

Comparative efficacy and tolerability of antidepressants for major depressive disorder in children and adolescents: a network meta-analysis



Andrea Cipriani*, Xinyu Zhou*, Cinzia Del Giovane, Sarah E Hetrick, Bin Qin, Craig Whittington, David Coghill, Yuqing Zhang, Philip Hazell, Stefan Leucht, Pim Cuijpers, Juncai Pu, David Cohen, Arun V Ravindran, Yiyun Liu, Kurt D Michael, Lining Yang, Lanxiang Liu, Peng Xie

Summary

Background Major depressive disorder is one of the most common mental disorders in children and adolescents. However, whether to use pharmacological interventions in this population and which drug should be preferred are still matters of controversy. Consequently, we aimed to compare and rank antidepressants and placebo for major depressive disorder in young people.

Methods We did a network meta-analysis to identify both direct and indirect evidence from relevant trials. We searched PubMed, the Cochrane Library, Web of Science, Embase, CINAHL, PsycINFO, LiLACS, regulatory agencies' websites, and international registers for published and unpublished, double-blind randomised controlled trials up to May 31, 2015, for the acute treatment of major depressive disorder in children and adolescents. We included trials of amitriptyline, citalopram, clomipramine, desipramine, duloxetine, escitalopram, fluoxetine, imipramine, mirtazapine, nefazodone, nortriptyline, paroxetine, sertraline, and venlafaxine. Trials recruiting participants with treatment-resistant depression, treatment duration of less than 4 weeks, or an overall sample size of less than ten patients were excluded. We extracted the relevant information from the published reports with a predefined data extraction sheet, and assessed the risk of bias with the Cochrane risk of bias tool. The primary outcomes were efficacy (change in depressive symptoms) and tolerability (discontinuations due to adverse events). We did pair-wise meta-analyses using the random-effects model and then did a random-effects network meta-analysis within a Bayesian framework. We assessed the quality of evidence contributing to each network estimate using the GRADE framework. This study is registered with PROSPERO, number CRD42015016023.

Findings We deemed 34 trials eligible, including 5260 participants and 14 antidepressant treatments. The quality of evidence was rated as very low in most comparisons. For efficacy, only fluoxetine was statistically significantly more effective than placebo (standardised mean difference -0.51 , 95% credible interval [CrI] -0.99 to -0.03). In terms of tolerability, fluoxetine was also better than duloxetine (odds ratio [OR] 0.31 , 95% CrI 0.13 to 0.95) and imipramine (0.23 , 0.04 to 0.78). Patients given imipramine, venlafaxine, and duloxetine had more discontinuations due to adverse events than did those given placebo (5.49 , 1.96 to 20.86 ; 3.19 , 1.01 to 18.70 ; and 2.80 , 1.20 to 9.42 , respectively). In terms of heterogeneity, the global I^2 values were 33.21% for efficacy and 0% for tolerability.

Interpretation When considering the risk–benefit profile of antidepressants in the acute treatment of major depressive disorder, these drugs do not seem to offer a clear advantage for children and adolescents. Fluoxetine is probably the best option to consider when a pharmacological treatment is indicated.

Funding National Basic Research Program of China (973 Program).

Introduction

Major depressive disorder is common in young people with an estimated point prevalence of 2.8% in school-age children (aged 6–12 years) and 5.6% in adolescents (aged 13–18 years).¹ Compared with adults, children and adolescents with major depressive disorder are still underdiagnosed and undertreated,² possibly because they tend to present with rather undifferentiated depressive symptoms—eg, irritability, aggressive behaviours, and school refusal.³ Consequences of depressive episodes in these patients include serious impairments in social functioning, and suicidal ideation and attempts.⁴ Even though psychological treatments are still considered the first-line treatment in many clinical guidelines,^{4,5}

antidepressants are widely used in the treatment of depression in children and adolescents and the rate of prescription has increased over time.⁶ However, in 2004, the US Food and Drug Administration (FDA) cautioned practitioners on the use of antidepressants in children and adolescents because of increased suicide risk.⁷ Consequently, the question of whether to use antidepressant drugs for the treatment of major depressive disorder in young people and, if so, which antidepressant would be preferred, remains controversial.³

Previous pairwise meta-analyses were done to evaluate the efficacy of all types of antidepressants^{8,9} and to identify factors associated with treatment efficacy.¹⁰ However, these studies were inconclusive because they

Published Online

June 8, 2016

[http://dx.doi.org/10.1016/S0140-6736\(16\)30385-3](http://dx.doi.org/10.1016/S0140-6736(16)30385-3)

See Online/Comment

[http://dx.doi.org/10.1016/S0140-6736\(16\)30585-2](http://dx.doi.org/10.1016/S0140-6736(16)30585-2)

*Contributed equally

Department of Psychiatry, University of Oxford, Oxford, UK (A Cipriani PhD); Department of Neurology and Psychiatry, The First Affiliated Hospital of Chongqing Medical University, Chongqing, China (X Zhou PhD, B Qin MD, Y Zhang MD, J Pu MD, Y Liu MD, L Yang MD, L Liu MD, Prof P Xie MD); Institute of Neuroscience and the Collaborative Innovation Center for Brain Science, Chongqing Medical University, Chongqing, China (X Zhou, B Qin, Y Zhang, J Pu, Y Liu, L Yang, L Liu, Prof P Xie); Department of Diagnostic, Clinical and Public Health Medicine, University of Modena and Reggio Emilia, Modena, Italy (C Del Giovane PhD); Orygen, The National Centre of Excellence in Youth Mental Health, University of Melbourne, Melbourne, VIC, Australia (S E Hetrick DPsych); Doctor Evidence, Santa Monica, CA, USA (C Whittington PhD); Division of Neuroscience, School of Medicine, University of Dundee, Dundee, UK (Prof D Coghill MD); Department of Paediatrics, and Department of Psychiatry, Faculty of Medicine, Dentistry and Health Sciences, University of Melbourne, Melbourne, VIC, Australia (Prof D Coghill); Discipline of Psychiatry, Sydney Medical School, Concord West, NSW, Australia (Prof P Hazell PhD); Department of Psychiatry and Psychotherapy, Technische Universität München, Munich,

Germany (Prof S Leucht MD);
Department of Clinical,
Neuro- and Developmental
Psychology, VU University
Amsterdam, Amsterdam,
Netherlands

(Prof P Cuijpers PhD);

Department of Child and
Adolescent Psychiatry, Hôpital
Pitié-Salpêtrière, Institut des
Systèmes Intelligents et
Robotiques, Université Pierre
et Marie Curie, Paris, France
(D Cohen PhD); Department of

Psychiatry, University of
Toronto, Toronto, ON, Canada
(Prof A V Ravindran PhD);

Campbell Family Mental Health
Research Institute, Centre for
Addiction and Mental Health,
Toronto, ON, Canada

(Prof A V Ravindran);

Department of Psychology,
Appalachian State University,
Boone, NC, USA

(Prof K D Michael PhD);

South Australian Health and
Medical Research Institute,
Adelaide, South Australia,
Australia (Prof P Xie)

Correspondence to:

Dr Andrea Cipriani, Department
of Psychiatry, University of
Oxford, Warneford Hospital
Oxford, OX3 7JX, UK

andrea.cipriani@psych.ox.ac.uk

or

Prof Peng Xie, Department of
Neurology and Psychiatry,
The First Affiliated Hospital of
Chongqing Medical University,
Yuzhong District, Chongqing,
China

xiepeng973@126.com

Research in context

Evidence before this study

Even though psychological treatments are still considered the first-line treatment, antidepressants are widely used in the treatment of depression in children and adolescents. We searched PubMed for previously published meta-analyses on antidepressants in children and adolescents, with the search terms “depressive disorder”, “child”, and “adolescent”. Previous pairwise meta-analyses have been inconclusive because they could not generate clear hierarchies among available treatments, because many antidepressants have not been directly compared.

Added value of this study

Our study provides the first comprehensive systematic review and network meta-analysis of all available double-blind randomised trials, comparing any antidepressant with placebo or another active antidepressant as oral monotherapy in the acute treatment of major depressive disorder in children and adolescents (mean age 9–18 years). Our findings emphasise that only fluoxetine is significantly more efficacious than placebo and some other active drugs at reducing depressive symptoms or the number of discontinuations due to adverse events over 8 weeks. Furthermore, we found robust evidence to suggest a significantly increased risk for suicidality (suicidal behaviour or ideation) for young people given venlafaxine.

were not able to generate clear hierarchies among available treatments, because many antidepressants have not been directly compared. Therefore, we did a network meta-analysis to comprehensively compare and rank antidepressants for the acute treatment of major depressive disorder in children and adolescents.

Methods

Search strategy and selection criteria

For this network meta-analysis, we searched PubMed, the Cochrane Central Register of Controlled Trials, Web of Science, Embase, CINAHL, PsycINFO, and LiLACS for double-blind randomised controlled trials (RCTs) published from the date of database inception to May 31, 2015, comparing any antidepressant with placebo or another active antidepressant as oral monotherapy in the acute treatment of children and adolescents (mean age 9–18 years at enrolment in the primary study), with a primary diagnosis of major depressive disorder according to standardised diagnostic criteria. We also screened international trial registers and relevant reports on the FDA website, and searched key scientific journals in the field for published and unpublished studies (see appendix p 4 for more details).¹¹ We put no restrictions on language. For data about suicidality, we referred to the original papers, the Columbia University re-analysis of the FDA report,¹² and additionally searched the Medicines and Healthcare Products Regulatory Agency

Implications of all the available evidence

Our study has several implications for clinical practice. First, our findings suggest that fluoxetine should be considered the best available choice when a pharmacological treatment is indicated for moderate-to-severe depression in people younger than 18 years who do not have access to psychotherapy or have not responded to non-pharmacological interventions. Other antidepressants do not seem to be suitable as routine treatment options. Second, venlafaxine was found to be associated with an increased risk of suicidality in the young population. Because of the absence of reliable data on suicidality for many antidepressants, we could not comprehensively assess the risk of suicidality for all drugs. However, from a clinical perspective, children and adolescents taking antidepressant drugs should be closely monitored regardless of the treatment chosen, particularly at the beginning of treatment. Finally, we found that the methods used in individual studies were poor. Together with selective reporting, these are important limitations to be considered when interpreting the results from studies in such a population. Without access to individual patient-level data, we cannot be completely confident about the accuracy of information contained in published studies or clinical study reports.

database and pharmaceutical company websites for unpublished data. Study authors and drug manufacturers were contacted to supplement incomplete reports of the original papers or provide data for unpublished studies.

We included the following antidepressant interventions, but only if administered within the therapeutic dose range: amitriptyline, citalopram, clomipramine, desipramine, duloxetine, escitalopram, fluoxetine, imipramine, mirtazapine, nefazodone, nortriptyline, paroxetine, sertraline, and venlafaxine. Trials involving patients with comorbid, non-affective psychiatric disorders were included; however, RCTs recruiting participants with treatment-resistant depression, with treatment duration of less than 4 weeks, or with an overall sample size of fewer than ten patients were excluded.

Data extraction and quality assessment

Four investigators (AC, SEH, BQ, and YL) selected studies independently, and four investigators (YZ, BQ, YL, and LY) independently reviewed the main reports and supplementary materials, extracted the relevant information from the included trials with a predefined data extraction sheet, and assessed the risk of bias with the Cochrane risk of bias tool.¹³ Any discrepancies were resolved by consensus and arbitration by a panel of other investigators within the review team (AC, SEH, DC, SL, XZ, and PX).

See Online for appendix

Outcomes

We considered the mean overall change in depressive symptoms from baseline to endpoint and the proportion of patients who discontinued treatment due to any adverse events for our primary analyses. To measure improvement in depressive symptoms, the Children's Depression Rating Scale Revised (CDRS-R; a clinician-rated scale adapted for children and adolescents from the Hamilton Depression Rating Scale [HAMD], which is a tool validated and commonly used in adults [both CDRS-R and HAMD have good reliability and validity; appendix p 2]), and the Beck Depression Inventory and the Children's Depression Inventory (the two most commonly used among depression symptom severity self-rated scales) were used by the study investigators and we extracted the score data. When depression symptoms had been measured with more than one standardised rating scale, we used a predefined hierarchy, based on psychometric properties, frequency of use in children and adolescents, and consistency of use across included trials (appendix p 6). Secondary outcomes included response rate (estimated as the proportion of patients who achieved a reduction of 50% or more in depression rating score, or who scored much or very much improved on the Clinical Global Impression scale), all-cause discontinuation, and suicidal behaviour or ideation. We defined acute treatment as 8 weeks of treatment for both efficacy and tolerability analyses.¹⁴ If data at 8 weeks were not available, we used data ranging between 4 and 16 weeks (we gave preference to the timepoint used in the original study as the study endpoint).

Statistical analysis

Details of the applied statistical approaches are provided in the appendix (p 2, 8). First, we did pairwise meta-analyses with the random-effects model with STATA (version 13.0). The standardised mean difference (SMD) was calculated as the effect size for continuous outcomes and the odds ratio (OR) was calculated for dichotomous outcomes, both with 95% CI. We assessed statistical heterogeneity in each pairwise comparison with the I^2 statistic and p value.¹³ We used the funnel plot and Egger's test to detect publication bias, if at least ten studies were available.¹³

Second, we did a random-effects network meta-analysis within a Bayesian framework with WinBUGS (version 1.4.3) and further analysis with STATA (version 13.0) and R (version 3.2.2). We summarised the results of network meta-analysis with effect sizes (SMD or OR) and their credible intervals (CrI).¹⁵ See appendix (p 8) for details about the WinBUGS codes used. The 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.¹⁶ A common heterogeneity parameter was assumed for all

comparisons and we assessed the global heterogeneity using the I^2 statistic with the GeMTC R package (version 3.2.2).¹⁷ Inconsistency 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).¹⁸ The node splitting method was used to calculate the inconsistency of the model, which separated evidence on a particular comparison into direct and indirect evidence.¹⁹ We estimated the ranking probabilities for all treatments of being at each possible rank for each intervention. The treatment hierarchy was summarised and reported as surface under the cumulative ranking curve (SUCRA).²⁰ We also plotted a comparison-adjusted funnel plot for the network meta-analysis, to detect the presence of any dominant publication bias in network meta-analysis.²⁰ Additionally, we assessed the quality of evidence contributing to each network estimate using the GRADE framework, which characterises the quality of a body of evidence on the basis of the study limitations, imprecision, inconsistency, indirectness, and publication bias for the primary outcomes.²¹

To determine whether the results were affected by study characteristics, we did subgroup network meta-analyses in WinBUGS for primary outcomes according to the following variables: sex ratio, age group, treatment duration, severity of symptoms (see appendix p 12 for details), comorbid psychiatric disorder, quality of study, sample size, and sponsorship. Additionally, we did sensitivity network meta-analyses for primary outcomes by omitting unpublished trials, and trials in which response was imputed using remission rate.

This study is registered with PROSPERO, number CRD42015016023. The full dataset is available online.

Role of the funding source

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

Results

Overall, 5794 citations were identified by the search and 165 potentially eligible articles were retrieved in full text (figure 1). We excluded 138 reports, but included four additional studies from trial registers and pharmaceutical company websites, resulting in 31 publications describing 34 parallel RCTs (5260 patients) published between 1986 and 2014 and comparing 14 antidepressants or placebo (figure 1 and table 1). Seven of the included trials were unpublished and two were not in English (full references for all trials are given in the appendix p 14). The mean study sample size was 159 participants,

For the full dataset see
<http://ora.ox.ac.uk/objects/uuid:a4d1186b-6be4-486a-b9ce-ce431ccd545b>

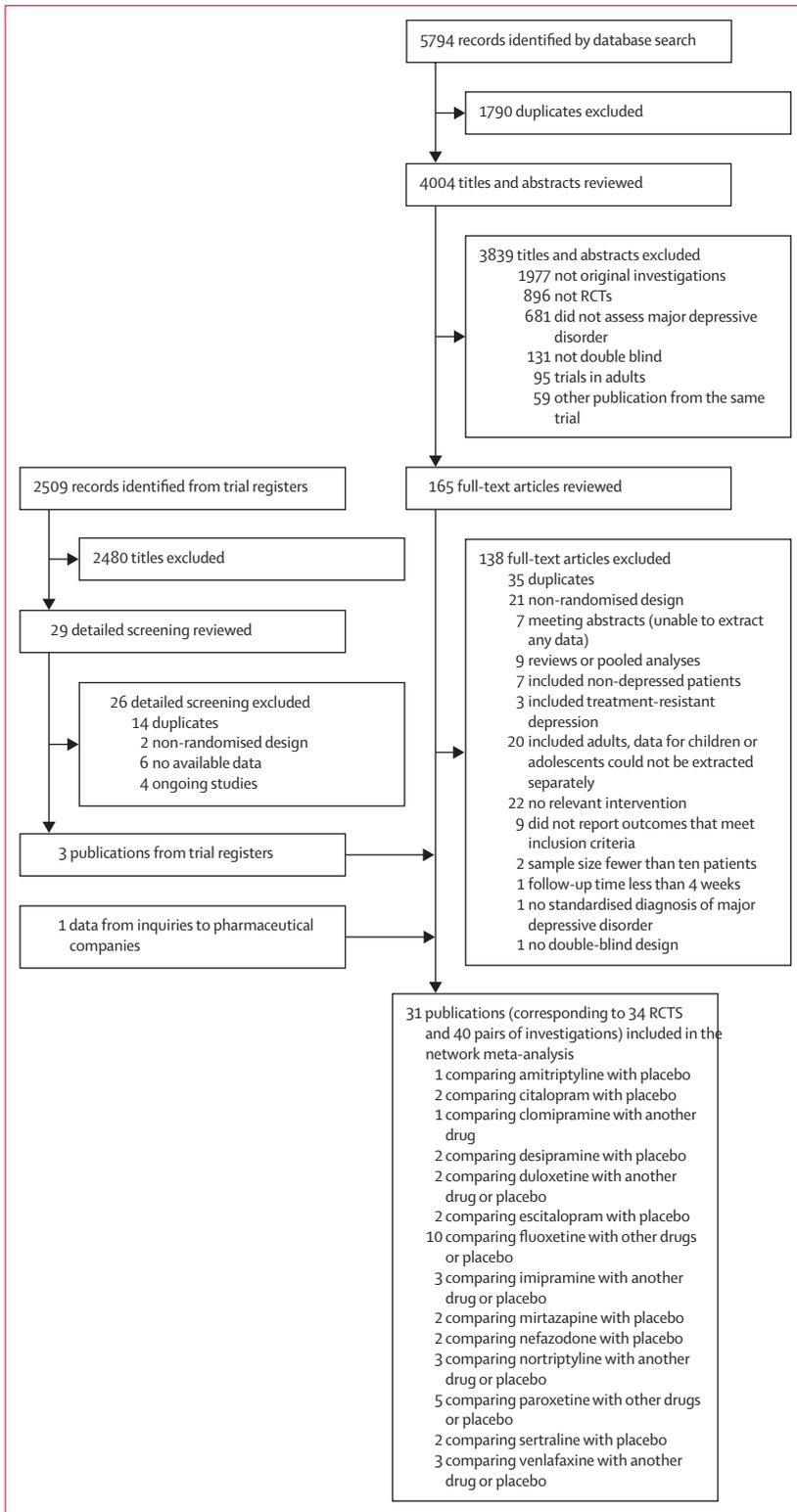


Figure 1: Study selection
RCT=randomised controlled trial.

ranging between 23 and 463 patients. Overall, 3106 participants were randomly assigned to an active drug and 2154 to placebo. Only 25 randomised controlled trials reported full clinical and demographic characteristics (table 1); about half of the sample population were female (2348 [53%] of 4432) and the mean age was 13.6 years (SD 2.87). The median duration of the acute treatment was 8 weeks (range 5–12) and 82% of trials (28 of 34) enrolled children and adolescents with moderate-to-severe depressive symptoms. 17 (50%) trials recruited patients from North America, five (15%) from across continents, and four (12%) from Europe. Of the remaining trials, four (12%) recruited participants from other countries and four (12%) trials did not specify. 22 (65%) trials were funded by pharmaceutical companies. In terms of study quality, ten (29%) trials were rated as high risk of bias, 20 (59%) trials as moderate, and four (12%) as low (appendix p 18).

Figure 2 shows the network of eligible comparisons for efficacy. For graphical representation of the other networks see appendix (p 22). All antidepressant drugs, except for clomipramine, had at least one placebo-controlled trial and five drugs were directly compared with at least one other active drug. Detailed results of pairwise meta-analyses are given in the appendix (p 25). Fluoxetine, sertraline, and escitalopram were statistically more efficacious than placebo in both continuous and dichotomous outcomes; fluoxetine was superior to nortriptyline, whereas paroxetine showed a significant benefit in terms of mean overall change in symptoms (not significant in terms of response rate) compared with clomipramine. For tolerability (as assessed by OR of discontinuation due to adverse effects), duloxetine, imipramine, sertraline, and venlafaxine were not as well tolerated as placebo; and paroxetine was not as well tolerated as imipramine (appendix p 29).

The results of the network meta-analyses for the primary outcomes is presented as a league table in figure 3. In terms of efficacy, only fluoxetine was better than placebo (SMD -0.51 , 95% CrI -0.99 to -0.03). Nortriptyline was significantly less effective than seven other antidepressants and placebo (SMDs ranging between -1.65 and -1.14). In terms of tolerability, fluoxetine was significantly better tolerated than duloxetine (OR 0.31, 95% CrI 0.13 to 0.95) and imipramine (0.23, 0.04 to 0.78), and citalopram and paroxetine were significantly better tolerated than imipramine alone (0.27, 0.04 to 0.96 and 0.22, 0.08 to 0.87, respectively). Imipramine was significantly less well tolerated than placebo (5.49, 1.96 to 20.86) as was venlafaxine (3.19, 1.01 to 18.70) and duloxetine (2.80, 1.20 to 9.42). Results for secondary outcomes of response rate and all-cause discontinuation were not materially different from, and lent support to, the findings for primary outcomes (appendix p 34). The only exception was nefazodone, which was more efficacious and better tolerated in the secondary analyses than in the primary analyses, but this evidence was based on only one

	Diagnostic criteria	Treatments, n (dose range)	Treatment duration (weeks)	Age range (mean)	Proportion of girls or women (%)	Recruiting area	Setting	Funder
Organon (2002a)	DSM-IV	Mirtazapine, 82 (15–45 mg/day) Placebo, 44	8	7–18 (12.3)	51%	Europe	Outpatients	Organon
Organon (2002b)	DSM-IV	Mirtazapine, 88 (15–45 mg/day) Placebo, 45	8	7–18 (12.0)	53%	Europe	Outpatients	Organon
Almeida-Montes et al (2005)	DSM-IV-TR	Fluoxetine, 12 (20 mg/day) Placebo, 11	6	8–14 (11.4)	35%	Mexico	Outpatients	None*
Atkinson et al (2014)	DSM-IV-TR	Duloxetine, 117 (60–120 mg/day) Fluoxetine, 117 (20–40 mg/day) Placebo, 103	10	7–17 (13.2)	52%	USA, Finland, France, Germany, Russia, Slovakia, Estonia, Ukraine, South Africa	Outpatients	Eli Lilly
Attari et al (2006)	DSM-IV	Fluoxetine, 20 (0.5–2 mg/day per kg) Nortriptyline, 20 (1–2 mg/day per kg)	8	7–16 (12.9)	50%	Iran	Outpatients	Not stated
Eli Lilly (1986)	DSM III	Fluoxetine, 21 (20–60 mg/day) Placebo, 19	6	12–17 (15.6)	55%	Canada	Inpatients and outpatients	Eli Lilly
Berard et al (2006)	DSM-IV	Paroxetine, 187 (20–40 mg/day) Placebo, 99	12	13–18 (15.6)	67%	Belgium, Italy, Mexico, UK, Spain, Netherlands, Canada, South Africa, United Arab Emirates, Argentina	Outpatients	GlaxoSmithKline
Braconnier et al (2003)	DSM-IV	Clomipramine, 58 (75–150 mg/day) Paroxetine, 63 (20–40 mg/day)	8	12–20 (16.1)	60%	France	..	GlaxoSmithKline
Bristol-Myers Squibb (2002)	DSM-IV	Nefazodone, 95 (100–300 mg/day) Nefazodone, 95 (200–600 mg/day) Placebo, 94	8	7–17 (..)	Bristol-Myers Squibb
Emslie et al (1997)	DSM III-R	Fluoxetine, 48 (20 mg/day) Placebo, 48	8	7–17 (12.4)	46%	USA	Outpatients	None
Emslie et al (2002a)	DSM-IV	Fluoxetine, 109 (10–20 mg/day) Placebo, 110	9	8–18 (12.7)	49%	USA	Outpatients	Eli Lilly
Emslie et al (2002b)	DSM-IV	Nefazodone, 99 (100–400 mg/day) Placebo, 96	8	12–17 (..)	59%	Bristol-Myers Squibb
Emslie et al (2006)	DSM-IV	Paroxetine, 104 (10–50 mg/day) Placebo, 102	8	7–17 (12.0)	47%	USA and Canada	..	GlaxoSmithKline
Emslie et al (2007)†	DSM-IV	Venlafaxine, 184 (37.5–225 mg/day) Placebo, 183	8	7–17 (12.3)	46%	USA	Outpatients	Wyeth Research
Emslie et al (2009)	DSM-IV	Escitalopram, 158 (10–20 mg/day) Placebo, 158	8	12–17 (14.6)	59%	USA	Outpatients	Forest Laboratories
Emslie et al (2014)	DSM-IV-TR	Duloxetine, 108 (60 mg/day) Duloxetine, 116 (30 mg/day) Fluoxetine, 117 (20 mg/day) Placebo, 122	10	7–17 (13.0)	51%	USA, Argentina, Canada, Mexico	Outpatients	Eli Lilly
Findling et al (2009)	DSM-IV	Fluoxetine, 18 (10–20 mg/day) Placebo, 16	8	12–17 (16.5)	15%	USA	Outpatients	Eli Lilly
Geller et al (1990)	DSM-III	Nortriptyline, 12 (45–140 mg/day) Placebo, 19	8	12–17 (14.3)	45%	USA	Outpatients	None
Geller et al (1992)	DSM III	Nortriptyline, 30 (10–140 mg/day) Placebo, 30	8	6–12 (9.7)	30%	..	Outpatients	None
Hongfen et al (2009)	CCMD-3	Fluoxetine, 30 (20 mg/day) Venlafaxine, 30 (150 mg/day)	8	12–18 (15.8)	47%	China	Inpatients and outpatients	Not stated
Hughes et al (1990)	DSM-III	Imipramine, 13 (..) Placebo, 14	6	6–12 (..)	Inpatients	None
Keller et al (2001)	DSM-IV	Imipramine, 95 (50–300 mg/day) Paroxetine, 93 (20–40 mg/day) Placebo, 87	8	12–18 (14.9)	62%	USA, Canada	..	GlaxoSmithKline
Klein et al (1998)	DSM-III-R	Desipramine, 23 (50–300 mg/day) Placebo, 22	6	13–18 (15.7)	67%	USA	Outpatients	None
Kutcher et al (1994)	DSM-III-R	Desipramine, 30 (200 mg/day) Placebo, 30	6	15–19 (17.8)	70%	Canada	Outpatients	None
Kye et al (1996)	K-SADS and RDC	Amitriptyline, 18 (5 mg/day per kg) Placebo, 13	8	12–17 (14.8)	29%	USA	Outpatients	None

(Table 1 continues on next page)

	Diagnostic criteria	Treatments, n (dose range)	Treatment duration (weeks)	Age range (mean)	Proportion of girls or women (%)	Recruiting area	Setting	Funder
(Continued from previous page)								
March et al (2004)	DSM-IV	Fluoxetine, 109 (10–40 mg/day) Placebo, 112	12	12–17 (14.6)	54%	USA	Outpatients	None
GlaxoSmithKline (2009)	DSM-IV-TR	Paroxetine, 29 (10–40 mg/day) Placebo, 27	8	7–17 (14.6)	61%	Japan	..	GlaxoSmithKline
Puig-Antich et al (1987)	K-SADS and RDC	Imipramine, 20 (3.25–5 mg/day per kg) Placebo, 22	5	6–12 (9.1)	40%	USA	Inpatients and outpatients	None
von Knorring et al (2006)	DSM-IV	Citalopram, 124 (10–40 mg/day) Placebo, 120	12	13–18 (16.0)	..	Europe	Inpatients and outpatients	Lundbeck
Wagner et al (2003a)	DSM-IV	Sertraline, 97 (50–200 mg/day) Placebo, 91	10	6–17 (..)	51%	USA, Costa Rica, India, Canada, Mexico	Outpatients	Pfizer
Wagner et al (2003b)	DSM-IV	Sertraline, 92 (50–200 mg/day) Placebo, 96	10	6–17 (..)	52%	USA, Costa Rica, India, Canada, Mexico	Outpatients	Pfizer
Wagner et al (2004)	DSM-IV	Citalopram, 93 (20–40 mg/day) Placebo, 85	8	7–17 (12.1)	53%	USA	..	Forest Laboratories
Wagner et al (2006)	DSM-IV	Escitalopram, 132 (10–20 mg/day) Placebo, 136	8	6–17 (12.3)	52%	USA	Outpatients	Forest Laboratories

Full references for all trials are given in the appendix (p 14). DSM=Diagnostic and Statistical Manual of Mental Disorders. DSM-IV-TR=Diagnostic and Statistical Manual of Mental Disorders fourth edition, text revision. CCMD-3=Chinese Classification of Mental Disorders third version. K-SADS=Kiddie-Schedule for Affective Disorders and Schizophrenia for School-Age Children. RDC=Research Diagnostic Criteria. *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 included in the systematic review and network meta-analysis

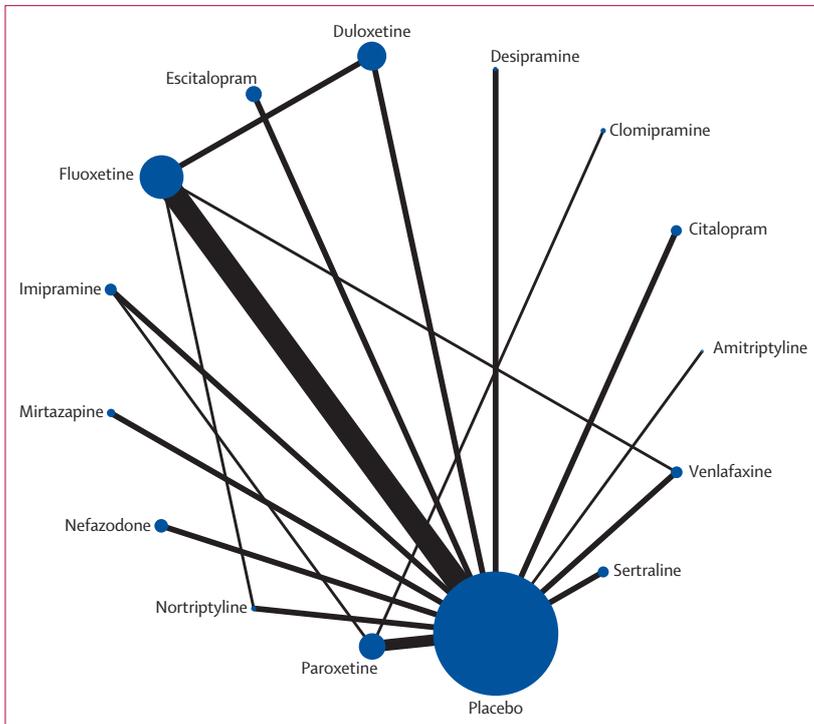


Figure 2: Network of eligible comparisons for efficacy
The width of the lines is proportional to the number of trials comparing every pair of treatments, and the size of every circle is proportional to the number of randomly assigned participants (sample size).

unpublished trial of poor quality, which is a major limitation for the reliability of findings (appendix p 35). Figure 4 shows the network meta-analyses results of the

suicide-related outcome (actual number of patients with suicidal behaviour or ideation are in table 2). Venlafaxine was associated with a significantly increased risk of suicidal behaviour or ideation compared with placebo (OR 0.13, 95% CrI 0.00–0.55) and five other antidepressants (escitalopram, imipramine, duloxetine, fluoxetine, and paroxetine).

The common heterogeneity SD was 0.68 (95% CrI 0.47–0.94) for efficacy, 0.42 (0.02–0.94) for tolerability, 0.22 (0.01–0.50) for response rate, 0.32 (0.03–0.67) for all-cause discontinuation, and 0.30 (0.01–0.84) for suicide-related outcome. The global I^2 values were 33.21% for efficacy and 0% for tolerability. The test of global inconsistency showed a significant difference between the consistency and inconsistency models for efficacy ($p < 0.0001$), but not for tolerability ($p = 0.8432$; appendix p 36). Tests of local inconsistency showed that the percentages for inconsistent loops were to be expected based on the empirical data (two of four comparison loops for the efficacy outcome and zero of two for tolerability outcome; for details of the assessments of consistency see appendix p 38). The test of inconsistency from the node-splitting model showed significant differences between some comparisons in primary efficacy, but not for tolerability (appendix p 41). The comparison-adjusted funnel plots of the network meta-analysis for primary outcomes were not suggestive of any publication bias (appendix p 44).

The ranking of treatments based on cumulative probability plots and SUCRAs is presented in the appendix (p 48). In terms of efficacy, the most effective

FLU	0.18 (0.04 to 1.75)	0.31 (0.13 to 0.95)	0.39 (0.05 to 1.47)	0.91 (0.09 to 3.49)	0.43 (0.06 to 1.58)	0.69 (0.24 to 3.50)	0.30 (0.07 to 3.06)	0.78 (0.21 to 2.18)	0.96 (0.05 to 4.56)	0.23 (0.04 to 0.78)	0.11 (0.03 to 77.0)	1.03 (0.50 to 2.70)	0.43 (0.11 to 3.98)	..
-0.06 (-1.23 to 1.11)	DES	2.05 (0.18 to 8.72)	1.79 (0.11 to 8.70)	4.23 (0.19 to 20.49)	1.96 (0.12 to 9.40)	1.90 (0.47 to 22.04)	3.55 (0.16 to 17.62)	3.61 (0.39 to 14.96)	4.38 (0.11 to 24.29)	1.09 (0.09 to 4.95)	0.32 (0.09 to 377.5)	2.85 (0.83 to 21.80)	1.17 (0.25 to 22.42)	..
-0.16 (-1.05 to 0.72)	-0.10 (-1.49 to 1.28)	DUL	1.17 (0.13 to 4.68)	2.73 (0.23 to 11.00)	1.27 (0.14 to 5.05)	1.89 (0.61 to 11.49)	2.27 (0.19 to 9.64)	2.32 (0.53 to 7.26)	2.88 (0.12 to 14.26)	0.68 (0.11 to 2.54)	0.33 (0.09 to 224.8)	2.80 (1.20 to 9.42)	1.17 (0.30 to 12.58)	..
-0.25 (-1.13 to 0.64)	-0.19 (-1.54 to 1.17)	-0.09 (-1.28 to 1.10)	VEN	1.17 (0.23 to 18.93)	0.61 (0.14 to 8.69)	2.13 (0.56 to 20.32)	0.92 (0.19 to 16.01)	1.69 (0.47 to 13.33)	0.71 (0.13 to 22.98)	0.40 (0.10 to 4.39)	0.37 (0.10 to 343.2)	3.19 (1.01 to 18.70)	1.30 (0.29 to 21.16)	..
-0.27 (-1.39 to 0.84)	-0.21 (-1.68 to 1.26)	-0.11 (-1.46 to 1.23)	-0.02 (-1.33 to 1.28)	MIR	0.93 (0.06 to 4.52)	0.91 (0.23 to 10.97)	0.40 (0.08 to 8.51)	1.71 (0.20 to 7.49)	2.12 (0.05 to 12.37)	0.17 (0.04 to 2.27)	0.18 (0.05 to 151.0)	1.36 (0.41 to 10.99)	0.56 (0.12 to 10.82)	..
-0.28 (-1.38 to 0.82)	-0.22 (-1.68 to 1.24)	-0.12 (-1.46 to 1.21)	-0.03 (-1.34 to 1.27)	-0.01 (-1.43 to 1.40)	SER	1.98 (0.52 to 18.57)	0.85 (0.17 to 15.06)	1.56 (0.44 to 12.09)	0.64 (0.11 to 21.50)	0.37 (0.09 to 4.05)	0.35 (0.09 to 304.7)	2.94 (0.94 to 17.19)	1.20 (0.27 to 18.95)	..
-0.33 (-1.43 to 0.78)	-0.27 (-1.72 to 1.20)	-0.17 (-1.50 to 1.17)	-0.08 (-1.38 to 1.22)	-0.06 (-1.47 to 1.35)	-0.05 (-1.45 to 1.35)	CIT	0.91 (0.07 to 3.89)	0.93 (0.20 to 2.77)	1.17 (0.05 to 5.60)	0.27 (0.04 to 0.96)	0.13 (0.03 to 91.24)	1.13 (0.45 to 3.66)	1.18 (0.11 to 4.76)	..
-0.34 (-1.44 to 0.75)	-0.28 (-1.73 to 1.17)	-0.18 (-1.51 to 1.15)	-0.09 (-1.39 to 1.20)	-0.07 (-1.48 to 1.34)	-0.06 (-1.45 to 1.34)	-0.01 (-1.41 to 1.40)	ESC	2.19 (0.22 to 9.23)	2.67 (0.06 to 14.62)	0.63 (0.05 to 2.87)	0.16 (0.05 to 196.9)	1.64 (0.46 to 13.49)	0.68 (0.14 to 13.64)	..
-0.35 (-1.19 to 0.50)	-0.29 (-1.56 to 0.99)	-0.19 (-1.32 to 0.94)	-0.10 (-1.19 to 0.99)	-0.07 (-1.30 to 1.15)	-0.07 (-1.28 to 1.16)	-0.02 (-1.23 to 1.19)	-0.01 (-1.21 to 1.20)	PAR	0.35 (0.07 to 6.80)	0.22 (0.08 to 0.87)	0.19 (0.05 to 115.6)	1.59 (0.77 to 3.95)	0.79 (0.26 to 3.77)	..
-0.36 (-1.46 to 0.74)	-0.30 (-1.76 to 1.15)	-0.20 (-1.54 to 1.13)	-0.11 (-1.42 to 1.18)	-0.09 (-1.50 to 1.32)	-0.08 (-1.48 to 1.32)	-0.03 (-1.44 to 1.37)	-0.02 (-1.42 to 1.37)	-0.01 (-1.23 to 1.19)	NEF	0.16 (0.03 to 4.50)	0.11 (0.04 to 241.2)	1.29 (0.30 to 21.89)	0.52 (0.10 to 20.79)	..
-0.49 (-1.57 to 0.58)	-0.44 (-1.88 to 1.01)	-0.33 (-1.65 to 0.98)	-0.25 (-1.53 to 1.03)	-0.22 (-1.61 to 1.17)	-0.22 (-1.60 to 1.17)	-0.17 (-1.55 to 1.22)	-0.16 (-1.54 to 1.22)	-0.15 (-1.21 to 0.91)	-0.13 (-1.52 to 1.26)	IMP	0.67 (0.17 to 471.9)	5.49 (1.96 to 20.86)	2.47 (0.62 to 21.47)	..
-0.59 (-2.21 to 1.01)	-0.53 (-2.39 to 1.33)	-0.43 (-2.20 to 1.34)	-0.34 (-2.09 to 1.40)	-0.32 (-2.14 to 1.52)	-0.31 (-2.13 to 1.52)	-0.26 (-2.10 to 1.57)	-0.25 (-2.08 to 1.57)	-0.24 (-1.92 to 1.43)	-0.23 (-2.05 to 1.59)	-0.10 (-1.92 to 1.71)	AMI	0.10 (0.02 to 32.16)	6.38 (0.01 to 24.56)	..
-0.51 (-0.99 to -0.03)	-0.45 (-1.52 to 0.62)	-0.35 (-1.24 to 0.54)	-0.26 (-1.10 to 0.58)	-0.24 (-1.25 to 0.77)	-0.23 (-1.21 to 0.77)	-0.18 (-1.18 to 0.82)	-0.17 (-1.15 to 0.81)	-0.16 (-0.86 to 0.54)	-0.15 (-1.14 to 0.85)	-0.01 (-0.98 to 0.95)	0.08 (-1.45 to 1.61)	PBO	0.79 (0.12 to 2.75)	..
-0.83 (-2.48 to 0.81)	-0.77 (-2.67 to 1.13)	-0.67 (-2.49 to 1.14)	-0.58 (-2.38 to 1.20)	-0.56 (-2.43 to 1.32)	-0.55 (-2.42 to 1.31)	-0.50 (-2.36 to 1.36)	-0.49 (-2.35 to 1.36)	-0.48 (-1.90 to 0.93)	-0.47 (-2.33 to 1.39)	-0.34 (-2.10 to 1.43)	-0.24 (-2.43 to 1.95)	-0.32 (-1.90 to 1.25)	CLO	..
-1.65 (-2.57 to -0.72)	-1.59 (-2.98 to -0.21)	-1.49 (-2.71 to -0.27)	-1.40 (-2.60 to -0.20)	-1.38 (-2.71 to -0.04)	-1.37 (-2.70 to -0.05)	-1.32 (-2.65 to 0.01)	-1.31 (-2.63 to 0.01)	-1.30 (-2.43 to -0.18)	-1.29 (-2.61 to 0.04)	-1.15 (-2.46 to 0.15)	-1.06 (-2.81 to 0.71)	-1.14 (-2.02 to -0.25)	-0.82 (-2.61 to 0.99)	NOR

■ Treatment □ Efficacy (mean overall change in symptoms, SMD [95% CrI]) □ Tolerability (discontinuation due to adverse events, OR [95% CrI])

Figure 3: Network meta-analysis of efficacy and tolerability

Drugs are reported in order of efficacy ranking according to SUCRAs. Comparisons should be read from left to right. The efficacy and tolerability estimate is located at the intersection of the column-defining treatment and the row-defining treatment. For efficacy (mean overall change in symptoms), an SMD below 0 favours the column-defining treatment. For tolerability (discontinuation due to adverse events), an OR below 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 and underlined. AMI=amitriptyline. CIT=citalopram. CLO=clomipramine. CrI=credibility interval. DES=desipramine. DUL=duloxetine. ESC=escitalopram. FLU=fluoxetine. IMP=imipramine. MIR=mirtazapine. NEF=nefazodone. NOR=nortriptyline. OR=odds ratio. PAR=paroxetine. PBO=placebo. SER=sertraline. SMD=standardised mean difference. SUCRA=surface under the cumulative ranking curve. VEN=venlafaxine.

treatment was fluoxetine (76.6%) and the least effective was nortriptyline (3.7%). In terms of tolerability, fluoxetine was the best drug (75.7%) and imipramine the worst (13.1%). According to GRADE, the quality of evidence for primary outcomes was rated as very low for most comparisons (appendix p 77–86). Quality of evidence was very low for overall ranking of treatment in terms of efficacy and low for tolerability (appendix p 82, 86). We also studied the effect of several potential moderator variables for the primary outcomes in subgroup analyses, the findings of which were not materially different from those of the primary analysis for most of these comparisons (appendix p 87). Preplanned sensitivity analyses did not affect the main results (appendix p 92).

Discussion

This network meta-analysis represents the most comprehensive synthesis of data for currently available pharmacological treatments for children and adolescents with acute major depressive disorder. We found that only fluoxetine was significantly more effective than placebo, and the corresponding SMD of 0.51 is considered to be a medium effect size.²² However, the large credible interval and its upper limit close to the point of no difference

IMP														
0.42 (0.09 to 5.35)	PBO													
0.40 (0.08 to 5.98)	1.07 (0.51 to 2.01)	DUL												
0.34 (0.07 to 6.37)	1.08 (0.33 to 2.57)	1.14 (0.27 to 3.18)	ESC											
0.32 (0.06 to 7.60)	1.41 (0.18 to 5.33)	0.61 (0.15 to 6.04)	0.61 (0.14 to 7.66)	CLO										
0.35 (0.07 to 4.88)	0.90 (0.49 to 1.49)	0.91 (0.43 to 1.70)	1.09 (0.29 to 2.98)	1.37 (0.15 to 5.34)	FLU									
0.96 (0.07 to 3.96)	0.89 (0.27 to 2.17)	0.94 (0.21 to 2.66)	1.09 (0.19 to 3.62)	1.02 (0.21 to 3.03)	1.08 (0.27 to 2.92)	PAR								
0.26 (0.05 to 5.66)	0.89 (0.22 to 2.53)	0.57 (0.19 to 2.98)	0.57 (0.17 to 4.00)	0.44 (0.10 to 6.02)	0.69 (0.23 to 3.31)	0.69 (0.20 to 4.73)	CIT							
0.73 (0.02 to 4.05)	0.57 (0.06 to 2.05)	0.60 (0.05 to 2.35)	0.69 (0.05 to 2.92)	0.87 (0.03 to 4.29)	0.69 (0.06 to 2.62)	0.85 (0.06 to 3.75)	0.93 (0.05 to 4.06)	SER						
0.16 (0.00 to 0.96)	0.13 (0.00 to 0.55)	0.14 (0.00 to 0.64)	0.16 (0.00 to 0.79)	0.19 (0.00 to 1.04)	0.15 (0.00 to 0.72)	0.19 (0.00 to 0.90)	0.21 (0.00 to 1.07)	0.51 (0.00 to 3.11)	VEN					

■ Treatment □ Suicidal behaviour or ideation (OR [95% CrI])

Figure 4: Network meta-analysis of suicide-related outcome

Drugs are reported in order of suicide-related outcome ranking according to SUCRAs. Comparisons should be read from left to right. The estimate is located at the intersection of the column-defining treatment and the row-defining treatment. For suicidal behaviour or ideation, an OR value below 1 favours the column-defining treatment. To obtain ORs for comparisons in the opposing direction, reciprocals should be taken. Significant results are in bold and underlined. CIT=citalopram. CLO=clomipramine. CrI=credibility interval. DUL=duloxetine. ESC=escitalopram. FLU=fluoxetine. IMP=imipramine. OR=odds ratio. PBO=placebo. PAR=paroxetine. SER=sertraline. SUCRA=surface under the cumulative ranking curve. VEN=venlafaxine.

raise the question of whether this estimate is robust enough to inform clinical practice. By comparison with placebo, fluoxetine was significantly more effective in

	Number of trials	Events/total (%)	
		Group 1	Group 2
Citalopram vs placebo	2	10/217 (5%)	7/205 (3%)
Duloxetine vs placebo	2	44/341 (13%)	32/225 (14%)
Escitalopram vs placebo	2	15/290 (5%)	15/294 (5%)
Fluoxetine vs placebo	7	51/521 (10%)	44/514 (9%)
Imipramine vs placebo	2	2/95 (2%)	1/87 (1%)
Mirtazapine vs placebo	2	1/170 (1%)	0/89
Nefazodone vs placebo	2	0/289	0/190
Paroxetine vs placebo	4	13/413 (3%)	7/315 (2%)
Sertraline vs placebo	2	5/189 (3%)	2/187 (1%)
Venlafaxine vs placebo	2	8/184 (4%)	0/183
Clomipramine vs paroxetine	1	7/58 (12%)	9/63 (14%)
Duloxetine vs fluoxetine	2	44/341 (13%)	33/234 (14%)
Imipramine vs paroxetine	1	2/95 (2%)	5/93 (5%)

Table 2: Number of patients with suicidal behaviour or ideation according to study treatment

trials without industry sponsors; however, a possible explanation is that trials without industry sponsors tend to have a smaller sample size, which might result in an exaggerated treatment effect.²³ By comparison with other antidepressants, fluoxetine was significantly more effective than nortriptyline and, in terms of discontinuations due to adverse events, fluoxetine was better tolerated than imipramine and duloxetine. However, the clinical interpretation of these findings is limited not only by the uncertainty around these estimates, but also by the potential bias due to selective reporting and the small number of trials in each node. 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 or that published reports might overestimate the efficacy of treatments.^{9,24} Moreover, poor methodology, risk of bias within individual studies, and potential selective reporting are important factors to be considered when interpreting the results from this meta-analysis. Without access to individual patient-level data, we cannot be confident about the accuracy of information contained in published studies or even clinical study reports.

Since 2003, many international agencies, including the European Medicines Agency, the US FDA, and the Medicines and Healthcare Products Regulatory Agency in the UK, have added a black box warning (the most serious type of warning) to the prescription drug labelling of antidepressants, indicating that they might increase the risk of suicidal thinking and behaviour in some children and adolescents with major depressive disorder.⁷ Our analysis found robust evidence to suggest a significantly increased risk for suicidality (suicidal behaviour or ideation) for young people given

venlafaxine. Unfortunately, due to the absence of reliable data on suicidality for many antidepressants, it was not possible to comprehensively assess the risk of suicidality for all drugs. However, from a clinical perspective, the decision maker should always consider the overall clinical picture, and patient management plans need to balance the risks and benefits. Children and adolescents taking antidepressant drugs should be closely monitored regardless of the treatment chosen, particularly at the beginning of treatment.⁴

This study has some limitations. First, in the GRADE framework, many comparisons were assessed as low or very low quality, which largely restricts the interpretation of these results. In the network, we found inconsistency for efficacy, which was mainly determined by the loop of fluoxetine–nortriptyline–placebo (we did not find heterogeneity for the tolerability outcome, probably because the proportion of patients who dropped out is a harder outcome than efficacy measured on a rating scale). We believe that this inconsistency might be a consequence of a cohort effect that relates to different methods used in the older studies compared with those done more recently. Some evidence suggests that quality of psychopharmacological clinical trials has substantially changed in the past 30 years²⁵ and other network meta-analyses confirmed similar findings.²⁶ Second, the review was restricted to trials involving children and adolescents with major depressive disorder. We excluded studies in which participants were described as having subsyndromal depressive symptoms, which is a significant proportion of patients seen in real-world, clinical settings. Similarly, we excluded patients with treatment-resistant depression. We did this to reduce heterogeneity and inconsistency among trials in the network meta-analysis, but acknowledge that it restricts the external validity of the results. Additionally, omission of trials of treatment-resistant depression might have led to an overestimation of efficacy in this meta-analysis, because patients who are treatment resistant are clearly a difficult-to-treat population. Third, we are aware of the Restoring Study 329,²⁴ which found different results to those of the original Study 329 when the original protocol was used to analyse the data. Because Restoring Study 329 was published after our last update of the search, we included the original study data, which were biased in favour of paroxetine over placebo. Findings from our review, however, were not affected by the results from this single study, because paroxetine overall did not show any statistical difference when compared with placebo in our analysis. The example of Restoring Study 329 supports the added value of network meta-analysis, which provides a more reliable estimate in terms of comparative efficacy.²⁷ Finally, too few studies were included to be able to do a network meta-analysis that addressed the clinically important issue of antidepressant therapy for preventing relapse of depression in children and adolescents. Some of the

adverse effects of antidepressants occur over a long period, meaning that positive results from short-term studies need to be interpreted with great caution. However, some long-term data suggest that fluoxetine can be well tolerated and effective in reducing the risk of relapse in children and adolescents with major depressive disorder after 32 weeks of treatment.²⁸ These data should be analysed and contextualised at the individual patient level (many adolescent patients with major depressive disorder are given drugs in adulthood and efficacy of drugs can change over time),¹⁴ but the data support the use of fluoxetine as a long-term treatment in principle, if effective in the acute phase for children and adolescents.

The findings of this comprehensive network meta-analysis provide some evidence that fluoxetine might reduce depressive symptoms in children and adolescents with major depressive disorder and the extent to which this reduction is clinically meaningful is still uncertain. Notwithstanding these caveats, fluoxetine might still be considered the best option among antidepressants when a pharmacological treatment is indicated. Other antidepressants do not seem to be suitable as routine treatment options. In the clinical care of young people with major depressive disorder, clinical guidelines recommend psychotherapy (especially cognitive-behavioural therapy or interpersonal therapy)^{4,5,29} as the first-line intervention, and fluoxetine should be considered only for patients with moderate-to-severe depression (especially adolescents)¹⁰ who do not have access to psychotherapy (eg, in low-income and middle-income countries)³⁰ or have not responded to non-pharmacological interventions. Antidepressants are not well studied in this population, and further research on moderators of treatment effect and possible new interventions are needed. In all these cases, however, clinicians should carefully look for the emergence or exacerbation of suicidality and balance the risk–benefit profile of antidepressants during the acute treatment phase.

Contributors

AC, XZ, CDG, BQ, CW, YZ, DCoh, YL, KDM, and PX conceived and designed the study. AC, XZ, BQ, YZ, JP, YL, LY, and LL selected the articles and extracted the data. AC, XZ, and CDG analysed the data. AC, XZ, and PX wrote the first draft of the manuscript. SEH, CW, DCog, PH, SL, PC, AVR, and KDM interpreted the data and wrote the final version. All authors read and met the ICMJE criteria for authorship and agree with the results and conclusions of this Article.

Declaration of interests

AC reports personal fees from Accord Healthcare as an expert witness for a patent issue about quetiapine extended release. SEH is an Editor of the Cochrane Common Mental Disorders Group and an author of the Cochrane systematic review of newer generation antidepressants for depression in children and adolescents. CW has worked for Doctor Evidence, a health care technology company, since May, 2015, that develops software and services related to evidence synthesis, and has clients that manufacture, or are developing, antidepressant drugs. DCog reports grants and personal fees from Shire; personal fees from Eli Lilly, Janssen Cilag, Novartis, Sandoz, Oxford University Press; and grants from European Union FP7, outside of the submitted work. PF reports fees paid to his employer for preparing and delivering educational

material from Eli Lilly and for participating in the advisory board and convening a conference from Shire, outside of the submitted work. SL reports honoraria for lectures from Eli Lilly, Lundbeck, Pfizer, Janssen, Bristol-Myers Squibb, Johnson & Johnson, Otsuka, Roche, Sanofi, ICON, AbbVie, AOP Orphan, and Servier; for consulting or advisory boards from Roche, Janssen, Lundbeck Institute, Eli Lilly, Otsuka, and Teva; and for the preparation of educational material and publications from Lundbeck Institute and Roche. Eli Lilly has provided drugs for a clinical trial led by SL as the principal investigator. DCoh reports personal fees, past consultation for, and honoraria from Otsuka, Shire, Lundbeck, and IntegraGen, outside of the submitted work. AVR reports personal fees from Bristol Myers Squibb, Pfizer, Sunovion; grants from Pfizer, Grand Challenges Canada, Canadian Institute of Health Research, and AstraZeneca, outside of the submitted work. XZ, CDG, BQ, YZ, PC, JP, YL, KDM, LY, LL, and PX declare no competing interests.

Acknowledgments

We thank Graham J Emslie and Taryn Mayes from the University of Texas Southwestern Medical Center in Dallas for providing unpublished data in this review, and Salvatore Gentile from the Department of Mental Health, ASL Salerno Mental Health Center, Italy, for his valuable advice. AC is supported by the National Institute for Health Research Oxford Cognitive Health Clinical Research Facility, National Basic Research Program of China (973 Program) (grant number 2009CB918300).

References

- 1 Costello JE, Erkanli A, Angold A. Is there an epidemic of child or adolescent depression? *J Child Psychol Psychiatry* 2006; **47**: 1263–71.
- 2 Olfson M, Blanco C, Wang S, Laje G, Correll CU. National trends in the mental health care of children, adolescents, and adults by office-based physicians. *JAMA Psychiatry* 2014; **71**: 81–90.
- 3 Thapar A, Collishaw S, Pine DS, Thapar AK. Depression in adolescence. *Lancet* 2012; **379**: 1056–67.
- 4 Hopkins K, Crosland P, Elliott N, Bewley S, for the Clinical Guidelines Update Committee B. Diagnosis and management of depression in children and young people: summary of updated NICE guidance. *BMJ* 2015; **350**: h824.
- 5 Malhi GS, Bassett D, Boyce P, et al. Royal Australian and New Zealand College of Psychiatrists clinical practice guidelines for mood disorders. *Aust N Z J Psychiatry* 2015; **49**: 1087–206.
- 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 US Food and Drug Administration. Suicidality in children and adolescents being treated with antidepressant medications. Oct 15, 2004. www.fda.gov/safety/medwatch/safetyinformation/safetyalertsforhumanmedicalproducts/ucm155488.htm (accessed Oct 18, 2015).
- 8 Hetrick SE, McKenzie JE, Cox GR, Simmons MB, Merry SN. Newer generation antidepressants for depressive disorders in children and adolescents. *Cochrane Database Syst Rev* 2012; **11**: CD004851.
- 9 Whittington CJ, Kendall T, Fonagy P, Cottrell D, Cotgrove A, Boddington E. Selective serotonin reuptake inhibitors in childhood depression: systematic review of published versus unpublished data. *Lancet* 2004; **363**: 1341–45.
- 10 Tsapakis EM, Soldani F, Tondo L, Baldessarini RJ. Efficacy of antidepressants in juvenile depression: meta-analysis. *Br J Psychiatry* 2008; **193**: 10–17.
- 11 Zhou X, Qin B, Whittington C, et al. Comparative efficacy and tolerability of first-generation and newer-generation antidepressant medications for depressive disorders in children and adolescents: study protocol for a systematic review and network meta-analysis. *BMJ Open* 2015; **5**: e007768.
- 12 Hammad TA. Relationship between psychotropic drugs and pediatric suicidality. 16 August 2004. <http://www.fda.gov/ohrms/dockets/ac/04/briefing/2004-4065b1-10-TAB08-Hammads-Review.pdf> (accessed May 12, 2016).
- 13 The Cochrane Collaboration. Cochrane handbook for systematic reviews of interventions (Version 5.1.0). March, 2011. <http://handbook.cochrane.org> (accessed Oct 20, 2015).
- 14 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.

- 15 Salanti G, Higgins JP, Ades AE, Ioannidis JP. Evaluation of networks of randomized trials. *Stat Methods Med Res* 2008; **17**: 279–301.
- 16 Brooks SP, Gelman A. General methods for monitoring convergence of iterative simulations. *J Comput Graph Stat* 1998; **7**: 434–55.
- 17 Neupane B, Richer D, Bonner AJ, Kibret T, Beyene J. Network meta-analysis using R: a review of currently available automated packages. *PLoS One* 2014; **9**: e115065.
- 18 Chaimani A, Higgins JP, Mavridis D, Spyridonos P, Salanti G. Graphical tools for network meta-analysis in STATA. *PLoS One* 2013; **8**: e76654.
- 19 Dias S, Welton NJ, Caldwell DM, Ades AE. Checking consistency in mixed treatment comparison meta-analysis. *Stat Med* 2010; **29**: 932–44.
- 20 Salanti G, Ades AE, Ioannidis JP. Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. *J Clin Epidemiol* 2011; **64**: 163–71.
- 21 Salanti G, Del Giovane C, Chaimani A, Caldwell DM, Higgins JP. Evaluating the quality of evidence from a network meta-analysis. *PLoS One* 2014; **9**: e99682.
- 22 Cohen J. *Statistical power analysis for the behavioral sciences* (2nd edn). Hillsdale, NJ: Lawrence Erlbaum Associates, 1988.
- 23 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.
- 24 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.
- 25 Brunoni AR, Tadini L, Fregni F. Changes in clinical trials methodology over time: a systematic review of six decades of research in psychopharmacology. *PLoS One* 2010; **5**: e9479.
- 26 Samara MT, Dold M, Gianatsi M, et al. Efficacy, acceptability, and tolerability of antipsychotics in treatment-resistant schizophrenia: a network meta-analysis. *JAMA Psychiatry* 2016; **73**: 199–210.
- 27 Cipriani A, Higgins JP, Geddes JR, Salanti G. Conceptual and technical challenges in network meta-analysis. *Ann Intern Med* 2013; **159**: 130–37.
- 28 Emslie GJ, Heiligenstein JH, Hoog SL, et al. Fluoxetine treatment for prevention of relapse of depression in children and adolescents: a double-blind, placebo-controlled study. *J Am Acad Child Adolesc Psychiatry* 2004; **43**: 1397–405.
- 29 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.
- 30 Kakuma R, Minas H, van Ginneken N, et al. Human resources for mental health care: current situation and strategies for action. *Lancet* 2011; **378**: 1654–63.

Antidepressants fail, but no cause for therapeutic gloom



The careful study by Andreas Cipriani and colleagues in *The Lancet*¹ has disturbing implications for clinical practice, concluding as it does that the risk-benefit profile of antidepressants in the acute treatment of depression does “not seem to offer a clear advantage for children and adolescents”.

The case for antidepressants, including fluoxetine, is, in fact, even weaker than their meta-analysis suggests. The authors do express appropriate scepticism about the quality and potential bias of data analysed. But, they were not able to factor in the additional problems and consequences of probable data misrepresentation by the companies that did the primary studies. While the common manoeuvre of changing nominated primary outcomes would not have an impact on the findings from this meta-analysis, other manipulations will, such as failing to exclude unblinded patients from the efficacy analysis.²

Similarly, although discontinuation due to adverse events is a relatively hard outcome, it can be underestimated in published papers and clinical study reports because of miscoding.³ Furthermore, the suicidal event data available to Cipriani and colleagues are likely to be substantially underestimated in the drug groups. In four trials of paroxetine versus placebo, only 13 (3%) of 413 events were reported in the paroxetine group; this seems implausible when individual patient-level data reanalysis of just one of those studies found ten events in only 93 patients given paroxetine (10.8%).³

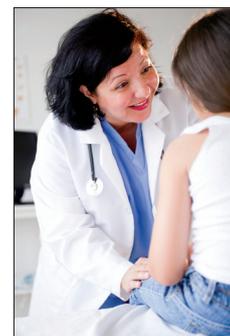
So what are the implications for clinicians? Every decision about whether and what to prescribe needs a complex and partly intuitive calculation of the balance between harms and benefits according to the patient's circumstances. With research evidence as an important part of that calculation, we now know that we need to make a conscious correction for favourable misrepresentation of outcomes in published and unpublished study reports. A reduction should be applied to the reported benefit of a drug, while routinely assuming that its harms are more serious and frequent than reported. Only if the discounted benefit outweighs the boosted harm should the treatment be prescribed. For antidepressants in adolescents, this equation will rarely favour prescribing; in younger children, almost never.

Opposing this approach is the fact that most psychiatrists and many general practitioners

(family doctors) have vast experience of prescribing antidepressants to adolescents, and many will believe that their clinical experience overrides any scepticism introduced by Cipriani and colleagues' study. Fair enough, so long as they are honest with themselves and their patients that such prescribing is unsupported by evidence from randomised controlled trials.

A second clinical implication is that we need an alternative to the guideline recommendation of prescribing antidepressants when evidence-based psychotherapy is unavailable. Cipriani and colleagues accept that careful prescribing of fluoxetine is justified because of the unacceptable consequences of not intervening at all—the association of depression with serious impairment. But association is not causation. Although short-term symptomatic relief is plausibly associated with better health, academic, and interpersonal outcomes, this association has not been shown. So, clinicians should not be pressured into prescribing just because they have no capacity to offer evidence-based psychotherapy. The effect of misreporting is that antidepressants, possibly including fluoxetine, are likely to be more dangerous and less effective treatments than has been previously recognised, so there is little reason to think that any antidepressant is better than nothing for young people. Prescribing might help the doctor feel like he or she is doing something, or help parents feel that something is being done, but the adolescent might feel it to be dismissive of their distress.⁴

Furthermore, it is a mistake to think that clinicians have nothing to offer other than drugs if evidence-based psychotherapy is not available within a health system. The therapeutic potency of a relationship with a benign, supportive clinician remains active outside of a formal psychotherapeutic framework, and careful prescribing of fluoxetine has not been established as more efficacious than such a relationship. An example of this approach is watchful waiting—a clinical posture shown to be better than nothing.⁵ Not shown to be inferior to fluoxetine, it is less toxic than fluoxetine, and available to competent clinicians at all levels of therapeutic skill. Watchful waiting invites the clinician to engage with a young person, and elucidate and, wherever possible, act on the predicament driving their distress.



Jodi/Jamie Grill/Getty Images

Published Online
June 8, 2016
[http://dx.doi.org/10.1016/S0140-6736\(16\)30585-2](http://dx.doi.org/10.1016/S0140-6736(16)30585-2)

See Online/Articles
[http://dx.doi.org/10.1016/S0140-6736\(16\)30385-3](http://dx.doi.org/10.1016/S0140-6736(16)30385-3)

Finally, Cipriani and colleagues provide a vital message about how data are managed. They report that they could not be confident in the accuracy of published information used for their analysis because they were unable to access individual patient-level data. This is a shocking conclusion that appears to be applicable to all pharmaceutical research reports, including unpublished clinical study reports. Patients who take part in randomised controlled trials have a right to expect that maximum benefit will come from the data they generate. We doctors and researchers are failing to meet our obligation to research participants and to our patients, and we will only succeed if independent researchers such as Cipriani and colleagues are able to analyse individual patient-level data. Claims that appropriate access to such data is incompatible with intellectual property constraints and patient privacy must be strongly resisted.

Jon Jureidini

Critical and Ethical Mental Health research group, Robinson Research Institute, University of Adelaide, North Adelaide, 5006, SA, Australia
jon.jureidini@adelaide.edu.au

In the last three years, University of Adelaide has been paid by Baum, Hedlund, Aristei and Goldman, Los Angeles for JJ to provide expert analysis and opinion about documents obtained from Forest Pharmaceuticals in relation to paediatric antidepressant randomised controlled trials. I thank Mickey Nardo and Matthew Dunbar for their helpful comments.

- 1 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; published online June 8. [http://dx.doi.org/10.1016/S0140-6736\(16\)30385-3](http://dx.doi.org/10.1016/S0140-6736(16)30385-3).
- 2 Jureidini J, Amsterdam J, McHenry LB. The Citalopram CIT-MD-18 Pediatric Depression Trial: deconstruction of medical ghostwriting, data mischaracterisation and academic malfeasance. *Int J Risk Saf Med* 2016; **28**: 33-43.
- 3 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.
- 4 Jureidini J. Let children cry. *Med J Aust* 2014; **201**: 612-13.
- 5 Batstra L, Frances A. Diagnostic inflation: causes and a suggested cure. *J Nerv Ment Dis* 2012; **200**: 474-79.