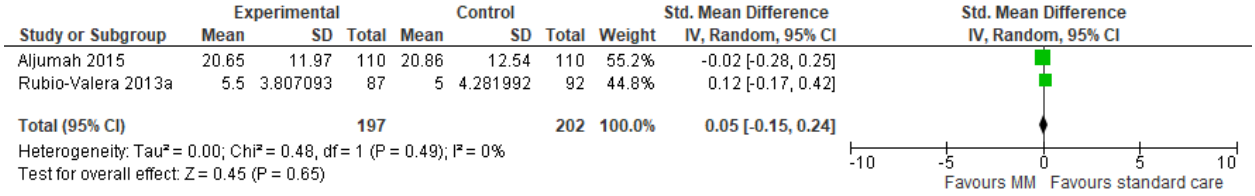
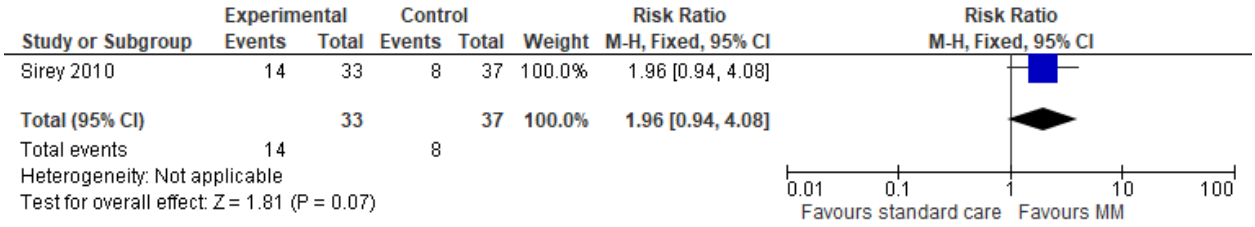


Figure 33: Response at 6 months



Important outcomes

Figure 34: Antidepressant use at 6 months

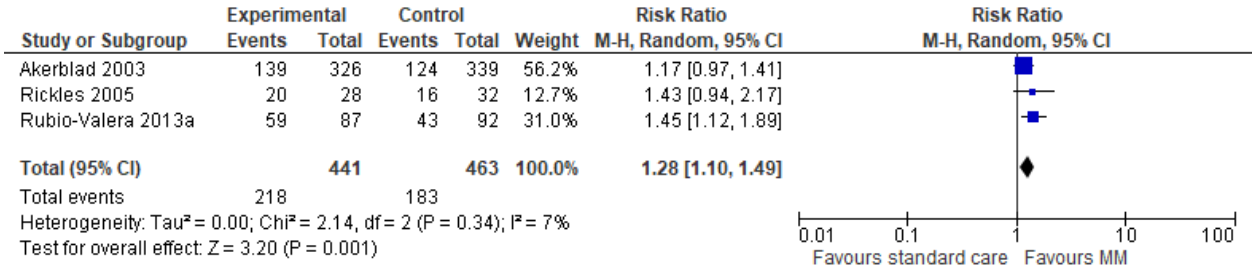
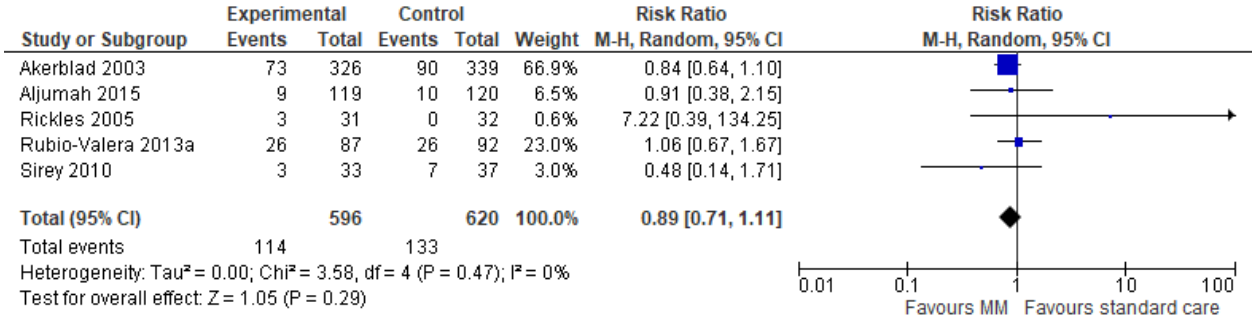


Figure 35: Discontinuation at 6 months



Comparison 6: Care coordination versus standard care/enhanced standard care

Critical outcomes

Figure 36: Depression symptomatology at 6 months

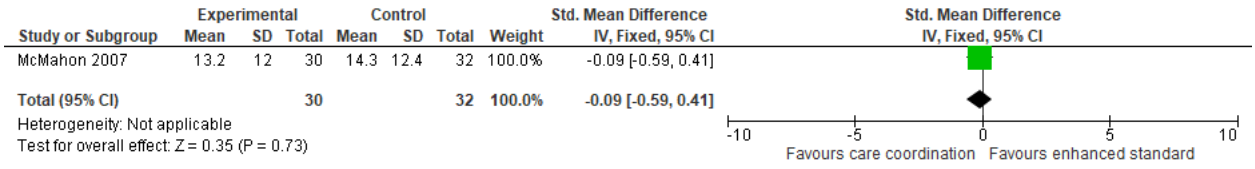


Figure 37: Depression symptomatology at 12 months

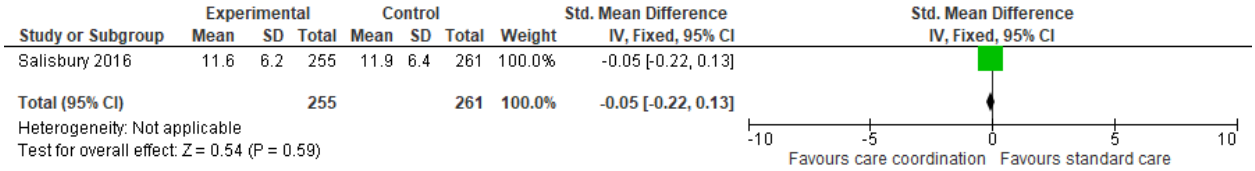
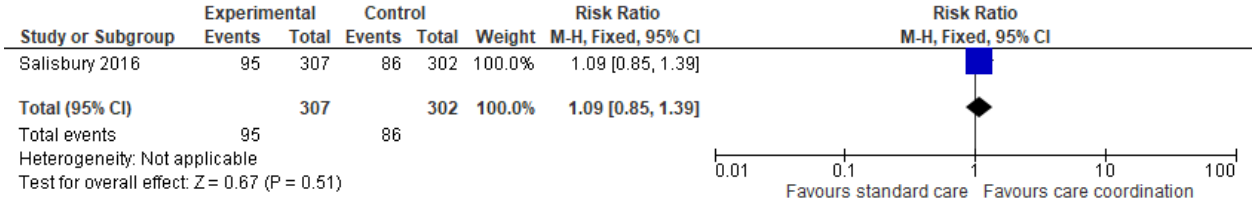


Figure 38: Remission at 12 months



Important outcomes

Figure 39: Discontinuation at 6 months

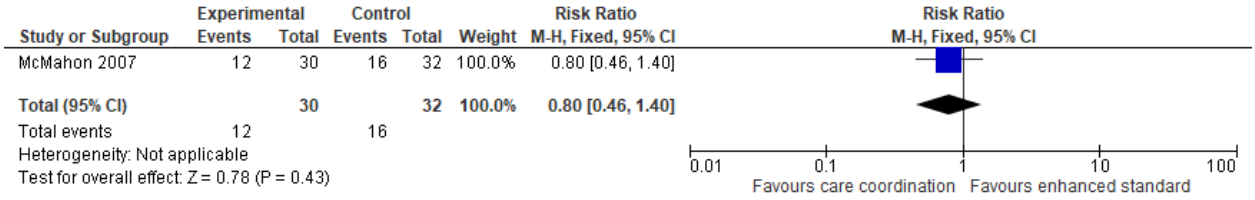
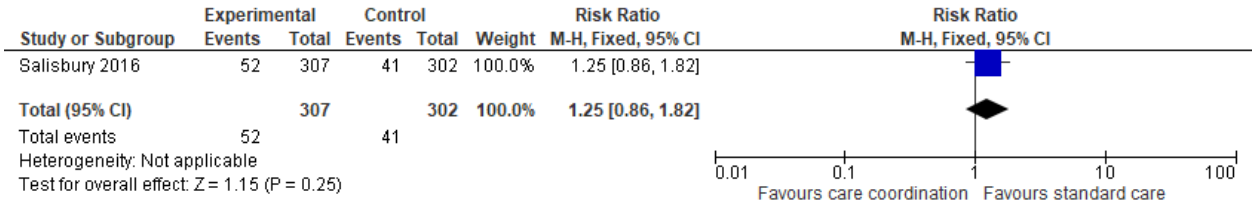


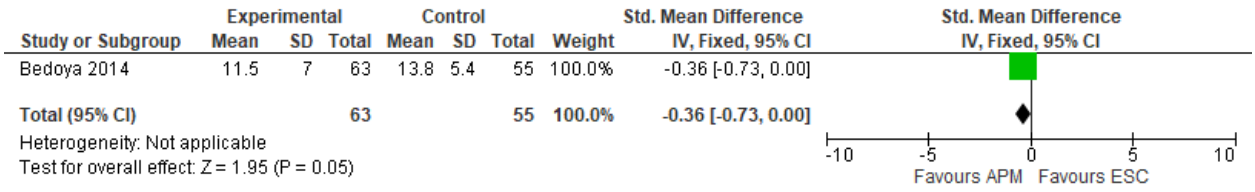
Figure 40: Discontinuation at 12 months



Comparison 7: Attached professional model versus enhanced standard care

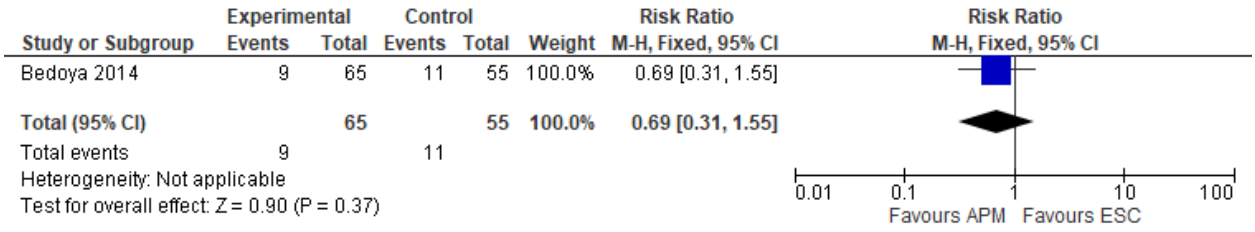
Critical outcomes

Figure 41: Depression symptomatology at 6 months



Important outcomes

Figure 42: Discontinuation at 6 months



Comparison 8: Shared care versus standard care

Critical outcomes

Figure 43: Depression symptomatology at 6 months

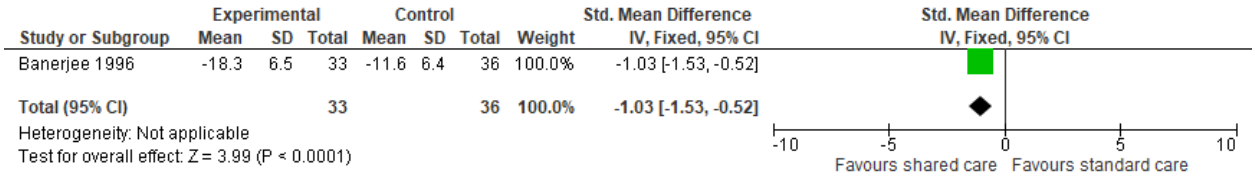
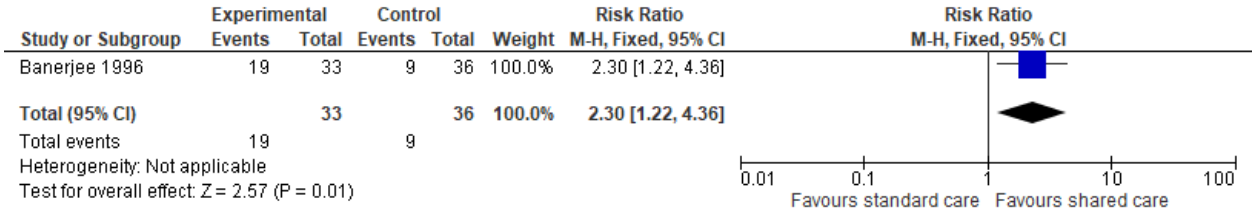


Figure 44: Remission at 6 months



Important outcomes

Figure 45: Antidepressant use at 6 months

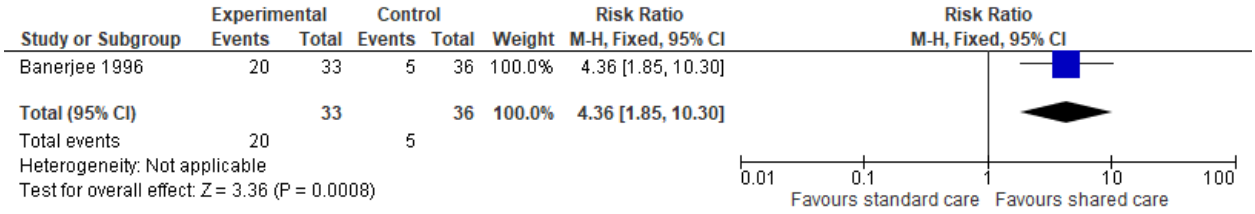
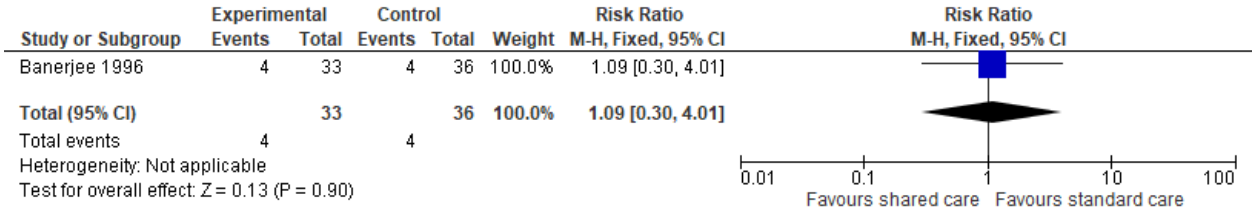


Figure 46: Discontinuation at 6 months



Comparison 9: Measurement-based care versus standard care

Critical outcomes

Figure 47: Depression symptomatology at 6 months

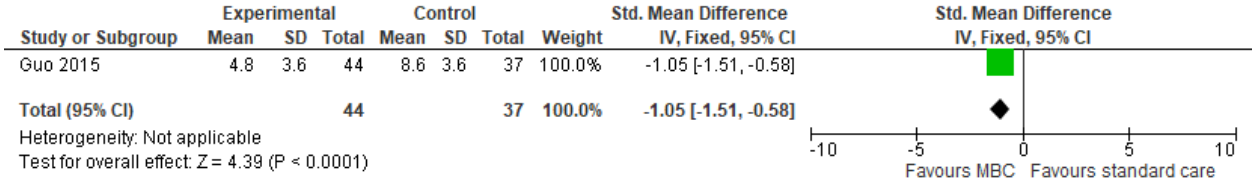


Figure 48: Response at 6 months

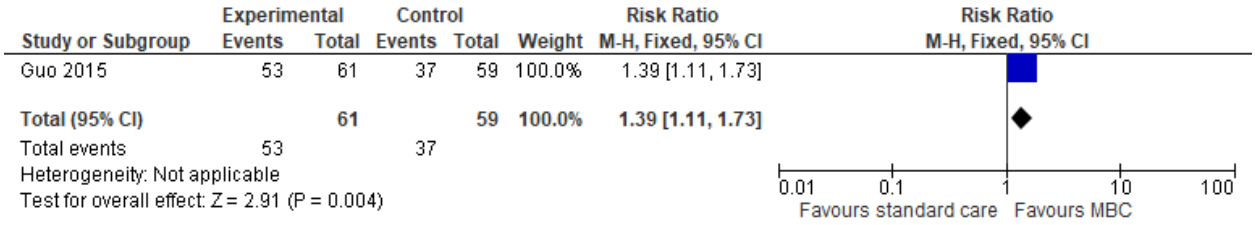
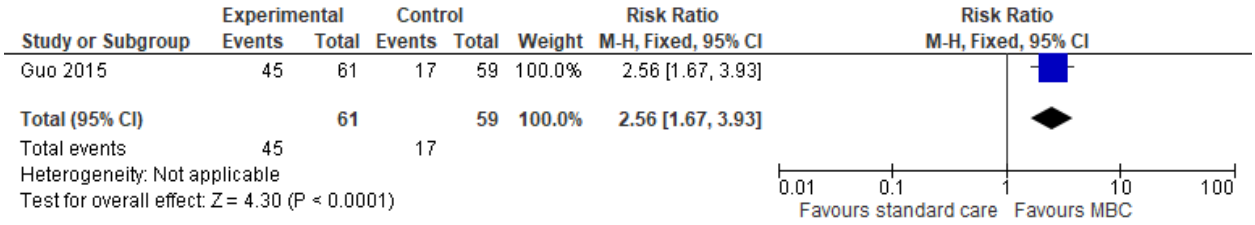


Figure 49: Remission at 6 months



Important outcomes

Figure 50: Discontinuation at 6 months

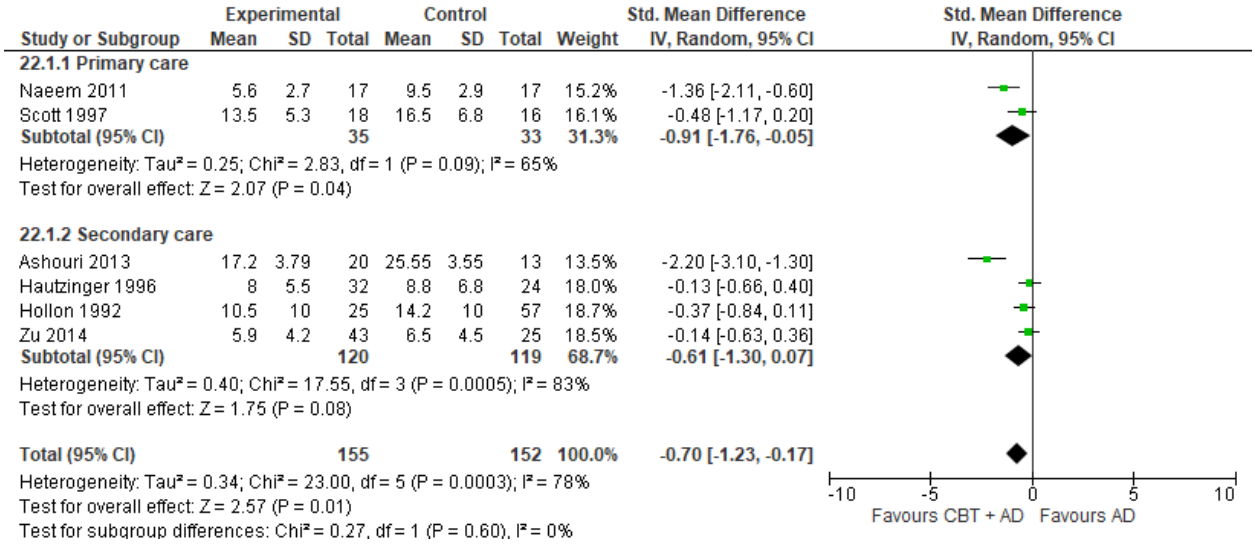
Forest plots for review question 1.2 For adults with depression, what are the relative benefits and harms associated with different settings for the delivery of care?

Comparison 1. Primary care versus secondary care

Primary care versus secondary care subgroup analysis for Comparison 1a Cognitive and cognitive behavioural therapies individual + antidepressant versus antidepressant

Critical outcomes

Figure 51: Depression symptomatology at endpoint



Primary care versus secondary care subgroup analysis for Comparison 1b. Selective serotonin reuptake inhibitors (SSRIs) versus placebo

Critical Outcomes

Figure 52: Depression symptomatology at endpoint

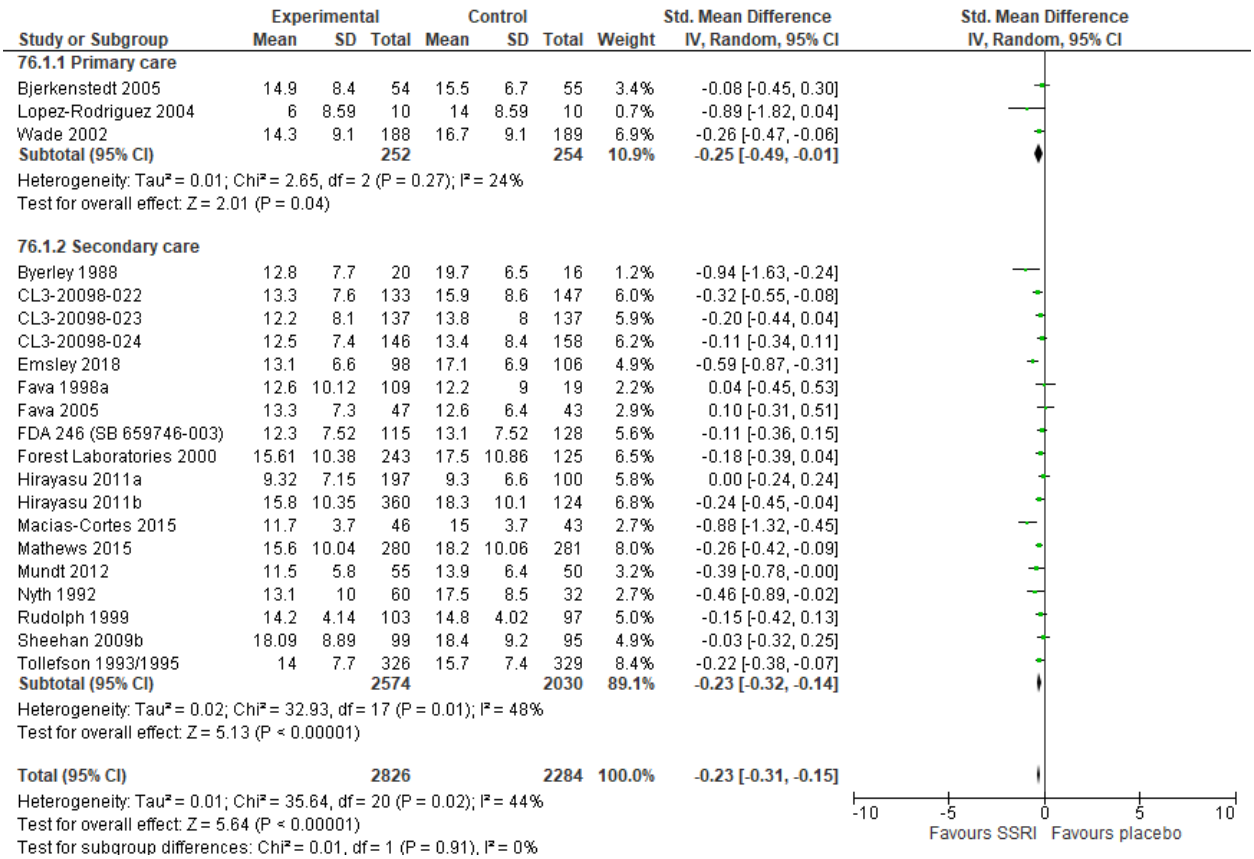


Figure 53: Depression symptomatology change score

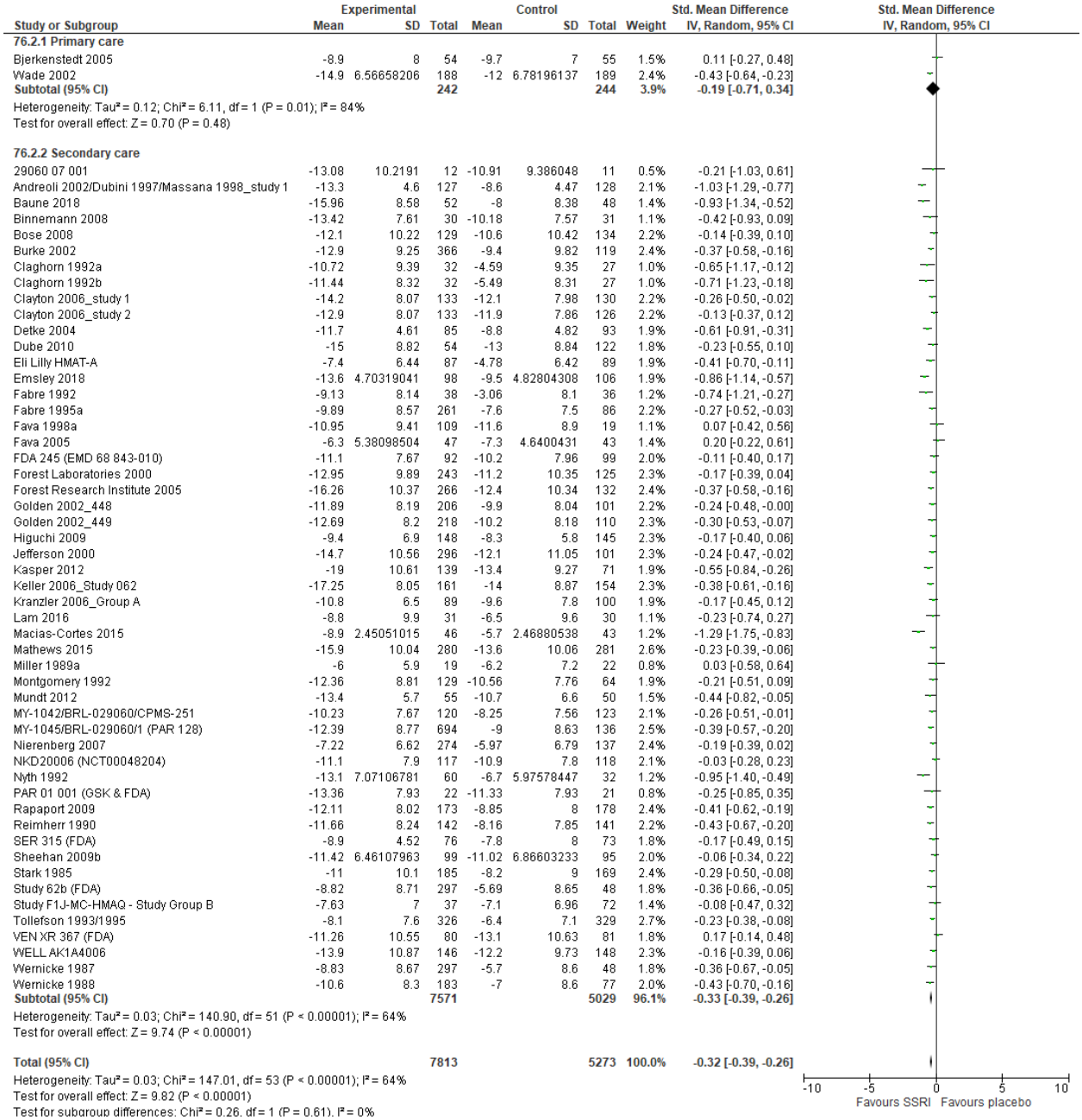
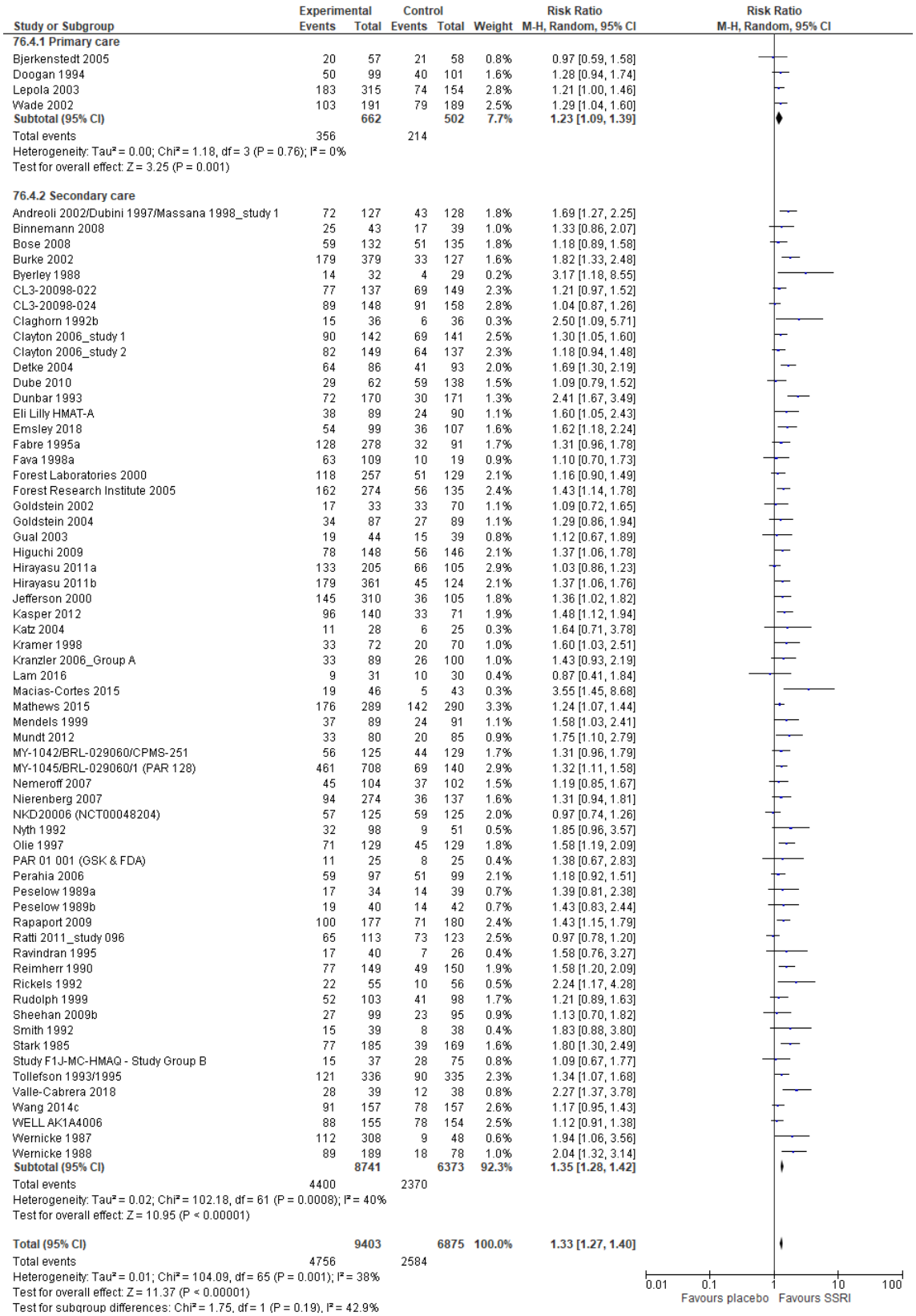


Figure 54: Response



Primary care versus secondary care subgroup analysis for Comparison 1c. SSRIs versus Tricyclic Antidepressants (TCAs)

Critical outcomes

Figure 55: Depression symptomatology at endpoint

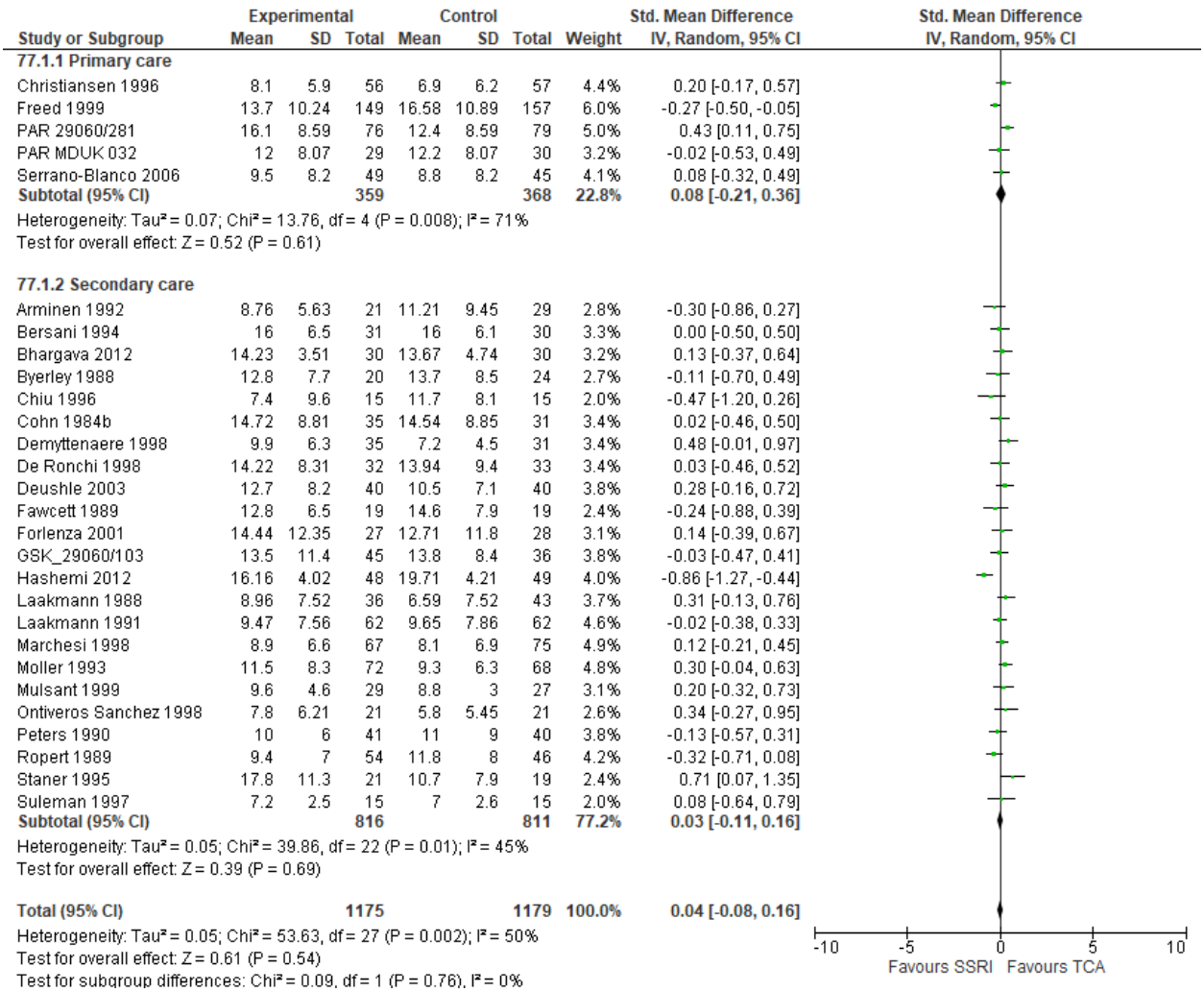


Figure 56: Depression symptomatology change score

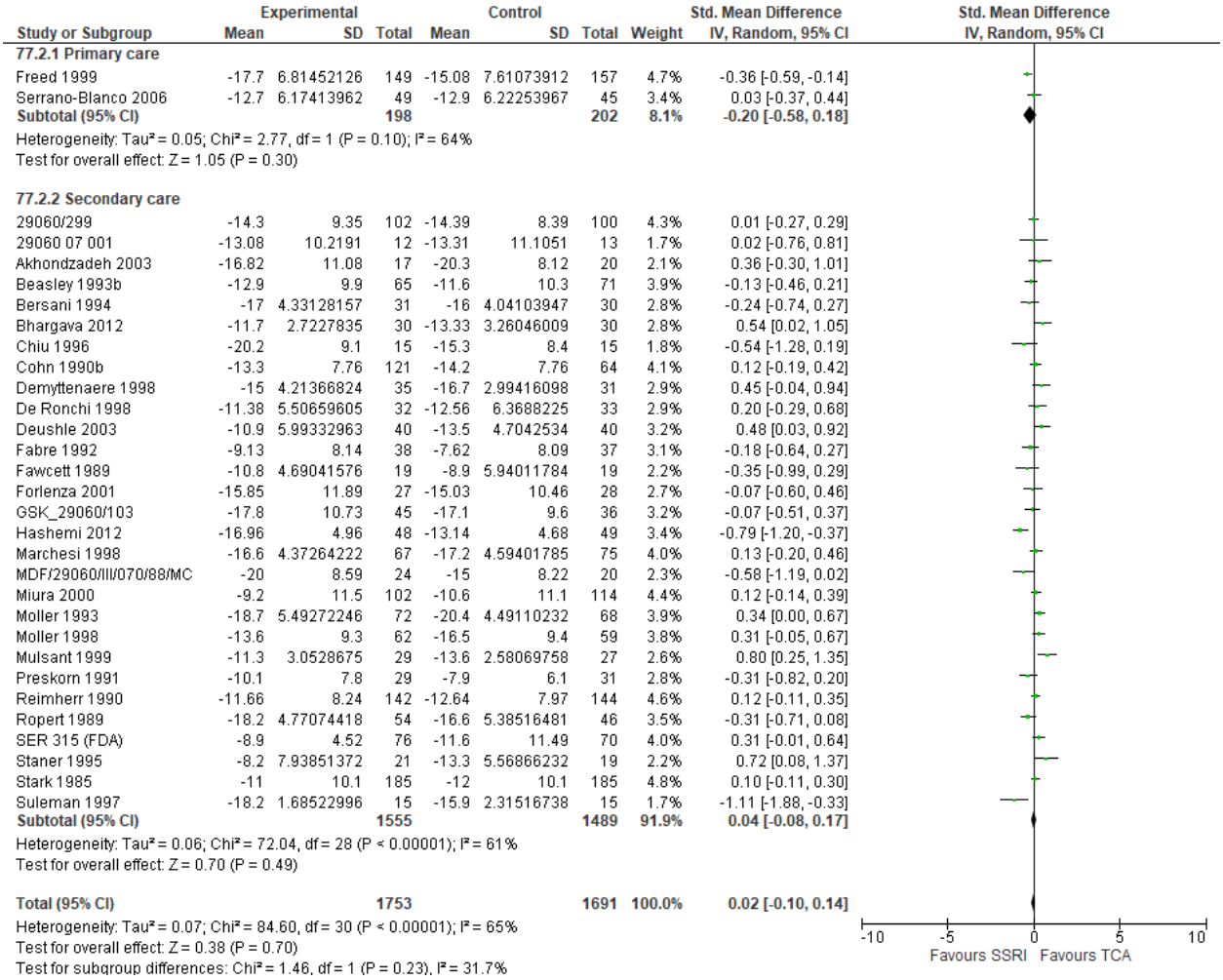


Figure 57: Remission

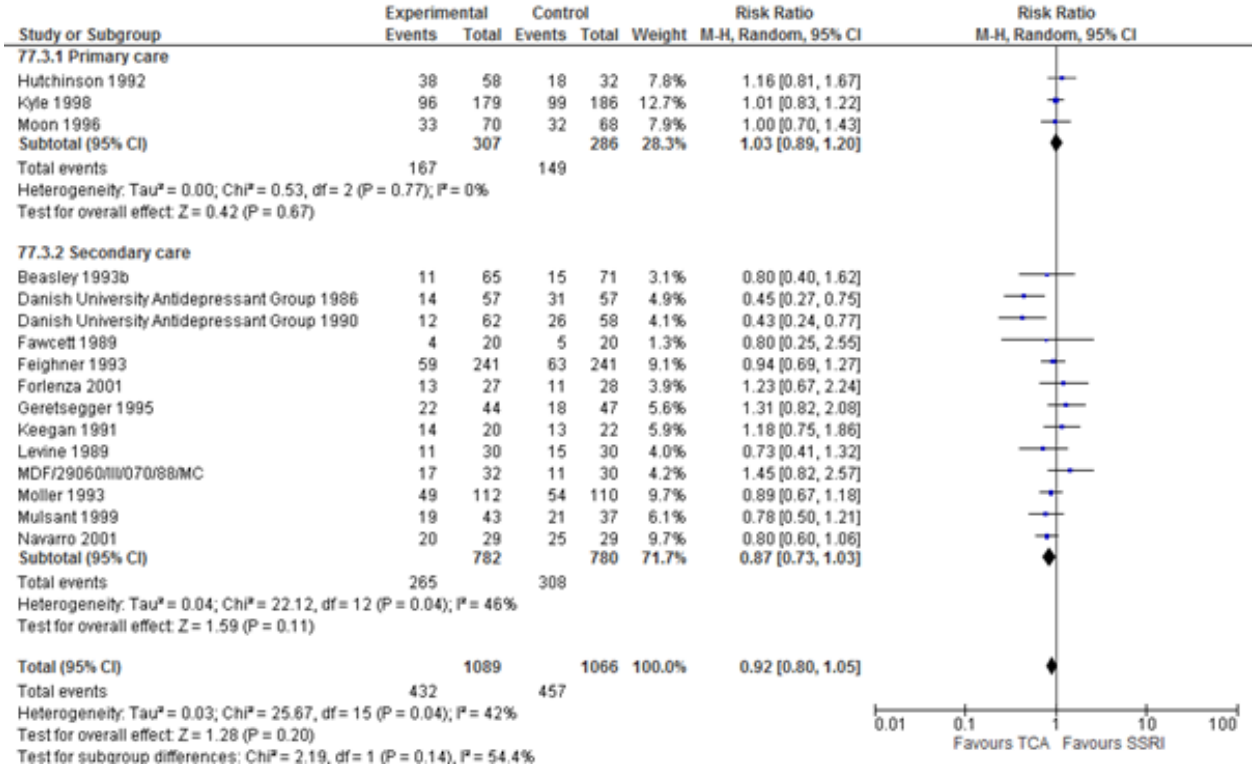
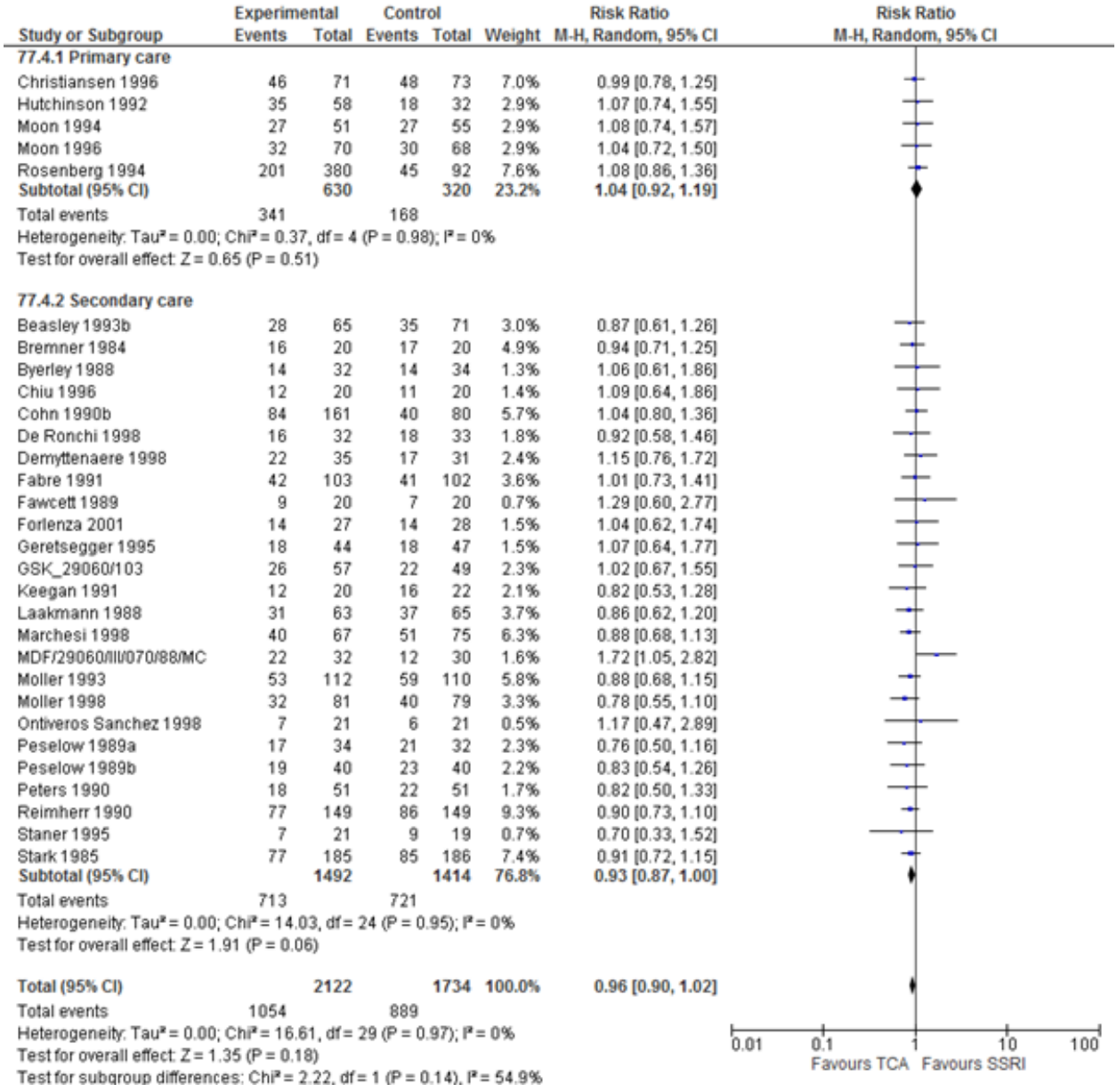


Figure 58: Response



Primary care versus secondary care subgroup analysis for Comparison 1d. TCAs versus placebo

Critical outcomes

Figure 59: Depression symptomatology at endpoint

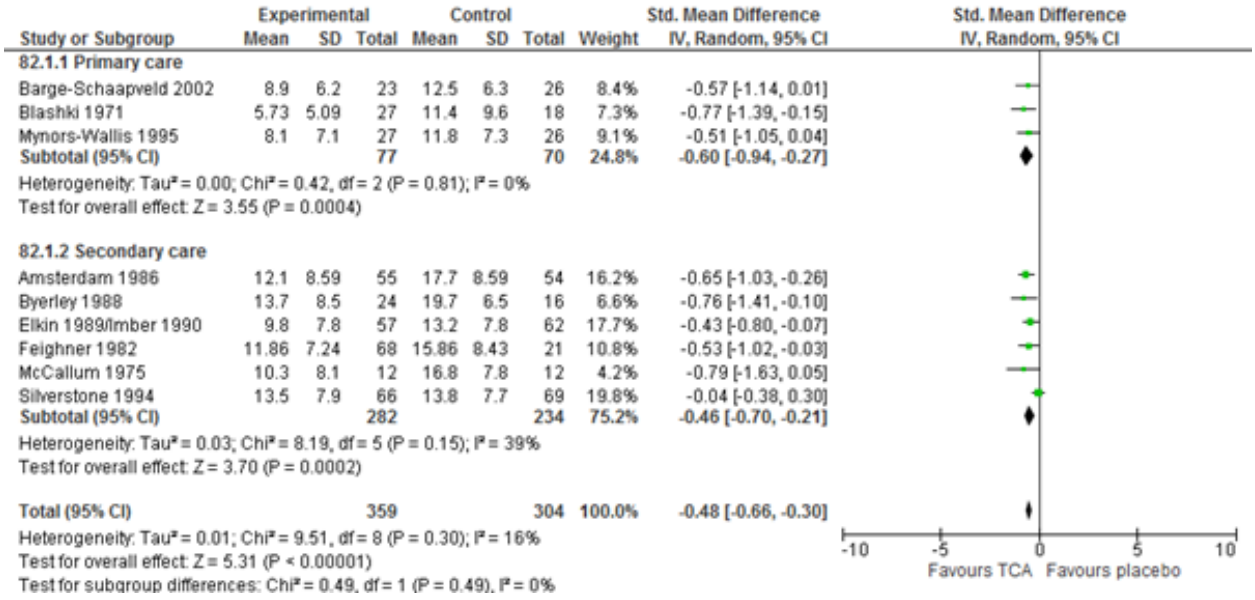


Figure 60: Depression symptomatology change score

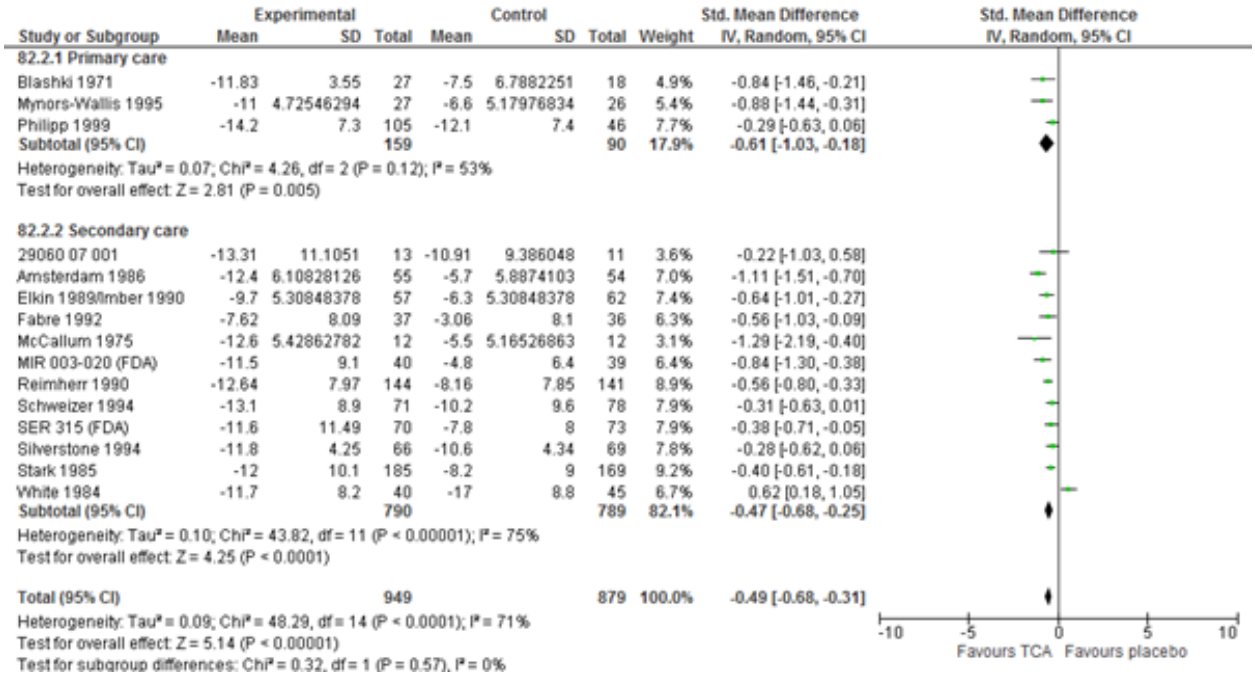
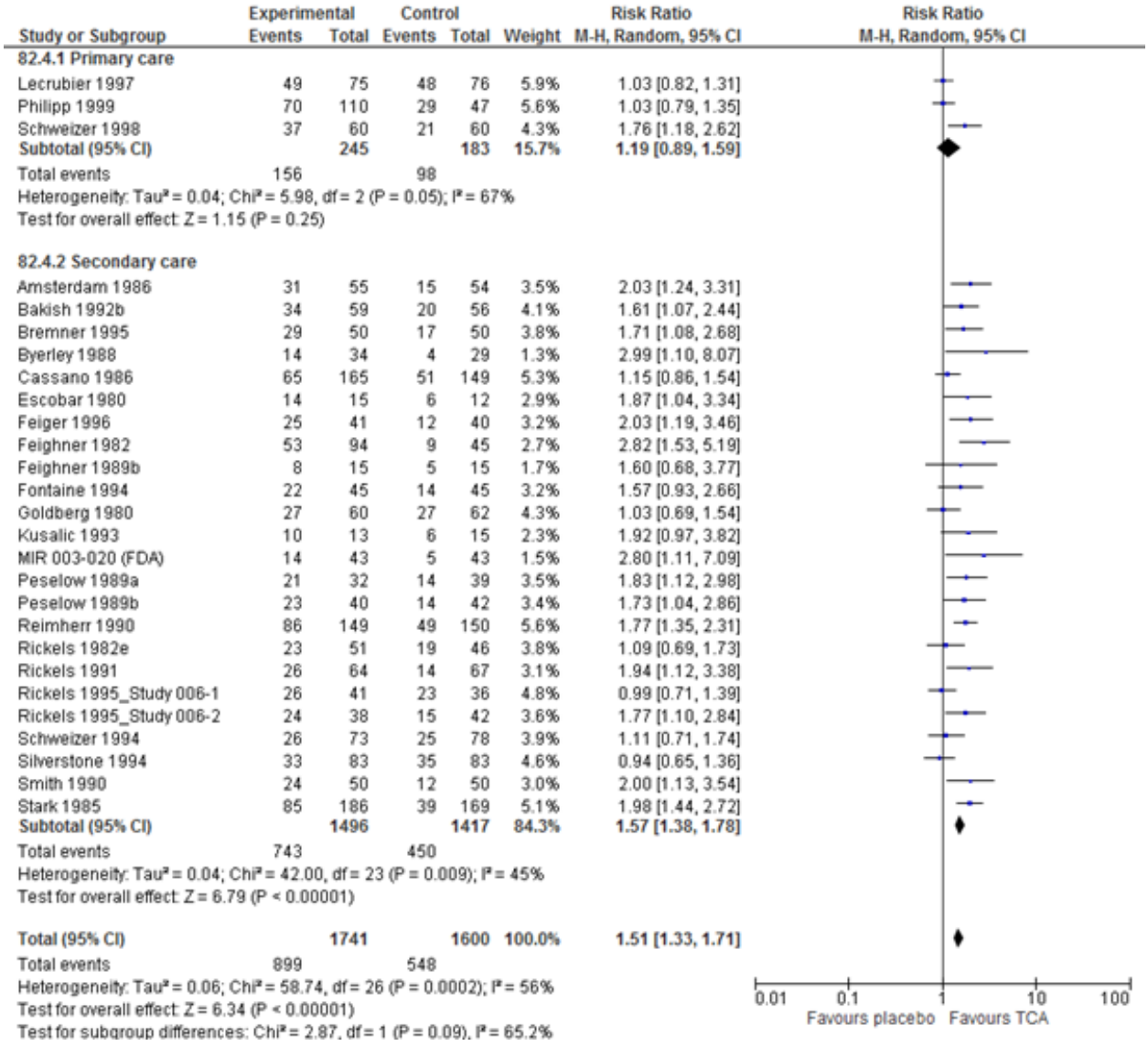


Figure 61: Response



Primary care versus secondary care subgroup analysis for Comparison 1e. Serotonin–norepinephrine reuptake inhibitors (SNRIs) versus SSRIs

Critical outcomes

Figure 62: Remission

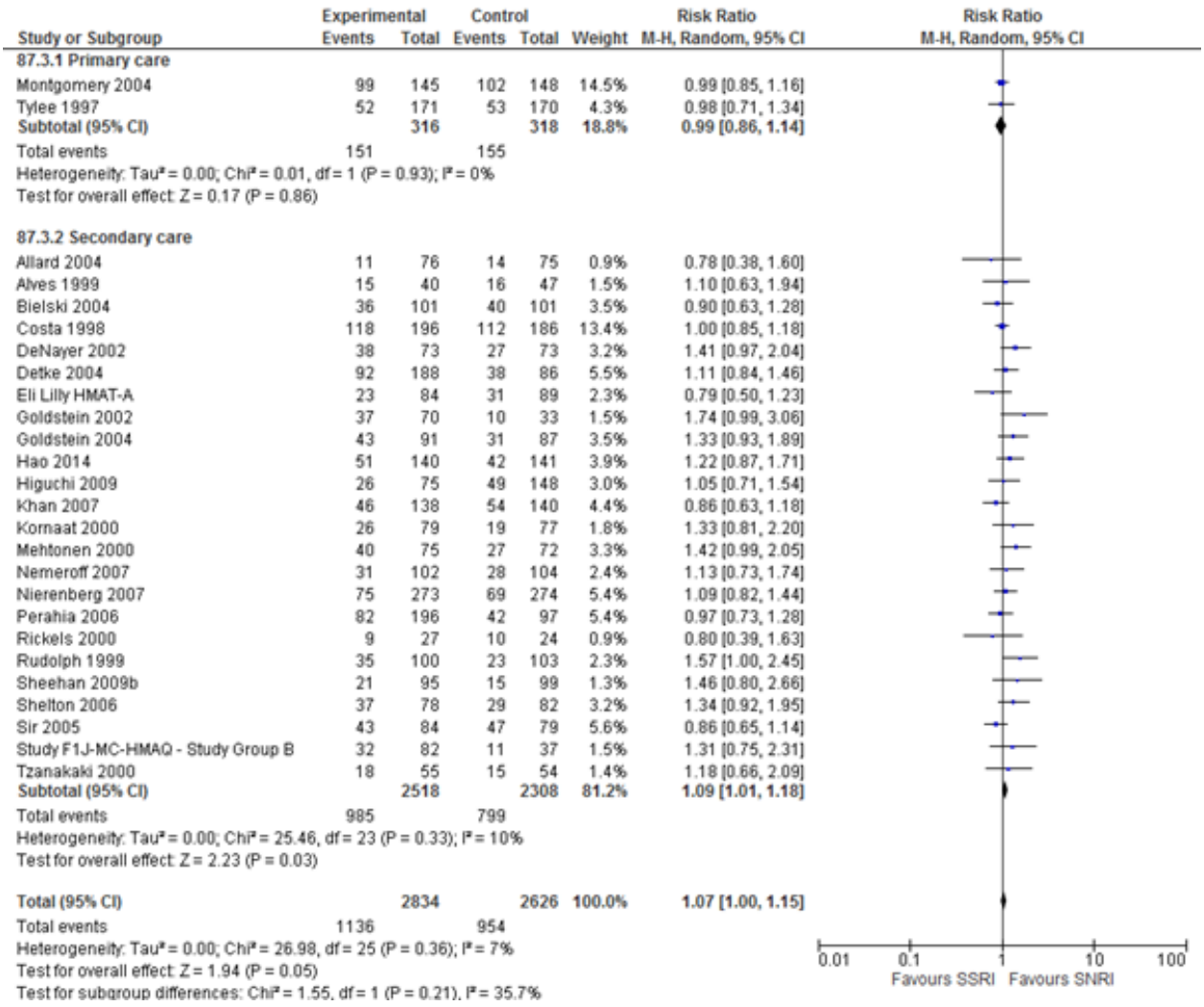
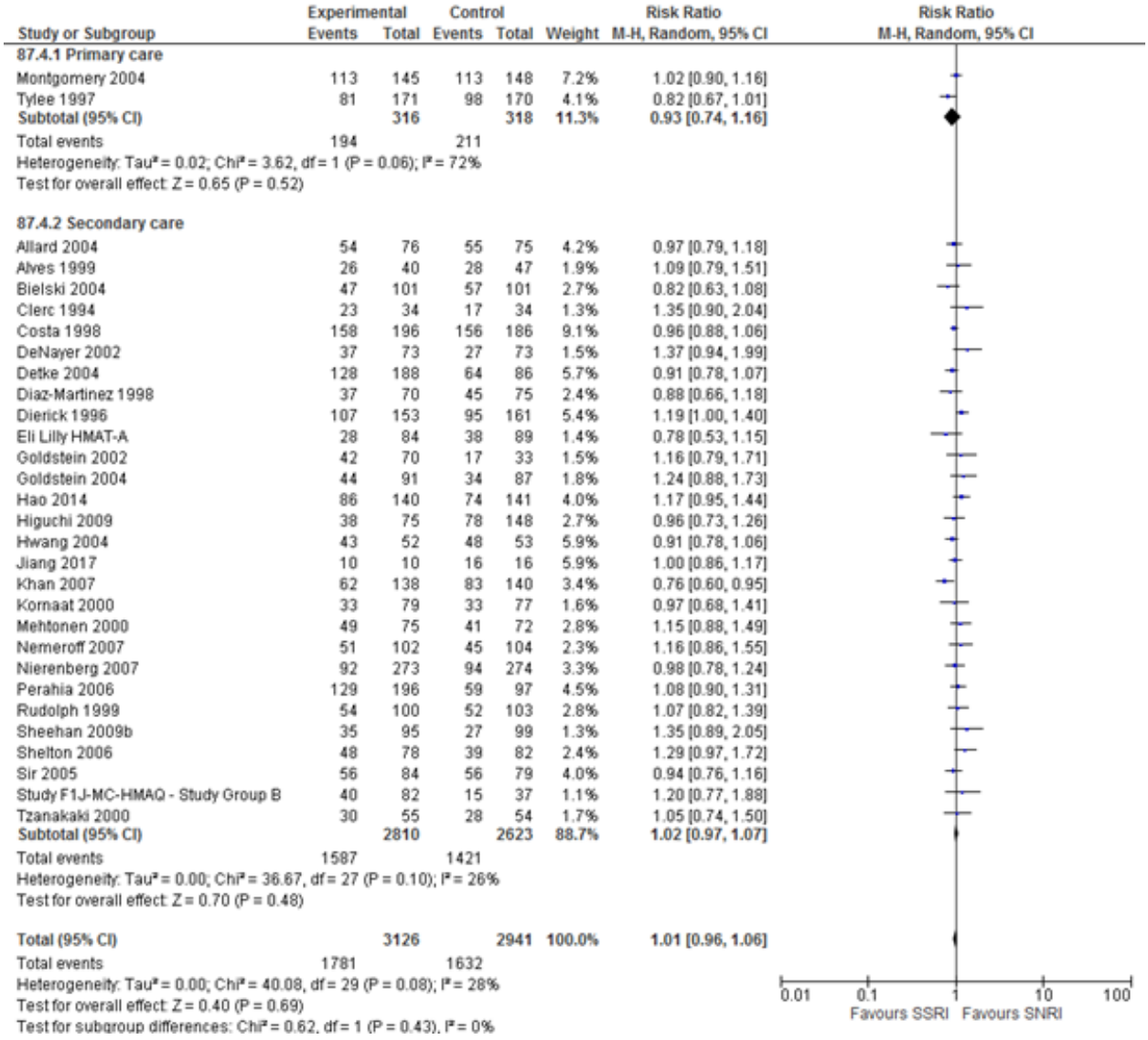


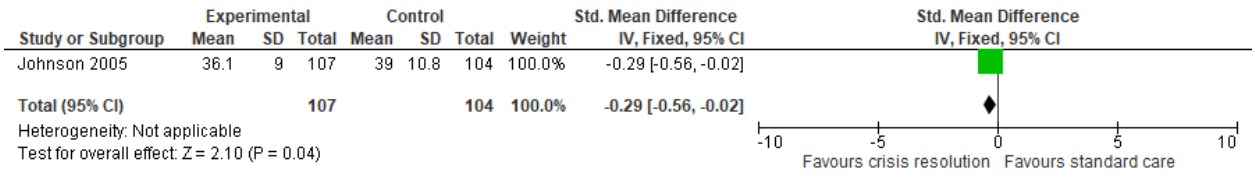
Figure 63: Response



Comparison 2. Crisis resolution team care versus standard care (for adults with non-psychotic severe mental illness)

Critical outcomes

Figure 64: Mental health symptomatology: Symptom severity (BPRS) 8 weeks after crisis



Important outcomes

Figure 65: Service utilisation: Admission as inpatient 6 months after crisis

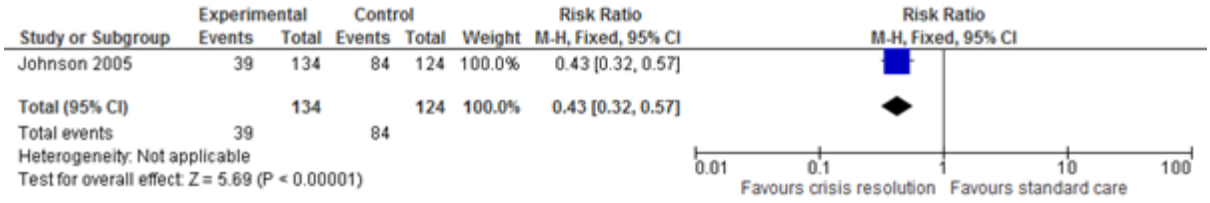


Figure 66: Service utilisation: Bed days in hospital 6 months after crisis

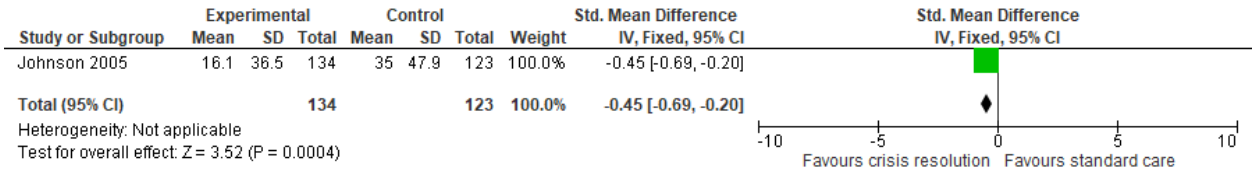


Figure 67: Psychological functioning: Quality of life (MANSA) 8 weeks after crisis

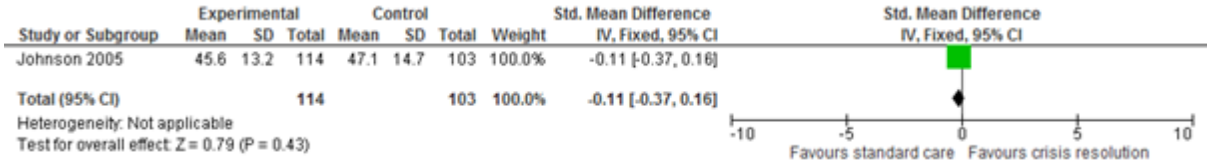


Figure 68: Social functioning: Social functioning (LSP) 8 weeks after crisis

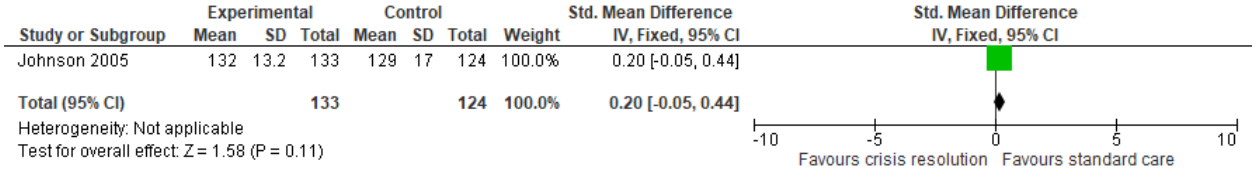


Figure 69: Social functioning: Social functioning (LSP) 6 months after crisis

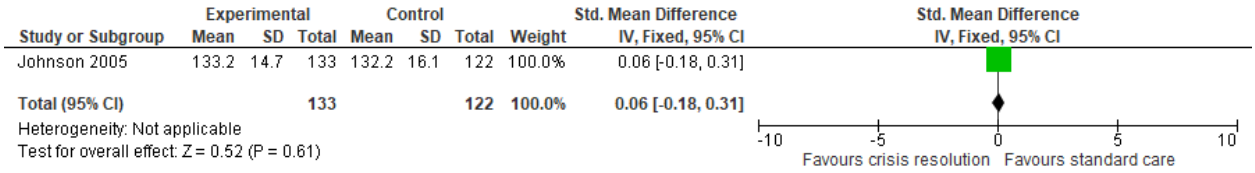
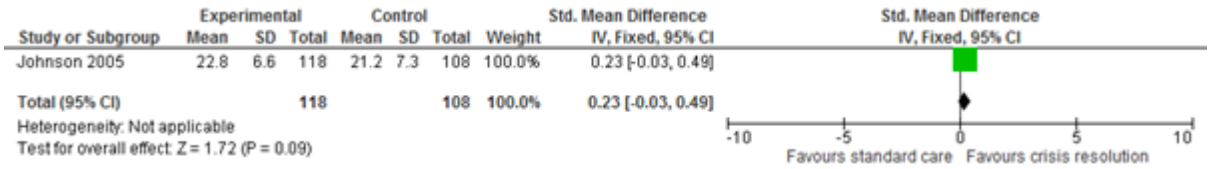


Figure 70: Satisfaction: Patient satisfaction (CSQ-8) 8 weeks after crisis



Comparison 3. Inpatient versus outpatient settings

Inpatient versus outpatient settings subgroup analysis for Comparison 3a. Selective serotonin reuptake inhibitors (SSRIs) versus placebo

Critical outcomes

Figure 71: Depression symptomatology change score

Study or Subgroup	Experimental Mean	Experimental SD	Experimental Total	Control Mean	Control SD	Control Total	Weight	Std. Mean Difference IV, Random, 95% CI
76.1.1 Inpatient								
29060 07 001	-13.08	10.2191	12	-10.91	9.386048	11	0.5%	-0.21 [-1.03, 0.61]
Sheehan 2009b	-11.42	6.46107963	99	-11.02	6.86603233	95	2.2%	-0.06 [-0.34, 0.22]
Subtotal (95% CI)			**111**			**106**	**2.6%**	**-0.08 [-0.34, 0.19]**
Heterogeneity: Tau² = 0.00; Chi² = 0.12, df = 1 (P = 0.73); I² = 0%								
Test for overall effect: Z = 0.56 (P = 0.58)								
76.1.2 Outpatient								
Baune 2018	-15.96	8.58	52	-8	8.38	48	1.4%	-0.93 [-1.34, -0.52]
Binnemann 2008	-13.42	7.61	30	-10.18	7.57	31	1.0%	-0.42 [-0.93, 0.09]
Bjerkenstedt 2005	-8.9	8	54	-9.7	7	55	1.6%	0.11 [-0.27, 0.48]
Blumenthal 2007/Hoffman 2011	-6.1	6.7	49	-6.1	7.3	49	1.5%	0.00 [-0.40, 0.40]
Bose 2008	-12.1	10.22	129	-10.6	10.42	134	2.5%	-0.14 [-0.39, 0.10]
Burke 2002	-12.9	9.25	366	-9.4	9.82	119	2.7%	-0.37 [-0.58, -0.16]
Claghorn 1992a	-10.72	9.39	32	-4.59	9.35	27	1.0%	-0.65 [-1.17, -0.12]
Claghorn 1992b	-11.44	8.32	32	-5.49	8.31	27	1.0%	-0.71 [-1.23, -0.18]
Clayton 2006_study 1	-14.2	8.07	133	-12.1	7.98	130	2.5%	-0.26 [-0.50, -0.02]
Clayton 2006_study 2	-12.9	8.07	133	-11.9	7.86	126	2.4%	-0.13 [-0.37, 0.12]
Detke 2004	-11.7	4.61	85	-8.8	4.82	93	2.0%	-0.61 [-0.91, -0.31]
Dube 2010	-15	8.82	54	-13	8.84	122	1.9%	-0.23 [-0.55, 0.10]
Eli Lilly HMAT-A	-7.4	6.44	87	-4.78	6.42	89	2.0%	-0.41 [-0.70, -0.11]
Emsley 2018	-13.6	4.70319041	98	-9.5	4.82804308	106	2.1%	-0.86 [-1.14, -0.57]
Fabre 1992	-9.13	8.14	38	-3.06	8.1	36	1.2%	-0.74 [-1.21, -0.27]
Fava 1998a	-10.95	9.41	109	-11.6	8.9	19	1.1%	0.07 [-0.42, 0.56]
Fava 2005	-6.3	5.38098504	47	-7.3	4.6400431	43	1.4%	0.20 [-0.22, 0.61]
FDA 245 (EMD 68 843-010)	-11.1	7.67	92	-10.2	7.96	99	2.1%	-0.11 [-0.40, 0.17]
Forest Laboratories 2000	-12.95	9.89	243	-11.2	10.35	125	2.7%	-0.17 [-0.39, 0.04]
Forest Research Institute 2005	-16.26	10.37	266	-12.4	10.34	132	2.7%	-0.37 [-0.58, -0.16]
Golden 2002_448	-11.89	8.19	206	-9.9	8.04	101	2.5%	-0.24 [-0.48, -0.00]
Golden 2002_449	-12.69	8.2	218	-10.2	8.18	110	2.6%	-0.30 [-0.53, -0.07]
Hunter 2011	-9.67	5.78727915	12	-8.64	5.99548163	11	0.5%	-0.17 [-0.99, 0.65]
Jefferson 2000	-14.7	10.56	296	-12.1	11.05	101	2.6%	-0.24 [-0.47, -0.02]
Keller 2006_Study 062	-17.25	8.05	161	-14	8.87	154	2.6%	-0.38 [-0.61, -0.16]
Komulainen 2018	-1.9	3.05569959	17	-2.2	3.29146624	15	0.6%	0.09 [-0.60, 0.79]
Kranzler 2006_Group A	-10.8	6.5	89	-9.6	7.8	100	2.1%	-0.17 [-0.45, 0.12]
Lam 2016	-8.8	9.9	31	-6.5	9.6	30	1.1%	-0.23 [-0.74, 0.27]
Macias-Cortes 2015	-8.9	2.45051015	46	-5.7	2.46880538	43	1.2%	-1.29 [-1.75, -0.83]
Mathews 2015	-15.9	10.04	280	-13.6	10.06	281	3.1%	-0.23 [-0.39, -0.06]
Miller 1989a	-6	5.9	19	-6.2	7.2	22	0.8%	0.03 [-0.58, 0.64]
Mundt 2012	-13.4	5.7	55	-10.7	6.6	50	1.5%	-0.44 [-0.82, -0.05]
MY-1042/BRL-029060/CPMS-251	-10.23	7.67	120	-8.25	7.56	123	2.4%	-0.26 [-0.51, -0.01]
MY-1045/BRL-029060/1 (PAR 128)	-12.39	8.77	694	-9	8.63	136	3.0%	-0.39 [-0.57, -0.20]
Nierenberg 2007	-7.22	6.62	274	-5.97	6.79	137	2.8%	-0.19 [-0.39, 0.02]
NKD20006 (NCT00048204)	-11.1	7.9	117	-10.9	7.8	118	2.4%	-0.03 [-0.28, 0.23]
PAR 01 001 (GSK & FDA)	-13.36	7.93	22	-11.33	7.93	21	0.8%	-0.25 [-0.85, 0.35]
Rapaport 2009	-12.11	8.02	173	-8.85	8	178	2.7%	-0.41 [-0.62, -0.19]
Reimherr 1990	-11.66	8.24	142	-8.16	7.85	141	2.5%	-0.43 [-0.67, -0.20]
SER 315 (FDA)	-8.9	4.52	76	-7.8	8	73	1.9%	-0.17 [-0.49, 0.15]
Stark 1985	-11	10.1	185	-8.2	9	169	2.7%	-0.29 [-0.50, -0.08]
Study 62b (FDA)	-8.82	8.71	297	-5.69	8.65	48	2.0%	-0.36 [-0.66, -0.05]
Study F1J-MC-HMAQ - Study Group B	-7.63	7	37	-7.1	6.96	72	1.5%	-0.08 [-0.47, 0.32]
Tollefson 1993/1995	-8.1	7.6	326	-6.4	7.1	329	3.2%	-0.23 [-0.38, -0.08]
VEN XR 367 (FDA)	-11.26	10.55	80	-13.1	10.63	81	2.0%	0.17 [-0.14, 0.48]
Wade 2002	-14.9	6.56658206	188	-12	6.78196137	189	2.8%	-0.43 [-0.64, -0.23]
WELL AK1A4006	-13.9	10.87	146	-12.2	9.73	148	2.6%	-0.16 [-0.39, 0.06]
Wernicke 1987	-8.83	8.67	297	-5.7	8.6	48	2.0%	-0.36 [-0.67, -0.05]
Wernicke 1988	-10.6	8.3	183	-7	8.6	77	2.3%	-0.43 [-0.70, -0.16]
Subtotal (95% CI)			**6916**			**4716**	**97.4%**	**-0.29 [-0.36, -0.23]**
Heterogeneity: Tau² = 0.02; Chi² = 105.28, df = 48 (P < 0.00001); I² = 54%								
Test for overall effect: Z = 9.46 (P < 0.00001)								
Total (95% CI)			**7027**			**4822**	**100.0%**	**-0.29 [-0.35, -0.23]**
Heterogeneity: Tau² = 0.02; Chi² = 107.84, df = 50 (P < 0.00001); I² = 54%								
Test for overall effect: Z = 9.44 (P < 0.00001)								
Test for subgroup differences: Chi² = 2.47, df = 1 (P = 0.12), I² = 59.4%								

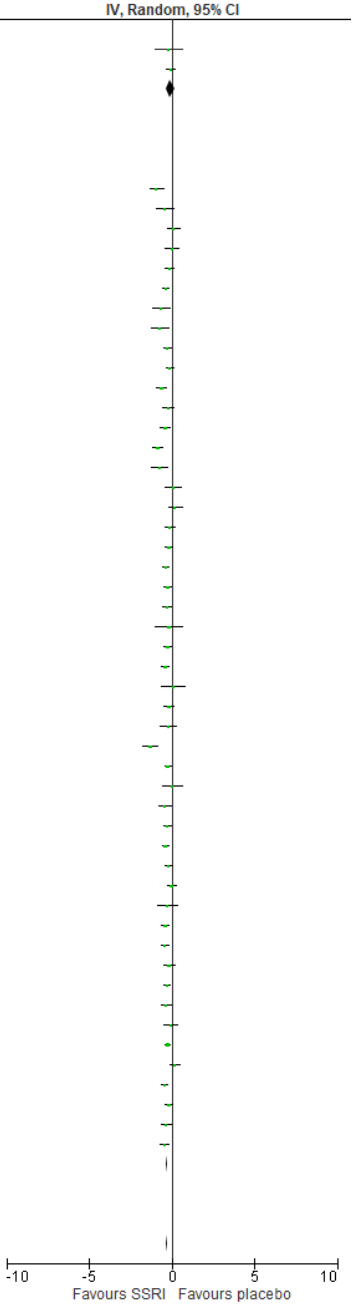
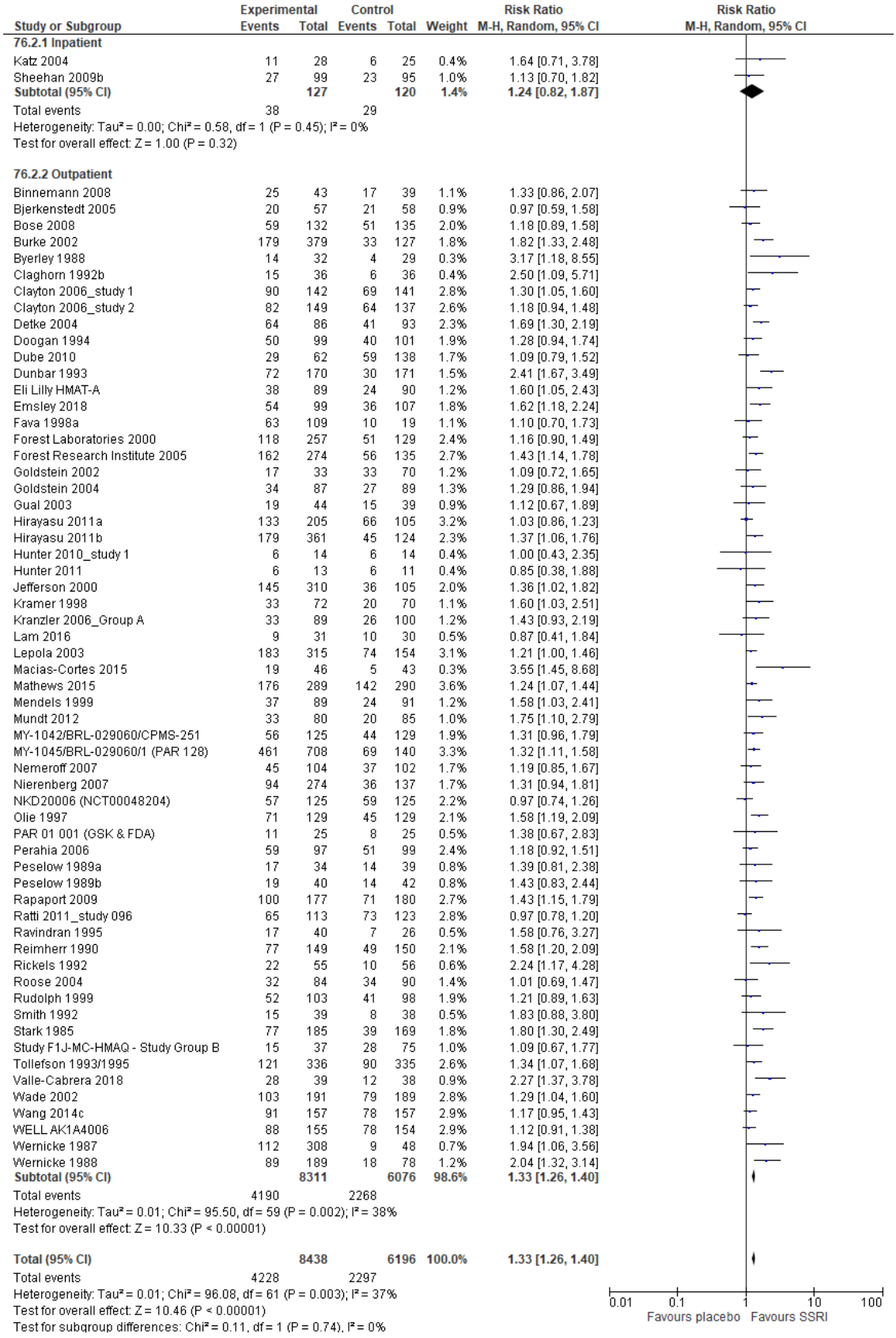


Figure 72: Response



Inpatient versus outpatient settings subgroup analysis for Comparison 3b. SSRIs versus tricyclic antidepressants (TCAs)

Critical outcomes

Figure 73: Depression symptomatology endpoint

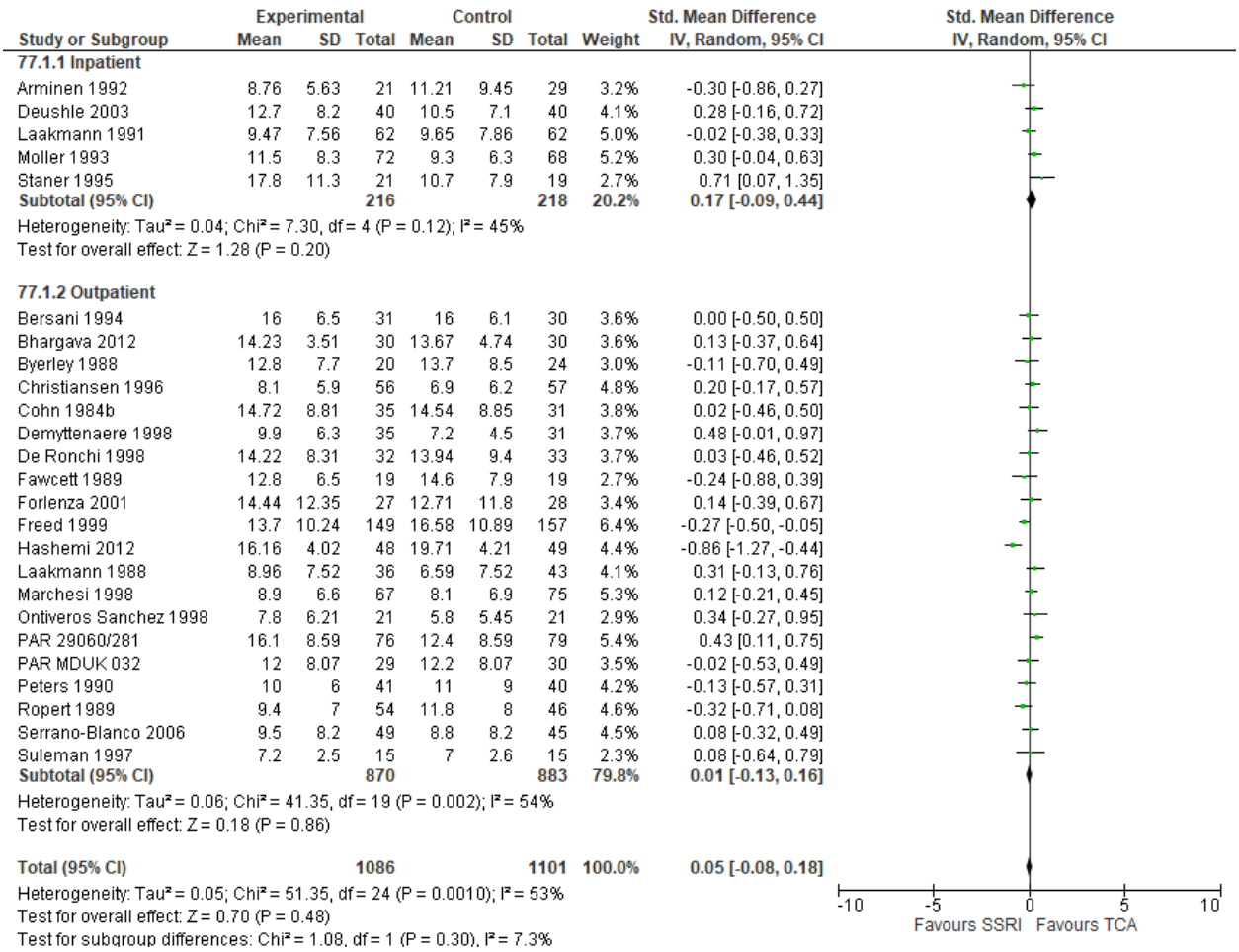


Figure 74: Depression symptomatology change score

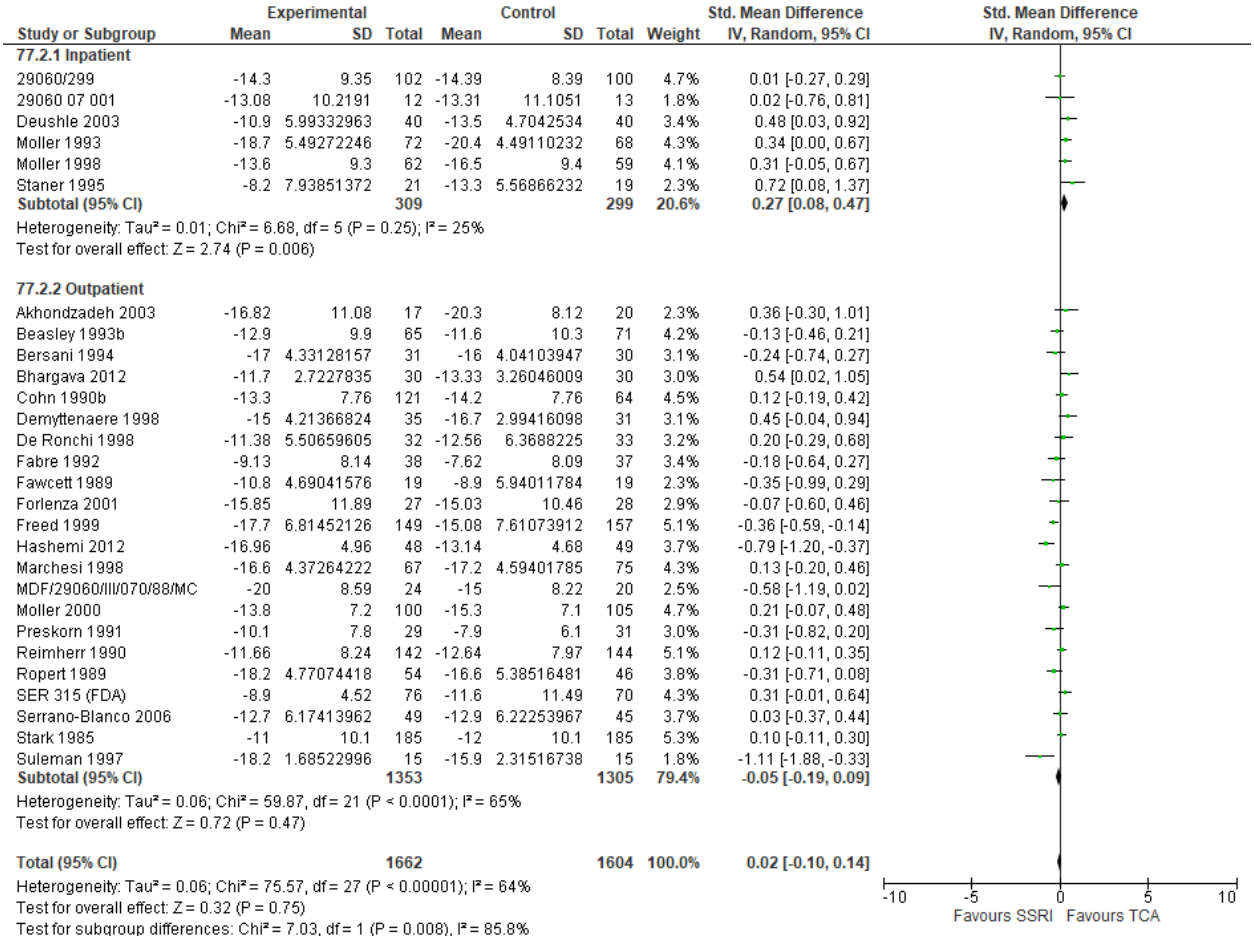


Figure 75: Remission

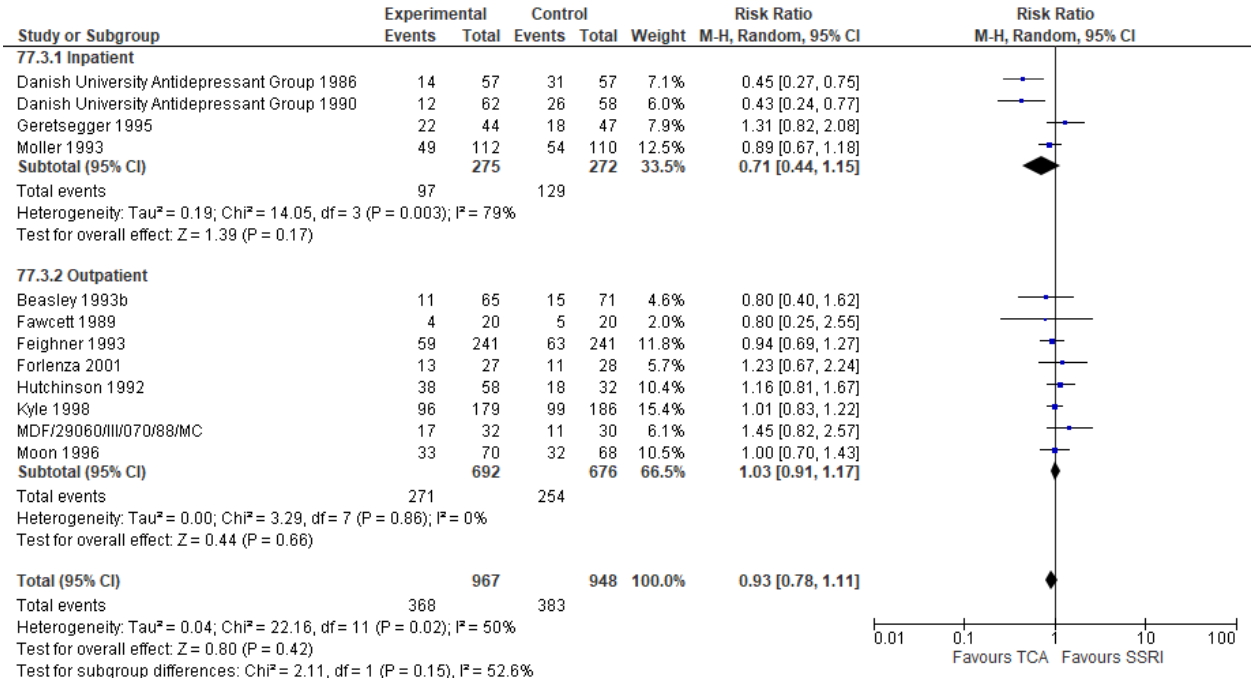
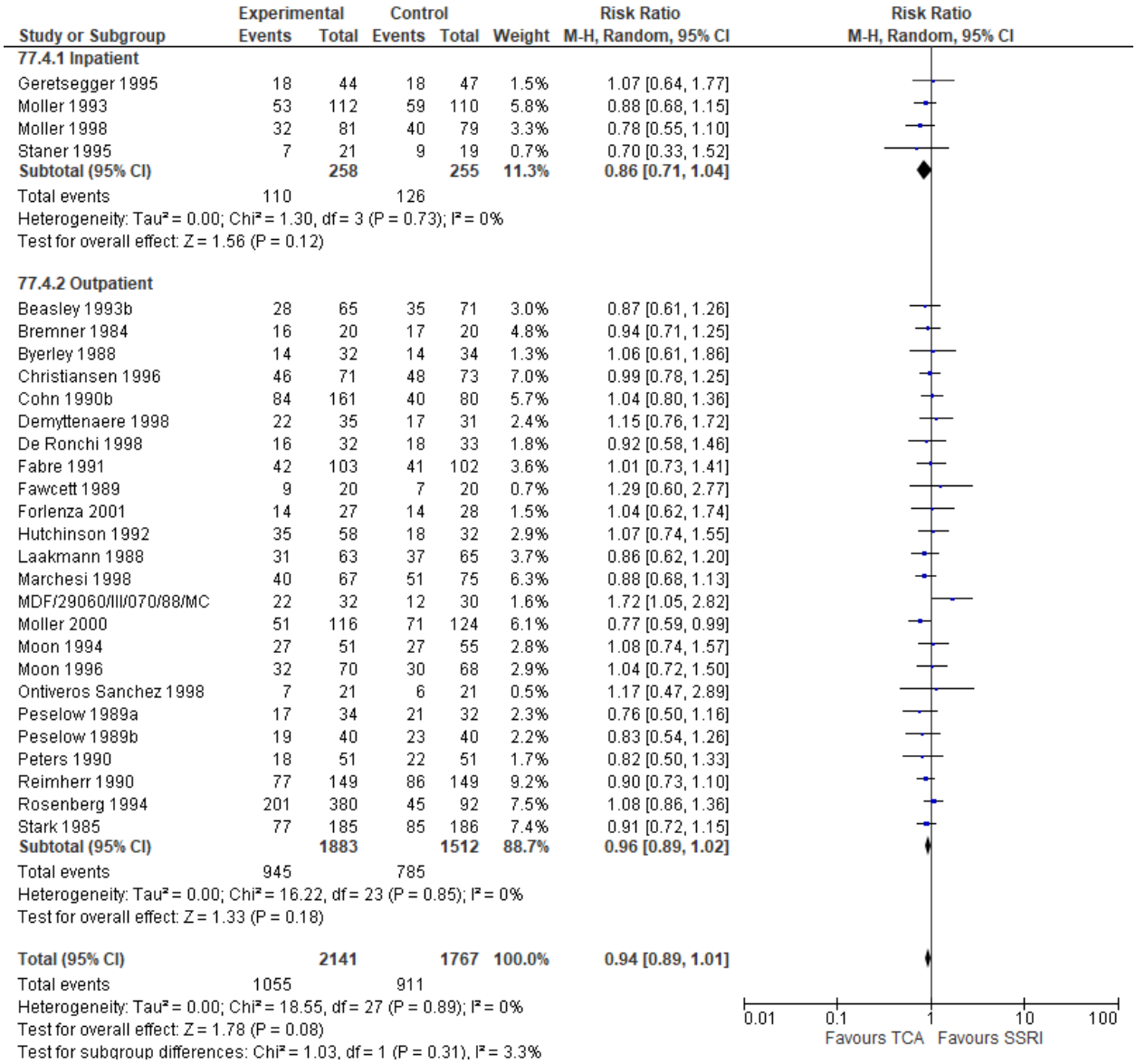


Figure 76: Response



Inpatient versus outpatient settings subgroup analysis for Comparison 3c. Serotonin–norepinephrine reuptake inhibitors (SNRIs) versus placebo

Critical outcomes

Figure 77: Depression symptomatology endpoint

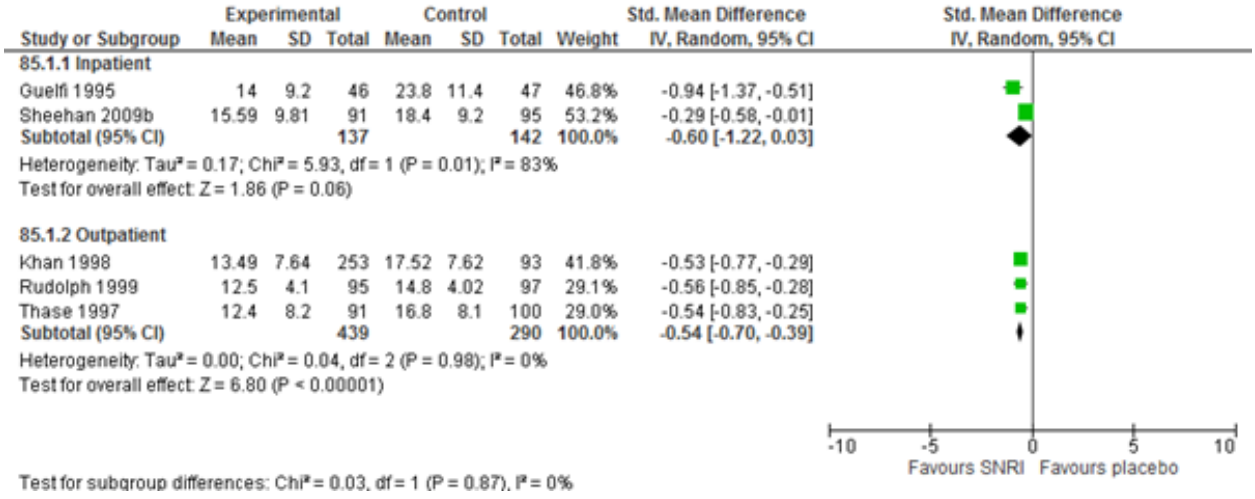


Figure 78: Depression symptomatology change score

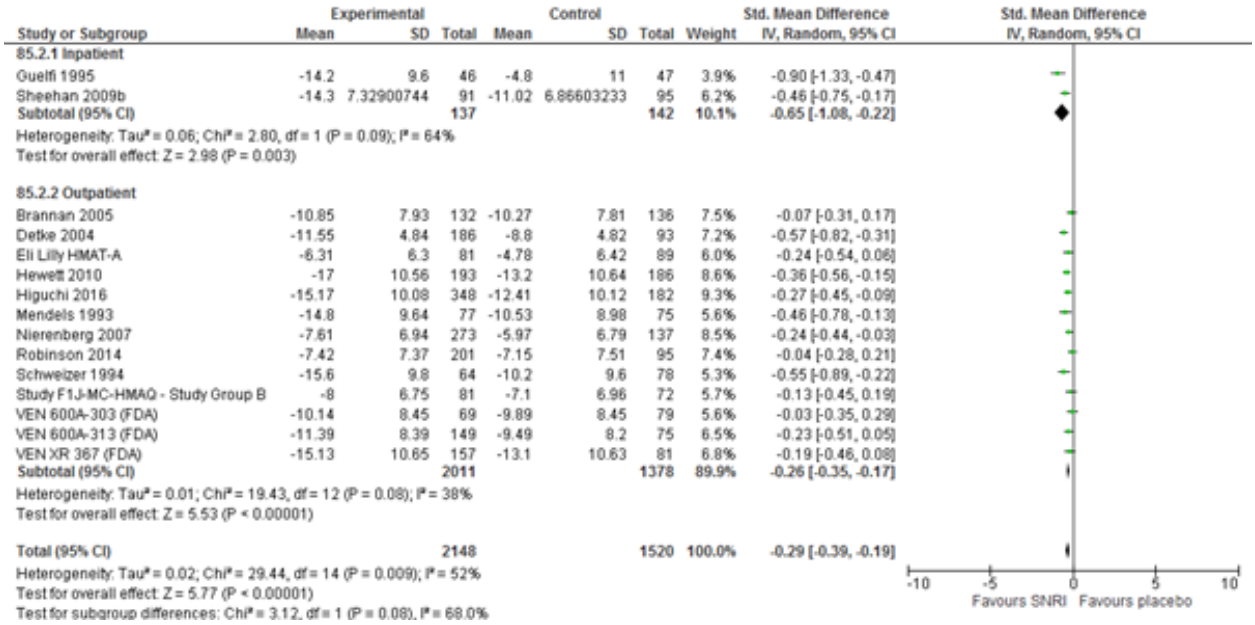
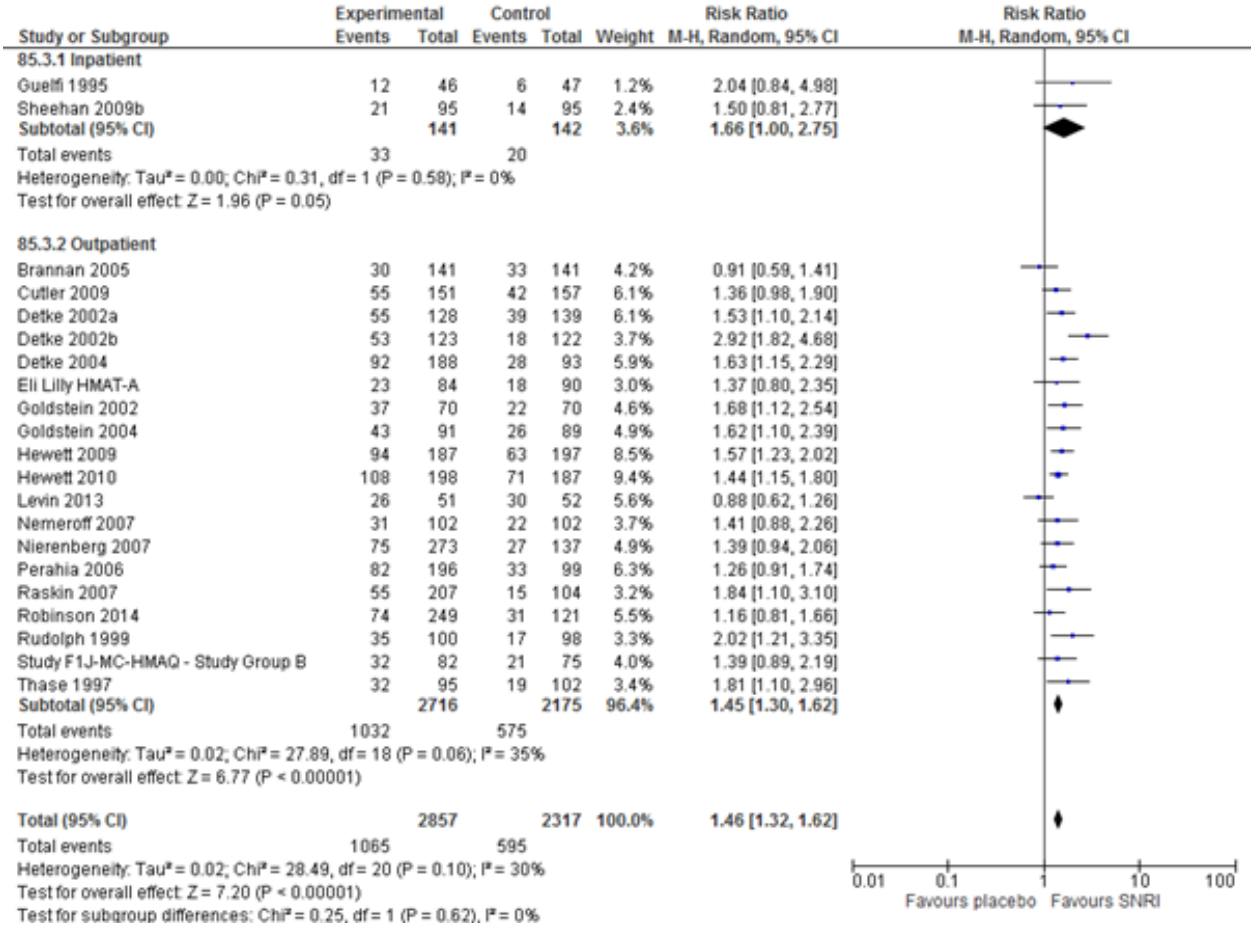


Figure 79: Remission



Inpatient versus outpatient settings subgroup analysis for Comparison 3d. SNRIs versus SSRIs

Critical outcomes

Figure 80: Depression symptomatology endpoint

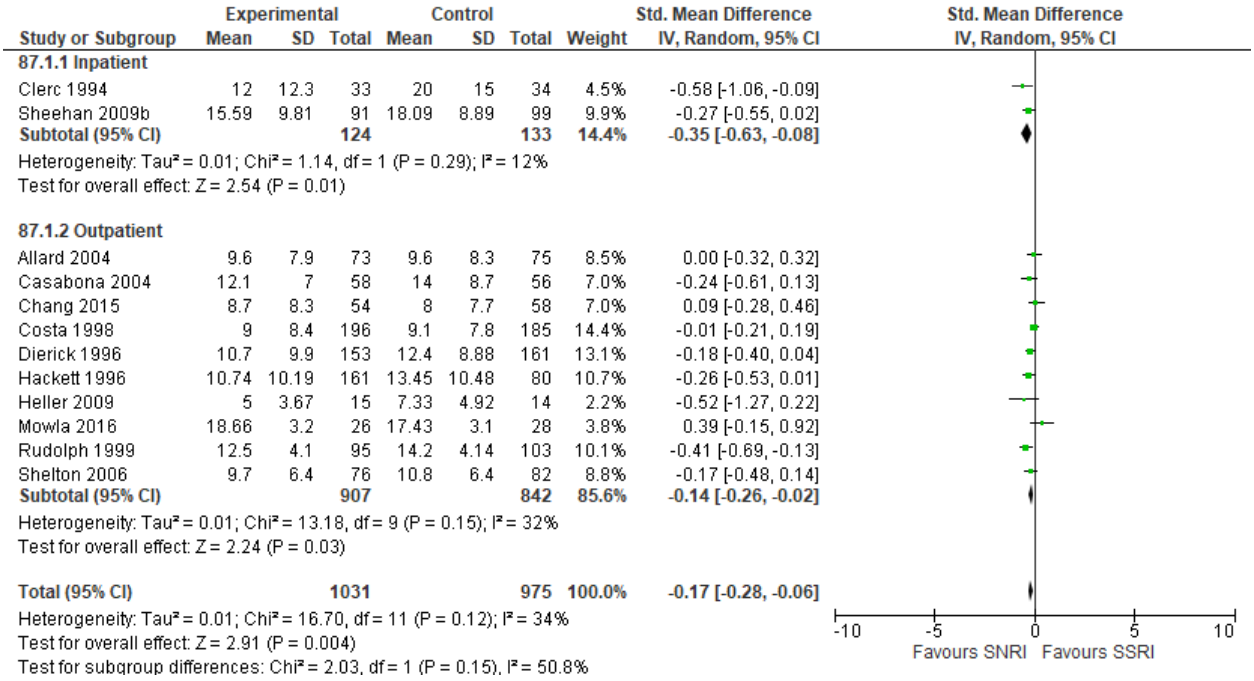


Figure 81: Depression symptomatology change score

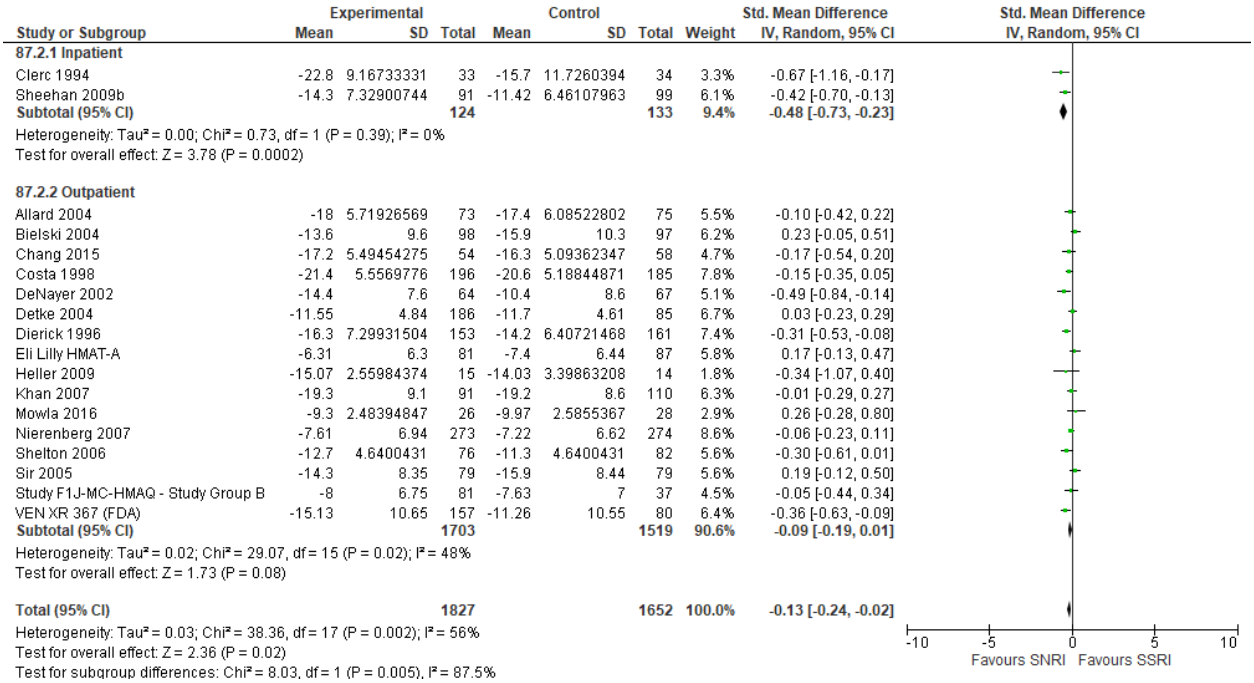


Figure 82: Remission

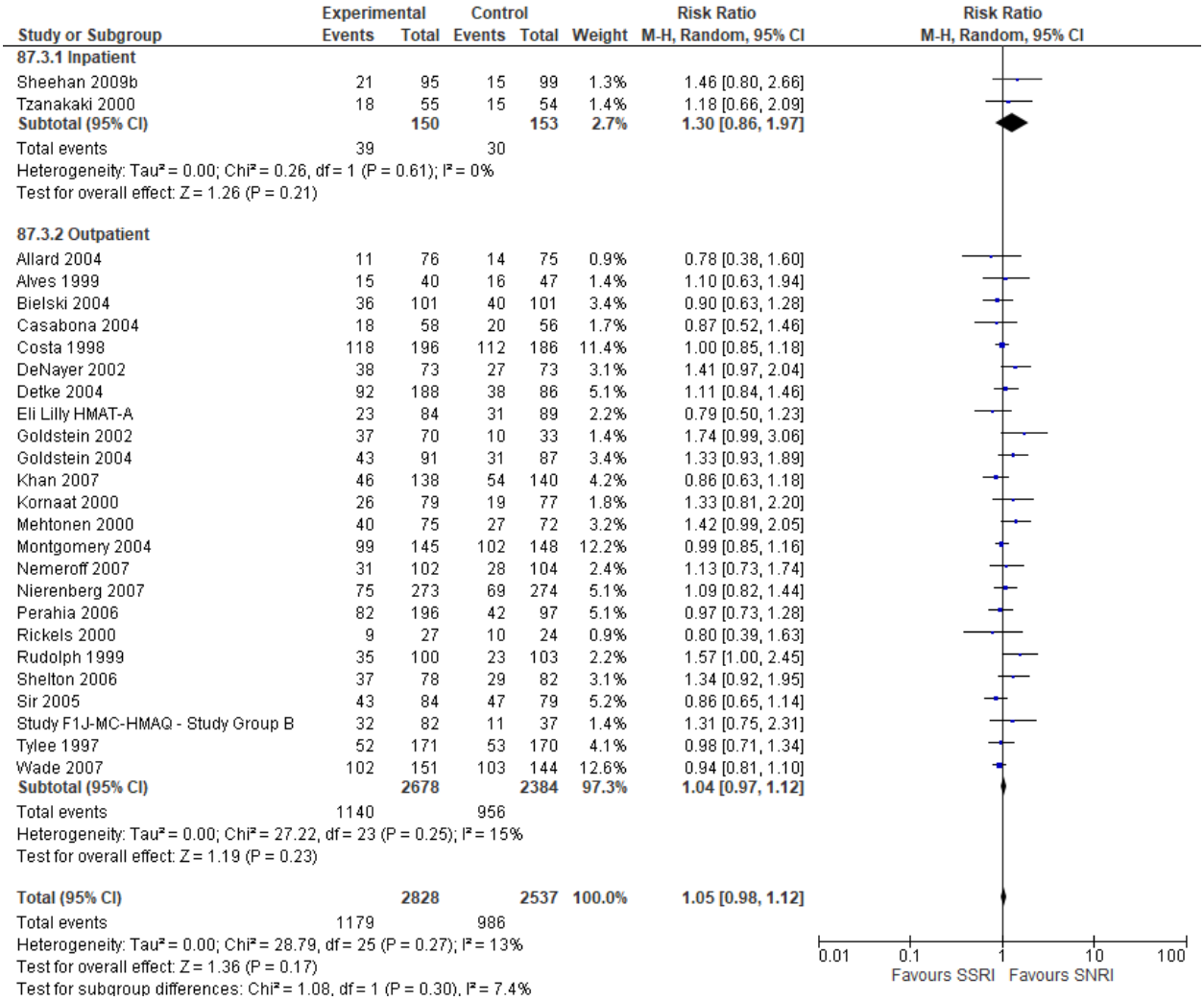
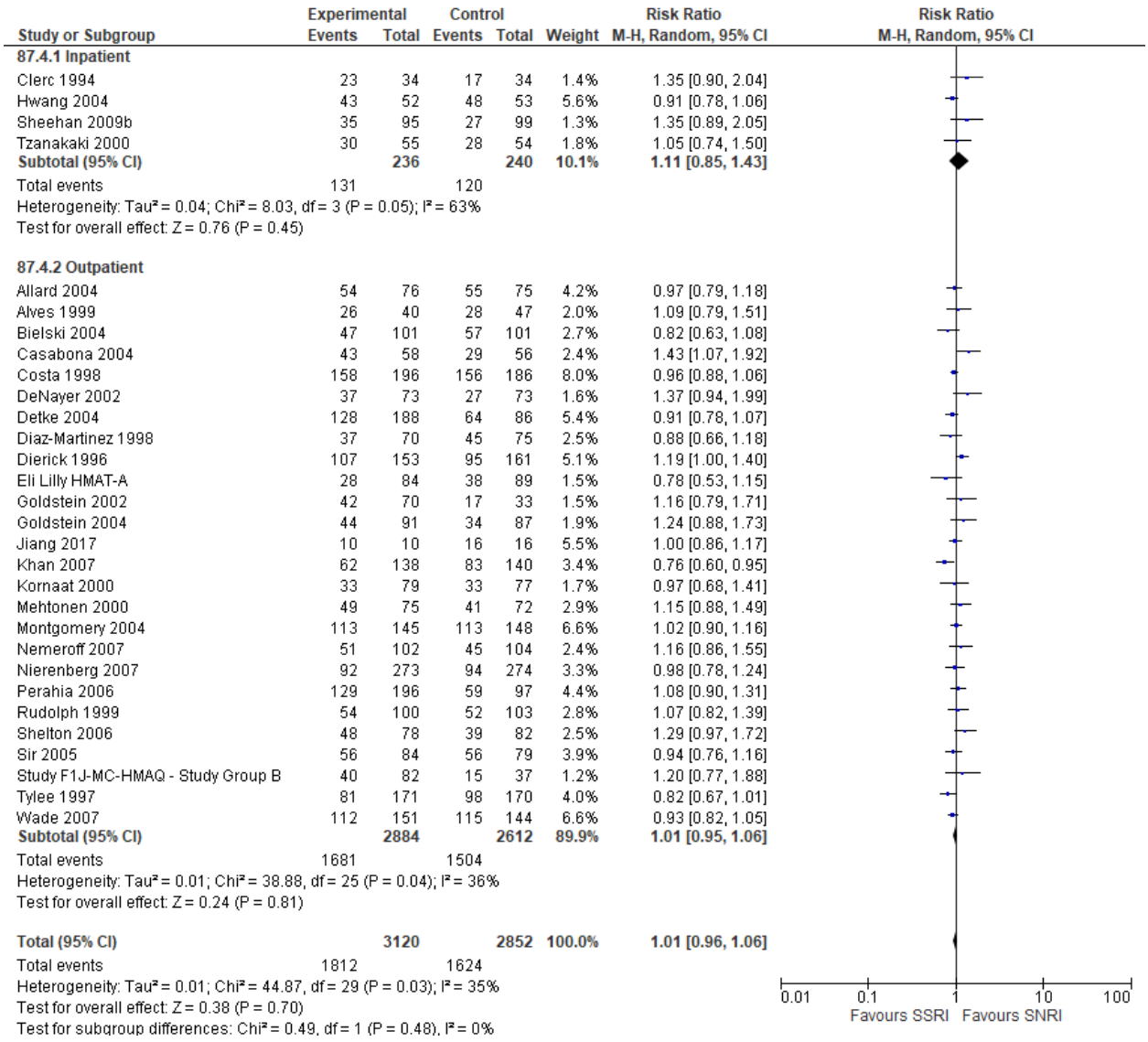


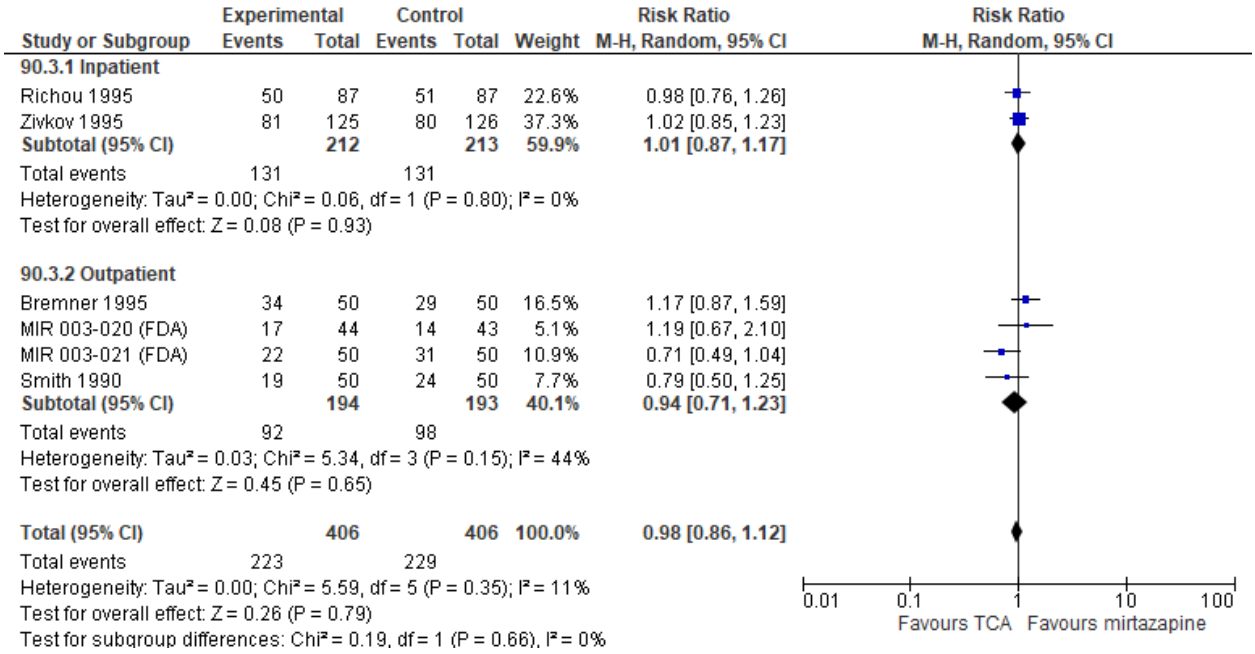
Figure 83: Response



Inpatient versus outpatient settings subgroup analysis for Comparison 3e. Mirtazapine versus TCAs

Critical outcomes

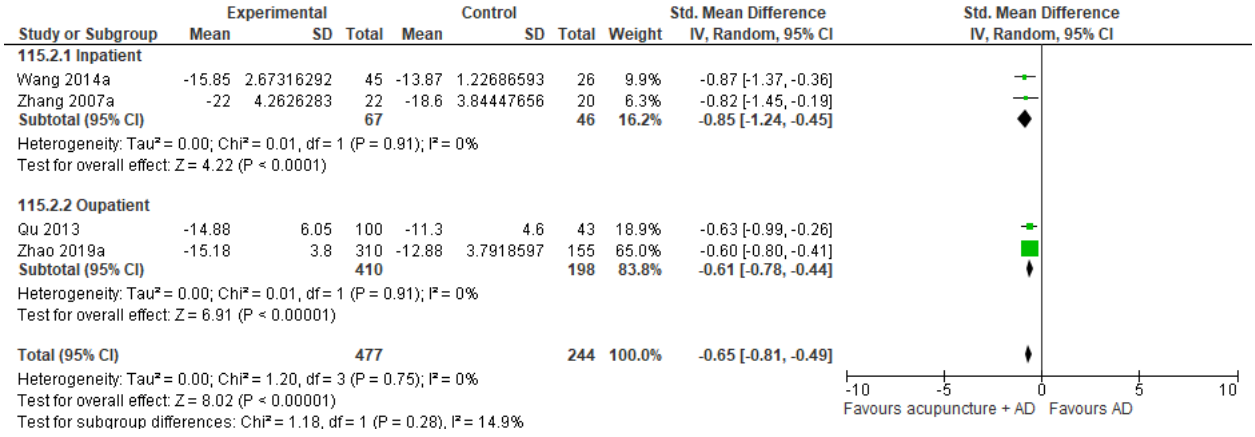
Figure 84: Response



Inpatient versus outpatient settings subgroup analysis for Comparison 3f. Acupuncture + antidepressants versus antidepressants

Critical outcomes

Figure 85: Depression symptomatology change score



Comparison 4. Acute psychiatric day hospital care versus inpatient care (for adults with depression and non-psychotic severe mental illness)

Critical outcomes

Figure 86: Psychiatric symptom severity at 2-3 months post-admission

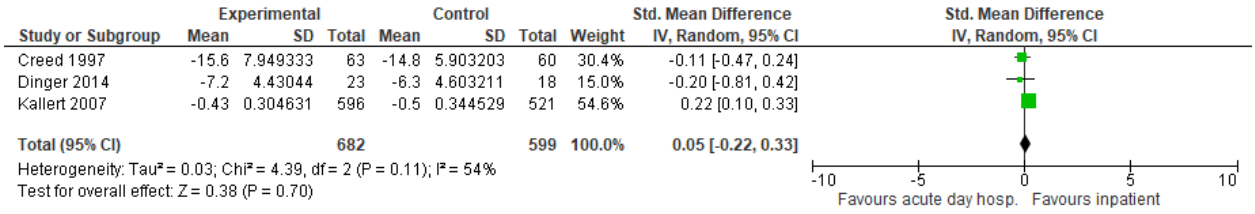


Figure 87: Psychiatric symptom severity at 12-14 months post-admission

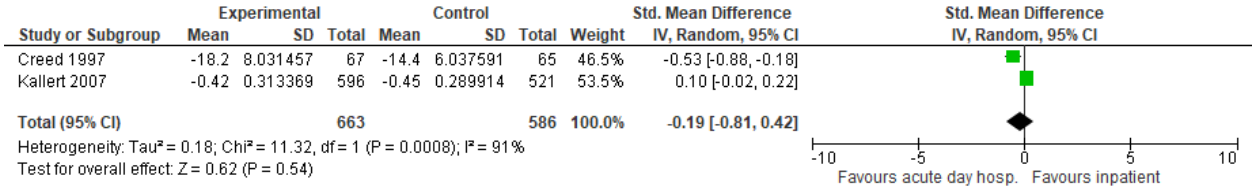


Figure 88: Remission (HAM-D<7/Present State Examination: Index of Definition ≤4)

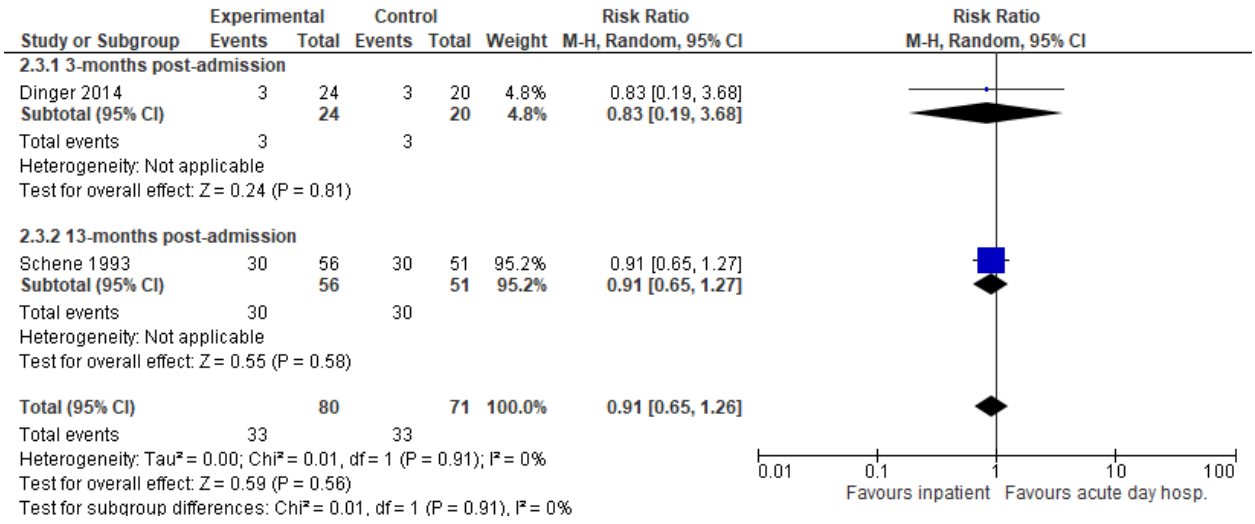
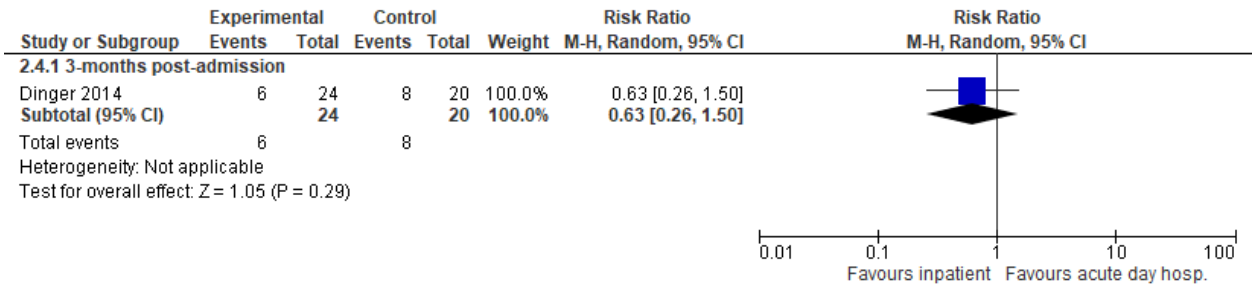


Figure 89: Response (at least 47% improvement on HAM-D)



Important outcomes

Figure 90: Duration of index admission

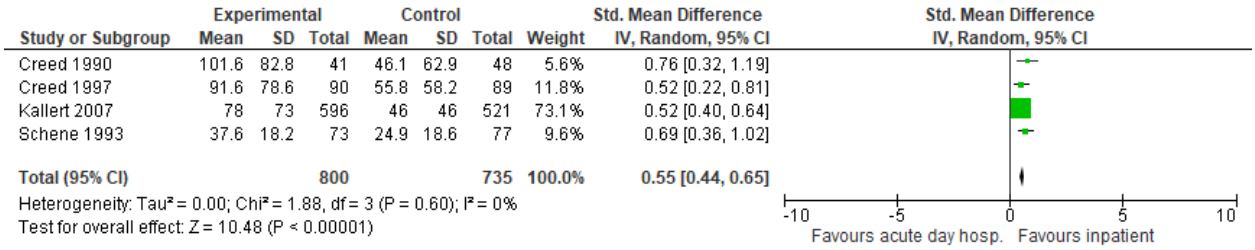


Figure 91: Readmission

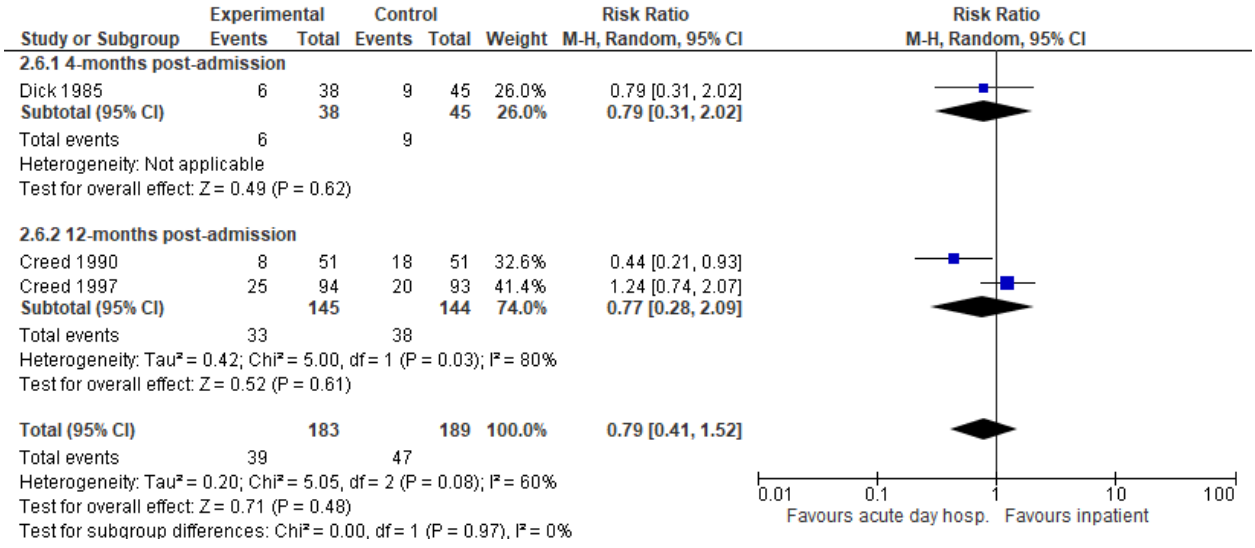


Figure 92: Service utilisation: Emergency contacts

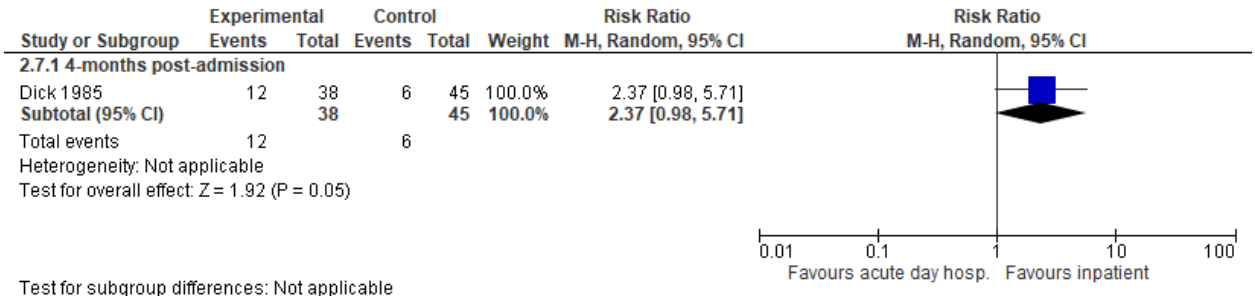


Figure 93: Service utilisation: Outpatient contact

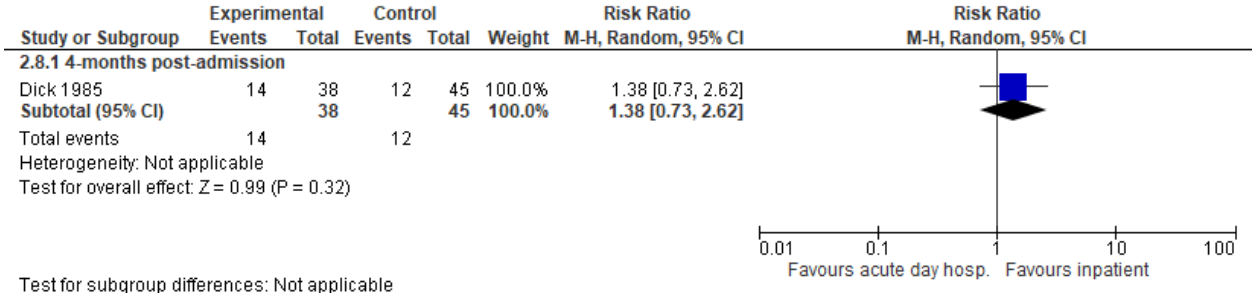


Figure 94: Quality of life (MANSA)

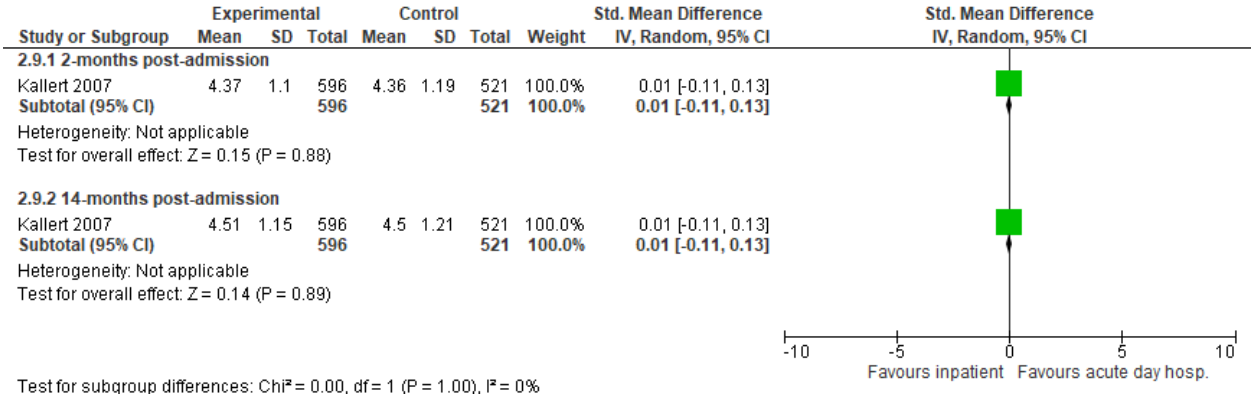


Figure 95: Social functioning impairment (GSDS-II)

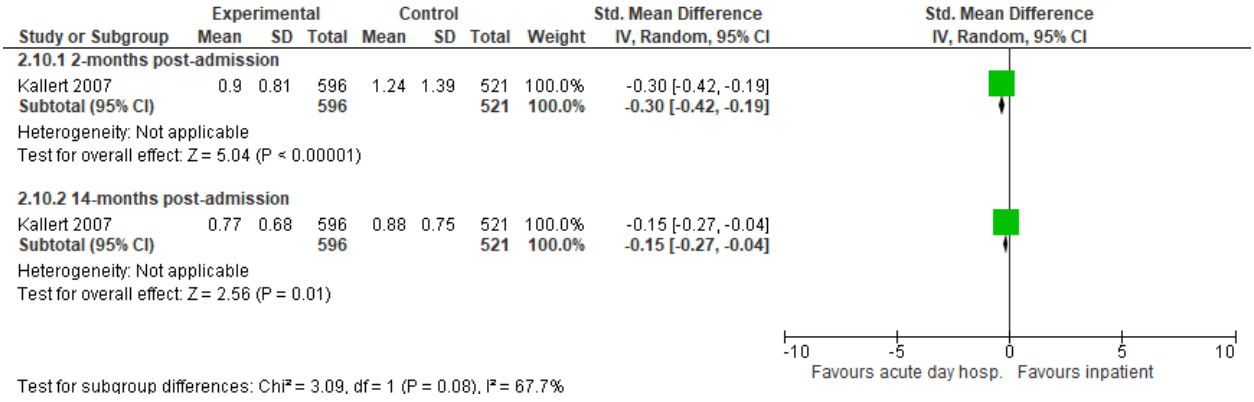


Figure 96: Social functioning response

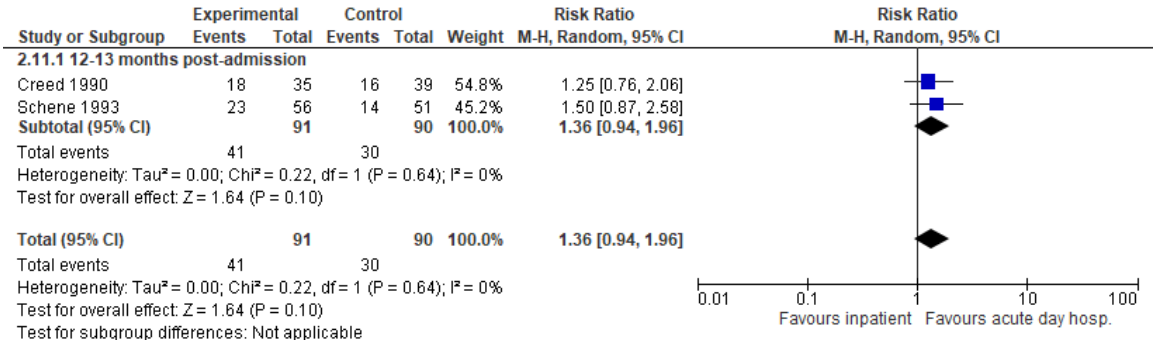


Figure 97: Satisfaction (number of participants satisfied or very satisfied with their treatment)

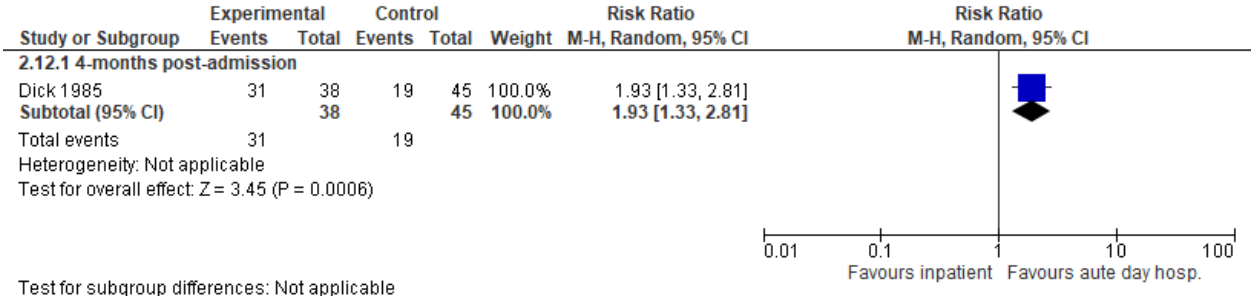


Figure 98: Satisfaction (CAT)

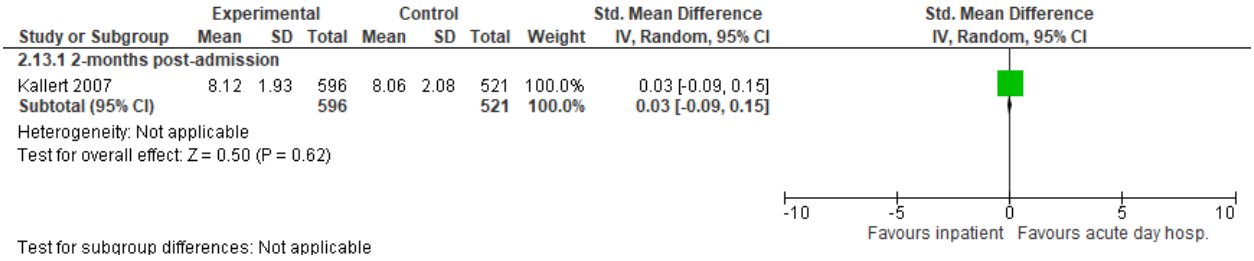
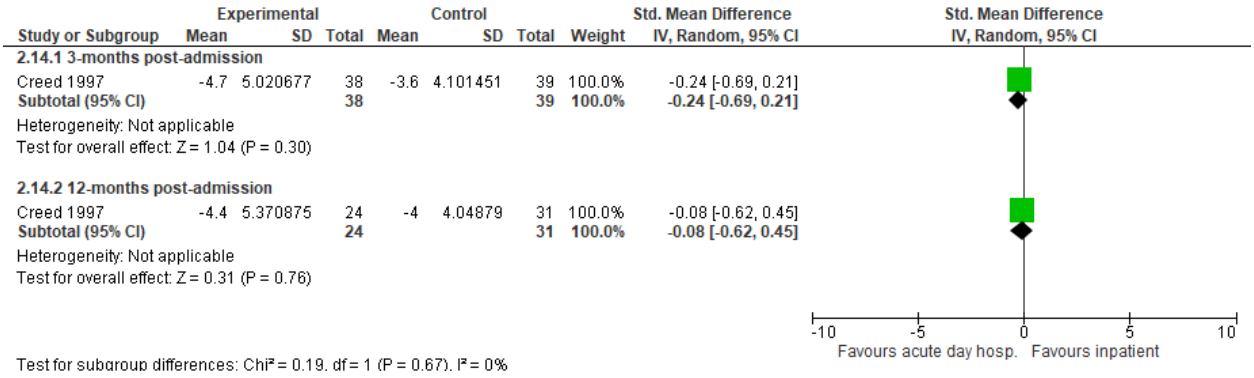


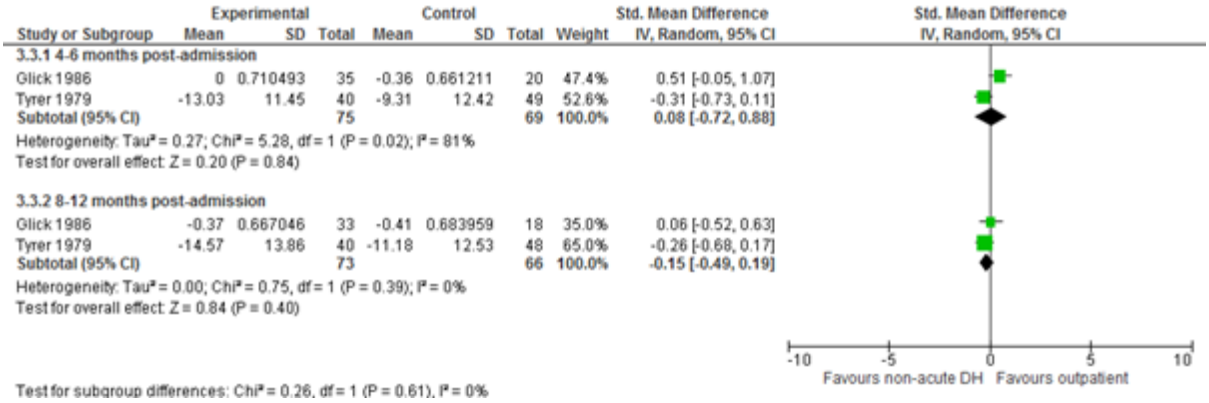
Figure 99: Carer distress (GHQ change score)



Comparison 5. Non-acute day hospital care versus outpatient care (for adults with depression and non-psychotic severe mental illness)

Critical outcomes

Figure 100: Psychiatric symptom severity (Psychiatric Evaluation Form/Present State Examination; change score)



Important outcomes

Figure 101: Service utilisation – admission as inpatient

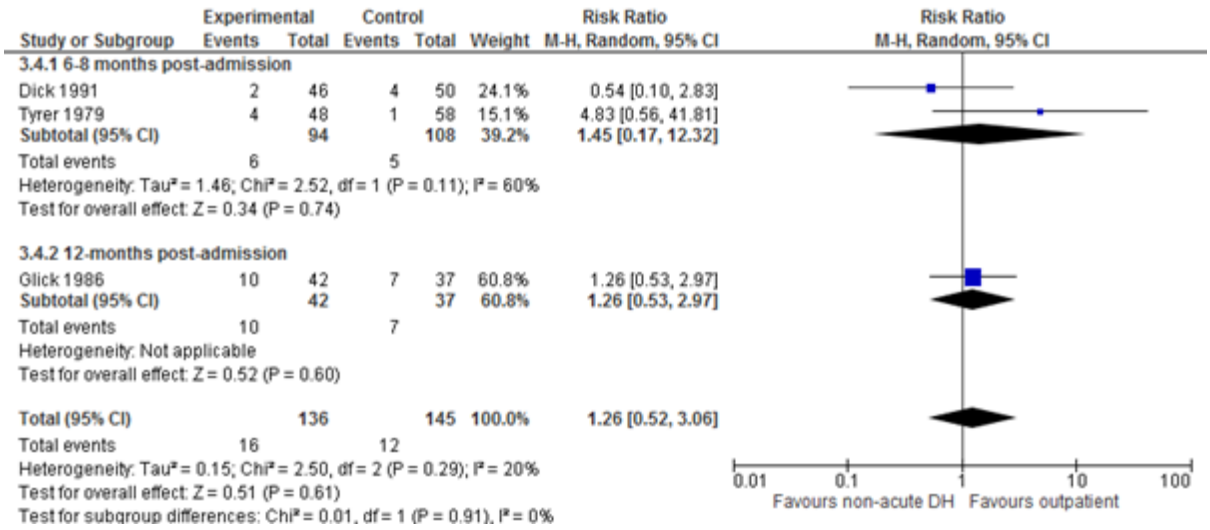


Figure 102: Social functioning (SAS-SR/SFS; change score)

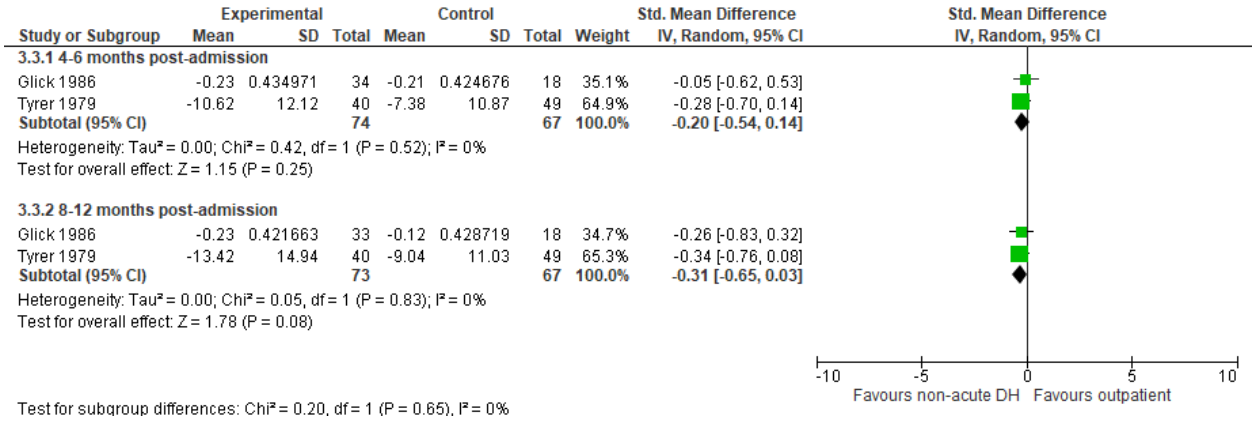


Figure 103: Global functioning (GAS; change score)

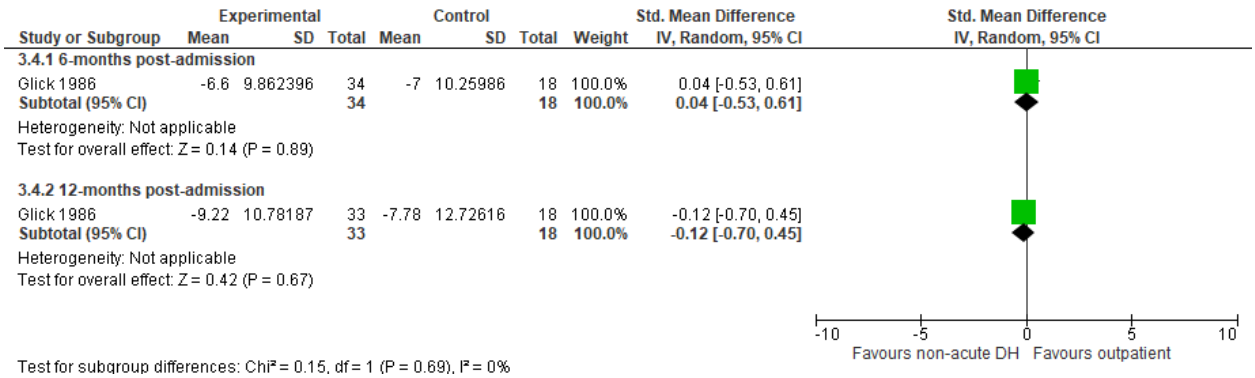
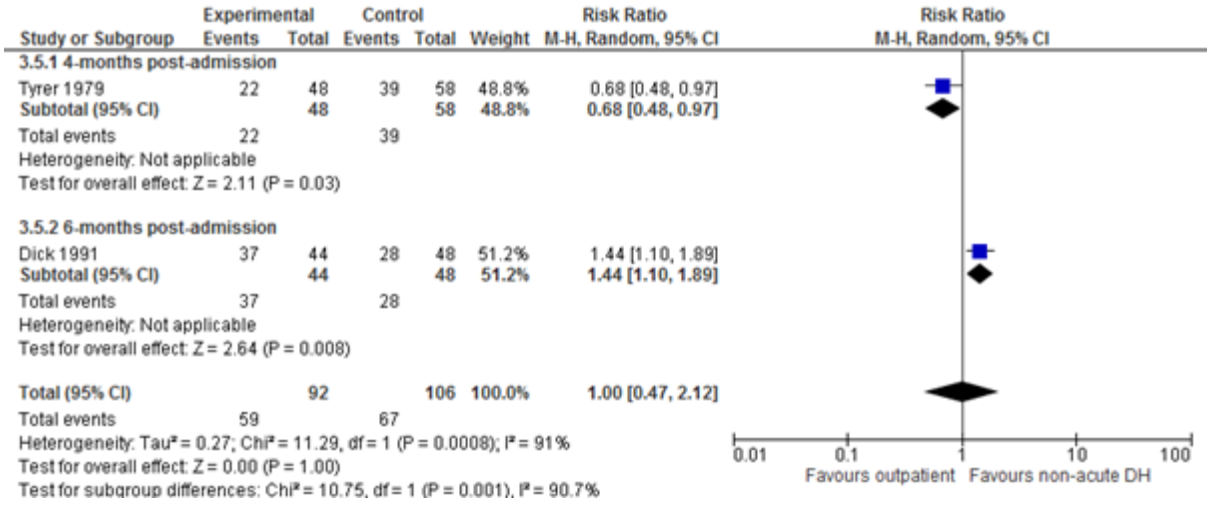


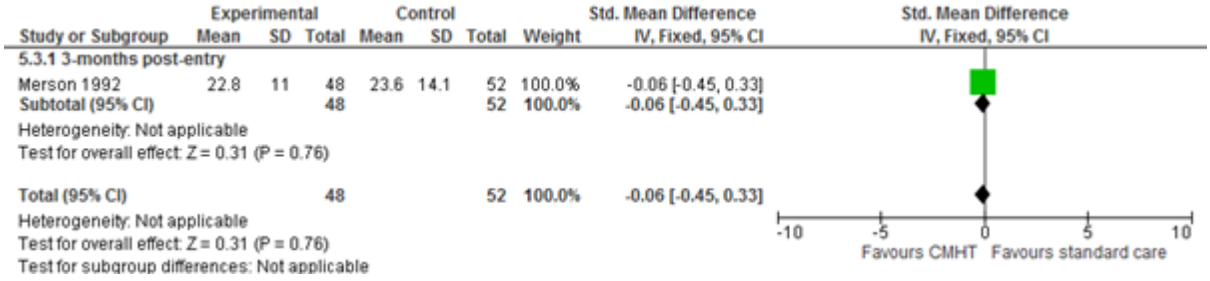
Figure 104: Satisfaction (number of participants satisfied or very satisfied with their treatment)



Comparison 6. Community mental health teams versus standard care (for adults with non-psychotic severe mental illness)

Critical outcomes

Figure 105: Psychiatric symptom severity (CPRS at endpoint)



Important outcomes

Figure 106: Service utilisation – admission as inpatient

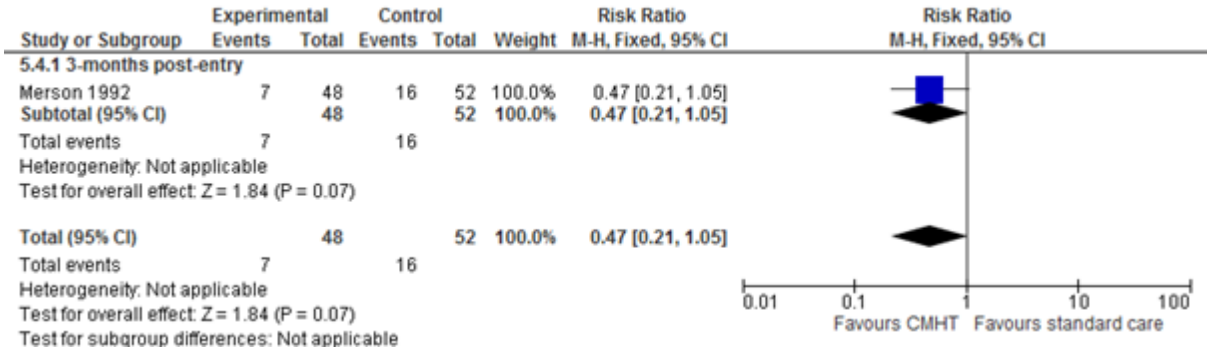


Figure 107: Service utilisation – admission as inpatient for >10 days

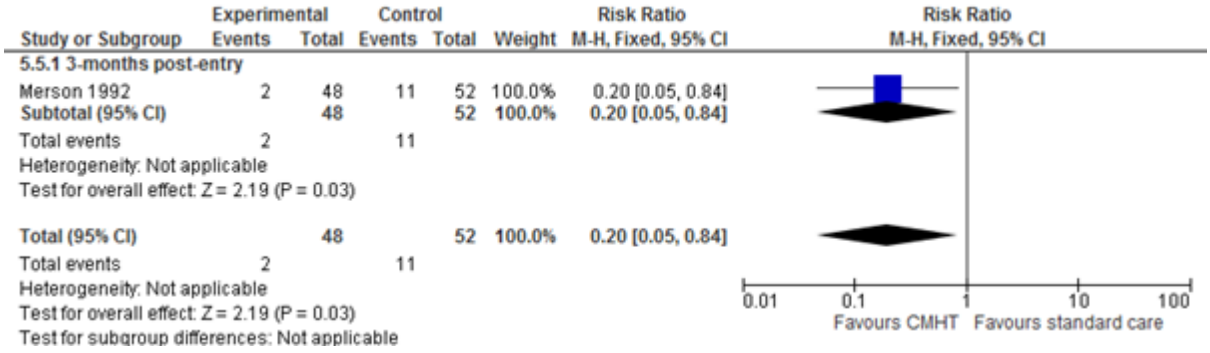


Figure 108: Satisfaction – number of participants satisfied with their treatment

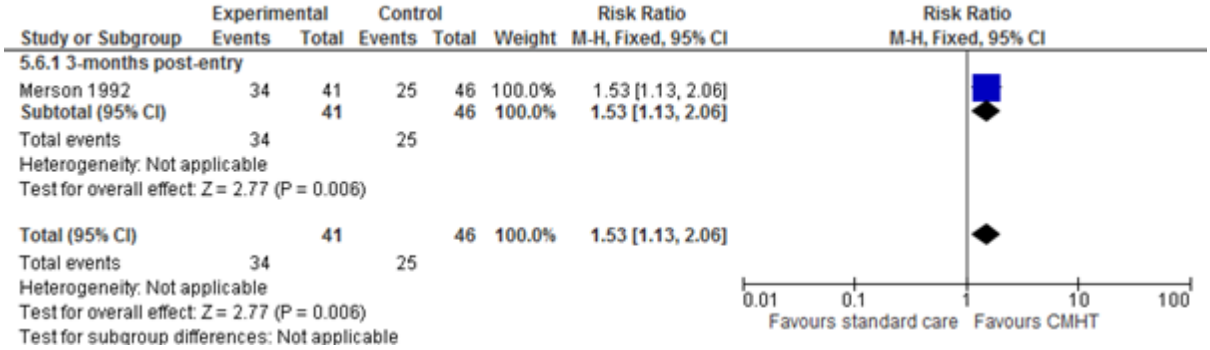
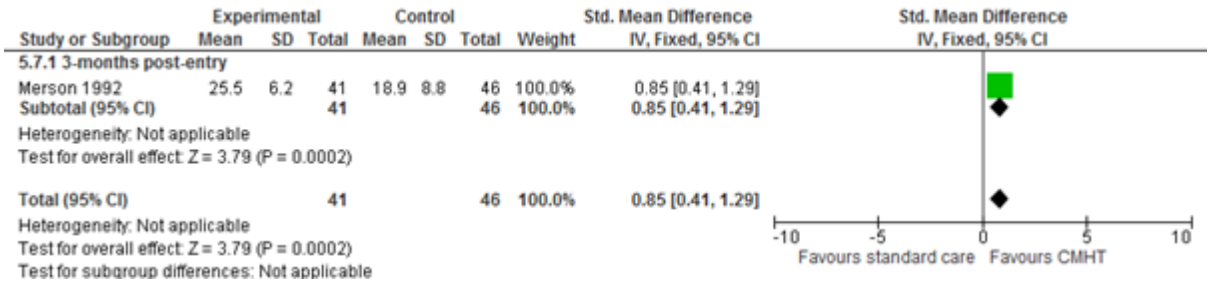


Figure 109: Satisfaction – service satisfaction score



Appendix F – GRADE tables

GRADE tables for review question 1.1 For adults with depression, what are the relative benefits and harms associated with different models for the coordination and delivery of services?

GRADE tables not provided for subgroup analyses.

Table 29: Clinical evidence profile for Comparison 1: Collaborative care (simple or complex) versus standard care/enhanced standard care.

Quality assessment							Number of participants		Effect			
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Collaborative care	Standard care/enhanced standard care	Relative (95% CI)	Absolute (95% CI)	Quality	Importance
Depression symptomatology at 6 months (assessed with: Hamilton Depression Rating Scale (HAMD)/Patient Health Questionnaire (PHQ-9)/Beck Depression Inventory-II (BDI-II))												
9 (Aragones 2012; Buszewicz 2016; Chen 2015; Curth 2020; Harter 2018; Huang 2018; Landis 2007; Ng 2020; Oladeji 2015)	randomised trials	serious[1]	very serious[2]	not serious	serious[3]	none	1781	1010	-	SMD 0.4 lower (0.71 lower to 0.09)	VERY LOW	CRITICAL
Depression symptomatology at 12 months (assessed with: Hamilton Depression Rating Scale (HAMD)/Patient Health Questionnaire (PHQ-9)/Beck Depression Inventory (BDI/BDI-II))												
13 (Aragones 2012; Bosan	randomised trials	serious[1]	very serious[2]	not serious	serious[3]	none	2957	2451	-	SMD 0.35 lower (0.53 lower to	VERY LOW	CRITICAL

Quality assessment							Number of participants		Effect			
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Collaborative care	Standard care/enhanced standard care	Relative (95% CI)	Absolute (95% CI)	Quality	Importance
quet 2017; Bruce 2004; Buszewicz 2016; Chen 2015; Gensichen 2009; Gilbody 2017/ Lewis 2017; Harter 2018; Holzel 2018; Morriss 2016; Ng 2020; Richards 2013/ 2016; Swindle 2003)										0.16 lower)		
Response at 6 months (assessed with: Number of participants whose scores improved by at least 50% on Hamilton Depression Rating Scale (HAMD)/Patient Health Questionnaire (PHQ-9))												
8 (Aragones 2012; Araya 2003; Berghofer 2012; Chen	randomised trials	serious[1]	serious[4]	not serious	not serious	none	411/885 (46.4%)	198/818 (24.2%)	RR 1.85 (1.34 to 2.56)	206 more per 1,000 (from 82 more to 378 more)	LOW	CRITICAL

Quality assessment							Number of participants		Effect			
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Collaborative care**	**Standard care/enhanced standard care**	**Relative (95% CI)**	**Absolute (95% CI)**	**Quality**	**Importance**
2015; Huijbregts 2013; Ng 2020; Yeung 2010; Yeung 2016)												
Response at 12 months (assessed with: Number of participants whose scores improved by at least 50% on Hamilton Depression Rating Scale (HAMD)/Patient Health Questionnaire (PHQ-9))												
13 (Aragones 2012; Berghofer 2012; Bruce 2004; Chen 2015; Ell 2007; Gensichen 2009; Harter 2018; Holzel 2018; Huijbregts 2013; Katzelnick 2000; Morriss 2016; Ng 2020; Richards	randomised trials	serious[1]	serious[4]	not serious	not serious	none	984/2744 (35.9%)	535/2166 (24.7%)	RR 1.51 (1.30 to 1.76)	126 more per 1,000 (from 74 more to 188 more)	LOW	CRITICAL

Quality assessment							Number of participants		Effect			
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Collaborative care	Standard care/enhanced standard care	Relative (95% CI)	Absolute (95% CI)	Quality	Importance
2013/ 2016)												
Remission at 6 months (assessed with: Number of participants showing Hamilton Depression Rating Scale (HAMD) score <7 or 8/Patient Health Questionnaire (PHQ-9) score <5/Montgomery-Asberg Depression Rating Scale-self report (MADRS-SR) score <13/Center for Epidemiologic Studies Depression (CES-D) score <20/loss of diagnosis)												
12 (Aragones 2012; Araya 2003; Bjorkelund 2018; Chen 2015; Huijbregts 2013; Jeong 2013; Katon 1999; Ng 2020; Smit 2006; Wells 2000; Yeung 2010; Yeung 2016	randomised trials	serious[1]	serious[4]	not serious	not serious	none	940/2313 (40.6%)	439/1620 (27.1%)	RR 1.63 (1.31 to 2.02)	171 more per 1,000 (from 84 more to 276 more)	LOW	CRITICAL
Remission at 12 months (assessed with: Number of participants showing Hamilton Depression Rating Scale (HAMD) score <7/Patient Health Questionnaire (PHQ-9) score <5 or 10/Center for Epidemiologic Studies Depression (CES-D) score <20/loss of diagnosis)												
14 (Aragones 2012; Bruce 2004; Chen 2015; Ell 2007;	randomised trials	serious[1]	serious[4]	not serious	serious[3]	none	1119/3664 (30.5%)	581/2591 (22.4%)	RR 1.49 (1.23 to 1.8)	110 more per 1,000 (from 52 more to 179 more)	VERY LOW	CRITICAL

Quality assessment							Number of participants		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Collaborative care	Standard care/enhanced standard care	Relative (95% CI)	Absolute (95% CI)		
Gensichen 2009; Harter 2018; Holzel 2018; Huijbregts 2013; Katzelnick 2000; Ludman 2007; Morriss 2016; Ng 2020; Richards 2013/ 2016; Wells 2000												
Antidepressant use at 6 months (assessed with: Number of participants adhering to or in receipt of antidepressants)												
11 (Aragones 2012; Araya 2003; Bjorkelund 2018; Finley 2003; Jeong 2013; Katon 1999; Simon 2004	randomised trials	serious[1]	very serious[2]	not serious	very serious[5]	none	1432/2204 (65.0%)	1007/1818 (55.4%)	RR 1.14 (0.91 to 1.43)	78 more per 1,000 (from 50 fewer to 238 more)	VERY LOW	IMPORTANT

Quality assessment							Number of participants		Effect			
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Collaborative care	Standard care/enhanced standard care	Relative (95% CI)	Absolute (95% CI)	Quality	Importance
(CM); Simon 2004 (CM + psych); Simon 2006; Smit 2006; Unutzer 2002/ Arean 2005)												
Antidepressant use at 12 months (assessed with: Number of participants adhering to or in receipt of antidepressants)												
13 (Aragones 2012; Bosanquet 2017; Bruce 2004; Capoccia 2004; Dobscha 2006; Ell 2007; Fortney 2007; Gensichen 2009; Gilbody 2017/ Lewis 2017; Jarjou	randomised trials	serious[1]	serious[4]	not serious	serious[3]	none	1679/2823 (59.5%)	1433/2843 (50.4%)	RR 1.14 (1.04 to 1.26)	71 more per 1,000 (from 20 more to 131 more)	VERY LOW	IMPORTANT

Quality assessment							Number of participants		Effect			
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Collaborative care	Standard care/enhanced standard care	Relative (95% CI)	Absolute (95% CI)	Quality	Importance
ra 2004; Ludman 2007; Richards 2013/ 2016 Unutzer 2002/ Arean 2005)												
Discontinuation at 6 months (assessed with: Number of participants who dropped out of the study for any reason)												
19 (Aragones 2012; Araya 2003; Bjorkelund 2018; Buszewicz 2016; Chen 2015; Curth 2020; Finley 2003; Harter 2018; Huang 2018; Huijbregts 2013; Jeong 2013; Ng 2020;	randomised trials	not serious	serious[4]	not serious	serious[3]	none	952/5008 (19%)	576/3297 (17.5%)	RR 0.94 (0.77 to 1.15)	10 fewer per 1,000 (from 40 fewer to 26 more)	LOW	IMPORTANT

Quality assessment							Number of participants		Effect			
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Collaborative care	Standard care/enhanced standard care	Relative (95% CI)	Absolute (95% CI)	Quality	Importance
Oladeji 2015; Simon 2004 (CM); Simon 2004 (CM + psych); Simon 2006; Smit 2006; Unutzer 2002/ Arean 2005; Wells 2000)												
Discontinuation at 12 months (assessed with: Number of participants who dropped out of the study for any reason)												
22 (Aragones 2012; Bosanquet 2017; Bruce 2004; Buszewicz 2016; Capoccia 2004; Chen 2015; Dobscha 2006; Ell 2007; Fortne	randomised trials	not serious	serious[4]	not serious	not serious	none	1381/5986 (23.1%)	1015/4930 (20.6%)	RR 1.06 (0.93 to 1.2)	12 more per 1,000 (from 14 fewer to 41 more)	MODERATE	IMPORTANT

Quality assessment							Number of participants		Effect			
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Collaborative care	Standard care/enhanced standard care	Relative (95% CI)	Absolute (95% CI)	Quality	Importance
y 2007; Gensichen 2009; Gilbody 2017/ Lewis 2017; Harter 2018; Holzel 2018; Huijbregts 2013; Katzelnick 2000; Ludman 2007; Morriss 2016; Ng 2020; Richards 2013/ 2016; Swindle 2003; Unutzer 2002/ Arean 2005; Wells 2000)												

CI: Confidence interval; SMD: Standardised mean difference; RR: Risk ratio
1. Risk of bias is high or unclear across multiple domains
2. I-squared>80%
3. 95% CI crosses 1 clinical decision threshold
4. I-squared>50%
5. 95% CI crosses 2 clinical decision thresholds

Table 30: Clinical evidence profile for Comparison 2: Collaborative care for relapse prevention versus standard care

Quality assessment							Number of participants		Effect			
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Collaborative care**	**Standard care**	**Relative (95% CI)**	**Absolute (95% CI)**	**Quality**	**Importance**
Relapse at 12 months (assessed with: Longitudinal Interval Follow-up Evaluation)												
1 (Katon 2001)	randomised trials	serious[1]	not serious	not serious	very serious[2]	none	68/194 (35.1%)	66/192 (34.4%)	RR 1.02 (0.78 to 1.34)	7 more per 1,000 (from 76 fewer to 117 more)	VERY LOW	CRITICAL
Antidepressant use at 6 months (assessed with: Number of participants receiving antidepressants)												
1 (Katon 2001)	randomised trials	serious[1]	not serious	not serious	serious[3]	none	139/194 (71.6%)	112/192 (58.3%)	RR 1.23 (1.06 to 1.43)	134 more per 1,000 (from 35 more to 251 more)	LOW	IMPORTANT
Antidepressant use at 12 months (assessed with: Number of participants receiving antidepressants)												
1 (Katon 2001)	randomised trials	serious[1]	not serious	not serious	serious[3]	none	123/194 (63.4%)	95/192 (49.5%)	RR 1.28 (1.07 to 1.53)	139 more per 1,000 (from 35 more to 262 more)	LOW	CRITICAL
Discontinuation at 12 months (assessed with: Number of participants who dropped out of the study for any reason)												
1 (Katon 2001)	randomised trials	serious[1]	not serious	not serious	serious[3]	none	20/194 (10.3%)	40/192 (20.8%)	RR 0.49 (0.30 to 0.81)	106 fewer per 1,000 (from 40 fewer to 146 fewer)	LOW	IMPORTANT

CI: Confidence interval; RR: Risk ratio
1. Risk of bias is high or unclear across multiple domains
2. 95% CI crosses 2 clinical decision thresholds
3. 95% CI crosses 1 clinical decision threshold

Table 31: Clinical evidence profile for Comparison 3. Stepped care versus standard care/enhanced standard care

Quality assessment							Number of participants		Effect			
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Stepped care	Standard care/enhanced standard care	Relative (95% CI)	Absolute (95% CI)	Quality	Importance
Depression symptomatology (endpoint score) at 6 months (assessed with: Patient Health Questionnaire (PHQ-9))												
2 (Gureje 2019; Knapstad 2020)	randomised trials	serious[1]	very serious[2]	not serious	not serious	none	959	655	-	SMD 0.36 lower (0.46 to 0.26 lower)	VERY LOW	CRITICAL
Depression symptomatology (change score) at 6 months (assessed with: Montgomery-Asberg Depression Rating Scale (MADRS)/Patient Health Questionnaire (PHQ-9) change from baseline to endpoint)												
2 (Knapstad 2020; Van Der Weele 2012)	randomised trials	serious[1]	very serious[2]	not serious	not serious	none	524	302	-	SMD 0.73 lower (0.89 to 0.58 lower)	VERY LOW	CRITICAL
Depression symptomatology (endpoint score) at 12 months (assessed with: Patient Health Questionnaire (PHQ-9))												
1 (Gureje 2019)	randomised trials	serious[1]	not serious	not serious	not serious	none	542	456	-	SMD 0.02 higher (0.1 lower to 0.15 higher)	MODERATE	CRITICAL
Depression symptomatology (change score) at 12 months (assessed with: Montgomery-Asberg Depression Rating Scale (MADRS) change from baseline to endpoint)												
1 (Van Der Weele 2012)	randomised trials	serious[1]	not serious	not serious	serious[3]	none	101	93	-	SMD 0.24 higher (0.04 lower to 0.53 higher)	LOW	CRITICAL
Response at 6 months (assessed with: Number of participants showing improvement of at least 50% on Montgomery-Asberg Depression Rating Scale (MADRS))												
1 (Van Der	randomised trials	serious[1]	not serious	not serious	very serious[4]	none	17/121 (14.0%)	23/118 (19.5%)	RR 0.72 (0.41 to 1.28)	55 fewer per	VERY LOW	CRITICAL

Quality assessment							Number of participants		Effect			
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Stepped care	Standard care/enhanced standard care	Relative (95% CI)	Absolute (95% CI)	Quality	Importance
Weele 2012)										1,000 (from 115 fewer to 55 more)		
Response at 12 months (assessed with: Number of participants showing improvement of at least 50% on Montgomery-Asberg Depression Rating Scale (MADRS))												
1 (Van Der Weele 2012)	randomised trials	serious[1]	not serious	not serious	serious[3]	none	21/121 (17.4%)	31/118 (26.3%)	RR 0.66 (0.40 to 1.08)	89 fewer per 1,000 (from 158 fewer to 21 more)	LOW	CRITICAL
Remission at 6 months (assessed with: Number of participants showing Hamilton Depression Rating Scale (HAMD) score < 11/ Patient Health Questionnaire (PHQ-9) score < 6)												
2 (Adewuya 2019; Callahan 1994)	randomised trials	serious[1]	serious[5]	not serious	not serious	none	259/556 (46.6%)	126/526 (24%)	RR 2 (1.69 to 2.38)	240 more per 1,000 (from 165 more to 331 more)	LOW	CRITICAL
Remission at 12 months (assessed with: Number of participants showing Patient Health Questionnaire (PHQ-9) score < 6)												
2 (Adewuya 2019; Gureje 2019)	randomised trials	serious[1]	very serious[2]	not serious	very serious[4]	none	756/1087 (69.5%)	502/998 (50.3%)	RR 1.81 (0.45 to 7.28)	407 more per 1,000 (from 277 fewer to 1000 more)	VERY LOW	CRITICAL
Antidepressant use at 6 months (assessed with: Number of participants receiving antidepressants)												
1 (Calla	randomised trials	serious[1]	not serious	not serious	not serious	none	27/100 (27.0%)	7/75 (9.3%)	RR 2.89 (1.33 to 6.28)	176 more per	MODERATE	IMPORTANT

Quality assessment							Number of participants		Effect			
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Stepped care	Standard care/enhanced standard care	Relative (95% CI)	Absolute (95% CI)	Quality	Importance
han 1994)										1,000 (from 31 more to 493 more)		
Discontinuation at 6 months (assessed with: Number of participants who dropped out of the study for any reason)												
5 (Adewuya 2019; Callahan 1994; Gureje 2019; Knapstad 2020; Van Der Weele 2012)	randomised trials	not serious	serious[5]	not serious	serious[3]	none	334/1771 (18.9%)	307/1409 (21.8%)	RR 0.75 (0.6 to 0.94)	54 fewer per 1,000 (from 13 fewer to 87 fewer)	LOW	IMPORTANT
Discontinuation at 12 months (assessed with: Number of participants who dropped out of the study for any reason)												
3 (Adewuya 2019; Gureje 2019; Van Der Weele 2012)	randomised trials	not serious	not serious	not serious	serious[3]	none	154/1208 (12.7%)	195/1116 (17.5%)	RR 0.74 (0.61 to 0.9)	45 fewer per 1,000 (from 17 fewer to 68 fewer)	MODERATE	IMPORTANT

CI: Confidence interval; SMD: Standardised mean difference; RR: Risk ratio
1. Risk of bias is high or unclear across multiple domains
2. I-squared>80%
3. 95% CI crosses 1 clinical decision threshold
4. 95% CI crosses 2 clinical decision thresholds
5. I-squared>50%

Table 32: Clinical evidence profile for Comparison 4. Stepped care for relapse prevention versus standard care

Quality assessment							Number of participants		Effect			
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Stepped care	Standard care	Relative (95% CI)	Absolute (95% CI)	Quality	Importance
Relapse at 12 months (assessed with: Number of participants who relapsed according to Mini-International Neuropsychiatric Interview (MINI))												
1 (Apil 2012)	randomised trials	serious[1]	not serious	not serious	serious[2]	none	19/74 (25.7%)	9/61 (14.8%)	RR 1.74 (0.85 to 3.56)	109 more per 1,000 (from 22 fewer to 378 more)	LOW	CRITICAL
Antidepressant use at 12 months (assessed with: Number of participants receiving antidepressants)												
1 (Apil 2012)	randomised trials	serious[1]	not serious	not serious	very serious[3]	none	25/49 (51.0%)	24/45 (53.3%)	RR 0.96 (0.65 to 1.41)	21 fewer per 1,000 (from 187 fewer to 219 more)	VERY LOW	IMPORTANT
Discontinuation at 12 months (assessed with: Number of participants who dropped out of the study for any reason)												
1 (Apil 2012)	randomised trials	not serious	not serious	not serious	very serious[3]	none	35/74 (47.3%)	30/62 (48.4%)	RR 0.98 (0.69 to 1.39)	10 fewer per 1,000 (from 150 fewer to 189 more)	LOW	IMPORTANT

CI: Confidence interval; RR: Risk ratio
1. Risk of bias is high or unclear across multiple domains
2. 95% CI crosses 1 clinical decision threshold
3. 95% CI crosses 2 clinical decision thresholds

Table 33: Clinical evidence profile for Comparison 5: Pure medication management versus standard care

Quality assessment							Number of participants		Effect			
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Pure medication management	Standard care	Relative (95% CI)	Absolute (95% CI)	Quality	Importance
Depression symptomatology at 6 months (assessed with: Montgomery-Asberg Depression Rating Scale (MADRS)/Patient Health Questionnaire (PHQ-9))												
2 (Aljumah 2015; Rubio-Valera 2013a)	randomised trials	not serious	not serious	not serious	not serious	none	197	202	-	SMD 0.05 higher (0.15 lower to 0.24 higher)	HIGH	CRITICAL
Response at 6 months (assessed with: Number of participants showing improvement of at least 50% on Hamilton Depression Rating Scale (HAMD))												
1 (Sirey 2010)	randomised trials	not serious	not serious	not serious	serious[1]	none	14/33 (42.4%)	8/37 (21.6%)	RR 1.96 (0.94 to 4.08)	208 more per 1,000 (from 13 fewer to 666 more)	MODERATE	CRITICAL
Antidepressant use at 6 months (assessed with: Number of participants adhering to antidepressant medication)												
3 (Akerblad 2003; Rickels 2005; Rubio-Valera 2013a)	randomised trials	serious[2]	not serious	not serious	serious[1]	none	218/441 (49.4%)	183/463 (39.5%)	RR 1.28 (1.10 to 1.49)	111 more per 1,000 (from 40 more to 194 more)	LOW	IMPORTANT
Discontinuation at 6 months (assessed with: Number of participants who dropped out of the study for any reason)												
5 (Akerblad	randomised trials	not serious	not serious	not serious	serious[1]	none	114/596 (19.1%)	133/620 (21.5%)	RR 0.89 (0.71 to 1.11)	24 fewer per	MODERATE	IMPORTANT

Quality assessment							Number of participants		Effect			
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Pure medication management	Standard care	Relative (95% CI)	Absolute (95% CI)	Quality	Importance
2003; Aljumah 2015; Rickles 2005; Rubio-Valera 2013a; Sirey 2010)										1,000 (from 62 fewer to 24 more)		

CI: Confidence interval; SMD: Standardised mean difference; RR: Risk ratio
1. 95% CI crosses 1 clinical decision threshold
2. Risk of bias is high or unclear across multiple domains

Table 34: Clinical evidence profile for Comparison 6: Care coordination versus standard care/enhanced standard care

Quality assessment							Number of participants		Effect			
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Care coordination	Standard care/enhanced standard care	Relative (95% CI)	Absolute (95% CI)	Quality	Importance
Depression symptomatology at 6 months (measured with: Montgomery-Asberg Depression Rating Scale (MADRS))												
1 (McMahon 2007)	randomised trials	serious[1]	not serious	not serious	serious[2]	reporting bias[3]	30	32	-	SMD 0.09 lower (0.59 lower to 0.41 higher)	VERY LOW	CRITICAL
Depression symptomatology at 12 months (measured with: Patient Health Questionnaire (PHQ-9))												
1 (Salisbury 2016)	randomised trials	serious[1]	not serious	not serious	not serious	none	255	261	-	SMD 0.05 lower (0.22 lower to 0.13 higher)	MODERATE	CRITICAL
Remission at 12 months (assessed with: Number of participants showing score < 10 on Patient Health Questionnaire (PHQ-9))												

Quality assessment							Number of participants		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Care coordination	Standard care/enhanced standard care	Relative (95% CI)	Absolute (95% CI)		
1 (Salisbury 2016)	randomised trials	serious[1]	not serious	not serious	serious[2]	none	95/307 (30.9%)	86/302 (28.5%)	RR 1.09 (0.85 to 1.39)	26 more per 1,000 (from 43 fewer to 111 more)	LOW	CRITICAL
Discontinuation at 6 months (assessed with: Number of participants who dropped out of the study for any reason)												
1 (McMahon 2007)	randomised trials	serious[1]	not serious	not serious	very serious[4]	reporting bias[3]	12/30 (40.0%)	16/32 (50.0%)	RR 0.80 (0.46 to 1.40)	100 fewer per 1,000 (from 270 fewer to 200 more)	VERY LOW	IMPORTANT
Discontinuation at 12 months (assessed with: Number of participants who dropped out of the study for any reason)												
1 (Salisbury 2016)	randomised trials	serious[1]	not serious	not serious	serious[2]	none	52/307 (16.9%)	41/302 (13.6%)	RR 1.25 (0.86 to 1.82)	34 more per 1,000 (from 19 fewer to 111 more)	LOW	IMPORTANT

CI: Confidence interval; SMD: Standardised mean difference; RR: Risk ratio
1. Risk of bias is high or unclear across multiple domains
2. 95% CI crosses 1 clinical decision threshold
3. Funding from pharmaceutical company
4. 95% CI crosses 2 clinical decision thresholds

Table 35: Clinical evidence profile for Comparison 7: Attached professional model versus enhanced standard care

Quality assessment							Number of participants		Effect			
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Attached professional model	Enhanced standard care	Relative (95% CI)	Absolute (95% CI)	Quality	Importance
Depression symptomatology at 6 months (measured with: Quick Inventory of Depressive Symptomatology (QIDS))												
1 (Bedoya 2014)	randomised trials	very serious[1]	not serious	not serious	serious[2]	none	63	55	-	SMD 0.36 lower (0.73 lower to 0 higher)	VERY LOW	CRITICAL
Discontinuation at 6 months (assessed with: Number of participants who dropped out of the study for any reason)												
1 (Bedoya 2014)	randomised trials	serious[1]	not serious	not serious	very serious[3]	none	9/65 (13.8%)	11/55 (20.0%)	RR 0.69 (0.31 to 1.55)	62 fewer per 1,000 (from 138 fewer to 110 more)	VERY LOW	IMPORTANT

CI: Confidence interval; SMD: Standardised mean difference; RR: Risk ratio
1. Risk of bias is high or unclear across multiple domains
2. 95% CI crosses 1 clinical decision threshold
3. 95% CI crosses 2 clinical decision thresholds

Table 36: Clinical evidence profile for Comparison 8: Shared care versus standard care

Quality assessment							Number of participants		Effect			
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Shared care	Standard care	Relative (95% CI)	Absolute (95% CI)	Quality	Importance
Depression symptomatology at 6 months (measured with: Montgomery-Asberg Depression Rating Scale (MADRS) change score)												
1 (Banerjee 1996)	randomised trials	not serious	not serious	not serious	not serious	none	33	36	-	SMD 1.03 lower (1.53 lower to 0.52 lower)	HIGH	CRITICAL
Remission at 6 months (assessed with: Number of participants who lost their diagnosis)												

Quality assessment							Number of participants		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Shared care	Standard care	Relative (95% CI)	Absolute (95% CI)		
1 (Banerjee 1996)	randomised trials	not serious	not serious	not serious	serious[1]	none	19/33 (57.6%)	9/36 (25.0%)	RR 2.30 (1.22 to 4.36)	325 more per 1,000 (from 55 more to 840 more)	MODERATE	CRITICAL
Antidepressant use at 6 months (assessed with: Number of participants receiving antidepressants)												
1 (Banerjee 1996)	randomised trials	not serious	not serious	not serious	not serious	none	20/33 (60.6%)	5/36 (13.9%)	RR 4.36 (1.85 to 10.30)	467 more per 1,000 (from 118 more to 1,000 more)	HIGH	IMPORTANT
Discontinuation at 6 months (assessed with: Number of participants who dropped out of the study for any reason)												
1 (Banerjee 1996)	randomised trials	not serious	not serious	not serious	very serious[2]	none	4/33 (12.1%)	4/36 (11.1%)	RR 1.09 (0.30 to 4.01)	10 more per 1,000 (from 78 fewer to 334 more)	LOW	IMPORTANT

CI: Confidence interval; SMD: Standardised mean difference; RR: Risk ratio
1. 95% CI crosses 1 clinical decision threshold
2. 95% CI crosses 2 clinical decision thresholds

Table 37: Clinical evidence profile for Comparison 9: Measurement-based care versus standard care

Quality assessment							Number of participants		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Measurement-based care	Standard care	Relative (95% CI)	Absolute (95% CI)		
Depression symptomatology at 6 months (measured with: Hamilton Depression Rating Scale (HAMD))												

Quality assessment							Number of participants		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Measurement-based care	Standard care	Relative (95% CI)	Absolute (95% CI)		
1 (Guo 2015)	randomised trials	serious[1]	not serious	not serious	not serious	none	44	37	-	SMD 1.05 lower (1.51 lower to 0.58 lower)	MODERATE	CRITICAL
Response at 6 months (assessed with: Number of participants showing improvement of at least 50% on Hamilton Depression Rating Scale (HAMD))												
1 (Guo 2015)	randomised trials	serious[1]	not serious	not serious	serious[2]	none	53/61 (86.9%)	37/59 (62.7%)	RR 1.39 (1.11 to 1.73)	245 more per 1,000 (from 69 more to 458 more)	LOW	CRITICAL
Remission at 6 months (assessed with: Number of participants showing score <8 on Hamilton Depression Rating Scale (HAMD))												
1 (Guo 2015)	randomised trials	serious[1]	not serious	not serious	not serious	none	45/61 (73.8%)	17/59 (28.8%)	RR 2.56 (1.67 to 3.93)	449 more per 1,000 (from 193 more to 844 more)	MODERATE	CRITICAL
Discontinuation at 6 months (assessed with: Number of participants who dropped out of the study for any reason)												
1 (Guo 2015)	randomised trials	serious[1]	not serious	not serious	very serious[3]	none	17/61 (27.9%)	22/59 (37.3%)	RR 0.75 (0.44 to 1.26)	93 fewer per 1,000 (from 209 fewer to 97 more)	VERY LOW	IMPORTANT

CI: Confidence interval; SMD: Standardised mean difference; RR: Risk ratio
1. Risk of bias is high or unclear across multiple domains
2. 95% CI crosses 1 clinical decision threshold
3. 95% CI crosses 2 clinical decision thresholds

GRADE tables for review question 1.2 For adults with depression, what are the relative benefits and harms associated with different settings for the delivery of care?

GRADE tables not provided for subgroup analyses of NMA dataset

Table 38: Clinical evidence profile for comparison 2 Crisis resolution team care versus standard care (for adults with non-psychotic severe mental illness)

Quality assessment							No of patients		Effect			
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Crisis resolution team care	Standard care	Relative (95% CI)	Absolute	Quality	Importance
Psychiatric symptom severity 8 weeks after crisis (measured with: Brief psychiatric rating scale (BPRS); Better indicated by lower values)												
1 (Johnson 2005)	randomised trials	very serious [1]	no serious inconsistency	serious[2]	serious[3]	none	107	104	-	SMD 0.29 lower (0.56 to 0.02 lower)	VERY LOW	CRITICAL
Service utilisation: Admission as inpatient 6 months after crisis (assessed with: Number of participants that had been admitted to a psychiatric ward within 6 months after crisis)												
1 (Johnson 2005)	randomised trials	very serious [1]	no serious inconsistency	serious[2]	no serious imprecision	none	39/134 (29.1%)	84/124 (67.7%)	RR 0.43 (0.32 to 0.57)	386 fewer per 1000 (from 291 fewer to 461 fewer)	VERY LOW	IMPORTANT
Service utilisation: Bed days in hospital 6 months after crisis (measured with: Number of bed days in hospital for those admitted within 6 months after crisis; Better indicated by lower values)												
1 (Johnson 2005)	randomised trials	very serious [1]	no serious inconsistency	serious[2]	serious[3]	none	134	123	-	SMD 0.45 lower (0.69 to	VERY LOW	IMPORTANT

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Crisis resolution team care	Standard care	Relative (95% CI)	Absolute		
										0.20 lower)		
Psychological functioning: Quality of life 8 weeks after crisis (measured with: Manchester short assessment of quality of life (MANSA) 8 weeks after crisis; Better indicated by higher values)												
1 (Johnson 2005)	randomised trials	very serious [1]	no serious inconsistency	serious[2]	no serious imprecision	none	114	103	-	SMD 0.11 lower (0.37 lower to 0.16 higher)	VERY LOW	IMPORTANT
Social functioning 8 weeks after crisis (measured with: Life Skills Profile (LSP); Better indicated by lower values)												
1 (Johnson 2005)	randomised trials	very serious [1]	no serious inconsistency	serious[2]	no serious imprecision	none	133	124	-	SMD 0.2 higher (0.05 lower to 0.44 higher)	VERY LOW	IMPORTANT
Social functioning 6 months after crisis (measured with: Life Skills Profile (LSP); Better indicated by lower values)												
1 (Johnson 2005)	randomised trials	very serious [1]	no serious inconsistency	serious[2]	no serious imprecision	none	133	122	-	SMD 0.06 higher (0.18 lower to 0.31 higher)	VERY LOW	IMPORTANT
Patient satisfaction 8 weeks after crisis (measured with: Client Satisfaction Questionnaire - 8 item version (CSQ-8); Better indicated by higher values)												

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Crisis resolution team care	Standard care	Relative (95% CI)	Absolute		
1 (Johnson 2005)	randomised trials	very serious [1]	no serious inconsistency	serious[2]	no serious imprecision	none	118	108	-	SMD 0.23 higher (0.03 lower to 0.49 higher)	VERY LOW	IMPORTANT

CI: Confidence interval; RR: Risk ratio; SMD: Standardised mean difference

- High risk of bias associated with randomisation method due to significant difference between groups at baseline and non-blind participants, intervention administrator(s) and outcome assessor(s)
- Not depression-specific population
- 95% CI crosses 1 clinical decision threshold

Table 39: Clinical evidence profile for comparison 4 Acute psychiatric day hospital care versus inpatient care (for adults with depression and non-psychotic severe mental illness)

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acute day hospital care	Inpatient care	Relative (95% CI)	Absolute		
Psychiatric symptom severity at 2-3 months post-admission (measured with: Comprehensive Psychopathological Rating Scale (CPRS; change score)/Brief Psychiatric Rating Scale (BPRS; change score)/Hamilton Rating Scale for Depression (HAM-D; change score); Better indicated by lower values)												

Quality assessment							No of patients		Effect			
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Acute day hospital care**	**Inpatient care**	**Relative (95% CI)**	**Absolute**	**Quality**	**Importance**
3	randomised trials	very serious[1]	serious[2]	serious[3]	no serious imprecision	none	682	599	-	SMD 0.05 higher (0.22 lower to 0.33 higher)	VERY LOW	CRITICAL
Psychiatric symptom severity at 12-14 months post-admission (measured with: Comprehensive Psychopathological Rating Scale (CPRS; change score)/Brief Psychiatric Rating Scale (BPRS; change score); Better indicated by lower values)												
2	randomised trials	very serious[1]	very serious[4]	serious[3]	serious[5]	none	663	586	-	SMD 0.19 lower (0.81 lower to 0.42 higher)	VERY LOW	CRITICAL
Response at 3 months post-admission (assessed with: Number of people showing ≥47% improvement on Hamilton Rating Scale for Depression (HAM-D))												
1	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	very serious[6]	none	6/24 (25%)	8/20 (40%)	RR 0.62 (0.26 to 1.5)	152 fewer per 1000 (from 296 fewer to 200 more)	VERY LOW	CRITICAL
Remission at 3-13 months post-admission (assessed with: Present State Examination: Index of Definition≤4/<7 on Hamilton Rating Scale for Depression (HAM-D))												
2	randomised trials	very serious[1]	no serious inconsistency	serious[3]	very serious[6]	none	33/80 (41.3%)	33/71 (46.5%)	RR 0.91 (0.65 to 1.26)	42 fewer per 1000 (from 163 fewer to 121 more)	VERY LOW	CRITICAL
Service utilisation: Duration of index admission (follow-up 12-14 months; measured with: Number of days/months in hospital; Better indicated by lower values)												
4	randomised trials	very serious[1]	no serious inconsistency	serious[3]	serious[5]	none	800	735	-	SMD 0.55 higher (0.44 to 0.65 higher)	VERY LOW	IMPORTANT
Service utilisation: Readmission at 4-12 months post-admission (assessed with: Number of patients readmitted to hospital)												

Quality assessment							No of patients		Effect			
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Acute day hospital care**	**Inpatient care**	**Relative (95% CI)**	**Absolute**	**Quality**	**Importance**
3	randomised trials	very serious[1]	serious[2]	serious[3]	very serious[6]	none	39/183 (21.3%)	47/189 (24.9%)	RR 0.79 (0.41 to 1.52)	52 fewer per 1000 (from 147 fewer to 129 more)	VERY LOW	IMPORTANT
Service utilisation: Emergency contacts 4 months post-admission (assessed with: Number of participants making emergency contacts within 4 months post-admission)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[3]	serious[5]	none	12/38 (31.6%)	6/45 (13.3%)	RR 2.37 (0.98 to 5.71)	183 more per 1000 (from 3 fewer to 628 more)	VERY LOW	IMPORTANT
Service utilisation: Outpatient contact 4 months post-admission (assessed with: Number of participants making outpatient contacts within 4 months post-admission)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[3]	very serious[6]	none	14/38 (36.8%)	12/45 (26.7%)	RR 1.38 (0.73 to 2.62)	101 more per 1000 (from 72 fewer to 432 more)	VERY LOW	IMPORTANT
Psychological functioning: Quality of life at 2-months post-admission (measured with: Manchester short assessment of quality of life (MANSA); Better indicated by higher values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[3]	no serious imprecision	none	596	521	-	SMD 0.01 higher (0.11 lower to 0.13 higher)	VERY LOW	IMPORTANT
Psychological functioning: Quality of life at 14-months post-admission (measured with: Manchester short assessment of quality of life (MANSA); Better indicated by higher values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[3]	no serious imprecision	none	596	521	-	SMD 0.01 higher (0.11 lower to 0.13 higher)	VERY LOW	IMPORTANT

Quality assessment							No of patients		Effect			
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acute day hospital care	Inpatient care	Relative (95% CI)	Absolute	Quality	Importance
Social functioning response at 12-13 months (assessed with: Number of participants scoring 2 role disabilities or less on Groningen Social Disabilities Schedule (GSDS)/Number of participants living in the community and social functioning at previous level (according to the social performance and behaviour assessment schedule))												
2	randomised trials	very serious[1]	no serious inconsistency	serious[3]	serious[5]	none	41/91 (45.1%)	30/90 (33.3%)	RR 1.36 (0.94 to 1.96)	120 more per 1000 (from 20 fewer to 320 more)	VERY LOW	IMPORTANT
Social functioning impairment at 2-months post-admission (measured with: Groningen Social Disabilities Schedule, Second revision (GSDS-II); Better indicated by lower values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[3]	no serious imprecision	none	596	521	-	SMD 0.3 lower (0.42 to 0.19 lower)	VERY LOW	IMPORTANT
Social functioning impairment at 14-months post-admission (measured with: Groningen Social Disabilities Schedule, Second revision (GSDS-II); Better indicated by lower values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[3]	no serious imprecision	none	596	521	-	SMD 0.15 lower (0.27 to 0.04 lower)	VERY LOW	IMPORTANT
Satisfaction at 4 months post-admission (assessed with: Number of participants satisfied or very satisfied with their treatment)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[3]	no serious imprecision	none	31/38 (81.6%)	19/45 (42.2%)	RR 1.93 (1.33 to 2.81)	393 more per 1000 (from 139 more to 764 more)	VERY LOW	IMPORTANT
Satisfaction at 2 months post-admission (measured with: Client Assessment of Treatment (CAT); Better indicated by higher values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[3]	no serious imprecision	none	596	521	-	SMD 0.03 higher (0.09 lower to 0.15 higher)	VERY LOW	IMPORTANT

Quality assessment							No of patients		Effect			
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Acute day hospital care**	**Inpatient care**	**Relative (95% CI)**	**Absolute**	**Quality**	**Importance**
Carer distress at 3-months post-admission (measured with: General Health Questionnaire (GHQ; change score); Better indicated by lower values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[3]	serious[5]	none	38	39	-	MD 1.1 lower (3.15 lower to 0.95 higher)	VERY LOW	IMPORTANT
Carer distress at 12-months post-admission (measured with: General Health Questionnaire (GHQ; change score); Better indicated by lower values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[3]	serious[5]	none	24	31	-	MD 0.4 lower (2.98 lower to 2.18 higher)	VERY LOW	

CI: Confidence interval; RR: Risk ratio; SMD: Standardised mean difference

1. Risk of bias is high or unclear across multiple domains

2. I-squared>50%
3. Non depression-specific population
4. I-squared>80%
5. 95% CI crosses 1 clinical decision threshold
6. 95% CI crosses 2 clinical decision thresholds

Table 40: Clinical evidence profile for comparison 5 Non-acute day hospital care versus outpatient care (for adults with depression and non-psychotic severe mental illness)

Quality assessment							No of patients		Effect			
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Non-acute day hospital care	Outpatient care	Relative (95% CI)	Absolute	Quality	Importance
Psychiatric symptom severity at 4-6 months post-admission (measured with: Psychiatric Evaluation Form (change score)/Present State Examination (change score); Better indicated by lower values)												
2	randomised trials	serious[1]	very serious[2]	serious[3]	very serious[4]	none	75	69	-	SMD 0.08 higher (0.72 lower to 0.88 higher)	VERY LOW	CRITICAL
Psychiatric symptom severity at 8-12 months post-admission (measured with: Psychiatric Evaluation Form (change score)/Present State Examination (change score); Better indicated by lower values)												
2	randomised trials	serious[1]	no serious inconsistency	serious[3]	no serious imprecision	none	73	66	-	SMD 0.15 lower (0.49 lower to 0.19 higher)	LOW	CRITICAL
Service utilisation: Admission as inpatient 6-12 months post-admission (assessed with: Number of participants admitted into inpatient care during the study period)												
3	randomised trials	serious[1]	no serious inconsistency	serious[3]	very serious[4]	none	16/136 (11.8%)	12/145 (8.3%)	RR 1.26 (0.52 to 3.06)	22 more per 1000 (from 40 fewer to 170 more)	VERY LOW	IMPORTANT
Global functioning at 6-months post-admission (measured with: Global Assessment Scale (GAS; change score); Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[3]	very serious[4]	none	34	18	-	SMD 0.04 higher (0.53 lower to 0.61 higher)	VERY LOW	IMPORTANT
Global functioning at 12-months post-admission (measured with: Global Assessment Scale (GAS; change score); Better indicated by lower values)												

Quality assessment							No of patients		Effect			
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Non-acute day hospital care**	**Outpatient care**	**Relative (95% CI)**	**Absolute**	**Quality**	**Importance**
1	randomised trials	serious[1]	no serious inconsistency	serious[3]	serious[5]	none	33	18	-	SMD 0.12 lower (0.7 lower to 0.45 higher)	VERY LOW	IMPORTANT
Social functioning at 4-6 months post-admission (measured with: Social Adjustment Scale-Self Report (SAS-SR; change score)/Social Functioning Scale (SFS; change score); Better indicated by lower values)												
2	randomised trials	serious[1]	no serious inconsistency	serious[3]	serious[5]	none	74	67	-	SMD 0.2 lower (0.54 lower to 0.14 higher)	VERY LOW	IMPORTANT
Social functioning at 8-12 months post-admission (measured with: Social Adjustment Scale-Self Report (SAS-SR; change score)/Social Functioning Scale (SFS; change score); Better indicated by lower values)												
2	randomised trials	serious[1]	no serious inconsistency	serious[3]	serious[5]	none	73	67	-	SMD 0.31 lower (0.65 lower to 0.03 higher)	VERY LOW	IMPORTANT
Satisfaction at 4-6 months post-admission (assessed with: Number of participants satisfied or very satisfied with their treatment)												
2	randomised trials	serious[1]	very serious[2]	serious[3]	very serious[4]	none	59/92 (64.1%)	67/106 (63.2%)	RR 1 (0.47 to 2.12)	0 fewer per 1000 (from 335 fewer to 708 more)	VERY LOW	IMPORTANT

CI: Confidence interval; RR: Risk ratio; SMD: Standardised mean difference

- *Risk of bias is high or unclear across multiple domains*
- *I-squared>80%*
- Non-depression specific population
- 95% CI crosses 2 clinical decision thresholds

- 95% CI crosses 1 clinical decision threshold

Table 41: Clinical evidence profile for comparison 6 Community mental health teams versus standard care (for adults with non-psychotic severe mental illness)

Quality assessment							No of patients		Effect			
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Community mental health teams (CMHTs)**	**Standard care**	**Relative (95% CI)**	**Absolute**	**Quality**	**Importance**
Psychiatric symptom severity at 3 months post-entry (measured with: Comprehensive Psychopathological Rating Scale (CPRS); Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	no serious imprecision	none	48	52	-	SMD 0.06 lower (0.45 lower to 0.33 higher)	LOW	CRITICAL
Service utilisation: Admission as inpatient at 3 months post-entry (assessed with: Number of participants admitted into inpatient care during the study period)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	serious[3]	none	7/48 (14.6%)	16/52 (30.8%)	RR 0.47 (0.21 to 1.05)	163 fewer per 1000 (from 243 fewer to 15 more)	VERY LOW	IMPORTANT
Service utilisation: Admission as inpatient for >10 days at 3 months post-entry (assessed with: Number of participants admitted into inpatient care for more than 10 days during the study period)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	serious[3]	none	2/48 (4.2%)	11/52 (21.2%)	RR 0.2 (0.05 to 0.84)	169 fewer per 1000 (from 34 fewer to 201 fewer)	VERY LOW	IMPORTANT
Satisfaction at 3 months post-entry (assessed with: Number of participants satisfied with their treatment)												

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Community mental health teams (CMHTs)	Standard care	Relative (95% CI)	Absolute		
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	serious[3]	none	34/41 (82.9%)	25/46 (54.3%)	RR 1.53 (1.13 to 2.06)	288 more per 1000 (from 71 more to 576 more)	VERY LOW	IMPORTANT
Satisfaction at 3 months post-entry (measured with: Service Satisfaction Score; Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	serious[3]	none	41	46	-	SMD 0.85 higher (0.41 to 1.29 higher)	VERY LOW	IMPORTANT

CI: Confidence interval; RR: Risk ratio; SMD: Standardised mean difference

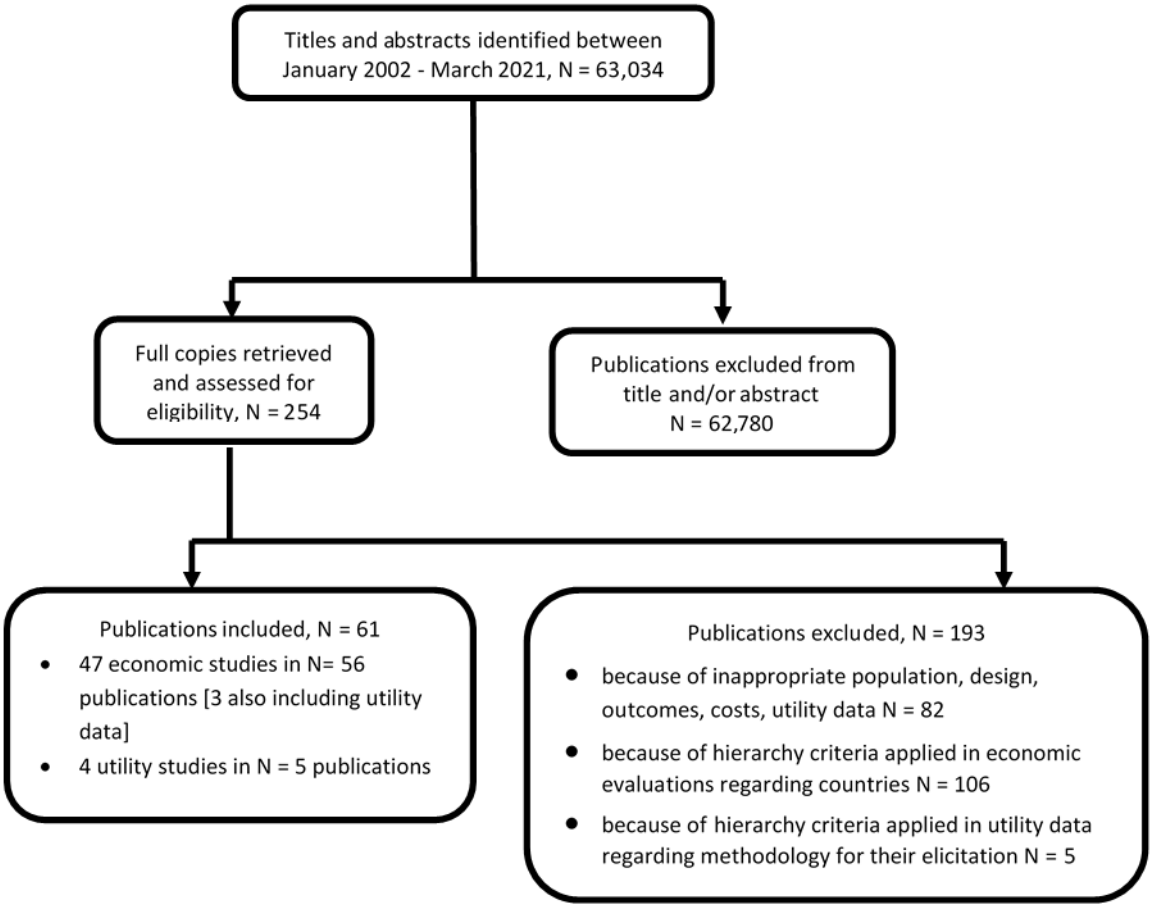
- Risk of bias is high or unclear across multiple domains
- Non-depression specific population

Appendix G – Economic evidence study selection

Economic evidence study selection for review question 1.1 For adults with depression, what are the relative benefits and harms associated with different models for the coordination and delivery of services?

A global health economics search was undertaken for all areas covered in the guideline. Figure 79 shows the flow diagram of the selection process for economic evaluations of interventions and strategies for adults with depression and studies reporting depression-related health state utility data.

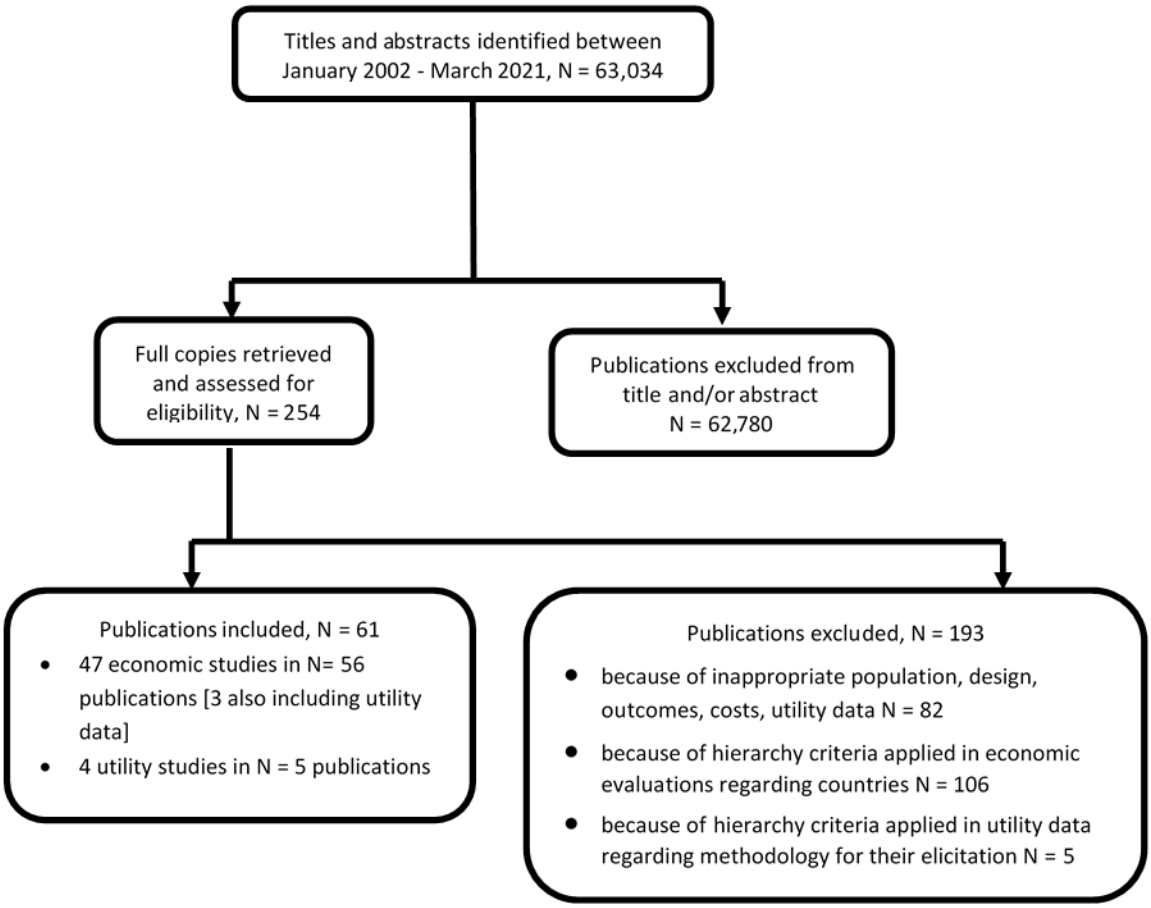
Figure 110. Flow diagram of selection process for economic evaluations of interventions and strategies for adults with depression and studies reporting depression-related health state utility data



Economic evidence study selection for review question 1.2 For adults with depression, what are the relative benefits and harms associated with different settings for the delivery of care?

A global health economics search was undertaken for all areas covered in the guideline. Figure 61 shows the flow diagram of the selection process for economic evaluations of interventions and strategies for adults with depression and studies reporting depression-related health state utility data.

Figure 111. Flow diagram of selection process for economic evaluations of interventions and strategies for adults with depression and studies reporting depression-related health state utility data.



Appendix H – Economic evidence tables

Economic evidence tables for review question 1.1 For adults with depression, what are the relative benefits and harms associated with different models for the coordination and delivery of services?

Table 42: Economic evidence table for simple collaborative care

Study Country Study type	Intervention details	Study population Study design Data sources	Costs and outcomes: description and values	Results: Cost-effectiveness	Comments
Bosanquet 2017 UK Cost-utility analysis	Interventions: Simple collaborative care (SCC), using behavioural activation, designed specifically for people aged ≥ 65 with depression, delivered over 8 sessions by a case manager (a primary care mental health / IAPT worker) for an average of 6 sessions over 7-8 weeks. SCC included telephone support, medication management, symptom monitoring and active surveillance, facilitated by a computerised case management. The first session was delivered face-to-face and subsequent sessions via telephone. SCC was provided in	Adults aged ≥ 65 years with major depressive disorder. Exclusion criteria: alcohol dependency; psychotic symptoms; recent suicidal risk/self-harm; significant cognitive impairment Pragmatic, multi-centre open RCT (N=485) Source of efficacy and resource use data: RCT (Bosanquet 2017); (N=485; at 18 months n=344; cost data available for n=447) Source of unit costs: national sources	Costs: intervention (case manager's time and supervision, as well as training including manual, supervision, travel and accommodation) and usual primary care (GP appointment, home visits and telephone consultation; practice nurse appointments and telephone consultations) Mean total cost per person (95% CI): SCC: £1,171 (£1,167 to £1176); TAU: £654 (£651 to £658) Adjusted difference £480 (£381 to £579). Primary outcome measure: QALY based on SF-6D ratings (UK tariff) Mean number of QALYs per person (SD): SCC: 0.900 (0.241); TAU: 0.889 (0.224) Adjusted difference 0.019 (95% CI -0.020 to 0.057, p=0.338)	ICER of SCC vs TAU: £26,010/QALY Probability of SCC being cost-effective: 0.39 and 0.55 at WTP £20,000 and £30,000/QALY, respectively. Sensitivity analysis: Including only participants who engaged with 5 or more sessions in the analysis: ICER £9,876/QALY	Perspective: NHS/PSS (intervention and primary care exclusively considered) Currency: GBP£ Cost year: 2012/13 Time horizon: 18 months Discounting: NA Applicability: directly applicable Quality: potentially serious limitations

Study Country Study type	Intervention details	Study population Study design Data sources	Costs and outcomes: description and values	Results: Cost-effectiveness	Comments
	addition to usual GP care. Treatment as usual, comprising GP care alone (TAU)				
Green 2014 UK Cost-utility analysis	Interventions: Simple collaborative care in addition to usual primary care (SCC), comprising care managers making 6-12 contacts with service users over 14 weeks; contacts involved education about depression, medication management, behavioural activation and relapse prevention instructions. Care managers provided GPs with advice on medication and regular updates on service user progress including medication adherence. Treatment as usual (TAU), defined as GP care that includes antidepressant treatment and referral for other treatments, including Improving Access to Psychological	Adults with depression Multi-centre cluster RCT (N=581) Source of efficacy data: RCT (Richards 2013); (data available for n=466) Source of resource use data: RCT (data available for n=447) Source of unit costs: national sources	Costs: intervention (care manager's time and supervision by specialists), staff time (GP, mental health nurse, practice nurse, counsellor, mental health worker, social worker, home care worker, occupational therapist, psychiatrist, psychologist, psychiatric nurse/care coordinator), walk-in-centre, voluntary group, inpatient psychiatric and general stay, A&E, day hospital, other outpatient contact, day care centre, drop-in club; informal care and service user expenses in sensitivity analysis Mean NHS/PSS cost per person (SD): SCC: £1,887 (£3,714); TAU: £1,571 (£2,442) Unadjusted difference: £316 Adjusted difference: £271 (95%CI: -£203 to £886) Primary outcome measure: QALY based on EQ-5D ratings (UK tariff); SF-6D (UK tariff) used in sensitivity analysis Mean number of QALYs per person (SD):	ICER of SCC vs TAU: £14,248/QALY Probability of SCC being cost-effective: 0.58 and 0.65 at WTP £20,000 and £30,000/QALY, respectively. Results robust to multiple imputation of missing data, use of SF-6D utility values, use of alternative SCC costs; SCC dominant using a broader perspective; excluding one participant with an extremely high level of self-reported resource use, ICER became £3,334/QALY and probability of cost effectiveness 0.76 and 0.79 at WTP £20,000 and £30,000 /QALY, respectively	Perspective: NHS/PSS; broader perspective (informal care costs and service user expenses) considered in sensitivity analysis Currency: GBP£ Cost year: 2011 Time horizon: 12 months Discounting: NA Applicability: directly applicable Quality: minor limitations

Study Country Study type	Intervention details	Study population Study design Data sources	Costs and outcomes: description and values	Results: Cost-effectiveness	Comments
	Therapies (IAPT) services		SCC: 0.605 (0.261); TAU: 0.554 (0.286) Unadjusted difference: 0.051 Adjusted difference: 0.019 (95%CI: -0.019 to 0.06)		
Lewis 2017 UK Cost-utility analysis	Interventions: Simple collaborative care (SCC), which included behavioural activation delivered by a case manager (a primary care mental health worker / Improving Access to Psychological Therapies (IAPT) worker) for an average of 7 sessions over 8–10 weeks, in addition to usual GP care. Collaborative care included telephone support, symptom monitoring and active surveillance, facilitated by computerised case management. Treatment as usual, comprising GP care alone (TAU)	Older adults who screened positive for subthreshold depression (≥ 75 years old during the pilot phase and ≥ 65 years old during the main trial) Pragmatic, multi-centre RCT (N=705) Source of efficacy and resource use data: RCT (Gilbody 2017); (N=705; complete data used in base-case economic analysis n=448) Source of unit costs: national sources	Costs: intervention (case manager's time and supervision, as well as training including manual, supervision, travel and accommodation) and usual primary care (GP appointment, home visits and telephone consultation; practice nurse appointments and telephone consultations) Mean NHS/PSS cost per person (SD): SCC: £894 (£391); TAU: £450 (£393) Unadjusted difference: £444 for n=620 Adjusted bootstrapped difference for n=448 sample included in economic analysis: £421 (95%CI: £348 to £494) Primary outcome measure: QALY based on EQ-5D ratings (UK tariff) Mean number of QALYs per person (SD): SCC: 0.756 (0.246); TAU: 0.660	ICER of SCC vs TAU: £9,633/QALY Probability of SCC being cost-effective: 0.92 and 0.97 at WTP £20,000 and £30,000/QALY, respectively. Sensitivity analysis: Accounting for the true observed SCC contact rate (rather than the expected SCC contact rate that was used in the base-case analysis), ICER became £3,328/QALY	Perspective: NHS/PSS (intervention and primary care exclusively considered) Currency: GBP£ Cost year: 2012/13 Time horizon: 12 months Discounting: NA Applicability: directly applicable Quality: potentially serious limitations

Study Country Study type	Intervention details	Study population Study design Data sources	Costs and outcomes: description and values	Results: Cost-effectiveness	Comments
			(0.247) Unadjusted difference: 0.096 Adjusted difference: 0.044 (95%CI: 0.015 to 0.072, p=0.003)		
Simon 2002 US Cost effectiveness analysis	Interventions: Simple collaborative care comprising an educational book and videotape on effective management of depression; 2 visits to a depression prevention specialist including shared decision making on maintenance antidepressant treatment; plus 3 scheduled telephone contacts and 4 personalised mailings for monitoring depressive symptoms and treatment adherence (SCC) Treatment as usual (TAU), including primary care and referral to specialty mental health care	Adults with a history of either recurrent major depression (i.e. at least 3 depressive episodes in the previous 5 years) or dysthymia (depressive symptoms present continuously for the past 2 years) that had recovered from a depressive episode following antidepressant treatment in primary care RCT (Katon 2001) Source of efficacy and resource use data: RCT; N=386, n=315 (82%) completed all follow-up assessments; n=377 (98%) remained enrolled throughout the follow-up period Source of unit costs: local data	Costs: medication, staff time, any inpatient and outpatient services for mental health or general medical care Mean total cost cost per person: SCC: $2,691 (95%CI $2,320 to $3,062) TAU: $2,619 (95%CI $2,139 to $3,099) Incremental $13 (95%CI -$584 to $511), after adjustment for gender, age, baseline Hopkins Symptoms Checklist (HSCL) depression score and chronic disease score Primary outcome measure: number of depression-free days, defined as days with a HSCL depression score ≤ 0.5; days with a HSCL score above 0.5 but < 2 were considered 50% depression free Number of depression-free days: SCC: 253.2 (95% CI 241.7 to 264.7) TAU: 239.4 (95% CI 227.3 to 251.4) Incremental 13.9 (95%CI -1.5 to 29.3, p=0.078), after adjustment for gender, age, baseline SCL	ICER of SCC vs. TAU $1 per depression-free day (95%CI -$134 to $344)	Perspective: 3rd party payer Currency: US$ Cost year: 1998 Time horizon: 12 months Discounting: NA Applicability: partially applicable Quality: potentially serious limitations

Study Country Study type	Intervention details	Study population Study design Data sources	Costs and outcomes: description and values	Results: Cost-effectiveness	Comments
			depression score and chronic disease score		

Table 43: Economic evidence table for complex collaborative care

Study Country Study type	Intervention details	Study population Study design Data sources	Costs and outcomes: description and values	Results: Cost-effectiveness	Comments
Morriss 2016 UK Cost-utility analysis	Interventions: Complex collaborative care, comprising secondary outpatient specialist depression services offering tailored integrated pharmacological and psychological (CBT, MBCT and compassion focused therapy, as appropriate) treatment within a collaborative care approach for 12-15 months (CCC) Usual secondary mental health care (TAU)	Adults with persistent unipolar moderate or severe depression, with HDRS total≥16, GAF≤60, that have received treatment for depression for at least 6 months and are currently receiving secondary mental healthcare Multi-site single-blind RCT (N=187) Source of efficacy and resource use data: RCT (Morriss 2016, N=187; 84% completed at 6 months, 72% at 12 months and 59% at 18 months) Source of unit costs: national sources	Costs: primary care (GP surgery and home attendances), practice / district / community psychiatric nurse, psychotherapist, inpatient and outpatient (psychiatric or other) care, A&E attendances, medication Mean total cost per person (95% CI): CCC: £9,315 (£7,547 to £11,084) TAU: £5,869 (£4,501 to £7,238) Incremental total cost (bias-corrected bootstrapped): £3,446 (£1,915 to £5,180) Primary outcome measure: QALYs based on EQ-5D-3L ratings (UK tariff) Mean QALYs per person (95% CI): CCC: 0.753 (0.659 to 0.847) TAU: 0.646 (0.538 to 0.754) Incremental QALYs (bias-corrected bootstrapped): 0.079 (0.007 to 0.149)	ICER of CCC vs. TAU £43,603/QALY Controlling for baseline differences and cluster effects: probability of CCC being cost-effective exceeds 0.50 at WTP of £42,000/QALY	Perspective: NHS and personal social services Currency: GBP£ Cost year: 2014 Time horizon: 18 months Discounting: NA Applicability: directly applicable Quality: minor limitations

Study Country Study type	Intervention details	Study population Study design Data sources	Costs and outcomes: description and values	Results: Cost-effectiveness	Comments
Goorden 2015 The Netherlands Cost-utility analysis	Interventions: Complex collaborative care (CCC) provided by a depression care manager, usually a qualified nurse, who collaborated with a GP and a liaison psychiatrist in order to provide and guide more structured and adherent depression treatment in primary care. Treatment consisted of problem solving, manual guided self-help (both provided by the care manager), and, if necessary, antidepressants (prescribed by the GP). Care managers and GPs received training in CCC. Treatment as usual (TAU) in primary care, comprising prescription of antidepressants or referral to psychotherapy	People aged ≥17 years with major depression according to the MINI. Exclusion criteria: being suicidal, psychotic symptoms, dementia, drug or alcohol dependence, already under specialty mental health treatment RCT (N=150; 93 identified by screening and 47 by GP referral) Source of efficacy and resource use data: RCT (Huijbregts 2013, n=93 identified by screeening) Source of unit costs: national sources	Costs: GP, psychiatric / mental health care practice nurse, psychiatric inpatient care, specialist outpatient care, private psychologist / psychiatrist, occupational physician, other specialist, paramedic, social worker, counselling centre for drugs, alcohol, etc, alternative medicine, self-help group, day care, psychotropic medication Mean total healthcare cost per person: CCC €4,011 (95% CI €,2679 to €,5513) TAU €2,838 (95% CI €,2463 to €,3244) Difference: €1,173 (95% CI, -€216 to €2726) Primary outcome measure: QALYs based on EQ-5D ratings (Dutch tariff) Mean total number of QALYs gained per person: CCC 0.07 (95% CI 0.05 to 0.09) TAU 0.05 (95% CI 0.03 to 0.06) Difference: 0.02 (95% CI −0.004 to 0.04)	ICER of TAU vs CCC €53,717/QALY Probability of CCC being cost-effective: 0.20 and 0.70 at WTP €20,000 and €80,000/QALY, respectively.	Perspective: healthcare system; productivity losses reported separately Currency: Euro (€) Cost year: 2013 Time horizon: 12 months Discounting: NA Applicability: partially applicable Quality: potentially serious limitations
Grochtdreis 2019 Germany	Interventions: Complex collaborative care (CCC) formed around a primary care physician (PCP);	Adults aged ≥ 60 years with moderate depressive symptoms; PHQ-9 score 10-14.	Costs: outpatient physician (e.g. PCP, specialist physician, psychotherapy) and non-physician services (e.g. physiotherapy, occupational therapy, massage), inpatient care,	ICER of CCC vs TAU €26.07/DFD €55,800/QALY	Perspective: healthcare system (informal care reported separately) Currency: Euro (€)

Study Country Study type	Intervention details	Study population Study design Data sources	Costs and outcomes: description and values	Results: Cost-effectiveness	Comments
Cost effectiveness	treatment evaluation occurred every 8 weeks. Intervention consisted of a patient manual, an initial face-to-face session and ongoing telephone sessions between the care manager and the patient every other week. Patients' depressive symptom severity was regularly assessed by the PHQ-9. Problem-solving techniques were optionally held. Treatment as usual (TAU) comprising regular PCP visits without involvement of a care manager. Depressive symptom severity not routinely assessed.	Exclusion criteria: alcohol/drug abuse, severe cognitive impairment, severe psychological disorders, suicidal ideation, active depression treatment Cluster RCT (N=246 from 71 clusters; ITT analysis) Source of efficacy and resource use data: RCT (Hölzel 2018) Source of unit costs: national sources	rehabilitation, formal nursing care (professional nurse or housekeeper), informal nursing care (family or friends), medication and medical devices. Mean total healthcare cost per person: CCC €6155; TAU €5674 Adjusted difference: €558; p = 0.532 Primary outcome measure: depression-free days (DFDs), based on PHQ-9 scores. PHQ-9 <5: depression-free; PHQ-9 ≥15: depressed; linear interpolation used for calculations. Secondary outcome measure: QALYs based on EQ-5D ratings (UK tariff) Mean total DFDs per person: CCC 207.1; TAU 185.8 Adjusted difference: 21.4; p = 0.022 Mean total QALYs per person: CCC 0.57; TAU 0.56 Adjusted difference: 0.01; p = 0.701	Probability of CCC being cost-effective: 0.95 for WTP of €200/DFD; 0.45 for WTP of €50,000/QALY	Cost year: 2013 Time horizon: 12 months Discounting: NA Applicability: partially applicable Quality: minor limitations

Table 44: Economic evidence table for stepped care

Study Country Study type	Intervention details	Study population Study design Data sources	Costs and outcomes: description and values	Results: Cost-effectiveness	Comments
Mukuria 2013 UK Cost effectiveness and cost-utility analysis	Interventions: Stepped care approach: Improving Access to Psychological Therapies (IAPT) service comprising: Step 1 watchful waiting; Step 2 guided self-help including bibliotherapy with support, computerised CBT (cCBT) with support and CBT-based telephone support for problem-solving; Step 3 CBT ± medication. IAPT was provided in addition to treatment as usual (TAU) TAU alone, comprising GP care, primary care counselling and referral to mental health professionals in secondary care. IAPT was evaluated in Doncaster demonstration site. Comparator sites were selected to match IAPT site regarding size & type of population served based on	People 16-64 years old with a new or recurrent episode of depression or anxiety, who were likely to benefit from psychological therapies. More than 95% of people in IAPT had a primary diagnosis of depression by their GP. Prospective cohort study with matched sites (N=403) Source of efficacy and resource use data: cohort study (N=403; available 8-month cost and QALY data for n=297) Source of unit costs: IAPT data and national sources	Costs: intervention (staff time, training, equipment, facilities and overheads), other mental healthcare (psychiatrist, psychologist, community psychiatric nurse, psychotherapist/ counsellor, other mental health professionals and voluntary sector services), primary and secondary care, social care; medication costs not considered Mean total cost per person (SD): IAPT: £1,190 (£2,193); TAU: £934 (£1,666) Unadjusted difference: £256 (95% CI: -£266 to £779) Adjusted difference: £236 (95%CI: -£214 to £689) Primary outcome measures: proportion of people with a reliable and clinically significant (RCS) improvement on the PHQ-9; QALY based on SF-6D ratings (UK tariff); QALYs based on predicted EQ-5D ratings (UK tariff), estimated from SF-6D using an empirical mapping function were used in sensitivity analysis Proportion of people with a PHQ-9 RCS significant improvement (95% CI):	ICER of IAPT vs. TAU £9,440 per participant with RCS improvement £29,500/QALY using SF-6D £16,857/QALY using predicted EQ-5D scores Probability of IAPT being cost-effective using SF-6D QALYs: <0.40 at WTP £30,000/QALY; using EQ-5D QALYs: 0.38 and 0.53 at WTP £20,000 and £30,000 / QALY, respectively. Using national unit costs instead of IAPT financial data resulted in an ICER of £3,800 per participant achieving RCS improvement and £11,875/QALY using SF-6D	Perspective: NHS and social services; productivity losses estimated separately Currency: GBP£ Cost year: 2008/09 Time horizon: 8 months Discounting: NA Applicability: directly applicable Quality: potentially serious limitations

Study Country Study type	Intervention details	Study population Study design Data sources	Costs and outcomes: description and values	Results: Cost-effectiveness	Comments
	deprivation, ethnicity and age; geographical location; local implementation of 'pathways to work'; ethnic diversity; recent changes in organisational structure. Also, comparator sites were selected based on how well they performed according to average Quality and Outcomes Framework points, a voluntary annual reward and incentive programme for all GPs in England that assesses areas of clinical care, organisation, patient experience & other services.		IAPT: 0.221 (0.164 to 0.278) TAU: 0.205 (0.116 to 0.293) Unadjusted difference: 0.016 (-0.089 to 0.122) Adjusted difference: 0.025 (-0.078 to 0.127) Mean number of SF-6D QALYs per person (95% CI): IAPT: 0.026 (0.018 to 0.033) TAU: 0.018 (0.007 to 0.029) Unadjusted difference 0.007 (-0.006 to 0.021) Adjusted difference 0.008 (-0.005 to 0.021) Mean number of EQ-5D QALYs per person (95% CI): IAPT: 0.038 (0.027 to 0.049) TAU: 0.025 (0.009 to 0.040) Unadjusted difference: 0.013 (-0.007 to 0.033) Adjusted difference: 0.014 (-0.005 to 0.032)		
Meeuwissen 2019 The Netherlands Cost-utility analysis	Interventions: Stepped care (SC) comprising a standardised stepwise treatment algorithm for mild or moderate/ severe depression; basic interventions (psychoeducation, active monitoring, structuring of the day) offered to all; self-help	Adults with mild, moderate or severe major depression without psychotic symptoms. Decision-analytic modelling Source of efficacy data: literature review Source of resource use data: published literature	Costs: health professional time (GP, psychologist, psychiatrist, psychotherapist, social worker, nurse), antidepressants, telephone consultation, self-help book or information leaflet, group therapy, crisis intervention, inpatient care, day care, homecare, other out-patient care Mean incremental cost/person:	ICER: Mild depression: SC dominant Moderate/severe depression: €3,166/QALY Probability of SC being dominant: Mild depression: 0.67	Perspective: healthcare Currency: Euro (€) Cost year: 2017 Time horizon: 5 years Discounting: 4% or costs, 1.5% for outcomes Applicability: partially applicable

Study Country Study type	Intervention details	Study population Study design Data sources	Costs and outcomes: description and values	Results: Cost-effectiveness	Comments
	may be added according to patient preference Treatment as usual (TAU) comprising all commonly available treatments in the health care system, often delivered in a mix of care .	(clinical trials and empirical studies) Source of unit costs: possibly national sources	Mild depression: -€36.72 Moderate/severe depression: €46.96 Primary outcome measure: QALY; effect size transformed into a utility increment. Mean incremental QALY/person: Mild depression: 0.014 Moderate/severe depression: 0.015	Moderate/severe depression: 0.33 Probability of SC being cost-effective at €20,000/QALY: >0.95 for both mild and moderate/ severe depression	Quality: minor limitations
Van Der Weele 2012 The Netherlands Cost effectiveness and cost-utility analysis	Interventions: Stepped care (SC) comprising step 1 individual counselling concerning treatment needs and motivation of the subjects during 1-2 home visits by a community psychiatric nurse; step 2 'Coping with Depression' course, based on CBT, by trained mental health professionals; if indicated, step 3 referral back to GP to discuss further treatment. Treatment as usual (TAU); GPs and participants in control	Adults ≥75 years old who screened positive for depressive symptoms in general practice, according to a ≥5 points score on an interviewer-administered 15-item version of the Geriatric Depression Scale (GDS-15) Exclusion criteria: current treatment for depression, clinical diagnosis of dementia or a Mini-Mental State Examination (MMSE) score <19, loss of partner or child in the preceding 3 months, life expectancy ≤3 months and not speaking Dutch.	Costs: intervention (individual consultation, course sessions, course instructors, room rental, refreshments, course materials), staff time (psychiatrist, psychologist, GP, physiotherapist), medication, hospitalisation (psychiatric & general), hospital day care, specialist care, paramedical care; service user costs (time & travel), informal care Mean healthcare cost per person: 75-79 years: SC €10,199, TAU €7,816 ≥80 years: SC €14,097, TAU €14,518 Mean total cost per person:	Under a healthcare perspective: 75-79 years: SC dominated using EQ-5D QALY ICER of SC vs. TAU €297,838/QALY using SF-6D ≥80 years: SC dominant using either EQ-5D or SF-6D QALY	Perspective: healthcare plus service user and informal care costs considered Currency: Euro (€) Cost year: likely 2004 Time horizon: 12 months Discounting: NA Applicability: partially applicable Quality: potentially serious limitations

Study Country Study type	Intervention details	Study population Study design Data sources	Costs and outcomes: description and values	Results: Cost-effectiveness	Comments
	arm were not informed about screen-positive results before the end of the study, except in case of a MADRS score >30 and/or suicidal ideation	Pragmatic cluster RCT (N=239) Source of efficacy and resource use data: RCT (Van Der Weele2012, N=239; completers n=194) Source of unit costs: national sources	75-79 years: SC €14,026, TAU €9,353; p=0.10 ≥80 years: SC €16,087, TAU €16,661; p=0.87 Primary outcome measures: MADRS change score, QALY based on EQ-5D and SF-6D ratings (UK tariff) Mean MADRS change score (SE): SC -3.1 (0.61); TAU: -4.6 (0.64); p=0.084 Mean EQ-5D QALYS per person: 75-79 years: SC 0.404; TAU 0.429; p=0.66 ≥80 years: SC 0.350; TAU 0.303; p=0.36 Mean SF-6D QALYs per person: 75-79 year: SC 0.624; TAU 0.616; p=0.78 ≥80 years: SC 0.588; TAU 0.568; p=0.46		
Health Quality Ontario 2019 Cost-utility analysis	Analysis A: Stepped care (SC1) comprising computerised CBT (cCBT) with support followed by individual CBT	Analysis A: adults with mild to moderate major depression Analysis B: adults with mild to moderate major depression who are likely to drop out of treatment	Costs: intervention (health professional time, training and supervision, equipment), assessment, medication, follow-up care with GP, psychiatrist time Mean cost/person:	Analysis A: SC dominant over TAU. ICER of SC1 vs SC2: $1,098/QALY. Results robust to change in efficacy, dropout rates, utilities,	Perspective: healthcare and long term care Currency: Can$ Cost year: 2018 Time horizon: Analysis A: lifetime

Study Country Study type	Intervention details	Study population Study design Data sources	Costs and outcomes: description and values	Results: Cost-effectiveness	Comments
	Stepped care (SC2) comprising cCBT with support followed by group CBT Treatment as usual (TAU) Analysis B: Stepped care (SC) comprising cCBT without support followed by cCBT with support Individual CBT Group CBT TAU	Decision-analytic modelling Source of efficacy data: systematic literature review Source of resource use data: published literature and expert opinion	Analysis A: SC1: $280,538; SC2: $280,498 TAU: $283,651 Analysis B: SC $715; group CBT $1,690; individual CBT $2,654; TAU $409 Primary outcome measure: QALY; utility data from literature review, ratings of various scales. Mean QALY/person: Analysis A: SC1: 18.33; SC2: 18.30; TAU: 18.09 Analysis B: SC 0.80; group CBT 0.82; individual CBT 0.83; TAU 0.79	medication costs, time horizon. Probability of SC1 being cost-effective at $50,000/QALY: 0.60 Analysis B ICERs: Indiv CBT vs group CBT: $100,316/QALY Group CBT vs SC: $67,161/QALY SC vs TAU: $19,454/QALY Probability of SC being cost-effective at $50,000/QALY: 0.48	Analysis B: 1 year Discounting: 1.5% for costs and outcomes Applicability: partially applicable Quality: minor limitations

Table 45: Economic evidence table for medication management

Study Country Study type	Intervention details	Study population Study design Data sources	Costs and outcomes: description and values	Results: Cost-effectiveness	Comments
Rubio-Valera 2013 Spain Cost effectiveness and cost-utility analysis	Interventions: Medication management (MM), comprising an educational intervention provided by the pharmacist, focusing on improving service users' knowledge of antidepressant medication, making them aware of the importance of compliance to the medication, reassuring them about possible side-effects, and stressing the importance of carrying out GPs' advice. In service users with a sceptical attitude towards antidepressants, the intervention aimed to reduce stigma. Pharmacists were trained for the intervention. Treatment as usual from GP and pharmacist (TAU), comprising filling the	Adults aged 18-75 years initiating treatment with antidepressants because of depression RCT (N=179) Source of efficacy and resource use data: RCT (Rubio-Valera 2013, N=179; 71% completed at 6 months; n=151 received intervention as allocated) Source of unit costs: regional sources	Costs: intervention (pharmacist time, pharmacist training), publicly funded healthcare services (GP, nurse, psychologist, psychiatrist, other medical specialists, social worker, hospital emergency visits, hospital stay, diagnostic tests, medication), privately funded healthcare services (psychiatrist, psychologist, medical specialist, GP), absenteeism from paid labour. Mean societal cost per person: MM: €1,091; TAU: €767 Mean difference €324 (95%CI –€97 to €745). Mean direct cost per person: MM: €444; TAU: €425 Mean difference €49 (95%CI not reported). Primary outcome measures: adherence to antidepressant treatment measured using electronic pharmacy records; remission of depressive symptoms defined as a reduction in the Patient Health Questionnaire 9-item (PHQ-9) of at least 50%; QALYs based on EQ-5D ratings (Spanish tariff)	Under a healthcare perspective: ICER of MM vs. TAU €962 per extra adherent service user €3,592/QALY TAU dominant in terms of remission Probability of MM being cost-effective 0.71 and 0.76 for WTP €6,000 /adherent service user and €30,000 /QALY, respectively. Using remission, maximum probability of MM being cost-effective 0.46. Results robust to per protocol or complete case analysis, use of DSM-IV criteria for depression, intervention costs or method for estimating indirect costs.	Perspective: societal and healthcare Currency: Euro (€) Cost year: 2009 Time horizon: 6 months Discounting: NA Applicability: partially applicable Quality: potentially serious limitations

Study Country Study type	Intervention details	Study population Study design Data sources	Costs and outcomes: description and values	Results: Cost-effectiveness	Comments
	prescriptions, addressing service users' questions about medication and giving basic advice about how to take the antidepressant.		Incremental probability of adherence per person: 0.04 (95%CI -0.2 to 0.1) Incremental probability of remission per person: -0.01 (95%CI -0.2 to 0.1) Incremental QALYs per person: 0.01 (95%CI -0.02 to 0.03)		

Table 46: Economic evidence table for shared care

Study Country Study type	Intervention details	Study population Study design Data sources	Costs and outcomes: description and values	Results: Cost-effectiveness	Comments
Wiley-Exley 2009 US Cost effectiveness and cost-utility analysis	Interventions: Integrated (shared) care (IC) comprising collaboration between primary and specialty mental health care; a behavioural health professional was co-located in the primary care setting and the primary care provider continued involvement in the mental health care of the service user Primary care with a specialty referral system (SRS) for referral to a behavioural	Adults above 65 years of age with depression (major or minor) Multi-site pragmatic RCT (N=840) Source of efficacy and resource use data: RCT (populations with various conditions. Subgroup with depression: N=840; within VA n=365, outside VA n=475; individuals with major depression within VA n=214, outside VA n=302) Source of unit costs: national sources	Costs: outpatient visits, inpatient care, nursing home, rehabilitation, emergency room, medication, service users' and caregivers' time and travel costs. Adjusted incremental total cost per person: All: VA: -$651, p=ns; Non-VA: $46, p=ns Major depression: VA: $877, p=ns; Non-VA: -$380, p=ns Primary outcome measures: Center for Epidemiologic Studies Depression Scale (CES-D) score; number of depression-free days (DFD) derived from the 20-item CES-D (score =0 indicated depression-free day, ≥ 16 full	Full VA sample: IC is dominant Probability of IC being cost-effective >0.70 for any WTP/QALY-SF Full non-VA sample: IC is dominated when using CES-D, DFD, QALY-DFD. When using QALY-SF, ICER of IC vs. SRS was $94,929/QALY Probability of IC being cost-effective <0.40 for any WTP/QALY-SF	Perspective: healthcare & service users' and carers' time and travel costs Currency: US$ Cost year: 2002 Time horizon: 6 months Discounting: NA Applicability: partially applicable Quality: potentially serious limitations

Study Country Study type	Intervention details	Study population Study design Data sources	Costs and outcomes: description and values	Results: Cost-effectiveness	Comments
	health provider outside the primary care setting, who had primary responsibility for the mental health needs of the service user. Both service delivery models were assessed within and outside the Veteran Affairs (VA) system.		symptoms and intermediate severity scores were assigned a value between depression-free and fully symptomatic by linear interpolation); QALYs estimated based on depression-free days (QALY-DFD), using utility weights of health=1, depression=0.59); QALYs estimated based on SF-36 (QALY-SF), using preferences for matched vignettes created following cluster analysis of SF-12 mental and physical component scores, elicited by US service users with depression using SG <u>Adjusted incremental CES-D score per person:</u> All: VA: -1.3, p=ns; Non-VA: 2.9, p<0.01 Major depression: VA: -2.8, p<0.05; Non-VA: 3.45, p<0.05 <u>Adjusted incremental DFDs per person:</u> All: VA: 3.89, p=ns; Non-VA: -5.73, p=ns Major depression: VA: 9.29, p=ns; Non-VA: -5.20, p<0.05 <u>Adjusted incremental QALY-DFD per person:</u> All: VA: 0.005, p=ns; Non-VA: -0.016, p<0.05 Major depression: VA: 0.019,	Major depression VA sample: ICER of IC vs. SRS: • $322/CES-D point change • $94/DFD • $45,965/QALY-DFD • $58,815/QALY-SF Probability of IC being cost-effective <0.50 for WTP of $40,000/QALY-SF and above Major depression non-VA sample: SRS is dominant in terms of CES-D ICER of SRS vs. IC: • $73/DFD • $34,167/QALY-DFD • $79,590/QALY-SF Probability of IC being cost-effective >0.50 for WTP $50,000/QALY-SF and above	

Study Country Study type	Intervention details	Study population Study design Data sources	Costs and outcomes: description and values	Results: Cost-effectiveness	Comments
			p=ns; Non-VA: -0.011, p<0.05 <u>Adjusted incremental QALY-SF per person:</u> All: VA: 0.007, p=ns; Non-VA: 0.0004, p=ns Major depression: VA: 0.015, p=ns; Non-VA: -0.005, p=ns		

Economic evidence tables for review question 1.2 For adults with depression, what are the relative benefits and harms associated with different settings for the delivery of care?

No economic evidence was identified which was applicable to this review question.

Appendix I – Economic evidence profiles

Economic evidence profiles for review question 1.1 For adults with depression, what are the relative benefits and harms associated with different models for the coordination and delivery of services?

Collaborative care

Table 47: Economic evidence profile for simple collaborative care alone or in addition to standard care versus standard care

Simple collaborative care alone or in addition to standard care versus standard care for adults with depression							
Study and country	**Limitations**	**Applicability**	**Other comments**	**Incremental cost (£)[1]**	**Incremental effect**	**ICER (£/effect)[1]**	**Uncertainty[1]**
Bosanquet 2017 UK	Potentially serious limitations[2]	Directly applicable[3]	Older adults Outcome: QALY	£531	0.019	£28,765	Probability of intervention being cost-effective: 0.39 and 0.55 at WTP £20,000 and £30,000/QALY, respectively. Including only participants who engaged with 5 or more sessions in the analysis, ICER fell at £10,922/QALY
Green 2014 UK	Minor limitations[4]	Directly applicable[5]	Outcome: QALY	£311	0.019	£16,361	Probability of intervention being cost-effective: 0.58 and 0.65 at WTP £20,000 and £30,000/QALY, respectively Results robust to multiple imputation of missing data, use of SF-6D utility values, use of alternative intervention costs
Lewis 2017 UK	Potentially serious limitations[6]	Directly applicable[7]	Older adults Outcome: QALY	£465	0.044	£10,653	Probability of intervention being cost-effective: 0.92 and 0.97 at WTP £20,000 and £30,000/QALY, respectively. Accounting for the true observed intervention contact rate (rather than the expected that was used in the base-case analysis), ICER fell at £3,681/QALY

ICER: incremental cost-effectiveness ratio; QALY: quality-adjusted life years; WTP: willingness to pay
1. Costs uplifted to 2020 UK pounds using the NHS cost inflation index (Curtis 2020).
2. Time horizon 18 months; analysis conducted alongside RCT (N=485; at 18 months n=344; cost data available for n=447); national unit costs used; statistical analyses conducted; CEACs presented; consideration of intervention and primary care costs only
3. UK study; NHS & PSS perspective; QALY estimates based on SF-6D (UK tariff)

Simple collaborative care alone or in addition to standard care versus standard care for adults with depression
4. Time horizon 12 months; analysis conducted alongside RCT (N=581; data available for cost analysis n=447); national unit costs used; statistical analyses conducted; CEACs presented. *5. UK study; NHS & PSS perspective; QALY estimates based on EQ-5D (UK tariff)* *6. Time horizon 12 months; analysis conducted alongside RCT (N=705; complete data used in base-case economic analysis n=448); national unit costs used; statistical analyses conducted; CEACs presented; high attrition that was markedly greater in the collaborative care arm; consideration of intervention and primary care costs only* *7. UK study; NHS & PSS perspective; QALY estimates based on EQ-5D (UK tariff)*

Table 48: Economic evidence profile for simple collaborative care for relapse prevention versus standard care

Simple collaborative care for relapse prevention versus standard care							
Study and country	**Limitations**	**Applicability**	**Other comments**	**Incremental cost (£)[1]**	**Incremental effect**	**ICER (£/effect)[1]**	**Uncertainty[1]**
Simon 2002 US	Potentially serious limitations[2]	Partially applicable[3]	Adults with recurrent depression Outcome: number of depression-free days (days with a Hopkins Symptoms Checklist (HSCL) depression score ≤ 0.5; days with a HSCL score above 0.5 but < 2 considered 50% depression free)	£15	13.9	£1	ICER 95% CI: -£155 to £399
ICER: incremental cost-effectiveness ratio *1. Costs converted and uplifted to 2020 UK pounds using purchasing power parity (PPP) exchange rates and the NHS cost inflation index (Curtis 2020).* *2. Time horizon 12 months; analysis conducted alongside RCT (N=386, n=377 used for cost analysis and n=315 used for clinical analysis); local prices used; statistical analyses conducted, including bootstrapping; analyses of clinical data included only those completing all blinded follow-up assessments; cost analyses included only those remaining enrolled throughout the follow-up period; participation in follow-up interviews was significantly greater in the intervention group than in usual care, introducing a possibility of bias.* *3. US study; 3rd party payer perspective; no QALYs estimated*							

Table 49: Economic evidence profile for complex collaborative care alone or in addition to standard care versus standard care

Complex collaborative care alone or in addition to standard care versus standard care							
Study and country	**Limitations**	**Applicability**	**Other comments**	**Incremental cost (£)[1]**	**Incremental effect**	**ICER (£/effect)[1]**	**Uncertainty[1]**
Morriss 2016 UK	Minor limitations[2]	Directly applicable[3]	Adults with persistent depression Outcome: QALY	£3,770	0.079	£47,690	Controlling for baseline differences and cluster effects: probability of complex collaborative care being

Complex collaborative care alone or in addition to standard care							
							cost-effective exceeds 0.50 at WTP of £45,500/QALY
Goorden 2015 The Netherlands	Potentially serious limitations[4]	Partially applicable[5]	Primary care setting Outcome: QALY	£1,181	0.02	£54,087	Probability of CCC being cost-effective: 0.20 and 0.70 at WTP £20,100 and £80,500/QALY, respectively.
Grochtdreis 2019 Germany	Minor limitations[6]	Partially applicable[7]	Older adults with late-life depression Primary care setting Outcome: Number of depression-free days (DFDs) and QALY	£561	21.4 DFDs 0.01 QALYs	£26/DFD £56,184/QALY	Probability of CCC being cost-effective: 0.95 for WTP of £204/DFD; 0.45 for WTP of £50,400/QALY

ICER: incremental cost-effectiveness ratio; QALY: quality-adjusted life years; WTP: willingness to pay
1. Costs converted and uplifted to 2020 UK pounds using purchasing power parity (PPP) exchange rates and the NHS cost inflation index (Curtis 2020).
2. Time horizon 18 months; analysis conducted alongside RCT (N=187; 84% completed at 6 months, 72% at 12 months and 59% at 18 months); national unit costs used; statistical analyses conducted; CEACs presented.
3. UK study; NHS & PSS perspective; QALY estimates based on EQ-5D (UK tariff)
4. Time horizon 12 months; analysis conducted alongside RCT (N=150; 93 identified by screening and 47 by GP referral; economic analysis based only on n=93 identified by screening); national unit costs used; CEACs presented
5. Dutch study; healthcare system perspective; QALY based on EQ-5D ratings but Dutch tariff
6.Time horizon 12 months; analysis conducted alongside RCT (N=246); national unit costs used; CEACs presented
7. German study; healthcare system perspective; QALY based on EQ-5D ratings and UK tariff

Stepped care

Table 50: Economic evidence profile for stepped care (± TAU) versus TAU

Stepped care (± TAU) versus TAU							
Study and country	**Limitations**	**Applicability**	**Other comments**	**Incremental cost (£)[1]**	**Incremental effect**	**ICER (£/effect)[1]**	**Uncertainty[1]**
Mukuria 2013 UK	Potentially serious limitations[2]	Directly applicable[3]	IAPT setting Outcomes: • proportion with reliable and clinically significant improvement on PHQ-9 • QALY - SF-6D (UK tariff)	£281	0.025 0.008 0.014	£11,234/ improved participant £35,106/QALY (SF-6D)	Probability of IAPT being cost-effective using SF-6D QALYs: <0.40 at WTP £30,000/QALY; using EQ-5D QALYs: 0.38 and 0.53 at WTP £20,000 and £30,000/QALY, respectively.

Stepped care (± TAU) versus TAU							
			• QALY - predicted EQ-5D (UK tariff), estimated from SF-6D using empirical mapping			£20,059/QALY (predicted EQ-5D)	Using national unit costs instead of IAPT financial data: £4,522/improved participant; £14,132/QALY using SF-6D
Meeuwissen 2019 The Netherlands	Minor limitations[4]	Partially applicable[5]	Outcome: QALY Separate analysis for mild depression and for moderate/severe depression	Mild: -£37 Moderate /severe: £47	Mild: 0.014 Moderate /severe: 0.015	Mild: dominant Moderate /severe: £3,159	Probability of intervention being dominant: Mild: 0.67; Moderate/severe: 0.33 Probability of intervention being cost-effective at £20,000/QALY: >0.95 for both Mild and Moderate/ severe
Van Der Weele 2012 The Netherlands	Potentially serious limitations[6]	Partially applicable[7]	Outcome: QALY Separate analysis for people aged 75-79 years on those ≥80 years	75-79 years: £2,133 ≥80 years: -£378	75-79 years: -0.025 ≥80 years: 0.047	75-79 years: SC dominated ≥80 years: SC dominant	No statistically significant differences in costs or outcomes
Health Quality Ontario 2019	Minor limitations[8]	Partially applicable[9]	Analysis A: adults with mild-to-moderate depression Interventions: SC1 comprising cCBT with support followed by individual CBT; SC2 comprising cCBT with support followed by group CBT; TAU Analysis B: adults with mild-to-moderate depression likely to drop out of treatment Interventions: SC comprising cCBT without support followed by cCBT with support;	Analysis A: Vs TAU: SC1: -£1,868; SC2: -£1,892 Analysis B: Vs TAU: SC: £183; group CBT: £769; individual CBT £1,346	Analysis A: SC1: 18.33; SC2: 18.30; TAU: 18.09 Analysis B: SC 0.80; group CBT 0.82; individual CBT 0.83; TAU 0.79	Analysis A: SC dominant over TAU; ICER of SC1 vs SC2: £659/QALY. Analysis B ICERs: Indiv CBT vs group CBT: £60,157/QALY Group CBT vs SC: £40,275/QALY SC vs TAU: £11,666/QALY	Analysis A: Results robust to change in efficacy, dropout rates, utilities, medication costs, time horizon. Probability of SC1 being cost-effective at £30,000/QALY: 0.60 Analysis B: Probability of SC being cost-effective at £30,000/QALY: 0.48

Stepped care (± TAU) versus TAU							
			individual CBT; group CBT; TAU				

cCBT: computerised Cognitive Behavioural therapy; CBT: cognitive behavioural therapy; ICER: incremental cost-effectiveness ratio; QALY: quality-adjusted life years; SC: stepped care; TAU: treatment as usual; WTP: willingness to pay
1. Costs converted and uplifted to 2020 UK pounds using PPP exchange rates and the NHS cost inflation index (Curtis 2020).
2. Time horizon 8 months; prospective cohort study with matched sites (N=403); low response rate at recruitment (403/3,391, 11.9%); IAPT service was assessed over the first 2 years of establishment, therefore costs associated with learning effects were likely; IAPT financial data used – results sensitive to the use of national unit costs; CEACs presented.
3. UK; NHS and social service perspective; QALY based on SF-6D (UK tariff); QALYs based on predicted EQ-5D ratings (UK tariff), estimated from SF-6D using an empirical mapping function, used in sensitivity analysis
4. Time horizon 5 years; modelling study; efficacy data from a guideline literature review; all relevant costs considered; CEAC presented; likely national unit costs used
5. Dutch study; healthcare perspective; QALYs estimated from translating effect size into utility increment
6. Time horizon 12 months; analysis based on cluster RCT (N=239); national unit costs used; statistical analyses conducted around differences in outcomes and costs; results not synthesised in ICERs therefore uncertainty in ICER not reported and not possible to estimate
7. Dutch study; healthcare perspective; QALYs based on EQ-5D (UK tariff) and SF-6D
8. Time horizon (A) lifetime and (B) 1 year; modelling study; efficacy data from a systematic literature review; all relevant costs considered; CEAC presented; national unit costs used
9. Canadian study; healthcare and long term care perspective; QALYs estimated using utility values from literature review – various scales used for rating of health-related quality of life

Medication management

Table 51: Economic evidence profile for medication management in addition to standard care versus standard care

Medication management in addition to standard care versus standard care							
Study and country	**Limitations**	**Applicability**	**Other comments**	**Incremental cost (£)[1]**	**Incremental effect**	**ICER (£/effect)[1]**	**Uncertainty[1]**
Rubio-Valera 2013 Spain	Potentially serious limitations[2]	Partially applicable[3]	Outcomes: Adherence; Remission; QALY	£45	0.04 -0.01 0.01	£935/extra adherence Dominated using remission as an outcome £3,495/QALY	Probability of intervention being cost-effective 0.71 and 0.76 for WTP £5,800 /adherent service user and £29,000/QALY, respectively. Using remission, maximum probability of intervention being cost-effective was 0.46

ICER: incremental cost-effectiveness ratio; QALY: quality-adjusted life years; WTP: willingness to pay
1. Costs converted and uplifted to 2020 UK pounds using PPP exchange rates and the NHS cost inflation index (Curtis 2020).
2. Time horizon 6 months; analysis conducted alongside RCT (N=179; 71% completed at 6 months; n=151 received intervention as allocated); regional unit costs used; CEACs presented; contradictory results depending on the outcome measure used
3. Spanish study; healthcare perspective; QALYs based on EQ-5D ratings, Spanish tariff

Integrated (shared) care

Table 52: Economic evidence profile for integrated (shared) care versus primary care with referral system to specialist care

Integrated (shared) care versus primary care with referral system to specialist care							
Study and country	**Limitations**	**Applicability**	**Other comments**	**Incremental cost (£)[1]**	**Incremental effect**	**ICER (£/effect)[1]**	**Uncertainty[1]**
Wiley-Exley 2009 US	Potentially serious limitations[2]	Partially applicable[3]	Separate analyses for: • Full (major and minor depression) VA sample • Full non-VA sample • Major depression VA sample • Major depression non-VA sample Outcomes used: CES-D score; number of depression-free days derived from CES-D; QALYs estimated based on depression-free days, using utility weights of health=1, depression=0.59; QALYs estimated based on SF-36, using preferences for matched vignettes created following cluster analysis of SF-12 mental and physical component scores, elicited by US service users with depression using SG. Only results for the latter presented here.	-£629 £44 £847 -£367	0.007 0.0004 0.015 -0.005	Dominant £91,674/QALY £56,799/QALY £76,861/QALY (less effective, less costly)	Probability of IC being cost-effective: >0.70 for any WTP/QALY <0.40 for any WTP/QALY <0.50 for WTP of £38,500/QALY and above >0.50 for WTP £48,200/QALY and above

ICER: incremental cost-effectiveness ratio; QALY: quality-adjusted life years; WTP: willingness to pay
1. Costs converted and uplifted to 2020 UK pounds using PPP exchange rates and the NHS cost inflation index (Curtis 2020).
2. Time horizon 6 months; analysis conducted alongside multi-site pragmatic RCT (N=840 with major or minor depression, assessed within and outside the Veteran Affairs (VA) system.; within VA n=365, outside VA n=475; individuals with major depression within VA n=214, outside VA n=302); national unit costs; bootstrapping conducted, CEACs presented
3. US study; health care provider perspective including service users' time and mileage; QALYs based on SF-36, using preferences for matched vignettes created following cluster analysis of SF-12 mental and physical component scores, elicited by US service users with depression using SG.

Economic evidence profiles for review question 1.2 For adults with depression, what are the relative benefits and harms associated with different settings for the delivery of care?

No economic evidence was identified which was applicable to this review question.

Appendix J – Economic analysis

Economic evidence analysis for review question 1.1 For adults with depression, what are the relative benefits and harms associated with different models for the coordination and delivery of services?

No economic analysis was conducted for this review question.

Economic evidence analysis for review question 1.2 For adults with depression, what are the relative benefits and harms associated with different settings for the delivery of care?

No economic analysis was conducted for this review question.

Appendix K – Excluded studies

Excluded clinical and economic studies for review question 1.1 For adults with depression, what are the relative benefits and harms associated with different models for the coordination and delivery of services?

Clinical studies

Please refer to the excluded studies in supplement A1 – Clinical evidence tables for review 1.1

Economic studies

Please refer to supplement 3 - Economic evidence included & excluded studies.

Excluded clinical and economic studies for review question 1.2 For adults with depression, what are the relative benefits and harms associated with different settings for the delivery of care?

Clinical studies

Please refer to the excluded studies in supplement A2 – Clinical evidence tables for review 1.2

Economic studies

Please refer to supplement 3 - Economic evidence included & excluded studies.

Appendix L – Research recommendations

Research recommendations for review question 1.1 For adults with depression, what are the relative benefits and harms associated with different models for the coordination and delivery of services?

No research recommendations were made for this review question.

Research recommendations for review question 1.2 For adults with depression, what are the relative benefits and harms associated with different settings for the delivery of care?

No research recommendations were made for this review question.