Comparative efficacy and acceptability of antidepressants, psychotherapies, and their combination for acute treatment of children and adolescents with depressive disorder: a systematic review and network meta-analysis

Xinyu Zhou, Teng Teng*, Yuqing Zhang*, Cinzia Del Giovane, Toshi A Furukawa, John R Weisz, Xuemei Li, Pim Cuijpers, David Coghill, Yajie Xiang, Sarah E Hetrick, Stefan Leucht, Mengchang Qin, Jürgen Barth, Arun V Ravindran, Lining Yang, John Curry, Li Fan, Susan G Silva, Andrea Ciprianit, Peng Xie†*

Summary

Background Depressive disorders are common in children and adolescents. Antidepressants, psychotherapies, and their combination are often used in routine clinical practice; however, available evidence on the comparative efficacy and safety of these interventions is inconclusive. Therefore, we sought to compare and rank all available treatment interventions for the acute treatment of depressive disorders in children and adolescents.

Methods We did a systematic review and network meta-analysis. We searched PubMed, Embase, the Cochrane Central Register of Controlled Trials, Web of Science, PsycINFO, ProQuest, CINAHL, LiLACS, international trial registries, and the websites of regulatory agencies for published and unpublished randomised controlled trials from database inception until Jan 1, 2019. We included placebo-controlled and head-to-head trials of 16 antidepressants, seven psychotherapies, and five combinations of antidepressant and psychotherapy that are used for the acute treatment of children and adolescents (≤18 years old and of both sexes) with depressive disorder diagnosed according to standard operationalised criteria. Trials recruiting participants with treatment-resistant depression, bipolar disorder, psychotic depression, treatment duration of less than 4 weeks, or an overall sample size of fewer than ten patients were excluded. We extracted data following a predefined hierarchy of outcome measures, and assessed risk of bias and certainty of evidence using validated methods. Primary outcomes were efficacy (change in depressive symptoms) and acceptability (treatment discontinuation due to any cause). We estimated summary standardised mean differences (SMDs) or odds ratios (ORs) with credible intervals (CrIs) using network meta-analysis with random effects. This study was registered with PROSPERO, number CRD42015020841.

Findings From 20 366 publications, we included 71 trials (9510 participants). Depressive disorders in most studies were moderate to severe. In terms of efficacy, fluoxetine plus cognitive behavioural therapy (CBT) was more effective than CBT alone (–0·78, 95% CrI –1·55 to –0·01) and psychodynamic therapy (–1·14, –2·20 to –0·08), but not more effective than fluoxetine alone (–0·22, –0·86 to 0·42). No pharmacotherapy alone was more effective than psychotherapy alone. Only fluoxetine plus CBT and fluoxetine were significantly more effective than pill placebo or psychological controls (SMDs ranged from –1·73 to –0·51); and only interpersonal therapy was more effective than all psychological controls (–1·37 to –0·66). Nortriptyline (SMDs ranged from 1·04 to 2·22) and waiting list (SMDs ranged from 0·67 to 2·08) were less effective than most active interventions. In terms of acceptability, nefazodone and fluoxetine were associated with fewer dropouts than sertraline, imipramine, and desipramine (ORs ranged from 0·17 to 0·50); imipramine was associated with more dropouts than pill placebo, desvenlafaxine, fluoxetine plus CBT, and vilazodone (2·51 to 5·06). Most of the results were rated as "low" to "very low" in terms of confidence of evidence according to Confidence In Network Meta-Analysis.

Interpretation Despite the scarcity of high-quality evidence, fluoxetine (alone or in combination with CBT) seems to be the best choice for the acute treatment of moderate-to-severe depressive disorder in children and adolescents. However, the effects of these interventions might vary between individuals, so patients, carers, and clinicians should carefully balance the risk-benefit profile of efficacy, acceptability, and suicide risk of all active interventions in young patients with depression on a case-by-case basis.

Funding National Key Research and Development Program of China.

Copyright © 2020 Elsevier Ltd. All rights reserved.

Lancet Psychiatry 2020; 7: 581–601

See Comment page 562

*These authors contributed equally

†These authors share the senior authorship

Department of Psychiatry (X Zhou PhD) **and Department of Neurology** (T Teng PhD, X Li MD, Y Xiang PhD, M Qin MD, L Yang MD, L Fan MD, P Xie MD), **The First Affiliated Hospital of Chongqing Medical University, Chongqing, China; Department of Neurology, The Second Affiliated Hospital of Chongqing Medical University, Chongqing, China** (Y Zhang PhD); **Institute of Primary Health Care, University of Bern, Bern, Switzerland** (C Del Giovane PhD); **Department of Health Promotion and Human Behavior, Kyoto University Graduate School of Medicine and School of Public Health, Kyoto, Japan** (T A Furukawa PhD); **Department of Psychology, Harvard University, Cambridge, MA, USA** (J R Weisz PhD); **Department of Clinical, Neuro and Developmental Psychology, Amsterdam Public Health Research Institute, Vrije Universiteit Amsterdam, Amsterdam, Netherlands** (P Cuijpers PhD); **Department of Paediatrics, University of Melbourne, Melbourne, VIC, Australia** (D Coghill MD); **Department of Psychological Medicine, University of Auckland, Auckland, New Zealand** (S E Hetrick DPsych); **Department of Psychiatry and Psychotherapy, School of Medicine, Technische Universität München, Munich, Germany** (S Leucht MD);

Introduction

Childhood and adolescence are risk periods for the development of psychiatric disorders, and major depressive disorder is a leading contributor to burden of disease in young people aged 10–24 years.[1] In England in 2017, major depressive disorder in children and

Institute for Complementary and Integrative Medicine, University Hospital Zurich and University of Zurich, Zurich, Switzerland (J Barth PhD); **Department of Psychiatry, University of Toronto and Division of Mood and Anxiety Disorders, Centre for Addiction and Mental Health, Toronto, ON, Canada** (A V Ravindran MD); **Department of Psychiatry, University of Toronto, Toronto, ON, Canada** (A V Ravindran); **Department of Psychiatry & Behavioral Sciences, Duke University School of Medicine, Durham, NC, USA** (J Curry PhD); **Duke University School of Nursing, Durham, NC, USA** (S G Silva PhD); **Department of Psychiatry, University of Oxford, Oxford, UK** (A Cipriani PhD); **and Oxford Health NHS Foundation Trust, Warneford Hospital, Oxford, UK** (A Cipriani)

Correspondence to:
Dr Peng Xie, Department of Neurology and Psychiatry, The First Affiliated Hospital of Chongqing Medical University, Chongqing, China
xiepeng973@126.com

Research in context

Evidence before this study

Antidepressants and psychotherapies are routinely used worldwide for the treatment of depressive disorder in children and adolescents. Several clinical practice guidelines recommend that psychotherapy should be considered as the first-line intervention for the management of depressive disorder in children and adolescents, whereas antidepressants are often reserved for more severe illness or when psychotherapy does not work or is not available. However, the evidence base has not been well established that psychotherapy is more effective and safer than antidepressants in the treatment of child and adolescent depressive disorder, and whether the combination of antidepressants and psychotherapies is more beneficial than antidepressants alone remains unknown. We searched for eligible trials of combinations of antidepressants and psychotherapy on PubMed, Embase, the Cochrane Central Register of Controlled Trials, Web of Science, PsycINFO, ProQuest, CINAHL, and LiLACS database for randomised controlled trials (RCTs) published from the date of their inception to Jan 1, 2019. Our two previous studies investigated the comparative efficacy and acceptability of 14 antidepressants and nine psychotherapies for depression. No network meta-analysis has examined the relative effects of psychotherapies, pharmacotherapies, and their combination in the treatment of depressive disorder in children and adolescents.

Added value of this study

Our study provides the first comprehensive systematic review and network meta-analysis of all available RCTs, comparing any active interventions (antidepressant, psychotherapy, and their combination) with another or control conditions for the acute treatment of depressive disorders in children and adolescents. Our findings suggest that, in terms of efficacy, only fluoxetine plus cognitive behavioural therapy and fluoxetine alone were more efficacious than pill placebo, psychological controls and some active treatments for the acute treatment of depressive disorder in children and adolescents. In terms of suicidality, our findings confirmed that venlafaxine is associated with an increased risk of suicidal behaviour or ideation compared with pill placebo and ten other interventions.

Implications of all the available evidence

Fluoxetine (alone or in combination with CBT) seems to be the best choice for the acute treatment of moderate-to-severe depressive disorder in children and adolescents but the quality of evidence is low. Patients, carers, and clinicians should carefully balance the risk-benefit profile of efficacy, acceptability, and suicide risk of all active interventions in young patients with depression on a case-by-case basis.

adolescents was common, with an estimated point prevalence of about 0·3% in children (5–10 years), 2·7% in younger adolescents (11–16 years), and 4·8% in older adolescents (17–19 years).[2] The course of this disorder is often characterised by heterogeneous symptoms (eg, irritability, aggressive behaviours, and school refusal), protracted episodes, frequent recurrence, and comorbid psychiatric disorders.[3] Young patients with depression have more serious impairments in social and educational functioning and have an increased risk of smoking, substance misuse, obesity, and suicide compared with adults with depression.[4] Moreover, depression is the second or third leading cause of death in adolescence.[4]

In the past two decades, pharmacological and psychological interventions have been widely used in the treatment of depressive disorder in children and adolescents worldwide.[5] In 2005–12, the prevalence of antidepressant use in children and adolescents increased from 1·3% to 1·6% in the USA and from 0·7% to 1·1% in the UK.[6] As the first-line treatment, psychotherapies, especially cognitive-behavioural therapy (CBT) and interpersonal psychotherapy, appeared to be more effective compared with psychological controls in previous meta-analyses.[7,8] The mean effects (standardised mean difference [SMD] –0·29) after treatment were more modest than those found for treatment of other youth problems, including anxiety (SMD –0·61), attention deficit hyperactivity disorder (SMD –0·34), and conduct-related problems and disorders (SMD –0·46).[9] Previous meta-analyses[10,11] have shown that antidepressants, except for fluoxetine, do not offer a clear advantage over pill placebo for many individuals, and some antidepressants might increase risk of suicidality. The mean effects of antidepressants for major depressive disorder compared with pill placebo (Hedges g 0·21 for selective serotonin reuptake inhibitor [SSRI] and 0·16 for serotonin–norepinephrine reuptake inhibitor [SNRI]) have been more modest than those found for treatment of other youth problems, including anxiety disorder (Hedges g 0·71 for SSRI and 0·41 for SNRI) and obsessive-compulsive disorder (Hedges g 0·39 for SSRI).[12]

Whether the combination of antidepressant and psychological interventions is more beneficial than antidepressants alone remains unclear.[13] The aim of this study was to synthesise all the available evidence on commonly used antidepressants, psychotherapies, and their combinations for the acute treatment of depressive disorder in children and adolescents.

Methods

Search strategy and selection criteria

In this systematic review and network meta-analysis, we updated the literature search from our two previous publications[7,10] for the identification of trials of antidepressants and psychotherapies monotherapy. We searched for eligible trials of combinations of antidepressants and

psychotherapy on PubMed, Embase, the Cochrane Central Register of Controlled Trials, Web of Science, PsycINFO, ProQuest, CINAHL, and LiLACS database for randomised controlled trials (RCTs) published from the date of their inception to Jan 1, 2019. We included studies comparing any active intervention (antidepressant, psychotherapy, and combination of antidepressant and psychotherapy) with any control condition or another active intervention for the acute treatment of children and adolescents (≤18 years old and of both sexes) with a primary diagnosis of depressive disorder, including major depressive disorder, dysthymia, and other specified types as defined by standard operationalised diagnostic criteria (Research Diagnostic Criteria, Schedule for Affective Disorders and Schizophrenia for School-Age Children—Present and Lifetime Version, DSM-III, DSM-III revised, DSM-IV, DSM-IV text revision, DSM-5, and ICD-10). The electronic database searches were supplemented with manual searches for published, unpublished, and ongoing RCTs in international trial registers (eg, ClinicalTrials.gov), websites of drug approval agencies (eg, US Food and Drug Administration [FDA] website), key scientific journals and conference proceedings in the field, and reference lists of relevant trials or reviews appendix pp 3–17).[14] We contacted study authors and drug manufacturers to request complete reports of the original papers or data from unpublished studies. There was no restriction on language.

We included any licensed oral antidepressants within the therapeutic dose range, including tricyclic antidepressants (amitriptyline, clomipramine, desipramine, imipramine, and nortriptyline), selective serotonin reuptake inhibitors (citalopram, escitalopram, fluoxetine, paroxetine, and sertraline), serotonin norepinephrine reuptake inhibitors (desvenlafaxine, duloxetine, and venlafaxine), and other drugs (mirtazapine, nefazodone, and vilazodone), as well as any manualised or structured psychotherapies, including behavioural therapy, CBT, family therapy, interpersonal psychotherapy, psychodynamic therapy, problem-solving therapy, supportive therapy, and others, regardless of the delivery format (eg, individual or group) or treatment medium (eg, face-to-face or online). We also included the combination of the above-mentioned antidepressants and psychotherapies. The pharmacological control condition was always a pill placebo, whereas the psychological control conditions were waiting list, treatment as usual, and psychological placebo (appendix pp 18–20). For trials of antidepressants alone, we included only double-blind RCTs (patients and raters blinded). For trials of psychotherapy alone or the combination of antidepressant and psychotherapy, we included trials in which observers or raters were masked or participants were assessed by self-rating depression scales, because participants and therapists cannot be blinded.[15,16] To reduce clinical heterogeneity, we excluded trials with quasi-randomised design, treatment duration of less than 4 weeks, and an overall sample size of fewer than ten patients. Trials involving patients with certain comorbid psychiatric disorders (eg, anxiety disorder or attention deficit hyperactivity disorder; appendix pp 21–24) were included, whereas trials that included participants with bipolar disorder, psychotic depression, depressive symptoms that did not meet the diagnostic criteria of depressive disorder, or treatment-resistant depression were excluded.

Two of four investigators (XZ, TT, YZ, and LY) independently selected the studies, reviewed the main reports

See Online for appendix

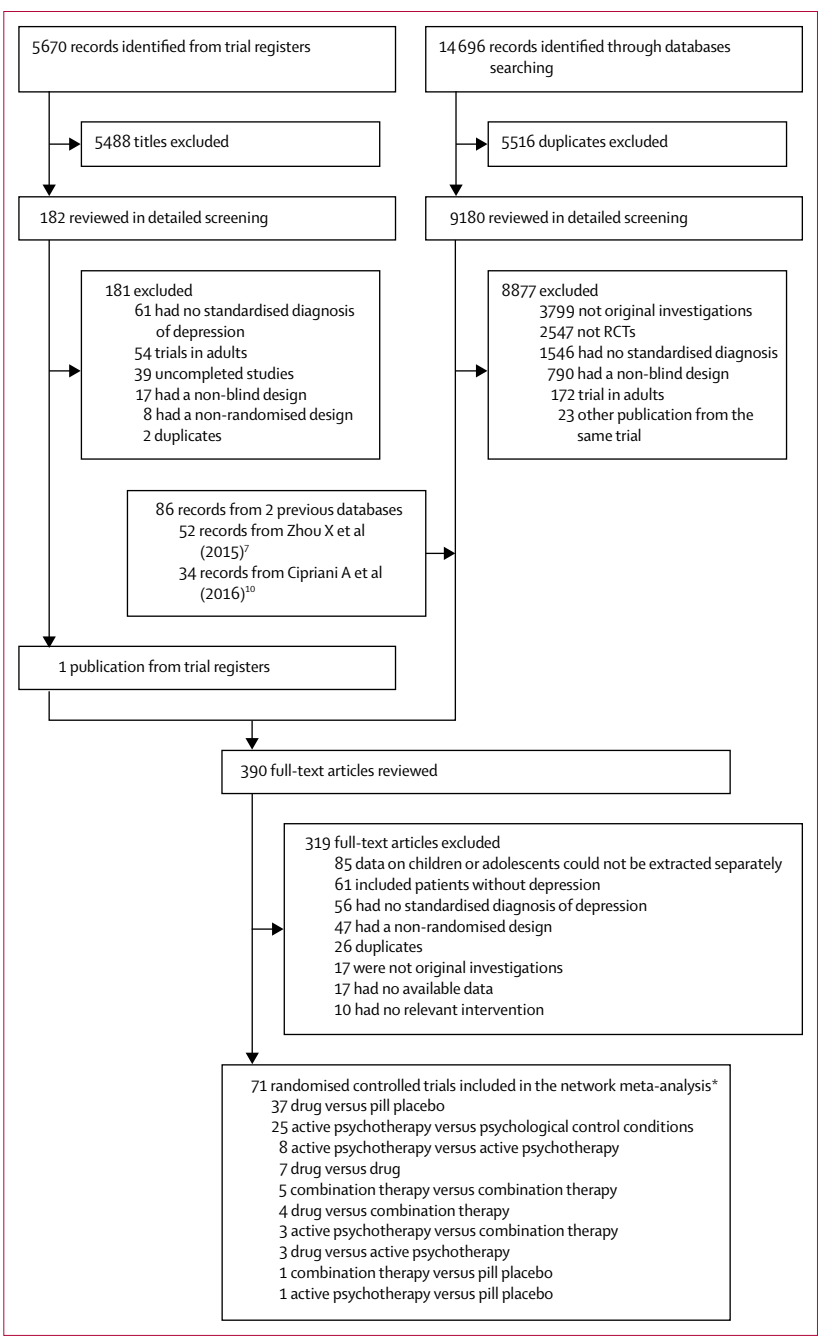


Figure 1: **Study profile**
*Descriptions are not mutually exclusive.

	Diagnostic criteria	Type of depression	Treatments (dose range)	Number randomly assigned to each group	Treatment duration (selected timepoint, weeks)	Age range, years (mean)	Proportion female	Area recruited from	Setting	Baseline severity scale; mean baseline severity (SD)	Transforming score of baseline* (SD)	Manufacturer funder	Type of publication	Type of blinding
Kye et al (1996)	K-SADS and RDC	MDD	Amitriptyline (5 mg/day per kg); pill placebo	18/13	8 (8)	12–17 (14·8)	29%	USA	Outpatients	HAMD (clinician-reported); 12·50 (4·31)	41·73 (8·52)	None	Published trial	Double-blind
Von Knorring et al (2006)	DSM-IV	MDD	Citalopram (10–40 mg/day); pill placebo	124/120	12 (12)	13–18 (16·0)	Not stated	Europe	Inpatients and outpatients	MADRS (clinician-reported); Not stated	Not stated	Lundbeck	Unpublished data from author	Double-blind
Wagner et al (2004)	DSM-IV	MDD	Citalopram (20–40 mg/day); pill placebo	93/85	8 (8)	7–17 (12·1)	53%	USA	Not stated	CDRS-R (clinician-reported); 58·32 (10·98)	58·32 (10·98)	Forest Laboratories	Published trial	Double-blind
Braconnier et al (2003)	DSM-IV	MDD	Clomipramine (75–150 mg/day); Paroxetine (20–40 mg/day)	58/63	8 (8)	12–20 (16·1)	60%	France	Not stated	MADRS (clinician-reported); 31·84 (4·63)	65·29 (7·02)	GlaxoSmithKline	Published trial	Double-blind
Klein et al (1998)	DSM-III-R	MDD	Desipramine (50–300 mg/day); pill placebo	23/22	6 (6)	13–18 (15·7)	67%	USA	Outpatients	HAMD-24 (clinician-reported); 21·39 (4·44)	42·61 (5·32)	None	Published trial	Double-blind
Kutcher et al (1994)	DSM-III-R	MDD	Desipramine (200 mg/day); pill placebo	30/30	6 (6)	15–19 (17·8)	70%	Canada	Outpatients	HAMD (clinician-reported); 23·20 (5·23)	57·60 (9·15)	None	Published trial	Double-blind
Atkinson et al (2018)	DSM-IV-TR	MDD	Desvenlafaxine (20–35 mg/day); Desvenlafaxine (25–50 mg/day); pill placebo	122/121/120	8 (8)	7–17 (13·0)	56%	USA and Chile	Outpatients	CDRS-R (clinician-reported); 58·09 (9·19)	58·09 (9·19)	Pfizer	Published trial	Double-blind
Weihs et al (2018)	DSM-IV-TR	MDD	Desvenlafaxine (25–50 mg/day); Fluoxetine (20 mg/day); pill placebo	115/113/112	8 (8)	7–17 (12·7)	54%	USA and Mexico	Outpatients	CDRS-R (clinician-reported); 56·53 (8·94)	56·53 (8·94)	Pfizer	Published trial	Double-blind
Emslie et al (2014)	DSM-IV-TR	MDD	Duloxetine (60 mg/day); Duloxetine (30 mg/day); Fluoxetine (20 mg/day); pill placebo	108/116/117/122	10 (10)	7–17 (13·0)	51%	Cross-continental	Outpatients	CDRS-R (clinician-reported); 58·78 (10·33)	58·78 (10·33)	Eli Lilly	Published trial	Double-blind
Atkinson et al (2014)	DSM-IV-TR	MDD	Duloxetine (60–120 mg/day); Fluoxetine (20–40 mg/day); pill placebo	117/117/103	10 (10)	7–17 (13·2)	52%	Cross-continental	Outpatients	CDRS-R (clinician-reported); 59·37 (10·90)	59·37 (10·90)	Eli Lilly	Published trial	Double-blind
Emslie et al (2009)	DSM-IV	MDD	Escitalopram (10–20 mg/day); pill placebo	158/158	8 (8)	12–17 (14·6)	59%	USA	Outpatients	CDRS-R (clinician-reported); 56·80 (8·26)	56·80 (8·26)	Forest Laboratories	Published trial	Double-blind

(Table 1 continues on next page)

	Diagnostic criteria	Type of depression	Treatments (dose range)	Number randomly assigned to each group	Treatment duration (selected timepoint, weeks)	Age range, years (mean)	Proportion female	Area recruited from	Setting	Baseline severity scale; mean baseline severity (SD)	Transforming score of baseline* (SD)	Manufacturer funder	Type of publication	Type of blinding
(Continued from previous page)														
Bristol-Myers Squibb (2002)	DSM-IV	MDD	Nefazodone (100–300 mg/day); Nefazodone (200–600 mg/day); pill placebo	95/95/94	8 (8)	7–17 (Not stated)	Not stated	Not stated	Not stated	CDRS-R (clinician-reported); 60·17 (Not stated)	60·17 (Not stated)	Bristol-Myers Squibb	Unpublished trial from FDA report	Double-blind
Emslie et al (2002)	DSM-IV	MDD	Nefazodone (100–400 mg/day); pill placebo	99/96	8 (8)	12–17 (Not stated)	59%	Not stated	Not stated	CDRS-R (clinician-reported); Not stated	Not stated	Bristol-Myers Squibb	Unpublished trial from abstract for conference	Double-blind
Geller et al (1990)	DSM-III	MDD	Nortriptyline (45–140 mg/day); pill placebo	12/19	8 (8)	12–17 (14·3)	45%	USA	Outpatients	CDRS (clinician-reported); 51·36 (3·91)	51·36 (3·91)	None	Published trial	Double-blind
Geller et al (1992)	DSM-III	MDD	Nortriptyline (10–140 mg/day); pill placebo	30/30	8 (8)	6–12 (9·7)	30%	Not stated	Outpatients	CDRS-R (clinician-reported); 49·75 (4·37)	49·75 (4·37)	None	Published trial	Double-blind
Berard et al (2006)	DSM-IV	MDD	Paroxetine (20–40 mg/day); pill placebo	187/99	12 (8)	13–18 (15·6)	67%	Cross-continental	Outpatients	MADRS (clinician-reported); 25·90 (6·42)	56·28 (9·74)	GlaxoSmithKline	Published trial	Double-blind
Emslie et al (2006)	DSM-IV	MDD	Paroxetine (10–50 mg/day); pill placebo	104/102	8 (8)	7–17 (12·0)	47%	USA and Canada	Not stated	CDRS-R (clinician-reported); 61·64 (9·20)	61·64 (9·20)	GlaxoSmithKline	Published trial	Double-blind
GlaxoSmith Kline (2009)	DSM-IV-TR	MDD	Paroxetine (10–40 mg/day); pill placebo	29/27	8 (8)	7–17 (14·6)	61%	Japan	Not stated	CDRS-R (clinician-reported); 56·08 (7·84)	56·08 (7·84)	GlaxoSmithKline	Unpublished trial from clinical trials.gov	Double-blind
Noury et al (2015)	DSM-III-R	MDD	Paroxetine (20–60 mg/day); Imipramine (200–300 mg/day); pill placebo	93/95/87	8 (8)	12–18 (14·9)	62%	USA	Not stated	HAMD-17 (clinician-reported); 18·66 (4·19)	49·65 (7·33)	None	Published trial	Double-blind
Wagner et al (2003)	DSM-IV	MDD	Sertraline (50–200 mg/day); pill placebo	97/91	10 (10)	6–17 (Not stated)	51%	Cross-continental	Outpatients	CDRS-R (clinician-reported); 64·01 (10·97)	64·01 (10·97)	Pfizer	Published trial	Double-blind
Wagner et al (2003)	DSM-IV	MDD	Sertraline (50–200 mg/day); pill placebo	92/96	10 (10)	6–17 (Not stated)	52%	Cross-continental	Outpatients	CDRS-R (clinician-reported); 64·91 (10·98)	64·91 (10·98)	Pfizer	Published trial	Double-blind

(Table 1 continues on next page)

	Diagnostic criteria	Type of depression	Treatments (dose range)	Number randomly assigned to each group	Treatment duration (selected timepoint, weeks)	Age range, years (mean)	Proportion female	Area recruited from	Setting	Baseline severity scale; mean baseline severity (SD)	Transforming score of baseline* (SD)	Manufacturer funder	Type of publication	Type of blinding
(Continued from previous page)														
Wagner et al (2006)	DSM-IV	MDD	Escitalopram (10–20 mg/day); pill placebo	132/136	8 (8)	6–17 (12·3)	52%	USA	Outpatients	CDRS-R (clinician-reported); 55·57 (Not stated)	55·57 (Not stated)	Forest Laboratories	Published trial	Double-blind
Attari et al (2006)	DSM-IV	MDD	Fluoxetine (0·5–2 mg/day per kg); Nortriptyline (1–2 mg/day per kg)	20/20	8 (8)	7–16 (12·9)	50%	Iran	Outpatients	CDI (self-reported); 28·65 (8·50)	65·28 (14·33)	Not stated	Published trial	Double-blind
Almeida-Montes et al (2005)	DSM-IV-TR	MDD	Fluoxetine (20 mg/day); pill placebo	12/11	6 (6)	8–14 (11·4)	35%	Mexico	Outpatients	DSRS (self-reported); Not stated	Not stated	None†	Published trial	Double-blind
Eli Lilly et al (1986)	DSM-III	MDD	Fluoxetine (20–60 mg/day); pill placebo	21/19	6 (6)	12–17 (15·6)	55%	Canada	Inpatients and outpatients	HAMD-17 (clinician-reported); 21·90 (3·46)	55·33 (6·06)	Eli Lilly	Unpublished trial from Eli Lilly company	Double-blind
Emslie et al (1997)	DSM-III-R	MDD	Fluoxetine (20 mg/day); pill placebo	48/48	8 (8)	7–17 (12·4)	46%	USA	Outpatients	CDRS-R (clinician-reported); 58·05 (10·40)	58·05 (10·40)	None	Published trial	Double-blind
Emslie et al (2002)	DSM-IV	MDD	Fluoxetine (10–20 mg/day); pill placebo	109/110	9 (9)	8–18 (12·7)	49%	USA	Outpatients	CDRS-R (clinician-reported); 56·10 (10·92)	56·10 (10·92)	Eli Lilly	Published trial	Double-blind
Findling et al (2009)	DSM-IV	MDD or other depressive disorder	Fluoxetine (10–20 mg/day); pill placebo	18/16	8 (8)	12–17 (16·5)	15%	USA	Outpatients	CDRS-R (clinician-reported); 53·44 (9·70)	53·44 (9·70)	Eli Lilly	Published trial	Double-blind
Hongfen et al (2009)	CCMD-3	MDD	Fluoxetine (20 mg/day); Venlafaxine (150 mg/day)	30/30	8 (8)	12–18 (15·8)	47%	China	Inpatients and outpatients	HAMD-17 (clinician-reported); 22·05 (2·34)	55·59 (4·10)	Not stated	Unpublished trial from abstract for conference	Double-blind
Puig-Antich et al (1987)	K-SADS and RDC	MDD	Imipramine (3·25–5 mg/day per kg); pill placebo	20/22	5 (5)	6–12 (9·1)	40%	USA	Inpatients and outpatients	K-SADS-9 (clinician-reported); 3·05 (0·56)	Not available	None	Published trial	Double-blind
Organon et al (2002)	DSM-IV	MDD	Mirtazapine (15–45 mg/day); pill placebo	82/44	8 (8)	7–18 (12·3)	51%	Europe	Outpatients	CDRS-R (clinician-reported); 51·28 (9·05)	51·28 (9·05)	Organon	Unpublished trial from FDA report	Double-blind
Organon et al (2002)	DSM-IV	MDD	Mirtazapine (15–45 mg/day); pill placebo	88/45	8 (8)	7–18 (12·0)	53%	Europe	Outpatients	CDRS-R (clinician-reported); 48·43 (10·56)	48·43 (10·56)	Organon	Unpublished trial from FDA report	Double-blind

(Table 1 continues on next page)

	Diagnostic criteria	Type of depression	Treatments (dose range)	Number randomly assigned to each group	Treatment duration (selected timepoint, weeks)	Age range, years (mean)	Proportion female	Area recruited from	Setting	Baseline severity scale; mean baseline severity (SD)	Transforming score of baseline* (SD)	Manufacturer funder	Type of publication	Type of blinding
(Continued from previous page)														
Emslie et al (2007)‡	DSM-IV	MDD	Venlafaxine (37·5–225 mg/day); pill placebo	184/183	8 (8)	7–17 (12·3)	46%	USA	Outpatients	CDRS-R (clinician-reported); 56·10 (8·80)	56·10 (8·80)	Wyeth Research	Unpublished data from author	Double-blind
Durgam et al (2018)	DSM-IV-TR	MDD	Vilazodone (15 mg/day); Vilazodone (30 mg/day); pill placebo	175/180/174	8 (8)	12–17 (14·8)	60%	USA	Outpatients	CDRS-R (clinician-reported); 57·36 (8·59)	57·36 (8·59)	Forest Research Institute	Published trial	Double-blind

References for included studies are provided in the appendix (pp 41–48). CCMD-3=Chinese Classification of Mental Disorders third version. CDI=Children's Depression Inventory. CDRS-R=Children's Depression Rating Scale-Revised. DSRS=Depression Self-Rating Scale. FDA=US Food and Drug Administration. HAMD=Hamilton Rating Scale for Depression. K-SADS=Kiddie-Schedule for Affective Disorders and Schizophrenia for School-Age Children. MADRS=Montgomery-Asberg Depression Rating Scale. MDD=major depressive disorder. RDC=Research Diagnostic Criteria. *The method for transforming other depressive scales to CDRS-R.[26] †The authors stated that fluoxetine and placebo were donated by Eli Lilly, but this company was not involved in the design, planning, implementation, collection, analysis, and presentation of the results of this study. ‡This publication reports the combined data from two similarly designed controlled studies comparing venlafaxine with placebo.

Table 1: **Randomised controlled trials of drugs included in the systematic review and network meta-analysis**

and supplementary materials, extracted the relevant information from the included trials, and assessed the risk of bias (κ range for interrater reliability 0·87–0·90). Any discrepancies were resolved by consensus and arbitration by a panel of investigators within the review team (PX, AC, TAF, and PC). The full protocol of this network meta-analysis has been published.[14] We assessed the studies' risk of bias in accordance with the Cochrane Handbook for Systematic Reviews of Interventions.[14] We assessed the confidence of evidence contributing to each network estimate using the Confidence In Network Meta-Analysis (CINeMA) software.[17]

For **Confidence In Network Meta-Analysis** see https://cinema.ispm.unibe.ch/

Outcomes

Our primary outcomes were efficacy (depressive symptoms measured by mean overall change scores from baseline to after completion of treatment on standardised depressive symptom scales) and acceptability (all-cause discontinuation measured by the proportion of patients who withdrew from the study for any reason). All-cause discontinuation was used as a measure of the acceptability of treatments because it encompasses efficacy and tolerability.[18] The secondary outcome was suicidality (measured by reported cases of suicidal behaviour or ideation). When depressive symptoms were measured with more than one standardised rating scale in the same trial, we used a predefined hierarchy (appendix p 26) based on psychometric properties and consistency of use across included trials.[14] We established a hierarchy of informants of depressive rating scales, giving priority to those that were clinician-reported then those that were self-reported. We recorded the outcomes as close to 8 weeks as possible for all analyses. If information at 8 weeks was not available, we used data from 4–16 weeks (we gave preference to the timepoint closest to 8 weeks; if equidistant, we took the longer outcome).[14]

Data analysis

We did a pairwise meta-analysis in STATA (version 15.1)[14] and network meta-analysis in OpenBUGS (version 3.2.3)[19] using the random-effects model by summary standardised mean differences (SMDs, Cohen's d) with 95% CIs for continuous outcomes and odds ratios (ORs) with credible intervals (CrIs) for dichotomous outcomes. Missing continuous outcome data were analysed using the last available follow-up data, and missing dichotomous outcome data were managed according to the intention-to-treat principle. Missing SDs were calculated from p values, *t* values, CIs, or standard errors.[20] Further details about statistical analyses are provided in the published protocol.[14]

To assess transitivity, we compared the distribution of clinical and methodological variables (eg, age, sex, depressive severity at baseline, and treatment duration) that could act as effect modifiers across treatment comparisons.[14] The variance in the random-effects distribution (heterogeneity variance) was considered to

measure the extent of cross-study and within-comparison variability of treatment effects. A common estimate for the heterogeneity variance was assumed for all comparisons in the entire network, and we assessed the presence of statistical heterogeneity using the magnitude of the heterogeneity variance parameter (τ^2) and total I^2 statistic. Incoherence between direct and indirect sources of evidence was statistically assessed globally, by comparison of the fit and parsimony of consistency and inconsistency models, and locally, by calculation of the difference between direct and indirect estimates in all closed loops in the network.[21] The node splitting method, which separated evidence on a particular comparison into direct and indirect evidence, was used to calculate the inconsistency of the model.[22] We estimated the ranking probabilities of being at each possible rank for each intervention. The treatment hierarchy was summarised and reported as surface under the cumulative ranking curve. To determine whether the results were affected by study characteristics, we did network meta-regression for primary outcomes according to the following variables: sex ratio, mean age, sponsorship, treatment duration, comorbid psychiatric disorder, risk of bias, sample size, rating scale, publication year, and mean baseline severity. We did prespecified sensitivity analyses for primary outcomes by omitting trials with unpublished data, trials with imputed data, trials with sample sizes smaller than 20, trials with inconsistent treatment durations and selected timepoints, and trials with non-blinding assessment. We used comparison-adjusted funnel plots to assess publication bias.[23]

We fitted all models of network meta-analysis with uninformative previous distributions for the treatment effects. The codes for the network meta-analysis models are listed in the appendix (pp 27–37). In the network meta-analysis, we used group-level data; the normal likelihood for continuous outcomes and the binomial likelihood were used for dichotomous outcomes. Pooled estimates were obtained using the Markov Chains Monte Carlo method. Two Markov chains were run simultaneously with different arbitrarily chosen initial values. To ensure convergence, trace plots and the Brooks-Gelman-Rubin statistic were assessed.[24] Statistical evaluation of inconsistency and production of network graphs and figures were done using the network and network graphs packages in STATA (version 15.1).[25] The appendix (p 39) lists the changes to the original protocol, which is registered with PROSPERO, number CRD42015020841.

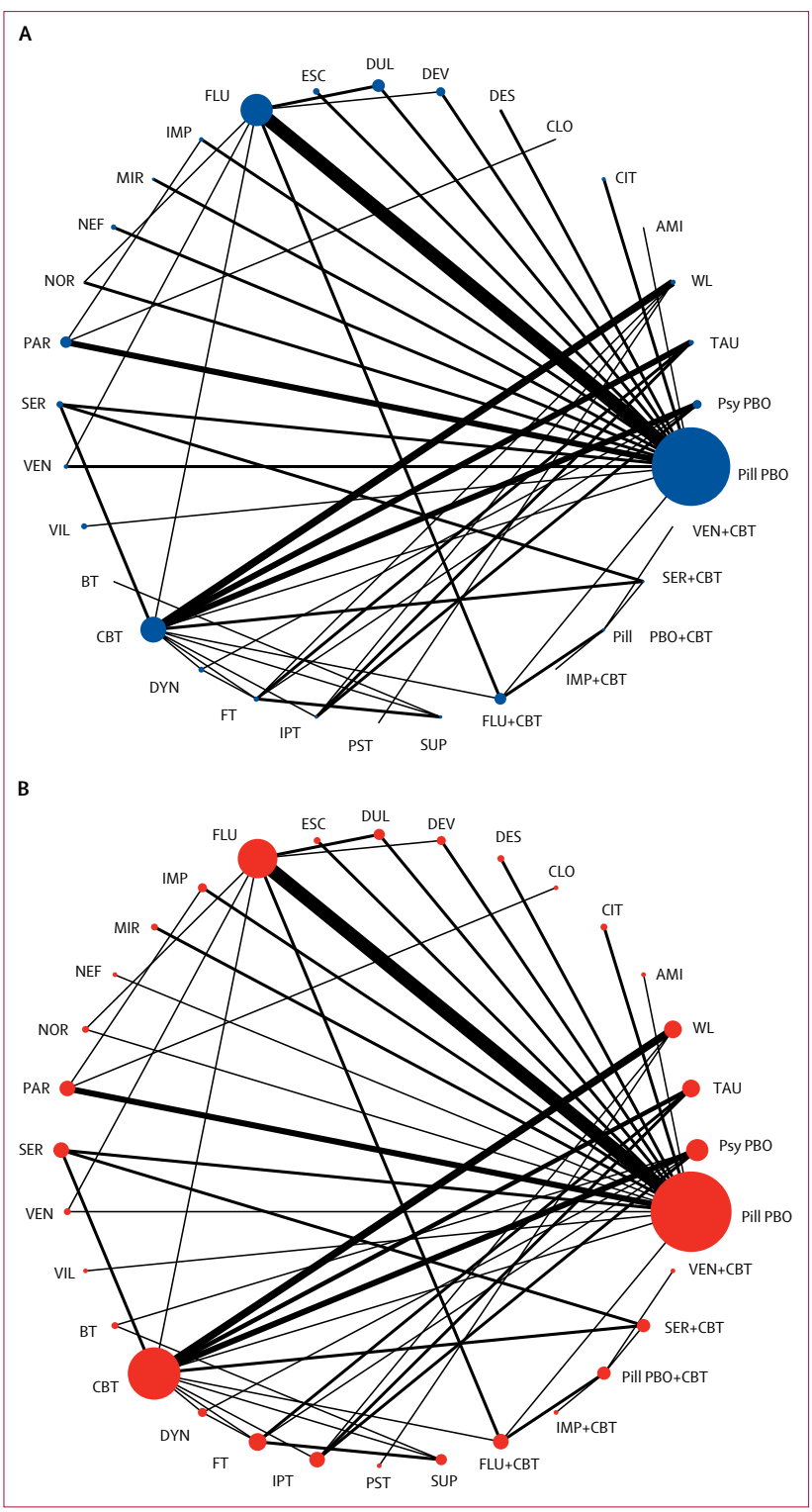


Figure 2: **Network of eligible comparisons**
(A) Efficacy. (B) Acceptability. The width of the lines is proportional to the number of trials comparing every pair of treatments, and the size of each node is proportional to the number of randomly assigned participants. AMI=Amitriptyline. BT=Behavioural therapy. CBT=cognitive-behavioural therapy. CIT=citalopram. CLO=clomipramine. DYN=psychodynamic therapy. DES=desipramine. DEV=desvenlafaxine. DUL=duloxetine. ESC=escitalopram. FT=family therapy. FLU=fluoxetine. IPT=interpersonal therapy. IMP=imipramine. MIR=mirtazapine. NEF=nefazodone. NOR=nortriptyline. PST=problem-solving therapy. PAR=paroxetine. Pill PBO=pill placebo. Psy PBO=psychological placebo. SUP=supportive therapy. SER=sertraline. TAU=treatment as usual. VEN=venlafaxine. VIL=vilazodone. WL=waiting list.

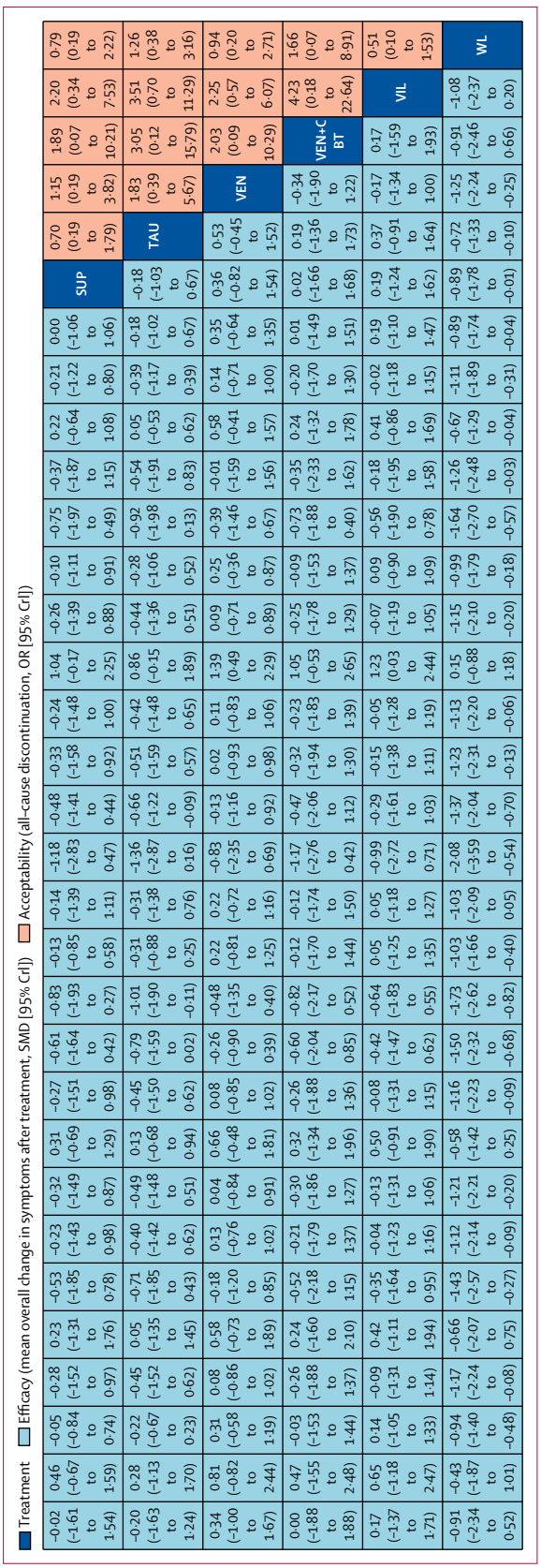
Treatment | Efficacy (mean overall change in symptoms after treatment, SMD [95% CrI]) | Acceptability (all-cause discontinuation, OR [95% CrI])

AMI	2·39 (0·09 to 13·10)	2·71 (0·22 to 12·18)	3·03 (0·25 to 13·34)	1·67 (0·10 to 8·08)	1·31 (0·09 to 6·21)	3·44 (0·29 to 15·24)	2·79 (0·24 to 12·14)	2·09 (0·13 to 10·16)	2·08 (0·17 to 9·06)	3·72 (0·35 to 15·81)	3·92 (0·31 to 17·39)	2·33 (0·15 to 11·17)	1·15 (0·09 to 5·07)	7·39 (0·23 to 41·97)	3·84 (0·22 to 18·88)	3·48 (0·26 to 15·70)	5·99 (0·42 to 28·03)	3·80 (0·21 to 18·25)	2·23 (0·20 to 9·65)	1·16 (0·29 to 12·13)	4·28 (0·24 to 21·12)	37·45 (0·12 to 213·60)	1·77 (0·12 to 8·26)	1·82 (0·15 to 8·05)	1·70 (0·11 to 8·22)	3·41 (0·19 to 16·75)	1·93 (0·12 to 9·39)	2·61 (0·20 to 11·91)	4·50 (0·08 to 27·81)	4·96 (0·37 to 22·73)	2·10 (0·14 to 9·99)
−0·48 (−2·42 to 1·46)	BT	2·05 (0·31 to 6·98)	2·69 (0·28 to 10·96)	1·48 (0·11 to 6·58)	1·19 (0·09 to 5·24)	3·00 (0·30 to 12·04)	2·45 (0·27 to 9·51)	1·52 (0·19 to 5·69)	1·83 (0·19 to 7·31)	3·24 (0·39 to 12·13)	3·29 (0·36 to 12·82)	1·58 (0·27 to 5·15)	1·02 (0·10 to 4·19)	6·26 (0·23 to 32·81)	2·82 (0·34 to 10·49)	3·13 (0·29 to 13·09)	5·26 (0·49 to 22·39)	3·31 (0·24 to 15·00)	1·98 (0·21 to 7·73)	1·20 (0·31 to 9·61)	3·57 (0·26 to 15·49)	27·40 (0·16 to 179·90)	1·27 (0·19 to 4·32)	1·54 (0·17 to 5·84)	1·37 (0·13 to 5·64)	2·03 (0·49 to 5·74)	1·40 (0·19 to 4·96)	2·29 (0·21 to 9·62)	3·73 (0·09 to 21·71)	4·38 (0·39 to 18·48)	1·59 (0·19 to 5·86)
0·03 (−1·33 to 1·39)	0·51 (−0·88 to 1·89)	CBT	1·29 (0·43 to 3·11)	0·71 (0·14 to 2·20)	0·57 (0·12 to 1·69)	1·45 (0·49 to 3·38)	1·18 (0·43 to 2·62)	0·76 (0·32 to 1·66)	0·89 (0·29 to 2·11)	1·57 (0·68 to 3·06)	1·60 (0·60 to 3·43)	0·85 (0·37 to 1·70)	0·50 (0·14 to 1·24)	2·98 (0·30 to 12·16)	1·43 (0·53 to 3·13)	1·50 (0·42 to 3·89)	2·54 (0·65 to 6·96)	1·62 (0·29 to 5·10)	0·95 (0·35 to 2·11)	1·08 (0·55 to 2·43)	1·73 (0·38 to 4·98)	13·19 (0·15 to 6·33)	0·65 (0·32 to 1·16)	0·75 (0·30 to 1·57)	0·66 (0·19 to 1·65)	1·27 (0·41 to 3·06)	0·72 (0·29 to 1·46)	1·10 (0·33 to 2·84)	1·83 (0·09 to 8·95)	2·11 (0·58 to 5·61)	0·77 (0·34 to 1·48)
0·26 (−1·13 to 1·65)	0·74 (−0·95 to 2·42)	0·23 (−0·75 to 1·20)	CIT	0·62 (0·13 to 1·86)	0·49 (0·11 to 1·41)	1·26 (0·44 to 2·84)	1·03 (0·38 to 2·22)	0·76 (0·17 to 2·24)	0·77 (0·27 to 1·74)	1·37 (0·59 to 2·67)	1·44 (0·48 to 3·37)	0·85 (0·19 to 2·42)	0·43 (0·13 to 1·03)	2·70 (0·25 to 11·26)	1·41 (0·29 to 4·21)	1·30 (0·38 to 3·26)	2·20 (0·59 to 5·83)	1·41 (0·26 to 4·33)	0·82 (0·32 to 1·73)	0·96 (0·52 to 1·97)	1·58 (0·31 to 4·83)	13·53 (0·11 to 86·21)	0·64 (0·16 to 1·73)	0·67 (0·24 to 1·50)	0·62 (0·14 to 1·77)	1·25 (0·23 to 4·02)	0·71 (0·15 to 1·99)	0·96 (0·30 to 2·41)	1·65 (0·07 to 8·27)	1·82 (0·53 to 4·67)	0·77 (0·18 to 2·12)
−0·25 (−1·91 to 1·41)	0·23 (−1·68 to 2·14)	−0·28 (−1·61 to 1·05)	−0·51 (−1·87 to 0·86)	CLO	1·12 (0·17 to 3·97)	2·88 (0·61 to 8·57)	2·35 (0·52 to 6·81)	1·75 (0·25 to 6·28)	1·76 (0·37 to 5·25)	3·14 (0·77 to 8·61)	3·29 (0·67 to 10·13)	1·95 (0·29 to 6·70)	0·95 (0·20 to 2·76)	6·17 (0·42 to 28·75)	3·24 (0·45 to 11·43)	2·96 (0·55 to 9·29)	5·02 (0·86 to 16·70)	3·21 (0·39 to 11·79)	1·76 (0·54 to 4·31)	1·75 (0·66 to 6·57)	3·60 (0·48 to 13·33)	31·67 (0·20 to 199·40)	1·48 (0·24 to 4·90)	1·53 (0·33 to 4·54)	1·41 (0·21 to 4·88)	2·87 (0·35 to 10·71)	1·63 (0·23 to 5·53)	2·20 (0·42 to 7·03)	3·78 (0·13 to 20·06)	4·18 (0·76 to 13·58)	1·76 (0·27 to 6·03)
0·51 (−0·93 to 1·96)	0·99 (−0·74 to 2·72)	0·49 (−0·57 to 1·54)	0·26 (−0·83 to 1·35)	0·76 (−0·67 to 2·18)	DES	3·49 (0·80 to 10·02)	2·85 (0·69 to 8·07)	1·13 (0·34 to 7·35)	2·13 (0·49 to 6·15)	3·82 (1·03 to 10·24)	3·99 (0·87 to 11·93)	2·38 (0·37 to 8·23)	1·18 (0·24 to 3·55)	7·43 (0·52 to 33·49)	3·98 (0·58 to 14·43)	3·57 (0·72 to 10·79)	6·05 (1·16 to 19·33)	3·89 (0·50 to 13·77)	2·28 (0·57 to 6·33)	2·21 (0·88 to 7·67)	4·35 (0·62 to 15·99)	35·71 (0·25 to 225·80)	1·80 (0·31 to 5·93)	1·85 (0·43 to 5·36)	1·71 (0·27 to 5·90)	3·50 (0·46 to 12·66)	1·99 (0·31 to 6·84)	2·65 (0·56 to 8·08)	4·59 (0·17 to 24·87)	5·04 (0·99 to 15·57)	2·12 (0·35 to 7·01)
0·21 (−1·16 to 1·57)	0·68 (−0·97 to 2·33)	0·18 (−0·74 to 1·09)	−0·05 (−1·04 to 0·93)	0·45 (−0·88 to 1·79)	−0·31 (−1·37 to 0·75)	DEV	0·91 (0·35 to 1·94)	0·44 (0·16 to 2·01)	0·68 (0·24 to 1·53)	1·21 (0·56 to 2·28)	1·27 (0·43 to 2·98)	0·75 (0·18 to 2·13)	0·38 (0·12 to 0·91)	2·36 (0·23 to 9·81)	1·25 (0·27 to 3·67)	1·15 (0·35 to 2·89)	1·94 (0·53 to 5·17)	1·24 (0·24 to 3·86)	0·73 (0·29 to 1·52)	0·85 (0·47 to 1·74)	1·38 (0·28 to 4·24)	11·75 (0·10 to 75·29)	0·57 (0·15 to 1·52)	0·59 (0·22 to 1·34)	0·54 (0·13 to 1·57)	1·10 (0·21 to 3·57)	0·63 (0·14 to 1·77)	0·85 (0·26 to 2·13)	1·46 (0·07 to 7·26)	1·62 (0·48 to 4·20)	0·68 (0·17 to 1·86)
0·30 (−1·05 to 1·65)	0·77 (−0·87 to 2·41)	0·27 (−0·63 to 1·15)	0·04 (−0·93 to 1·00)	0·54 (−0·78 to 1·86)	−0·22 (−1·26 to 0·82)	0·09 (−0·82 to 0·99)	DUL	0·54 (0·20 to 2·34)	0·81 (0·31 to 1·79)	1·43 (0·77 to 2·45)	1·51 (0·57 to 3·34)	0·90 (0·22 to 2·48)	0·45 (0·15 to 1·05)	2·83 (0·28 to 11·85)	1·49 (0·34 to 4·33)	1·37 (0·45 to 3·29)	2·32 (0·67 to 6·01)	1·47 (0·29 to 4·38)	0·87 (0·38 to 1·76)	1·04 (0·62 to 1·96)	1·65 (0·36 to 4·94)	14·17 (0·12 to 90·87)	0·68 (0·19 to 1·76)	0·71 (0·28 to 1·53)	0·65 (0·16 to 1·80)	1·32 (0·27 to 4·15)	0·75 (0·18 to 2·05)	1·01 (0·34 to 2·49)	1·75 (0·08 to 8·72)	1·93 (0·61 to 4·83)	0·81 (0·21 to 2·17)
−0·33 (−1·88 to 1·21)	0·15 (−1·35 to 1·66)	−0·35 (−1·09 to 0·38)	−0·59 (−1·80 to 0·64)	−0·08 (−1·60 to 1·44)	−0·84 (−2·12 to 0·44)	−0·53 (−1·70 to 0·64)	−0·62 (−1·77 to 0·53)	DYN	1·40 (0·30 to 3·99)	2·47 (0·66 to 6·20)	2·51 (0·61 to 6·80)	1·29 (0·41 to 2·96)	0·78 (0·15 to 2·28)	4·68 (0·36 to 20·70)	2·18 (0·58 to 5·63)	2·35 (0·44 to 7·15)	3·97 (0·69 to 12·42)	2·54 (0·33 to 8·88)	1·50 (0·35 to 4·08)	1·42 (0·54 to 4·92)	2·72 (0·42 to 8·97)	20·82 (0·19 to 160·70)	0·96 (0·37 to 1·93)	1·17 (0·29 to 3·10)	1·04 (0·20 to 3·11)	1·94 (0·46 to 5·39)	1·11 (0·31 to 2·77)	1·73 (0·34 to 5·21)	2·85 (0·11 to 14·49)	3·32 (0·61 to 10·27)	1·21 (0·33 to 3·15)
0·25 (−1·14 to 1·64)	0·73 (−0·95 to 2·41)	0·22 (−0·75 to 1·19)	−0·01 (−1·02 to 1·01)	0·50 (−0·85 to 1·84)	−0·26 (−1·35 to 0·82)	0·05 (−0·92 to 1·02)	−0·05 (−1·00 to 0·91)	0·58 (−0·63 to 1·78)	ESC	2·00 (0·87 to 3·92)	2·09 (0·70 to 4·93)	1·24 (0·29 to 3·55)	0·62 (0·19 to 1·49)	3·91 (0·37 to 16·40)	2·06 (0·43 to 6·25)	1·88 (0·56 to 4·67)	3·19 (0·88 to 8·49)	2·04 (0·38 to 6·37)	1·20 (0·48 to 2·51)	1·40 (0·77 to 2·86)	2·29 (0·45 to 7·12)	18·77 (0·16 to 124·40)	0·94 (0·24 to 2·55)	0·97 (0·34 to 2·20)	0·90 (0·20 to 2·59)	1·82 (0·34 to 5·84)	1·04 (0·23 to 2·96)	1·39 (0·43 to 3·53)	2·42 (0·11 to 12·17)	2·66 (0·78 to 6·88)	1·11 (0·27 to 3·12)
0·59 (−0·64 to 1·82)	1·07 (−0·46 to 2·60)	0·56 (−0·11 to 1·23)	0·33 (−0·46 to 1·13)	0·84 (−0·36 to 2·03)	0·08 (−0·81 to 0·96)	0·38 (−0·31 to 1·08)	0·29 (−0·34 to 0·93)	0·92 (−0·08 to 1·91)	0·34 (−0·44 to 1·12)	FLU	1·06 (0·49 to 2·03)	0·47 (0·19 to 1·61)	0·32 (0·12 to 0·68)	1·99 (0·22 to 7·88)	1·05 (0·28 to 2·83)	0·97 (0·36 to 2·16)	1·65 (0·54 to 3·97)	1·04 (0·23 to 2·95)	0·62 (0·32 to 1·10)	0·78 (0·56 to 1·15)	1·16 (0·29 to 3·22)	9·88 (0·09 to 63·82)	0·48 (0·16 to 1·14)	0·50 (0·23 to 0·97)	0·46 (0·13 to 1·20)	0·93 (0·22 to 2·75)	0·53 (0·15 to 1·33)	0·72 (0·28 to 1·61)	1·23 (0·07 to 5·98)	1·37 (0·49 to 3·20)	0·57 (0·18 to 1·41)
0·81 (−0·55 to 2·18)	1·29 (−0·29 to 2·87)	**0·78 (0·01 to 1·55)**	0·55 (−0·43 to 1·53)	1·06 (−0·26 to 2·39)	0·30 (−0·76 to 1·35)	0·61 (−0·31 to 1·51)	0·52 (−0·36 to 1·40)	1·14 (0·08 to 2·20)	0·56 (−0·41 to 1·53)	0·22 (−0·42 to 0·86)	FLU+CBT	0·46 (0·17 to 1·77)	0·34 (0·10 to 0·84)	1·90 (0·25 to 7·00)	1·08 (0·25 to 3·12)	1·03 (0·29 to 2·63)	1·75 (0·45 to 4·71)	1·11 (0·20 to 3·47)	0·66 (0·24 to 1·43)	0·75 (0·39 to 1·65)	1·11 (0·36 to 2·62)	10·32 (0·09 to 64·31)	0·49 (0·14 to 1·25)	0·52 (0·19 to 1·17)	0·47 (0·12 to 1·30)	0·96 (0·20 to 3·02)	0·55 (0·14 to 1·45)	0·76 (0·23 to 1·94)	1·16 (0·07 to 5·39)	1·45 (0·41 to 3·87)	0·59 (0·16 to 1·52)
0·11 (−1·35 to 1·56)	0·59 (−0·75 to 1·93)	0·09 (−0·43 to 0·61)	−0·14 (−1·24 to 0·97)	0·36 (−1·06 to 1·79)	−0·40 (−1·57 to 0·78)	−0·09 (−1·14 to 0·96)	−0·18 (−1·20 to 0·85)	0·44 (−0·30 to 1·19)	−0·14 (−1·22 to 0·96)	−0·48 (−1·32 to 0·37)	−0·70 (−1·63 to 0·23)	FT	0·68 (0·14 to 2·01)	4·07 (0·32 to 17·80)	1·86 (0·57 to 4·57)	2·05 (0·42 to 6·24)	3·47 (0·65 to 10·98)	2·21 (0·31 to 7·77)	1·31 (0·33 to 3·55)	1·26 (0·49 to 4·31)	2·35 (0·38 to 7·80)	17·87 (0·18 to 121·30)	0·84 (0·35 to 1·72)	1·02 (0·27 to 2·68)	0·90 (0·19 to 2·66)	1·54 (0·60 to 3·33)	0·92 (0·35 to 1·93)	1·51 (0·32 to 4·64)	2·50 (0·10 to 13·06)	2·90 (0·58 to 9·17)	1·05 (0·32 to 2·54)

(Figure 3 continues on next page)

Treatment ☐ Efficacy (mean overall change in symptoms after treatment, SMD [95% CrI]) ☐ Acceptability (all-cause discontinuation, OR [95% CrI])

| | | |



Role of the funding source

The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. PX, AC, and XZ had full access to all the data in the study, and PX had final responsibility for the decision to submit for publication.

Results

We identified 20 366 citations, retrieved the full text of 390 potentially eligible articles, and included 71 RCTs (9510 patients) published between 1986 and 2018 (figure 1). These trials compared 16 antidepressants, seven psychotherapies, five combinations of antidepressants and psychotherapy, and three psychological controls, or pill placebo (figure 1; appendix pp 40–48). 4081 participants were randomly assigned to antidepressants, 1575 to psychotherapy, 553 to a combination treatment, and 3301 to a psychological control or pill placebo. The mean study sample size was 136 participants and ranged from 10 to 529 (tables 1, 2, 3). The age range was from 3 years to 20 years (mean age 14·0 years, SD 2·6); two studies included participants up to 20 years of age, but were included, because the majority of participants and the mean age were younger than 18 years. 5051 (57·2%) of the sample population were female. The median duration of the acute treatment was 8 weeks (IQR 8–12). Participants were randomly assigned to three or more groups in ten (14·1%) of 71 studies. Only outpatients were recruited in 41 (55·7%) of 71 studies. 41 (57·7%) studies were done in North America, 12 (16·9%) in Europe, five (7·0%) in Asia, two (2·8%) in Australia, and one (1·4%) in South America, seven (9·9%) trials were cross-continental, and the remaining three (4·2%) were either from other regions or did not specify. 7179 (75·5%) of 9510 patients had moderate-to-severe major depressive disorder, with a mean reported baseline severity score on the Children's Depression Rating Scale-Revised of 58·5 (SD 10·1), Children's Depression Inventory of 23·3 (SD 8·8), or Beck Depression Inventory of 24·7 (11·4). Pharmaceutical

Figure 3: **Network meta-analysis of efficacy and acceptability**
Interventions are reported in alphabetical order. Comparisons between treatments should be read from left to right, and the estimate is in the cell in common between the column-defining treatment and the row-defining treatment. For efficacy (blue), a SMD less than 0 favours the column-defining treatment. For acceptability (red), an OR less than 1 favours the row-defining treatment. To obtain SMDs for comparisons in the opposing direction, negative values should be converted into positive values and vice versa. To obtain ORs for comparisons in the opposing direction, reciprocals should be taken. Significant results are in bold. AMI=amitriptyline. BT=behavioural therapy. CBT=cognitive-behavioural therapy. CIT=citalopram. CLO=clomipramine. DYN=psychodynamic therapy. DES=desipramine. DEV=desvenlafaxine. DUL=duloxetine. ESC=escitalopram. FT=family therapy. FLU=fluoxetine. IPT=interpersonal therapy. IMP=imipramine. MIR=mirtazapine. NEF=nefazodone. NOR=nortriptyline. OR=odds ratio. PST=problem-solving therapy. PAR=paroxetine. Pill PBO=pill placebo. Psy PBO=psychological placebo. SUP=supportive therapy. SER=sertraline. SMD=standardised mean difference. TAU=treatment as usual. VEN=venlafaxine. VIL=vilazodone. WL=waiting list.

	Diagnostic criteria	Type of depression	Treatments (dose range)	Number randomly assigned to each group	Treatment duration (selected timepoint, weeks)	Age range, years (mean)	Proportion female	Area recruited from	Setting	Baseline severity scale; mean baseline severity (SD)	Transforming score of baseline* (SD)	Manufacturer funder	Type of publication	Type of blinding
Reed et al (1994)	DSM-III-R	MDD or dysthymia	Behavioural therapy (6 sessions); psychological placebo	12/6	12 (no data available)	14–19 (Not stated)	50%	USA	Not stated	CDI (self-reported); Not stated	Not stated	None	Published trial	Not stated (self reported scale)
Fine et al (1991)	DSM-III-R	MDD or dysthymia	Behavioural therapy (Not stated); supportive therapy (Not Stated)	30/36	12 (12)	13–17 (15·1)	83%	Canada	Outpatients	CDI (self-reported); 20·16 (7·52)	50·97 (12·68)	None	Published trial	Not stated (self reported scale)
Charkhandeh et al (2016)	DSM-IV-TR	MDD	CBT (12 sessions); waitlist	65/60	12 (12)	12–17 (Not stated)	54%	Iran	Outpatients	CDI (self-reported); 29·89 (5·46)	67·37 (9·20)	Not stated	Published trial	Not stated (self reported scale)
Clarke et al (1999)	DSM-III-R	MDD or dysthymia	CBT (16 sessions); waitlist	87/36	8 (8)	14–18 (16·2)	71%	USA	Not stated	HAMD-14 (clinician-reported); 14·15 (5·75)	45·00 (11·37)	None	Published trial	Single-blind (assessor-blind)
Curtis et al (1992)	DSM-III-R	MDD or dysthymia	CBT (12 sessions); waitlist	12/11	8 (8)	High school students (15·8)	89%	USA	School	BDI-21 (self-reported); 25·64 (8·47)	54·04 (12·23)	Not stated	Unpublished trial from doctoral dissertation	Non-blind (self reported scale)
Lewinsohn et al (1990)	DSM-III	MDD, MinDD, or intermittent depression	CBT (14 sessions); waitlist	45/24	7 (7)	14–18 (16·2)	61%	USA	Not stated	BDI-21 (self-reported); 22·30 (11·26)	49·21 (16·27)	None	Published trial	Not stated (self reported scale)
Brent et al (1997)	DSM-III-R	MDD	CBT (12–16 sessions); family therapy (12–16 sessions); supportive therapy (12–16 sessions)	37/35/35	12–16 (12–16)	13–18 (15·6)	76%	USA	Outpatients	BDI-21 (self-reported); 24·20 (8·06)	51·96 (11·64)	None	Published trial	Not stated (self reported scale)
Rossello et al (1999)	DSM-III-R	MDD or dysthymia	CBT (12 sessions); interpersonal therapy (12 sessions); waitlist	25/23/23	12 (12)	13–18 (14·7)	54%	USA	School	CDI (self-reported); 20·48 (6·78)	51·51 (11·42)	None	Published trial	Not stated (self reported scale)
Rohde et al (2004)	DSM-IV	MDD	CBT (16 sessions); psychological placebo	45/48	8 (8)	13–17 (15·1)	48%	USA	Outpatients	HAMD-17 (clinician-reported); 13·99 (5·18)	41·49 (9·06)	None	Published trial	Single-blind (assessor-blind)

(Table 2 continues on next page)

	Diagnostic criteria	Type of depression	Treatments (dose range)	Number randomly assigned to each group	Treatment duration (selected timepoint, weeks)	Age range, years (mean)	Proportion female	Area recruited from	Setting	Baseline severity scale; mean baseline severity (SD)	Transforming score of baseline* (SD)	Manufacturer funder	Type of publication	Type of blinding
(Continued from previous page)														
Goodyer et al (2017)	DSM-IV	MDD	CBT (20 sessions); psychodynamic therapy (28 sessions); psychological placebo	155/157/158	Mean 24·9 (12)/ mean 27·9 (12)/ mean 27·5 (12)	11–17 (15·0)	75%	UK	Outpatients	MFQ (self-reported); 45·93 (10·55)	58·95 (9·99)	None	Published trial	Single-blind (assessor-blind)
Vostanis et al (1996)	DSM-III-R	MDD, dysthymia, or MinDD	CBT (9 sessions); psychological placebo	31/30	18 (18)	8–17 (12·7)	56%	UK	Not stated	MFQ (self-reported); 31·04 (13·43)	59·80 (18·52)	None	Published trial	Single-blind (assessor-blind)
Wood et al (1996)	DSM-III-R	MDD, or MinDD	CBT (6·4 sessions); psychological placebo	26/27	Mean 9·2 (9·2)/ Mean 8·4 (8·4)	9–17 (14·2)	69%	UK	Outpatients	MFQ (self-reported); 27·28 (10·75)	54·61 (14·82)	None	Published trial	Single-blind (assessor-blind)
Clarke et al (2002)	DSM-III-R	MDD or dysthymia	CBT (16 sessions); treatment as usual	41/47	8 (8)	13–18 (15·3)	69%	USA	Health maintenance organisation	HAMD-14 (clinician-reported); 11·68 (5·12)	40·11 (10·13)	None	Published trial	Single-blind (assessor-blind)
Kobak et al (2015)	DSM-5	MDD, dysthymia, other specified depressive disorder, or DDNOS	CBT (Not stated); treatment as usual	39/37	12 (12)	12–17 (15·4)	66%	USA	Not Stated	QIDS-A-Pat (self-reported); Not stated	Not stated	None	Published trial	Not stated (self reported scale)
Shirk et al (2014)	K-SADS-LS	MDD, dysthymia, or DDNOS	CBT (12 sessions); treatment as usual	20/23	16 (16)	13–17 (15·5)	84%	USA	Outpatients	BDI-21 (self-reported); 31·11 (11·84)	61·94 (17·11)	None	Published trial	Not stated (self reported scale)
Weisz et al (2009)	DSM-IV	MDD, dysthymia, or MinDD	CBT (15 sessions); treatment as usual	32/25	24 (24)/39 (39)	8–15 (11·8)	56%	USA	Not stated	CDI (self-reported); 11·06 (7·85)	35·64 (13·23)	None	Published trial	Single-blind (assessor-blind)
Trowell et al (2007)	K-SADS	MDD or dysthymia	Psychodynamic therapy (24·7 sessions); family therapy (11 sessions)	35/37	36 (36)	9–15 (11·7)	38%	Europe	Mental health service	CDI (self-reported); 23·43 (7·27)	56·49 (12·26)	None	Published trial	Not stated (self reported scale)
Diamond et al (2002)	DSM-III-R	MDD	Family therapy (12 sessions); waitlist	16/16	12 (12)	13–17 (14·9)	78%	USA	School	HAMD-24 (clinician-reported); 18·60 (6·42)	39·27 (7·69)	None	Published trial	Single-blind (assessor-blind)
Luby et al (2012)	RDC	MDD	Family therapy (14 sessions); psychological placebo	27/27	12 (12)	3–7 (Not stated)	37%	USA	Outpatients	BDI-21 (self-reported); 12·60 (8·49)	35·20 (12·26)	None	Published trial	Single-blind (assessor-blind)

(Table 2 continues on next page)

	Diagnostic criteria	Type of depression	Treatments (dose range)	Number randomly assigned to each group	Treatment duration (selected timepoint, weeks)	Age range, years (mean)	Proportion female	Area recruited from	Setting	Baseline severity scale; mean baseline severity (SD)	Transforming score of baseline* (SD)	Manufacturer funder	Type of publication	Type of blinding
(Continued from previous page)														
Tompson et al (2017)	DSM-IV-TR	MDD, dysthymia, double depression, or DDNOS	Family therapy (15 sessions); supportive therapy (15 sessions);	67/67	22 (22)	7–14 (10·8)	56%	USA	Not stated	CDRS-R (clinician-reported); 53·59 (11·37)	53·59 (11·37)	None	Published trial	Single-blind (assessor-blind)
Israel et al (2013)	K-SADS-PL	MDD	Family therapy (12 sessions); treatment as usual	11/9	12 (12)	13–17 (15·6)	55%	Norway	Outpatients	HAMD-17 (clinician-reported); 20·20 (4·91)	52·34 (8·59)	None	Published trial	Single-blind (assessor-blind)
Poole et al (2018)	DSM-IV	MDD, dysthymia, or MinDD	Family therapy (8 sessions); treatment as usual	31/33	8 (8)	12–18 (15·2)	73%	Australia	Mental health service	SMFQ (self-reported); 18·13 (7·69)	80·45 (26·91)	None	Published trial	Single-blind (assessor-blind)
Dietz et al (2015)	DSM-IV	MDD, dysthymia, or DDNOS	Interpersonal therapy (14 sessions); psychological placebo	29/13	14 (14)	7–12 (10·8)	67%	USA	Outpatients	CDRS-R (clinician-reported); 45·20 (7·81)	45·20 (7·81)	None	Published trial	Single-blind (assessor-blind)
Mufson et al (1999)	DSM-III-R	MDD	Interpersonal therapy (12 sessions); psychological placebo	24/24	12 (12)	12–18 (15·8)	73%	USA	Outpatients	HAMD-24 (clinician-reported); 18·95 (7·99)	39·69 (9·56)	None	Published trial	Single-blind (assessor-blind)
Mufson et al (2004)	DSM-IV	MDD, dysthymia, or DDNOS	Interpersonal therapy (12 sessions); treatment as usual	34/30	12–16 (12–16)	15–18 (15·1)	84%	USA	School	HAMD-24 (clinician-reported); 18·62 (5·46)	39·29 (6·54)	None	Published trial	Single-blind (assessor-blind)
Tang et al (2009)	DSM-IV-TR	MDD	Interpersonal therapy (12 sessions); treatment as usual	35/38	6 (6)	12–18 (15·3)	66%	China	School	BDI-21 (self-reported); 32·48 (9·31)	63·92 (13·45)	Not stated	Published trial	Single-blind (assessor-blind)
Eskin et al (2008)	DSM-IV	MDD	Problem-solving therapy (6 sessions); waitlist	13/10	6 (6)	15–18 (16·3)	65%	Turkey	School	HAMD-17 (clinician-reported); 15·17 (6·40)	43·55 (11·20)	Not stated	Unpublished data from author	Non-blind (self reported scale)

BDI=Beck Depression Inventory. CBT=cognitive-behavioural therapy. CCMD-3=Chinese Classification of Mental Disorders third version. CDI=Children's Depression Inventory. CDRS-R=Children's Depression Rating Scale-Revised. DDNOS=Depressive disorder-not otherwise specified. DSRS=Depression Self-Rating Scale. FDA=US Food and Drug Administration. HAMD=Hamilton Rating Scale for Depression. K-SADS=Kiddie-Schedule for Affective Disorders and Schizophrenia for School-Age Children. MADRS=Montgomery-Asberg Depression Rating Scale. MDD=major depressive disorder. MFQ=Mood and Feelings Questionnaire. MinDD=minor depressive disorder. QIDS-A-Pat=Quick inventory of depressive symptomatology-adolescent version. RADS=Reynolds Adolescent Depression Scale. RDC=Research Diagnostic Criteria. SMFQ=The Short Moods and Feelings Questionnaire. *The method for transforming other depressive scales to CDRS-R.[26]

Table 2: **Randomised controlled trials of psychotherapy included in the systematic review and network meta-analysis**

	Diagnostic criteria	Type of depression	Treatments (dose range)	Number randomly assigned to each group	Treatment duration (selected timepoint, weeks)	Age range, years (mean)	Proportion female	Area recruited from	Setting	Baseline severity scale; mean baseline severity (SD)	Transforming score of baseline* (SD)	Manufacturer funder	Type of publication	Type of blinding
Cornelius et al (2009)	DSM-IV	MDD	Fluoxetine plus CBT (10–20 mg/day; 9 sessions); pill placebo plus CBT (9 sessions)	24/26	12 (12)	15–20 (Not stated)	56%	USA	Not stated	HAMD-27 (clinician-reported); 20·00 (8·50)	40·95 (10·17)	None	Published trial	Double-blind
March et al (2004)	DSM-IV	MDD	Fluoxetine plus CBT (10–40 mg/day; 15 sessions); Fluoxetine (10–40 mg/day); CBT (15 sessions); pill placebo	107/109/111/112	12 (12)	12–17 (14·6)	54%	USA	Outpatients	CDRS-R (clinician-reported); 60·14 (4·49)	60·14 (4·49)	None	Published trial	Double-blind (fluoxetine, placebo); assessor-blind (CBT, fluoxetine plus CBT)
Goodyer et al (2008)	DSM-IV	MDD	Fluoxetine plus CBT (10–60mg/day; 12 sessions); Fluoxetine (10–60 mg/day)	105/103	12 (12)	11–17 (14·0)	74%	UK	Outpatients	CDRS-R (clinician-reported); 58·95 (9·99)	58·95 (9·99)	None	Published trial	Single-blind (assessor-blind)
Riggs et al (2007)	DSM-IV	MDD	Fluoxetine plus CBT (20 mg/day; 16 sessions); pill placebo plus CBT (16 sessions)	63/63	16 (8)	13–19 (17·2)	33%	USA	Outpatients	CDRS-R (clinician-reported); 56·84 (13·42)	56·84 (13·42)	None	Published trial	Double-blind
Bernstein et al (2000)	DSM-III-R	MDD	Imipramine plus CBT (3 mg/day per kg; 8 sessions); pill placebo plus CBT (8 sessions)	31/32	8 (8)	12–18 (13·9)	60%	USA	Not stated	CDRS-R (clinician-reported); 49·70 (10·50)	49·70 (10·50)	None	Published trial	Double-blind
Melvin et al (2006)	DSM-IV	MDD, dysthymia, or DDNOS	Sertraline plus CBT (25–100 mg/day; 12 sessions); Sertraline (25–100 mg/day per kg); CBT (12 sessions)	25/26/22	12 (12)	12–18 (15·3)	66%	Australia	Outpatients	RADS (self-reported); 84·24 (13·21)	72·18 (13·36)	None	Published trial	Non-blind (self-reported scale)
Deas et al (2000)	DSM-IV	MDD	Sertraline plus CBT (25–100 mg/day; 12 sessions); pill placebo plus CBT (12 sessions)	5/5	12 (12)	15–18 (16·6)	20%	USA	Outpatients	HAMD-24 (clinician-reported); 20·60 (5·19)	41·67 (6·21)	None	Published trial	Double-blind

(Table 3 continues on next page)

	Diagnostic criteria	Type of depression	Treatments (dose range)	Number randomly assigned to each group	Treatment duration (selected timepoint, weeks)	Age range, years (mean)	Proportion female	Area recruited from	Setting	Baseline severity scale; mean baseline severity (SD)	Transforming score of baseline* (SD)	Manufacturer funder	Type of publication	Type of blinding
(Continued from previous page)														
Iftene et al (2015)	DSM-IV	MDD	Sertraline plus CBT (25–50 mg/day; 16 sessions); Sertraline (25–50 mg/day); CBT (16 sessions)	27/33/28	16 (8)	11–17 (15·3)	56%	Romania	Mental health services	CDI (self-reported); 24·01 (5·79)	57·45 (9·76)	None	Published trial	Not stated (self-reported scale)
Mandoki et al (1997)	DSM-IV	MDD	Venlafaxine plus CBT (12·5–75 mg/day; 6 sessions); pill placebo plus CBT (6 sessions)	20/20	6 (6)	8–17 (12·8)	76%	USA	Outpatients	CDRS (clinician-reported); 34·83 (Not stated)	34·83 (Not stated)	Not stated	Published trial	Double-blind

CBT=cognitive-behavioural therapy. CDI=Children's Depression Inventory. CDRS-R=Children's Depression Rating Scale-Revised. DDNOS=Depressive disorder-not otherwise specified. HAMD=Hamilton Rating Scale for Depression. MDD=major depressive disorder. RADS=Reynolds Adolescent Depression Scale. *The method for transforming other depressive scales to CDRS-R.[26]

Table 3: **Randomised controlled trials of combinations of drugs and psychotherapy included in the systematic review and network meta-analysis**

companies funded 24 (33·8%) of 71 studies. We retrieved unpublished information for 11 (15·5%) of the 71 included trials. 32 trials (45·1%) were rated high on risk of bias, 32 (45·1%) as moderate, and seven (9·9%) as low (appendix pp 49–53).

In terms of efficacy (70 RCTs, comprising 8906 patients), only fluoxetine plus CBT (SMD –0·73, 95% CrI –1·39 to –0·07) and fluoxetine (–0·51, –0·84 to –0·18) were more effective than both pill placebo and psychological controls (SMDs ranged from –1·73 to –0·83; figures 2A, 3, 4A; appendix pp 56–65). Fluoxetine plus CBT was more effective than CBT (SMDs –0·78, 95% CrI –1·55 to –0·01) and psychodynamic therapy (–1·14, –2·20 to –0·08); and interpersonal psychotherapy was more effective than all psychological controls (SMDs ranged from –1·37 to –0·66; figures 2A, 3, 4A; appendix pp 56–65). By contrast, nortriptyline (SMDs ranged from 1·04 to 2·22) and waiting list (SMDs ranged from 0·67 to 2·08) were worse than most active interventions.

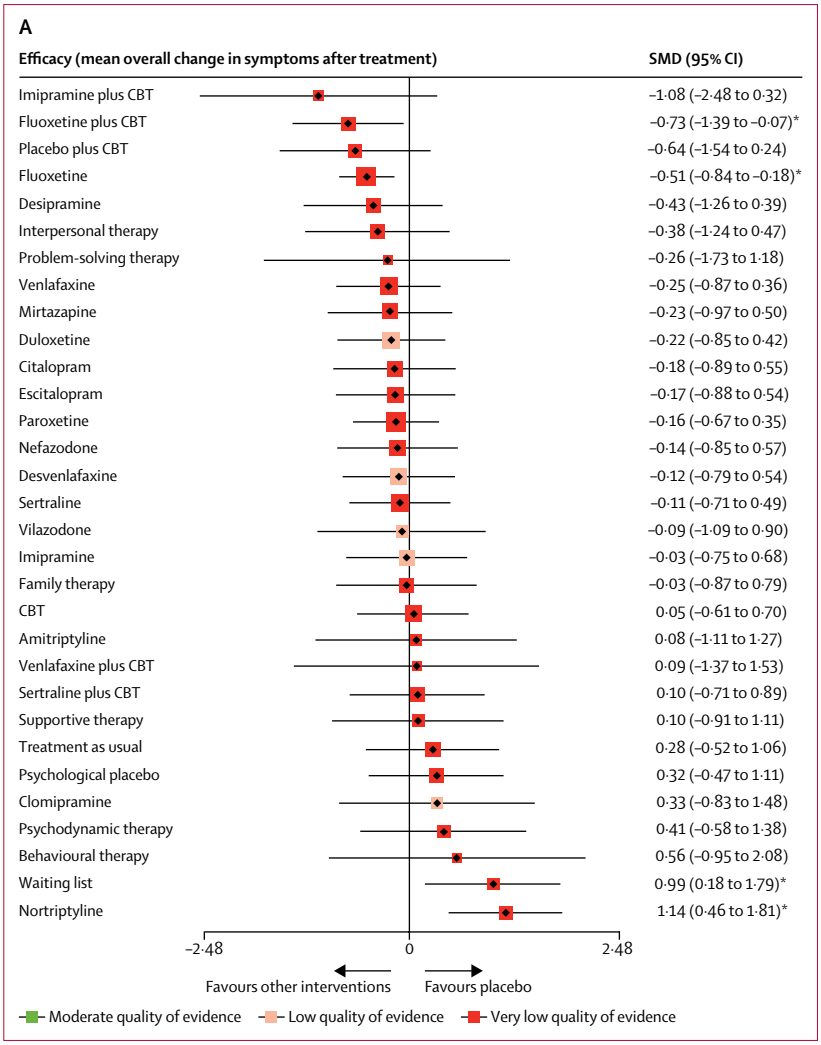
In terms of acceptability (66 RCTs, comprising 9075 patients), nefazodone and fluoxetine were associated with fewer dropouts than sertraline, imipramine, and desipramine (ORs ranged from 0·17 to 0·50; figure 2B, 3, 4B). Imipramine was associated with more dropouts than pill placebo, desvenlafaxine, fluoxetine plus CBT, and vilazodone (ORs ranged from 2·51 to 5·06; figure 2B, 3, 4B).

Venlafaxine was associated with a significantly increased risk of suicidal behaviour or ideation compared with pill placebo (OR 8·31, 95% CrI 1·92–343·17) and ten other interventions (citalopram, escitalopram, fluoxetine, fluoxetine plus CBT, duloxetine, imipramine, family therapy, desvenlafaxine, CBT, and pill placebo plus CBT; ORs ranged from 5·07 to 18·98; figure 4C; appendix pp 67–69).

The median heterogeneity variances were estimated at 0·49 (95% CrI 0·37–0·64) for efficacy and 0·32 (0·04–0·61) for acceptability. The global I^2 values were 56% for efficacy and 14% for acceptability. The assessment of transitivity showed most of the comparisons had variable baseline severity, mean age, sex ratio, and treatment duration. For example, one comparison of psychodynamic therapy with family therapy showed that it had a relatively long treatment duration of 36 weeks (appendix pp 70–72). The test of global incoherence showed a significant difference between the consistency and inconsistency models for efficacy ($p<0·0001$), but not for acceptability ($p=0·5531$; appendix p 74). Tests of local incoherence showed that the percentages for inconsistent loops were within the expected ranges based on the empirical data (six of 25 loops for the efficacy outcome and one of 24 for the acceptability outcome; appendix pp 74–77). The test of incoherence from the node-splitting model showed significant differences between some comparisons in efficacy and acceptability (appendix pp 78–81). The comparison-adjusted funnel plots of the network meta-analysis were suggestive of

publication bias for efficacy outcome in psychotherapy trials, but not for acceptability (appendix pp 82–88).

Network meta-regression analyses showed that most modifiers (appendix pp 90) did not significantly affect the efficacy and acceptability of interventions; however, we found that studies in which participants had more severe depressive symptoms at baseline were associated with larger treatment effects, and that studies with high risk of bias were associated with a lower drop-out rate. These findings might result from the fact that most psychotherapy trials, which were assessed as high risk of bias due to non-blinding of performance and personnel, had relatively lower drop-out rates and baseline severity scores than the pharmacological trials (appendix pp 91–96). The sensitivity analyses did not materially affect the relative treatment effects (appendix pp 97–100). The ranking of treatments based on cumulative probability plots and surfaces under the cumulative ranking curve are presented in the appendix (pp 101–106). According to CINeMA, nine (12·5%) of 72 comparisons for the efficacy outcome were rated as low confidence of evidence and 63 (87·5%) as very low, and for the acceptability outcome, one (1·3%) was rated as high confidence of evidence, three (4·0%) as moderate, 13 (17·3%) as low, and 58 (77·3%) as very low (appendix pp 107–125).



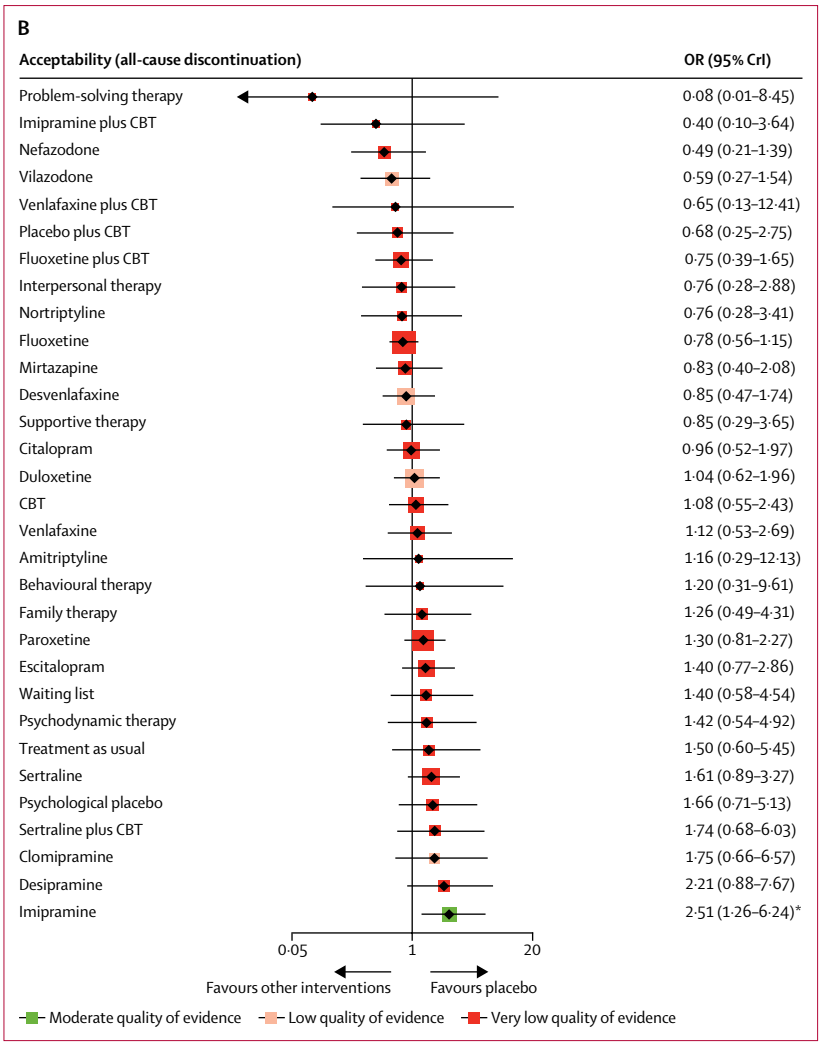
(Figure 4 continues on next page)

Discussion

This updated analysis is based on 71 RCTs, which included 9510 children and adolescents with depressive disorders randomly assigned to 28 active interventions or four control conditions. To our knowledge, this is the first time that psychological intervention, pharmacological intervention, and their combination for depressive disorder in children and adolescents have been compared in a network meta-analysis.

We found that, of all the included active interventions, only fluoxetine plus CBT and fluoxetine were significantly more efficacious than pill placebo in children and adolescents with depressive disorders. We also found that interpersonal psychotherapy was more efficacious than all psychological controls, but with very low confidence of evidence. Fluoxetine plus CBT was associated with a greater reduction in depressive symptoms than either CBT or psychodynamic psychotherapy, with very low confidence of evidence. Nortriptyline was worse than most active interventions; however, the interpretation of this result was limited by the inconsistent loop of nortriptyline versus fluoxetine versus pill placebo. These summary effect sizes were mostly medium to large with some uncertainty, which might result from the small number of patients included, and wide credible intervals. Thus, statistical indications of clinical superiority in this study should be interpreted cautiously.

Our findings in children and adolescents contrast with findings on the efficacy of antidepressants and psychological interventions in adults with major depressive disorder, for whom all antidepressants were more efficacious than pill placebo[27] and all psychotherapeutic interventions were superior to psychological control conditions.[28] There are several possible explanations for this considerable difference. First, neurodevelopmental mechanisms, including robust changes in hormones and hormonal receptors in adolescent depression, could exacerbate emotional responses to negative social stimuli by dysregulation of the hypothalamic–pituitary–adrenal axis.[29] Second, the smaller number of trials and smaller sample sizes for young patients with depression decreases statistical power for each comparison.[30] Third, different design methods between adult and paediatric trials could lead to a higher placebo response rate in children and adolescents (45%) than adults (36%) based on clinician ratings, hindering the detection of positive results for depression in children and adolescents.[31] It is also possible that the psychotherapies used with young patients with depression, which are largely adaptations of treatments developed for adults, might not be ideally



(Figure 4 continues on next page)

suited to the cognitive, behavioural, and emotional characteristics of young people, and that innovations in treatment design and content will be needed to produce stronger treatment effects.

In 2004, the FDA placed a boxed warning on antidepressants for risk of suicidal thoughts and behaviour in children and adolescents on the basis of results of clinical trials.[32] In our analysis, suicidality data on psychological and combination interventions were, for the first time, systematically investigated using the same approach used for medication alone. We found that venlafaxine had a significantly increased risk for suicidality (suicidal behaviour or ideation) for young people, which is in line with previous reviews.[10,11] Two US medical claims databases that contain data on 221 028 young people with depression for the period 2004–09 showed that, after accounting for the time varying effect of confounders, the apparent association between antidepressant use and suicide attempts and self-inflicted injury was diminished and not statistically significant.[33] Antidepressant use by adolescents had previously been increasing but declined abruptly after the warnings were introduced.[34] Our evidence linked venlafaxine alone to an increased effect on suicidal behaviour or ideation, which might be due to better reporting of venlafaxine data. Owing to the absence of reliable data on suicidality for many antidepressants, comprehensive assessment of the risk of suicidality for all interventions was not possible. Prescribers should closely monitor suicide risk when children and adolescents take any antidepressant drugs, particularly at the beginning of treatment.[5]

Our review has several limitations. First, according to the CINeMA assessment, the quality of most comparisons was low or very low. Many trials did not report adequate information about allocation concealment, and it is difficult to use a double-blind design for patients in trials of psychotherapy, which would affect the transitivity of the whole network and restricts the interpretation of these results.[15] We did a sensitivity analysis excluding non-blinded psychotherapy trials, the findings of which were not materially different from those of the primary analysis. Additionally, different outcomes from the same trials can be a source of pharmaceutical marketing bias.[35] However, before the study, we established a hierarchy of informants of depressive rating scales, which could reduce this type of outcome bias. Second, in the network, we found some global and local inconsistencies in efficacy outcomes, but few in acceptability outcomes, perhaps because the proportion of patients who withdrew was a more consistently measured outcome across studies than efficacy, which was measured using various rating scales. Moreover, this inconsistency in efficacy outcomes might be a consequence of the decrease in antidepressant–placebo differences in antidepressant clinical trials in the past three decades, which could be explained by changes in study design.[36] Although the meta-regression analyses of modifiers did not materially affect the outcomes, we found that some comparisons had relatively low or high values in the transitivity assessment; thus, we downgraded the confidence of these comparisons. Third, in order to support transitivity assumption in the network, the review was restricted to trials involving children and adolescents with depressive disorder. We excluded studies in which participants were described as having subsyndromal depressive symptoms, because antidepressants are not recommended in this group of patients. They do, however, form a substantial proportion of the patients seen in real-world, clinical settings.[37] We also excluded patients with psychotic or treatment-resistant depression. Augmentation therapy is usually required for these patients, and including them would have violated transitivity required of the network meta-analysis. Fourth, despite the Egger's test showing no publication bias for most outcomes, we found some potential asymmetry of funnel plots in this network

meta-analysis. Thus, the clinical interpretation of these findings is limited by the potential bias from selective reporting. We did our best to retrieve all available unpublished information and contacted study authors for supplementary data, but we cannot rule out the possibility that some unpublished studies are still missing.[38] Fifth, the Restoring Study 329,[39] which reanalysed the data and protocol of SmithKline Beecham's Study 329,[40] showed different and even opposite results of efficacy and tolerability of paroxetine and imipramine. We have selected the data from Restoring Study 329 for this review, but we could not assure the accuracy of the data in the other included trials. Although we have checked the published data with their protocols or trial register reports, we were not able to investigate these main outcomes at the individual patient level. Researchers and clinicians should recognise the potential biases in published studies, especially with regard to the potential barriers that have led to inaccurate reporting of harm outcomes.[39] Sixth, antidepressants with different doses might produce different treatment effects.[41] Although we included antidepressants without therapeutic dose ranges, we should consider the potential dose effects in this review. Moreover, various antidepressants have a wide range of half-lives, from 5 h to 5 days. Antidepressants with a long half-life (ie, fluoxetine and paroxetine) need to be titrated over 3 or 4 weeks, whereas antidepressants with a short half-life (ie, venlafaxine) do not.[42] These titrations might confuse the outcomes from the short trials. In this review, we have excluded trials with treatment duration of less than 4 weeks, which could reduce the effect for the final analysis. Seventh, because of the paucity of information reported in the original studies, we were not able to quantify some outcomes, such as adverse events discontinuation and global functioning. Some of the adverse effects would also be expected in psychotherapy trials, including the emergence of new symptoms and strains in the patient-therapist relationship,[43] however, few psychotherapy trials report data on adverse events and suicidality.[44] The current report summarises evidence of efficacy and acceptability of active interventions when prescribed in acute treatment. Relatively few studies addressed the issue of preventing relapse of depression in children and adolescents, and some of the adverse effects of antidepressants and response to psychotherapy occur over a prolonged period, meaning that positive results need to be interpreted with caution. Finally, there were some limitations in the network meta-analysis method. In this network meta-analysis, a small number of trials compared the same treatments, and the assumption of transitivity over various control conditions was understated. These control conditions can lead to reduced network connectivity in network meta-analyses and therefore low statistical power.[45] We excluded observational studies to decrease the heterogeneity in the network meta-analysis; however, observational studies can provide more information about real-world evidence on antidepressant effectiveness in the studied population group.[46]

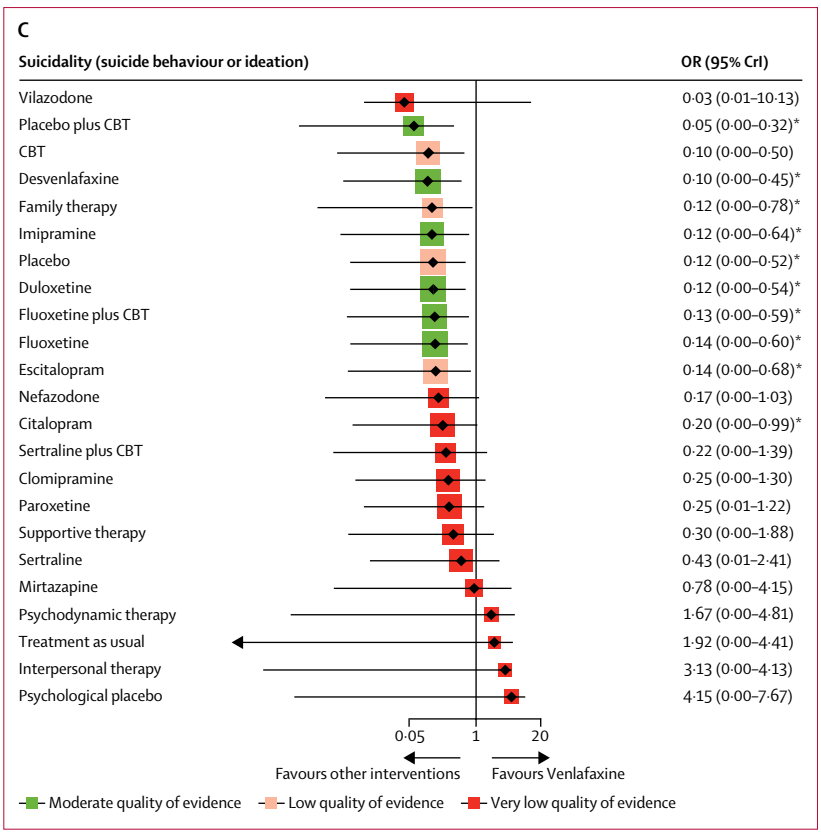


Figure 4: **Forest plots of network meta-analysis**
(A) Efficacy. (B) Acceptability. (C) Suicidality. Interventions were compared with pill placebo for efficacy and acceptability and with venlafaxine for suicidality. CBT=Cognitive-behavioural therapy. CrI=credible interval. OR=odds ratio. SMD=standardised mean difference. *Significant results.

Despite these limitations, the findings from this network meta-analysis represent the most comprehensive analysis of the available evidence. The findings suggest that fluoxetine (alone or in combination with CBT) might be considered the best option to treat acute symptoms in children and adolescents with major depression. Future guidelines and daily clinical decision making on the choice of interventions for acute treatment of young patients with depression should account for these results. Academia, industry, and study authors should collaborate to produce more research that analyses individual patient data in network meta-analyses. Such analyses will enable the prediction of personalised clinical outcomes, including specific side-effects, comparative efficacy at multiple timepoints, and different baseline severities.

Contributors

PX, AC, and XZ had full access to all data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. XZ, YZ, CDG, TAF, JRW, PC, DC, SEH, SL, JB, AVR, LY, AC, and PX conceived and designed the study. XZ, TT, YZ, XL, YX, MQ, LY, and LF selected the articles and extracted the data. XZ, TT, YZ and CDG analysed the data. XZ, AC, and PX interpreted the data and wrote the

first draft of the manuscript. All authors contributed to critical revision of the report for important intellectual content. All authors read and met the ICMJE criteria for authorship and agree with the results and conclusions of this Article.

Declaration of interests

XZ reports travel and accommodation expenses from the Chinese Society of Psychiatry (CSP) for lectures delivered for CSP, outside the submitted work. TAF reports grants and personal fees from Mitsubishi-Tanabe and personal fees from MSD and Shionogi and has a pending patent (2018-177688), outside the submitted work. DC reports grants and personal fees from Shire-Takeda and personal fees from Medice, Servier, and Oxford University Press, outside the submitted work. SEH reports that she is the joint coordinating editor of the Cochrane Common Mental Disorders Group and manages the Children and Young People Satellite. She has funding from the Royal Society, the Faculty of Medical and Health Sciences at the University of Auckland, and Cochrane to pursue this work, including systematic reviews in the area of children and young people's mental health. She is funded by the Auckland Medical Research Foundation to develop and test an application that delivers goal setting for young people with mental health and related difficulties, such as self-harm. She is a CureKids Research Fellow, working on developing digital tools to support parents to support children with mental health and related difficulties. SL reports personal fees from LB Pharma, Otsuka, Lundbeck, Boehringer Ingelheim, LTS Lohmann, Janssen, Johnson & Johnson, TEVA, Merck Sharp & Dohme, Sandoz, Sanofi-Aventis, Angelini, Recordati, and Gedeon Richter, outside the submitted work. AVR reports grants and non-financial support from Janssen Canada and personal fees from Abilify Maintena, MDD National Advisory Board, Mental Health National Advisory, Allergan National Advisory Board, Brexpiprazole Advisory Board, and Bipolar Disorder Advisory Board, outside the submitted work. AC reports personal fees from the CARIPLO Foundation, Angelini Pharma for consultancy and paid peer reviewing of grant applications, and Italian Network for Paediatric Clinical Trials, outside the submitted work. PX reports speaker's honoraria from Janssen, outside the submitted work. All other authors declare no competing interests.

Data sharing

With the publication of this Article, the full dataset will be freely available online in Mendeley Data with the digital object identifier 10.17632/kw6nmfn2tb.1.

Acknowledgments

This study was funded by the National Key Research and Development Program of China (2017YFA0505700), We thank Sofia Dias (Health Technology Assessment, University of York, UK), Alex Sutton (Department of Health Sciences, College of Life Sciences, University of Leicester, UK) and Nicky Welton (Department of Population Health Sciences, Bristol Medical School, University of Bristol, UK) for providing statistical guidance. We are grateful to Mehmet Eskin (Department of Psychology, College of Social Sciences and Humanities, Koc University, Turkey), Graham J Emslie (Department of Psychiatry, University of Texas Southwestern Medical Center, USA), and Taryn Mayes (Department of Psychiatry, University of Texas Southwestern Medical Center, USA) for providing unpublished data in this review, and Salvatore Gentile (Department of Neurosciences, University of Naples, Italy) for his valuable advice. We also thank Ian M Goodyer (Department of Psychiatry, University of Cambridge, UK), Giovanna Porta (Department of Psychiatry, University of Pittsburgh School of Medicine, USA), David Brent (Department of Psychiatry, University of Pittsburgh School of Medicine, USA), Greg Clarke (Kaiser Permanente Center for Health Research, Portland), Paul Wilkinson (Department of Psychiatry, University of Cambridge, UK), and Glenn Melvin (School of Psychology, University of Deakin, Australia) for replying to our requests. PX is supported by the National Key Research and Development Program of China (2017YFA0505700), the Non-profit Central Research Institute Fund of Chinese Academy of Medical Sciences (2019PT320002), and the Natural Science Foundation Project of China (81820108015). XZ is supported by the National Natural Science Foundation of China (81873800 and 81701342), High-level Talents Special Support Plan of Chongqing (T04040016), and Science and Technology Research Project of Chongqing Education Commission (KJQN201800415). AC is supported by the NIHR Oxford Cognitive Health Clinical Research Facility, by an NIHR Research Professorship (RP-2017-08-ST2-006), by the NIHR Collaboration for Leadership in Applied Health Research and Care Oxford, now recommissioned as NIHR Applied Research Collaboration Oxford and Thames Valley, and by the NIHR Oxford Health Biomedical Research Centre (BRC-1215-20005). The views expressed are those of the authors and not necessarily those of the UK National Health Service, the NIHR, or the UK Department of Health and Social Care.

References

1 Mokdad AH, Forouzanfar MH, Daoud F, et al. Global burden of diseases, injuries, and risk factors for young people's health during 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet* 2016; **387:** 2383–401.

2 Sadler K, Vizard T, Ford T, et al. Mental health of children and young people in England, 2017: summary of key findings. NHS Digital, 2018. https://digital.nhs.uk/data-and-information/publications/statistical/mental-health-of-children-and-young-people-in-england/2017/2017 (accessed March 3, 2020).

3 Ryan ND. Treatment of depression in children and adolescents. *Lancet* 2005; **366:** 933–40.

4 Thapar A, Collishaw S, Pine DS, Thapar AK. Depression in adolescence. *Lancet* 2012; **379:** 1056–67.

5 Hopkins K, Crosland P, Elliott N, Bewley S. Diagnosis and management of depression in children and young people: summary of updated NICE guidance. *BMJ* 2015; **350:** h824.

6 Bachmann CJ, Aagaard L, Burcu M, et al. Trends and patterns of antidepressant use in children and adolescents from five western countries, 2005-2012. *Eur Neuropsychopharmacol* 2016; **26:** 411–19.

7 Zhou X, Hetrick SE, Cuijpers P, et al. Comparative efficacy and acceptability of psychotherapies for depression in children and adolescents: a systematic review and network meta-analysis. *World Psychiatry* 2015; **14:** 207–22.

8 Eckshtain D, Kuppens S, Ugueto A, et al. Meta-analysis: 13-year follow-up of psychotherapy effects on youth depression. *J Am Acad Child Adolesc Psychiatry* 2020; **59:** 45–63.

9 Weisz JR, Kuppens S, Ng MY, et al. What five decades of research tells us about the effects of youth psychological therapy: a multilevel meta-analysis and implications for science and practice. *Am Psychol* 2017; **72:** 79–117.

10 Cipriani A, Zhou X, Del Giovane C, et al. Comparative efficacy and tolerability of antidepressants for major depressive disorder in children and adolescents: a network meta-analysis. *Lancet* 2016; **388:** 881–90.

11 Vitiello B, Davico C. Twenty years of progress in paediatric psychopharmacology: accomplishments and unmet needs. *Evid Based Ment Health* 2018; **21:** e10.

12 Locher C, Koechlin H, Zion SR, et al. Efficacy and safety of selective serotonin reuptake inhibitors, serotonin-norepinephrine reuptake inhibitors, and placebo for common psychiatric disorders among children and adolescents: a systematic review and meta-analysis. *JAMA Psychiatry* 2017; **74:** 1011–20.

13 Cox GR, Callahan P, Churchill R, et al. Psychological therapies versus antidepressant medication, alone and in combination for depression in children and adolescents. *Cochrane Database Syst Rev* 2012; **11:** CD008324.

14 Zhou X, Cipriani A, Zhang Y, et al. Comparative efficacy and acceptability of antidepressants, psychological interventions, and their combination for depressive disorder in children and adolescents: protocol for a network meta-analysis. *BMJ Open* 2017; **7:** e016608.

15 Boutron I, Guittet L, Estellat C, Moher D, Hróbjartsson A, Ravaud P. Reporting methods of blinding in randomized trials assessing nonpharmacological treatments. *PLoS Med* 2007; **4:** e61.

16 Del Giovane C, Cortese S, Cipriani A. Combining pharmacological and nonpharmacological interventions in network meta-analysis in psychiatry. *JAMA Psychiatry* 2019; **76:** 867–68.

17 Salanti G, Del Giovane C, Chaimani A, Caldwell DM, Higgins JPT. Evaluating the quality of evidence from a network meta-analysis. *PLoS One* 2014; **9:** e99682.

18 Cipriani A, Furukawa TA, Salanti G, et al. Comparative efficacy and acceptability of 12 new-generation antidepressants: a multiple-treatments meta-analysis. *Lancet* 2009; **373:** 746–58.

19 Lunn D, Spiegelhalter D, Thomas A, Best N. The BUGS project: evolution, critique and future directions. *Stat Med* 2009; **28:** 3049–67.

20 Furukawa TA, Barbui C, Cipriani A, Brambilla P, Watanabe N. Imputing missing standard deviations in meta-analyses can provide accurate results. *J Clin Epidemiol* 2006; **59:** 7–10.

21 Higgins JP, Jackson D, Barrett JK, Lu G, Ades AE, White IR. Consistency and inconsistency in network meta-analysis: concepts and models for multi-arm studies. *Res Synth Methods* 2012; **3:** 98–110.

22 Dias S, Welton NJ, Caldwell DM, Ades AE. Checking consistency in mixed treatment comparison meta-analysis. *Stat Med* 2010; **29:** 932–44.

23 Chaimani A, Higgins JP, Mavridis D, Spyridonos P, Salanti G. Graphical tools for network meta-analysis in STATA. *PLoS One* 2013; **8:** e76654.

24 Brooks SP, Gelman A. General methods for monitoring convergence of iterative simulations. *J Comput Graph Stat* 1998; **7:** 434–55.

25 Chaimani A, Salanti G. Visualizing assumptions and results in network meta-analysis: the network graphs package. *Stata J* 2015; **15:** 905–50.

26 Schünemann HJ, Oxman AD, Higgins JPT, et al. Presenting results and "Summary of findings" tables. In: Higgins JPT, Green S, eds. Cochrane handbook for systematic reviews of interventions. Hoboken, NJ: Wiley, 2008: 335–58.

27 Cipriani A, Furukawa TA, Salanti G, et al. Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis. *Lancet* 2018; **391:** 1357–66.

28 Barth J, Munder T, Gerger H, et al. Comparative efficacy of seven psychotherapeutic interventions for patients with depression: a network meta-analysis. *PLoS Med* 2013; **10:** e1001454.

29 Cousins L, Goodyer IM. Antidepressants and the adolescent brain. *J Psychopharmacol* 2015; **29:** 545–55.

30 Biau DJ, Kernéis S, Porcher R. Statistics in brief: the importance of sample size in the planning and interpretation of medical research. *Clin Orthop Relat Res* 2008; **466:** 2282–88.

31 Meister R, Abbas M, Antel J, et al. Placebo response rates and potential modifiers in doubleblind randomized controlled trials of second and newer generation antidepressants for major depressive disorder in children and adolescents: a systematic review and metaregression analysis. *Eur Child Adolesc Psychiatry* 2020; **9:** 253–73.

32 US Food and Drug Administration. Suicidality in children and adolescents being treated with antidepressant medications. US Food and Drug Administration, Oct 15, 2004. https://www.fda.gov/drugs/postmarket-drug-safety-information-patients-and-providers/suicidality-children-and-adolescents-being-treated-antidepressant-medications (accessed March 3, 2020).

33 Gibbons RD, Coca Perraillon M, Hur K, Conti RM, Valuck RJ, Brent DA. Antidepressant treatment and suicide attempts and self-inflicted injury in children and adolescents. *Pharmacoepidemiol Drug Saf* 2015; **24:** 208–14.

34 Lu CY, Zhang F, Lakoma MD, et al. Changes in antidepressant use by young people and suicidal behavior after FDA warnings and media coverage: quasi-experimental study. *BMJ* 2014; **348:** g3596.

35 Lexchin J, Bero LA, Djulbegovic B, Clark O. Pharmaceutical industry sponsorship and research outcome and quality: systematic review. *BMJ* 2003; **326:** 1167–70.

36 Khan A, Bhat A, Kolts R, Thase ME, Brown W. Why has the antidepressant-placebo difference in antidepressant clinical trials diminished over the past three decades? *CNS Neurosci Ther* 2010; **16:** 217–26.

37 Cheung AH, Zuckerbrot RA, Jensen PS, Laraque D, Stein REK. Guidelines for adolescent depression in primary care (GLAD-PC): part II. Treatment and ongoing management. *Pediatrics* 2018; **141:** e20174082.

38 de Vries YA, Roest AM, Turner EH, de Jonge P. Hiding negative trials by pooling them: a secondary analysis of pooled-trials publication bias in FDA-registered antidepressant trials. *Psychol Med* 2019; **49:** 2020–26.

39 Le Noury J, Nardo JM, Healy D, et al. Restoring Study 329: efficacy and harms of paroxetine and imipramine in treatment of major depression in adolescence. *BMJ* 2015; **351:** h4320.

40 Keller MB, Ryan ND, Strober M, et al. Efficacy of paroxetine in the treatment of adolescent major depression: a randomized, controlled trial. *J Am Acad Child Adolesc Psychiatry* 2001; **40:** 762–72.

41 Berney P. Dose-response relationship of recent antidepressants in the short-term treatment of depression. *Dialogues Clin Neurosci* 2005; **7:** 249–62.

42 Richelson E. Pharmacology of antidepressants. *Mayo Clin Proc* 2001; **76:** 511–27.

43 Berk M, Parker G. The elephant on the couch: side-effects of psychotherapy. *Aust N Z J Psychiatry* 2009; **43:** 787–94.

44 Linden M, Schermuly-Haupt ML. Definition, assessment and rate of psychotherapy side effects. *World Psychiatry* 2014; **13:** 306–09.

45 Cipriani A, Higgins JP, Geddes JR, Salanti G. Conceptual and technical challenges in network meta-analysis. *Ann Intern Med* 2013; **159:** 130–37.

46 Kernot C, Tomlinson A, Chevance A, Cipriani A. One step closer to personalised prescribing of antidepressants: using real-world data together with patients and clinicians' preferences. *Evid Based Ment Health* 2019; **22:** 91–92.