

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Hetrick, SE;McKenzie, JE;Cox, GR;Simmons, MB;Merry, SN

Title:

Newer generation antidepressants for depressive disorders in children and adolescents

Date:

2012-11-14

Citation:

Hetrick, S. E., McKenzie, J. E., Cox, G. R., Simmons, M. B. & Merry, S. N. (2012). Newer generation antidepressants for depressive disorders in children and adolescents. Cochrane Database of Systematic Reviews, 2021 (9), <https://doi.org/10.1002/14651858.CD004851.pub3>.

Persistent Link:

<https://hdl.handle.net/11343/59246>

Newer generation antidepressants for depressive disorders in children and adolescents (Review)

Hetrick SE, McKenzie JE, Cox GR, Simmons MB, Merry SN



**THE COCHRANE
COLLABORATION®**

This is a reprint of a Cochrane review, prepared and maintained by The Cochrane Collaboration and published in *The Cochrane Library* 2012, Issue 11

<http://www.thecochranelibrary.com>



Newer generation antidepressants for depressive disorders in children and adolescents (Review)
Copyright © 2012 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

TABLE OF CONTENTS

HEADER	1
ABSTRACT	1
PLAIN LANGUAGE SUMMARY	2
SUMMARY OF FINDINGS FOR THE MAIN COMPARISON	3
BACKGROUND	6
OBJECTIVES	7
METHODS	7
RESULTS	12
Figure 1.	14
Figure 2.	21
Figure 3.	22
Figure 4.	24
Figure 5.	25
Figure 6.	26
Figure 7.	27
Figure 8.	29
DISCUSSION	36
AUTHORS' CONCLUSIONS	40
ACKNOWLEDGEMENTS	41
REFERENCES	42
CHARACTERISTICS OF STUDIES	52
DATA AND ANALYSES	114
ADDITIONAL TABLES	116
WHAT'S NEW	127
HISTORY	128
CONTRIBUTIONS OF AUTHORS	128
DECLARATIONS OF INTEREST	129
SOURCES OF SUPPORT	129
DIFFERENCES BETWEEN PROTOCOL AND REVIEW	130
NOTES	130
INDEX TERMS	130

Newer generation antidepressants for depressive disorders in children and adolescents

Sarah E Hetrick^{1,2}, Joanne E McKenzie³, Georgina R Cox¹, Magenta B Simmons^{1,2}, Sally N Merry⁴

¹Orygen Youth Health Research Centre, Centre for Youth Mental Health, University of Melbourne, Melbourne, Australia. ²headspace Centre of Excellence, Centre for Youth Mental Health, University of Melbourne, Melbourne, Australia. ³School of Public Health & Preventive Medicine, Monash University, Melbourne, Australia. ⁴Department of Psychological Medicine, University of Auckland, Auckland, New Zealand

Contact address: Sarah E Hetrick, shetrick@unimelb.edu.au.

Editorial group: Cochrane Depression, Anxiety and Neurosis Group.

Publication status and date: New search for studies and content updated (no change to conclusions), published in Issue 11, 2012.

Review content assessed as up-to-date: 28 October 2011.

Citation: Hetrick SE, McKenzie JE, Cox GR, Simmons MB, Merry SN. Newer generation antidepressants for depressive disorders in children and adolescents. *Cochrane Database of Systematic Reviews* 2012, Issue 11. Art. No.: CD004851. DOI: 10.1002/14651858.CD004851.pub3.

Copyright © 2012 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

ABSTRACT

Background

Depressive disorders are common in young people and are associated with significant negative impacts. Newer generation antidepressants, particularly selective serotonin reuptake inhibitors (SSRIs), are often used, however evidence of their effectiveness in children and adolescents is not clear. Furthermore, there have been warnings against their use in this population due to concerns about increased risk of suicidal ideation and behaviour.

Objectives

To determine the efficacy and adverse outcomes, including definitive suicidal behaviour and suicidal ideation, of newer generation antidepressants compared with placebo in the treatment of depressive disorders in children and adolescents.

Search methods

For this update of the review, we searched the Cochrane Depression, Anxiety and Neurosis Review Group's Specialised Register (CCDANCTR) to October 2011. The CCDANCTR includes relevant randomised controlled trials from the following bibliographic databases: CENTRAL (the Cochrane Central Register of Controlled Trials) (all years), EMBASE (1974 -), MEDLINE (1950 -) and PsycINFO (1967 -). We searched clinical trial registries and pharmaceutical company websites. We checked reference lists of included trials and other reviews, and sent letters to key researchers and the pharmaceutical companies of included trials from January to August 2011.

Selection criteria

Published and unpublished randomised controlled trials (RCTs), cross-over trials and cluster trials comparing a newer generation antidepressant with a placebo in children and adolescents aged 6 to 18 years old and diagnosed with a depressive disorder were eligible for inclusion. In this update, we amended the selection criteria to include newer generation antidepressants rather than SSRIs only.

Data collection and analysis

Two or three review authors selected the trials, assessed their quality, and extracted trial and outcome data. We used a random-effects meta-analysis. We used risk ratio (RR) to summarise dichotomous outcomes and mean difference (MD) to summarise continuous measures.

Main results

Nineteen trials of a range of newer antidepressants compared with placebo, containing 3335 participants, were included. The trials excluded young people at high risk of suicide and many co-morbid conditions and the participants are likely to be less unwell than those seen in clinical practice. We judged none of these trials to be at low risk of bias, with limited information about many aspects of risk of bias, high drop out rates and issues regarding measurement instruments and the clinical usefulness of outcomes, which were often variously defined across trials. Overall, there was evidence that those treated with an antidepressant had lower depression severity scores and higher rates of response/remission than those on placebo. However, the size of these effects was small with a reduction in depression symptoms of 3.51 on a scale from 17 to 113 (14 trials; N = 2490; MD -3.51; 95% confidence interval (CI) -4.55 to -2.47). Remission rates increased from 380 per 1000 to 448 per 1000 for those treated with an antidepressant. There was evidence of an increased risk (58%) of suicide-related outcome for those on antidepressants compared with a placebo (17 trials; N = 3229; RR 1.58; 95% CI 1.02 to 2.45). This equates to an increased risk in a group with a median baseline risk from 25 in 1000 to 40 in 1000. Where rates of adverse events were reported, this was higher for those prescribed an antidepressant. There was no evidence that the magnitude of intervention effects (compared with placebo) were modified by individual drug class.

Authors' conclusions

Caution is required in interpreting the results given the methodological limitations of the included trials in terms of internal and external validity. Further, the size and clinical meaningfulness of statistically significant results are uncertain. However, given the risks of untreated depression in terms of completed suicide and impacts on functioning, if a decision to use medication is agreed, then fluoxetine might be the medication of first choice given guideline recommendations. Clinicians need to keep in mind that there is evidence of an increased risk of suicide-related outcomes in those treated with antidepressant medications.

PLAIN LANGUAGE SUMMARY

Newer antidepressants for depression in children and adolescents

Depression is common in young people and can contribute to a variety of negative outcomes, such as poor academic functioning, difficulties in peer and family relationships, increases in substance use, and both attempted and completed suicide. This review contained 19 trials (with a total of 3353 participants) testing the effectiveness of newer generation antidepressants (these are antidepressants developed and used since tricyclic antidepressants were developed). These include the well-known selective serotonin reuptake inhibitors that have an impact primarily on the brain chemical called serotonin, as well as several other newer classes of antidepressants now being used, which aim to target noradrenaline and dopamine as well as serotonin and include selective norepinephrine reuptake inhibitors (SNRIs), norepinephrine reuptake inhibitors (NRIs), norepinephrine dopamine reuptake inhibitors (NDRIs), norepinephrine dopamine disinhibitors (NDDIs) and tetracyclic antidepressants (TeCAs) for the treatment of depression in children and adolescents. Based on 14 of the trials (2490 participants in total), there was evidence that those treated with an antidepressant had lower depression severity scores than those on placebo, however, the size of this difference was small. Based on 17 trials (3229 participants in total), there was evidence of an increased risk (64%) of suicide-related outcomes for those on antidepressants compared with those given placebo. Where rates of adverse events were reported, this was higher for those prescribed an antidepressant. There was no evidence that one particular type of newer generation antidepressant had a larger effect than the others when compared to placebo.

It is unclear how meaningful the results regarding the effectiveness of these medications are in terms of a young person's day-to-day functioning. Children and adolescents with other conditions (such as anxiety, substance use disorder or a conduct disorder) as well as depression, and those at risk of suicide, were often excluded from trials. However, these young people are more representative of the population who present to clinical services, therefore it is not possible to predict how they would respond to antidepressants. There was often insufficient information to judge the quality of the trials accurately. With these limitations, it is difficult to answer questions about the effectiveness and safety of antidepressants for treating depression in children and adolescents. Clinicians need to provide accurate information to children and adolescents, and their families, about the uncertainties regarding the benefits and risks of newer generation antidepressant medication as a treatment option for depression. If a decision to use medication is agreed then fluoxetine might be

the medication of first choice given guideline recommendations and, if used, the risk of suicide should be assessed and monitored particularly closely.

SUMMARY OF FINDINGS FOR THE MAIN COMPARISON [\[Explanation\]](#)

Second generation antidepressant versus placebo (by medication) for depressive disorders in children and adolescents					
Patient or population: patients with depressive disorders in children and adolescents Settings: outpatient clinics Intervention: second generation antidepressant versus placebo (by medication)					
Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of participants (studies)	Quality of the evidence (GRADE)
	Assumed risk	Corresponding risk			
	Control	Second generation antidepressant versus placebo (by medication)			
Depressive disorder according to DSM or ICD criteria Clinician diagnostic interview	Study population		Not estimable	0 (0)	See comment
	See comment	See comment			
	Moderate				
Suicide completion Completed suicide rates	Study population		Not estimable	0 (0)	See comment
	See comment	See comment			
	Moderate				
Remission or response (as defined in trial)	Low response		RR 1.18 (1.08 to 1.28)	2924 (16 studies)	⊕⊕⊕⊕ moderate ¹

	170 per 1000	201 per 1000 (184 to 218)		
	Median response			
	380 per 1000	448 per 1000 (410 to 486)		
	High response			
	670 per 1000	791 per 1000 (724 to 858)		
Suicide-related outcome	Low risk²		RR 1.58 (1.02 to 2.45)	 low^{3,4}
	5 per 1000	8 per 1000 (5 to 12)		
	Median risk²			
	26 per 1000	40 per 1000 (26 to 62)		
	High risk²			
	111 per 1000	175 per 1000 (113 to 272)		
Depressive symptom severity CDRS-R. Scale from: 17 to 113.		The mean depressive symptom severity in the intervention groups was 3.51 lower (4.55 to 2.47 lower)	2490 (14 studies)	 low^{1,5}
*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). CI: confidence interval; RR: risk ratio				

GRADE Working Group grades of evidence

High quality: Further research is very unlikely to change our confidence in the estimate of effect.

Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

Very low quality: We are very uncertain about the estimate.

¹ Randomised trials were judged to be at a low risk of bias if they met the following criteria: a low risk of bias for sequence generation, allocation concealment, blinding of participants and assessors, incomplete outcome data, and selective outcome reporting. No trial met this criteria.

² Low, median and high risk rates were derived from the lowest, the median and the highest placebo risk of suicide-related outcomes rate.

³ Randomised trials were judged to be at a low risk of bias if they met the following criteria: a low risk of bias for sequence generation, allocation concealment, blinding of participants and assessors, incomplete outcome data, and selective outcome reporting. No trial met this criteria.

⁴ The total number of events is less than 300.

⁵ Randomised trials were judged to be at a low risk of bias if they met the following criteria: a low risk of bias for sequence generation, allocation concealment, blinding of participants and assessors, incomplete outcome data, and selective outcome reporting. No trial met this criteria.

BACKGROUND

Description of the condition

Depressive disorders are common in young people (Kaufman 2001; Pine 1998) and are diagnosed according to the Diagnostic and Statistical Manual of Mental Disorders (DSM) or International Classification of Diseases (ICD) criteria, usually by a clinician conducting a structured or semi-structured diagnostic interview. The core features of depressive disorder include persistent low mood, loss of enjoyment in once pleasurable activities, and a negative view of ones self, the future and others. These features are generally similar in children, adolescents and adults (Carlson 1988; Marttunen 1998). The DSM includes criteria changes for children and adolescents such as the presence of irritability as an alternative to a depressed mood for this age group (Angold 1988; Essau 1999). Generally, anhedonia and psychomotor retardation are less common in the younger age group where clinical phenotypes can be indistinct with presentations including an admixture of anxiety, depressive and somatic symptoms (Axelson 2001; Rivas-Vasquez 2004). Low self esteem, concentration and thinking problems, and behaviour difficulties are more frequent (Carlson 1988). In adolescents the presentation of a depressive disorder may include substance abuse, antisocial behaviour, social withdrawal and academic failure (Masi 1998) with suicide attempts and ideation also common in adolescents (Marttunen 1998).

Prevalence estimates are higher in studies with a six to 12 rather than three-month time frame, with meta-analysis of prevalence estimates across different time points giving estimates of 2.8% (standard error (SE) 0.5%) for children, and 5.7% (SE 0.3%) for adolescents (Costello 2006). Life-time estimates range between 15% and 20% (Birmaher 1996). Incidence rates (rate of new diagnoses during a particular time period) range from 3.3% to 7.8% over a year for major depressive disorder (MDD) (Garrison 1997; Lewinsohn 1998). The length of a depressive episode is often greater than 12 months: about 50% of children and adolescents remain clinically depressed at 12 months, and 20% to 40% at 24 months (Birmaher 1996; Harrington 2001; Kovacs 1984). Between 30% and 70% of cases have recurrences within five years, and many of these develop episodes into adult life (Fombonne 2001a; Fombonne 2001b; Lewinsohn 1998; Richmond 2005; Weissman 1999). In the longer term, those children and adolescents who develop a recurrent or chronic disorder extending into adulthood are likely to suffer considerable disability and impairment, high rates of co-morbid disorders with poor academic functioning, difficulties in peer and family relationships, increases in substance use, and attempted and completed suicide (Brent 2002; Ebmeier 2006; Fleming 1993; Harrington 1990; Lewinsohn 1998; NHMRC 1997; Rao 1995). The publication of the Global Burden of Disease (Murray 1996) included adolescents aged 15 to 18 years of age and showed depression as the fourth most important disease in the estimation of disease burden.

Description of the intervention

Overall there are relatively few long-term studies on depressive disorders in children and adolescents. Despite what is known about its prevalence and impacts, there is relatively little evidence for effective treatments or their impact on prognosis (NICE 2005). While a range of psychotherapies are effective (NICE 2005), adolescents' response to psychotherapies may be weaker than adults (Cuijpers 2005; Gloaguen 1998; Weisz 2006) and overall more research is required (NICE 2005).

Tricyclic antidepressants (TCAs) have not been shown to be an effective pharmacological treatment for depressive disorder in young people (Hazell 2002; Weller 2000). Newer generation antidepressants, especially selective serotonin reuptake inhibitors (SSRIs), have been increasingly used in recent decades (Vitiello 2006), with initial studies showing they were well tolerated (Cooper 1988). However, concerns about the increased risk of suicide and suicide attempt on SSRIs were first raised in 2003 (Healy 2003). Meta-analyses examining the risks of suicide-related behaviour and suicidal ideation combined (Hammad 2006) or separately (Dubicka 2006) have shown a consistent and modest increased risk of suicide for those taking SSRIs compared with placebo. Although debate about this issue is ongoing (e.g. Goodyer 2010; Hetrick 2010), the evidence for such risks has lead to action by regulatory bodies. The Committee on Safety of Medicines, Medicines and Healthcare Products Regulatory Agency (MHRA) in the UK (CSM 2004), the European Medicines Agency (EMA 2005) and the Food and Drug Administration (FDA 2004) have cautioned practitioners in the use of SSRIs in children and adolescents, including an FDA 'black box' warning label issued 14 September 2004 (FDA 2004). In addition to SSRIs, several other classes of antidepressants are now being used, including selective norepinephrine reuptake inhibitors (SNRIs), norepinephrine reuptake inhibitors (NRIs), norepinephrine dopamine reuptake inhibitors (NDRIs), norepinephrine dopamine disinhibitors (NDDIs) and tetracyclic antidepressants (TeCAs). SSRIs are sometimes referred to as 'second generation' antidepressants, and these newer additional classes are sometimes referred to as 'third generation' antidepressants. Rather than being a homogenous group based on mechanisms of action, however, third generation antidepressants are classed together because they are modified versions of first and second generation antidepressants (Olver 2001). Given the use of these classes of antidepressants, this updated version of the review includes these and throughout the review, second and third generation antidepressants are referred to as 'newer generation' antidepressants.

How the intervention might work

Antidepressant medication has evolved over the past 50 years. Until recently, a widely held belief was that dysfunction in serotonergic neurons and their targets may underlie depressive symptomatology (van Praag 1987). The dopaminergic system has also been

implicated, given its association with reward and appetitive motivation, whereby depression is characterised by a diminished ability to experience pleasure. Serotonin does have modulatory effects on dopamine, either increasing or decreasing its activity depending on the concomitant action of other neurotransmitters and the receptor subtype it is acting on.

SSRIs cause an initial inhibition of the reuptake of serotonin (5-hydroxytryptamine, or 5-HT) (Lenox 2008), but while this effect occurs within hours of taking the medication, the clinical effects are not evident for some weeks. This suggests it may be the downstream effects of reuptake inhibition on gene expression and receptor regulation that are chiefly responsible for the clinical effects of SSRIs (Castren 2005). SSRIs also affect other neurotransmitters, including noradrenaline and dopamine (Healy 1997).

A third generation of antidepressants target the noradrenaline and dopamine systems to a greater degree than the SSRIs, though most also have an effect on the serotonergic system (Healy 1997). There is significant interplay between the monoaminergic systems, rendering efforts to reduce antidepressant action to their effects on a single system simplistic.

This review refers to 'newer generation antidepressants' which encompasses both SSRIs and third generation antidepressant medication. It is unlikely that even within class these antidepressants will have similar effectiveness; similar to the multi-treatments meta-analysis by Cipriani 2009 we have pooled all individual compounds (rather than pooling within class) in order to look at whether the effect of newer antidepressants is modifiable by individual compound type (but we cannot assess comparative effectiveness given this is not a multi-treatments meta-analysis).

Why it is important to do this review

The American Academy of Child and Adolescent Psychiatrists, in response to the initial black box warnings, expressed concern about a stand that will deprive young people of effective treatment for a condition that carries with it considerable morbidity and mortality (Brent 2004; Findling 2004). Similarly, reviews examining the risks and benefits of antidepressants, most commonly SSRIs, consistently highlight the potentially serious consequences of untreated depression in children and adolescents. While modest treatment benefits have been shown for fluoxetine in children and adolescents (Hetrick 2007; Whittington 2004), there is contention about the effectiveness of other newer generation antidepressants. This review updates the previous version of the review to include not only SSRIs but newer antidepressants, to ensure information about all antidepressants is available and attempts to investigate issues of effectiveness and risk for children and adolescents with depression treated with these newer generation antidepressants.

OBJECTIVES

This update of the review aims to investigate not only the effects of SSRIs (as in the previous version) but also the effects of newer generation antidepressants compared with placebo, and factors which may modify the effects, in children and adolescents with a diagnosed depressive disorder. Specific objectives are to:

1. estimate the pooled effect of newer generation antidepressants on depression, function and adverse outcomes and whether this effect is modified by drug and age (children versus adolescents);
2. estimate the effect of each newer generation antidepressant, compared with placebo, on depression, function and adverse outcomes.

Results from this review do not address questions of comparative effectiveness of each of the newer generation drugs. However, in a future review we plan to undertake a multiple-treatments meta-analysis comparing the effectiveness of each individual drug which meets the scope of the current review.

METHODS

Criteria for considering studies for this review

Types of studies

Three types of published (including internet publication) and unpublished trials were eligible for inclusion in the review: parallel-group individually randomised controlled trials (RCTs), cross-over trials (where cross-over trials with less than one-week washout are treated as parallel-group trials including only the first period) and cluster trials. In the original review only SSRIs were included; however, including newer antidepressants was considered important to ensure information about all antidepressants was available. Information on adverse effects from other types of studies were not included in the review. We applied no language restrictions.

Types of participants

Children and adolescents aged 6 to 18 years old, both in and outpatients, who were diagnosed by a clinician and met DSM (APA 2000) or ICD (WHO 2004) criteria for a primary diagnosis of depressive disorder in its acute phase were eligible for inclusion. In the original version of the review, we intended to include trials of sub-syndromal depressive disorders. However, in this update we restricted the inclusion criteria to depressive disorder given that guidelines for the treatment of youth depression do not recommend medication is used in this group (McDermott 2011; NICE 2005).

Trials where both adults and children/adolescents were treated were eligible for inclusion if data on the children/adolescents could be extracted separately or obtained from trial authors.

Trials that did not exclude participants with co-morbid conditions secondary to a depressive disorder were eligible for inclusion and, if included in the future, a separate analysis is planned for those with only depressive disorder and those with a depressive disorder and a co-morbid condition(s).

Trials of children and adolescents with an intellectual quotient (IQ) of less than 70, organic brain injury or serious medical condition (that might result in risk to the participant or compliance issues) were not eligible for inclusion.

Types of interventions

In this updated version of the review, trials were eligible for inclusion if they compared the effectiveness of newer generation antidepressants with a placebo. It was considered important to include all newer antidepressants to ensure information about all antidepressants was available. These antidepressants were those consistent with the medications included in the equivalent Cochrane Depression, Anxiety and Neurosis (CCDAN) Group Meta-Analysis of New Generation Antidepressants (MANGA) reviews for adult depressive disorders (Churchill 2010; Cipriani 2005; Cipriani 2009a; Cipriani 2009b; Cipriani 2009c; Cipriani 2010; Guaiana 2010; Imperadore 2009; Nakagawa 2009; Nosè 2009; Omori 2010; Watanabe 2011). We have grouped them according to class as follows.

- Selective serotonin reuptake inhibitors (SNRIs): fluoxetine, fluvoxamine, sertraline, paroxetine, escitalopram, citalopram
- Selective norepinephrine reuptake inhibitors (SNRIs): venlafaxine, duloxetine, desvenlafaxine, milnacipran
- Norepinephrine reuptake inhibitors (NRIs): reboxetine
- Norepinephrine dopamine reuptake inhibitors (NDRIs): bupropion
- Norepinephrine dopamine disinhibitors (NDDIs): agomelatine
- Tetracyclic antidepressants (TeCAs): mirtazapine

We placed no restrictions on the dose or pattern of administering these antidepressants for the purposes of inclusion in the review. Trials where newer generation antidepressants were used in combination with another pharmacological intervention or psychological intervention exclusively were not eligible for inclusion. Trials with multiple comparison arms were eligible for inclusion, with only data from relevant treatment arms to be extracted.

Types of outcome measures

Primary outcomes

1. Depressive disorder according to DSM or ICD criteria and established by a clinician conducting a structured or semi-structured

diagnostic interview such as the Schedule for Affective Disorders and Schizophrenia for School Aged Children, Present Episode Version (K-SADS-P) (Chambers 1985). This was chosen as the most robust approach to establishing the resolution of a depressive episode.

2. Suicide completion established via recording of adverse outcome within the trial period or by medical record or direct inquiry with appropriate contact person at follow-up.

Secondary outcomes

1. Efficacy outcomes

1.1 Depression symptom severity (clinician-rated) using the Children's Depression Rating Scale (CDRS-R). This outcome was chosen on the basis of a hierarchy of rating scales based on psychometric properties and appropriateness for use with children and adolescents and for consistency of use across trials (the most commonly used tool) (see Appendix 1). The CDRS-R was adapted for children and adolescents from the Hamilton Depression Rating Scale (HAM-D), a tool validated and commonly used in adult populations (Brooks 2001). Both the CDRS-R and HAM-D have good reliability and validity (Brooks 2001). The Montgomery-Asberg Depression Rating Scale (MADRS) was also based on the HAM-D but designed to better assess sensitivity to change. However, it was not designed specifically for children and adolescents (Brooks 2001)

1.2. Remission or response as defined by trialists. 'Remission' and 'response' are commonly defined by dichotomising a continuous measure of clinician-rated depression symptoms. The labelling of remission and response varied across trials, with the labelling being different even although the cut-point was the same. For consistency across trials, we chose the most commonly reported cut-point, which was generally referred to as 'remission' (CDRS-R $\leq$ 28). When 'remission' was not reported, we used 'response' if available (we have used last observation carried forward (LOCF) data in the first instance (see [Dealing with missing data](#)) so that if remission was only available from observed case (OC) data but response data were available from LOCF data, we have used response data). The outcome represents some measure of improvement, but there are difficulties in interpretation because of the inconsistencies in scales and cut-points (Hetrick 2010).

We have chosen to include both continuous and dichotomised measures of clinician-rated depression symptoms (outcomes 1.1 and 1.2/1.3 respectively), since there are advantages and disadvantages to each. Responder analyses (based on the dichotomised continuous outcomes) are well known to be problematic (Kieser 2004), with arbitrariness in the choice of cut-point, loss of power resulting from the dichotomisation (Altman 2006) and difficulties in interpretation (as outlined above). However, synthesising continuous outcomes is not without its difficulties. The scales used to measure depression symptoms vary across trials; there is inconsistency in the analytical methods employed (e.g. analyses of change scores, regression models), which can preclude the use of the stan-

standardised mean difference; and there are also interpretational difficulties.

1.4 Depression symptom severity - self rated (on standardised, validated, reliable depression rating scales). The Beck Depression Inventory (BDI)/Childrens Depression Inventory (CDI) were the most commonly used across trials and ranked the highest in the hierarchy (see Appendix 1), therefore we meta-analysed this outcome. Results based on other scales are reported in [Table 1](#) and [Table 2](#).

1.5 Functioning (on standardised, validated, reliable global functioning rating scales). The Children's Global Assessment Scale (CGAS) was the most commonly used and, therefore, meta-analysis was based on this outcome. Results based on other scales are reported in [Table 3](#).

2. Suicide-related outcomes - where possible data based on the definitions used in the FDA review using the Columbia Classification system ([Hammad 2004](#)) have been chosen, again to maximise consistency across trials. In addition, we have collected data on suicidal ideation as a continuous outcome where a standardised, validated and reliable rating scale has been used.

3. Overall adverse outcomes (number with any adverse outcome reported).

4. Completion of trial protocol (the percentage of participants completing a trial).

Search methods for identification of studies

CCDAN's Specialised Register (CCDANCTR) searches for the original review

The Cochrane Depression, Anxiety and Neurosis Group (CCDAN) maintain two clinical trials registers at their editorial base in Bristol, UK, a references register and a studies-based register. The CCDANCTR-References Register contains over 30,000 reports of randomised controlled trials in depression, anxiety and neurosis. Approximately 65% of these references have been tagged to individual, coded trials. The coded trials are held in the CCDANCTR-Studies Register and records are linked between the two registers through the use of unique Study ID tags. Coding of trials is based on the EU-Psi coding manual. Please contact the CCDAN Trials Search Co-ordinator for further details.

Reports of trials for inclusion in the Group's registers are collated from routine (weekly), generic searches of MEDLINE (1950 -), EMBASE (1974 -) and PsycINFO (1967 -); quarterly searches of the Cochrane Central Register of Controlled Trials (CENTRAL) and review-specific searches of additional databases. Reports of trials are also sourced from international trials registers c/o the World Health Organization's trials portal ([ICTRP](#)), [ClinicalTrials.gov](#), drug companies, the handsearching of key journals, conference proceedings and other (non Cochrane) systematic reviews and

meta-analyses. Details of [CCDAN's generic search strategies](#) can be found on the Group's website.

Electronic searches

Searches for the original review

The original searches of MEDLINE, EMBASE and PsycInfo were undertaken to October 2005 and of CENTRAL (the Cochrane Central Register of Controlled Trials) to Issue 2, 2004 (Appendix 2).

CCDAN's Specialized Register, the CCDANCTR, was also searched at this time by the Trials Search Co-ordinator using the following terms:

Diagnosis = (Depress* or Dysthymi*) AND Intervention = ("Selective Serotonin Reuptake Inhibitors" or Alaproclate or Citalopram or Escitalopram or Femoxetine or Fluoxetine or Fluvoxamine or Paroxetine or Sertraline) AND Age Group = (Child or Adolescent) Other databases searched were the National Research Register (now archived), [ClinicalTrials.gov](#) and [Controlled-Trials.com](#). Additionally, the trial databases of pharmaceutical companies were searched.

Updated searches

Update searches were conducted on the CCDANCTR (to 28 October 2011) using additional terms for newer generation antidepressants:

CCDANCTR-Studies Register

Diagnosis = (depress* or dysthymi*) AND Intervention = ("Selective Serotonin Reuptake Inhibitors" or Agomelatine or Alaproclate or Bupropion or Citalopram or Desvenlafaxine or Duloxetine or Escitalopram or Fluoxetine or Fluvoxamine or Milnacipran or Mirtazapine or Paroxetine or Reboxetine or Sertraline or Venlafaxine) AND Age Group = (child* or adolescent* or "not stated" or unclear)

CCDANCTR-References Register

The CCDANCTR-References register was searched using a more sensitive set of terms to identify additional untagged/uncoded references:

Title/Abstract/Keywords = (depress* or dysthymi*) AND Free-Text=(Agomelatine or Alaproclate or Bupropion or Citalopram or Desvenlafaxine or Duloxetine or Escitalopram or Fluoxetine or Fluvoxamine or Milnacipran or Mirtazapine or Paroxetine or Reboxetine or Sertraline or Venlafaxine or (serotonin and (uptake or reuptake or re-uptake)) or SSRI*) AND Free-Text=(adolesc* or child* or boys or girls or juvenil* or minors or paediatric* or pediatric* or pubescen* or school* or students or teen* or young or youth*)

An additional search of international trials registries was conducted via the World Health Organization's trial portal (ICTRP), which at the time covered the following data providers:

- Australian New Zealand Clinical Trials Registry
- ClinicalTrials.gov
- ISRCTN (ControlledTrials.com)
- Chinese Clinical Trial Registry
- Clinical Trials Registry - India
- German Clinical Trials Register
- Iranian Registry of Clinical Trials
- Japan Primary Registries Network
- Pan African Clinical Trial Registry
- Sri Lanka Clinical Trials Registry
- Netherlands National Trial Register

Searching other resources

Reference lists

We screened the reference lists of included articles and other reviews retrieved in the search.

Handsearches

We searched pharmaceutical company websites, including Eli Lilly and Company, Forest Laboratories, Merck Pharmaceuticals, Lundbeck Pharmaceuticals, GlaxoSmithKline, Bristol-Myers Squibb and Pfizer Pharmaceuticals (the company Wyeth that was searched in original review has been subsumed by Pfizer).

Conference abstracts

We searched conference abstracts for the American Academy of Child and Adolescent Psychiatry (2003 to 2005) for the original review.

Personal communication

We consulted the authors of the included trials and other individuals with expertise in this field to find out if they knew of any potentially relevant published or unpublished RCTs/controlled clinical trials (CCTs).

Data collection and analysis

Selection of studies

Two review authors (MS and GC) performed the selection of trials for inclusion in the updated review independently after employing the search strategy described above. Where a title or abstract appeared to describe a trial eligible for inclusion, we obtained the full article to assess whether it met the inclusion criteria. We have

reported the reasons for exclusion of trials in the [Characteristics of excluded studies](#) tables.

Data extraction and management

Two review authors (SH and GC) independently extracted information on each trial, including 'Risk of bias' criteria and details of participants, interventions, comparisons, outcomes and potential modifying factors (age, individual drug type). Discrepancies were resolved by a third review author (MS) ([Characteristics of included studies](#)). These data form the basis for discussing the internal and external validity of results.

Two review authors (SH and GC) independently extracted post intervention outcome data for the primary and secondary outcomes and discrepancies were resolved by a third review author (JM). When estimates of treatment effect or standard errors were not directly reported, we calculated these, where possible, through algebraic manipulation of available statistics (e.g. means, confidence interval limits, exact P values).

For the first version of the review we decided post hoc to extract suicide-related outcomes from the Medicines and Healthcare Products Regulatory Agency (MHRA) rather than from the individual trial reports retrieved in the search for the current review. The MHRA has produced a web-based report (www.mhra.gov.uk/) that summarises the results of the majority of the trials included in the original review. We used two additional reports, one on suicide-related outcomes ([Hammad 2004](#)) and one on trial characteristics ([Dubitsky 2004](#)) giving details of outcomes for 25 SSRI trials for a range of disorders in children and adolescents. For this version of the review, we have again used the data from the FDA report of suicide-related outcomes where it is available and where it is not, extracted data as similar in definition to that used in this report from individual trials.

Assessment of risk of bias in included studies

In the original review we assessed the risk of bias in the included trials using the quality of trials ratings devised by Moncrieff and colleagues ([Moncrieff 2001](#)). For this update we used the Cochrane Collaboration's new 'Risk of bias' tool ([Higgins 2009](#)). Specifically, for each trial, we assessed the domains sequence generation, allocation concealment, blinding of participants and assessors (separately for efficacy outcomes and adverse outcomes), incomplete outcome data and selective reporting. We sought to assess trial protocols where published/available in the first instance and also sought clarification from trial authors in the case of suspicion of reporting bias) and any other sources of possible bias that which might affect the outcome of the trial results. We judged each domain as being at a low, high or unclear risk of bias. We also extracted relevant text which underpinned our judgement and this is presented in the 'Risk of bias' tables in [Characteristics of included studies](#).

Two review authors (SH and GC) performed all assessments of the quality of trials independently, with discrepancies resolved by JM.

Measures of treatment effect

For dichotomous outcomes, we measured treatment effects using risk ratios (RR) (e.g. response rates and adverse effects). In the original version of this review, a post hoc decision was made not to calculate number needed to treat to benefit (NNTB) and number needed to treat to harm (NNTH) primarily because these effect measures have undesirable statistical properties and can be difficult to interpret (Julious 2005; Wisloff 2011). Instead, in our [Summary of findings for the main comparison](#), we present estimates of risk difference, in addition to risk ratios, to aid interpretation, for a range of control group rates (lowest, highest and median rate derived from the placebo groups).

For each drug, we estimated the risk ratio of experiencing adverse events (where a count of any adverse event was reported). However, given inconsistencies in data collection of these events between trials, the risk ratios may not be comparable between drugs; therefore, information about adverse events is also presented in [Table 4](#).

For continuous outcomes (such as clinician-rated depression symptom severity), we have used the mean difference (MD). In the majority of trials, multiple linear regression models had been fitted, and 'adjusted' estimates of treatment effects from these models were reported (often as least square means or least square mean differences). These models adjusted for varying factors such as age, sex, investigator site and baseline of the outcome (details available from the review authors). We were not able to use the standardised mean difference (SMD) where the same outcome was measured across trials, but using a different scale because there was inconsistency in the analytical methods employed (e.g. analyses of final values, change scores and regression models).

We presented data that could not be meta-analysed, such as information on co-morbidity, and treatment effects for trials that have used different scales than our chosen scale, in tables.

Unit of analysis issues

There were no trials with more than one intervention arm compared with the placebo group. In future updates, if this occurs, we will compare data from each arm with the placebo group. We will divide the sample size in the placebo arm by the number of intervention arms included in the meta-analysis to preclude multiple counting of trial participants. While this approach offers some solution, it does not completely account for the correlation arising from using the same set of participants in the placebo comparison group (Chapter 16, [Higgins 2009](#)).

No cross-over trials were included, but if they are located in future updates, if the appropriate data for a paired t-test analysis is not available and cannot be obtained from trial authors, we will impute

missing statistics (e.g. missing standard deviation, correlation) using data available from other trials included within the meta-analysis, or trials outside the meta-analysis (Chapter 16, [Higgins 2009](#); [Elbourne 2002](#)). We will use sensitivity analyses to assess the robustness of the pooled treatment effect to assumptions made regarding missing statistics.

If cluster-RCTs are included in future updates, which have not appropriately adjusted for the correlation between participants within clusters, we will contact trial authors to obtain an estimate of the intra-cluster correlation (ICC), or imputed using estimates from the other included trials or from similar external trials. We will inflate the trial standard errors (Chapter 16 [Higgins 2009](#)).

Dealing with missing data

We sought additional data from the principal authors and pharmaceutical companies of trials (the latter approached by the CC-DAN group on our behalf) (and the National Institutes for Mental Health (NIMH) in the case [TADS 2004](#)) that met the eligibility criteria where the data were missing, or were in a form unsuitable for meta-analysis. We also searched the pharmaceutical company websites for additional data on included trials.

Most trials used the last observation carried forward (LOCF) method of data imputation for the majority of outcomes, that is, the last observed value for a participant lost to follow-up is assigned as the follow-up value. We chose to pool LOCF data (rather than mix LOCF and OC data).

The observed case remission/response data were sought since we planned to investigate the effect of missing data using the "informative missingness odds ratios" (IMOR) framework ([Higgins 2008](#)). However, only a few trials reported observed case data for this outcome (paroxetine trials, sertraline trials, one trial of citalopram and one small trial of fluoxetine) and we were unsuccessful in obtaining observed case data for the other trials. We have undertaken sensitivity analysis using OC data where available. Estimates of treatment effect based on either LOCF or OC data can result in serious bias ([Sterne 2009](#)).

In some trials, least squares means and their standard errors were reported from regression models by treatment group, but no contrast between groups was reported. For these trials, we estimated the variance of the treatment effect by summing the square of the standard errors in each treatment group. There may be some inaccuracy in this approach when there is imbalance in the covariates being adjusted for.

Assessment of heterogeneity

We assessed heterogeneity of intervention effects by visually inspecting the overlap of confidence intervals on the forest plots, tested for heterogeneity using the Chi² test, and quantified heterogeneity using the I² statistic ([Higgins 2003](#)). Categories suggested in [Higgins 2009](#) (Chapter 9) are used to help interpret the degree of heterogeneity (0% to 40% might not be important; 30% to

60% may represent moderate heterogeneity; 50% to 90% may represent substantial heterogeneity; 75% to 100% considerable heterogeneity). We calculated 95% confidence intervals for values of I^2 using the non central χ^2 approximation implemented in the Stata module *heterogi* (Orsini 2006).

Assessment of reporting biases

We investigated the potential for small-study effects using funnel plots and contour-enhanced funnel plots. Contour-enhanced funnel plots aid in determining if funnel plot asymmetry is due to publication bias or other factors (Peters 2008; Sterne 2011). The outcomes 'remission/response' and suicide-related behaviour were chosen since they were available for the majority of included trials. We have used the statistical test proposed by Harbord et al (Harbord 2006) (as implemented in the Stata module *metabias* (Harbord 2009)) to test for small-study effects.

Data synthesis

We pooled estimates of treatment effect using inverse variance weighting, using a random-effects model. DerSimonian and Laird's method of moments estimator was used to estimate between-trial variance (DerSimonian 1986). We made the decision to use a random-effects model since there was expected clinical diversity in the antidepressant medications, given their differing actions on various monoaminergic systems. Where trials did not report data suitable for meta-analysis, treatment estimates or raw data (as appropriate for each outcome) for each individual study are reported in additional tables.

We judged randomised trials to be at a low risk of bias if they met the following criteria: a low risk of bias for sequence generation, allocation concealment, blinding of participants and assessors, incomplete outcome data, and selective outcome reporting. We summarised the risk of bias across the trials in the text of the review. As part of the sensitivity analyses, we present meta-analyses of trials that are judged to be at a low risk of bias for the outcomes depression symptom severity (CDRS-R), remission/response and suicide-related outcome.

Subgroup analysis and investigation of heterogeneity

We undertook a subgroup analysis based on the individual compounds to examine if the effect of newer generation antidepressants was modified by individual drug. Clinically, these drugs are used individually, with clinicians often trying to make a decision about which one is best to use for a particular individual.

There is evidence that children and adolescents may respond differently to pharmacological intervention, e.g. oral tricyclic antidepressants versus placebo significantly reduce symptoms in adolescents but not in children (Hazell 2002). For this reason we conducted subgroup analyses by age, where children and adolescents were defined as those aged approximately 6 to 12 and 13 to 18

years respectively (Characteristics of included studies). When estimates of treatment effect were not presented for children and adolescents separately, we created another subgroup which contained both children and adolescents.

In the first version of the review, we had planned a priori to undertake subgroup analyses based on depressive disorder (major versus dysthymic disorder, depressive disorder not otherwise specified and 'double depression'), sex and co-morbidity. However, due to limited data, we did not carry out analyses on these subgroups, nor was it possible in this update.

Sensitivity analysis

In the original review, we did not carry out pre-planned sensitivity analyses based on pharmaceutical funding (financial support of trials by pharmaceutical companies has been shown to influence trial results (Schott 2010)) and inclusion criteria (clinical diagnosis versus depression rating scales) since the majority of trials were pharmaceutically funded, and no trials used rating scales as inclusion criteria. We did not undertake a pre-planned sensitivity analysis based on attrition rates since there were reasonably high attrition rates in all included RCTs (19% to 38%, Table 5).

For this version of the review, we undertook sensitivity analyses based on the assessment of risk of bias in the trials (see Data synthesis), whereby trials deemed to be at a high risk of bias were excluded from the analysis.

We undertook sensitivity analysis based on methods of imputation for missing data used by trial authors such that we substituted LOCF data for OC data for the outcome 'remission/response'.

Finally, given the inclusion in this review of third generation antidepressants that have a different proposed mechanism of action from SSRIs, we have undertaken sensitivity analysis on the basis of drug class, excluding third generation antidepressants from the overall pooled analysis.

Timeline

Ongoing updates of this review will usually be submitted for editorial review within two years of publication of the review.

RESULTS

Description of studies

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

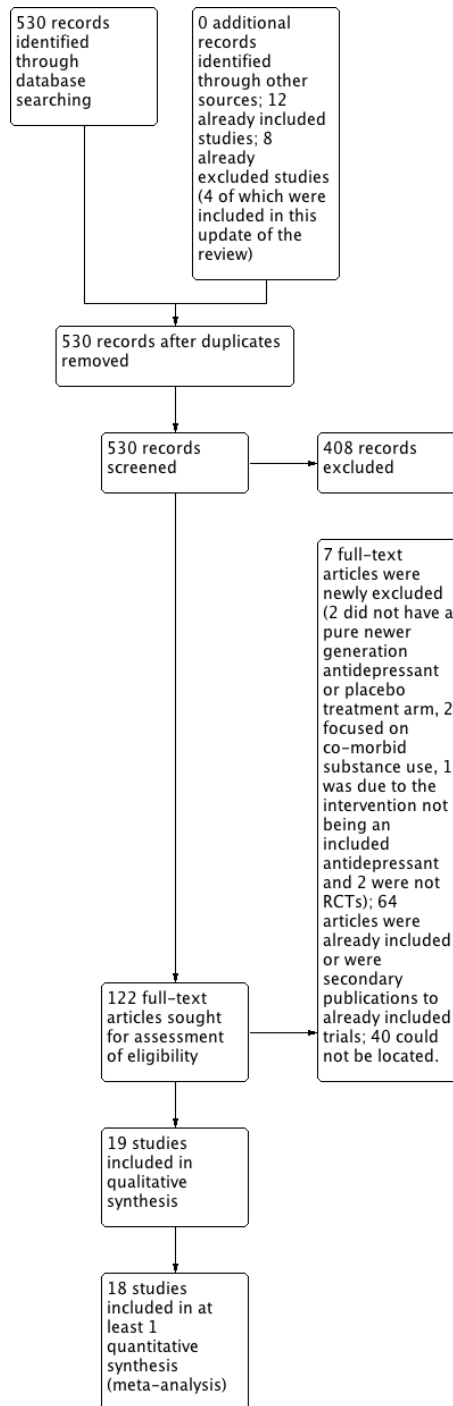
Results of the search

Twelve trials were included in the original review ([Hetrick 2007](#)), with data extracted from 10 trials and pooled in one or more meta-analyses. As our inclusion criteria have been expanded to include newer classes of antidepressants, four trials excluded from the original review have been included in this version of the review ([Emslie 2007 Trial 1](#); [Emslie 2007 Trial 2](#); [Mirtazapine Trial 1](#); [Mirtazapine Trial 2](#)).

In this update, 530 trials were retrieved in the search, of which 408 were excluded on the basis of title and abstract. We attempted to retrieve 122 full-text articles for full inspection. Forty publi-

cations could not be located (the majority of these were conference proceedings that had not been subsequently published). Of the 82 that were obtained, 64 were either already included trials or were secondary publications from already included trials (35 of these were for [TADS 2004](#) and are not listed under the main reference for this trial). Seven new trials were included, four of which are ongoing trials ([Duloxetine NCT00849693](#); [Duloxetine NCT00849901](#); [Glod 2004](#); [Solvay NCT00353028](#)), leaving a total of 19 trials included in this updated version of the review ([Figure 1](#)). In addition, seven studies were newly excluded (see [Excluded studies](#)).

Figure 1. Study flow diagram.



The three trials newly included from the updated search include: A new trial on fluoxetine (Almeida-Montes 2005). This trial was published in Spanish and data extracted by a colleague who speaks Spanish, with additional data sought via this colleague.

A new trial of escitalopram (Emslie 2009) and a new trial of paroxetine (Paroxetine Trial 1).

In the original review, two trials were initially unpublished, and in this update they are referred to by their subsequent publication trial identification: Berard 2006 (was Milin 2004); Emslie 2006 (was paroxetine study 3).

For the included trials, we had retrieved additional reports during preparation of the original review, including the web-based report of the Medicines and Healthcare Product Regulatory Agency (MHRA), summarising the majority of clinical trials on SSRIs for major depressive disorder in children and adolescents at the time. When we wrote to trial authors for additional data during preparation of the original review, in many cases trial authors did not have access to any additional data. We accessed the trial reports published online by SmithKline Beecham on paroxetine (Berard 2006; Emslie 2006; Keller 2001) (<http://www.gsk.com/media/paroxetine.htm>). For one paroxetine trial (Paroxetine Trial 1) we only had access to a brief trial report from this website. We had also accessed trial reports published online by Forest Laboratories for escitalopram and citalopram and for this update located the report for a newly located trial of escitalopram. Eli Lilly provided additional data for a trial on fluoxetine (Emslie 2002) during preparation of the original review. For this update, the trial authors for this trial and for Emslie 1997 were able to provide additional data in response to our requests. For this update we accessed trial reports of venlafaxine from CCDAN who had accessed them from Pfizer.

Two published trial reports include the results of two trials Wagner Trial 1&2 (2003) and Emslie 2007. Data for the individual trials for Wagner Trial 1&2 (2003) were only available from the MHRA report. Emslie 2007 provides some data separately for each of the two trials included in the publication.

There was no published report for the mirtazapine trials; data were available from the MHRA report and from two reports to regulatory agencies.

The trial by Simeon was discontinued early due to slow enrolment, with some information about the trial from the written report and some from the MHRA report (Dubitsky 2004).

Included studies

Design

The trials were all individual patient parallel-group randomised trials. All were multicentre with the exception of Almeida-Montes 2005, Emslie 1997 and Simeon 1990. Most trials had two arms,

comparing a newer generation antidepressant with placebo. The Treatment for Adolescents with Depression trial (TADS 2004) includes four comparison groups: an SSRI group, a cognitive behavioural group (CBT), a combined SSRI and CBT group and a placebo group. We only extracted data from the SSRI and placebo groups. The Keller 2001 study included three comparison groups: a TCA group, an SSRI group and a placebo group. We only extracted data from the SSRI and placebo groups.

Sample sizes

The number of participants randomised to the relevant arms in these trials ranged from 23 to 367 (median 188).

Setting

The included trials were undertaken in many countries (Denmark, Estonia, Germany, Norway, Sweden, Switzerland, Argentina, Belgium, Holland, Italy, Japan, Mexico, South Africa, Spain, United Arab Emirates, UK, India, Costa Rica, USA, Canada).

Most of the trials gave little information on their recruitment strategies. Of those that did, Emslie 2002, TADS 2004 and Mirtazapine Trial 1 & 2 used media advertising. Emslie 1997 stated that media recruitment was not used (Characteristics of included studies).

Only two trials (Simeon 1990; Von Knorring 2006) stated that inpatients were included, although the MHRA report (Dubitsky 2004) states that Simeon 1990 only included outpatients. We have attempted to contact this author but have not had a reply.

Participants

There were six trials in adolescents only (Berard 2006; Emslie 2009; Keller 2001; Simeon 1990 TADS 2004; Von Knorring 2006) with an age range of 12 or 13 to 17 or 18, and 13 trials of children and adolescents (Almeida-Montes 2005; Emslie 1997; Emslie 2002; Emslie 2006; Emslie 2007 Trial 1; Emslie 2007 Trial 2; Mirtazapine Trial 1 & 2; Paroxetine Trial 1; Wagner 2006; Wagner 2004; Wagner Trial 1&2 (2003)) with a lower age limit of between six to eight years. The mean age ranged from 14.4 to 16.0 years and 11.5 to 13.3 years in the adolescent, and child and adolescent, trials respectively (Characteristics of included studies). There were similar proportions of females and males in Emslie 2002; similar numbers but slightly more males than females in five trials (Emslie 1997; Emslie 2006; Emslie 2007 Trial 1; Emslie 2007 Trial 2; Mirtazapine Trial 1); and similar numbers but slightly more females than males in two trials (Mirtazapine Trial 2; Wagner 2006). The proportion of females was greater than males but balanced across groups in one trial (Emslie 2009). There were nearly

twice as many females in two trials (Keller 2001; Berard 2006). In Wagner 2004 and Wagner Trial 1 (two trials reported together), there was imbalance in the proportion of females between groups, with a greater proportion of females in the treatment group; in Paroxetine Trial 1 the proportion of females was greater in the placebo group. The TADS 2004 and Simeon 1990 trials did not provide information on sex by treatment arm but overall there were somewhat more females than males. Two trials provided no information on sex (Almeida-Montes 2005; Von Knorring 2006) (Characteristics of included studies).

All trials were of major depressive disorder with three trials basing diagnoses on DSM-III or DSM-III-R criteria (Emslie 1997; Keller 2001; Simeon 1990) and the remainder on DSM-IV criteria. The majority used a semi-structured clinical interview (K-SADS and Almeida-Montes 2005 used the MINI-KID); one trial gave no information (Paroxetine Trial 1). Von Knorring 2006, in contrast to all the other trials, used only a five-minute clinical interview with parents. In addition to a diagnostic interview, the majority of trials (except Emslie 2006) used a cut-off score on a measure of depression symptom severity to establish eligibility. Emslie 1997; Emslie 2002; Emslie 2007 Trial 1; Emslie 2007 Trial 2; Mirtazapine Trial 1 & 2; Wagner 2006 and Wagner 2004 used a cut-off of greater than 40 on the CDRS-R; while for Emslie 2009; Paroxetine Trial 1; TADS 2004 and Wagner Trial 1 the cut-off was 45. In Von Knorring 2006 the Children's Depression Inventory (CDI) was used with cut-offs greater than 21 and 16 for girls and boys respectively. A score greater than 12 or 20 on the HAM-D scale was used in Keller 2001 and Simeon 1990 respectively, and a score greater than 16 on the MADRS scale was used in Berard 2006. Almeida-Montes 2005 used a score greater than 13 on the DSDR. Some trials also used a measure of functioning to confirm diagnosis (Berard 2006; Keller 2001; Mirtazapine Trial 1 & 2; Von Knorring 2006; Wagner Trial 1).

Some trials included a screening process that was undertaken over a period of one to three weeks (Emslie 1997; Emslie 2002; Emslie 2006; Emslie 2007 Trial 1; Emslie 2007 Trial 2; Emslie 2009; Keller 2001; Paroxetine Trial 1; TADS 2004; Wagner Trial 1). A report by the MHRA described the process as more extensive for three of these trials (Emslie 1997; Emslie 2002; Keller 2001) but did not describe what this meant. Further investigation revealed a screening process that included up to three independent diagnostic interviews, taking place over a period of up to three weeks. In 11 trials all participants were treated with placebo for a lead-in period and those whose depressive disorder improved during this time were excluded (Almeida-Montes 2005; Berard 2006; Emslie 1997; Emslie 2002; Emslie 2007 Trial 1; Emslie 2007 Trial 2; Emslie 2009; Paroxetine Trial 1; Simeon 1990; Wagner 2004; Wagner 2006).

Authors of all trial reports, except two (Almeida-Montes 2005; Simeon 1990), describe depression symptom severity at baseline for the treatment and placebo groups. Mean severity scores at baseline from the individual trials range from 47.6 to 65.5 on the

CDRS-R (range 17 to 113) (equivalent to T-Scores that correspond to a depressive disorder being likely or very likely to be confirmed) and from 25.9 to 32.5 on the K-SADS nine-item depression score (range 9 to 56) (Keller 2001; Von Knorring 2006) and 25.9 on the MADRS (Berard 2006). For all trials, there was no clinically important imbalance between treatment groups in depression symptom severity at baseline. Clinical Global Impression (CGI) scores were reported in 11 trials and ranged from a mean of 3.9 to 4.8 (with a median of four being reported in Emslie 2006), which is in the moderately ill range. Of those trials that reported on the percentage of young people who were experiencing a first episode of depression (12 trials), the rates varied from 42.7% to 95% in the intervention group and 42.8% to 95% in the placebo group. The length of the current episode was reported variously in 10 trials and ranged from approximately 15 weeks to 108 weeks in the intervention group and 14 to 100 weeks in the placebo group, with the majority of trials reporting episode lengths of over 35 weeks.

Seven trials provided no detail about the co-morbid conditions of participants (Almeida-Montes 2005; Emslie 2007; Mirtazapine Trial 1 & 2; Paroxetine Trial 1; Simeon 1990; Von Knorring 2006). Wagner Trial 1&2 (2003) provided data for sertraline and placebo groups combined for both trials combined. In this trial the co-morbid conditions that were most common (> 5%) were anxiety, phobic disorder, adjustment reaction and ODD.

Five trials (Emslie 1997; Emslie 2006; Emslie 2009; Keller 2001; TADS 2004) provided data on the number of young people who had any co-morbid condition with rates very variable across trials (Emslie 2006 and Emslie 2009 having the lowest rates), ranging from 12.9% to 85.4% in the intervention arms and 16.6% to 77.1% in the placebo arms (see Table 6).

In eight trials the percentage of young people experiencing various different types of co-morbid conditions was provided and has been presented in Table 6. The percentages vary markedly between trials, with Emslie 1997; Emslie 2002; Keller 2001 and TADS 2004 having the highest rates of various co-morbid conditions. It is clear from research that co-morbidity may affect the clinical outcome (Birmaher 1996; Kovacs 1989); however it is difficult to examine this, given the non standard way in which co-morbidity is reported and because some co-morbid disorders form part of the exclusion criteria in some trials.

In all trials exclusion criteria included psychotic features or disorder, and all but two excluded substance abuse or dependence (Paroxetine Trial 1; Wagner Trial 1&2 (2003), who did not state that substance abuse or dependence was excluded). In all but seven trials (Almeida-Montes 2005; Emslie 2006; Paroxetine Trial 1; Simeon 1990; TADS 2004; Wagner Trial 1&2 (2003)) anorexia nervosa and bulimia nervosa were excluded; in all but nine trials (Almeida-Montes 2005; Emslie 1997; Emslie 2002; Emslie 2007 Trial 1; Emslie 2007 Trial 2; Mirtazapine Trial 1 & 2; Paroxetine Trial 1; Simeon 1990) pervasive developmental disorders were excluded. Externalising disorders (disruptive behaviour disorder,

Oppositional Defiant Disorder (ODD) and Conduct Disorder (CD)) and/or attention deficit-hyperactivity disorder (ADHD) were excluded in 10 trials (Almeida-Montes 2005; Berard 2006; Emslie 2009; Emslie 2007 Trial 1; Emslie 2007 Trial 2; TADS 2004; Von Knorring 2006; Wagner Trial 1&2 (2003); Wagner 2004; Wagner 2006).

Obsessive compulsive disorder was excluded from all but 10 trials (Almeida-Montes 2005; Emslie 1997; Emslie 2002; Emslie 2006; Paroxetine Trial 1; Simeon 1990; TADS 2004; Von Knorring 2006; Wagner 2004; Wagner 2006). Other anxiety disorders excluded were post-traumatic stress disorder (Berard 2006; Emslie 2009; Keller 2001; Wagner 2004; Wagner 2006) and panic disorder (Berard 2006; Emslie 2007 Trial 1; Emslie 2007 Trial 2; Wagner Trial 1&2 (2003)), social phobia (Berard 2006) and 'severe anxiety disorder' (Almeida-Montes 2005).

Bipolar was excluded from all but two trials (Berard 2006; Simeon 1990) and patients with a first-degree relative with a history of bipolar were excluded from seven trials (Emslie 1997; Emslie 2002; Emslie 2007 Trial 1; Emslie 2007 Trial 2; Emslie 2009; Mirtazapine Trial 1 & 2).

Participants who were considered at risk for suicide at baseline were specifically excluded in all but three trials (Emslie 1997; Paroxetine Trial 1 for which no statement was explicitly made; Von Knorring 2006). The method to define risk varied across trials; some gave no definition of 'serious suicidality risk' (Emslie 2002; Emslie 2006; Simeon 1990) or acute suicidality (Emslie 2007 Trial 1; Emslie 2007 Trial 2) or defined it as "in the opinion of the investigator" (Almeida-Montes 2005); previous attempt was exclusionary in nine trials (Emslie 2009; Keller 2001; Mirtazapine Trial 1 & 2; TADS 2004; Wagner Trial 1&2 (2003); Wagner 2004; Wagner 2006); one trial stated that attempt was not exclusionary, only current suicidal ideation with intent or plan (Berard 2006), although this was contradicted by the FDA report (Hammad 2004) that stated that a history of suicidal ideation was not an exclusion criteria in any of the trials included in its report. Along with suicide attempt, Keller 2001 also stated those with serious suicidal ideation with intent or a specific plan were excluded. Subsequent to the report by Hammad 2004, two trials have also included suicidal ideation as an exclusion criteria: TADS 2004 states that "clear intent or an active plan to commit suicide, or suicidal ideation with a disorganized family unable to guarantee adequate safety monitoring" and Emslie 2009 states that those who had "active suicidal ideation" was exclusionary. The FDA carried out a stratified analysis of those trials included at the time based on history of suicide attempt or ideation to investigate if risk of suicide attempt or ideation for those receiving SSRIs varied by stratum. They concluded that there was no evidence of this (Hammad 2004).

Interventions

In the SSRI class, there were four trials of paroxetine (Berard 2006; Emslie 2006; Keller 2001; Paroxetine Trial 1), five trials of fluoxe-

tine (Almeida-Montes 2005; Emslie 1997; Emslie 2002; Simeon 1990; TADS 2004), two trials of citalopram (Von Knorring 2006; Wagner 2004), two trials of escitalopram oxalate (the therapeutically active component of citalopram) (Emslie 2009; Wagner 2006), and two trials of sertraline (Wagner Trial 1&2 (2003)). In the SNRI class there were two trials of venlafaxine (both reported in Emslie 2007), and in the TeCA class there were two trials of mirtazapine (Mirtazapine Trial 1 & 2).

Ten trials explicitly excluded those who had previously not responded to antidepressant treatment (Emslie 2002; Emslie 2006; Emslie 2009; Mirtazapine Trial 1 & 2; TADS 2004; Wagner Trial 1&2 (2003); Wagner 2004; Wagner 2006).

The treatment period of the included trials was between 6 and 12 weeks.

Efficacy measures were collected throughout the treatment period and at completion of the trial. For four trials this was described as weekly (Almeida-Montes 2005; Emslie 1997; Emslie 2002; Keller 2001); for 11 trials this was near weekly (Berard 2006; Emslie 2006; Emslie 2007 Trial 1; Emslie 2007 Trial 2; Emslie 2009; Mirtazapine Trial 1 & 2; Paroxetine Trial 1; Von Knorring 2006; Wagner 2004; Wagner 2006). Wagner Trial 1&2 (2003) stated that there were frequent follow-up visits but no further details were given. TADS 2004 described assessments at baseline, 6, 12, 18, 24, 30 and 36 weeks. There was no detail given for Simeon 1990. Emslie 2009 specifically states that the high placebo response rate may be due to "extensive contact" (pg 728), which refers to this regular assessment.

Five trials (Emslie 1997; Emslie 2002; Keller 2001; TADS 2004; Wagner 2004) described a continuation phase. Emslie 1997 stated that after the eight weeks of acute treatment, treatment was not controlled and participants were followed up at 6 and 12 months. In a later report of Emslie 2002, two additional phases are described, one for non responders and one for relapse prevention, both of which were blinded. Keller 2001 stated that at the end of acute treatment (eight weeks), responders continued on blinded treatment (paroxetine, imipramine or placebo) for a further six months and non responders were tapered off medication and terminated from the trial. In TADS, after 12 weeks of acute treatment (after which the placebo group was unblinded and responders and partial responders given their choice of the three active intervention arms for 12 weeks), there was a six-week 'consolidation' phase for responders and partial responders, followed by an 18-week maintenance phase after which there was a 12-month naturalistic follow-up. Wagner 2004 stated that there was a 24-week open label extension trial.

With the exception of four trials (Almeida-Montes 2005; Emslie 1997; Emslie 2002; Emslie 2009), a flexible dosing scheme was used.

Outcomes

Clinician-rated depression symptom severity

Depression symptom severity was measured in a variety of ways, including:

1. The CDRS-R (Emslie 1997; Emslie 2002; Emslie 2006; Emslie 2007 Trial 1; Emslie 2007 Trial 2; Emslie 2009; Mirtazapine Trial 1 & 2; Paroxetine Trial 1; TADS 2004; Wagner Trial 1&2 (2003); Wagner 2004; Wagner 2006).

The Children's Depression Rating Scale (CDRS-R) (Poznanski 1984) was adapted for children and adolescents from the Hamilton Depression Rating Scale (HAM-D) and assesses 17 symptom areas. The first 14 items are rated on the basis of responses to interview questions by the child or an adult informant who knows the child well and are rated for the past two weeks and currently. The remaining three symptom areas (depressed facial affect, listless speech and hypoactivity) are rated by the clinician on the basis of the child's non verbal behaviour in the room. Each symptom is graded on a five or seven-point scale. For items 1 to 14 the highest ratings from the child, parent or other caretaker are taken as the *item* scores. The total score, or CDRS-R score, is the sum of all 17 item scores and has a range of 17 to 113. In samples studied in the development of the scale (Poznanski 1996) mean CDRS-R T scores (standardised scores) were 71, 58 and 53 for those with a depressive disorder (based on DSM-III criteria), other psychiatric disorder (outpatient) and no disorder respectively. This scale is used widely, has adequate internal reliability, good test-retest reliability, good to excellent inter-rater reliability and is sensitive to treatment effects (Myers 2002). We chose this measure for clinician-rated depressive symptom severity.

2. K-SADS (Von Knorring 2006; Keller 2001; Berard 2006).

The Schedule for Affective Disorders and Schizophrenia for School Aged Children, Present Episode Version (K-SADS-P) (Chambers 1985) depression module has nine items on an ordinal scale, four of which consist of two to three sub-items. Each of the items or sub-items is rated from zero to either four or six with higher numbers corresponding to greater severity. The score range is 9 to 56. The various items rated are for the last two weeks in order to enable diagnosis according to DSM-IV criteria. It has reasonable reliability but is not often used to assess treatment outcome (Brooks 2001). Data based on the KSADS are reported in an additional table but not included in meta-analysis.

3. The HAM-D was used in the Simeon trial. Almeida-Montes 2005 used both the HAM-D and the Depression Self Rating Scale (DSRS) (the later was translated into Spanish) .

The Hamilton Psychiatric Rating Scale for Depression (HAM-D) (Hamilton 1960) consists of 17 multiple choice questions each of which is numerically scored on a scale of zero to two or four. The score can range from 0 to 42 with a higher score indicating more severe depression. The various items refer to depressive disorder severity over the last week. Reliability is reported as excellent with trials suggesting sensitivity to treatment effects although further research is required (Myers 2002). There were no data for this outcome.

4. The MADRS was also used in Berard 2006 and Von Knorring 2006. The continuous data based on this measure was not meta-analysed. However, the dichotomised measure of this scale indicating response was meta-analysed (see below).

The Montgomery-Asberg Depression Rating Scale (MADRS) (Montgomery 1979) was also based on the HAM-D and is a clinician-rated scale that assesses depressive disorder symptoms in the last week or the last three days. This scale is less commonly used compared to other depression rating scales. It consists of 10 items covering apparent sadness, reported sadness, inner tension, reduced sleep, reduced appetite, concentration difficulties, lassitude, inability to feel, pessimistic thoughts and suicidal thoughts. Each of these items is scored between 0 and 6 based on severity. The possible range of scores is 0 to 60 with a higher score indicating more severe depression. This scale was specifically designed to assess the effect of treatment, but its psychometric properties have not been specifically examined in adolescents (Brooks 2001).

The CGI was also used in many trials, though not as a continuous measure of symptom severity outcome; rather it was often used to establish 'response'.

The Clinical Global Impression of Improvement (CGI-I) (Guy 1976) is a clinician-rated seven-point scale that assesses global improvement from baseline to the current state. The scale is:

1. very much improved
2. much improved
3. minimally improved
4. no change
5. minimally worse
6. much worse
7. very much worse

Response/remission

Remission was measured in a variety of ways both across and within trials; few trials used diagnostic interviews to establish remission. More often it was defined by the trial authors as a level of improvement in depression symptoms on clinician-rated scales. The scale and the cut-point used to define this level of improvement varied between trials. Response was reported in all trials, however, again the scale used and the cut-point was variously defined between trials and usually, but not always, was of a smaller magnitude compared to remission. Therefore, as stated in the methods, for this version of the review, we have used remission in the first instance, but response if remission was not available. We have used LOCF data for the primary analysis of this outcome, and in one case this meant using response data rather than remission data (Wagner Trial 1&2 (2003)). In one trial OC data based on response was all that was available and we have used this (Almeida-Montes 2005). A large number of trials reported remission based on CDRS-R ≤ 28 (Emslie 1997; Emslie 2002; Emslie 2006 (for total population); Emslie 2009; TADS 2004). Two trials used this cut-point definition but labelled it 'response' (Wagner 2004; Wagner 2006).

Data from [Von Knorring 2006](#) were based on what the trial authors called “remission”, defined as a MADRS score of 12 or less. Data from [Berard 2006](#) were based on what they called “response”, defined as $\geq 50\%$ reduction in baseline MADRS scores, and data from [Keller 2001](#) were based on what they called “response”, defined as HAM-D ≤ 8 . Similarly, data from [Almeida-Montes 2005](#) were based on what they called “response”, defined as $\geq 50\%$ reduction in DSRS and HAM-D score. Data from [Emslie 2007 Trial 1](#) and [Emslie 2007 Trial 2](#) were based on what they called “response” defined as $\geq 35\%$ reduction in baseline CDRS-R scores. The data from [Paroxetine Trial 1](#) were based on what they call “response” defined as a CGI score of 1 or 2 (much improved or very much improved). The paper publication of the sertraline trials ([Wagner Trial 1&2 \(2003\)](#)) did not include remission, however the MHRA reported these data using a definition of those who no longer met DSM-IV criteria for MDD; however because these MHRA data were only OC data, we have used response data defined as $\geq 40\%$ reduction in baseline CDRS-R scores. [Simeon 1990](#) and [Mirtazapine Trial 1 & 2](#) did not provide any data for this outcome.

Self rated depression symptom severity

Self rated depression symptom severity was measured in four trials. Two trials ([Berard 2006](#); [Von Knorring 2006](#)) used the Beck Depression Inventory (BDI) ([Beck 1961](#)); one trial ([Emslie 1997](#)) used the Children’s Depression Inventory (CDI) ([Kovacs 1992](#)) for child-aged participants and the BDI ([Beck 1961](#)) for adolescents and combined these into one score; and one trial ([TADS 2004](#)) used the Reynolds Adolescent Depression Scale (RADs) ([Reynolds 1986](#)).

The BDI ([Beck 1961](#)) is the first and most widely used self rated tool to assess depression severity, consisting of 21 items of which there are four item choices related to how the respondent has felt in the last week. The range of possible scores is from 0 to 63, with higher scores indicating worse depression severity. The tool was developed for adults ([Brooks 2001](#)). The CDI ([Kovacs 1992](#)) was based on the BDI but specifically developed for children aged seven years and older and has been frequently used in younger adolescents. It has 27 items, with the first 18 very similar to the BDI, and there are three response options for each item. The range of possible scores is 0 to 54, with higher scores indicating greater severity ([Brooks 2001](#)). The RADs ([Reynolds 1986](#)) was developed for adolescents aged 13 to 18 years. It has 30 items with four response options per item to gauge the current severity of symptoms of depression with the range of possible scores from 30 to 120 (higher scores indicating greater severity) ([Brooks 2001](#)).

Functioning

Measures of function included the Children’s Global Assessment Scale (CGAS) ([Almeida-Montes 2005](#); [Berard 2006](#); [Emslie 1997](#);

[Emslie 2009](#); [TADS 2004](#); [Wagner Trial 1](#); [Wagner Trial 2](#); [Wagner 2004](#); [Wagner 2006](#)), the Global Assessment of Functioning (GAF) ([Emslie 2002](#); [Emslie 2006](#); [Von Knorring 2006](#)) and the Autonomous Functioning Checklist ([Keller 2001](#)). One trial on paroxetine ([Paroxetine Trial 1](#)) and the trials of venlafaxine ([Emslie 2007 Trial 1](#); [Emslie 2007 Trial 2](#)) and mirtazapine ([Mirtazapine Trial 1 & 2](#)) did not report data on functioning.

The Children’s Global Assessment Scale (CGAS) ([Shaffer 1985](#)) measures a child’s current general functioning and is completed by the clinician. The range is from 1 to 100 with a score of 1 to 10 indicating a need for constant supervision and a 90 to 100 indicating superior functioning. The CGAS has adequate reliability and is sensitive to change ([Myers 2002](#)).

The Global Assessment of Functioning (GAF) ([APA 2000](#)) measure is a clinician-rated scale that assesses the patient’s current level of functioning. Scores range from 1 to 90 (90 indicates good functioning in all areas). There are few trials of its psychometric properties in child and adolescent populations; however, a recent review suggests it is likely to be reliable when used in research given the training and motivation of raters ([Schorre 2004](#)).

The Autonomous Functioning Checklist (AFC) ([Sigafos 1988](#)) is completed by the parent, and assesses the child’s autonomy in performing daily activities. It consists of 78 questions grouped into four categories; 22 questions on self and family care; 20 questions on management; 16 questions on recreational activities; and 20 questions on social and vocational activities. The first three categories are rated on an ordinal scale ranging from 0 (“does not do”) to 4 (“does every time there is an opportunity”). While the last categories consist of “yes” (coded 1) and “no” (coded 0) questions. A total score and a sub-score for each of the four categories are calculated, with higher values indicating a greater degree of autonomy. There has been little psychometric investigation of this measure.

Suicide-related outcomes

Suicide-related outcomes were classified and reported in various ways in each of the trials. As described in the methods section, a post hoc decision was made in the first version of this review to use the data provided in an FDA report ([Hammad 2004](#)) in order to overcome inconsistent reporting of these outcomes across trial reports.

The process of the FDA in establishing the rate of suicide-related outcomes for each trial was based on the following process. A group of 10 suicidology experts were assembled by Columbia University (led by Dr Kelly Posner). Suicide-related outcomes were defined after careful deliberation by this expert panel as including ‘definitive suicidal behaviour/ideation’ (pg 8, [Hammad 2004](#)) and where more than one event was recorded for an individual, the most severe event was used. The group of experts reviewed all of the suicide-related adverse events, all serious adverse events and all accidental injuries identified by the sponsors of SSRI trials. There

was some discrepancy between the sponsors' classifications and the expert panel classification (with 22 new events added, and 26 old events removed). Overall there were no completed suicides in any of the trials. The report highlighted the important point that none of these trials had adequate power for safety analysis.

For this version of the review we have again extracted suicide-related outcomes from trials of depressive disorders from the FDA report (Hammad 2004) where they were available. Where the FDA did not include data for trials included in our review, we included data from the trial report (for TADS 2004 data were extracted from the Emslie 2006 report where it is stated that rates are based on a reanalysis by Columbia Group using the Columbia-Classification Algorithm for Suicidal Assessment); Emslie 2009 states their data are based on an increase in suicidal ideation and behaviour on the Modified Columbia Suicide Severity Rating Scale (MC-SSRS), a clinician-rated instrument; Wagner 2006 states that "potential suicide-related events were identified" and describes these in the results as adverse events, which is not equivalent to the data based on the Columbia Classification; for the mirtazapine trials (Mirtazapine Trial 1 & 2), the MHRA report gave a description of events stating there was one case of suicidal ideation in the mirtazapine group (both trials combined) and one case of self mutilation in the placebo group (both trials combined). The data for Paroxetine Trial 1 were suicidal ideation reported as an adverse event. The trial by Almeida-Montes 2005 did not provide data for this outcome.

We have also included, in this version of the review, suicidal ideation as a continuous outcome, which was measured by the Suicidal Ideation Questionnaire (Reynolds 1987) in both trials in which this instrument was used (Emslie 2009; TADS 2004).

Ongoing studies

Four included trials have been categorised as ongoing (Duloxetine NCT00849693; Duloxetine NCT00849901; Glod 2004; Solvay NCT00353028) (see Characteristics of ongoing studies), although the trials' registration documentation suggests the duloxetine trials are completed, and the abstract by Glod 2004 was published some years ago. These are all individual patient parallel-group randomised trials. The two duloxetine and fluvoxamine (Solvay NCT00353028) trials are multicentre; it is unclear whether Glod 2004 is a single site or multicentre trial.

The duloxetine (SNRI) trials include two arms of duloxetine, one a flexible dose and one fixed, as well as a fluoxetine arm and a placebo arm all delivered for 10 weeks. These trials include children and adolescents treated in an outpatient setting with MDD according to DSM-IV (APA 2000) with additional inclusion criteria including a CDRS-R score of ≥ 45 and a CGI score of ≥ 4 . The trials exclude participants with psychotic features or disorder, substance use or dependence, eating disorders, pervasive developmental disorder, obsessive compulsive disorder, a history of bipolar disorder, and those with first degree relatives with bipolar

disorder. Those who had made a significant suicide attempt within one year of screening or were currently at risk of suicide in the opinion of the investigator were excluded. It is unclear how long the screening for inclusion takes or whether there is a placebo lead-in phase. There appears to be a continuation period for those on duloxetine with a 36-week follow-up. The CDRS-R is used as the measure of clinician-rated depression symptoms.

The Glod 2004 trial, published as a conference abstract, includes three arms: citalopram (SSRI), bupropion (NDRI) and placebo, delivered for eight weeks. It is unclear whether flexible or fixed dosing is used. The trials include adolescents treated in an outpatient setting with MDD according to DSM-IV (APA 2000) with no additional inclusion criteria stated. It is unclear what the screening procedure for trial entry was, or whether there was a placebo lead-in. In addition, it was unclear what psychiatric conditions were excluded or whether those with suicide risk were excluded. The conference abstract did provide some information on the first 18 young people randomised into the trial. Their mean age was 15.5; six were males and 12 females and their baseline level of depression symptom severity was 20.3 on the HAM-D. Few other details were provided.

The Solvay NCT00353028 trial undertaken in Japan includes two arms: fluvoxamine maleate (flexible dose) (SSRI) and placebo delivered for 38 weeks. Flexible dosing was used (25 mg to 150 mg). The trials include children and adolescents with MDD diagnosed using the Japanese Version of the Structured Interview Guide for the Hamilton Depression (JSIGH-D) (minimum total score of 18 on the JSIGH-D). An additional inclusion criteria was weight within the standard weight. Exclusion criteria were a predominant psychiatric diagnosis of schizophrenia, or previously been treated with fluvoxamine maleate.

Excluded studies

In the original review, there were eight excluded studies. These were excluded due to the intervention not being an SSRI; one trial was a head-to-head trial of antidepressants that did not include a placebo; one trial was a case-control trial and one trial included participants with bipolar disorder not depressive disorder.

Of these, the two trials on venlafaxine and the two trials on mirtazapine are now included in the update due to the inclusion criteria changing to include SSRIs as well as newer generation antidepressants.

In this version of the review, based on the updated search, seven new studies were excluded. The primary reasons for exclusion included the following: two did not have a pure newer generation antidepressant or placebo treatment arm; two focused on co-morbid substance use; one used an antidepressant that did not meet the inclusion criteria for antidepressants considered in this review; and two were not randomised trials.

Risk of bias in included studies

See [Figure 2](#) for the 'Risk of bias' graph that shows the proportion of studies with each of the judgements and [Figure 3](#) for the 'Risk of bias' summary showing all the judgements in a cross-tabulation by trial.

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

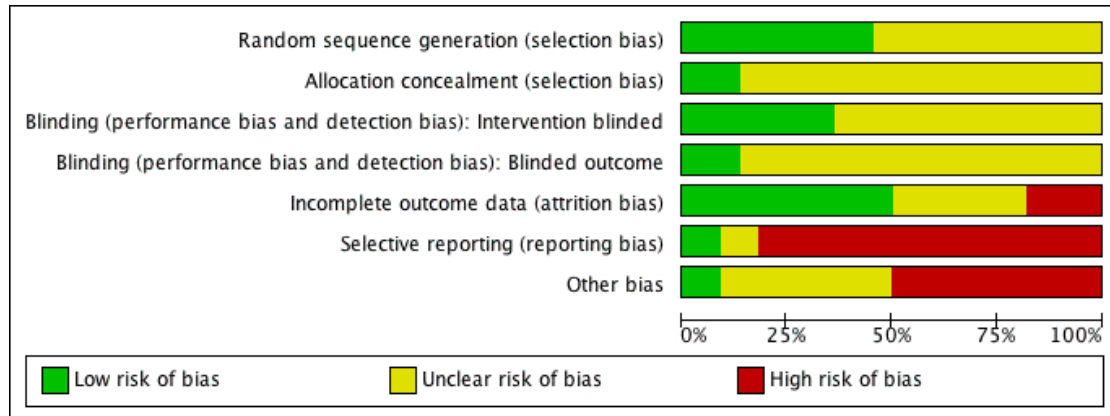


Figure 3. 'Risk of bias' summary: review authors' judgements about each risk of bias item for each included study.

	Random sequence generation (selection bias)	Allocation concealment (selection bias)	Blinding (performance bias and detection bias): Intervention blinded	Blinding (performance bias and detection bias): Blinded outcome	Incomplete outcome data (attrition bias)	Selective reporting (reporting bias)	Other bias
Almeida-Montes 2005	+	+	?	+	-	?	?
Berard 2006	+	+	+	?	?	+	?
Emslie 1997	+	?	+	+	+	-	?
Emslie 2002	+	?	+	?	?	-	-
Emslie 2006	+	?	?	?	+	+	?
Emslie 2007	?	?	?	?	+	-	-
Emslie 2007 Trial 1	?	?	?	?	+	-	-
Emslie 2007 Trial 2	?	?	?	?	+	-	-
Emslie 2009	?	?	?	?	+	-	+
Keller 2001	+	?	+	?	+	-	-
Mirtazapine Trial 1	?	?	?	?	?	-	-
Mirtazapine Trial 1 & 2	?	?	?	?	?	-	-
Mirtazapine Trial 2	?	?	?	?	?	-	?
Paroxetine Trial 1	?	?	?	?	?	-	?
Simeon 1990	?	?	?	?	-	-	-
TADS 2004	+	+	+	+	+	-	?
Von Knorring 2006	?	?	?	?	-	-	?
Wagner 2004	?	?	?	?	-	-	?
Wagner 2006	?	?	?	?	?	?	+
Wagner Trial 1	+	?	+	?	+	-	-
Wagner Trial 1&2 (2003)	+	?	+	?	+	-	-
Wagner Trial 2	+	?	+	?	+	-	-

Allocation

There were no full reports of allocation concealment in any of the included trials.

Blinding

All trials were described as being “double-blind” or of having the relevant treatment arms double-blind (TADS). In two trials (Emslie 1997; Emslie 2002) the description of blinding indicates that the antidepressant and placebo medications were identical. There is little description of the blinding in 10 trials, so that it is unclear what “double-blind” refers to (Emslie 2007 Trial 1; Emslie 2007 Trial 2; Emslie 2009; Mirtazapine Trial 1 & 2; Paroxetine Trial 1; Simeon 1990; Von Knorring 2006; Wagner 2004; Wagner 2006). Almeida-Montes 2005 and TADS 2004 state that there were independent evaluators who were also “blind”. Emslie 1997 mentions that the pharmacy staff were blind. There were no reports on the success of blinding in any of the trials, and the possibility of clinicians or patients guessing the nature of the intervention from side effects was not discussed. Given outcomes were based on ratings by participants and clinicians, this could be an important omission, although the updated CONSORT guidelines highlight that asking participants or healthcare providers what intervention they received as a test of blinding at the end of the trial is confounded because usually by this stage they know what intervention they received (Moher 2010).

Incomplete outcome data

One trial was discontinued early (Simeon 1990) and it is unclear whether this is also the case for Glod 2004. One trial of paroxetine was aiming to recruit 65 participants in each of the treatment and placebo arms, however, it appeared to cease recruitment with less than half this number recruited to each group. The attrition rate for the 18 trials varied between 11% and 82% in the control groups and 14% and 58% in the intervention groups (see Table 4). The disparity in attrition between treatment arms was of particular concern in the trials of fluoxetine (Table 4).

All authors stated that intention-to-treat analyses (ITT) had been undertaken. However, a full application of the intention-to-treat principle is only possible when complete outcome data are available for all randomised participants (Hollis 1999). Only two trials (Emslie 1997; TADS 2004) appear to include all patients randomised in their analyses (the Paroxetine Trial 1 appears to include all randomised patients only in their primary analysis). In the other trials, analyses are carried out on fewer patients than the number randomised. For the majority of trials, only those who received at least one dose of medication or placebo, or had at least one post baseline efficacy or safety evaluation were included in the analyses.

Selective reporting

There is some evidence of reporting bias in some of the trials, though this is difficult to assess in most trials, since it was not possible to obtain the trial protocol. The trial report by Emslie 2002 emphasises CDRS-R scores and remission rates rather than response rate, even although response rate was specified as the primary outcome in the methods section. Additionally, the cut-off used for remission rate differed from that stated in the methods section. Emslie 1997 reports outcomes at five weeks rather than at the completion of the trial. In a letter to the editor, Keller 2001 was criticised for changing the definition of response post data analysis to a cut-off that showed treatment effectiveness (Jureidini 2003). In response, Keller 2001 changed their claim of finding a significant effect to stating that the findings showed a strong signal for efficacy (Jureidini 2004; Keller 2003). In many trials response/remission is defined, measured and reported in many different ways within the trial, without it being clear what the primary outcome is, e.g. Emslie 2009 reports two different results for response using two different definitions. The report by Wagner Trial 1&2 (2003) and Emslie 2007 combines the results of two trials and in most cases reports the overall outcomes. Wagner 2006 emphasises post hoc subgroup analyses. The outcomes for TADS 2004 have been reported in multiple publications with the reporting of outcome results that are not consistent across papers. For trials where results are only reported in the MHRA report, there are few data reported (e.g. Mirtazapine Trial 1 & 2). For one trial of paroxetine there is no publication except a brief pharmaceutical company trial report. In many cases, trials appear not to have measured or reported the outcomes specified as outcomes for this review, or have reported data in a way that means they cannot be used in meta-analysis, so that there are data missing from the meta-analyses.

It appears that two of the included trials have been stopped early Simeon 1990 and Glod 2004. There were no data reported from Simeon on the 40 participants who were included. Glod 2004 reports data on depression symptom severity (but not by group) on the first 18 participants and we have been unable to find publication of the full trial, despite our efforts to contact the author and pharmaceutical company.

Reporting bias

The funnel plots for the outcomes remission (Figure 4) and suicide-related behaviour (Figure 5) were not suggestive of small-study effects. In addition, the contour-enhanced funnel plots (Figure 6; Figure 7) did not indicate that statistically significant results were more likely to be reported (publication bias). Harbord's test for small-study effects was non significant for both outcomes (remission P value = 0.364; suicide-related behaviour P = 0.275).

Figure 4. Funnel plot of comparison: I Newer generation antidepressant versus placebo (by medication), outcome: I.2 Remission or response (as defined in trial) LOCF only.

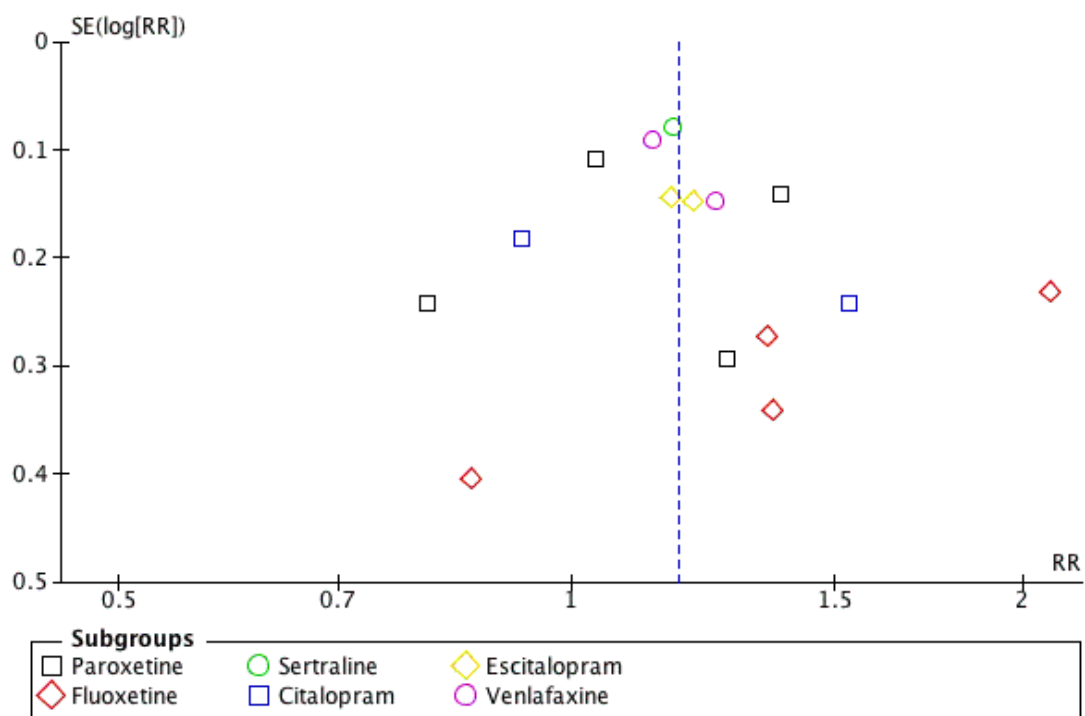


Figure 5. Funnel plot of comparison: I Newer generation antidepressant versus placebo (by medication), outcome: I.6 Suicide-related outcome.

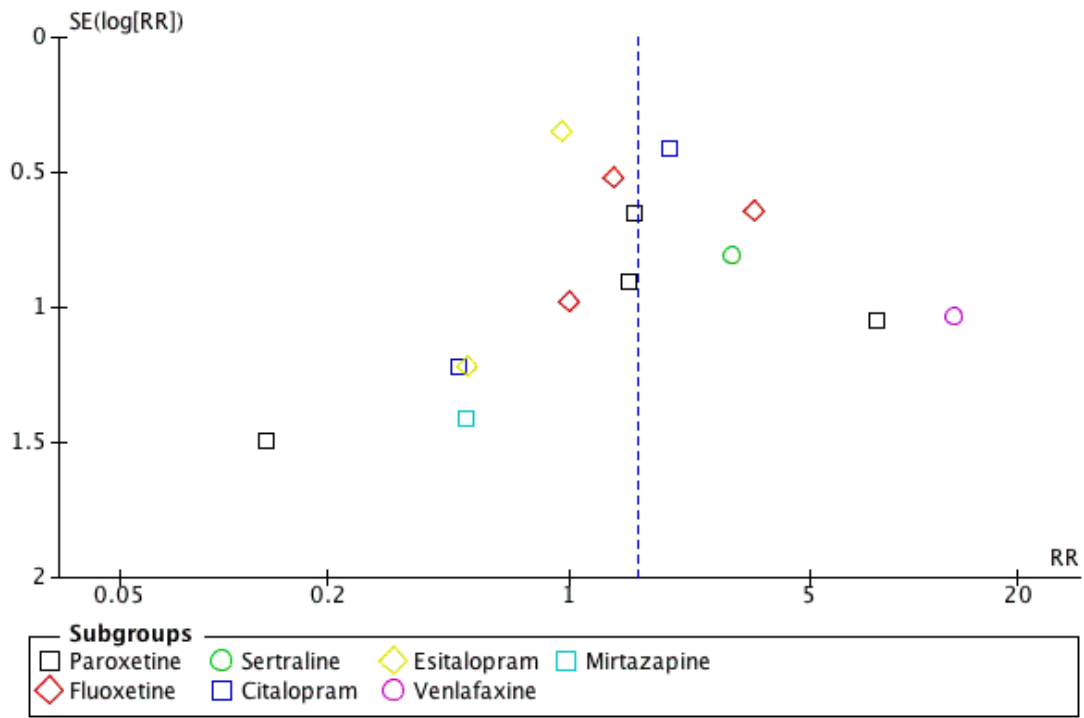


Figure 6. Contour-enhanced funnel plot for the outcome remission or response (mix of OC and LOCF).
The dashed vertical line represents the pooled random-effects estimate

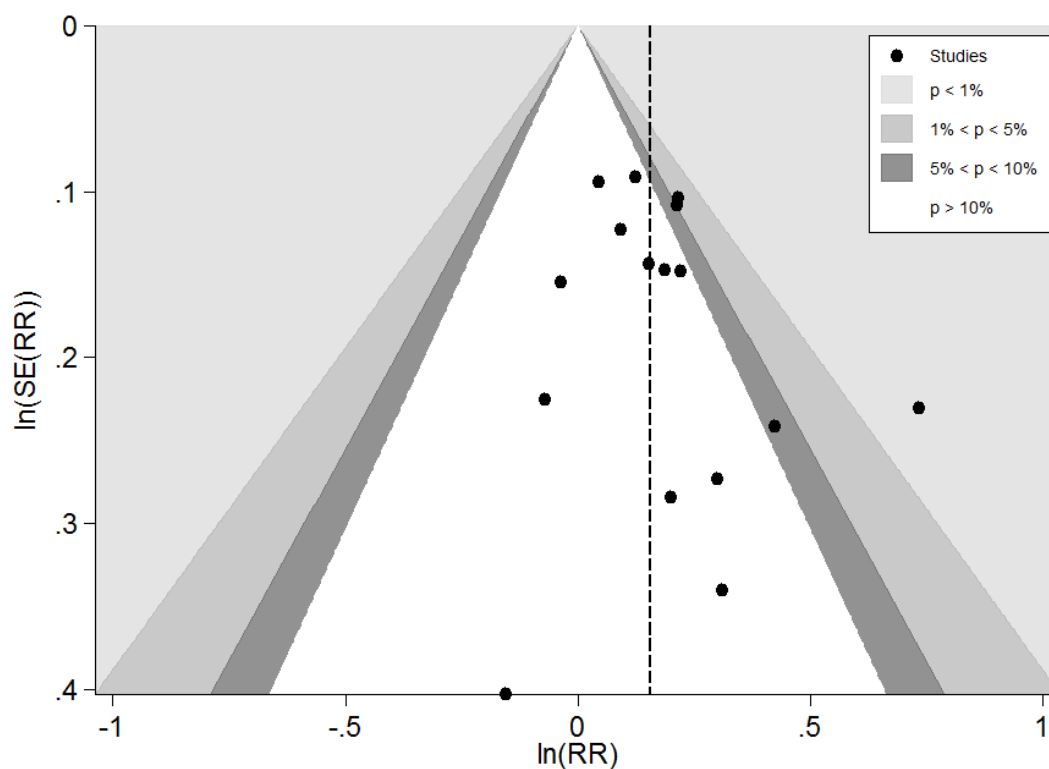
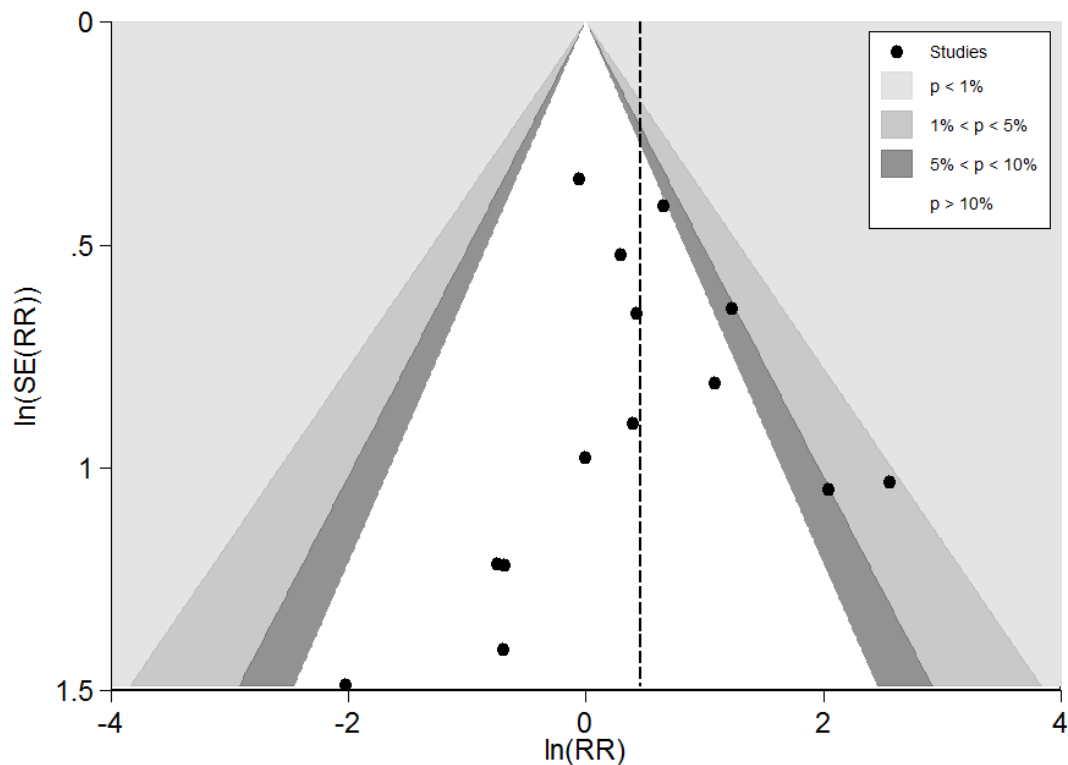


Figure 7. Contour-enhanced funnel plot for the suicide-related outcome. The dashed vertical line represents the pooled random-effects estimate



Other potential sources of bias

Funding

Most trials, with the exception of [Emslie 1997](#), were pharmaceutically funded. The [TADS 2004](#) trial was funded by an NIMH contract but had an “unrestricted educational grant from Eli Lilly” (pg 531 of the 2003 publication).

Compliance

Eleven trial reports did not describe any method for assessing compliance with intervention. Three trials (all of paroxetine) attempted to assess compliance by pill count ([Berard 2006](#); [Emslie 2006](#); [Keller 2001](#)) and six trials assessed plasma blood levels of the investigative trial medication ([Emslie 1997](#); [Mirtazapine Trial 1](#); [Mirtazapine Trial 2](#); [Paroxetine Trial 1](#); [Simeon 1990](#); [Von Knorring 2006](#)).

Additional therapy

Some trials gave details about additional support or psychotherapy provided to participants in the medication and placebo arms of trials. Psychotherapy was not permitted in [Wagner 2004](#); [Wagner 2006](#) and the mirtazapine trials ([Mirtazapine Trial 1 & 2](#)), although in the mirtazapine trials ‘supportive care’ was permitted, with no detail about how many received this. Non directive supportive therapy was permitted in [Berard 2006](#) but again no details are provided about how many young people received this. Supportive case management (including CBT and interpersonal therapy interventions) was provided to all participants in [Keller 2001](#). Therapy was permitted in the sertraline trials [Wagner Trial 1&2 \(2003\)](#) and it is unclear how many received this; [Von Knorring 2006](#) reports that psychotherapy was permitted and three-quarters of participants received it. In [TADS 2004](#) each participant received six 20 to 30-minute medication visits spread across 12 weeks of treatment (pg 809) during which their pharmacotherapist monitored their clinical status and medication effects, and offered general encouragement about the effectiveness of pharmacother-

apy for MDD. For the remainder of the trials there is no detail given about the provision of support or therapy (Almeida-Montes 2005; Emslie 1997; Emslie 2002; Emslie 2006; Emslie 2007 Trial 1; Emslie 2007 Trial 2; Emslie 2009; Simeon 1990).

Effects of interventions

See: [Summary of findings for the main comparison](#) Second generation antidepressant versus placebo (by medication) for depressive disorders in children and adolescents

A. Comparison 1: Newer generation antidepressants versus placebo

Primary outcomes

1. Depressive disorder according to DSM or ICD criteria and established by a clinician conducting a structured or semi-structured diagnostic interview

No data were provided for this outcome.

2. Suicide completion

No data were provided for this outcome.

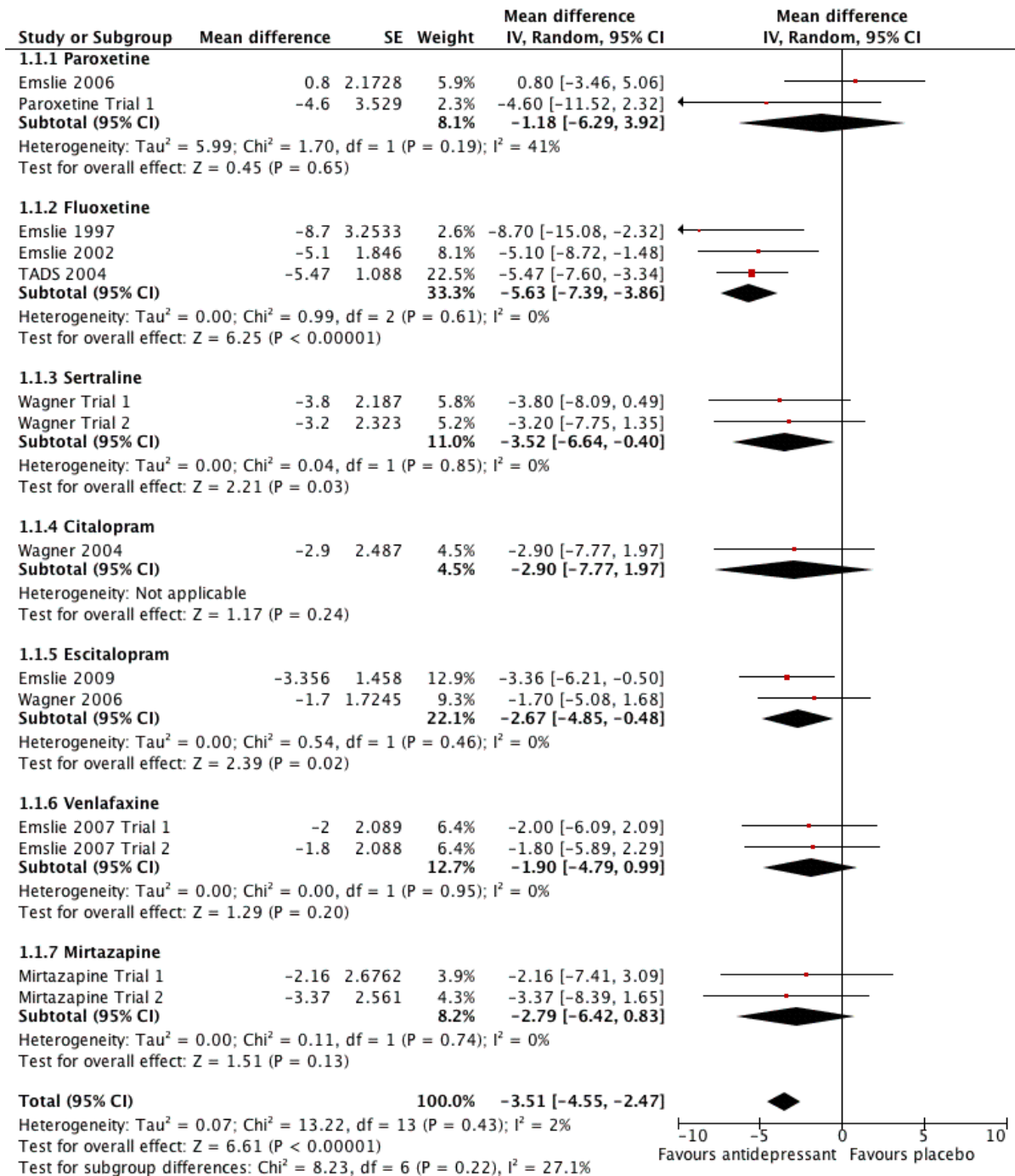
Secondary outcomes

1. Efficacy outcomes

1.1 Depression symptom severity (clinician-rated)

Compared with placebo, there was a small, statistically significant reduction in depression symptoms on the Children's Depression Rating Scale (CDRS-R) for those taking an antidepressant (14 trials; N = 2490; mean difference (MD) -3.51; 95% confidence interval (CI) -4.55 to -2.47) (Analysis 1.1). The majority of estimates were in the same direction, favouring antidepressants, with little heterogeneity ($\text{Chi}^2 = 13.22$, $\text{df} = 13$ ($P = 0.43$); $I^2 = 2\%$ (95% CI 0% to 48%)). There was no evidence that the individual drug class modified the effect of the newer generation antidepressants ($\text{Chi}^2 = 8.23$; $\text{df} = 6$, $P = 0.22$) (Figure 8). Three trials measured clinician-rated symptoms using the Schedule for Affective Disorders and Schizophrenia for School Aged Children (K-SADS); two of which tested paroxetine (Berard 2006; Keller 2001) and one which tested citalopram (Von Knorring 2006). Data have been extracted for this measure and are available in Table 1.

Figure 8. Forest plot of comparison: I Second generation antidepressant versus placebo (by medication), outcome: I.I Depressive symptom severity (CDRS-R).



We judged all trials to be at a high risk of bias, precluding a sensitivity analysis investigating how trials at a high risk of bias modify the pooled treatment effect.

The pooled treatment effect was not modified in a clinically important way with removal of the antidepressant classes selective norepinephrine reuptake inhibitors (SNRI) (venlafaxine) and tetracyclic antidepressants (TeCA) (mirtazapine) (sensitivity analysis pooled treatment effect: MD -3.73; 95% CI -5.09 to -2.36).

1.2 and 1.3 Remission or response

Compared with placebo, there was a statistically significant increase in the percentage of those who remitted/responded when taking an antidepressant (16 trials; N = 2924; risk ratio (RR) 1.18; 95% CI 1.08 to 1.28) (Analysis 1.2). There was generally consistency in the direction of effect across the trials, favouring antidepressant ($\text{Chi}^2 = 15.87$, $\text{df} = 14$ ($P = 0.32$); $I^2 = 3\%$ (95% CI 0 to 53)). There was no evidence that individual drug class modified the effect of the intervention, compared with placebo ($\text{Chi}^2 = 1.79$; $\text{df} = 5$; $P = 0.88$). The two trials testing mirtazapine did not provide data on response or remission.

We undertook sensitivity analysis to test the impact of using observed case (OC) data rather than last observation carried forward (LOCF) data. We had OC data for the trials of sertraline (Wagner Trial 1&2 (2003)), the four trials of paroxetine (Berard 2006; Emslie 2006; Keller 2001; Paroxetine Trial 1) and one trial of citalopram (Von Knorring 2006). Including OC data did not modify the pooled effect in an important way (RR 1.17; 95% CI 1.09 to 1.25) and there was no evidence of heterogeneity ($\text{Chi}^2 = 14.12$, $\text{df} = 15$ ($P = 0.52$); $I^2 = 0\%$ (95% CI 0 to 45)) (Analysis 1.3).

We judged all trials to be at a high risk of bias, precluding a sensitivity analysis investigating how trials at a high risk of bias modify the pooled treatment effect.

The pooled treatment effect was not modified in a clinically important way with the removal of the antidepressant class SNRI (venlafaxine) (sensitivity analysis pooled treatment effect: RR 1.19; 95% CI 1.07 to 1.32).

1.4 Depression symptom severity (self rated)

For self report depression scores, only five trials measured or reported this. One of these only reported median scores (Emslie 2006) and another did not report any measure of variance (Von Knorring 2006); however, we included Von Knorring 2006 assuming the baseline standard deviations from Berard 2006 as the follow-up standard deviations of Von Knorring 2006.

Three trials used the Beck Depression Inventory (BDI)/Childrens Depression Inventory (CDI) (Berard 2006; Emslie 1997; Von Knorring 2006). There was no evidence that self report depression scores were reduced for those taking an antidepressant (paroxetine and fluoxetine respectively) compared with placebo (N = 564;

MD -0.53; 95% CI -2.37 to 1.31) (Analysis 1.4). There was little variability and consistency in the direction of treatment effects between trials; however, with few trials, the confidence interval for I^2 was wide ($\text{Chi}^2 = 0.14$, $\text{df} = 2$ ($P = 0.93$), $I^2 = 0\%$ (95% CI 0% to 73%).

TADS 2004 reported data based on Reynolds Adolescent Depression Scale (RADS) final scores, with results available in Table 2.

1.5 Functioning

Nine trials measured functioning using the Children's Global Assessment Scale (CGAS) with only one (Almeida-Montes 2005) that did not provide usable data. There was evidence that antidepressants improved functioning compared with placebo (N = 1593; MD 2.20; 95% CI 0.90 to 3.49) (Analysis 1.5). There was consistency in the direction of treatment effects and little between-trial variability ($\text{Chi}^2 = 0.91$, $\text{df} = 6$ ($P = 0.99$); $I^2 = 0\%$ (95% CI 0% to 58%)). There was no evidence that the effect of newer generation antidepressants was modified by individual drug class ($\text{Chi}^2 = 0.81$; $\text{df} = 4$; $P = 0.94$). Three trials used Global Assessment of Functioning (GAF) to measure functioning with only two providing data (Emslie 2006 (paroxetine) and Emslie 2002 (fluoxetine); Von Knorring 2006 (citalopram) did not provide usable data. One trial (Keller 2001 (paroxetine) used the Autonomous Functioning Checklist. We have extracted data for these other measures and they are available in Table 3. The two trials of venlafaxine and the two trials of mirtazapine did not provide evidence that this outcome was measured.

2. Suicide-related outcomes (suicide-related behaviour and suicidal ideation)

Seventeen trials had data on suicide-related outcomes.

There was evidence of an increased risk in suicide-related outcomes for those receiving antidepressants compared with placebo (17 trials; N = 3229; RR 1.58; 95% CI 1.02 to 2.45) (Analysis 1.6). There was little heterogeneity and generally consistency in the direction of effects ($\text{Chi}^2 = 16.62$, $\text{df} = 13$ ($P = 0.22$); $I^2 = 22\%$ (95% CI 0% to 58%)). There was no evidence that the effect of new generation antidepressants on suicide-related outcomes was modified by drug class ($\text{Chi}^2 = 8.06$; $\text{df} = 6$; $P = 0.23$).

We judged all trials to be at a high risk of bias, precluding a sensitivity analysis investigating how trials at a high risk of bias modify the pooled treatment effect.

The pooled treatment effect was not modified in a clinically important way with the removal of the antidepressant class SNRI (venlafaxine) (sensitivity analysis pooled treatment effect: RR 1.46; 95% CI 1.00 to 2.12).

Two trials of adolescents only (TADS 2004 (fluoxetine; N = 221) and Emslie 2009 (escitalopram; N = 311)) measured suicidal

ideation on the Suicidal Ideation Questionnaire self report measure and showed no statistically significant evidence of effect on suicidal ideation for those taking an antidepressant compared with placebo (N = 532; MD 0.70; 95% CI -1.50 to 2.91) (Analysis 1.7). From the Medicines and Healthcare Products Regulatory Agency (MHRA) report, it was stated that [Wagner 2004](#) (child and adolescent) showed that citalopram was superior to placebo in reducing suicidal ideation based on item 13 of the CDRS-R; however, the report did not provide an estimate of the effect or measure of precision. Likewise, the MHRA report stated that the trial by [Von Knorring 2006](#) (adolescents only) showed that those taking citalopram had a slightly greater improvement in suicidal ideation than placebo, based on item nine of the K-SADS-P. However, an estimate of intervention effect (or precision) was not provided. No other trials reported a measure suicidal ideation.

3. Adverse outcomes

Eleven trials reported the number of children, adolescents or both who experienced any adverse event. These data were not available from the trials of sertraline, venlafaxine or mirtazapine or for two of the fluoxetine trials ([Almeida-Montes 2005](#); [Emslie 1997](#)). There was an increased risk of adverse events in those taking antidepressants compared with placebo (11 trials, N = 2136; RR 1.11; 95% CI 1.05 to 1.17) (Analysis 1.8). There was little heterogeneity and generally consistency in the direction of effects ($\text{Chi}^2 = 9.42$ df = 9 (P = 0.40); $I^2 = 4\%$ (95% CI 0% to 55%)). There was no evidence that the effect of new generation antidepressants on suicide-related outcomes was modified by drug class ($\text{Chi}^2 = 3.31$; df = 3; P = 0.35).

Side effects were reported differently in each trial, and a summary of the adverse events reported is available in [Table 4](#).

4. Completion of trial protocol

The rates of participants completing the trial did not differ between groups (18 trials; N = 3290; RR 0.99; 95% CI 0.94 to 1.05) (Analysis 1.9). However, there was some evidence of moderate heterogeneity in the relative rates of completion ($\text{Chi}^2 = 25.82$, df = 15 (P = 0.04); $I^2 = 42\%$ (95% CI 0% to 67%)). Although not statistically significant ($\text{Chi}^2 = 9.66$; df = 6; P = 0.14), there was moderate variability in relative rates of completion among the individual drug class ($I^2 = 37.9\%$ (95% CI 0% to 74%)). [Table 5](#) contains the full attrition rates for all trials.

B. Comparison 1: Subgroup analyses: newer generation antidepressants versus placebo by individual drug

1. Paroxetine

There were four trials on paroxetine that provided some outcome data.

Depression symptom severity (clinician-rated)

Meta-analysis of two trials reporting CDRS-R scores showed a non statistically significant reduction in depression symptom severity scores for those taking paroxetine compared with those taking placebo (MD -1.18; 95% CI -6.29 to 3.92) ([Figure 8](#)). Data based on the K-SADS are available in [Table 1](#).

Remission or response

There was no statistically significant increase in the percentage of those who met the criteria for response/remission when taking paroxetine compared with placebo (four trials; N = 704; RR 1.12; 95% CI 0.90 to 1.38) (on the criteria of 'response' for [Berard 2006](#), [Keller 2001](#) and [Paroxetine Trial 1](#) and remission for [Emslie 2006](#)). The response rate in the treatment groups in the three trials varied between 52% and 63%, and the response rates in the placebo groups varied between 41% and 58%; the remission rate in [Emslie 2006](#) was 23% in the treatment group and 28% in the placebo group.

We undertook sensitivity analysis to test the impact of using OC data rather than LOCF data. Including OC data did not modify the pooled effect in an important way (four trials; N = 522; RR 1.11; 95% CI 0.98 to 1.27) (on the criteria of 'response' for [Berard 2006](#), [Keller 2001](#) and [Paroxetine Trial 1](#) and remission for [Emslie 2006](#)). The response rate in the treatment groups in the three trials varied between 56% and 81% and the response rates in the placebo groups varied between 46% and 71%; the remission rates in [Emslie 2006](#) were 34% in the treatment group and 36% in the placebo group. It is clear that high attrition combined with the imputation strategy in the component trials impacts on the absolute rates.

Depression symptom severity (self rated)

One trial ([Berard 2006](#)) measured self report depression symptom severity and showed no statistically significant difference between groups (MD -0.43; 95% CI -2.91 to 2.05).

Functioning

There was no evidence that paroxetine improved functioning compared with placebo on the Children's Global Assessment Scale (C-GAS) in [Berard 2006](#) (MD 1.60; 95% CI -2.48 to 5.68). Data based on other measures used are presented in [Table 3](#).

Suicide-related outcomes

While there was evidence of an increased risk in suicide-related outcomes for those receiving paroxetine compared with placebo (Berard 2006; Emslie 2006; Keller 2001), the confidence interval was wide and did not exclude the possibility of no difference in the rates of suicide-related outcomes between groups (RR 1.57; 95% CI 0.46 to 5.31).

Adverse outcomes

There was no statistically significant increase in risk of adverse events in those taking paroxetine compared with placebo (RR 1.11; 95% CI 0.98 to 1.25). However, there was some evidence of moderate heterogeneity ($\text{Chi}^2 = 4.69$, $\text{df} = 3$ ($P = 0.20$); $I^2 = 36\%$ (95% CI 0% to 79%)) with three trials (Berard 2006; Emslie 2006; Keller 2001) demonstrating that adverse events were more common for those receiving paroxetine, but one demonstrating the opposite effect (Paroxetine Trial 1). Headaches were common side effects in both groups, as were nausea and dizziness. Somnolence, insomnia and emotional lability were also noted. Side effects were reported differently in each trial, and a summary of the adverse events reported is in Table 4.

Completion of trial protocol

The rates of participants completing the trial did not differ between the paroxetine and placebo groups (RR 0.94; 95% CI 0.86 to 1.02).

2. Fluoxetine

Four of five trials investigating the effectiveness of fluoxetine provided some outcome data.

Depression symptom severity (clinician-rated)

There was a statistically significant reduction in depression symptoms on the CDRS-R for those taking fluoxetine compared to placebo (three trials; MD -5.63; 95% CI -7.39 to -3.86).

Remission or response

Compared with placebo, there was a statistically significant increase in the percentage of those who met the criteria for remission (Emslie 1997; Emslie 2002; TADS 2004) or response (Almeida-Montes 2005) when taking fluoxetine (RR 1.47; 95% CI 1.03 to 2.08) (Figure 8). The remission/response rates in the fluoxetine groups in the four trials varied between 23% and 57% and in the placebo groups between 17% and 67%.

Depression symptom severity (self rated)

There was no evidence that self report depression scores were reduced for those taking fluoxetine compared with placebo (MD -1.30; 95% CI -5.87 to 3.27) on the CDI/BDI in one trial (Emslie 1997).

TADS 2004 reported data based on RADS final scores with results available in Table 2.

Functioning

Fluoxetine improved functioning when measured by the C-GAS from two trials (Emslie 1997; TADS 2004) (MD 3.08; 95% CI 0.14 to 6.02). Emslie 2002 measured functioning using the GAF and data based on this measure are available in Table 3.

Suicide-related outcomes

While there was evidence of an increased risk in suicide-related outcomes for those receiving fluoxetine compared with placebo, the difference between groups was not statistically significant (RR 1.77; 95% CI 0.85 to 3.69). There was no difference between groups in suicidal ideation scores post intervention in the one trial that measured this (TADS 2004).

Adverse outcomes

Data on adverse outcomes could not be extracted for Almeida-Montes 2005 or Emslie 1997. Based on two trials there was an increased risk of adverse events in those taking fluoxetine compared with placebo (RR 1.19; 95% CI 1.05 to 1.35).

Details about particular adverse events were only reported for TADS, in which the following adverse events were reported: headaches, diarrhoea, somnolence, insomnia, emotional lability, and mania or hypomania. Headache was most commonly reported with 12% and 9% of those in the fluoxetine and placebo groups experiencing these respectively. All other adverse events were reported in less than 3% of the participants, with similar rates observed in both groups. Details can be found in Table 4.

Completion of trial protocol

Based on four trials, the rates of participants completing the trial did not differ between those taking fluoxetine compared with placebo (RR 1.16; 95% CI 0.97 to 1.40).

3. Sertraline

There were two trials of sertraline. These trials were combined and treated as a single trial in the publication (Wagner Trial 1). The trial author states that "data were pooled in a prospectively defined combined analysis" (Wagner Trial 1, pg 1035). The author states that this pooling was planned a priori. The MHRA report,

however, reported results for the two trials separately. In total, 376 young people were randomised into the two trials.

Depression symptom severity (clinician-rated)

Compared with placebo, there was a statistically significant reduction in depression symptoms on the CDRS-R for those receiving sertraline (MD -3.52; 95% CI -6.64 to -0.40) (Figure 8).

Remission or response

Compared with placebo, there was no statistically significant increase in the percentage of those who responded when taking sertraline (two trials; N = 364; RR 1.17; 95% CI 1.00 to 1.36). The percentage of young people who met the criteria for response in the sertraline groups (two trials combined) was 69% and in the placebo groups was 59%.

We undertook sensitivity analysis to test the impact of using OC data rather than LOCF data. Including OC data did not modify the effect in an important way (RR 1.18; 95% CI 1.01 to 1.38). The percentage of young people who met the criteria for response in the sertraline groups (two trials combined) was 69% and in the placebo groups was 59%.

Depression symptom severity (self rated)

Self rated depressive symptom severity was not measured in these trials.

Functioning

Functioning was measured using the CGAS scale in both trials and there was no evidence that antidepressants improved functioning for those taking sertraline compared with placebo (MD 1.31; 95% CI -1.61 to 4.23).

Suicide-related outcomes

While there was evidence of an increased risk in suicide-related outcomes for those receiving sertraline compared with placebo, the difference between groups was not statistically significant and the confidence interval was wide (RR 2.97; 95% CI 0.61 to 14.52).

Adverse outcomes

Data were not provided for the number of overall adverse events in each group. Nausea was commonly reported in both groups. Adverse effects that were reported relatively frequently included diarrhoea, vomiting and insomnia. Emotional lability but not mania/hypomania was reported (see Table 4).

Completion of trial protocol

The rates of participants completing the trial did not differ between groups (RR 0.91; 95% CI 0.82 to 1.01).

4. Citalopram

There were two trials of citalopram.

Depression symptom severity (clinician-rated)

There was no statistically significant difference in the reduction of depression symptom severity scores (CDRS-R) for those taking citalopram compared with placebo based on one study (Wagner 2004) (MD -2.90; 95% CI -7.77 to 1.97) (Figure 8). Data from Von Knorring 2006 using the K-SADS measure are available in (Table 1).

Remission or response

Compared with placebo, there was a non statistically significant increase in the percentage of those who met the criteria for remission when taking citalopram (two trials; N = 407; RR 1.16; 95% CI 0.71 to 1.89). In the two trials of citalopram (Von Knorring 2006; Wagner 2004), the percentage of participants remitting in the citalopram groups varied between 33% and 36% and in the placebo groups between 24% and 36%.

We undertook sensitivity analysis to test the impact of using OC data rather than LOCF data. Including OC data from Von Knorring 2006 did not modify the pooled effect in an important way (RR 1.17; 95% CI 0.74 to 1.85). The response rate based on OC for those on citalopram was 51% (compared to 33% based on LOCF data) and 53% (compared to 36% based on LOCF data) for those in the placebo group.

Depression symptom severity (self rated)

One trial (Von Knorring 2006) provided data on self reported depression symptom severity and showed no difference between groups (MD -0.28; 95% CI -3.72 to 3.16).

Functioning

Data on improvement in functioning (CGAS) were available from one trial. There was no evidence that citalopram improved functioning compared with placebo (MD 2.50; 95% CI -1.52 to 6.52).

Suicide-related outcomes

While there was evidence of an increased risk in suicide-related outcomes for those receiving citalopram compared with placebo, the difference between groups was not statistically significant (RR 1.53; 95% CI 0.55 to 4.27).

Adverse outcomes

There was no increased risk of adverse events in those taking citalopram compared with placebo (two trials; RR 1.13; 95% CI 1.00 to 1.29). Headache was a commonly reported adverse outcome, as well as nausea, diarrhoea and insomnia. Dizziness and respiratory adverse events were also recorded. Emotional lability and mania/hypomania were not reported (see [Table 4](#)).

Completion of trial protocol

The rates of participants completing the trial did not differ between groups (two trials; RR 1.03; 95% CI 0.91 to 1.16).

5. Escitalopram

There were two trials of escitalopram.

Depression symptom severity (clinician-rated)

Compared with placebo, there was a statistically significant reduction in depression symptoms on the CDRS-R for those taking an escitalopram (two trials; MD -2.67; 95% CI -4.85 to -0.48) ([Figure 8](#)).

Remission or response

There was no statistically significant increase in the percentage of those who met the criteria for remission when taking escitalopram compared with placebo (RR 1.19; 95% CI 0.97 to 1.45). In the two trials, the percentage of participants who met the criteria for remission was 42% in the escitalopram group and 36% in the placebo group in [Emslie 2009](#) and 46% in the escitalopram group and 38% in the placebo group in [Wagner 2006](#).

Depression symptom severity (self rated)

There were no data on self reported depression symptom severity.

Functioning

Data on improvement in functioning (CGAS) were available from both trials. There was evidence that escitalopram improved functioning compared with placebo (MD 2.28; 95% CI 0.23 to 4.32).

Suicide-related outcomes

There was evidence of reduced risk in suicide-related outcomes for those receiving escitalopram compared with placebo, however the confidence interval did not exclude the possibility of no difference in risks between groups (RR 0.91; 95% CI 0.47 to 1.76).

Adverse outcomes

There was no difference in the risk of adverse events in those taking escitalopram compared with placebo (RR 1.03; 95% CI 0.94 to 1.14). In the [Wagner 2006](#) trial, headache was a commonly reported adverse outcome, as well as nausea, diarrhoea and insomnia. Dizziness and respiratory adverse events were also recorded. Emotional lability and mania/hypomania were not reported. In the [Emslie 2009](#) trial, abdominal pain, nausea, vomiting and dizziness were reported more often in the escitalopram group; headache, abdominal pain, diarrhoea, pharyngitis and rhinitis were reported in both groups ([Table 4](#)).

Completion of trial protocol

The rates of participants completing the trial did not differ between groups (RR 0.93; 95% CI 0.86 to 1.01).

6. Venlafaxine

There were two trials of venlafaxine.

Depression symptom severity (clinician-rated)

There was no statistically significant reduction of depression symptom severity scores (CDRS-R) for those taking venlafaxine compared with placebo in two trials (MD -1.90; 95% CI -4.79 to 0.99) ([Figure 8](#)).

Remission or response

There was no statistically significant increase in the percentage of those who met the criteria for response when taking venlafaxine compared with placebo (RR 1.16; 95% CI 1.00 to 1.35). In [Emslie 2007 Trial 1](#), the percentage of participants responding in the venlafaxine group was 63% compared with 51% in the placebo group and in [Emslie 2007 Trial 2](#) the percentage of participants responding in the venlafaxine group was 76% compared with 67% in the placebo group.

Depression symptom severity (self rated)

There were no data on self reported depression symptom severity.

Functioning

There were no data on functioning.

Suicide-related outcomes

There was evidence of an increased risk of suicide-related outcomes for those receiving venlafaxine compared with placebo, although there were few suicide-related events and the resulting confidence interval was very wide (one trial; RR 12.93; 95% CI 1.71 to 97.82).

Adverse outcomes

There were no data on the number of overall adverse events experienced by young people in these trials. Data on individual adverse events highlighted that abdominal pain and dizziness were reported more often in the treatment than the placebo group (see [Table 4](#)).

Completion of trial protocol

The rates of participants completing the trial did not differ between groups (RR 0.93; 95% CI 0.82 to 1.07).

7. Mirtazapine

There were two trials of mirtazapine.

Depression symptom severity (clinician-rated)

There was no statistically significant reduction of depression symptom severity scores (CDRS-R) for those taking mirtazapine compared with placebo (MD -2.79; 95% CI -6.42 to 0.83) ([Figure 8](#)).

Remission or response

No data on rates of remission or response were reported.

Depression symptom severity (self rated)

No data on self reported depression symptom severity were reported.

Functioning

No data on functioning were reported.

Suicide-related outcomes

There was evidence of reduced risk of suicide-related outcomes for those receiving mirtazapine compared with placebo, however the confidence interval did not exclude the possibility of no difference in risks between groups (RR 0.50; 95% CI 0.03 to 7.90).

Adverse outcomes

There were no data on the number of overall adverse events experienced by young people in these trials. Reports highlight a side effect profile for this medication that includes significant metabolic side effects including greater weight gain, increased appetite and hypertriglyceridaemia. The mirtazapine group also experienced greater somnolence, headaches, fatigue and urticaria (hives).

Completion of trial protocol

The rates of participants completing the trial did not differ between groups (RR 1.05; 95% CI 0.93 to 1.19).

C. Comparison 1: Subgroup analyses: newer generation antidepressants versus placebo by age (children versus adolescents)

Depression symptom severity (clinician-rated)

There was no evidence that age modified the effect of the newer generation antidepressants for any outcome including the following.

- Depression symptom severity ($\text{Chi}^2 = 0.44$; $\text{df} = 1$, $P = 0.51$).
- Remission or response ($\text{Chi}^2 = 1.54$; $\text{df} = 1$, $P = 0.21$).
- Functioning ($\text{Chi}^2 = 0.11$; $\text{df} = 1$, $P = 0.74$).
- Suicide-related outcomes ($\text{Chi}^2 = 0.31$; $\text{df} = 1$, $P = 0.58$).
- Adverse outcomes ($\text{Chi}^2 = 0.02$; $\text{df} = 1$, $P = 0.89$).

I. Children

Depression symptom severity (clinician-rated)

Compared with placebo, there was a statistically significant reduction in depression symptoms on the CDRS-R for children taking an antidepressant (eight trials; MD -3.18; 95% CI -5.95 to -0.41) (Analysis 2.1).

Remission or response

Only two trials provided data on children separately ([Emslie 2002](#); [Emslie 2006](#)). There was no statistically significant increase in the percentage of those who responded/remitted when receiving an antidepressant compared with placebo (RR 1.71; 95% CI 0.90 to 3.25) (Analysis 2.2).

Depression symptom severity (self rated)

Data were not available separately for children in the one trial ([Emslie 1997](#)) which reported self rated depression symptom severity that included this age group.

Functioning

Compared with placebo, there was no evidence that antidepressants improved functioning in children in two trials ([Emslie 1997](#); [Wagner 2006](#)) (MD 1.41; 95% CI -6.74 to 9.57) (Analysis 2.4). Functioning data were not available separately for children for the trials of sertraline ([Wagner Trial 1&2 \(2003\)](#)) or the citalopram trial ([Wagner 2004](#)).

Suicide-related outcomes

While there was evidence of an increased risk in suicide-related outcomes for children receiving an antidepressant compared with placebo, the difference between groups was not statistically significant (RR 2.07; 95% CI 0.67 to 6.35) (Analysis 2.5).

Adverse outcomes

Only one trial of children and adolescents (Emslie 2006) provided data on adverse outcomes separately for children and there was no evidence of an increased risk of adverse events in children taking antidepressants compared with placebo (RR 1.09; 95% CI 0.82 to 1.44) (Analysis 2.6).

Completion of trial protocol

Data on trial completion by age group were not available.

2. Adolescents

Depression symptom severity (clinician-rated)

Compared with placebo, there was a statistically significant reduction in depression symptoms on the CDRS-R for adolescents taking an antidepressant (10 trials; MD -4.21; 95% CI -5.50 to -2.92) (Analysis 2.1).

Remission or response

Compared with placebo, there was a statistically significant increase in the percentage of adolescents who met the criteria for remission or response on an antidepressant (seven trials; RR 1.13; 95% CI 1.02 to 1.26) (Analysis 2.2).

Depression symptom severity (self rated)

Data were not available separately for adolescents in one trial (Emslie 1997) reporting self rated depression symptom severity. The other trial, which included only adolescents, measured symptom severity using the BDI and observed no evidence that scores were reduced for those taking an antidepressant compared with placebo (MD -0.43; 95% CI -3.09 to 2.23) (Analysis 2.3).

Functioning

In the five trials of adolescents, or trials that provided data on functioning using the CGAS separately for adolescents, there was evidence that antidepressants improved functioning compared with placebo (MD 2.82; 95% CI 1.17 to 4.47) (Analysis 2.4). Functioning data were not available separately for adolescents for the trials of sertraline (Wagner Trial 1&2 (2003)) or for the citalopram trial (Wagner 2004).

Suicide-related outcomes

While there was evidence of an increased risk of suicide-related outcomes for adolescents receiving an antidepressant compared with placebo, the difference between groups was not statistically significant, although the effects encompassed by the confidence interval were generally suggestive of an increased risk (RR 1.47; 95% CI 0.99 to 2.19) (Analysis 2.5).

Adverse outcomes

There were six trials of adolescents or trials that reported data separately for adolescents and combined. There was an increased risk of adverse events in those taking antidepressants compared with placebo (RR 1.11; 95% CI 1.04 to 1.19) (Analysis 2.6).

Completion of trial protocol

Data on trial completion by age group were not available.

DISCUSSION

Summary of main results

We cannot make any comment on the comparative efficacy of individual antidepressant compounds. However, based on our meta-analyses, there was no evidence that the magnitude of intervention effects (compared with placebo) was modified by individual drug class. Pooled results showed some evidence that depression symptoms were reduced and that remission/response rates were higher in those taking a newer generation antidepressant compared to placebo. The clinical importance of these effects is unclear. Overall, the reduction in depression symptoms was relatively small at 3.51 (on a scale with a range from 17 to 113) (Summary of findings for the main comparison) and remission rates overall increased from 380 per 1000 (based on the median risk in the placebo group at post intervention) to 448 per 1000 (Summary of findings for the main comparison). Even for those drugs that had consistent evidence of efficacy across more than one outcome (fluoxetine and escitalopram), the reduction in average scores was 5.63 lower for those taking fluoxetine and 2.67 lower for those on escitalopram compared with those on placebo. Poznanski 1996 found a difference of 25 points on the CDRS-R scale between a clinically referred depressed group (n = 60) and a non clinical group (n = 223), and a difference of 19 points between clinically referred groups with (n = 60) and without depressive disorder (n = 18) (Poznanski 1996, pg 53). There was also a relatively small effect on functioning with those receiving fluoxetine and escitalopram having a score on average 3.08 higher and 2.28 points higher respectively than placebo on the CGAS (range 0 to 100). For fluoxetine, remission rates overall increased from 214 per 1000 (based on the median

risk in the placebo groups at post intervention) to 315 per 1000 (a difference of 101 with 95% confidence interval (CI) from 6 more to 231 more). While significantly more young people recover if on fluoxetine, the remission rates are relatively low overall indicating many young people do not experience a clinically meaningful reduction in symptoms. Methodological issues including large attrition rates and inappropriate outcome measurement, with the associated potential for selective reporting, make it difficult to draw conclusions about the clinical benefit.

In addition, the efficacy results need to be balanced with evidence about adverse outcomes. Over all drugs, the risk of a suicide-related outcome for those taking antidepressants was 58% (95% CI 2% to 245%) higher compared to those taking placebo. This equates to an increased of risk in a group with a median baseline risk from 25 in 1000 to 40 in 1000 ([Summary of findings for the main comparison](#)). There was some variability in the risk of suicide-related outcomes across the drugs ($I^2 = 22\%$; 95% CI 0% to 58%), but this was not statistically significantly different. Suicidal ideation was only measured in two trials with no effect of antidepressant treatment evident. Adverse events were greater for those taking an antidepressant, and there was no effect of antidepressant on the rate of trial completion.

There were few data available to investigate whether age modified the intervention effect. There was no evidence from the available data that age modified the effects of antidepressant medication. The evidence of the effectiveness of antidepressants for children was inconsistent, with a small overall reduction in depression symptoms (3.18) and remission rates in one study ([Emslie 2002](#)) of 44% for those receiving fluoxetine compared to 18% for those receiving placebo and in the other ([Emslie 2006](#)) 69% for those receiving escitalopram and 52% for those receiving placebo. For adolescents, there was consistent evidence across both these outcomes of antidepressants compared with placebo. However, the same cautions apply in terms of methodological considerations and the size of the effects.

Overall completeness and applicability of evidence

In this review, to our knowledge, we have presented data on efficacy and adverse outcomes, including suicide-related outcomes, from all published and unpublished trials examining the use of newer antidepressants for child and adolescent depressive disorder. Despite attempts to contact all trial authors, as well as pharmaceutical companies responsible for funding the included trials, there are many instances of missing data in terms of effect estimates. In two cases there is very limited reporting of trials by the Medicines and Healthcare Products Regulatory Agency (MHRA), with publication in peer-reviewed journals not yet available ([Mirtazapine Trial 1 & 2](#)). Likewise the trial by [Simeon 1990](#), which was stopped early, has never been published. We were unable to obtain any

further report of the trial [Glod 2004](#), the preliminary findings of which were published in a conference abstract.

The comparator group has an impact on the size of the effect and, in this case, the size of the effect in the placebo groups in these trials has been commented on (e.g. [Jureidini 2004](#)) and authors of the included trial reports themselves have commented, suggesting, for example, that this may be to do with the large amount of contact trial participants receive. Trials consistently include regular (often weekly) assessments and in a significant number of trials some sort of supportive contact or therapy was allowed. The interaction between participants and trial investigators was seldom standardised, as shown in a trial specifically investigating this issue, and while this interaction is typical of what takes place in real world clinical encounters, it impacts in an unknown way on the detection of differences between the active drug and placebo ([Dunlop 2010](#)). The characteristics of the participants also impact on the intervention effect. It is possible, for example, that the effect shown in the fluoxetine trials may be due to exclusion of placebo responders in the lead-in time to the start of the trial. The placebo remission rate was higher in trials of other compounds and has previously been cited as a cause for concern regarding the efficacy of selective serotonin reuptake inhibitors (SSRIs) ([Newman 2004](#)). Again, it is notable that in the trials of fluoxetine the attrition rate was consistently higher in the placebo group compared with the fluoxetine group. There were generally very high attrition rates in all of the trials and it is unclear what effect this has on treatment estimates. The young people in the trials are not likely to be representative of clinical populations typically seen in public child and adolescent mental health services as most trials included recruitment using advertisement, excluded many comorbid disorders, and in many trials excluded those who responded to placebo in the lead-in stage of the trial. Also excluded were those at risk of suicide despite suicide-related behaviours and comorbidity being significant features of clinically referred young people with depressive disorders ([Birmaher 1996](#); [Kovacs 1984](#); [Marttunen 1998](#); [Petersen 1993](#)). The baseline severity of young people included was in the moderately severely ill range, with a large proportion of included participants experiencing their first episode of depression (although in many cases this first episode was of a long duration). The effectiveness of newer generation antidepressants in young people with more severe disorders and complex presentations, including comorbid conditions and suicide risk, is therefore unknown.

Quality of the evidence

There was limited information on the conduct of trials in relation to allocation concealment, blinding and compliance. Blinding is an issue when clinician-rated scales are the main outcome, particularly in the context of an inactive placebo where it may be possible to guess the assigned treatment group given side and other physiological effects likely in this group ([Moncrieff 2004](#)). In most cases the write-up simply indicated that the trial was 'dou-

ble-blind', some stating that the placebo and medication capsules were identical and some stating that clinicians or trial personnel and participants were blinded. Only one trial specifically stated that outcome assessors were blinded.

The issue of reporting bias is important. [Kirsch 2008](#) highlight in their meta-analysis of all trials of antidepressants submitted to the Food and Drug Administration (FDA) that effect sizes are smaller when unpublished studies are included and [Turner 2008](#) showed that whether and how trials of antidepressants were published depended on the outcome of the trial. We located several trials that have not been fully reported or published at all. It is also worth noting that in many cases we were unable to obtain the required data from the published paper but had to contact authors or the pharmaceutical company. Reporting bias within the published reports of included trials was difficult to assess given the conduct of a trial can be obscured in the write-up for publication. Full and explicit reporting of changes in outcome definition was only undertaken by one investigator, however the primary outcome was reported and findings discussed ([Emslie 1997](#); [Jureidini 2004](#)). The possibility of reporting bias was highlighted in a letter to the editor regarding post hoc alterations of response definitions in the trial by [Keller 2001](#) ([Jureidini 2003](#)).

What level of improvement constitutes a meaningful clinical outcome is uncertain given response and remission were defined and reported variously both within and across trials, with the noted possibility of alteration of this definition, and the possibility of reporting bias as a result.

There was evidence of inappropriate methods of imputation with trialists often using last observation carried forward data ([Sterne 2009](#)). It was often the case that some randomised patients were not included in the final analysis.

The majority of trials were pharmaceutically funded. Two of the four fluoxetine trials were not pharmaceutically funded ([Emslie 1997](#); [TADS 2004](#)) (the TADS trial had an unrestricted education grant from Eli Lilly). Research has shown that across different health fields, pharmaceutically funded studies are more likely to have results favouring the pharmaceutical company's product ([Lexchin 2003](#); [Sismondo 2008](#)).

Measurement issues

A standard definition of remission or response would have been ideal; however, to calculate this individual patient data would have been necessary. Further, given a diagnosis of major depressive disorder on DSM criteria was an entry criterion for most of the trials, this may be considered the most desirable outcome measure.

Additionally, given symptom improvement or resolution does not necessarily correlate with improvements in functioning (see [Winters 2005](#) for a review), the latter would seem a more clinically important patient outcome to collect. However, this was inconsistently measured and reported on scales that often did not have established psychometric properties. In trials that used the

Children's Global Assessment Scale (CGAS) (an adaptation of the Global Assessment of Functioning (GAF) used in adults), there was some evidence that functioning was improved for those on an antidepressant compared with placebo. Self report data may also tap an outcome that is meaningful to the young person, however, few individual trials reported this.

The trials are designed only to examine the short-term effects of antidepressant medication, however, this does not preclude the possibility that the effectiveness of treatment is only apparent over a longer period of time. Long-term follow-up would be required to assess this.

Potential biases in the review process

It should be noted that the review process included collection of data from various sources. There were more complete data for the trials on paroxetine due to publication of trial reports by SmithKline Beecham on the internet. Details of aspects of trial methodology were relatively brief even in this case. Information and data from other trials were taken variously from scientific journal publications, from the MHRA data and, in some cases, obtained directly from trial authors and pharmaceutical companies.

We were unable to obtain a large number of potentially relevant publications to screen for inclusion. The majority of these were conference proceedings and there is some evidence that such data are often not published and that publication is associated with positive results ([Sherer 2007](#)). However, funnel plots for the outcomes remission and suicide-related behaviour were not suggestive of small-study effects and, moreover, the contour-enhanced funnel plots did not indicate that statistically significant results were more likely to be reported.

Agreements and disagreements with other studies or reviews

The issue of the efficacy of antidepressants and the associated suicide risk has been the topic of much research, review and debate in the literature.

Healy was one of the first to publish regarding the possibility of an increased risk of suicide associated with SSRIs ([Healy 2003](#)). Since then, observational studies have highlighted an association between antidepressant use and suicide risk in adolescents. For example, [Olfson 2006](#) showed an association between use of antidepressants and suicide attempt in children and adolescents and [Simon 2006b](#) showed a greater risk of suicide attempt that resulted in hospitalisation for children and adolescents compared with adults.

Data on risk of suicide for children and adolescents have also come from meta-analyses based on trial data submitted to the FDA ([Hammad 2006](#)) and the Committee on Safety in Medicines in the UK ([Dubicka 2006](#)) that show similar results to our review.

These findings have more recently been extended to young adults (18 to 25-year olds) (Stone 2009). Additionally, Dubicka 2006 analysed these data by type of suicide-related outcome showing suicidal ideation in 1.2% of those on SSRIs compared with 0.8% on placebo (odds ratio (OR) 1.45; 95% CI 0.54 to 3.88), self harm in 3.3% of those on SSRIs versus 2.6% on placebo (OR 1.44; 95% CI 0.70 to 2.97) and attempts in 1.9% of those on an SSRI versus 1.2% on placebo (OR 1.70; 95% CI 0.76 to 3.81) (Dubicka 2006). Another review (Bridge 2007) calculated the risk difference for suicide-related behaviours for antidepressant versus placebo for major depressive disorder (MDD) (13 trials), obsessive-compulsive disorder (OCD) (six trials) and non OCD anxiety disorders (six trials). The risk difference for suicide-related behaviours was not statistically significant for MDD alone, although the risk ratio was 1.6; 95% CI 1.0 to 2.7 (Bridge 2007). The risk difference calculated by Bridge 2007 was 1% compared with a risk difference of 1.5% based on risk ratio (assuming a median baseline risk) calculated in this review.

Counter to this evidence showing an increase in suicide-related behaviours, several arguments have been mounted.

Firstly, observational and ecological studies have shown no increased suicide risk related to antidepressant use. For example, observational studies by Valuck 2004 (a retrospective cohort study) and Sondergard 2006 (a register-linkage study) showed no statistically significant increases in suicide attempt (hazard ratio (HR) 1.43; 95% CI 0.89 to 2.82) and completed suicide (RR 4.47; 95% CI 0.95 to 20.96) (respectively) for children and adolescents taking antidepressants. Two reviews of studies investigating this link highlighted that observational studies have not shown a strong association between antidepressants and suicide risk (Hall 2006; Isacson 2010). Hall 2006 highlighted that observational studies have tended to show that those at increased risk of suicide were more likely to be prescribed an antidepressant (Hall 2006) and therefore the association between antidepressant use and adolescent suicide-related behaviours may be more likely to be explained by severity of depression, and not caused by the antidepressant (Friedman 2007; Simon 2006). However, issues of potential confounding bias arising in observational studies are minimised in trials where, on average, those in the intervention and control groups are comparable in both measured and unmeasured confounders. Hall 2006 and Isacson 2010 have highlighted that the majority of ecological studies have shown that suicide mortality has decreased when SSRI use has increased in various counties or regions (Hall 2006). However, these studies either have not included children and adolescents or analysed this age group separately, with the exception of Olfson 2003 in the US. Olfson 2003 showed that increasing use of antidepressants was associated with a decreased suicide rate. A subsequent study using US and Dutch data showed a decrease in SSRI prescriptions associated with an increased suicide rate in both countries (Gibbons 2007).

There are shortcomings to these types of ecological studies, which have been discussed in the debate in this area. It is commonly ac-

knowledge that ecological studies are not able to establish a causal link. For example, it could be that there are lower suicide rates in countries with less stigma associated with seeking intervention for mental health disorders, where antidepressant prescriptions rates are therefore higher (Gunnell 2004; Simon 2006). Ecological studies have focused on short time periods, for example, Healy 2009 and Reseland 2008 contend that ecological studies have omitted data from the 1960s and 1970s when suicide rates climbed despite the largest increase in antidepressant use seen in the last few decades; Safer 2007 similarly argues that time periods prior to the introduction of SSRIs have not been included when drops in suicide rates were observed. Another issue of ecological studies is that the relationships observed between variables at the aggregated level may not be the same as those observed at the individual level. An example of this is the possibility that increased use of antidepressants may be due to chronic users not new users, whose risk of treatment-induced suicide is less (Reseland 2008). There are issues of confounding, for example fluctuations in suicide rates may be due to cultural and religious factors (De Leo 2002). Further, there is the issue of redefinition of the outcome over time. For example, reanalysis of data on suicide rates, antidepressant sales and autopsy in Nordic countries highlighted the possibility that observed decreases in suicide rates may be due to falling autopsy rates, and an associated increase in the rate of deaths from unspecified causes (Reseland 2008).

Finally, the observation of increased suicide rates associated with a decrease in the rate of SSRI prescriptions may be because the recent censure of SSRIs has resulted in pessimism about being able to provide effective treatment. This may have caused both a decrease in help seeking and a disinclination to diagnose depressive disorders, supported by data showing decreased rates of diagnosis for child and adolescent depression in the USA (Libby 2007).

The fact that there were no reports of completed suicide in a total sample of 2240 young people has been highlighted. However, suicide is a rare event and much larger sample sizes with longer follow-up would be needed to assess the risk fully. Healy 2009 cautions that it is more accurate to state that no suicides were recorded and notes that large loss to follow-up makes it possible that suicides did occur. Research suggests that the best predictor of eventual suicide completion is previous suicidal ideation and suicidal behaviour so that an increase in rates of these outcomes may increase the risk of future suicide completion (Andrews 1992; Brent 1986; Brent 1993; Lewinsohn 1996). The trials included in our review and others do not provide long enough follow-up to determine whether short-term increased risk of suicide-related outcomes is subsequently followed by a longer-term risk reduction. An increase in risk followed by a steadily decreasing risk has been shown within one month of starting antidepressant use in one trial of adults (Jick 2004). Data on timing of suicidal risk has been analysed from the TADS 2004 trial and showed that there was no decrease in risk after one month as seen in adult data, again meaning it is unclear if there is a longer-term risk reduction.

In all reviews the rates of suicide-related behaviour are much lower than reports of suicidal ideation and behaviour in adolescents from the community. In some population-based studies suicidal ideation is reported by as many as 12% to 23% of young people and suicide attempts by 4% to 8% of adolescents aged 13 to 18 years (AHRG 2003; Grunbaum 2002; Lewinsohn 1996; Sawyer 2001). This could be explained by exclusion of those at risk of suicide from all but two trials (Emslie 2002; Von Knorring 2006). Most trials also excluded those with many co-morbid conditions. Co-morbidity is related to an increased risk of suicidal ideation and suicidal behaviours (Andrews 1992; Asarnow 1992; Brent 1986; Esposito 2002; Kovacs 1993; Shaffer 1996; Wetzler 1996).

The findings regarding an increased risk of suicide-related behaviours need to be balanced with concerns regarding the risk of untreated depression and the need to balance benefits and harms (AACAP 2004; Ebmeier 2006; Simon 2006).

Several reviews of the efficacy of SSRIs in children and adolescents with depression have been published. Earlier reviews based conclusions on narrative summaries of individual trial results (Brent 2004; Cheung 2005; Wagner 2005) not meta-analyses. Of these, the conclusions of Cheung 2005 are the most conservative, stating data are limited, commenting specifically on the inappropriateness of the outcomes used and cautioning clinicians to consider the use of SSRIs carefully.

A number of meta-analyses have been also published. In earlier meta-analyses Jureidini 2004 highlighted methodological problems, particularly regarding reporting, and concluded cautiously that there is only a small benefit of SSRIs that should be balanced with the risks. In contrast Cohen 2004 concluded SSRIs are an effective treatment for adolescent depression based on data published up to that point. In the same year, a review by Whittington 2004 highlighted the change in risk-benefit profile when unpublished data were included; their inclusion resulted in a favourable risk-benefit profile only for fluoxetine, with a number needed to treat to benefit (NNTB) of 6 (95% CI 4 to 15) for remission, and 5 (95% CI 4 to 13) for response. More recently, Tsapakis 2008 has published a systematic review including meta-analyses of any antidepressant (including tricyclic antidepressants (TCAs), SSRIs, monoamine oxidase inhibitors (MAOIs) and new antidepressants) for children and adolescents up to the age of 20 with a depressive disorder. The authors' conclusions were cautious, highlighting limited efficacy, particularly for children and in studies that combine adolescents and children. They found little appreciable difference between different types of antidepressants, however, suggesting that fluoxetine may be more effective, especially for adolescents. Similar to our findings, they questioned the appropriateness of outcome measures and noted that the young people included in trials may not be representative of more classic 'adult major depression' and may be less severe than young people typically seen in clinical practice.

In a meta-analysis of antidepressants for all age groups, Ioannidis 2008 questioned the appropriateness of outcome measures used

and highlighted issues related to selective and distorted reporting and interpretation, calling into the question the effectiveness of antidepressants even in adult populations.

The results of our review present a dilemma for those treating young people with depressive disorders. While overall the reduction in depression symptoms is statistically significant, there are questions about the clinical effectiveness. The trials are of young people not representative of those typically presenting for treatment in clinics and furthermore they had some significant methodological shortcomings, making it difficult to draw firm conclusions. Potential benefit must be balanced with the finding that SSRIs are associated with a statistically significant increased risk of suicide-related behaviour (a combination of suicidal ideation and definitive suicidal behaviour). Much debate on this matter has taken place, with variable views.

While there is a lack of power to adequately examine rare events such as suicide-related outcomes in trials, meta-analysis offers a method for obtaining reliable evidence for these rare outcomes across trials. Given a lack of consistent measurement of various aspects of suicide risk, the clinical implications of the findings are still somewhat unclear.

It is unknown how children and adolescents with a depressive disorder and co-morbid conditions, who are at risk of suicide (i.e. those more typical of the young people who present at health services), would respond to SSRIs. Overall, the data regarding the benefits of SSRIs for child and adolescent depression are far from compelling while the information on the risks is limited.

Clinically, it must be kept in mind that there are significant risks in not treating depressive disorder, which has an increased risk of suicide completion, as well as impacts on academic and social functioning (Brent 2002; Ebmeier 2006; Fleming 1993; Lewinsohn 1998; NHMRC 1997; Rao 1995).

AUTHORS' CONCLUSIONS

Implications for practice

The evidence for the effectiveness of newer generation antidepressants compared with placebo in the treatment of depressive disorder in children and adolescents is not compelling; the overall effects of antidepressants compared with placebo were not modified by drug. The methodological shortcomings of the trials make it difficult to interpret outcome data on the efficacy of newer generation antidepressant medications. This is a particularly unsatisfactory since large numbers of children and adolescents have participated in trials, and we are still unable to answer the important clinical question of whether antidepressant medications are effective in treating depressive disorders. The search for treatments that reliably lead to good clinical outcomes must continue.

Even when there is evidence that newer generation antidepressants reduce depression symptom severity, it is unclear whether the dif-

ference in effect between antidepressants and placebo reflects a difference that is of clinical importance to patients. As trials have largely been done in children and adolescents with no co-morbid conditions and with no significant suicidal ideation, it is unclear how children and adolescents with more serious difficulties and those at risk of suicide would respond. There is evidence to suggest an increased risk of suicide-related behaviours (combined suicidal ideation and definitive suicidal behaviour) in those treated with antidepressant medication, but the importance of this is unclear as is the association between antidepressant medication and suicide completion. Untreated depressive disorder is associated with the risk of completed suicide and impacts on academic and social functioning.

For clinicians, the results of the review may mean that the threshold of severity for treatment of a depressive disorder with antidepressant medication is raised. Clinicians should make every effort to present the information on the potential benefits and risks of SSRIs, including the risks of untreated depression, and together with the child or adolescent and their family, consider the various options for treatment. This should include consideration of psychological treatments such as cognitive behavioural therapy and other non medication options. The risk of suicide should be assessed and, if medication is used, this should be monitored particularly closely. If a decision to use medication is agreed, then fluoxetine might be considered given it is recommended as first-line medication in guidelines (McDermott 2011; NICE 2005), although based on this review no definitive recommendations can potentially be made until the comparative effectiveness of individual drugs is examined in a multiple-treatment meta-analysis.

Given the evidence does not clearly answer questions about the effectiveness and harms of newer generation antidepressants, there is a need for further research.

Implications for research

It is clear from the results that we need more effective treatments for depressive disorders in young people. The debate about efficacy and risk of suicide illustrates that there is no one clearly effective treatment. Young people with a depressive disorder who present for treatment are likely to be different to those in the included trials in this review; moreover those presenting for treatment in clinical

services are heterogeneous. Trials should include young people more typical of those presenting to clinical services and allow analysis of different subgroups of these young people. This may be an important avenue of new research, for example, there may be a genetic underpinning to medication responsiveness.

Trials that are undertaken should address the methodological shortcomings identified in this review, including adequate blinding, particularly of outcome assessors, consistent definition and use of clinically important outcomes and longer-term follow-up.

Within the context of the currently available trial data, individual patient data meta-analyses may be useful in examining whether the effect of treatment differs in particular subgroups.

In the meantime, given clinicians require guidance in terms of which individual antidepressant medication to use, in order to more robustly examine the comparative efficacy of these compounds, a multiple-treatment meta-analysis should be undertaken.

ACKNOWLEDGEMENTS

Alison Ward for her contribution to the protocol, data extraction and some input on the write-up of the review.

Steve McDonald, Australasian Cochrane Centre assisted with the development of the search strategy for the original review.

The CCDAN editorial base team assisted at all stages with helpful comments and guidance, including double data checking for the earlier version of the review; in this version of the review they have provided support in contacting pharmaceutical companies for additional data.

The editors and peer reviewers assisted with editorial comments at all stages of the review.

Trialists who provided data including Dr Graham Emslie, Dr Robert Milin, Professor Martin Keller and Professor John March.

Pharmaceutical companies who provided data on the web (SmithKline Beecham) and provided data as requested (Dr Jennie Jacobson from Eli Lilly).

MHRA for providing reports of trial data.

REFERENCES

References to studies included in this review

Almeida-Montes 2005 *{published data only}*

Almeida-Montes LG, Friederichsen A. Treatment of major depressive disorder with fluoxetine in children and adolescents. A double-blind, placebo-controlled study [Tratamiento del trastorno depresivo mayor con fluoxetina en niños y adolescentes. Estudio doble ciego, controlado con placebo]. *Psiquiatria Biologica* 2005;**12**(5):198–205.

Berard 2006 *{published and unpublished data}*

Berard R, Fong R, Carpenter DJ, Thomason C, Wilkinson C. An international, multicentre, placebo-controlled trial of paroxetine in adolescents with major depressive disorder. *Journal of Child and Adolescent Psychopharmacology* 2004;**16**(1-2):59–75.
GlaxoSmithKline. A double-blind, multicentre placebo controlled study of paroxetine in adolescents with unipolar major depression [Study 377]. www.gsk.com/media/paroxetine.htm (accessed 16 August 2004).
Medicines and Healthcare products Regulatory Agency (MHRA). Paroxetine Study 2 ID#377. <http://medicines.mhra.gov.uk> (accessed 20 June 2004).
Milin R, Simeon J, Spenst W. Double-blind study of paroxetine in adolescents with unipolar major depression. Unpublished manuscript.

Emslie 1997 *{published data only}*

Emslie GJ, Rush AJ, Weinberg WA, Kowatch RA, Carmody T, Mayes TL. Fluoxetine in child and adolescent depression: acute and maintenance treatment. *Depression and Anxiety* 1998;**7**:32–9.
Emslie GJ, Rush AJ, Weinberg WA, Kowatch RA, Hughes CW, Carmody T. A double-blind placebo controlled trial of fluoxetine in depressed children and adolescents. Sixth World Congress of Biological Psychiatry, Nice, France. June 22–27. 1997.
* Emslie GJ, Rush R, Weinberg WA, Kowatch WA, Hughes CW, Carmody T, et al. A double blind, randomized, placebo-controlled trial of fluoxetine in children and adolescents with depression. *Archives of General Psychiatry* 1997;**54**(11):1031–7.
Emslie GJ, Weinberg WA, Mayes TL. Treatment of children with antidepressants: focus on selective serotonin reuptake inhibitors. *Depression and Anxiety* 1998;**8**(Suppl 1):13–7.
Hughes CW, Emslie G, Kowatch R, Weinberg W, Rintelmann J, Rush AJ. Clinician, parent, and child prediction of medication or placebo in double-blind depression study. *Neuropsychopharmacology* 2000;**23**(5):591–4.
Kowatch RA, Carmody TJ, Emslie GJ, Rintelmann JW, Hughes CW, Rush AJ. Prediction of response to fluoxetine and placebo in children and adolescents with major depressive disorder: a hypothesis generating study. *Journal of Affective Disorders* 1999;**54**:269–76.
Medicines and Healthcare products Regulatory Agency (MHRA). Fluoxetine Study 1 ID# X065. <http://medicines.mhra.gov.uk> (assessed 20 June 2004).

medicines.mhra.gov.uk (assessed 20 June 2004).

Rintelmann JW, Emslie GJ, Rush AJ, Varghese T, Gullion CM, Kowatch RA, et al. The effects of extended evaluation on depressive symptoms in children and adolescents. *Journal of Affective Disorders* 1996;**41**:149–56.

Emslie 2002 *{published data only}*

Emslie GJ, Heiligenstein JH, Hoog SL, Judge R, Brown EB, Nilsson M. Fluoxetine for acute treatment of depression in children and adolescents: a placebo-controlled, randomized clinical trial. Proceedings of the 39th Annual Meeting of the American College of Neuropsychopharmacology; 2000 Dec 10–14; San Juan, Puerto Rico. 2000:172.
* Emslie GJ, Heiligenstein JH, Wagner KD, Hoog SI, Ernest DE, Brown E, et al. Fluoxetine for acute treatment of depression in children and adolescents: a placebo-controlled randomized clinical trial. *Journal of the American Academy of Child and Adolescent Psychiatry* 2002;**41**(10):1205–15.
Jain S, Carmody TJ, Trivedi MH, Hughes C, Bernstein IH, Morris D, et al. A psychometric evaluation of the CDRS and MADRS in assessing depressive symptoms in children. *Journal of the American Academy of Child & Adolescent Psychiatry* 2007; Vol. 46, issue 9:1204–12.
Medicines and Healthcare products Regulatory Agency (MHRA). Fluoxetine Study 2 ID# HCJE. <http://medicines.mhra.gov.uk> (accessed 20 June 2004).
Nilsson M, Joliat MJ, Miner CM, Brown EB, Heiligenstein JH. Safety of subchronic treatment with fluoxetine for major depressive disorder in children and adolescents. *Journal of Child and Adolescent Psychopharmacology* 2004;**14**(3):412–7.

Emslie 2006 *{published and unpublished data}*

Brooks SJ, Krulwicz MA, Kutcher S. The Kutcher Adolescent Depression Scale: assessment of its evaluative properties over the course of an 8-week pediatric pharmacotherapy trial. *Journal of Child and Adolescent Psychopharmacology* 2003;**13**(3):227–349.
Emslie GJ, Wagner KD, Kutcher S, Krulwicz S, Fong R, Carpenter DJ, et al. Paroxetine treatment in children and adolescents with major depressive disorder: a randomized, multicentre, double-blind, placebo-controlled trial. *Journal of the American Academy of Child and Adolescent Psychiatry* 2006;**45**(6):709–19.
* GlaxoSmithKline. A randomized multicentre, 8-week, double-blind, placebo-controlled flexible dose study to evaluate the efficacy and safety of paroxetine in children and adolescents with major depressive disorder (29060/701). www.gsk.com/media/paroxetine.htm (accessed 16 August 2004).
Medicines and Healthcare products Regulatory Agency (MHRA). Paroxetine Study 3 ID#701. <http://medicines.mhra.gov.uk> (accessed 20 June 2004).

Emslie 2007 *{published data only}*

Emslie GJ, Finding RL, Yeung PP, Kunz NR, Li Y, Durn BL. Efficacy and safety of venlafaxine ER in children and adolescents with major depressive disorder (poster

- presentation). Proceedings of the American Psychiatric Association; 2004 May 1-6; New York. 2004.
- Emslie GJ, Findling RL, Yeung PP, Kunz NR, Li Y. Venlafaxine ER for the treatment of pediatric subjects with depression: results of two placebo-controlled trials. *Journal of the American Academy of Child and Adolescent Psychiatry* 2007;**46**(4):479-88.
- Emslie GJ, Yeung PR, Kunz NR, Li Y. Long-term efficacy and safety of venlafaxine ER in children and adolescents with major depressive disorder. Proceedings of the American Psychiatric Association; 2004 May 1-6; New York. 2004.
- Emslie 2007 Trial 1 {published data only}**
- Wyeth. Double-blind, placebo-controlled study of venlafaxine ER in children and adolescents with major depression. (Protocol 0600B1-382-US (Study Period: October 1997 to September 2000)). <http://www.clinicalstudyresults.org/documents/company-study-44-0.pdf> 2004 (accessed 24 November 2011).
- Emslie 2007 Trial 2 {published data only}**
- Wyeth. Double-blind, placebo-controlled study of venlafaxine ER in children and adolescents with major depressive disorder. (Protocol 0600B1-394-US) (Study Period: August 2000 to August 2001). <http://www.clinicalstudyresults.org/documents/company-study-2252-0.pdf> 2004 (accessed 24 November 2011).
- Emslie 2009 {published data only}**
- Emslie GJ, Ventura D, Korotzer A, Tourkodimitris S. Escitalopram in the treatment of adolescent depression: a randomized placebo-controlled multisite trial. *Journal of the American Academy of Child and Adolescent Psychiatry* 2009;**48**(7):721-9.
- Keller 2001 {published data only}**
- GlaxoSmithKline. A multi-centre, double-blind, placebo controlled study of paroxetine and imipramine in adolescents with unipolar major depression - acute phase (29060/329). www.gsk.com/media/paroxetine.htm (accessed 16 August 2004).
- Jureidini J, Tonkin A. Paroxetine in major depression. *Journal of the American Academy of Child and Adolescent Psychiatry* 2003;**42**(5):514.
- Keller MB, Ryan ND, Birmaher B, Klein RG, Strober M, Wagner KD, et al. Paroxetine and imipramine in the treatment of adolescent depression [Abstract NR206]. Proceedings of the 151st Annual Meeting of the American Psychiatric Association: New Research Program and Abstracts; 1998 May 30 - June 4; Toronto, Ontario, Canada 1998:123.
- * Keller MB, Ryan ND, Strober M, Klein RG, Kutcher SP, Birmaher B, et al. Efficacy of paroxetine in the treatment of adolescent major depression: a randomised controlled trial. *Journal of the American Academy of Child and Adolescent Psychiatry* 2001;**40**(7):762-72.
- Medicines and Healthcare products Regulatory Agency (MHRA). Paroxetine Study 1 ID#329. <http://medicines.mhra.gov.uk> (accessed 20 June 2004).
- Mirtazapine Trial 1 {published data only}**
- * www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/DevelopmentResources/UCM164066.pdf.
- Rapporteurs Public Assessment Report for Paediatric Data in EU Worksharing Procedure: Mirtazapine (Remeron). UK/H/0016/pdWS/001 (http://www.hma.eu/fileadmin/dateien/Human_Medicines/CMD_h/Paediatric_Regulation/Assessment_Reports/Article_45-work-sharing/Mirtazapine_2010-07-45PaedPAR.pdf) 18 June 2010.
- Mirtazapine Trial 1 & 2 {published data only}**
- www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/DevelopmentResources/UCM164066.pdf.
- Rapporteurs Public Assessment Report for Paediatric Data in EU Worksharing Procedure: Mirtazapine (Remeron). UK/H/0016/pdWS/001 (http://www.hma.eu/fileadmin/dateien/Human_Medicines/CMD_h/Paediatric_Regulation/Assessment_Reports/Article_45-work-sharing/Mirtazapine_2010-07-45PaedPAR.pdf) 18 June 2010.
- Organon Pharmaceuticals USA Inc. A multicenter, randomized, double-blind, placebo-controlled, efficacy and safety study of Remeron® in outpatient children and adolescents with major depressive disorder [Studies 1 and 2]. <http://www.clinicalstudyresults.org/documents/company-study-51-0.pdf> 2001 (accessed 24 November 2011).
- Mirtazapine Trial 2 {published data only}**
- www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/DevelopmentResources/UCM164066.pdf.
- Rapporteurs Public Assessment Report for Paediatric Data in EU Worksharing Procedure: Mirtazapine (Remeron). UK/H/0016/pdWS/001 (http://www.hma.eu/fileadmin/dateien/Human_Medicines/CMD_h/Paediatric_Regulation/Assessment_Reports/Article_45-work-sharing/Mirtazapine_2010-07-45PaedPAR.pdf) 18 June 2010.
- Paroxetine Trial 1 {published data only}**
- GlaxoSmithKline. A randomised, double-blind, placebo controlled, parallel group, flexible dose study to evaluate the efficacy and safety of Paxil® tablets in children and adolescents with major depressive disorder <Post-marketing clinical study>. GSK - Clinical Study Register (<http://www.gsk-clinicalstudyregister.com>) 2010.
- Simeon 1990 {published data only}**
- Simeon JG, Dinicola VF, Ferguson HB, Copping W. Adolescent depression: a placebo controlled fluoxetine treatment study and follow-up. *Progress in Neuro-Psychopharmacology and Biological Psychiatry* 1990;**14**: 791-5.
- TADS 2004 {published data only}**
- Emslie G, Kratochvil C, Vitiello B, Silva S, Mayes T, McNulty S, et al. The Columbia Classification Team, The TADS Team. Treatment for Adolescents With Depression

- Study (TADS): safety results. *Journal of the American Academy of Child Adolescent Psychiatry* 2006;**45**(12): 1440–55.
- * March J, Silva S, Petrycki S, Curry J, Wells K, Fairbank J, et al. Fluoxetine, cognitive-behavioral therapy, and their combination for adolescents with depression: Treatment for Adolescents with Depression Study (TADS) randomized controlled trial. *JAMA* 2004;**292**:807–20.
- Treatment for Adolescents with Depression Study (TADS) Team. Treatment for Adolescents with Depression Study (TADS): rationale, design, and methods. *Journal of the American Academy of Child and Adolescent Psychiatry* 2003; **42**:531–42.
- Vitiello B, Silva S, Rohde P, Kratochvil C, Kennard B, Reinecke M, et al. Suicidal events in the Treatment for Adolescents with Depression Study (TADS). *Journal of Clinical Psychiatry* 2009;**70**(5):741–7.
- Von Knorring 2006** {published and unpublished data}
Medicines and Healthcare products Regulatory Agency (MHRA). Citalopram Study 2 ID#94404. <http://medicines.mhra.gov.uk> (accessed 20 June 2004).
- * von Knorring A, Olsson GI, Thomson PH, Lemming OM, Hultén A. A randomized, double blind, placebo controlled study of citalopram in adolescents with major depressive disorder. *Journal of Clinical Psychopharmacology* 2006;**26**(3):311–5.
- Wagner 2004** {published data only}
Forest Laboratories. A randomized, double-blind, placebo-controlled evaluation of the safety and efficacy of citalopram in children and adolescent outpatients diagnosed with major depressive disorders (MDD) [CIT-MD-18]. Forest Clinical Trial Registry (<http://www.forestclinicaltrials.com/CTR/CTRController/CTRViewPdf?file=id=scsr/SCSR/CIT-MD-18final.pdf>) 2005 (accessed 24 November 2011).
- Medicines and Healthcare products Regulatory Agency (MHRA). Citalopram Study 1 ID# CIT-MD-18. <http://medicines.mhra.gov.uk> (accessed 20 June 2004).
- Wagner KD, Robb AS, Findling R, Tiseo P. Citalopram is effective in the treatment of major depressive disorder in children and adolescents: results of a placebo-controlled trial. *International Journal of Neuropsychopharmacology* 2002;**5**(Suppl 1):161.
- * Wagner KD, Robb AS, Findling RL, Jin J, Gutierrez MM, Heydorn WE. A randomized, placebo-controlled trial of citalopram for the treatment of major depression in children and adolescents. *American Journal of Psychiatry* 2004;**161**: 1079–83.
- Wagner KD, Robb AS, Findling RL, Jin J, Tiseo PJ. Efficacy of citalopram in the treatment of MDD in children and adolescents [Abstract NR328]. Proceedings of the 155th Annual Meeting of the American Psychiatric Association: New Research Program and Abstracts; 2002 May 18–23; Philadelphia, PA, USA 2002:90.
- Wagner 2006** {published and unpublished data}
Forest Laboratories Ltd. Forest announces results of recently completed Lexapro pediatric depression clinical trial. <http://www.healthplace.com/Communities/Depression/news/children/antidepressants/lexapro%20.asp> OR <http://www.frx.com/> (accessed October 2004).
- * Wagner KD, Jonas J, Findling RL, Ventura D, Saikali K. A double-blind, randomized, placebo-controlled trial of escitalopram in the treatment of pediatric depression. *Journal of the American Academy of Child and Adolescent Psychiatry* 2006;**45**(3):280–8.
- Wagner Trial 1** {published data only}
Pfizer Pharmaceuticals Inc. Study A050100. A multicenter 10-week randomized double-blind placebo-controlled flexible dose outpatient study of sertraline in children and adolescents with major depressive disorder. <http://www.clinicalstudyresults.org/documents/company-study%2098%20.pdf> 2004 (accessed 24 November 2011).
- Wagner Trial 1&2 (2003)** {published data only}
Medicines and Healthcare products Regulatory Agency (MHRA). Sertraline Study 1 ID#1001. <http://medicines.mhra.gov.uk> (accessed 20 June 2004).
- Medicines and Healthcare products Regulatory Agency (MHRA). Sertraline Study 2 ID # 1017. <http://medicines.mhra.gov.uk> (accessed 20 June 2004).
- Wagner KD, Ambrosini P, Rynn M, Wohlberg C, Yang R, Greenbaum MS, et al. Efficacy of sertraline in the treatment of children and adolescents with major depressive disorder. *JAMA* 2003;**290**(8):1033–41.
- Wagner KD, Wohlberg C. Efficacy and safety of sertraline for treatment of pediatric major depressive disorder [Abstract NR327]. 155th Annual Meeting of the American Psychiatric Association; 2002 May 18–23; Philadelphia, PA. 2002.
- Wagner KD, Wohlberg CJ, Yang R. “Efficacy of sertraline in the treatment of children and adolescents with major depressive disorder”: Reply. *JAMA* 2004; Vol. 291, issue 1: 42.
- Wagner Trial 2** {published data only}
Pfizer Pharmaceuticals Inc. Study A0501017A. Multicenter 10-week randomized double-blind placebo-controlled flexible dose outpatient study of sertraline in children and adolescents with major depressive disorder. <http://www.clinicalstudyresults.org/documents/company-study%201981%20.pdf> 2004 (accessed 24 November 2011).
- References to studies excluded from this review**
- Braconnier 2003** {published data only}
Braconnier A, Coent R, Cohen D. Paroxetine versus clomipramine in adolescents with severe major depression: a double blind, randomized, multicentre trial. *Journal of the American Academy of Child and Adolescent Psychiatry* 2003; **42**(1):22–9.
- Cornelius 2009** {published data only}
Cornelius JR, Bukstein OG, Wood DS, Kirisci L, Douaihy A, Clark DB. Double-blind placebo-controlled trial of fluoxetine in adolescents with comorbid major depression and an alcohol use disorder. *Addictive Behaviors* 2009;**34** (10):905–9.
- Cornelius JR, Clark DB, Bukstein OG, Salloum IM, Matta J, Wood DS. Alcohol use but not cannabis use reported

- to contribute to depression in treatment trial of comorbid adolescents. Proceedings of the 68th Annual Scientific Meeting of the College on Problems of Drug Dependence; June 17-22 2006. Scottsdale, Arizona, USA, 2006.
- Cornelius 2010** *{published data only}*
Cornelius JR, Bukstein OG, Douaihy AB, Clark DB, Chung TA, Daley DC, et al. Double-blind fluoxetine trial in comorbid MDD-CUD youth and young adults. *Drug and Alcohol Dependence* 2010;**112**(1-2):39-45.
Cornelius JR, Clark DB, Bukstein OG, Salloum IM, Matta J, Wood DS. Alcohol use but not cannabis use reported to contribute to depression in treatment trial of comorbid adolescents. Proceedings of the 68th Annual Scientific Meeting of the College on Problems of Drug Dependence; June 17-22. Scottsdale, Arizona, USA, 2006.
- Cosgrove 1994** *{published data only}*
Cosgrove PVE. Fluvoxamine in the treatment of depressive illness in children and adolescents. *Journal of Psychopharmacology* 1994;**8**(2):118-23.
- Findling 2009** *{published data only}*
Findling RL, Pagano ME, McNamara NK, Stansbrey RJ, Faber JE, Lingler J, et al. The short-term safety and efficacy of fluoxetine in depressed adolescents with alcohol and cannabis use disorders: a pilot randomized placebo-controlled trial. *Child and Adolescent Psychiatry and Mental Health* 2009;**3**(1):11.
- Mandoki 1997** *{published data only}*
Mandoki MW, Tappia MR, Tapia MA, Sumner GS, Parker JL. Venlafaxine in the treatment of children and adolescents with major depression. *Psychopharmacology Bulletin* 1997;**33**(1):149-54.
- NIMH 2000** *{unpublished data only}*
National Institute of Mental Health. Fluoxetine for the treatment of major depression in youth with bipolar disorder [Study numbers N01-MH-70008A; N01-MH-70008]. <http://www.clinicaltrials.gov/ct2/show/NCT00005015> 2000 (accessed 24 November 2011).
- Riggs 2007** *{published data only}*
Riggs PD, Lohman M, Davies R, Mikulich-Gilbertson S, Laudenslager M, Klein C, et al. Randomized controlled trial of fluoxetine/placebo and cognitive behavioral treatments in depressed adolescents with substance use disorders. Proceedings of the 67th Annual Scientific Meeting of the College on Problems of Drug Dependence; June 19-23 2005. Orlando, Florida, USA, 2005.
Riggs PD, Mikulich-Gilbertson SK, Davies RD, Lohman M, Klein C, Stover SK. A randomized controlled trial of fluoxetine and cognitive behavioral therapy in adolescents with major depression, behavior problems, and substance use disorders. *Archives of Pediatrics and Adolescent Medicine* 2007;**161**(11):1026-34.
Thurstone C, Riggs PD, Orton HD, Libby AM. Fluoxetine for cannabis withdrawal in depressed adolescents. Proceedings of the 67th Annual Scientific Meeting of the College on Problems of Drug Dependence; June 19-23 2005. Orlando, Florida, USA, 2005.
- Sallee 1997** *{published data only}*
Chabrol H, Peresson G. Pulse clomipramine for depressed adolescents [letter]. *American Journal of Psychiatry* 1998;**155**(7):995.
Sallee FR. Dr. Sallee replies [letter]. *American Journal of Psychiatry* 1998;**155**(7):995.
* Sallee FR, Vrindavanam NS, Deas-Nesmith D, Carson SW, Sethuraman G. Pulse intravenous clomipramine for depressed adolescents: double-blind, controlled trial. *American Journal of Psychiatry* 1997;**154**(5):668-73.
- Tashakori 1997** *{published data only}*
Tashakori A, Arabgol F, Panaghi L, Davari R. The effect of reboxetine in the treatment of depression in children and adolescents. *Tehran University Medical Journal* 2007;**65**(8):40-8.
- Wohlfarth 2007** *{published data only}*
Wohlfarth T, Boer F, van den Brink W. Withdrawal of attention rather than pharmacological treatment affects suicide rates in depressed children and adolescents. *American Journal of Psychiatry* 2007;**164**(12):1908-10.

References to ongoing studies

- Duloxetine NCT00849693** *{published data only}*
Eli Lilly. A double-blind, efficacy and safety study of duloxetine versus placebo in the treatment of children and adolescents with major depressive disorder. <http://www.clinicaltrials.gov/ct2/show/NCT00849693> 2009 (accessed 24 November 2011).
- Duloxetine NCT00849901** *{published data only}*
Eli Lilly. A double-blind, efficacy and safety study of duloxetine versus placebo in the treatment of children and adolescents with major depressive disorder. <http://www.clinicaltrials.gov/ct2/show/NCT00849901> 2009 (accessed 24 Nov 2011).
- Glod 2004** *{published data only}*
Glod CA, Lynch A, Berkowitz C, Hennen J, Baldessarini RJ. Bupropion versus citalopram versus placebo in adolescents with major depression [Abstract NR484]. 157th Annual Meeting of the American Psychiatric Association; May 1-6 2004 2004.
- Solvay NCT00353028** *{published data only (unpublished sought but not used)}*
Yamaguchi T. SME3110 (fluvoxamine maleate) in the treatment of depression/depressive state: a post-marketing clinical study in children and adolescents (8 through 18 years of age) - a double-blind, randomized, placebo-controlled study. <http://www.clinicaltrials.gov/ct2/show/NCT00353028> 2010 (accessed 24 November 2011).

Additional references

- AACAP 2004**
American Academy of Child, Adolescent Psychiatry. Supplementary talking points for child and adolescent psychiatrists: regarding the FDA black box warning on the use of antidepressants for pediatric patients. <http://>

- www.aacap.org/press' release/2004/1101.htm 31 October 2004.
- AHRG 2003**
Adolescent Health Research Group. A health profile of New Zealand youth who attend secondary school. *New Zealand Medical Journal* 2003;**116**(1171):3–12.
- Altman 2006**
Altman DG, Royston P. The cost of dichotomising continuous variables. *BMJ* 2006;**332**(7549):1080.
- Andrews 1992**
Andrews JA, Lewinsohn PM. Suicide attempts among older adolescents: prevalence and co-occurrence with psychiatric disorder. *Journal of the American Academy of Child and Adolescent Psychiatry* 1992;**31**(4):655–62.
- Angold 1988**
Angold A. Childhood and adolescent depression II: Research in clinical populations. *British Journal of Psychiatry* 1988;**153**:476–92.
- APA 2000**
American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 4th Edition. Washington, DC: American Psychiatric Association, 2000.
- Asarnow 1992**
Asarnow JR. Suicidal ideation and attempts during middle childhood: associations with perceived family stress and depression among child psychiatric inpatients. *Journal of Clinical Child Psychology* 1992;**21**(1):35–40.
- Axelson 2001**
Axelson DA, Birmaher B. Relation between anxiety and depressive disorders in childhood and adolescence. *Depression and Anxiety* 2001;**14**(2):67–78.
- Beck 1961**
Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J. An inventory for measuring depression. *Archives of General Psychiatry* 1961;**4**:561–71.
- Birmaher 1996**
Birmaher B, Ryan ND, Williamson DE, Brent DA, Kaufman J, Dahl RE, et al. Childhood and adolescent depression: a review of the past 10 years: Part 1. *Journal of the American Academy of Child and Adolescent Psychiatry* 1996;**35**(11):1427–39.
- Brent 1986**
Brent DA, Kalas R, Edelbrock C, Costello AJ, Dulcan MK, Conover N. Psychopathology and its relationship to suicidal ideation in childhood and adolescence. *Journal of the American Academy of Child and Adolescent Psychiatry* 1986;**25**(5):666–73.
- Brent 1993**
Brent DA, Perper JA, Moritz G, Allman C, Friend A, Roth C, et al. Psychiatric risk factors for adolescent suicide: a case control study. *Journal of the American Academy of Child and Adolescent Psychiatry* 1993;**32**(3):521–9.
- Brent 2002**
Brent DA, Birmaher B. Adolescent depression. *New England Journal of Medicine* 2002;**347**(9):667–71.
- Brent 2004**
Brent. Antidepressants and pediatric depression - the risk of doing nothing. *New England Journal of Medicine* 2004;**351**(16):1598–601.
- Bridge 2007**
Bridge JA, Iyengar S, Salary CB, Barbe RP, Birmaher B, Pincus HA, et al. Clinical response and risk for reported suicidal ideation and suicide attempts in pediatric antidepressant treatment: a meta-analysis of randomized controlled trials. *JAMA* 2007;**297**:1683–96.
- Brooks 2001**
Brooks SJ, Kutcher S. Diagnosis and measurement of depression: a review of commonly utilized instruments. *Journal of Child and Adolescent Psychopharmacology* 2001;**11**(4):341–76.
- Carlson 1988**
Carlson GA, Kashani JH. Phenomenology of major depression from childhood through adulthood: analysis of three studies. *American Journal of Psychiatry* 1988;**145**(10):1222–5.
- Castren 2005**
Castren E. Is mood chemistry?. *Nature Reviews Neuroscience* 2005;**6**(3):241–6.
- Chambers 1985**
Chambers WJ, Puig-Antich J, Hirsch M, Paez P, Ambrosini PJ, Tabrizi MA, et al. The assessment of affective disorders in children and adolescents by semi-structured interview: test-retest reliability of the Schedule for Affective Disorders and Schizophrenia for school-age children. Present episode version. *Archives of General Psychiatry* 1985;**42**:696–702.
- Cheung 2005**
Cheung AH, Emslie GJ, Mayes TL. Review of the efficacy and safety of antidepressants in youth depression. *Journal of Child Psychology and Psychiatry* 2005;**46**(7):735–54.
- Churchill 2010**
Churchill R, Caldwell D, McGuire H, Barbui C, Cipriani A, Furukawa TA, et al. Reboxetine versus other antidepressive agents for depression. *Cochrane Database of Systematic Reviews* 2010, Issue 6. [DOI: 10.1002/14651858.CD007852]
- Cipriani 2005**
Cipriani A, Brambilla P, Furukawa T, Geddes J, Gregis M, Hotopf M, et al. Fluoxetine versus other types of pharmacotherapy for depression. *Cochrane Database of Systematic Reviews* 2005, Issue 5. [DOI: 10.1002/14651858.CD004185.pub2]
- Cipriani 2009**
Cipriani A, Furukawa TA, Salanti G, Geddes JR, Higgins JPT, Churchill R, et al. Comparative efficacy and acceptability of 12 new-generation antidepressants: a multiple-treatments meta-analysis. *The Lancet* 2009;**373**:746–58.
- Cipriani 2009a**
Cipriani A, Furukawa TA, Veronese A, Watanabe N, Churchill R, McGuire H, et al. Paroxetine versus other

- anti-depressive agents for depression. *Cochrane Database of Systematic Reviews* 2009, Issue 3. [DOI: 10.1002/14651858.CD006531]
- Cipriani 2009b**
Cipriani A, Signoretti A, Furukawa TA, Churchill R, Tomelleri S, Omori IM, et al. Venlafaxine versus other anti-depressive agents for depression. *Cochrane Database of Systematic Reviews* 2009, Issue 3. [DOI: 10.1002/14651858.CD006531]
- Cipriani 2009c**
Cipriani A, Santilli C, Furukawa TA, Signoretti A, Nakagawa A, McGuire H, et al. Escitalopram versus other antidepressive agents for depression. *Cochrane Database of Systematic Reviews* 2009, Issue 2. [DOI: 10.1002/14651858.CD006532.pub2]
- Cipriani 2010**
Cipriani A, La Ferla T, Furukawa TA, Signoretti A, Nakagawa A, Churchill R, et al. Sertraline versus other antidepressive agents for depression. *Cochrane Database of Systematic Reviews* 2010, Issue 4. [DOI: 10.1002/14651858.CD006117.pub4]
- Cohen 2004**
Cohen D, Gerardin P, Mazet P, Purper-Ouakil D, Flament MF. Pharmacological treatment of adolescent major depression. *Journal of Child and Adolescent Psychopharmacology* 2004;**14**(1):19–31.
- Cooper 1988**
Cooper GL. The safety of fluoxetine - an update. *British Journal of Psychiatry. Supplement* 1988;**3**:77–86.
- Costello 2006**
Costello JE, Erkanli A, Angold A. Is there an epidemic of child or adolescent depression. *Journal of Child Psychology and Psychiatry* 2006;**47**(12):1263–71.
- CSM 2004**
Committee on Safety of Medicines (CSM). Report of the CSM expert working group on the safety of selective serotonin reuptake inhibitors antidepressants. <http://medicines.mhra.gov.uk/ourwork/monitorsafequalmed/safetymessages/SSRIfinal.pdf> 2004.
- Cuijpers 2005**
Cuijpers P, Dekker J. Psychological treatment of depression; a systematic review of meta-analyses [Psychologische behandeling van depressie; een systematisch overzicht van meta-analysen]. *Nederlands Tijdschrift Voor Geneeskunde* 2005;**149**:1892–7.
- De Leo 2002**
De Leo D. Why are we not getting any closer to preventing suicide?. *British Journal of Psychiatry* 2002;**181**:372–4.
- DerSimonian 1986**
DerSimonian R, Laird N. Meta-analysis in clinical trials. *Controlled Clinical Trials* 1986;**7**:177–88.
- Dubicka 2006**
Dubicka B, Hadley S, Roberts C. Suicidal behaviour in youths diagnosed with depression treated with new-generation antidepressants. *British Journal of Psychiatry* 2006;**189**:393–98.
- Dubitsky 2004**
Dubitsky GM. Review and Evaluation of Clinical Data: Placebo-Controlled Antidepressant Studies In Pediatric Patients. Medicines and Healthcare Products Regulatory Agency 6 August 2004.
- Dunlop 2010**
Dunlop BW, Vaughan CL. Survey of investigators opinions on the acceptability of interactions with patients participating in clinical trials. *Journal of Clinical Psychopharmacology* 2010;**30**(3):323–7.
- Ebmeier 2006**
Ebmeier KP, Donaghey C, Steele JD. Recent developments and current controversies in depression. *Lancet* 2006;**367**:153–67.
- Elbourne 2002**
Elbourne DR, Altman DG, Higgins JPT, Curtin F, Worthington HV, Vail A. Meta-analyses involving cross-over trials: methodological issues. *International Journal of Epidemiology* 2002;**31**(1):140–9.
- EMA 2005**
European Medicines Agency. European Medicines Agency finalises review of antidepressants in children and adolescents. <http://www.emea.eu.int/pdfs/human/press/pr/12891805en.pdf> 25 April 2005.
- Esposito 2002**
Esposito CL, Clum GA. Psychiatric symptoms and their relationship to suicidal ideation in a high risk adolescent community sample. *Journal of the American Academy of Child and Adolescent Psychiatry* 2002;**41**(1):44–51.
- Essau 1999**
Essau CF, Petermann F, Reynolds WM. In: *Essau FPCA (Ed). Depressive Disorders in Children and Adolescents: Epidemiology, Risk Factors and Treatment*. USA: Jason Aronson Inc, 1999:326.
- FDA 2004**
FDA Public Health Advisory. Suicidality in children and adolescents being treated with antidepressant medications. <http://www.fda.gov/cder/drug/antidepressants/SSRIPHA200410.htm> 15 October 2004.
- Findling 2004**
Findling R, Greenhill L. Message to AACAP members. President of the American Academy of Child and Adolescent Psychiatrists 2004.
- Fleming 1993**
Fleming JE, Boyle MH, Offord DR. The outcome of adolescent depression in the Ontario Child Health Study Follow-up. *Journal of the American Academy of Child and Adolescent Psychiatry* 1993;**32**(1):28–33.
- Fombonne 2001a**
Fombonne E, Wostear G, Cooper V, Harrington R, Rutter M. The Maudsley long-term follow-up of child and adolescent depression: I. Psychiatric outcomes in adulthood. *British Journal of Psychiatry* 2001;**179**:210–7.

Fombonne 2001b

Fombonne E, Wostear G, Cooper V, Harrington R, Rutter M. The Maudsley long-term follow-up of child and adolescent depression. 2. Suicidality, criminality and social dysfunction in adulthood. *British Journal of Psychiatry* 2001;**179**:218–23.

Friedman 2007

Friedman RA, Leon AC. Expanding the black box - depression, antidepressants, and the risk of suicide. *New England Journal of Medicine* 7 May 2007 [Epub ahead of print].

Garrison 1997

Garrison CZ, Waller JL, Cuffe SP, McKeown RE, Addy CL, Jackson KL. Incidence of major depressive disorder and dysthymia in young adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry* 1997;**36**(4): 458–65.

Gibbons 2007

Gibbons RD, Brown CH, Hur K, Marcus SM, Bhaumik DK, Erkens JA, et al. Early evidence on the effects of regulators' suicidality warnings on SSRI prescriptions and suicide in children and adolescents. *American Journal of Psychiatry* 2007;**164**(9):1356–63.

Gloaguen 1998

Gloaguen V, Cottraux J, Cucherat M, Blackburn I-M. A meta-analysis of the effects of cognitive therapy in depressed patients. *Journal of Affective Disorders* 1998;**49**:59–72.

Goodyer 2010

Goodyer IM, Wilkinson P, Dubicka B, Kelvin R. Forum: The use of selective serotonin reuptake inhibitors in depressed children and adolescents: commentary on the meta-analysis of Hetrick et al. *Current Opinion in Psychiatry* 2010;**23**:58–61.

Grunbaum 2002

Grunbaum J, Kann L, Kinchen SA, Williams B, Ross JG, Lowry R, et al. Morbidity and mortality weekly report. Centres for Disease Control surveillance summaries. Youth Risk Behaviour Surveillance. Centres for Disease Control, 2002.

Guaiana 2010

Guaiana G, Gupta S, Chiodo D, Davies SJC. Agomelatine versus other antidepressive agents for major depression. *Cochrane Database of Systematic Reviews* 2010, Issue 11. [DOI: 10.1002/14651858.CD008851]

Gunnell 2004

Gunnell D, Ashby D. Antidepressants and suicide: what is the balance of benefit and harm. *BMJ* 2004;**329**(7456): 34–8.

Guy 1976

Guy W. *ECDEU Assessment Manual for Psychopharmacology*. 2nd Edition. Washington, DC: US Government Printing Office, 1976:76–388.

Hall 2006

Hall WD. How have the SSRI antidepressants affected suicide risk?. *The Lancet* 2006;**367**:1959–62.

Hamilton 1960

Hamilton M. A rating scale for depression. *Journal of Neurology, Neurosurgery and Psychiatry* 1960;**23**:56–62.

Hammad 2004

Hammad TA. Relationship between psychotropic drugs and pediatric suicidality. US Food and Drug Administration (FDA) 16 August 2004.

Hammad 2006

Hammad TA, Laugren T, Racoosin J. Suicidality in pediatric patients treated with antidepressant drugs. *Archives of General Psychiatry* 2006;**63**:332–9.

Harbord 2006

Harbord RM, Egger M, Sterne JA. A modified test for small-study effects in meta-analyses of controlled trials with binary endpoints. *Statistics in Medicine* 2006;**25**(20): 3443–57.

Harbord 2009

Harbord RM, Harris RJ, Sterne JAC. Updated tests for small-study effects in meta-analyses. *Stata Journal* 2009;**9**(2):197–210.

Harrington 1990

Harrington R, Fudge H, Rutter M, Pickles A, Hill J. Adult outcomes of childhood and adolescent depression. I. Psychiatric status. *Archives of General Psychiatry* 1990;**47**(5): 465–73.

Harrington 2001

Harrington R, Dubicka B. Natural history of mood disorders in children and adolescents. In: Goodyear IM editor(s). *The Depressed Child and Adolescent. Cambridge Child and Adolescent Psychiatry series*. Cambridge: Cambridge University Press, 2001:311–43.

Hazell 2002

Hazell P, O'Connell D, Heathcote D, Henry D. Tricyclic drugs for depression in children and adolescents. *Cochrane Database of Systematic Reviews* 2002, Issue 2. [DOI: 10.1002/14651858.CD002317]

Healy 1997

Healy D. *The Anti-Depressant Era*. Massachusetts: Harvard University Press, 1997.

Healy 2003

Healy D. Lines of evidence on the risks of suicide with selective serotonin reuptake inhibitors. *Psychotherapy and Psychosomatics* 2003;**72**(2):71–9.

Healy 2009

Healy D. Are selective serotonin reuptake inhibitors a risk factor for adolescent suicide?. *Canadian Journal of Psychiatry* 2009;**54**(2):69–71.

Hetrick 2007

Hetrick SE, Merry SN, McKenzie JE, Proctor M, Simmons M. Selective serotonin reuptake inhibitors (SSRIs) for depressive disorders in children and adolescents. *Cochrane Database of Systematic Reviews* 2007, Issue 3. [DOI: 10.1002/14651858.CD004851.pub2]

Hetrick 2010

Hetrick SE, McKenzie JE, Merry SN. The use of SSRIs in children and adolescents. *Current Opinion in Psychiatry* 2010;**23**:53–7.

Higgins 2003

Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;**327** (7414):557–60.

Higgins 2008

Higgins JP, White IR, Wood AM. Imputation methods for missing outcome data in meta-analysis of clinical trials. *Clinical Trials* 2008;**5**(3):225–39.

Higgins 2009

Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.0.2 [updated September 2009]. The Cochrane Collaboration, 2009. Available from www.cochrane-handbook.org.

Hollis 1999

Hollis S, Campbell F. What is meant by intention to treat analysis? Survey of published randomised controlled trials. *BMJ* 1999;**319**:670–4.

Imperadore 2009

Imperadore G, Cipriani A, Signoretti A, Furukawa TA, Watanabe N, Churchill R, et al. Citalopram versus other anti-depressive agents for depression. *Cochrane Database of Systematic Reviews* 2009, Issue 3. [DOI: 10.1002/14651858.CD006534]

Ioannidis 2008

Ioannidis JP. Effectiveness of antidepressants: an evidence myth constructed from a thousand randomized trials?. *Philosophy Ethics and Humanities in Medicine* 2008;**3**:14.

Isacsson 2010

Isacsson G, Rich CL, Jureidini J, Raven M. The increased use of antidepressants has contributed to the worldwide reduction in suicide rates. *British Journal of Psychiatry* 2010;**196**:429–33.

Jick 2004

Jick H, Kaye JA, Jick SS. Antidepressants and the risk of suicidal behaviours. *JAMA* 2004;**292**(3):338–43.

Julious 2005

Julious SA. Issues with number needed to treat. *Statistics in Medicine* 2005;**24**(20):3233–5.

Jureidini 2003

Jureidini J, Tonkin A. Paroxetine in major depression (letter). *Journal of the American Academy of Child and Adolescent Psychiatry* 2003;**42**(5):514.

Jureidini 2004

Jureidini JN, Doecke CJ, Mansfield PR, Haby MM, Menkes DB, Tonkin AL. Efficacy and safety of antidepressants for children and adolescents. *BMJ* 2004;**328**:879–83.

Kaufman 2001

Kaufman J, Martin A, King RA, Charney D. Are child-, adolescent-, and adult-onset depression one and the same disorder?. *Biological Psychiatry* 2001;**49**:980–1001.

Keller 2003

Keller MB, Ryan ND, Strober M, Weller EB, McCafferty JP, Hagino OR, et al. Paroxetine in major depression [reply to Jureidini]. *Journal of the American Academy of Child and Adolescent Psychiatry* 2003; Vol. 42, issue 5:514–5.

Kieser 2004

Kieser M, Rohmel J, Friede T. Power and sample size determination when assessing the clinical relevance of trial results by 'responder analyses'. *Statistics in Medicine* 2004;**23**(21):3287–305.

Kirsch 2008

Kirsch I, Deacon BJ, Huedo-Medina TB, Scoboria A, Moore TJ, Johnson BT. Initial severity and antidepressant benefits: a meta-analysis of data submitted to the Food and Drug Administration. *PLoS Medicine* 2008;**5**(2):e45.

Kovacs 1984

Kovacs M, Feinberg TL, Crouse-Novak MA, Paulauskas SL, Finkelstein R. Depressive disorders in childhood I: A longitudinal prospective study of characteristics and recovery. *Archives of General Psychiatry* 1984;**41**(3):229–37.

Kovacs 1989

Kovacs M, Gatsonis C, Paulauskas SL, Richards C. Depressive disorders in childhood: IV. A longitudinal study of comorbidity with and risk for anxiety disorders. *Archives of General Psychiatry* 1989;**46**(9):776–82.

Kovacs 1992

Kovacs M. *The Children's Depression Inventory*. North Tonawanda, New York: Mental Health Systems, 1992.

Kovacs 1993

Kovacs M, Goldston D, Gatsonis C. Suicidal behaviours in child onset depressive disorders: a longitudinal investigation. *Journal of the American Academy of Child and Adolescent Psychiatry* 1993;**32**(1):8–20.

Lenox 2008

Lenox RH, Frazer A. Mechanism of action of antidepressants and mood stabilizers. *Neuropsychopharmacology: The Fifth Generation of Progress*. American College of Neuropsychopharmacology: <http://www.acnp.org/publications/neuro5thgeneration.aspx>, 2008:1139–63.

Lewinsohn 1996

Lewinsohn PM, Rohde P, Seeley JR. Adolescent suicidal ideation and attempts: prevalence, risk factors and clinical implications. *Clinical Psychology: Science and Practice* 1996;**3**:25–46.

Lewinsohn 1998

Lewinsohn PM, Rohde P, Seeley JR. Major depressive disorder in older adolescents: prevalence, risk factors and clinical implications. *Clinical Psychology Review* 1998;**18** (7):765–94.

Lexchin 2003

Lexchin J, Bero LA, Djulbegovic B, Clark O. Pharmaceutical industry sponsorship and research outcome and quality: systematic review. *BMJ* 2003;**326**(7400):1167–70.

Libby 2007

Libby AM, Brent DA, Morrato EH, Orton HD, Allen R, Valuck RJ. Decline in treatment of pediatric depression after FDA advisory on risk of suicidality with SSRIs. *American Journal of Human Genetics* 2007;**164**(6):884–91.

Marttunen 1998

Marttunen M, Pelkonen M. Phenomenology of adolescent depression. *Psychiatria Fennica* 1998;**29**:29–39.

Masi 1998

Masi G, Favilla L, Mucci M. Depressive disorder in children and adolescents. *European Journal of Paediatric Neurology* 1998;**2**:287–95.

McDermott 2011

McDermott B, Baigent M, Chanen A, Fraser L, Graetz B, Hayman N, et al. *beyondblue Expert Working Committee (2010) Clinical practice guidelines: Depression in adolescents and young adults*. Melbourne: beyondblue: the National Depression Initiative, 2011.

Moher 2010

Moher D, Hopewell S, Schulz KF, Montori V, Gotzsche PC, Devereaux PJ, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *Journal of Clinical Epidemiology* 2010;**63**(8):e1–37.

Moncrieff 2001

Moncrieff J, Churchill R, Drummond C, McGuire H. Development of a quality assessment instrument for trials of treatments for depression and neurosis. *International Journal of Methods in Psychiatric Research* 2001;**10**(3):126–33.

Moncrieff 2004

Moncrieff J, Wessely S, Hardy R. Active placebo versus antidepressants for depression. *Cochrane Database of Systematic Reviews* 2005, Issue 1. [DOI: 10.1002/14651858.CD003012.pub2]

Montgomery 1979

Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. *British Journal of Psychiatry* 1979;**134**:382–9.

Murray 1996

Murray CJ, Lopez AD. *The Global Burden of Disease. A comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020*. Harvard School of Public Health. Cambridge, MA: Harvard University Press, 1996.

Myers 2002

Myers K, Winters NC. Ten-year review of rating scales: II: scales for internalizing disorders. *Journal of the American Academy of Child and Adolescent Psychiatry* 2002;**41**(6): 634–59.

Nakagawa 2009

Nakagawa A, Watanabe N, Omori IM, Barbui C, Cipriani A, McGuire H, et al. Milnacipran versus other antidepressive agents for depression. *Cochrane Database of Systematic Reviews* 2009, Issue 3. [DOI: 10.1002/14651858.CD006529.pub2]

Newman 2004

Newman TB. A black-box warning for antidepressants in children?. *New England Journal of Medicine* 2004;**351**(16): 1595–9.

NHMRC 1997

National Health and Medical Research Council. *Depression in Young People: Clinical Practice Guidelines*. Canberra: Australian Publishing Services, 1997.

NICE 2005

National Institute for Health and Clinical Excellence (NICE). *Depression in Children and Young People: Identification and management in primary, community and secondary care*. National Clinical Practice Guideline 28. Leicester: The British Psychological Society, 2005.

Nosè 2009

Nosè M, Cipriani A, Furukawa TA, Omori IM, Churchill R, McGuire H, et al. Duloxetine versus other antidepressive agents for depression. *Cochrane Database of Systematic Reviews* 2009, Issue 3. [DOI: 10.1002/14651858.CD006533]

Olson 2003

Olson M, Shaffer D, Marcus SC, Greenberg T. Relationship between antidepressant medication treatment and suicide in adolescents. *Archives of General Psychiatry* 2003;**60**:978–82.

Olson 2006

Olson M, Marcus SC, Shaffer D. Antidepressant drug therapy and suicide in severely depressed children and adults: a case-control study. *Archives of General Psychiatry* 2006;**63**(8):865–72.

Olver 2001

Olver JS, Burrows GD, Norman TR. Third-generation antidepressants - do they offer advantages over the SSRIs?. *CNS Drugs* 2001;**15**(12):941–54.

Omori 2010

Omori I, Watanabe N, Nakagawa A, Cipriani A, Barbui C, McGuire C, et al. Fluvoxamine versus other antidepressive agents for depression. *Cochrane Database of Systematic Reviews* 2010, Issue 3. [DOI: 10.1002/14651858.CD006114.pub2]

Orsini 2006

Orsini N, Bottai M, Higgins J, Buchan I. HETEROGI: Stata module to quantify heterogeneity in a meta-analysis, Statistical Software Components. Boston College Department of Economics, <http://econpapers.repec.org/RePEc:boc:bocode:s449201>, 2006.

Peters 2008

Peters JL, Sutton AJ, Jones DR, Abrams KR, Rushton L. Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. *Journal of Clinical Epidemiology* 2008;**61**(10):991–6.

Petersen 1993

Petersen AC, Compas BE, Brooks-Gunn J, Stemmler M, Ey S, Grant KE. Depression in adolescence. *American Psychologist* 1993;**48**(2):155–68.

Petti 1985

Petti T. Scales of potential use in the psychopharmacological treatment of depressed children and adolescents. *Psychopharmacology Bulletin* 1985;**21**(4):951–6.

Pine 1998

Pine DS, Cohen P, Gurley D, Brook J, Ma Y. The risk for early-adulthood anxiety and depressive disorders in adolescents with anxiety and depressive disorders. *Archives of General Psychiatry* 1998;**55**:56–64.

Poznanski 1984

Poznanski EO, Grossman JA, Buchsbaum Y, Banegas M, Freeman L, Gibbons R. Preliminary studies of the reliability and validity of the children's depression rating scale. *Journal of the American Academy of Child and Adolescent Psychiatry* 1984;**23**:191–7.

Poznanski 1996

Poznanski EO, Mokros HB. *Children's Depression Rating Scale Revised*. Los Angeles, CA: WPS, 1996.

Rao 1995

Rao U, Ryan ND, Birmaher B, Dahl RE, Williamson DE, Kaufman J, et al. Unipolar depression in adolescence: clinical outcome in adulthood. *Journal of the American Academy of Child and Adolescent Psychiatry* 1995;**34**(5): 566–78.

Reseland 2008

Reseland S, Le Noury J, Aldred G, Healy D. Contributions to national suicide rates: autopsies and antidepressants. *Journal of Psychotherapy and Psychosomatics* 2008;**77**:78–82.

Reynolds 1986

Reynolds WM. *Reynolds Adolescent Depression Scale*. Odessa, FL: Psychological Assessment Resources, 1986.

Reynolds 1987

Reynolds WM. *Suicidal Ideation Questionnaire*. Odessa, FL: Psychological Assessment Resources, 1987.

Richmond 2005

Richmond TK, Rosen DS. The treatment of adolescent depression in the era of the black box warning. *Current Opinion in Pediatrics* 2005;**17**(4):466–72.

Rivas-Vasquez 2004

Rivas-Vazquez RA, Saffa-Biller D, Ruiz I, Blais MA, Rivas-Vaquez A. Current issues in anxiety and depression: comorbid, mixed, and subthreshold disorders. *Professional Psychology: Research and Practice* 2004;**35**(1):74–83.

Safer 2007

Safer DJ, Zito JM. Do antidepressants reduce suicide rates?. *Public Health* 2007;**121**(4):274–7.

Sawyer 2001

Sawyer MG, Arney FM, Baghurst PA, Clark JJ, Graetz RJ, Kosky R, et al. The mental health of young people in Australia: key findings from the child and adolescent component of the national survey of mental health and well-being. *Australian and New Zealand Journal of Psychiatry* 2001;**35**:806–14.

Schorre 2004

Schorre BEH, Vandvik IH. Global assessment of functioning in child and adolescent psychiatry. *European Child and Adolescent Psychiatry* 2004;**13**(5):273–86.

Schott 2010

Schott G, Pacht H, Limbach U, Gundert-Remy U, Lieb K, Ludwig W. The financing of drug trials by pharmaceutical companies and its consequences. *Deutsches Ärzteblatt International* 2010;**107**(17):295–301.

Shaffer 1985

Shaffer D, Gould MS, Brasic J, Ambrosini P, Fisher P, Bird H, et al. A children's global assessment scale (CGAS). *Archives of General Psychiatry* 1985;**40**:1228–31.

Shaffer 1996

Shaffer D, Gould MS, Fisher P, Trautman P, Moreau D, Kleinman MM, et al. Psychiatric diagnosis in child and adolescent suicide. *Archives of General Psychiatry* 1996;**53**(4):338–48.

Sherer 2007

Scherer RW, Langenberg P, von Elm E. Full publication of results initially presented in abstracts. *Cochrane Database of Systematic Reviews* 2007, Issue 2. [DOI: 10.1002/14651858.MR000005.pub3]

Sigafoos 1988

Sigafoos AD, Feinstein CB, Damond M, Reiss D. The measurement of behavioral autonomy in adolescence: the autonomous functioning checklist. *Adolescent Psychiatry* 1988;**15**:432–62.

Simon 2006

Simon GE. How can we know whether antidepressants increase suicide risk. *American Journal of Psychiatry* 2006;**163**(11):1861–3.

Simon 2006b

Simon GE, Savarino J, Operskalski B, Wang PS. Suicide risk during antidepressant treatment. *American Journal of Psychiatry* 2006;**163**(1):41–7.

Sismondo 2008

Sismondo S. Pharmaceutical company funding and its consequences: a qualitative systematic review. *Contemporary Clinical Trials* 2008;**29**:109–13.

Sondergard 2006

Sondergard L, Kvist K, Andersen PL, Kessing LV. Do antidepressants precipitate youth suicide? A nationwide pharmacoepidemiological study. *European Child and Adolescent Psychiatry* 2006;**15**:232–40.

Sterne 2009

Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ* 2009;**338**:b2393.

Sterne 2011

Sterne JA, Sutton AJ, Ioannidis JP, Terrin N, Jones DR, Lau J, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ* 2011;**343**:d4002.

Stone 2009

Stone M, Laughren T, Jones ML, Levenson M, Holland PC, Hughes A, et al. Risk of suicidality in clinical trials of antidepressants in adults: analysis of proprietary data submitted to US Food and Drug Administration. *BMJ* 2009;**339**:b2880.

Tsapakis 2008

Tsapakis EM, Soldani F, Tondo L, Baldessarini RJ. Efficacy of antidepressants in juvenile depression: meta-analysis. *British Journal of Psychiatry* 2008;**193**:10–7.

Turner 2008

Turner EH, Matthews AM, Linardatos E, Tell RA, Rosenthal R. Selective publication of antidepressant trials and its influence on apparent efficacy. *New England Journal of Medicine* 2008;**358**:252–60.

Valuck 2004

Valuck RJ, Libby AM, Sills MR, Giese AA, Allen RR. Antidepressant treatment and the risk of suicide attempt by adolescents with major depressive disorder. *CNS Drugs* 2004;**18**(15):1119–32.

van Praag 1987

van Praag HM, Kahn RS, Asnis GM, Wetzler S, Brown SL, Bleich A, et al. Denosologization of biological psychiatry or the specificity of 5-HT disturbances in psychiatric disorders. *Journal of Affective Disorders* 1987;**13**:1–8.

Vitiello 2006

Vitiello B, Zukeas SH, Norquist GS. National estimates of antidepressant medication use among US children. *Journal of the American Academy of Child and Adolescent Psychiatry* 2006;**45**(3):271–9.

Wagner 2005

Wagner KD. Pharmacotherapy for major depression in children and adolescents. *Progress in Neuro-psychopharmacology and Biological Psychiatry* 2005;**29**: 819–26.

Watanabe 2011

Watanabe N, Omori IM, Nakagawa A, Cipriani A, Barbui C, Churchill R, et al. Mirtazapine versus other antidepressive agents for depression. *Cochrane Database*

of Systematic Reviews 2011, Issue 12. [DOI: 10.1002/14651858.CD006528.pub2]

Weissman 1999

Weissman MM, Wolk S, Goldstein RB, Moreau D, Adams P, Greenwald S, et al. Depressed adolescents grown up. *JAMA* 1999;**281**:1707–13.

Weisz 2006

Weisz JR, McCarty CA, Valeri SM. Effects of psychotherapy for depression in children and adolescents: a meta-analysis. *Psychological Bulletin* 2006;**132**:132–49.

Weller 2000

Weller EB, Weller RA. Treatment options in the management of adolescent depression. *Journal of Affective Disorders* 2000;**61**(Suppl):23–8.

Wetzler 1996

Wetzler S, Asnis GM, Hyman RB, Virtue C, Zimmerman J, Rathus JH. Characteristic of suicidality among adolescents. *Suicide and Life-Threatening Behaviour* 1996;**26**(1):37–45.

Whittington 2004

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**(9418):1341–5.

WHO 2004

World Health Organization. *ICD-10: International statistical classification of diseases and related health problems*. 2nd Edition. Vol. 3, Geneva, Switzerland: World Health Organization, 2004.

Winters 2005

Winters NC, Collett BR, Myers KM. Ten-year review of rating scales, VII: scales assessing functional impairment. *Journal of the American Academy of Child and Adolescent Psychiatry* 2005;**44**(4):309–38.

Wisloff 2011

Wisloff T, Aalen OO, Kristiansen IS. Considerable variation in NNT - a study based on Monte Carlo simulations. *Journal of Clinical Epidemiology* 2011;**64**(4):444–50.

* Indicates the major publication for the study

CHARACTERISTICS OF STUDIES

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

Almeida-Montes 2005

Methods	Trial design: randomised controlled trial; single site Power calculation: 0.90 power to detect a large effect size (0.80) Use of diagnostic criteria (or clear specification of inclusion criteria): DSM-IV-TR criteria for depressive disorder plus a score of 13 in the DSDR Intervention integrity: not described Outcome measures described or validated measures used: yes Follow-up assessment points: weekly for 7 weeks No. crossed: none Funded by: Eli Lilly provided fluoxetine and placebo
Participants	Setting of care: outpatient Recruitment: no statement Mean age (SD): intervention = 13.3 (3.16); control = 11.5 (1.58) Age range: 8 to 14 Gender (F:M): intervention = not stated; control = not stated Methods used to diagnose: DSM-IV using semi-structured interview; The Mini International Neuropsychiatric Interview for children and adolescents (MINI-KID) Diagnosis: MDD Baseline severity of depression: not reported Length of current episode: not reported % first episode: not reported Co-morbidity (intervention): not reported Co-morbidity (control): not reported Location: Mexico Inclusion criteria: major depressive disorder (DSM-IV-TR) plus a score of 13 in the DSDR Exclusion criteria: History of chronic physical illness, intellectual disability, bipolar disorder, substance use or dependence, Attention Deficit Hyperactivity Disorder, severe anxiety, behavioural disorder, hospital admission or increased treatment intensity due to a depressive episode in the preceding 4 weeks, antidepressive treatment in the preceding 4 weeks, any lab test which was considered to be abnormal by the clinician, oppositional defiant disorder Exclusion of suicidality: suicide attempt in the preceding 4 weeks
Interventions	Intervention group Drug: fluoxetine Dosage: 20 mg Regimen: daily Length of treatment: 6 weeks Control group: placebo
Outcomes	Definition and assessment of response: we used OC response data defined as 50% reduction in HDRS scores (they stated they used CGI-I score of 1 or 2; 50% reduction in DSRs and HAM-D scores

	Depressive symptoms: DSRS , HAM-D Functioning: Children's Global Assessment Scale (C-GAS) Suicidal behaviour: no report Other measures: Hamilton Anxiety Rating Scale HARS	
Notes	Type of data used for remission/response: observed case	
<i>Risk of bias</i>		
Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computer-generated random numbers using SPSS, pg 34
Allocation concealment (selection bias)	Low risk	An independent clinician who was not part of the trial allocated the 2 treatment conditions to either '0' or '1'. The trial researchers remained blind to treatment allocation throughout the course of the trial, pg 34
Blinding (performance bias and detection bias) Intervention blinded	Unclear risk	No statement
Blinding (performance bias and detection bias) Blinded outcome	Low risk	2 clinicians who remained blind to treatment allocation assessed the participants weekly, pg 34
Incomplete outcome data (attrition bias) All outcomes	High risk	Number eligible: 38 Number randomised: 23 Fluoxetine: 12 (inconsistencies in reporting noted); placebo: 11; total: 23 Number started trial: 23 Fluoxetine: 12 (inconsistencies); placebo: 11; total: 23 Number of withdrawals: Fluoxetine: 7; placebo: 9; total: 16 Number analysed post intervention: Fluoxetine: ITT = 7, LOCF = 10; placebo: ITT = 9, LOCF = 10; total: ITT = 16, LOCF = 20 Reasons for drop out: fluoxetine group lost to follow-up N = 5, withdrawn due to suicide risk N = 1, did not complete N = 7; placebo lost to follow-up N = 2, withdrawn due to suicide risk N = 0, did not complete N = 9

		ITT analysis: additionally we analysed outcomes using ITT analysis in the following way: we divided the number of patients who completed the trial and were considered to be 'responders' by the total sample pg 34; ITT population does not include all who were randomised Statistical analysis: LOCF
Selective reporting (reporting bias)	Unclear risk	All outcomes specified in methods were reported, however 2 outcomes (adverse and clinician-reported depression symptoms) were reported in a graph. No access to trial protocol
Other bias	Unclear risk	Contact: assessment undertaken weekly Screening: unclear Placebo lead-in: 1 week Baseline imbalance: data not reported

Berard 2006

Methods	Trial design: randomised controlled trial; multicentre Power calculation: yes Use of diagnostic criteria (or clear specification of inclusion criteria): yes Intervention integrity: yes - returned pill pack. "Every effort was to be made to encourage patient compliance with the dosage regimen as per protocol. All patients were instructed to return their medication pack, with any unused drug, to the investigator at their next visit. A record of the supplies dispensed, taken and returned was made in the Case Report Form (CRF) at each visit". Section 3.6, Final report Outcome measures described or validated measures used: yes Follow-up assessment points: post intervention No. crossed over: none Funded by: SmithKline Beecham
Participants	Setting of care: not stated Recruitment: no information Mean age (SD): intervention = 15.5 (SD 1.6); control = 15.8 (SD 1.6) Age range: 13 to 19 years Gender (F:M): intervention = 122:65; control = 61:38 Methods used to diagnose: DSM-IV; GAS < 69; Montgomery-Asberg Depression Rating Scale (MADRS) $\geq$ 16; after screening 14-day single-blind run-in period Diagnosis: MDD Baseline severity of depression: MADRS mean (SE) score: intervention = 25.9 (0.5); control: 25.9 (0.6) (both groups moderately to severely ill); CGI Intervention 4.2 (0.1); CGI Placebo 4.2 (0.1) Length of current episode: not reported % first episode: Intervention 70.9%; Placebo 68.8%

	Co-morbidity (intervention): specific phobia 6; separation anxiety 5; panic disorder 3; social phobia 3; Generalised Anxiety Disorder (GAD) 13; Post Traumatic Stress Disorder (PTSD) 1; Attention Deficit Hyperactivity Disorder (ADHA) 3; Oppositional Deficient Disorder (ODD) 1; Anorexia Nervosa (AN) 1; Bulimia Nervosa (BN)2; substance abuse 0 Co-morbidity (control): specific phobia 3; separation anxiety; panic disorder 0; social phobia 4; GAD 4; PTSD 3; ADHD 0; ODD 1; AN 0; BN 0; substance abuse 1 Location: Argentina, Belgium, Canada, Holland, Italy, Mexico, South Afric, Spain, United Arab Emirates, United Kingdom Inclusion criteria: unipolar MDD for at least 8 weeks' duration; negative pregnancy test Exclusion criteria: prepubertal; diagnosis of Conduct Disorder, autism, Pervasive Developmental Disorder, organic psychiatric disorder including schizophrenia and epilepsy; Obsessive Compulsive Disorder, panic disorder, social phobia, PTSD that preceded MDD; medical illness that contraindicated use of paroxetine; previous response to psychotherapy; planned long-term psychotherapy; Electroconvulsive Therapy (ECT) in previous 3 months or planned for trial period; drug or alcohol dependency; concomitant psychotropic medication or other drugs interfering with Central Nervous System (CNS) activity; use of sumatriptan, oral anticoagulants or type 1C antiarrhythmics, i. e. encainide, flecainide, lorcinide and propafenone; previous use of paroxetine or other SSRI; sensitivity to SSRI; sexually active and not using contraceptive or pregnant or lactating; use of other investigational drug Exclusion of suicidality: although a history of suicide attempt(s) was not exclusionary, patients with current serious suicidal ideation were excluded
Interventions	Intervention group Drug: paroxetine Dosage: 20 to 40 mg Regimen: daily Length of treatment: 12 weeks Control group: placebo pill
Outcomes	Definition and assessment of response: we used OC response defined as $\geq 50\%$ MADRS (they used responders defined as at least a 50% reduction on MADRS. Post hoc analysis on responder rate based on a CGI-I score of 1 or 2 was also conducted). Depressive symptoms: change from baseline in the K-SADS-L depression sub-scale score Functioning: Children's Global Assessment Scale (C-GAS) Suicidal behaviour: FDA data; no report of continuous measure Other measures: Montgomery-Asberg Depression Rating Scale (MADRS); K-SADS-L; Clinical Global Impressions Scale Improvement (CGI - Improvement); Mood and Feeling Questionnaire; adverse events
Notes	Additional data were sought and received from the authors MHRA # 377 MHRA contacted for additional data some of which were provided Data in MA taken from GlaxoKline Beecham web-based report Type of data used for remission/response: last observation carried forward
<i>Risk of bias</i>	

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"a computer generated randomization list" GlaxoKline Beecham "centralised computer-generated randomisation list" pg 61
Allocation concealment (selection bias)	Low risk	"masterlist held by SB...individual sealed code breaks held by investigators...could be broken in case of emergency" GlaxoKline Beecham "centralised" pg 61
Blinding (performance bias and detection bias) Intervention blinded	Low risk	"paroxetine and placebo capsules were identical and all packaging maintained the double blind nature of the trial" GlaxoKline Beecham pg 32 "placebo and paroxetine capsules were centrally prepared and packaged and were identical in appearance so that all trial personnel and patients were blind" pg 61
Blinding (performance bias and detection bias) Blinded outcome	Unclear risk	No statement
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Number eligible: 286 Number randomised: paroxetine: 187; placebo: 99; total: 286 Number started trial: paroxetine: 187; placebo: 99; total: 286 Number of withdrawals: paroxetine: 60; placebo: 30; total: 90 Number analysed post intervention: paroxetine: 182; placebo: 93; total: 275 Reasons for drop out: figure 1 in Berard 2006 publication. Higher rate of drop out in the paroxetine group including higher rate of discontinuation due to adverse events and lost to follow-up. Higher rate of drop out due to lack of efficacy in the placebo group ITT analysis: intention-to-treat (ITT) population was all patients randomised who received at least one dose of double blind medication and at least one treatment assessment was available. GlaxoKline Beecham and pg 63 of Berard 2006 publi-

		cation Statistical methods: last observation carried forward (LOCF) analysis was used, but authors also did OC analysis. Used logistic regression and analysis of covariance; included treatment group, country group and covariates of age and baseline scores (pg 63)
Selective reporting (reporting bias)	Low risk	Authors undertake analysis of response in multiple ways. Authors undertake a post hoc analysis of > 16 years of age. Report least square means and SEs for depression scores. Trial protocol contained in final report
Other bias	Unclear risk	Contact: assessment undertaken at weeks 1, 2, 3, 4, 6, 8 and 12 weeks. Participants were able to have non directive supportive therapy during treatment Screening: 2 weeks screening period from screening assessment to baseline assessment Placebo lead-in: 2-week single-blind Baseline imbalance: none reported - authors state baseline characteristics were similar. Table 1 reports all demographic and clinical characteristics.

Emslie 1997

Methods	Trial design: randomised controlled trial; single site Power calculation: not reported Use of diagnostic criteria (or clear specification of inclusion criteria): yes Intervention integrity: assessed by clinical chemistry profile Outcome measures described or validated measures used: yes Follow-up assessment points: post intervention No. crossed over: none Funded by: National Institute of Mental Health
Participants	Setting of care: outpatients Recruitment: self referred or referred to mood disorders programme; none were recruited by media Mean age (SD): intervention = 12.2 (2.7); control: 12.5 (2.6) Age range: 7 to 17 years Gender (F:M): intervention = 22:26; control = 22:26 Methods used to diagnose: DSM-II-RK-SADS depressive items; CDSR-R ≥ 40; 3 independent diagnostic interviews and a 1-week placebo lead-in Diagnosis: MDD Baseline severity of depression: CDRS-R mean (SD) score: intervention = 58.5 (10.5); control = 57.6 (10.4); CGI not reported at baseline

	Length of current episode: (mean weeks) intervention 14.6 (9.7); placebo 13.7 (7.5) % first episode: intervention 47.9%; placebo 47.9% Co-morbidity (intervention): none 7; dysthymia 20; anxiety disorders 32; ADHD 16; ODD/CD 13 Co-morbidity (control): none 11; dysthymia 14; anxiety disorders 22; ADHD 13; ODD/CD 16 Location: USA Inclusion criteria: non psychotic MDD, single or recurrent; good general medical health; normal intelligence Exclusion criteria: bipolar I and II; psychotic depression; independent sleep-wake disorder; alcohol and other substance abuse; anorexia nervosa; bulimia nervosa; previous adequate treatment with fluoxetine; at least 1 first-degree relative with bipolar I disorder Exclusion of suicidality: not specifically stated	
Interventions	Intervention group Drug: fluoxetine Dosage: 20 mg Regimen: taken daily Length of treatment: 8 weeks (following acute treatment, participants were given the option to continue treatment blindly or be treated openly) Control group: placebo pill	
Outcomes	Definition and assessment of response: we used remission CDRS-R ≤ 28 (they used responders defined as CGI improvement rating of 1 or 2) Depressive symptoms: Children's Depression Rating Scale - Revised (CDRS-R) Functioning: Children's Global Assessment Scale (CGAS) Suicidal behaviour: FDA report; no report of continuous measure Other outcomes: Clinical Global Impressions Scale Improvement (CGI); Children's Depression Inventory (CDI); Beck Depression Inventory (BDI); Weinberg Screening Affective Scale (WSAS); Brief Psychiatry Rating Scale - Children's (BPRS-C)	
Notes	Additional data were sought and supplied by the authors. Data in the MA for child, adolescent and total populations taken from paper publication and these additional data. Child and adolescent data from author. MHRA # X065 MHRA contacted for additional data some of which was provided Type of data used for remission/response: last observation carried forward	
<i>Risk of bias</i>		
Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"randomisation was by a table of random numbers" pg 1032
Allocation concealment (selection bias)	Unclear risk	"randomisation was conducted by the pharmacy and clinicians who remained blind to assignment until the end of the trial" pg 1032; "pharmacy provided blinded medication" pg 1033; MHRA re

		port states that an interactive voice response system was used to maintain blinding through follow-up phase
Blinding (performance bias and detection bias) Intervention blinded	Low risk	“clinicians who remained blind to assignment” pg 1032; “pharmacy provided blinded medication”; “results of blood chemistry levels not provided to clinicians” pg 1033
Blinding (performance bias and detection bias) Blinded outcome	Low risk	“clinicians who remained blind to assignment” pg 1032 “blood chemistry levels were not provided to clinicians” pg 1032
Incomplete outcome data (attrition bias) All outcomes	Low risk	Number eligible: 106 Number randomised: fluoxetine: 48; placebo: 48; total: 96 Number started trial: fluoxetine: 48; placebo: 48; total: 96 Number of withdrawals: fluoxetine: 14; placebo: 22; total: 36 Number analysed post intervention: fluoxetine: 48; placebo: 48; total: 96 Reasons for drop out: numbers of drop outs and reasons for drop out described in Table 2 . There were greater numbers in the placebo group who dropped out due to lack of efficacy (19 versus 7) and greater numbers in the fluoxetine group who dropped out due to side effects (4 versus 1) ITT analysis: all those randomised completed and were included in responder outcome, pg 1033 Statistical methods: last observation carried forward (LOCF) used for all 96 subjects randomised for Child Depression Rating Scale-Revised (CDRS-R) outcome. Undertook linear regression and analysis of covariance with baseline measurement as covariate for secondary outcomes
Selective reporting (reporting bias)	High risk	pg 1033 states that a secondary analysis to explore time to remission was intended; this is never reported nor are remission rates reported Overall adverse outcomes are not reported

Emslie 1997 (Continued)

Other bias	Unclear risk	Contact: participants were seen weekly for 8 weeks and outcomes were measured at each visit (pg 1033). No other details given Screening: phone screening followed by 3 independent full evaluations over 3 weeks Placebo lead-in: 1 week single-blind Baseline imbalance: pg1033 states there were not differences on any clinical or demographic features except the fluoxetine group had a greater incidence of life time anxiety disorders Other: small trial; Beck Depression Inventory and Childrens Depression Inventory scores combined to give a total score; while stratified for age, no outcome reporting by age was given
------------	--------------	--

Emslie 2002

Methods	Trial design: randomised controlled trial; multicentre Power calculation: yes Use of diagnostic criteria (or clear specification of inclusion criteria): yes Intervention integrity: not described Outcome measures described or validated measures used: yes Follow-up assessment points: post assessment No. crossed over: none Funded by: Eli Lilly
Participants	Setting of care: outpatients Recruitment: academic hospitals and private research psychiatric clinics as well as newspaper and radio recruitment Mean age (SD): intervention = 12.70 (2.46); control = 12.69 (2.67) Age range: 8 to < 18 years Gender (F:M): intervention = 54:55; control = 54:56 Methods used to diagnose: DSM-IV Diagnostic Interview for Children and Adolescents (DICA) interview, CDRS-R ≥ 40 and CGI = 4, 3 independent diagnostic interviews and a 1-week placebo lead in Diagnosis: MDD Baseline severity of depression: CDRS-R mean (SD) score: intervention = 57.1 (9.9); control = 55.1 (11.8); CGI intervention 4.5 (0.6); placebo 4.4 (0.6) Length of current episode: (mean weeks) intervention: 60.44; placebo: 61.29 % first episode: intervention 79.8%; placebo 78.2% Co-morbidity (intervention): ADHD 16; ODD 17; CD 3 Co-morbidity (control): ADHD 15; ODD 17; CD 1 Location: USA Inclusion criteria: outpatients; aged 8 to < 18; primary diagnosis of non psychotic major depressive disorder, single or recurrent; depressive symptoms of at least moderate severity; no clinically significant ECG abnormalities; able to keep appointments; normal

	intelligence as judged by investigator Exclusion criteria: serious illness that was not stabilised; abnormal thyroid function; seizure disorder; bipolar I or II; sleep-wake disorder; psychotic depression; anorexia nervosa; bulimia nervosa; borderline personality disorder; substance abuse disorder; 1 or more first degree relatives with bipolar I disorder; organic brain diseases; previous failed response to antidepressant medication; prior adequate treatment with fluoxetine; receipt of fluoxetine within 3 months prior to trial entry; regular use of other psychotropic drugs Exclusion of suicidality: serious suicide risk (no further definition)	
Interventions	Intervention group Drug: fluoxetine Dosage: 20 mg Regimen: 1 week 10 mg daily, then 20 mg daily for 8 weeks Length of treatment: 9 weeks Control group: placebo pill	
Outcomes	Definition and assessment of response: we used remission CDRS-R ≤ 28 (they used responders defined as CGI improvement rating of 1 or 2 or at least a 30% reduction on CDRS-R. Remission was defined as a score of ≤ 28 on the CDRS-R. Depressive symptoms: Children's Depression Rating Scale - Revised (CDRS-R) Functioning: Global Assessment of Functioning Scale (GAF) Suicidal behaviour: FDA report; no report of continuous measure Adverse events Other outcomes: Clinical Global Impressions Scale Severity (CGI-Severity), Clinical Global Impressions Scale Improvement (CGI - Improvement), Hamilton Anxiety Rating Scale (HAMA), Beck Depression Inventory (BDI) Children's Depression Inventory (CDI), Montgomery-Asberg Depression Rating Scale (MADRS)	
Notes	Additional data were sought from authors. They did not have the additional data but gave a contact in Eli Lily. Eli Lily provided additional data. Data in the MA from the paper and from additional data supplied by Eli Lily. MHRA # HCJE MHRA contacted for additional data some of which were provided All data from paper (Table 3) Assume the P value (that goes with the adjusted treatment effect of 7.1; effect size 0.51; CI 3.3, 10.9) is adjusted but the means presented in table 3 and provided by the author are probably not. JM calculated SE from SDs (in STATA file) for depression symptom outcome Type of data used for remission/response: last observation carried forward	
Risk of bias		
Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"computer generated random sequence" pg 1206
Allocation concealment (selection bias)	Unclear risk	No statement

Blinding (performance bias and detection bias) Intervention blinded	Low risk	“both groups took three capsules daily” pg 1209
Blinding (performance bias and detection bias) Blinded outcome	Unclear risk	No complete statement “clinicians who were blinded to treatment group” pg 1209 plus patient and parent report pg 1206
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Number randomised: fluoxetine: 109; placebo: 110; total: 219 Number started trial: fluoxetine: 109; placebo: 110; total: 219 Number of withdrawals: fluoxetine: 19; placebo: 42; total: 61 Number analysed post intervention: fluoxetine: 109; placebo: 110; total: 219 Reasons for drop out: given in Figure 1 , pg 1208. More drop outs were due to lack of efficacy in the placebo versus the fluoxetine group (12 versus 5); clinician decision (11 versus 3); lost to follow-up (7 versus 1) and adverse events (9 versus 5) ITT analysis: “analysis of response and remission included only those patients treatment at least two weeks with trial drug” pg 1208 Statistical methods: last observation carried forward (LCOF); ANOVA for CDRS-R total score with baseline and each post baseline visit included as dependent variables pg 1208
Selective reporting (reporting bias)	High risk	Remission data and overall suicide-related event data were not reported in the paper but obtained from Eli Lilly
Other bias	High risk	Contact: each patient had 6 visits over the 8-week treatment period with outcome data collected at each visit Screening: 3-week screening period with 3 independent evaluations Placebo lead-in: 1-week single-blind placebo lead-in Baseline imbalance: report that there were no statistically significant differences in baseline patient characteristics (Table 2) and reasonably balanced for current comorbidities except for conduct disorder Other: none noted

Methods	Trial design: randomised controlled trial; multicentre Power calculation: yes Use of diagnostic criteria (or clear specification of inclusion criteria): yes Intervention integrity: yes - "Every effort was made to encourage patient compliance with the dosing regimen as per protocol. All patients were instructed to return their medication bottles with any unused drug to the investigator when they returned for each visit". Section 3.6 compliance with trial medication, Final Report Outcome measures described or validated measures used: yes Follow-up assessment points: post intervention No. crossed over: none Funded by: GlaxoSmithKline
Participants	Setting of care: outpatient Recruitment: no information Mean age (SD): intervention = 11.9 (3.00); control = 12.1 (2.95) Age range: 7 to 17 years Gender (F:M): intervention = 48:53; control = 47:55 Methods used to diagnose: DSM-IV, K-SADS-PL using 1-week screening phase Diagnosis: MDD Baseline severity of depression: CDRS-R mean (SD) score: intervention = 60.7 (9.37); control = 62.6 (8.96); CGI intervention 4; CGI placebo 4 Length of current episode: (mean months) intervention: 26.9 (28.62); placebo: 24.9 (27.08) % first episode: intervention 53.5%; placebo 52.9% Co-morbidity (intervention): ODD 5; GAD 4; overanxious disorder 3; attention deficit disorder 3; separation anxiety disorder 2; simple phobia 1; PTSD 1; enuresis 1; adjustment disorder with depressed mood 0 Co-morbidity (control): ODD 4; GAD 1; overanxious disorder 1; attention deficit disorder 1; separation anxiety disorder 0; simple phobia 0; PTSD 0; enuresis 0; adjustment disorder with depressed mood 1 Location: USA and Canada Inclusion criteria: 7 to 17 years; MDD Exclusion criteria: clinically predominant Axis I disorder other than MDD; history of psychotic episode or disorder; bipolar disorder; mental retardation or Pervasive Developmental Disorder; substance abuse or dependence within 3 months of screening or current positive test on drug screen; epilepsy; ECT within 3 months of screening; lactating or pregnant; sexually active female and not using contraception; requirement of concurrent psychotherapy; clear history of non response to SSRIs Exclusion of suicidality: suicidal or homicidal risk (no further definition)
Interventions	Intervention group Drug: paroxetine Dosage: 10 to 50 mg Regimen: week one 10 mg daily with option to increase up to 10 mg weekly to a maximum of 50 mg; reduction/tapering over 4 weeks post 8-week treatment Length of treatment: 8 weeks Control group: placebo pill

Outcomes	Definition and assessment of response: we used OC remission CDRS-R ≤ 28 for total; OC response for child and adolescent data CGI ≤ 2 (they used response defined as CGI Improvement of 1 or 2) Depressive symptoms: Children's Depression Rating Scale - Revised (CDRS-R) Functioning: Global Assessment of Functioning (GAF) Suicidal behaviour: report of events based on Columbia classification ; no report of continuous measure Adverse events: gathered by spontaneous report from patient and family Other outcomes: Clinical Global Impressions Scale Severity (CGI-Severity); Clinical Global Impressions Scale Improvement (CGI-Improvement); Kutcher Adolescent Depression Rating Scale (KADS)	
Notes	MHRA #701 MHRA contacted for additional data some of which were provided Data in MA taken from GlaxoKline Beecham web-based report Type of data used for remission/response: last observation carried forward	
<i>Risk of bias</i>		
Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"a computer generated randomisation list was generated...stratified by age subgroup and performed in blocks" GlaxoKlineBeecham report "computer generated randomization list" pg 711
Allocation concealment (selection bias)	Unclear risk	"individual sealed envelopes indicating treatment assigned to each patient at a particular visit were lodged with the investigators/pharmacist....the master randomisation list was held by the sponsor" The investigators were blind to the trial medication except in the instance of a serious adverse event. GlaxoKlineBeecham report
Blinding (performance bias and detection bias) Intervention blinded	Unclear risk	"double blind. GlaxoKlineBeecham report. ..paroxetine and placebo...identical in size shape and colour...blinding of trial medication was maintained by referring to daily medication dose as dose level" pg 33
Blinding (performance bias and detection bias) Blinded outcome	Unclear risk	No statement

Incomplete outcome data (attrition bias) All outcomes	Low risk	Number eligible: 305 Number randomised: paroxetine: 104; placebo: 102; total: 206 Number started trial: paroxetine: 104; placebo: 102; total: 206 Number of withdrawals: paroxetine: 34; placebo: 23; total: 57 Number analysed post intervention: paroxetine: 101; placebo: 102; total: 203 Reasons for drop out: were reported in Figure 1. There were more drop outs in the paroxetine group, including more adverse events, more lost to follow-up and more who withdrew for any reason. There were more drop outs due to lack of efficacy ITT analysis: "the Intention-to-Treat (ITT) population...was all patients...who received at least one dose of randomised double blind treatment, and for whom at least one valid post-baseline evaluation was available" GlaxoKlineBeecham pg55; "All of patients who were randomised to the treatment phase, received at least one dose of trial medication and had at least one post baseline safety or efficacy assessment were included in the ITT population" pg 711 Statistical methods: last observation carried forward (LOCF) using the ITT population and observed case (OC) data analysis undertaken. Analysis of variance techniques and logistic regression. Adjusted for age group, gender, baseline scores and presence/absence of psychiatric comorbidity pg 711-712
Selective reporting (reporting bias)	Low risk	Remission appears to be a post hoc analysis
Other bias	Unclear risk	Contact: assessments undertaken at week 1, 2, 3, 4, 6 and 8 Screening: 1-week screening phase Placebo lead-in: no Baseline imbalance: states that 2 groups were similar at baseline. Reported in Table 1.

Methods	Trial design: randomised controlled trial; multicentre Power calculation: yes Use of diagnostic criteria (or clear specification of inclusion criteria): yes Intervention integrity: not described Outcome measures described or validated measures used: yes Follow-up assessment points: post-intervention No. crossed over: none Funded by: Wyeth Research
Participants	Setting of care: outpatient (consisting of academic and clinical sites) Recruitment: no information Mean age (SD): intervention = 12.2 (2.6); control = 12.3 (2.6) Age range: 7 to 17 years Gender (F:M): intervention = 78:101; control: 83:92 Methods used to diagnose: DSM-IV Schedule for Affective Disorder and Schizophrenia for School-Age Children-present and Lifetime version (K-SADS-PL), at pre-trial and baseline a CDRS-R score of ≥ 40, and CGI-S score of ≥ 4 and depressive symptoms for at least 1 month before trial entry. Single-blind placebo run-in period of 14 days (+/- 3) for trial 1 and 7 days (+/-3) for trial 2. Diagnosis: MDD Baseline severity of depression: CDRS-R mean (SD) score: intervention = 56.4 (9.2); control = 55.8 (8.4); CGI intervention 4.5 (0.6); CGI placebo 4.5 (0.7) Length of current episode: (mean weeks) intervention 91.1 (88.2); placebo 92.5 (91.3) % first episode: intervention 84.9%; placebo 86.9% Co-morbidity: not stated for either group Location: USA Inclusion criteria: DSM-IV criteria for MDD, pre-trial and baseline scores > 40 on the CDRS-R with ≤ 30 decrease between pretrial and baseline, a CGI-S score of ≥ 4 at pretrial and baseline, and depressive symptoms for at least 1 month prior to trial entry Exclusion criteria: history of any psychotic disorder or bipolar disorder; MDD with psychotic features, anorexia or bulimia, conduct disorder, panic disorder, or obsessive-compulsive disorder; first degree relative with bipolar disorder; recent drug or alcohol dependence or abuse; mental disorder caused by medical condition Exclusion of suicidality: acute suicidality (no further definition)
Interventions	Intervention group Drug: venlafaxine extended release Dosage: flexible dose based on body weight (37.5 mg/day to 225 mg/day). Mean daily dose was 109.2 mg/day for adolescents and 80.4 mg/day for children. Regimen: delivered once daily for 8 weeks followed by a taper period of up to 14 days Length of treatment: 8 weeks Control group: placebo pill
Outcomes	Definition and assessment of response: we used response $> 35\%$ decrease in CDRS-R (they used $\geq 35\%$ decrease in CDRS-R scores, $\geq 50\%$ decrease in HAM-D or MADRS or Clinical Global Impression Severity Scale (CGI-I; Guy 1976). A score of 1 (very much improved) or 2 (much improved) defining response) Depressive symptoms: Childhood Depression Rating Scale (CDRS-R; Poznanski 1996) Suicidal behaviour: FDA data; no report of continuous measure Functioning: GAF used but no report of data

	Adverse events	
Notes	Type of data used for remission/response: last observation carried forward	
<i>Risk of bias</i>		
Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	“Two similarly designed multi-centre, randomized...eligible subjects were randomly assigned to receive venlafaxine ER or placebo...” pg 480 (Method)
Allocation concealment (selection bias)	Unclear risk	No statement
Blinding (performance bias and detection bias) Intervention blinded	Unclear risk	“Two similarly designed multi-center, randomized, double-blind...” pg 480 (trial design)
Blinding (performance bias and detection bias) Blinded outcome	Unclear risk	”Two similarly designed multi-center, randomized, double-blind...” pg 480 (trial design)
Incomplete outcome data (attrition bias) All outcomes	Low risk	Number eligible: no statement (575 screened in total) Number randomised: venlafaxine ER: 184; placebo: 183; total: 367 Number started trial: venlafaxine ER: 182; placebo: 179; total: 361 Number of withdrawals: venlafaxine ER: 59; placebo: 50; total: 109 Number analysed post intervention: venlafaxine ER: 169; placebo: 165; total: 334 Reasons for drop out: drop outs and reasons were reported in Figure 1 . There were many more drop outs due to adverse events in the venlafaxine arms; and more in the venlafaxine arm who failed to return ITT analysis: the primary efficacy analysis population was the intent-to-treat population, which included all randomised subjects who had taken at least 1 dose of assigned medication and were evaluated for the primary efficacy outcome measure at baseline and at least once during therapy or within 3 days of the last full dose of treatment, pg 481 Statistical methods: last observation carried

Emslie 2007 (Continued)

		forward on therapy evaluation; also did observed case analysis. Parametric 2-way analysis of covariance with treatment and investigator as main effects and baseline score as covariate.
Selective reporting (reporting bias)	High risk	Overall adverse events were not reported; standard errors rather than standard deviations were reported; functioning and remission, apparently not trial outcomes
Other bias	High risk	Contact: visits on day 4, 7, 14, 21, 28, 42, 56 and at end of taper Screening: yes; describes a pretrial and baseline visit (unclear what time points these were) Placebo lead-in: trial 1: single-blind for 14 days (+/-3 days); trial 2: single-blind for 7 days (+/-3 days) Baseline imbalance: authors report that there were no statistically significant differences seen in demographic or clinical characteristics. Baseline characteristics not reported by individual trial Other: there were 2 studies reported in the one paper. Trial 1 had a higher drop out. One site was excluded from trial 1. Results are inconsistently reported by trial

Emslie 2007 Trial 1

Methods	See Emslie 2007	
Participants	See Emslie 2007	
Interventions	See Emslie 2007	
Outcomes	See Emslie 2007	
Notes	See Emslie 2007	
<i>Risk of bias</i>		
Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	“Two similarly designed multi-center, randomized...eligible subjects were randomly assigned to receive venlafaxine ER or placebo...” pg 480 (Method)

Allocation concealment (selection bias)	Unclear risk	No statement
Blinding (performance bias and detection bias) Intervention blinded	Unclear risk	"Two similarly designed multi-centre, randomized, double-blind..." pg 480 (trial design)
Blinding (performance bias and detection bias) Blinded outcome	Unclear risk	"Two similarly designed multi-center, randomized, double-blind..." pg 480 (trial design)
Incomplete outcome data (attrition bias) All outcomes	Low risk	Number eligible: no statement (575 screened in total) Number randomised: venlafaxine ER: 184; placebo: 183; total: 367 Number started trial: venlafaxine ER: 182; placebo: 179; total: 361 Number of withdrawals: venlafaxine ER: 59; placebo: 50; total: 109 Number analysed post intervention: venlafaxine ER: 169; placebo: 165; total: 334 Reasons for drop out: drop outs and reasons were reported in Figure 1. There were many more drop outs due to adverse events in the venlafaxine arms; and more in the venlafaxine arm who failed to return ITT analysis: the primary efficacy analysis population was the intent-to-treat population, which included all randomised subjects who had taken at least 1 dose of assigned medication and were evaluated for the primary efficacy outcome measure at baseline and at least once during therapy or within 3 days of the last full dose of treatment, pg 481 Statistical methods: last observation carried forward on therapy evaluation; also did observed case analysis. Parametric 2-way analysis of covariance with treatment and investigator as main effects and baseline score as covariate
Selective reporting (reporting bias)	High risk	Overall adverse events were not reported; standard errors rather than standard deviations were reported; functioning, remission, apparently not trial outcomes
Other bias	High risk	Contact: visits on day 4, 7, 14, 21, 28, 42, 56 and at end of taper Screening: not reported Placebo lead-in: trial 1: single-blind for 14 days (+/-3 days); trial 2: single-blind for 7 days (+/-3 days) Baseline imbalance: authors report that there were no statistically significant differences seen in demographic or clinical characteristics. Baseline characteristics not reported by individual trial Other: there were 2 studies reported in the one paper. Trial 1 had a higher drop out. One site was excluded from trial 1. Results are inconsistently reported by trial

Emslie 2007 Trial 2

Methods	See Emslie 2007
Participants	See Emslie 2007
Interventions	See Emslie 2007
Outcomes	See Emslie 2007
Notes	See Emslie 2007

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	"Two similarly designed multi-centre, randomized...eligible subjects were randomly assigned to receive venlafaxine ER or placebo..." pg 480 (Method)
Allocation concealment (selection bias)	Unclear risk	No statement
Blinding (performance bias and detection bias) Intervention blinded	Unclear risk	"Two similarly designed multi-center, randomized, double-blind..." pg 480 (trial design)
Blinding (performance bias and detection bias) Blinded outcome	Unclear risk	"Two similarly designed multi-centre, randomized, double-blind..." pg 480 (trial design)
Incomplete outcome data (attrition bias) All outcomes	Low risk	Number eligible: no statement (575 screened in total) Number randomised: venlafaxine ER: 184; placebo: 183; total: 367 Number started trial: venlafaxine ER: 182; placebo: 179; total: 361 Number of withdrawals: venlafaxine ER: 59; placebo: 50; total: 109 Number analysed post intervention: venlafaxine ER: 169; placebo: 165; total: 334 Reasons for drop out: drop outs and reasons were reported in Figure 1. There were many more drop outs due to adverse events in the venlafaxine arms; and more in the venlafaxine arm who failed to return ITT analysis: the primary efficacy analysis population was the intent-to-treat population, which included all randomised subjects who had taken at least 1 dose of assigned medication and were evaluated for the primary efficacy outcome measure at baseline and at least once during therapy or within 3 days of the last full dose of treatment, pg 481 Statistical methods: last observation carried forward on therapy evaluation; also did observed case analysis. Parametric 2-way

Emslie 2007 Trial 2 (Continued)

		analysis of covariance with treatment and investigator as main effects and baseline score as covariate
Selective reporting (reporting bias)	High risk	Overall adverse events were not reported; standard errors rather than standard deviations were reported; functioning, remission, apparently not trial outcomes
Other bias	High risk	Contact: visits on day 4, 7, 14, 21, 28, 42, 56 and at end of taper Screening: not reported Placebo lead-in: trial 1: single-blind for 14 days (+/-3 days); trial 2: single-blind for 7 days (+/-3 days) Baseline imbalance: authors report that there were no statistically significant differences seen in demographic or clinical characteristics. Baseline characteristics not reported by individual trial Other: there were 2 studies reported in the 1 paper. Trial 1 had a higher drop out. One site was excluded from trial 1. Results are inconsistently reported by trial

Emslie 2009

Methods	Trial design: randomised controlled trial; multicentre Power calculation: yes Use of diagnostic criteria (or clear specification of inclusion criteria): yes Intervention integrity: not described Outcome measures described or validated measures used: yes Follow-up assessment points: post-intervention No. crossed over: none Funded by: Forest laboratories
Participants	Setting of care: outpatient Recruitment: no information Mean age (SD): intervention = 14.7 (1.6); control = 14.5 (1.5) Age range: 12 to 17 years Gender (F:M): intervention = 92:63; control = 92:65 Methods used to diagnose: DSM-IV with duration of current episode at least 12 weeks at screening confirmed by K-SADS. At screening and baseline a CDRS-R score of ≥ 45 and a CGI-S score of ≥ 4. Screening period of 2 weeks, and a single-blind placebo run-in of 1 week during 2nd week of screening. Diagnosis: MDD Baseline severity of depression: CDRS-R mean (SD) score: intervention = 56.0 (0.66); control = 57.6 (0.66); CGI intervention 4.6 (0.05); CGI placebo 4.4 (0.04) Length of current episode: (mean months) intervention 15.7 (17.4); placebo 16.5 (15.4) % first episode: intervention 70.3%; placebo 72% Co-morbidity (intervention): previous and/or ongoing secondary psychiatric disorder 16.6% Co-morbidity (control): previous and/or ongoing secondary psychiatric disorder 12.9% Location: USA

	Inclusion criteria: current DSM-IV defined MDD episode of at least 12 weeks, CDRS-R score ≥ 45 at screening and baseline visits, CGI-S score of ≥ 4 and a score of ≥ 80 on the Kaufman Brief Intelligence Test Exclusion criteria: a principal diagnosis meeting DSM-IV criteria for an Axis 1 disorder other than MDD or who currently met DSM-IV criteria at screening for attention-deficit/hyperactivity disorder, obsessive-compulsive disorder, post-traumatic stress disorder, bipolar disorder, pervasive developmental disorder, mental retardation, conduct disorder, or oppositional defiant disorder; or who had any psychotic features or a history of any psychotic disorder; or any personality disorder that as judged by the investigator, would interfere with participation in the trial; a history of manic, or hypomanic episodes, a history of bulimia anorexia nervosa or substance abuse or dependence within the last year Exclusion of suicidality: patients considered a suicide risk by the investigator, including those who had active suicidal ideation, had made a suicide attempt, or had ever been hospitalised because of a suicide attempt	
Interventions	Intervention group Drug: escitalopram Dosage: 10 to 20 mg/day Regimen: daily Length of treatment: 8 weeks Control group: placebo pill	
Outcomes	Definition and assessment of response: we used remission CDRS-R ≤ 28 (they used the Clinical Global Impressions-Improvement Scale (CGI-I). A score of 1 (very much improved) or 2 (much improved) or CDRS-R reduction of $\geq 40\%$ defined response (remission CDRS-R ≤ 28)) Depressive symptoms: the Childrens Depression Rating Scale (CDRS-R; Poznanski 1996) Functioning: Children's Global Assessment Scale (C-GAS; Shaffer 1985) Suicidal behaviours: Modified Columbia Suicide Severity Rating Scale (MC-SSRS) report of suicidal ideation, presence and type of suicidal behaviour since last visit; continuous measure using the Suicidal Ideation Questionnaire-Junior High School Version (SIQ-JR; Reynolds 1987) Adverse events: either spontaneously reported by patient or parent, or noted by investigator	
Notes	Type of data used for remission/response: last observation carried forward	
<i>Risk of bias</i>		
Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	"...a prospective, randomized, double-blind placebo controlled trial..." pg 721 (Abstract)
Allocation concealment (selection bias)	Unclear risk	No statement

Blinding (performance bias and detection bias) Intervention blinded	Unclear risk	"This was a randomized, double-blind, placebo controlled trial..." pg 722 (Method)
Blinding (performance bias and detection bias) Blinded outcome	Unclear risk	"Evaluations were scheduled at the end of ..weeks of double-blind treatment" pg 723 (trial design)
Incomplete outcome data (attrition bias) All outcomes	Low risk	Number eligible: not stated Number randomised: escitalopram: 158; placebo: 158; total: 316 Number started trial: escitalopram: 155; placebo: 157; total: 312 Number of withdrawals: escitalopram: 32; placebo: 25; total: 57 Number analysed post intervention: escitalopram: 154; placebo: 157; total: 311 Reasons for drop out: described in Figure 1 and appear relatively well balanced, with slightly more drop outs due to adverse events in the escitalopram group. Authors report no statistically significant differences ITT analysis: efficacy analyses were performed on the intent-to-treat (ITT) population, which included all patients in the safety population who had at least 1 post-baseline CDRS-R assessment pg 723 Statistical methods: reported results are LOCF unless otherwise specified, pg 724 (Statistical Methods); baseline imbalance was tested. 2-way analysis of variance model with treatment and trial centre as factors for continuous variables and a Cochrane-Mantel Haenzel test controlling for trial centre for categorical variables. ANCOVA with treatment group and trial centre as factors and baseline scores as covariate. Logistic regression with treatment group and baseline score as explanatory variables. Dichotomous data. Also did mixed modelling for repeated measures for primary efficacy variable
Selective reporting (reporting bias)	High risk	2 measures of response were reported (one significant and one not)
Other bias	Low risk	Contact: evaluation at the end of week 1, 2, 3, 4, 6 and 8. In page 728 authors explain

		the placebo response as being due to the “extensive contact”. Psychotherapy was not allowed Screening: 2-week screening with 2 visits Placebo lead-in: 1-week single-blind Baseline imbalance: authors report no statistically significant differences, although there appears to be a difference in the baseline CGI-S score Other: none noted
--	--	--

Keller 2001

Methods	Trial design: randomised controlled trial; multicentre Power calculation: yes Use of diagnostic criteria (or clear specification of inclusion criteria): yes Intervention integrity: yes - “Compliance with taking trial medication was assessed by recording the amount of drug dispensed, taken, and returned in the CRF for each patient”. Section 3.6 Final report. Outcome measures described or validated measures used: yes Follow-up assessment points: post intervention No. crossed over: none Funded by: GlaxoSmithKline
Participants	Setting of care: outpatient Recruitment: no information Mean age (SD): intervention = 14.8 (1.6); control = 15.1 (1.6) Age range: 12 to 18 years Gender (F:M): intervention = 58:35; control: 57:30 Methods used to diagnose: DSM-IV diagnosis confirmed by K-SADS-L and current duration of episode at least 8 weeks, a score of ≥ 12 on the HAM-D, a CGAS score of ≥ 60; screening period of 7 to 14 days, no placebo run-in phase Diagnosis: MDD Baseline severity of depression: K-SADS 9-item depression score; intervention = 28.25; control = 28.84. C-GAS mean (SD) score: intervention = 42.7; control = 42.8; CGI not reported Length of current episode: (mean months) intervention: 14 (18); placebo: 13 (17) % first episode: intervention 81%; placebo 77% Co-morbidity (intervention): any diagnosis 41; anxiety disorder 19; externalising disorder 25 Co-morbidity (control): any diagnosis 50; anxiety disorder 26; externalising disorder 26 Location: USA and Canada Inclusion criteria: MDD of at least 8 weeks duration; at least 80 on the Peabody Picture Completion Task; Medically Healthy Exclusion criteria: bipolar disorder; schizoaffective disorder; eating disorder; alcohol or substance abuse disorder; Obsessive Compulsive Disorder; autism/pervasive developmental disorder; organic brain disorder; PTSD within 12 months of trial entry; current psychotropic drug use; trial of antidepressant medication within 6 months of trial entry

	Exclusion of suicidality: current suicidal ideation with intent or specific plan; history of suicide attempt by drug overdose	
Interventions	Intervention group Drug: paroxetine Dosage: 20 to 40 mg Regimen: 20 mg daily in week 1 to 4 with optional increase to 30 mg in week 5 and 40 mg in week 6 Length of treatment: 8 weeks Control group: placebo pill Comparison group: imipramine (gradual upward titration to 200 to 300 mg)	
Outcomes	Definition and assessment of response: we used OC response HAM-D ≤ 8 or $\geq 50\%$ reduction in baseline HAM-D (they used responders defined as ≤ 8 or less on HAM-D or at least 50% decrease from baseline) Depressive symptoms: depression items from K-SADS-L Functioning: Autonomous Function Checklist Suicidal behaviours: FDA data; no report of continuous measure Adverse events Other measures: HAM-D; Clinical Global Impressions Scale Improvement (CGI - Improvement); Self Perception Profile; Sickness Impact Scale	
Notes	Additional data were sought from the authors. They did not have the data required but provided a contact from GlaxoSmithKline who responded to inform us of the trial information now published on the web. MHRA # 329 MHRA contacted for additional data some of which were provided Data in MA taken from GlaxoSmithKline Beecham web-based report GlaxoSmithKline web publication Type of data used for remission/response: last observation carried forward	
<i>Risk of bias</i>		
Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	“computer generated list” pg 764
Allocation concealment (selection bias)	Unclear risk	No statement. GlaxoSmithKline Beecham states randomisation codes were stored at SB clinical safety department, pg 35
Blinding (performance bias and detection bias) Intervention blinded	Low risk	“Tablets overencapsulated in matching supro B locking capsules to preserve medication blinding”; “ number of capsules. ..identical for each...group during forced titration” pg 764

Blinding (performance bias and detection bias) Blinded outcome	Unclear risk	No statement
Incomplete outcome data (attrition bias) All outcomes	Low risk	Number eligible: 275 Number randomised: paroxetine: 93 placebo: 87 imipramine: 95 total: 275 Number started trial: paroxetine: 93 placebo: 87 imipramine: 95 total: 275 Number of withdrawals: paroxetine: 26 placebo: 21 imipramine: 38 total: 85 Number analysed post intervention: paroxetine: 67 placebo: 66 imipramine: 57 total: 190 Reasons for drop out: some information was provided (pg 765) about drops outs, but only about premature trial discontinuation due to adverse effects, which was 6.9% in the placebo group and 9.7% in the placebo group (P = 0.50) and described protocol violation as the most common reason for withdrawal in the placebo group (pg 765) ITT analysis: "efficacy analysis based on patients who were randomised and had at least one post baseline efficacy analysis evaluation" pg 76 Statistical methods: both last observation carried forward (LOCF) and observed case (OC) data analysis undertaken. 2 factor analysis of variance using general linear models with terms for treatment and investigator. Logistic analysis implemented in the categorical modelling procedure including effects for investigator and treatment
Selective reporting (reporting bias)	High risk	Multiple measurement of depression outcome (HAM-D, HAM-D depressed mood item, depression item of K-SADS-L and 9-item depression sub-scale of the K-SADS) . Response data given as percentages. In a letter to the editor Jureidini 2003 states that the definition of response is changed so that a significant result can be reported. Overall adverse event rate not described. Kennard 2006 (TADS) states that Keller had remission definition of HAM-D < 8, although Keller describes this in the methods section

		as 'response'
Other bias	High risk	Contact: participants were seen weekly and undertook assessments at each visit. Supportive case management was provided to all subjects at each visit (interpersonal or cognitive behavioural psychotherapeutic interventions were strictly prohibited) (pg 764) Screening: 7 to 14-day screening period with no detail about number of assessments during this screening phase Placebo lead-in: no Baseline imbalance: treatment groups stated to be similar at baseline for demographic and psychiatric profile (pg 765). These features are described in Table 1.

Mirtazapine Trial 1

Methods	Trial design: randomised controlled trial; multicentre Power calculation: not stated Use of diagnostic criteria (or clear specification of inclusion criteria): yes Intervention integrity: yes - "Plasma samples, for the purpose of measuring mirtazapine, (Org 3770) concentrations, were to be collected on trial Days 28 and 56 (or the subject's final day of treatment)" pg 4. Company trial report Outcome measures described or validated measures used: yes Follow-up assessment points: post intervention No. crossed over: none Funded by: Organon International
Participants	Setting of care: outpatients Recruitment: through clinical practice of investigators, referrals and/or advertisements for volunteers Mean age (SD): intervention = 12.3; control = 12.4 Age range: 8 to 18 years Gender (F:M): intervention = 39:43; control: 25:19 Methods used to diagnose: DSM-IV diagnosis confirmed by K-SADS-L and baseline score of ≥ 15 on 1st 17 items of HAM-D (21-item), a CGAS score of < 70; CDRS-R ≥ 40; screening period not stated Diagnosis: MDD Baseline severity of depression: CDRS-R mean (SD) score: intervention = 50.93; control = 51.93 Length of current episode: not stated % first episode: not stated Co-morbidity (intervention): not stated Co-morbidity (control): not stated Location: USA

Mirtazapine Trial 1 (Continued)

	Inclusion criteria: current episode of MDD (as defined by DSM-IV criteria, with a primary diagnosis of major depressive disorder on the Kiddie-SADS P-L (Kiddie schedule for affective disorders and schizophrenia - present and lifetime). Baseline score of >15 on the 1st 17 items of the Hamilton Scale for Depression, 21 items (HAM-D 21), <70 on the Children's Global Assessment Scale (C-GAS), and a Children's Depression Rating Scale-Revised (CDRS-R) score of ≥ 40 Exclusion criteria: concurrent psychiatric diagnosis of anorexia or bulimia, past history of eating disorder, concurrent diagnosis of obsessive compulsive disorder or schizophrenia, bipolar disorder (I or II) or parental history of bipolar I disorder, drug/and or alcohol abuse Exclusion of suicidality: serious suicide attempt during the current major depressive episode, or any previous suicide attempt resulting in hospitalisation	
Interventions	Intervention group Drug: mirtazapine Dosage: 15 to 45 mg Regimen: starting dose 15 mg with increase to 30 to 45 mg in 15 mg increments during subsequent weeks (to 28 days) Length of treatment: 8 weeks Control group: placebo pill	
Outcomes	Definition and assessment of response: not stated Depressive symptoms: CDRS-R clinician rating; HAM-D 21 self rating Functioning: C-GAS used but no report of data Suicidal behaviours: events reported as adverse events; no report of continuous measure Adverse events Other measures: Clinical Global Impressions (CGI), Self Report Childhood Anxiety Related Disorder (SCARED), Connors' Global Index (Parent and Teacher Versions)	
Notes		
<i>Risk of bias</i>		
Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No statement
Allocation concealment (selection bias)	Unclear risk	No statement
Blinding (performance bias and detection bias) Intervention blinded	Unclear risk	MHRA report states double-blind
Blinding (performance bias and detection bias) Blinded outcome	Unclear risk	MHRA report states double-blind

Mirtazapine Trial 1 (Continued)

Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Number eligible: not stated Number randomised: mirtazapine: 82; placebo: 44; total: 126 Number started trial: mirtazapine: 82; placebo: 44; total: 126 Number of withdrawals: mirtazapine: 13; placebo: 9; total 22 Number analysed post intervention: mirtazapine: 82; placebo: 44; total: 126 Reasons for drop out: MHRA reports drops across the 2 mirtazapine trials: 9 (5.3%) patients discontinued due to an adverse event in the mirtazapine group compared with 3 (3.4%) in the placebo-treated group. The most common adverse treated event leading to discontinuation in the acute phase in the mirtazapine treated group was weight gain. Weight gain (31.8% versus 3.4%), somnolence (38.8% versus 6.8%), headache (35% versus 23%), fatigue (19.4% versus 11.4%), increased appetite (8.8% versus 2.3%), urticaria (11.8% versus 6.8%) and hypertriglyceridaemia (2.9% versus 0%) were reported more often for mirtazapine-treated patients than by placebo-treated patients ITT analysis: state ITT using LOCF was used Statistical methods: not stated
Selective reporting (reporting bias)	High risk	Only 1 outcome reported in MHRA report; Rapporteurs report gives safety outcomes in addition
Other bias	High risk	Contact: weekly visits (week 5 and 7 optional); psychotherapy could not be started during the trial, but 'supportive care' as defined in the protocol was permitted Screening: unclear Placebo lead-in: no Baseline imbalance: data not reported States it was initially 2 trials that were amalgamated a few months after trial initiation

Mirtazapine Trial 1 & 2

Methods	Information provided separately for each trial	
Participants	Information provided separately for each trial	
Interventions	Information provided separately for each trial	
Outcomes	Information provided separately for each trial	
Notes	Information provided separately for each trial	
<i>Risk of bias</i>		
Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No statement
Allocation concealment (selection bias)	Unclear risk	No statement
Blinding (performance bias and detection bias) Intervention blinded	Unclear risk	MHRA report states double-blind
Blinding (performance bias and detection bias) Blinded outcome	Unclear risk	MHRA report states double-blind
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Reasons for drop out: MHRA reports drops across the 2 mirtazapine trials: 9 (5.3%) patients discontinued due to an adverse event in the mirtazapine group compared with 3 (3.4%) in the placebo-treated group. The most common adverse treated event leading to discontinuation in the acute phase in the mirtazapine-treated group was weight gain. Weight gain (31.8 versus 3.4%), somnolence (38.8% versus 6.8%), headache (35% versus 23%), fatigue (19.4% versus 11.4%), increased appetite (8.8% versus 2.3%), urticaria (11.8 versus 6.8%) and hypertriglyceridaemia (2.9 versus 0%) were reported more often for mirtazapine-treated patients than by placebo treated patients ITT analysis: state ITT using LOCF was used Statistical methods: not stated

Mirtazapine Trial 1 & 2 (Continued)

Selective reporting (reporting bias)	High risk	Only 1 outcome reported in MHRA report; Rapporteurs report gives safety outcomes in addition
Other bias	High risk	Contact: psychotherapy could not be started during the trial, but 'supportive care' as defined in the protocol was permitted Screening: unclear Placebo lead-in: unclear Baseline imbalance: data not reported States it was initially 2 trials that were amalgamated a few months after trial initiation

Mirtazapine Trial 2

Methods	Trial design: randomised controlled trial; multicentre Power calculation: not stated Use of diagnostic criteria (or clear specification of inclusion criteria): yes Intervention integrity: yes - "Plasma samples, for the purpose of measuring mirtazapine, (Org 3770) concentrations, were to be collected on trial Days 28 and 56 (or the subject's final day of treatment)". Pg 4. Company trial report Outcome measures described or validated measures used: yes Follow-up assessment points: post intervention No. crossed over: none Funded by: Organon International
Participants	Setting of care: outpatients Recruitment: through clinical practice of investigators, referrals and/or advertisements for volunteers Mean age (SD): intervention = 11.9; control = 12.3 Age range: 8 to 18 years Gender (F:M): intervention = 46:42; control: 24:21 Methods used to diagnose: DSM-IV diagnosis confirmed by K-SADS-L and baseline score of ≥ 15 on 1st 17 items of HAM-D (21 item), a CGAS score of < 70; CDRS-R ≥ 40; screening period not stated Diagnosis: MDD Baseline severity of depression: CDRS-R mean (SD) score: intervention = 48.87; control = 47.57 Length of current episode: not stated % first episode: not stated Co-morbidity (intervention): not stated Co-morbidity (control): not stated Location: USA Inclusion criteria: current episode of MDD (as defined by DSM-IV criteria, with a primary diagnosis of major depressive disorder on the Kiddie-SADS P-L (Kiddie schedule for affective disorders and schizophrenia - present and lifetime).

Mirtazapine Trial 2 (Continued)

	Baseline score of >15 on the 1st 17 items of the Hamilton Scale for Depression, 21 items (HAM-D 21), <70 on the Children's Global Assessment Scale (C-GAS), and a Children's Depression Rating Scale-Revised (CDRS-R) score of ≥ 40 Exclusion criteria: serious suicide attempt during the current major depressive episode, or any previous suicide attempt resulting in hospitalisation; concurrent psychiatric diagnosis of anorexia or bulimia, past history of eating disorder, concurrent diagnosis of obsessive compulsive disorder or schizophrenia, bipolar disorder (I or II) or parental history of bipolar I disorder	
Interventions	Intervention group Drug: mirtazapine Dosage: 15 to 45 mg Regimen: starting dose 15 mg with increase to 30 to 45 mg in 15 mg increments during subsequent weeks (to 28 days) Length of treatment: 8 weeks Control group: placebo pill	
Outcomes	Definition and assessment of response: not stated Depressive symptoms: CDRS-R clinician rating; HAM-D 21 self rating Functioning: C-GAS used but no report of data Suicidal behaviours: events reported as adverse events; no report of continuous measure Adverse events Other measures: Clinical Global Impressions (CGI), Self Report Childhood Anxiety Related Disorder (SCARED), Connors' Global Index (Parent and Teacher Versions)	
Notes		
<i>Risk of bias</i>		
Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No statement
Allocation concealment (selection bias)	Unclear risk	No statement
Blinding (performance bias and detection bias) Intervention blinded	Unclear risk	MHRA states double-blind
Blinding (performance bias and detection bias) Blinded outcome	Unclear risk	MHRA states double-blind
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Number eligible: not stated Number randomised: mirtazapine: 88; placebo: 45; total: 133 Number started trial: mirtazapine: 88;

Mirtazapine Trial 2 (Continued)

		placebo: 45 total: 133 Number of withdrawals: mirtazapine: 19; placebo: 8; total: 27 Number analysed post intervention: mirtazapine: 83; placebo: 41; total: 124 Reasons for drop out: MHRA reports drops across the 2 mirtazapine trials: 9 (5.3%) patients discontinued due to an adverse event in the mirtazapine group compared with 3 (3.4%) in the placebo-treated group. The most common adverse treated event leading to discontinuation in the acute phase in the mirtazapine treated group was weight gain. Weight gain (31.8% versus 3.4%), somnolence (38.8% versus 6.8%), headache (35% versus 23%), fatigue (19.4% versus 11.4%), increased appetite (8.8% versus 2.3%), urticaria (11.8% versus 6.8%) and hypertriglyceridaemia (2.9% versus 0%) were reported more often for mirtazapine-treated patients than by placebo-treated patients ITT analysis: states ITT done using LOCF but table of participants shows ITT analysis did not include all randomised patients Statistical methods: not stated
Selective reporting (reporting bias)	High risk	Only 1 outcome reported in MHRA report; Rapporteurs report gives safety outcomes in addition
Other bias	Unclear risk	Contact: weekly visits (week 5 and 7 optional); psychotherapy could not be started during the trial, but 'supportive care' as defined in the protocol was permitted Screening: unclear Placebo lead-in: no Baseline imbalance: data not reported States it was initially 2 trials that were amalgamated a few months after trial initiation

Paroxetine Trial 1

Methods	Trial design: randomised controlled trial; multi-site Power calculation: yes Use of diagnostic criteria (or clear specification of inclusion criteria): yes Intervention integrity: yes. Plasma concentration monitored. Outcome measures described or validated measures used: yes Follow-up assessment points: post-intervention No. crossed over: none Funded by: GSK
Participants	Setting of care: unclear Recruitment: not stated Mean age: intervention = 14.4 years (SD = 1.99); placebo = 14.8 years (SD = 2.62) Age range: 7 to 17 years Gender (F:M): intervention = 18:9; placebo = 16:13; total female = 34, male = 22 Methods used to diagnose: DSM-IV; CDRS-R score of ≥ 45 Diagnosis: MDD Baseline severity of depression: CDRS-R mean (SD) intervention = 55.4 (7.3); placebo = 56.8 (8.46) Length of current episode: not stated % first episode: not stated Co-morbidity: not stated Location: Japan Inclusion criteria: single episode of MDD or recurrent symptoms of depression or depressed state Exclusion criteria: primary diagnosis of an axis 1 disorder other than MDD, those with a history of psychotic episode or psychotic disorder or bipolar disorder Exclusion of suicidality: not stated
Interventions	Intervention group Drug: paroxetine Dosage: 10 to 40 mg dependent on age Regimen: 10 mg for 2 weeks and 10 to 20 mg for next 6 weeks for 7 to 11 year olds and 10 to 40 mg for the next 6 weeks for 12 to 17 year olds. The dose described at week 6 was maintained for the last 2 weeks Length of treatment: 8 weeks Control group: placebo pill
Outcomes	Definition and assessment of response: Clinical Global Impressions (CGI) of 1 or 2 Depressive symptoms: Change from baseline CDRS-R Functioning: no report of measure used Suicidal Behaviours: events reported as adverse events; no report of continuous data Other: Change from baseline CGI score; Incidence of adverse events
Notes	Type of data used for remission/response: last observation carried forward
<i>Risk of bias</i>	

Paroxetine Trial 1 (Continued)

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No information to make a judgement
Allocation concealment (selection bias)	Unclear risk	No information to make a judgement
Blinding (performance bias and detection bias) Intervention blinded	Unclear risk	Double-blind stated but no other details
Blinding (performance bias and detection bias) Blinded outcome	Unclear risk	Double-blind stated but no other details
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Number eligible: not stated Number randomised: paroxetine: 29; placebo: 27; total: 56 Number started trial: paroxetine: 29; placebo: 27; total: 56 Number of withdrawals: paroxetine: 4; placebo: 3; total: 7 Number analysed post intervention: paroxetine: 29; placebo: 27; total: 56 Reasons for drop out: broad reasons described ITT analysis: yes. Observed case data were used in some secondary analyses Statistical methods: last observation carried forward (LOCF) analysis was used for the primary outcome, analysis of covariance was used with CDRS-R total score at week 1 with total score as a covariate
Selective reporting (reporting bias)	High risk	High risk as no published peer-reviewed data. Drug company report available only
Other bias	Unclear risk	Contact: assessment at week 1, 2, 3, 4 and 6, no other details Screening: 2-week screening, a CDRS-R score of ≥ 45 at week -2 and week 0 Placebo lead-in: there was a 2-week placebo lead-in period Baseline imbalance: no specific statement, however proportion of female to male is different in the placebo group, and fewer children in the paroxetine group Other: did not reach planned recruitment numbers

Simeon 1990

Methods	Trial design: randomised controlled trial; single site Power calculation: not stated Use of diagnostic criteria (or clear specification of inclusion criteria): yes Intervention integrity: yes - assessed by clinical chemistry profile Outcome measures described or validated measures used: yes Follow-up assessment points: weekly visits, post intervention and long-term follow-up on average 24 months post trial termination No. crossed over: none Funded by: not stated
Participants	Setting of care: outpatient Recruitment: no information Mean age: 16 (group ages not stated) Age range: actual range not stated Gender(total): female = 22; male = 18 (group gender not stated) Methods used to diagnose: DSM-III criteria with HAM-D score of ≥ 20, 1-week placebo run-in period Diagnosis: MDD Baseline severity of depression: not stated for either group Length of current episode: not stated % first episode: not stated Co-morbidity: not stated for either group Location: Canada Inclusion criteria: 13 to 18 years; MDD with a HAM-D score > 20, a Raskin Depression Scale score of > 8, a Raskin Depression Score that must exceed the Covi Anxiety Scale Score, an outpatient Exclusion criteria: history of seizures, schizophrenia or other psychotic illnesses, girls who were sexually active and not using medically accepted means of contraception, patients with a recent drug or alcohol abuse Exclusion of suicidality: serious suicidal risk (no further definition)
Interventions	Intervention group Drug: fluoxetine Dosage: 20 to 60 mg Regimen: initial dose 20 mg daily increased to 40 mg after 4 to 7 days, and up to 60 mg in the second week Length of treatment: 7 weeks Control group: placebo pill
Outcomes	Definition and assessment of response: not stated Depressive symptoms: HAM-D; Raskin Depression Scale; Functioning: no report Suicidal behaviours: no report of events or continuous measure Other: Clinical Global Impressions Scale (CGI); Covi Anxiety Scale; Hopkins Symptom Checklist Follow-up assessment included semi-structured interviews by a nurse to obtain treatment subsequent to the trial, current activities and functioning with family and peers, and follow-up interview with parents using the HAM-D, Raskin, Covi and a DSM-III

	checklist for MDD and an adaptive functioning scale	
Notes	Letter requesting additional data sent. Data have not been received	
<i>Risk of bias</i>		
Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	“randomly assigned” pg 792 no other statement
Allocation concealment (selection bias)	Unclear risk	No statement
Blinding (performance bias and detection bias) Intervention blinded	Unclear risk	“double-blind” pg 792
Blinding (performance bias and detection bias) Blinded outcome	Unclear risk	No statement
Incomplete outcome data (attrition bias) All outcomes	High risk	Number eligible: not stated Number randomised: 40, group Ns not stated Number started trial: 40, group Ns not stated Number of withdrawals: 8, group Ns not stated Number analysed post intervention: fluoxetine: 16; placebo: 16; total: 32 Reasons for drop out: not stated ITT analysis: not stated Statistical methods: little detail provided; pg 792 states Wilcoxon's Rank Sum Test and Chi ² test used
Selective reporting (reporting bias)	High risk	No outcome data are reported
Other bias	High risk	Contact: no details are given of the contact time with clinicians in either group Screening: no details of screening procedure given Placebo lead-in: there was a 1-week single-blind placebo lead-in (pg 792) Baseline imbalance: pg 792 states there were no significant difference between groups at baseline; however, no demographic or clinical data are provided by group

Other: [Hammad 2004](#) reports that this trial was “terminated early” pg28

TADS 2004

Methods	Trial design: randomised controlled trial; multicentre Power calculation: yes Use of diagnostic criteria (or clear specification of inclusion criteria): yes Intervention integrity: not described for fluoxetine and placebo arms Outcome measures described or validated measures used: yes Follow-up assessment points: post intervention No. crossed over: none Funded by: NIMH
Participants	Setting of care: outpatient Recruitment: included newspaper, TV and radio advertising Mean age (total): 14.6 (SD 1.5) Age range (actual): 12 to 18 years Gender (F:M): 239:200 Methods used to diagnose: DSM-IV confirmed using K-SADS-PL and a CDRS-R score of ≥ 45; assessment (not interview) at consent and baseline Diagnosis: MDD Baseline severity of depression: CDRS-R raw mean (SD) score: intervention: 58.96 (10.16) (T-score 74.73 (6.74)); control: 61.11 (10.50) (T-score 76.14 (6.11)); CGI intervention 4.66; CGI placebo 4.84 Length of current episode: (median weeks) intervention 38 weeks; placebo: 35.5 weeks % first episode: 86% of total (not reported by group) Co-morbidity (intervention): any 47; dysthymia 6; anxiety 26; OCD/tic 2; ADHD 13; substance use 3; disruptive behaviour 25 Co-morbidity (control): any 57; dysthymia 12; anxiety 28; OCD/tic 4; ADHD 19; substance use 0; disruptive behaviour 28 Location: USA Inclusion criteria: outpatient; age 12 to 17; Full Scale IQ > 80; antidepressant-free before trial Exclusion criteria: bipolar disorder; severe Conduct Disorder; substance abuse; pervasive developmental disorder; thought disorder; use of psychotropic medication or psychotherapy (stable stimulants permitted for ADHD); 2 previous failed SSRI trials or a failed trial of CBT; confounding medical condition; non English speaking Exclusion of suicidality: suicidality or homicidality (patients were excluded for dangerousness to self or others if they had been hospitalised for dangerousness within 3 months of consent or were deemed by a cross-site panel to be “high risk” because of a suicide attempt requiring medical attention within 6 months, clear intent or an active plan to commit suicide, or suicidal ideation with a disorganized family unable to guarantee adequate safety monitoring)
Interventions	Intervention group Drug: fluoxetine Dosage: 20 to 40 mg Regimen: 10 mg daily to start; increase to 20 mg daily in week 1 with increase to a

	maximum of 40 mg daily thereafter Length of treatment: 12 weeks Control group: placebo Comparison group 1: CBT Comparison group 2: CBT plus fluoxetine
Outcomes	Definition of response and assessment: we used remission CDRS-R ≤ 28 (they used a range of outcomes including response and remission, using different definitions: In the main results paper they use response defined as a CGI improvement of 1 or 2) Depressive symptoms: Children's Depression Rating Scale - Revised (CDRS-R) Functioning: C - GAS Suicidal behaviours: report of events based on Columbia classification; continuous measure using the Suicidal Ideation Questionnaire-Junior High School Version (SIQ-Jr) Adverse events Other outcomes: Clinical Global Impressions Scale Improvement (CGI-Improvement) ; Reynolds Adolescent Depression Scale (RADS)
Notes	Additional trial information was sought and received from the author. Data in the MA from the paper. All young people in the trial were included as adolescents Type of data used for remission/response: last observation carried forward

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"computer stratified randomisation" pg 808 in 2004 publication
Allocation concealment (selection bias)	Low risk	"centralized IVRS service. Eligibility was assessed by same i.e. as did dependent variable assessments. trial coordinator not independent evaluator interfaced with IVRS and primary clinician for that patient revealed randomization status at Gate C2 after having first confirmed that patient/parent understood and were willing to accept randomization to any TADS treatment" from personal correspondence
Blinding (performance bias and detection bias) Intervention blinded	Low risk	"except in emergencies participants and clinicians remained blind in fluoxetine alone and placebo" groups pg 808 in 2004 publication
Blinding (performance bias and detection bias) Blinded outcome	Low risk	"as rated by an independent evaluator pg 535 in the 2003 publication; "masking was maintained for the primary dependent measures by means of independent evalu-

		ators blind to treatment assignment. Specific instructions were provided to parents, participants and the independent evaluator not to disclose treatment assignment“ pg 808 in the 2004 publication
Incomplete outcome data (attrition bias) All outcomes	Low risk	Number eligible: 1088 Number randomised: fluoxetine: 109; placebo: 112; total: 439 (including additional 2 trial arms) Number started trial: fluoxetine: 109; placebo: 112; total: 439 (including additional 2 trial arms) Number of withdrawals: fluoxetine: 18; placebo: 23; total: 90 (including additional 2 trial arms) Number analysed post intervention: fluoxetine: 109; placebo: 112; total: 439 (including additional 2 trial arms) Reasons for drop out: full table of number of drop outs and reason for drop outs given pg 811. Reasons for drop out are not specific e.g. terminated prematurely. Similar reasons in each group except 10 participants in the placebo group withdrew consent, compared with 5 in the fluoxetine group ITT analysis: "all analyses were conducted using an intent-to-treat analysis"; "primary intent to treat, all patients regardless of treatment status return for all scheduled assessments" pg 535 in the 2003 publication Statistical methods: for the CDRS -R results linear random coefficient regression model; used random-effects for participants and clinical site (but site interaction omitted). Responder (CGI-I) used logistic regression model for last available assessment point (LOCF) with site as covariate
Selective reporting (reporting bias)	High risk	Percentages given for CGI-I response rates. Multiple publications report varying outcome results that are not consistent across papers. In the 2004 paper presenting the main results, functioning was not reported
Other bias	Unclear risk	Contact: "Patients have one pharmacotherapist throughout the trial who, in addition to monitoring clinical status and medica-

		tion effects, offers general encouragement about the effectiveness of pharmacotherapy for MDD. Major assessments undertaken at baseline, 12 weeks, 24 weeks and 36 weeks with minor assessments at 6 weeks, 18 weeks and 30 weeks“ 2003 publication pg 537.” “six 20 to 30 minute medication visits spread across 12 weeks of treatment” 2004 publication pg 809 Screening: phone screening assessment followed by 1 full assessment to determine ‘caseness’, which on average takes 3 weeks (range 2 to 8 weeks) Placebo lead-in: no Baseline imbalance: for main results paper (2004) there were none reported; no demographic information given by group; Table 1 reports baseline clinical information with no significant differences reported across the four treatment groups TADS 2005 paper on demographics does not report demographic and clinical characteristics by group
--	--	---

Von Knorring 2006

Methods	Trial design: randomised controlled trial; multicentre Power calculation: not stated Use of diagnostic criteria (or clear specification of inclusion criteria): yes Intervention integrity: yes - non compliance assessed by blood levels of citalopram Outcome measures described or validated measures used: yes Follow-up assessment points: post intervention Funded by: pharmaceutical company not stated
Participants	Setting of care: in and outpatient (14% of participants hospitalised at entry to trial) Recruitment: no information Mean age (SD): 16 (1) Age range: 13 to 18 years Gender: not stated Methods used to diagnose: DSM-IV including 5-minute interview with parents. Global assessment of functioning less than 60 on either symptoms, activities, relationships or personal care, BDI less than 21 for girls and less than 16 for boys Diagnosis: MDD Baseline severity of depression: K-SADS-P intervention 32.5; control = 32.3 and totals only for MADRS 30 (SD = 5/6), GAF 55 (SD = 7); CGI not reported Length of current episode: not reported % first episode: intervention 72%; placebo 64% Co-morbidity: not stated for either group

	Location: Denmark, Estonia, Finland, Germany, Norway, Sweden, Switzerland Inclusion criteria: DSM-IV MDD current episode of greater than 4 weeks but less than 1 year duration; in or outpatient plus score of at least 21 or 16 on BDI and at least 60 on the GAF; 13 to 18 years inclusive; Tanner Stage III (commencement of puberty) Exclusion criteria: bipolar disorder including hypermania; ongoing DSM-IV Attention Deficit Disorder or disruptive behaviour disorder; DSM-IV psychotic disorder; progressive neurological disorder; drug or alcohol abuse that influences daily functioning; primary anorexia nervosa or bulimia nervosa; attends special school for mentally retarded; pervasive developmental disorders Exclusion of suicidality: not explicitly stated
Interventions	Intervention group Drug: citalopram Dosage: 10 to 40 mg Regimen: 10 mg for the first week with dose increases at the end of the week 1, 2, 5 or 9 weeks of 10 mg if GAF decreased by 10 points or unchanged to a maximum of 40 mg Length of treatment: 12 weeks Control group: placebo pill
Outcomes	Definition of response and assessment: we used OC remission MADRS < 12 (they used responders defined as those with a score of 2 or less on the Kiddie-SADS-P depression and anhedonia items or with a reduction of at least 50% from baseline of the MADRS total score) Functioning: Global Assessment of Functioning (GAF) Suicidal behaviours: FDA data; no report of continuous measure Adverse outcomes Other outcomes: K-SADS-P total score; Montgomery-Asberg Depression Rating Scale (MADRS); Beck Depression Inventory (BDI)
Notes	MHRA #94404 MHRA contacted for additional data some of which were provided We included data on self report depression from Von Knorring 2006 assuming the baseline standard deviations from Berard 2006 as the follow-up standard deviations of Von Knorring 2006 Type of data used for remission/response: last observation carried forward

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	"randomized" pg 311
Allocation concealment (selection bias)	Unclear risk	No statement
Blinding (performance bias and detection bias) Intervention blinded	Unclear risk	"double blind"

Blinding (performance bias and detection bias) Blinded outcome	Unclear risk	No statement
Incomplete outcome data (attrition bias) All outcomes	High risk	Number eligible: not stated Number randomised: citalopram: 124; placebo: 120; total: 244 Number started trial: citalopram: 121; placebo: 112; total: 233 Number of withdrawals: citalopram: 45; placebo: 46; total: 91 Number analysed post intervention: citalopram: 121; placebo: 112; total: 233 Reasons for drop out: full table of number of drop outs but full description of reasons for drop outs not given. More withdrew from the placebo group due to lack of efficacy and more withdrew from the citalopram group due to adverse effects ITT analysis: efficacy analyses were conducted on an intent-to-treat population, which included all randomised patients who took at least 1 dose of double-blind medication and who had at least 1 valid post assessment kiddie-SADs-P assessment pg 312 Statistical methods: primary analysis based on adjusted mean change of observed case data using ANCOVA (analysis of covariance). Dichotomous data analysed using LOCF
Selective reporting (reporting bias)	High risk	Error in the von Knorring paper when describing response data where it is reported twice and both times as OC data. Both response and remission data are only reported as percentages and when calculating these out using both the ITT population and the OC population the whole numbers do not match. Results only (no data) are reported for functioning, depression severity (clinician and self rated)
Other bias	Unclear risk	Contact: evaluation undertaken at 1, 2, 5, 9 and 12 weeks. Psychotherapy was allowed and three-quarters of the participants received it Screening: there was 1 screening visit and then a baseline visit

		Placebo lead-in: no Baseline imbalance: authors state that baseline data were similar for the 2 treatment groups, however, much baseline data (e.g. depression severity, age) was not reported by group. There were more patients in the citalopram group hospitalised for a psychiatric disorder and with a first episode Other: after recruitment of 15% of the population the trialists changed the inclusion criteria to ≥ 16 on the BDI for boys and added the MADRS. Post hoc analysis of high versus low baseline scores and of those receiving psychotherapy versus not receiving psychotherapy
--	--	--

Wagner 2004

Methods	Trial design: randomised controlled trial; multicentre Power calculation: not reported Use of diagnostic criteria (or clear specification of inclusion criteria): yes Intervention integrity: not described Outcome measures described or validated measures used: yes Follow-up assessment points: post intervention No. crossed over: none Funded by: Forest Pharmaceuticals
Participants	Setting of care: outpatients Recruitment: no information Mean age (SD): Intervention = 12.1 (2.8); control = 12.1 (3.1) Age range: 7 to 17 years Gender (F:M): intervention = 54:39; control = 43:42 Methods used to diagnose: DSM-IV confirmed using The Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS-P and L) and a CDRS-R score of ≥ 40 Diagnosis: MDD Baseline severity of depression: CDRS-R mean (SD) score: intervention = 58.8 (10.9); control = 57.8 (11.1); CGI not reported Length of current episode: (mean months) intervention: 20.8 (21.4); placebo: 18.6 (16.4) % first episode: intervention 78.7%; placebo 82.4% Co-morbidity (intervention): dysthymia 5; enuresis 4; previous ADHD 4 Co-morbidity (control): dysthymia 1; enuresis 3; previous ADHD 1 Location: USA Inclusion criteria: MDD of at least 4 weeks' duration; normal physical exam, laboratory tests and Electrocardiography (ECG); parent available to accompany child Exclusion criteria: primary psychiatric diagnosis other than MDD; ADHD; PTSD; bipolar disorder; pervasive developmental disorder; mental retardation; CD; ODD; any

	psychotic features; any personality disorder that would interfere with treatment; alcohol or substance abuse; anorexia or bulimia nervosa; initiation of psychotherapy or behaviour therapy 3 months prior to trial entry; and antidepressant or anxiolytic medication in 2 weeks prior to trial entry; neuroleptic or stimulant medication within 6 months of trial entry Exclusion of suicidality: suicide risk or previous active attempt in previous year or hospitalised due to attempt
Interventions	Intervention group Drug: citalopram Dosage: 20 mg to 40 mg Regimen: 20 mg daily for 4 weeks with option to increase to 40 mg daily Length of treatment: 8 weeks Control group: placebo pill
Outcomes	Definition of response and assessment: we used what they call response (called remission in other trials) CDRS-R $\leq$ 28 (they use responders defined as at least = 28 on Children's Depression Rating Scale - Revised (CDRS-R)) Depressive symptoms: CDRS-R Functioning: Children's Global Assessment Scale (CGAS) Suicidal behaviours: FDA data; no report of continuous outcome Adverse events Other outcomes: Clinical Global Impressions Scale Improvement (CGI - Improvement) ; Clinical Global Impressions Scale Severity (CGI - Severity)
Notes	Additional data were sought from authors. No response was received. MHRA # CIT-MD-18 MHRA contacted for additional data some of which were provided Type of data used for remission/response: last observation carried forward

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	"randomly assigned" but no statement how
Allocation concealment (selection bias)	Unclear risk	Unclear
Blinding (performance bias and detection bias) Intervention blinded	Unclear risk	No statement
Blinding (performance bias and detection bias) Blinded outcome	Unclear risk	"in a double-blind fashion" pg 1080; different colour coating was used for placebo and citalopram pills with 9 patients were dispensed medication that potentially unblinded treatment assignment

Incomplete outcome data (attrition bias) All outcomes	High risk	Number eligible: 178 Number randomised: citalopram: 93; placebo: 85; total: 178 Number started trial: citalopram: 93; placebo: 85; total: 178 Number of withdrawals: citalopram: 4; placebo: 0; total: 36 Number analysed post intervention: citalopram: 89; placebo: 85; total: 174 Reasons for drop out: 4 patients all randomly assigned to citalopram group were lost to follow-up and did not receive trial medication. "These patients were not included in the Intention-to-Treat (ITT) analysis "...of these (ITT population) 18 patients from each group discontinued double-blind treatment prematurely pg 1080. Reasons for drop out are not described ITT analysis: "These (4 patients in the citalopram group who were lost to follow-up) patients were not included in the Intention-to-Treat (ITT) analysis"...of these (ITT population) Statistical methods: analysis of covariance with treatment, trial centre, and age as factors and baseline scores as covariate. Cochrane-Mantel Haenszel test controlling for centre and age group. Used LOCF
Selective reporting (reporting bias)	High risk	Percentages only given for response data. Response in this trial is defined in the same way as remission is defined in many other SSRI trials (TADS, Emslie 1997 and 2002 Emslie 2006) but remission itself is not included as an outcome in this trial. Depression symptom severity means and standard deviations were not reported but represented in a figure with a result only reported (MHRA report change scores)
Other bias	Unclear risk	Contact: evaluation undertaken at 1, 2, 4, 6 and 8 weeks. Psychotherapy was not allowed pg 1080 Screening: there was 1 screening visit and then a baseline visit Placebo lead-in: 1 week single-blind in between screening visit and baseline visit Baseline imbalance: authors report no sig-

		nificant differences (report data in Table 1) Other: data not reported by child versus adolescent
--	--	---

Wagner 2006

Methods	Trial design: randomised controlled trial; multicentre Power calculation: not stated Use of diagnostic criteria (or clear specification of inclusion criteria): yes Intervention integrity: not described Outcome measures described or validated measures used: yes Follow-up assessment points: post assessment No. crossed over: none Funded by: Forest Laboratories, Inc
Participants	Setting of care: outpatients Recruitment: no information Mean age: intervention = 12.2 (2.9); control = 12.4 (3.0) Age range: 6 to 17 years Gender (F:M): intervention = 68:63; control = 69:64 Methods used to diagnose: DSM-IV confirmed using K-SADS-PL and a CDRS-R score of ≥ 40; 1-week placebo run-in period Diagnosis: MDD Baseline severity of depression: CDRS-R mean score: intervention = 54.5; control = 56.6; CGI intervention 4.4; CGI placebo 4.2 Length of current episode: (mean months) intervention 16.7 (15.3); placebo 15.6 (13.6) % first episode: not reported Co-morbidity (intervention): 6 had an ongoing anxiety disorder; none had ADHD Co-morbidity (control): 10 had an ongoing anxiety disorder; none had ADHD Location: 25 centres in the USA Inclusion criteria: MDD of at least a 4-week duration, normal results at screening from physical examination, laboratory tests and electrocardiography Exclusion criteria: any primary psychiatric diagnosis apart from MDD; any psychotic features; any severe personality disorder; met DSM-IV criteria for Attention Deficit Hyperactivity Disorder, Post Traumatic Stress Disorder, bipolar disorder, Pervasive Developmental Disorder, mental retardation, conduct or oppositional defiant disorder; females not practising or willing to practise a reliable method of birth control; history of Anorexia Nervosa, Bulimia Nervosa, substance abuse; initiation of psychotherapy was not allowed during the trial of within 3 months before the screening visit; previous treatment failure on SSRI Exclusion of suicidality: suicide risk based on clinical judgement of investigator or ever hospitalised for suicide attempt or had made a suicide attempt within the past year
Interventions	Intervention group Drug: escitalopram oxalate Dosage: fixed dose of 10 mg for the first 4 weeks; thereafter flexibly dosed from 10 to 20 mg based on clinical response Regimen: taken daily

	Length of treatment: 8 weeks Control group: placebo pill	
Outcomes	Definition and assessment of response: we used what they call response (called remission in other trials) CDRS-R $\leq$ 2 (they did 2 separate analyses of response data were undertaken using 2 different definitions of response: CDRS-R score of less than or equal to 28; or CGI-I of less than or equal to 2) Depressive symptoms: Children’s Depression Rating Scale - Revised (CDRS-R) Functioning: Children’s Global Assessment Scale (C-GAS) Suicidal behaviour: events reported as adverse events; no report of continuous measure Adverse outcomes Other outcomes: Clinical Global Impressions Scale Severity (CGI-Severity); Clinical Global Impressions Scale Improvement (CGI-Improvement)	
Notes	Forest pharmaceutical ID is SCT MD 15 Data in the MA from the web-based publication. Subsequent to this Wagner 2006 was published and data checked against this publication with child and adolescent data added to the MA Type of data used for remission/response: last observation carried forward	
<i>Risk of bias</i>		
Bias	Authors’ judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	pg 282 computer-generated randomisation sequence. Patient randomisation numbers were allocated to each site in ascending sequence in blocks of 4. Randomisation was not stratified by age
Allocation concealment (selection bias)	Unclear risk	No statement
Blinding (performance bias and detection bias) Intervention blinded	Unclear risk	Stated to be “double blind” with tablets identical indicating participants may be blinded
Blinding (performance bias and detection bias) Blinded outcome	Unclear risk	No statement but clinicians and subjects completed measures and both of these were probably blind
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Number eligible: 268 Number randomised: escitalopram: 132; placebo: 136; total: 268 Number started trial: escitalopram: 131; placebo: 133; total: 264 Number of withdrawals: escitalopram: 29; placebo: 18; total: 48 Number analysed post intervention: escitalopram: 129; placebo: 132; total: 261

Wagner 2006 (Continued)

		Reasons for drop out: full list of drop outs and reasons for drop out figure 1 pg 283. Trial authors state no significant differences in specific reasons for premature discontinuation; appear to be more withdrawing consent from escitalopram group ITT analysis: efficacy analyses were performed on the intent-to-treat population, which included all patients in the safety population (i.e. received at least 1 dose of trial medication) who had at least 1 post-baseline CDRS-R assessment pg 282 Statistical methods: LOCF was used (as well as some OC analysis). Analysis of covariance (treatment group and trial centre as factors and baseline scores as covariate) . Logistic regression with treatment as the factor and baseline scores as covariate pg 282
Selective reporting (reporting bias)	Unclear risk	2 prospective definitions of response were used. A post hoc analysis of suicide-related outcomes was undertaken (pg 282). Only P values are provided for clinician-rated depression symptoms
Other bias	Low risk	Contact: evaluations at end of 1, 2, 4, 6 and 8 weeks; psychotherapy was not allowed (pg 281) Screening: diagnostic criteria have to be met at the screening visit and then again at the baseline visit after the 1-week placebo lead-in Baseline imbalance: authors state there were no significant differences between the groups Other: not noted

Wagner Trial 1

Methods	See Wagner Trial 1&2 (2003) entry
Participants	See Wagner Trial 1&2 (2003) entry
Interventions	See Wagner Trial 1&2 (2003) entry
Outcomes	See Wagner Trial 1&2 (2003) entry

Wagner Trial 1 (Continued)

Notes	See Wagner Trial 1&2 (2003) entry	
<i>Risk of bias</i>		
Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"using a computer generated randomisation code" pg 1034
Allocation concealment (selection bias)	Unclear risk	No statement
Blinding (performance bias and detection bias) Intervention blinded	Low risk	"double blind receipt of sertraline or matching placebo" pg 1034; "trial drug was packaged in identical blister packs...both patients and clinicians were blinded to group assignment" pg 1035
Blinding (performance bias and detection bias) Blinded outcome	Unclear risk	No statement
Incomplete outcome data (attrition bias) All outcomes	Low risk	Number randomised: 188 Number of withdrawals: 46 Number analysed post intervention: 142 Reasons for drop out: full list of drop outs and reasons for drop outs figure 1 pg 1036. There were more drop outs due to adverse events reported in the sertraline group ITT analysis: intention-to-treat population was modified... post randomisation efficacy data collected...problems with data collection pg 1036. Only those who received at least 1 dose of trial medication were included in the efficacy analyses pg 1036 Statistical methods: used repeated measures mixed-model analysis with the model including baseline effect as a covariate, random subject effect and fixed-effect of site, treatment, age group, week and week by treatment interaction. Response data were analysed using Cochrane-Mantel Haenszel methods with centres as strata. Last observation carried forward (LOCF) analysis for responder outcome but not clear for Child Depression Rating Scale-Revised (CDRS-R). LOCF data were used in analysis of covariance with treatment group, age and baseline effects as covariates
Selective reporting (reporting bias)	High risk	Data are not given separately for each individual trial. The trial reports several response data sets,

Wagner Trial 1 (Continued)

		some weekly data and they looked at individual items in their measures. While the paper does not report on remission as an outcome the MHRA report does have these data by group. Response data are given as percentages in the paper and these data do not match MHRA data. Denominators for response and remission in the MHRA data are different. They do not report total adverse event rate
Other bias	High risk	Contact: authors state there were “frequent follow-up visits” pg 1039 and regular measurements taken. They were also allowed to receive therapy pg 1035 Screening: diagnostic criteria have to be met at the first and third visits during a 2-week screening period (total of 3 visits in the screening period) Placebo lead-in: no Baseline imbalance: authors state there were no differences between the groups except for gender (more females than males) Other: this is mostly a first episode population; there were 2 studies reported in the one paper; trial 2 had much higher response and remission rates, but data are not reported separately in the published paper

Wagner Trial 1&2 (2003)

Methods	Trial design: randomised controlled trial; multicentre Power calculation: yes Use of diagnostic criteria (or clear specification of inclusion criteria): yes Intervention integrity: not described Outcome measures described or validated measures used: yes Follow-up assessment points: post intervention No. crossed over: none Funded by: Pfizer
Participants	Setting of care: outpatient Recruitment: no information Mean age: not stated for either group Age range: 6 to 17 years Gender (F:M): intervention = 108:81; control = 84:103 Methods used to diagnose: DSM-IV confirmed using K-SADS-PL, a CDRS-R score of ≥ 45 and a CGI-S score of ≥ 4 During 2-week screen had to meet these criteria at first and third visit Diagnosis: MDD Baseline severity of depression: CDRS-R mean (SD) score intervention = 64.3 (11.0); control = 64.6 (11.0); CGI intervention 4.6 (0.6); CGI placebo 4.5 (0.7)

	Length of current episode: not reported % first episode: intervention 95%; placebo 95% Co-morbidity (intervention and control): 40% of participants had at least 1 co-morbid condition; the conditions that occurred in at least 5% of patients included anxiety; phobic disorder; adjustment reaction; ODD Location: USA, India, Canada, Costa Rica, Mexico Inclusion criteria: outpatients; aged 6 to 17; MDD at the first and third visits during a 2-week screen and current episode had to be of at least 6 weeks duration; illness of at least moderate severity Exclusion criteria: Attention Deficit Hyperactivity Disorder; Conduct Disorder; Obsessive Compulsive Disorder; panic disorder; history of bipolar or current psychotic features; history of psychotic disorders or autistic spectrum disorders; current anorexia nervosa or bulimia nervosa; drug or alcohol abuse/dependence within 6 months or current positive drug screen; pregnant or breast feeding; abnormal Electrocardiography (ECG), laboratory test results, vital signs or body weight; current use of other psychotropic medication; intention to commence psychotherapy; requirement of concomitant psychotropic therapy; previous failed response to an SSRI; additionally trial 2 stated it excluded those requiring inpatient admission Exclusion of suicidality: previous suicide attempt or current significant suicidal or homicidal risk
Interventions	Intervention group Drug: sertraline Dosage: flexible dosage 25 to 200 mg Regimen: 25 mg for 3 days; 50 mg till the end of the second week; increases as indicated by 50 mg per day to a maximum of 200 mg Length of treatment: 10 weeks Control group: placebo pill
Outcomes	Definition and assessment of response: we used OC remission: subjects who no longer met DSM-IV criteria for a current major depression episode at endpoint from MHRA (they used responders defined as at least 40% decrease on Children's Depression Rating Scale - Revised (CDRS-R)) Depression symptoms: Childrens Depression Rating Scale (CDRS-R) Functioning: Childrens Global Assessment Scale (C-GAS) Suicidal behaviour: events reported as adverse events; no report of continuous outcome Other outcomes: Clinical Global Impressions Scale Severity (CGI-Severity); Clinical Global Impressions Scale Improvement (CGI - Improvement); clinician-rated severity; Multidimensional Anxiety Scale for Children (MASC); Pediatric Quality of Life Enjoyment and Satisfaction Questionnaire (PQ-LES-Q); adverse events
Notes	Additional data were sought from authors. No response was received. MHRA contacted for additional data for #1001 and 1017, some of which were provided MHRA data used in MA as it gave data for each separate trial and separately for child and adolescent Type of data used for remission/response: last observation carried forward
<i>Risk of bias</i>	

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"using a computer generated randomisation code" pg 1034
Allocation concealment (selection bias)	Unclear risk	No statement
Blinding (performance bias and detection bias) Intervention blinded	Low risk	"double blind receipt of sertraline or matching placebo" pg 1034; "trial drug was packaged in identical blister packs... both patients and clinicians were blinded to group assignment" pg 1035
Blinding (performance bias and detection bias) Blinded outcome	Unclear risk	No statement
Incomplete outcome data (attrition bias) All outcomes	Low risk	Number eligible: 376 Number randomised: sertraline: 189; placebo: 187; total: 376 Number started trial: sertraline: 189; placebo: 187; total: 376 Number of withdrawals: sertraline: 46; placebo: 31; total: 77 Number analysed post intervention: sertraline: 185; placebo: 179; total: 364 <u>Trial 1</u> Number randomised: 188 Number of withdrawals: 46 Number analysed post intervention: 142 <u>Trial 2</u> Number randomised: 188 Number of withdrawals: 31 Number analysed post intervention: 157 Reasons for drop out: full list of drop outs and reasons for drop out figure 1 pg 1036. There were more drop outs due to adverse events reported in the sertraline group ITT analysis: intention-to-treat population was modified... post randomisation efficacy data collected...problems with data collection pg 1036. Only those who received at least one dose of trial medication were included in the efficacy analyses pg 1036 Statistical methods: used repeated measures mixed-model analysis with the model including baseline effect as a covariate, random subject effect and fixed-effect of site,

		treatment, age group, week and week by treatment interaction. Response data were analysed using Cochrane-Mantel Haenszel methods with centres as strata. Last observation carried forward (LOCF) analysis for responder outcome but not clear for Child Depression Rating Scale-Revised (CDRS-R). LOCF data were used in analysis of covariance with treatment group, age and baseline effects as covariates
Selective reporting (reporting bias)	High risk	Data are not given separately for each individual trial. The trial reports several response data sets, some weekly data and they looked at individual items in their measures. While the paper does not report on remission as an outcome the MHRA report does have these data by group. Response data are given as percentages in the paper and these data do not match MHRA data. Denominators for response and remission in the MHRA data are different. They do not report total adverse event rate
Other bias	High risk	Contact: authors state there were "frequent follow-up visits" pg 1039 and regular measurements taken. They were also allowed to receive therapy pg 1035 Screening: diagnostic criteria has to be met at the first and third visits during a 2-week screening period (total of 3 visits in the screening period) Placebo lead-in: no Baseline imbalance: authors state there were no differences between the groups except for gender (more females than males) Other: this is mostly a first episode population; there were 2 studies reported in the 1 paper; trial 2 had much higher response and remission rates, but data are not reported separately in the published paper

Wagner Trial 2

Methods	See Wagner Trial 1&2 (2003) entry
Participants	See Wagner Trial 1&2 (2003) entry
Interventions	See Wagner Trial 1&2 (2003) entry
Outcomes	See Wagner Trial 1&2 (2003) entry
Notes	See Wagner Trial 1&2 (2003) entry

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"using a computer generated randomisation code" pg 1034
Allocation concealment (selection bias)	Unclear risk	No statement
Blinding (performance bias and detection bias) Intervention blinded	Low risk	"double blind receipt of sertraline or matching placebo" pg 1034; "trial drug was packaged in identical blister packs...both patients and clinicians were blinded to group assignment" pg 1035
Blinding (performance bias and detection bias) Blinded outcome	Unclear risk	No statement
Incomplete outcome data (attrition bias) All outcomes	Low risk	Number randomised: 188 Number of withdrawals: 31 Number analysed post intervention: 157 Reasons for drop out: full list of drop outs and reasons for drop out figure 1 pg 1036. There were more drop outs due to adverse events reported in the sertraline group ITT analysis: intention-to-treat population was modified... post randomisation efficacy data collected...problems with data collection pg 1036. Only those who received at least one dose of trial medication were included in the efficacy analyses pg 1036 Statistical methods: used repeated measures mixed-model analysis with the model including baseline effect as a covariate, random subject effect and fixed-effect of site, treatment, age group, week and week by treatment interaction. Response data were analysed using Cochrane-Mantel Haenszel methods with centres as strata. Last

		observation carried forward (LOCF) analysis for responder outcome but not clear for Child Depression Rating Scale-Revised (CDRS-R). LOCF data were used in analysis of covariance with treatment group, age and baseline effects as covariates
Selective reporting (reporting bias)	High risk	Data are not given separately for each individual trial. The trial reports several response data sets, some weekly data and they looked at individual items in their measures. While the paper does not report on remission as an outcome the MHRA report does have these data by group. Response data are given as percentages in the paper and these data do not match MHRA data. Denominators for response and remission in the MHRA data are different. They do not report total adverse event rate
Other bias	High risk	Contact: authors state there were “frequent follow-up visits” pg 1039 and regular measurements taken. They were also allowed to receive therapy pg 1035 Screening: diagnostic criteria has to be met at the first and third visits during a 2-week screening period (total of 3 visits in the screening period) Placebo lead-in: no Baseline imbalance: authors state there were no differences between the groups except for gender (more females than males) Other: this is mostly a first episode population; there were 2 studies reported

ADD: attention deficit disorder; ADHD: attention deficit hyperactivity disorder; AN: Anorexia Nervosa; BDI: Beck Depression Inventory; BN: Bulimia Nervosa; BPRS-C: Brief Psychiatry Rating Scale - Children's ; CD: Conduct Disorder; CDI: Children's Depression Inventory; CDRS-R: Children's Depression Rating Scale - Revised; C-GAS: Children's Global Assessment Scale; CGI: Clinical Global Impressions Scale; CGI-I: Clinical Global Impressions Scale - Improvement; CGI-S: Clinical Global Impressions Scale - Severity; CNS: central nervous system; DICA: Diagnostic Interview for Children and Adolescents; DSM-IV-TR: Diagnostic and Statistical Manual of Mental Disorders IV - Text Revision; DSRS: Depression Self Assessment Scale; ECT: electroconvulsive therapy; FSIQ: Full Scale Intelligence Quotient; GAS: Global Assessment Scale; GAD: Generalised Anxiety Disorder; GSK: GlaxoSmithKline; HAMA: Hamilton Anxiety Rating Scale; HAM-D: Hamilton Rating Scale for Depression; HARS: Hamilton Anxiety Rating Scale; HDRS: Hamilton Depression Rating Scale; ITT: intention-to-treat; K-SADS: Schedule for Affective Disorders and Schizophrenia for School Aged Children; LOCF: last observation carried forward; MA: meta-analysis; MADRS: Montgomery-Asberg Depression Rating Scale; MC-SSRS: Modified Columbia Suicide Severity Rating Scale; MDD: major depressive disorder; MHRA: Committee on Safety of Medicines, Medicines and Healthcare Products Regulatory Agency; MINI KIDS: The Mini International Neuropsychiatric Interview for children and adolescents ; NIMH: National Institutes for Mental Health; OC: observed case; OCD: obsessive-compulsive disorder; ODD: Oppositional Defiant Disorder; PDD: Pervasive Developmental Disorder; PTSD: post-traumatic stress disorder; SD: standard deviation; SE: standard error; SIQ-Jr: Suicidal Ideation Questionnaire-Junior High School Version; TADS: Treatment for Adolescents with Depression (study); WSAS: Weinberg Screening Affective Scale

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

Study	Reason for exclusion
Braconnier 2003	Comparison is not placebo; paroxetine is compared with clomipramine
Cornelius 2009	No pure fluoxetine or placebo treatment arm
Cornelius 2010	No pure fluoxetine or placebo treatment arm
Cosgrove 1994	Case trial design
Findling 2009	Focus of the intervention was co-morbid substance use rather than depression
Mandoki 1997	Comparison of venlafaxine plus psychotherapy with placebo and psychotherapy
NIMH 2000	Primary diagnosis of bipolar disorder; trial discontinued
Riggs 2007	No pure fluoxetine or placebo treatment arm
Sallee 1997	Antidepressant not on our list of included compounds
Tashakori 1997	Not a RCT
Wohlfarth 2007	Not a RCT

RCT: randomised controlled trial

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

Duloxetine NCT00849693

Trial name or title	Duloxetine NCT00849693
Methods	Trial design: randomised controlled trial; multicentre Power calculation: not stated Use of diagnostic criteria (or clear specification of inclusion criteria): yes Intervention integrity: NA Outcome measures described or validated measures used: yes Follow-up assessment points: 10 weeks; 36 weeks No. crossed over: none Funded by: Eli Lilly
Participants	Setting of care: outpatient Recruitment: not stated

Duloxetine NCT00849693 (Continued)

	Mean age: NA Age range: 7 to 17 years Gender (F:M): NA Methods used to diagnose: DSM-IV-TR confirmed using MINI-KID; CDRS-R ≥ 45 and CGI ≥ 4 Diagnosis: MDD Baseline severity of depression: NA Co-morbidity intervention and control: NA Location: USA, Canada, Mexico Inclusion criteria: moderate or greater severity of MDD as determined by CDRS-R ≥ 40 at screen, and randomisation and a CGI-Severity rating ≥ 4 at screen, and randomisation Exclusion criteria: current or previous diagnosis of bipolar disorder, psychotic depression, schizophrenia or other psychotic disorder, anorexia, bulimia, obsessive compulsive disorder, or pervasive development disorder, as judged by the investigator; history of DSM-IV-TR defined substance abuse or dependence within the past year; current primary DSM-IV-TR Axis I disorder other than MDD or a current secondary DSM-IV-TR Axis I disorder that requires any pharmacologic treatment; 1 or more first-degree relatives with diagnosed bipolar I disorder; significant suicide attempt within 1 year of screening or are currently at risk of suicide in the opinion of the investigator; significant suicide attempt within 1 year of screening or are currently at risk of suicide in the opinion of the investigator
Interventions	Intervention group Drug arm 1: duloxetine Dosage: 30 to 120 mg Regimen: once daily Length of treatment: 38 weeks Drug arm 2: duloxetine Dosage: 30 mg Regimen: once daily Length of treatment: 38 weeks Control group: placebo Comparison group: fluoxetine 10 to 40 mg Evaluations at baseline, 10 weeks, 38 weeks
Outcomes	Definition and assessment of response: not stated Depressive symptoms: CDRS-R Functioning: not stated Adverse outcomes: Columbia-Suicide Severity Rating Scale (C-SSRS); hepatic laboratory results, systolic blood pressure, diastolic blood pressure, pulse and weight Other outcomes: CGI-S
Starting date	
Contact information	
Notes	

Duloxetine NCT00849901

Trial name or title	Duloxetine NCT00849901
Methods	Trial design: randomised controlled trial Power calculation: not stated Use of diagnostic criteria (or clear specification of inclusion criteria): yes Intervention integrity: NA Outcome measures described or validated measures used: yes Follow-up assessment points: 10 weeks; 36 weeks No. crossed over: none Funded by: Eli Lilly
Participants	Setting of care: outpatient Recruitment: not stated Mean age: NA Age range: 7 to 17 years Gender (F:M): NA Methods used to diagnose: DSM-IV TR confirmed using MINI-KID; CDRS-R ≥ 45 and CGI ≥ 4 Diagnosis: MDD Baseline severity of depression: NA Co-morbidity intervention and control: NA Location: USA, Canada, Mexico Inclusion criteria: moderate or greater severity of MDD as determined by CDRS-R ≥ 40 at screen, and randomisation and a CGI-Severity rating ≥ 4 at screen, and randomisation Exclusion criteria: current or previous diagnosis of bipolar disorder, psychotic depression, schizophrenia or other psychotic disorder, anorexia, bulimia, obsessive compulsive disorder, or pervasive development disorder, as judged by the investigator; history of DSM-IV-TR-defined substance abuse or dependence within the past year; current primary DSM-IV-TR Axis I disorder other than MDD or a current secondary DSM-IV-TR Axis I disorder that requires any pharmacologic treatment; 1 or more first-degree relatives with diagnosed bipolar I disorder; significant suicide attempt within 1 year of screening or are currently at risk of suicide in the opinion of the investigator; significant suicide attempt within 1 year of screening or are currently at risk of suicide in the opinion of the investigator
Interventions	Intervention group Drug: duloxetine Dosage: 30 to 120 mg Regimen: once daily Length of treatment: 36 weeks Control group: placebo Comparison group: fluoxetine 10 to 40 mg Evaluations at baseline, 10 weeks, 36 weeks
Outcomes	Definition and assessment of response: Depressive symptoms: CDRS-R Functioning: Adverse outcomes: Columbia-Suicide Severity Rating Scale (C-SSRS), hepatic laboratory results, systolic blood pressure, diastolic blood pressure, pulse and weight Other outcomes: CGI
Starting date	

Contact information	
Notes	

Glod 2004

Trial name or title	Glod 2004
Methods	Trial design: randomised controlled trial Power calculation: not stated Use of diagnostic criteria (or clear specification of inclusion criteria): Intervention integrity: not stated. Outcome measures described or validated measures used: Follow-up assessment points: not stated No. crossed over: not stated Funded by: not stated
Participants	Setting of care: outpatient Recruitment: not stated Mean age: 15.5 years (1.9) Age range: 12 to 19 years Gender (F:M): 12:6 Methods used to diagnose: semi-structured clinical interview (K-SADS-E) Diagnosis: DSM-IV defined MDD Baseline severity of depression: 20.3 (3.7) on the Hamilton Depression Rating Scale Co-morbidity intervention and control: not stated Location: not stated Inclusion criteria: MDD; no further details stated Exclusion criteria: not stated.
Interventions	Intervention group: citalopram Drug arm 1: not stated Dosage: not stated Regimen: not stated Length of treatment: 8 weeks Drug arm 2: bupropion Dosage: not stated Regiman: not stated Length of treatment: 8 weeks Control group: placebo Evaluations at: not stated
Outcomes	Depressive symptoms: change in Hamilton Depression Rating Scale score
Starting date	
Contact information	

Notes	Dr Carol Glod contacted on 28 November 2011 for additional information
-------	--

Solvay NCT00353028

Trial name or title	Solvay NCT00353028
Methods	Trial design: randomised controlled trial; multicentre Power calculation: not stated Use of diagnostic criteria (or clear specification of inclusion criteria): yes Intervention integrity: NA Outcome measures described or validated measures used: yes Follow-up assessment points: 8 weeks No. crossed over: not stated Funded by: Solvay Pharmaceuticals
Participants	Setting of care: not stated Recruitment: not stated Mean age: not stated Age range: 8 to 18 years Gender (F:M): NA Methods used to diagnose: the Japanese Version of the Structured Interview Guide for the Hamilton Depression Rating Scale (JSIGH-D) 17-item total score Diagnosis: depression or depressive state Baseline severity of depression: not stated Co-morbidity intervention and control: NA Location: Japan Inclusion criteria: a minimum total score of 18 on the JSIGH-D, weight within the standard weight ± 2 SD based on the standard weight for each age in the School Health Statistical Survey Exclusion criteria: predominant psychiatric diagnosis - schizophrenia, or previously been treated with fluvoxamine maleate
Interventions	Intervention group: Drug arm 1: fluvoxamine maleate Dosage: 25 mg to 150 mg (1 to 6 tablets) Regimen: once daily Length of treatment: 8 weeks Control group: placebo Evaluations at baseline, 8 weeks
Outcomes	Depressive symptoms: time of onset of 50% decrease from baseline in the JSIGH-D Functioning: the Clinical Global Impression scale (CGI)
Starting date	
Contact information	

Notes	Contacted Toshiaki Yamaguchi, trial director at Solvay Pharmaceuticals on 13 October 2011 regarding trial status, however no reply received at time of publication
-------	--

CDRS-R: Children's Depression Rating Scale - Revised; CGI: Clinical Global Impressions Scale; DSM-IV-TR: Diagnostic and Statistical Manual of Mental Disorders IV - Text Revision; K-SADS: Schedule for Affective Disorders and Schizophrenia for School Aged Children; JSIGH-D: Japanese Version of the Structured Interview Guide for the Hamilton Depression Rating Scale ; MDD: major depressive disorder; NA: not applicable; OC: observed case; SD: standard deviation

DATA AND ANALYSES

Comparison 1. Newer generation antidepressant versus placebo (by drug)

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Depressive symptom severity (CDRS-R)	14		Mean difference (Random, 95% CI)	-3.51 [-4.55, -2.47]
1.1 Paroxetine	2		Mean difference (Random, 95% CI)	-1.18 [-6.29, 3.92]
1.2 Fluoxetine	3		Mean difference (Random, 95% CI)	-5.63 [-7.39, -3.86]
1.3 Sertraline	2		Mean difference (Random, 95% CI)	-3.52 [-6.64, -0.40]
1.4 Citalopram	1		Mean difference (Random, 95% CI)	-2.9 [-7.77, 1.97]
1.5 Escitalopram	2		Mean difference (Random, 95% CI)	-2.67 [-4.85, -0.48]
1.6 Venlafaxine	2		Mean difference (Random, 95% CI)	-1.90 [-4.79, 0.99]
1.7 Mirtazapine	2		Mean difference (Random, 95% CI)	-2.79 [-6.42, 0.83]
2 Remission or response (as defined in trial) LOCF only	15	2924	Risk Ratio (M-H, Random, 95% CI)	1.18 [1.08, 1.28]
2.1 Paroxetine	4	704	Risk Ratio (M-H, Random, 95% CI)	1.12 [0.90, 1.38]
2.2 Fluoxetine	4	543	Risk Ratio (M-H, Random, 95% CI)	1.47 [1.03, 2.08]
2.3 Sertraline	1	364	Risk Ratio (M-H, Random, 95% CI)	1.17 [1.00, 1.36]
2.4 Citalopram	2	407	Risk Ratio (M-H, Random, 95% CI)	1.16 [0.71, 1.89]
2.5 Escitalopram	2	572	Risk Ratio (M-H, Random, 95% CI)	1.19 [0.97, 1.45]
2.6 Venlafaxine	2	334	Risk Ratio (M-H, Random, 95% CI)	1.16 [1.00, 1.35]
3 Remission or response (as defined in trial) mix OC and LOCF	16	2662	Risk Ratio (M-H, Random, 95% CI)	1.17 [1.09, 1.25]
3.1 Paroxetine OC	4	522	Risk Ratio (M-H, Random, 95% CI)	1.11 [0.98, 1.27]
3.2 Fluoxetine	4	543	Risk Ratio (M-H, Random, 95% CI)	1.47 [1.03, 2.08]
3.3 Sertraline OC	2	364	Risk Ratio (M-H, Random, 95% CI)	1.18 [1.01, 1.38]
3.4 Citalopram OC (von Knorring only)	2	327	Risk Ratio (M-H, Random, 95% CI)	1.17 [0.74, 1.85]
3.5 Escitalopram	2	572	Risk Ratio (M-H, Random, 95% CI)	1.19 [0.97, 1.45]
3.6 Venlafaxine	2	334	Risk Ratio (M-H, Random, 95% CI)	1.16 [1.00, 1.35]
4 Depressive symptom severity self rated	3		Mean difference (Random, 95% CI)	-0.53 [-2.37, 1.31]
4.1 Paroxetine	1		Mean difference (Random, 95% CI)	-0.43 [-2.91, 2.05]
4.2 Fluoxetine	1		Mean difference (Random, 95% CI)	-1.3 [-5.87, 3.27]
4.3 Citalopram	1		Mean difference (Random, 95% CI)	-0.28 [-3.72, 3.16]
5 Functioning CGAS	7		Mean difference (Random, 95% CI)	2.20 [0.90, 3.49]
5.1 Paroxetine	1		Mean difference (Random, 95% CI)	1.6 [-2.48, 5.68]
5.2 Fluoxetine	2		Mean difference (Random, 95% CI)	3.08 [0.14, 6.02]
5.3 Sertraline	1		Mean difference (Random, 95% CI)	1.31 [-1.61, 4.23]
5.4 Citalopram	1		Mean difference (Random, 95% CI)	2.5 [-1.52, 6.52]
5.5 Escitalopram	2		Mean difference (Random, 95% CI)	2.28 [0.23, 4.32]
6 Suicide-related outcome	14	3229	Risk Ratio (M-H, Random, 95% CI)	1.58 [1.02, 2.45]
6.1 Paroxetine	4	702	Risk Ratio (M-H, Random, 95% CI)	1.57 [0.46, 5.31]
6.2 Fluoxetine	3	536	Risk Ratio (M-H, Random, 95% CI)	1.77 [0.85, 3.69]
6.3 Sertraline	1	376	Risk Ratio (M-H, Random, 95% CI)	2.97 [0.61, 14.52]
6.4 Citalopram	2	418	Risk Ratio (M-H, Random, 95% CI)	1.53 [0.55, 4.27]

6.5 Escitalopram	2	575	Risk Ratio (M-H, Random, 95% CI)	0.91 [0.47, 1.76]
6.6 Venlafaxine	1	367	Risk Ratio (M-H, Random, 95% CI)	12.93 [1.71, 97.82]
6.7 Mirtazapine	1	255	Risk Ratio (M-H, Random, 95% CI)	0.5 [0.03, 7.90]
7 Suicidal ideation SIQ-JR	2		Mean difference (Random, 95% CI)	0.70 [-1.50, 2.91]
7.1 Paroxetine	0		Mean difference (Random, 95% CI)	0.0 [0.0, 0.0]
7.2 Fluoxetine	1		Mean difference (Random, 95% CI)	-0.57 [-3.49, 2.35]
7.3 Sertraline	0		Mean difference (Random, 95% CI)	0.0 [0.0, 0.0]
7.4 Citalopram	0		Mean difference (Random, 95% CI)	0.0 [0.0, 0.0]
7.5 Escitalopram	1		Mean difference (Random, 95% CI)	1.7 [-0.77, 4.17]
7.6 Venlafaxine	0		Mean difference (Random, 95% CI)	0.0 [0.0, 0.0]
8 Adverse events	10	2136	Risk Ratio (M-H, Random, 95% CI)	1.11 [1.05, 1.17]
8.1 Paroxetine	4	714	Risk Ratio (M-H, Random, 95% CI)	1.11 [0.98, 1.25]
8.2 Fluoxetine	2	440	Risk Ratio (M-H, Random, 95% CI)	1.19 [1.05, 1.35]
8.3 Citalopram	2	407	Risk Ratio (M-H, Random, 95% CI)	1.13 [1.00, 1.29]
8.4 Escitalopram	2	575	Risk Ratio (M-H, Random, 95% CI)	1.03 [0.94, 1.14]
9 Completion of trial protocol	16	3290	Risk Ratio (M-H, Random, 95% CI)	0.99 [0.94, 1.05]
9.1 Paroxetine	4	728	Risk Ratio (M-H, Random, 95% CI)	0.94 [0.86, 1.02]
9.2 Fluoxetine	4	559	Risk Ratio (M-H, Random, 95% CI)	1.16 [0.97, 1.40]
9.3 Sertraline	1	376	Risk Ratio (M-H, Random, 95% CI)	0.91 [0.82, 1.01]
9.4 Citalopram	2	422	Risk Ratio (M-H, Random, 95% CI)	1.03 [0.91, 1.16]
9.5 Escitalopram	2	584	Risk Ratio (M-H, Random, 95% CI)	0.93 [0.86, 1.01]
9.6 Venlafaxine	1	367	Risk Ratio (M-H, Random, 95% CI)	0.93 [0.82, 1.07]
9.7 Mirtazapine	2	254	Risk Ratio (M-H, Random, 95% CI)	1.05 [0.93, 1.19]

Comparison 2. Newer generation antidepressant versus placebo (child vs adolescent)

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Depressive symptom severity (CDRS-R)	10		Mean difference (Random, 95% CI)	-3.71 [-4.93, -2.48]
1.1 Child	8		Mean difference (Random, 95% CI)	-3.18 [-5.95, -0.41]
1.2 Adolescent	10		Mean difference (Random, 95% CI)	-4.21 [-5.50, -2.92]
2 Remission or response (as defined in trial)	7	1368	Risk Ratio (M-H, Random, 95% CI)	1.19 [1.05, 1.35]
2.1 Child	2	187	Risk Ratio (M-H, Random, 95% CI)	1.71 [0.90, 3.25]
2.2 Adolescents	7	1181	Risk Ratio (M-H, Random, 95% CI)	1.13 [1.02, 1.26]
3 Depressive symptom severity self rated	1		Mean difference (Random, 95% CI)	-0.43 [-3.09, 2.23]
3.1 Adolescent	1		Mean difference (Random, 95% CI)	-0.43 [-3.09, 2.23]
4 Functioning CGAS	5		Mean difference (Random, 95% CI)	2.49 [0.85, 4.12]
4.1 Child	2		Mean difference (Random, 95% CI)	1.41 [-6.74, 9.57]
4.2 Adolescent	5		Mean difference (Random, 95% CI)	2.82 [1.17, 4.47]
5 Suicide-related outcome	10	2056	Risk Ratio (M-H, Random, 95% CI)	1.53 [1.05, 2.22]
5.1 Child	4	389	Risk Ratio (M-H, Random, 95% CI)	2.07 [0.67, 6.35]
5.2 Adolescents	10	1667	Risk Ratio (M-H, Random, 95% CI)	1.47 [0.99, 2.19]
6 Adverse events	6	1423	Risk Ratio (M-H, Random, 95% CI)	1.11 [1.04, 1.18]
6.1 Child	1	96	Risk Ratio (M-H, Random, 95% CI)	1.09 [0.82, 1.44]
6.2 Adolescents	6	1327	Risk Ratio (M-H, Random, 95% CI)	1.11 [1.04, 1.19]

ADDITIONAL TABLES

Table 1. Depressive symptom severity based on K-SADS clinician measure

Paroxetine	Berard 2006	Mean difference -0.41 (95% CI -2.00 to 1.18)
	Keller 2001	Mean difference -2.10 (95% CI -4.37 to 0.17)
Citalopram	Von Knorring 2006	Mean difference 0.28 (95% CI -2.20 to 2.76)

CI: confidence interval

K-SADS: Schedule for Affective Disorders and Schizophrenia for School Aged Children

Table 2. Depressive symptom severity based on RADS

Fluoxetine	TADS 2004	Mean difference -6.10 (95% CI -9.34 to -2.86)
------------	---------------------------	---

CI: confidence interval

Table 3. Functioning based on GAF and Autonomous Functioning Checklist

GAF	Paroxetine	Emslie 2006	Mean difference 1.40 (95% CI -2.44 to 5.24)
	Fluoxetine	Emslie 2002	Mean difference 0.90 (95% CI -2.26 to 4.06)
Autonomous Functioning Checklist	Paroxetine	Keller 2001	Mean difference 5.40 (95% CI -2.29 to 13.09)

GAF: Global Assessment of Functioning

Table 4. Adverse outcomes

Body Group	Specific outcome	Paroxetine (Be-rard 2006)	Paroxetine (Be-rard 2006)	Paroxetine (Keller 2001)	Paroxetine (Em-slie 2006)	Fluoxetine (Em-slie 1997; Em-slie 2002)	Fluoxetine (TADS 2004)	Escitalopram (Em-slie 2009)	Escitalopram (Wag-ner 2006)	Citalopram (Von Knor-ring 2006)	Citalopram (Wag-ner 2004)	Ser-traline (Wag-ner Trial 1&2 (2003))	Ven-lafax-ine (Em-slie 2007) *only includes symptoms reported by more than 10% of sample	Mir-taza-pine
Body as a whole	Ab-domi-nal pain treat-ment		6	10	4		6	14	14	11	10		23	
	Ab-domi-nal pain placebo		9	10	3		2	11	7	6	6		18	
	As-thenia treat-ment		12	10	7									
	As-thenia placebo		9	10	9									
	Headac treatment	2	34	32	20			39	30	32				60

Table 4. Adverse outcomes (Continued)

		0	21	34	20			40	29	28				20
	Headache placebo													
	Infection treatment		14	10	7									
	Infection placebo		6	9	6									
	Trauma treatment			2	13									
	Trauma placebo			6	8									
	Fatigue treatment						1	12		7	5			33
	Fatigue placebo						2	13		1	1			10
	Weight gain treatment													54
	Weight gain placebo													3
	Hypertriglyceridaemia treatment													5

Table 4. Adverse outcomes (Continued)

	Hy-per-triglyc-eri-daemia placebo													0
	Ur-ticaria (hives) treat-ment													20
	Ur-ticaria (hives) placebo													6
Di-ges-tive sys-tem	Con-stipa-tion treat-ment			5										
	Con-stipa-tion placebo			4										
	De-creased ap-petite treat-ment		14	7								10		
	De-creased ap-petite placebo		3	4								2		
	In-creased ap-petite treat-ment													15

Table 4. Adverse outcomes (Continued)

	De-creased ap-petite placebo													2
	Diar-rhea treat-ment		4	7			2	8	5	7	5	18		
	Diar-rhea placebo		3	7			1	5	8	3	1	3		
	Dry mouth treat-ment			19										
	Dry mouth placebo			12										
	Dys-pepsia treat-ment			6	6									
	Dys-pepsia placebo			4	3									
	Nau-sea treat-ment		44	22	13			16	10	23	12	27		
	Nau-sea placebo		14	17	9			13	6	17	3	3		
	Vom-iting treat-ment		7	3	6		2	10	7			8		

Table 4. Adverse outcomes (Continued)

	Vomiting placebo		3	6	2		1	9	5			4		
Nervous system	Dizziness treatment		19	22	5				6	10			22	
	Dizziness placebo		7	16	1				3	6			11	
	Emotional lability treatment		8	6			0					12		
	Emotional lability placebo		3	3			1					2		
	Hostility/anger treatment			7			1							
	Hostility/anger placebo			0			0							
	Mania/hypomania treatment						3							
	Mania/hypomania placebo						2							

Table 4. Adverse outcomes (Continued)

	mania placebo													
	Nervousness treatment		2	8	6		0							
	Nervousness placebo		3	5	4		1							
	Somnolence treatment		17	8	10		0							66
	Somnolence placebo		6	5	7		1							6
	Tremor treatment		6	10			2							
	Tremor placebo		1	2			0							
	Insomnia treatment		9	14	11		3	16				26		
	Insomnia placebo		3	4	7		1	10				18		
Respiratory system	Pharyngitis treatment	6	2	5	8			13	7					

Table 4. Adverse outcomes (Continued)

	Pharyn- gitis placebo	4	5	8	6			15	8					
	Respi- ratory disor- der treat- ment		5	10	11			8						
	Respi- ratory dis- order placebo		3	11	11			12						
	Rhini- tis treat- ment		3	7	5			11	8	12				
	Rhini- tis placebo		3	5	3			14	8	9				
	Si- nusi- tis treat- ment			6	6		4							
	Si- nusi- tis placebo			7	4		2							
	Cough in- creased treat- ment			5	6									

Table 4. Adverse outcomes (Continued)

	Cough in- creased placebo			6	3									
	In- fluenza type symp- toms treat- ment	1			7			11	3	7	5			
	In- fluenza type symp- toms con- trol	2			4			5	8	3	1			
	Total par- tici- pants per group		Treat- ment = 183 Placebo = 93	Treat- ment = 93 Placebo = 87	Treat- ment = 101 Placebo = 102		Treat- ment = 109 Placebo = 112	Treat- ment = 158 Placebo = 147	Treat- ment = 131 Placebo = 133	Treat- ment = 121 Placebo = 112	Treat- ment = 89 Placebo = 85	Treat- ment = 189 Placebo = 184	Treat- ment = 45 Placebo = 29	

Table 5. Attrition (drop out) rates

Medica- tion	Trial ID	Total at- trition	Total ran- domised	Percent- age attri- tion	Attrition SSRI group	SSRI to- tal N	% attri- tion SSRI	Attrition placebo group	Placebo total N	% attri- tion placebo
Paroxe- tine	Berard 2006	90	286	31%	60	187	32%	30	99	30%
	Emslie 2006	57	206	28%	34	104	33%	23	102	23%
	Keller 2001	47	180	26%	26	93	28%	21	87	24%
	Paroxe- tine Trial 1	7	56	12.5%	4	29	14%	3	27	11%

Table 5. Attrition (drop out) rates *(Continued)*

Fluoxetine	Almeida-Montes 2005	16	23	70%	7	12	58%	9	11	82%
	Emslie 1997	36	96	38%	14	48	29%	22	48	46%
	Emslie 2002	61	219	28%	19	109	17%	42	110	38%
	Simeon 1990	8	40	20%						
	TADS 2004	41	221	19%	18	109	17%	23	112	21%
Sertraline	Wagner Trial 1&2 (2003)	77	376	20%	46	189	24%	31	187	17%
Citalopram	Von Knorring 2006	91	242	38%	50	124	40%	41	120	34%
	Wagner 2004	40	178	22%	22	93	24%	18	85	21%
Escitalopram	Emslie 2009	57	316	18%	32	158	20%	25	158	16%
	Wagner 2006	91	244	37%	45	124	36%	46	120	38%
Venlafaxine	Emslie 2007	109	367	30%	59	184	35%	50	183	27%
Mirtazapine	Mirtazapine Trial 1	22	126	18%	13	82	16%	9	44	21%
	Mirtazapine Trial 2	27	133	20%	19	88	22%	8	45	18%

SSRI: selective serotonin reuptake inhibitor

Table 6. Concurrent comorbid conditions in the treatment or control groups

	Almeida Montes 2005	Emslie 1997	Emslie 2002	Emslie 2006	Emslie 2007	Emslie 2009	Mir- tazap- ine Trial 1 & 2	TADS 2004	Berard 2006	Keller 2001	Wag- ner Trial 1&2 (2003)	Wag- ner 2004	Wag- ner 2006
Pres- ence of dys- thymia (treat- ment group)		41.7%		2%				5.5%				5.6%	
Pres- ence of dys- thymia (con- trol group)		29.2%		0%				10.7%				1.2%	
Pres- ence of anxiety (treat- ment group)		66.7%		10.9%				25.7%	17.0%	20.4%			4.5%
Pres- ence of anxiety (con- trol group)		45.8%		2%				28.6%	18.3%	32.2%			7.5%
Pres- ence of ADHD (treat- ment group)		33.3%	14.7%	3%				11.9%	1.6%			4.5%	
Pres- ence of ADHD (con- trol group)		27.1%	13.6%	1%				16.7%	0%			1.2%	

Table 6. Concurrent comorbid conditions in the treatment or control groups (Continued)

group)													
Pres- ence of ODD/ CD (treat- ment group)		27.1%	20.2%	4.9%				22.9%	0.5%	26.9%			
Pres- ence of ODD/ CD (con- trol group)		33.3%	16.4%	3.9%				25.0%	1.1%	23.0%			
Pres- ence of 'any' diag- nosis (treat- ment group)		85.4%		27.7%		12.9%		56.9%		44.1%			
Pres- ence of 'any' diag- nosis (con- trol group)		77.1%		17.6%		16.6%		48.7%		51.7%			

ADHD: attention deficit hyperactivity disorder; CD: Conduct Disorder; ODD: Oppositional Defiant Disorder

WHAT'S NEW

Last assessed as up-to-date: 28 October 2011.

Date	Event	Description
11 September 2012	New citation required but conclusions have not changed	Review updated to include new studies.
12 December 2011	New search has been performed	Search updated October 2011. Inclusion criteria amended to include new generation antidepressant medications

HISTORY

Protocol first published: Issue 3, 2004

Review first published: Issue 3, 2007

Date	Event	Description
5 November 2008	Amended	Converted to new review format.
30 March 2007	New citation required and conclusions have changed	Substantive amendment.

CONTRIBUTIONS OF AUTHORS

Sarah Hetrick

Conceived the review.

Co-ordinated preparation of the protocol and the review.

Contributed to the search strategy, extraction of information about the trials and outcome data, 'Risk of bias' assessment, data analysis and preparation of the text of the review.

Jo McKenzie

Assisted with the assessment of risk of bias of the included trials. Advised on the analysis of the data and made a significant contribution to writing the results, including input to the discussion and conclusions.

Georgina Cox

Extracted data, double entered data, updated tables of included studies and additional tables, assisted the write-up of some of the results, and reviewed the final manuscript.

Magenta Simmons

Screened studies for in and exclusion, updated the background section, assisted with the discussion and reviewed the final manuscript for the updated version of the review.

Sally Merry

Help to conceive the review.

Provided general advice on the protocol.

Helped prepare the text of the review, with a significant contribution to the discussion and conclusion of the review.

Per Sindahl

Provided general advice on the protocol. Contributed to extraction of information about the trials and outcome data, and preparation of the text of the review for the first published version of the review. He was unable to contribute to the update of the review.

Michelle Proctor

Assisted with preparation of the protocol and made significant contribution to running the search strategy, contributed to extraction of information about the trials and outcome data, and commented on the text of the first published version of the first published version of the review. She was unable to contribute to the update of the review.

Alison Ward

Provided with general advice on the protocol, contributed to extraction of information about the trials and outcome data and commented on the text of the review of the first published version of the review. She was unable to contribute to the update of the review.

DECLARATIONS OF INTEREST

Per Sindahl works at Leo Pharma A/S, however, his main areas are dermatology and critical care. He is not involved in psychiatric disorders.

There is no conflict of interest for any author.

SOURCES OF SUPPORT**Internal sources**

- Colonial Foundation, Australia.

Orygen Youth Health Research Centre receives funds from the Colonial Foundation

- Australian National Health and Medical Research Council (NHMRC) Clinical Training Fellowship, Australia.

Person support for Sarah Hetrick for research focused on implementation of evidence for youth depression

External sources

- No sources of support supplied

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

Inclusion criteria have changed for this version of the review and now include all 'second and third' or newer antidepressant compounds. Additionally, sub-syndromal depression was not included.

We have added as an objective the aim to look at effectiveness by drug type; this analysis was undertaken in the first review but not explicitly described in the objectives.

Analysis has been reorganised such that there is not a separate meta-analysis for each antidepressant compound, rather, these have become one of two subgroup analyses.

NOTES

The results for the two duloxetine studies became available in July 2012, just as the review was being finalised for publication. The results for these studies will be incorporated in an update of this review in early 2013.

INDEX TERMS

Medical Subject Headings (MeSH)

Antidepressive Agents [adverse effects; *therapeutic use]; Citalopram [administration & dosage; therapeutic use]; Depressive Disorder [*drug therapy; psychology]; Fluoxetine [adverse effects; therapeutic use]; Induction Chemotherapy; Paroxetine [adverse effects; therapeutic use]; Randomized Controlled Trials as Topic; Serotonin Uptake Inhibitors [adverse effects; *therapeutic use]; Sertraline [administration & dosage; therapeutic use]; Suicide [*psychology]

MeSH check words

Adolescent; Child; Humans