Neuropsychopharmacology (2012) 37, 851–864

www.neuropsychopharmacology.org

Review

Randomized, Placebo-Controlled Trials of Antidepressants for Acute Major Depression: Thirty-Year Meta-Analytic Review

Juan Undurraga[1,2] and Ross J Baldessarini*[,1]

[1]Department of Psychiatry, Harvard Medical School, Psychopharmacology Program and International Consortium for Psychotic and Mood Disorders Research, Mailman Research Center, McLean Division of Massachusetts General Hospital, Boston, MA, USA; [2]Bipolar Disorders Program, Department of Psychiatry, Hospital Clinic, University of Barcelona, IDIBAPS, CIBERSAM, Barcelona, Spain

Antidepressant–placebo response-differences (RDs) in controlled trials have been declining, potentially confounding comparisons among older and newer drugs. For clinically employed antidepressants, we carried out a meta-analytic review of placebo-controlled trials in acute, unipolar, major depressive episodes reported over the past three decades to compare efficacy (drug–placebo RDs) of individual antidepressants and classes, and to consider factors associated with year-of-reporting by bivariate and multivariate regression modeling. Observed drug–placebo differences were moderate and generally similar among specific drugs, but larger among older antidepressants, notably tricyclics, than most newer agents. This outcome parallels selective increases in placebo-associated responses as trial-size has increased in recent years. Study findings generally support moderate efficacy of clinically employed antidepressants for acute major depression, but underscore limitations of meta-analyses of controlled trials for ranking drugs by efficacy. We suggest that efficiency and drug–placebo differences may be improved with fewer sites and subjects, and better quality-control of diagnostic and clinical assessments. *Neuropsychopharmacology* (2012) **37,** 851–864; doi:10.1038/npp.2011.306; published online 14 December 2011

Keywords: antidepressants; evidence-based medicine; depression; meta-analysis; controlled trials; secular changes

INTRODUCTION

Efficacy of antidepressant drugs for treatment of acute, unipolar, major depressive episodes continues to be a research and clinical topic of considerable interest. There have been major changes in clinical practice involving antidepressants since the early 1990s (Healy, 1997; Baldessarini, 2005; Ghaemi, 2008). Currently favored drugs include the serotonin reuptake inhibitors (SRIs) introduced since the late 1980s, and a series of additional modern, 'second-generation' antidepressants. These include agents with mixed inhibitory actions on the neuronal-uptake and inactivation of serotonin and norepinephrine (SNRIs, including desvenlafaxine, duloxetine, milnacipran, venlafaxine, and others), and 'atypical' agents with other actions (such as bupropion, nefazodone, mirtazapine, and vilazodone). These modern or 'second-generation' antidepressants have largely displaced older antidepressants including tricyclics (TCAs) and monoamine oxidase (MAO) inhibitors (Baldessarini, 2005, 2012).

*Correspondence: Dr RJ Baldessarini, Mailman Research Center 312, McLean Hospital, 115 Mill Street, Belmont, MA 02478-9106 USA, Tel: +1 617 855 3203, Fax: +1 617 855 3479,
E-mail: rbaldessarini@mclean.harvard.edu
Received 27 July 2011; revised 27 September 2011; accepted 3 November 2011

The superiority of most clinically employed antidepressants over placebos in controlled trials has been modest in adult patients diagnosed with major depression, even lower in juvenile depressed patients, and probably has declined in recent years (Walsh *et al*, 2002; Baldessarini, 2005; Cipriani *et al*, 2007; Papakostas *et al*, 2007; Gartlehner *et al*, 2008; Kirsch *et al*, 2008; Tsapakis *et al*, 2008; Bridge *et al*, 2009; Wooley *et al*, 2009; Masi *et al*, 2010; Pigott *et al*, 2010; Khin *et al*, 2011). Evident decline in superiority of drugs over placebos has occurred despite evidence of selective reporting of positive findings of potential commercial interest from therapeutic trials (Ioannidis, 2008; Turner *et al*, 2008).

Moreover, there is little evidence that one antidepressant or pharmacological class of antidepressants is clearly and convincingly more effective than others (Anderson, 2001; Baldessarini, 2005, 2012; Cipriani *et al*, 2007; Papakostas *et al*, 2007; Gartlehner *et al*, 2008; Kirsch *et al*, 2008; Khin *et al*, 2011). In part, this lack of clear differentiation may arise from the modest drug–placebo differences in many controlled trials of antidepressants, which, in turn, may reflect broad clinical heterogeneity arising from the current broad concept of 'major depression' (Healy, 1997; Ghaemi, 2008). Possible differentiation of efficacy among antidepressants was a lively question soon after introduction of the SRIs and SNRIs (Healy, 1997; Baldessarini, 2005, 2012). However, the popularity of most modern antidepressants

owes far more to their perceived safety, relative ease of use, and broad clinical utility rather than to well-demonstrated superior efficacy in major depressive disorder compared with older agents (Baldessarini, 2005, 2012; Cipriani *et al*, 2007; Papakostas *et al*, 2007; Wooley *et al*, 2009; Pigott *et al*, 2010; Khin *et al*, 2011).

Given current uncertainties regarding the relative efficacy of specific drugs and pharmacological classes of antidepressants, we carried out a systematic, meta-analytic review of peer-reviewed, placebo-controlled trials, reported since 1980, limiting inclusion to drugs with regulatory approval for major depression that are currently employed clinically in the United States. Specific aims were to: (a) compare the efficacy of older and modern antidepressants compared with placebo; (b) further test the apparently widely held assumption that modern agents are at least equivalent in efficacy to older antidepressants (specifically TCAs); and (c) examine factors associated with anticipated declining differences in drug- *vs* placebo-associated responses in randomized, placebo-controlled trials of antidepressants.

MATERIALS AND METHODS

Search Strategy

We conducted a computerized literature review using Medline, CINAH Library, Cochrane Library, and PsycINFO literature databases using the following search-terms: 'antidepressant, amitriptyline, amoxapine, bupropion, citalopram, clomipramine, desipramine, depression (or major depression), desmethylvenlafaxine, duloxetine, escitalopram, fluoxetine, imipramine, isocarboxazid, mirtazapine, maprotiline, monoamine oxidase (or MAO) inhibitors, nortriptyline, phenelzine, paroxetine, *S*-citalopram, selegiline, sertraline, tranylcypromine, trazodone, tricyclic antidepressants, trimipramine, and venlafaxine,' alone and in various combinations. Also, reference lists of articles and reviews on antidepressant efficacy were hand-searched for relevant reports. The search was limited to peer-reviewed, published, randomized, placebo-controlled trials (RCTs) in acute episodes of adult major depressive disorder diagnosed by standardized criteria, and reported from 1980 through August 2011.

Eligibility Criteria

Included were reports of randomized, double-blind, placebo-controlled trials in adults in an acute, apparently unipolar, major depressive episode (or with $\leqslant$10% identified cases of bipolar depression or diagnoses other that major depression) based on DSM-III, III-R, or -IV, ICD-9 or -10, or RDC diagnostic criteria, and with at least 20 subjects per arm. We excluded trials of drugs that are not US FDA-approved and indicated in the United States for treatment of acute episodes of major depressive disorder, as well as reports involving special populations, such as juvenile or geriatric patients, treatment-resistant depression or depression associated with major neuromedical or other psychiatric disorders. Only monotherapy trials were included; antidepressant doses could be fixed or flexible, with or without low-doses (below the approximate equivalent (Baldessarini, 2005) of 2 mg/day of lorazepam) of supplemental sedative or hypnotic agents. For 35 trials with three randomized treatment conditions involving two active agents and a placebo arm, we compared each drug–placebo pair separately; in some three-arm trials involving an experimental agent and a standard comparator, we considered only a marketed agent *vs* placebo. When an active agent was used in different doses in the same trial, we calculated mean doses and outcome measures, all considered as a single drug-arm. Total daily drug doses (mg/day) were converted to approximate imipramine-equivalents (IMI-eq), based on the median of the range of clinical doses recommended by the manufacturers as summarized elsewhere (Baldessarini *et al*, 2010), so as to permit comparisons of agents of dissimilar potency.

Outcome Measures

The primary outcome measure was categorical 'response,' usually defined as $\geqslant$50% reduction in initial depression rating-scale scores. Most often, ratings were based on the Hamilton (HDRS) or Montgomery–Åsberg (MADRS) Depression Rating Scales (Hamilton, 1960; Montgomery and Åsberg, 1979), or Clinical Global Impression (CGI) ratings (Guy, 1976) when these measures were not available. Scores employed for analyses were standardized as the percentage of maximum attainable scores on each rating scale (eg, 48 for 17-item HDRS, 60 for the MADRS, and 61 for 21-item HDRS). When the number of items in the HDRS was not specified by the investigators, we considered it to be the most commonly employed 17-item version. When more than one depression rating scale was employed, we gave priority to results obtained with the HDRS for greater comparability. All measures of initial depression severity and its change by end-point were standardized by use of percentages of observed ratings to the maximum attainable score with each rating scale employed. Continuous measures of change in depression ratings with drug *vs* placebo were considered as secondary measures, because lack of variance measures in most trials precluded formal meta-analysis. We considered factors that might influence outcomes, including numbers of subjects and collaborating sites, percentage women, initial depression ratings, IMI-eq daily drug doses, trial-duration, dropout rates, specific drugs and types, and year of reporting. As manufacturers of the drugs involved sponsored almost all trials, sources of support were not further considered.

Data Analysis

Averaged data are means with SD, unless stated otherwise. Meta-analyses based on Stata *metan* programs, used random-effects modeling to limit effects of inter-trial variance; responder rates for each drug–placebo pair yielded pooled rate ratios (RRs) and rate differences (RDs) with their computed 95% confidence intervals (CIs) (Tsapakis *et al*, 2008; Yildiz *et al*, 2011a, b). Percentage-improvement in depression for drug–placebo pairs was compared by paired-*t* testing and averaged to provide overall estimates of response differences (RDs). We also carried out bivariate and multiple linear regression modeling from these analyses to evaluate associations of selected

covariates with reporting year. Correlations employed nonparametric Spearman rank methods (r_s) to avoid effects of non-normally distributed data and potential nonlinear relationships. The primary study-hypothesis was that all marketed antidepressants would be statistically more effective than placebo, on average, with only minor differences among specific drugs or types. Analyses were based on standard commercial software (Stata.8; StataCorp, College Station, TX; Statview.5; SAS Institute, Cary, NC).

RESULTS

Trials Characteristics

Initially, we screened >2000 potentially relevant reports appearing between 1980 and 2011. Based on reviewing abstracts, 179 reports appeared to meet selection criteria and not to include multiple reports of the same trials. Exclusions (71/179) were as follow: (a) 17 studies involved <20 patients per arm; (b) another 17 included >10% of subjects with diagnoses other than major depressive episode; (c) 11 studies involved special populations; (d) 4 trials did not include a placebo control arm; (e) another 22 reports were excluded for various other reasons, including outcomes that were not quantified or did not include responder rates or improvement in depression ratings, represented subpopulations of larger trials already considered, or involved unapproved drugs. Detailed review of entire reports led to inclusion of 107; they involved 142 drug–placebo comparisons (Table 1), owing to 35 trials arising from studies with three randomized arms (for which a total of 3677 placebo-treated subjects were considered twice). There were 27 127 non-duplicated adult subjects (17 059 randomized to an antidepressant, 9925 to placebo), of average age 40 years (62.0 ± 9.9% women). Antidepressants tested ($n = 19$) ranked by trial-count as: imipramine (23 trials), fluoxetine (17), venlafaxine (15), paroxetine (14), amitriptyline (12), duloxetine (10), bupropion (9), desvenlafaxine (8), sertraline (8), *R,S*-citalopram (7), *S*-citalopram (5), mirtazapine (4), selegiline (3), desipramine (2), clomipramine (1), nortriptyline (1), phenelzine (1), tranylcypromine (1), and trazodone (1). Types of antidepressants ranked: SRIs (52 trials (36.6%)), TCAs (38 (26.8%)), SNRIs (33 (23.2%)), atypical agents (bupropion, mirtazapine, trazodone; 14 (9.9%)), and MAO-inhibitors (5 (3.5%)). Subjects per trial ranked: SNRIs (288 ± 118) >SRIs (230 ± 146) > atypical agents (224 ± 144) >MAO-inhibitors (181 ± 98) > TCAs (139 ± 101); there were far more sites per trial since the median reporting-year of 1998 (range 1983–2010): 22.7 ± 16.8 *vs* 7.22 ± 5.98, as well as more subjects per trial: 270 ± 114 *vs* 181 ± 122, indicating a major secular trend toward increasing trial-size.

Initial depression scores (as percentage of scale maxima were similar in drug- (45.6 ± 6.5%) and placebo-arms (48.6 ± 8.4%)). There were 120 ± 90 subjects (range: 11–521) per antidepressant arm and 96 ± 57 per placebo arm (range: 18–273) or 216 ± 136 participants per trial, and 16.1 ± 15.3 collaborating sites per trial. Treatment lasted approximately 7.2 ± 1.8 weeks, uncorrected for early dropouts at unspecified times, at rates of approximately 29.8 ± 12.3% or 4.54 ± 2.58% per week with drugs, and 33.3 ± 15.7% or 5.06 ± 3.02% per week with placebos (paired-$t = 2.71$, $p = 0.007$). Supplemental use of moderate doses of sedative-anxiolytics was permitted in 59.1% of all trials. Most trials (81.7%) included at least brief periods to allow previously administered drugs to 'wash-out,' and most (78.9%) employed intention-to-treat methods; 97.4% of trials were sponsored by pharmaceutical manufacturers. The overall estimated IMI-eq standardized dose was 158 ± 68 mg/day, and did not differ by drug-type or between older (TCAs, MAO-inhibitors: 155 ± 49 mg/day) and modern antidepressants (SRIs, SNRIs, and atypical agents: 159 ± 77 mg/day).

Meta-Analysis

Meta-analyses with the 122 trials reporting on responder rates yielded pooled drug–placebo RRs (RR with CIs) for each agent, and an overall pooled RR value of 1.42. (95% CI: 1.38–1.48; $z = 16.3$, $p < 0.0001$). Among agents with more than one trial, amitriptyline ranked highest in apparent efficacy, and bupropion lowest; however, CIs for most agents overlapped, indicating the need for caution in attempting to rank drugs by efficacy (Figure 1). Single-trial data available for phenelzine, clomipramine, nortriptyline, trazodone, and tranylcypromine are likely to be unstable and unreliable (Figure 1). Construction of a 'funnel plot' (1/standard-error-of-RR *vs* 1/RR) for all reports with data on responder rates yielded a V-shaped distribution of values that was symmetrically distributed around the pooled value of 1/RR (not shown); this finding may provide evidence against selective reporting of positive trials results.

We also compared antidepressants by types with pooled data, and compared apparent efficacy by three outcome measures. These included meta-analytically computed response RRs and responder rate-differences (RD), as well as relative differences (RD) in changes in depression ratings with drug–placebo pairs. Although these outcome measures yielded slightly different rankings, TCAs consistently ranked as the most effective antidepressants considered, and atypical agents, seemingly least effective (Table 2). Trials carried out before the median reporting year (1998) yielded higher values of all efficacy measures (Table 2). Median years of trial-reporting ranked: TCAs (1991) < MAO-inhibitors (1997) < atypical agents (1998) = SRIs (1998) < SNRIs (2003). Efficacy based on responder-rate RR values was much greater for TCAs than other types of antidepressants (1.83 ± 0.62 *vs* 1.48 ± 0.41; $F = 11.8$, $p = 0.0008$). Moreover, when the numbers of placebo-responders and nonresponders in the TCA trials were substituted for corresponding placebo data for trials of modern antidepressants, the meta-analytically pooled RR value was identical to that found in the TCA trials, supporting the impression that apparent differences response rates with the two classes of antidepressants was accounted for by secular changes in placebo responses.

Factors Associated with Trials Results

Given the preceding findings suggesting that older agents, specifically TCAs, might appear to be somewhat more effective than modern antidepressants in general, and that older trials yielded consistently greater drug–placebo differences, we carried out several correlational analyses to further examine effects of reporting-year on numbers of sites and

Table 1 Characteristics of Placebo-Controlled Trials of Antidepressants in Major Depression

Trial (reference)	Drug	mg/day	IMI-eq	Weeks	Total N	Sites	% Women	Response	N Rx	% Resp Rx	N Pbo	% Resp Pbo	RR Resp	Ratings	Initial Dep Rx	Change Rx (%)	Initial Dep Pbo	Change Pbo (%)	RD % Change	Dropout Rx (%)	Dropout Pbo (%)	ITT	Washout
Claghorn *et al* (1983)	AMI	180	180	4	172	3	56	HDRS≥50%	85	62.4	87	42.5	1.47	HDRS21	42.6	—	45.9	—	—	58	58	Yes	Yes
Feighner *et al* (1983)	IMI	163	163	6	487	5	71	CGI	244	57.8	243	32.5	1.78	HDRS21	42.6	—	42.6	—	—	24	40	Yes	Yes
Itil *et al* (1983)	IMI	127	127	4	47	—	44	CGI	25	44	22	22.7	1.94	HDRS16	47.8	53	43.5	18	34.8	48	50	No	Yes
Pitts *et al* (1983)	BUP	525	262	4	59	—	34	HDRS≥50%	34	—	25	—	—	HDRS21	50.8	48	50.8	29	19.3	—	—	No	Yes
White *et al* (1984)	NRT	109	136	4	120	1	45	CGI	61	41	59	32.2	1.27	HDRS	52.1	54	56.2	37	16.6	34	24	Yes	Yes
White *et al* (1984)	TCP	44	145	4	122	1	45	CGI	63	39.7	59	32.2	1.23	HDRS	56.3	45	56.2	37	8.1	41	24	Yes	Yes
Cohn and Wilcox (1985)	FLX	70	350	6	112	—	62	HDRS≥50%	54	74.1	58	20.7	3.58	HDRS21	42.6	55	41	16	39.3	35	72	Yes	Yes
Cohn and Wilcox (1985)	IMI	152	152	6	112	—	53	HDRS≥50%	54	40.7	58	20.7	1.97	HDRS21	42.6	34	41	16	17.9	63	72	Yes	Yes
Rickels *et al* (1985)	AMI	148	148	6	254	—	66	HDRS≥50%	124	53.2	130	26.9	1.98	HDRS21	41	42	42.6	28	13.6	27	45	No	Yes
Mendels and Schless (1986)	IMI	167	167	6	68	—	46	HDRS≥50%	34	38.2	34	17.6	2.17	HDRS17	50	40	50	23	16.7	52	52	Yes	Yes
Rickels *et al* (1987)	IMI	143	143	6	124	—	62	HDRS≥50%	63	69.8	61	37.7	1.85	HDRS21	39.3	42	41	20	21.2	41	39	No	Yes
Wernicke *et al* (1987)	FLX	40	200	6	240	10	57	HDRS≥50%	207	54.1	33	—	—	—	—	—	0	—	—	44	44	No	Yes
Hollyman *et al* (1988)	AMI	110	119	6	178	—	83	CGI	90	58.9	88	44.3	1.33	HDRS	31.3	62	31.2	41	21.3	26	16	No	No
Wernicke *et al* (1988)	FLX	22	110	—	363	—	61	HDRS≥50%	285	46.3	78	23.1	2	HDRS21	41	43	42.6	27	15.7	37	46	Yes	Yes
Feighner *et al* (1989a)	FLX	80	400	6	99	—	75	CGI	51	—	48	—	—	HDRS21	42.6	31	42.6	22	8.5	51	68	No	Yes
Feighner *et al* (1989a)	IMI	150	159	6	55	—	74	CGI	36	—	19	—	—	HDRS21	42.6	38	42.6	22	2	48	68	No	Yes
Feighner *et al* (1989b)	IMI	159	150	6	94	—	89	HDRS≥50%	46	—	48	—	—	HDRS21	44.3	39	41	37	15.9	—	—	No	Yes
Larsen *et al* (1989)	CMI	150	165	6	38	—	66	HDRS<9	20	55	18	22.2	2.48	HDRS17	37.5	58	37.5	32	26.2	15	28	Yes	No
Miller *et al* (1989)	PRX	30	150	4	47	—	71	CGI	22	45.5	25	36	1.26	HDRS21	39.3	25	39.3	26	–0.50	45	20	Yes	Yes
Quitkin *et al* (1989)	IMI	150	150	6	54	2	55.9	HDRS≥50%	27	51.9	27	18.5	2.81	HDRS17	30.2	65	30.2	26	39.3	26	30	Yes	Yes
Quitkin *et al* (1989)	PNZ	75	150	6	53	2	55.9	HDRS≥50%	26	57.7	27	18.5	3.12	HDRS17	30.2	50	30.2	26	24.8	26	27	Yes	Yes
Gelenberg *et al* (1990)	IMI	175	175	4	43	—	32	HDRS≥50%	22	—	21	×	×	HDRS27	33.3	51	34	35	15.9	36	23	No	Yes
Lineberry *et al* (1990)	BUP	287	144	6	219	5	65	HDRS≥50%	110	50.9	109	33.9	1.5	HDRS21	—	—	44.3	38	—	23	29	Yes	Yes
Reimherr *et al* (1990)	AMI	104	104	8	299	8	54	HDRS≥50%	149	57.7	150	32.7	1.77	HDRS18	44.2	61	44.2	37	23.5	42	—	—	—
Reimherr *et al* (1990)	SRT	145	181	8	299	8	54	HDRS≥50%	149	51.7	150	32.7	1.58	HDRS18	44.2	53	44.2	37	15.9	41	37	Yes	Yes
Roth *et al* (1990)	DMI	224	168	6	53	2	59	CGI	24	62.5	29	37.9	1.65	HDRS17	62.5	38	60.4	29	8.6	—	—	No	Yes
Smith *et al* (1990)	AMI	111	111	6	100	—	57	HDRS≥50%	50	56	50	30	1.87	HDRS17	50	54	47.9	29	24.8	30	—	Yes	—
Smith *et al* (1990)	MTZ	18	77.4	6	100	—	57	HDRS≥50%	50	54	50	30	1.8	HDRS17	47.9	47	47.9	29	17.8	40	50	Yes	Yes
Carman *et al* (1991)	AMI	200	200	6	150	—	—	HDRS≥50%	—	—	—	—	—	HDRS21	45.9	51	44.3	26	24.5	4	4	No	Yes
Khan *et al* (1991)	VNX	74	66.6	6	93	—	60	HDRS≥50%	67	—	26	—	—	HDRS21	41	56	42.6	31	25.7	21	15	—	Yes
Bakish *et al* (1992)	AMI	112	112	7	112	5	43	HDRS≥50%	57	50.9	55	34.5	1.47	HDRS17	47.9	—	47.9	—	—	32	49	Yes	Yes
Claghorn *et al* (1992)	PRX	38	190	6	337	4	52	HDRS<10	168	38.1	169	24.3	1.57	HDRS21	44.3	48	42.6	33	15.2	35	44	Yes	Yes
Cohn and Wilcox (1992)	PRX	37	175	6	67	—	58	HDRS≥50%	31	—	36	—	—	HDRS17	52.1	34	54.2	20	20.9	31	67	Yes	Yes
Cohn and Wilcox (1992)	IMI	175	185	6	71	—	54	HDRS≥50%	35	—	36	—	—	HDRS17	52.1	41	54.2	20	14.2	26	67	Yes	Yes
Fabre (1992)	IMI	135	135	6	80	1	62	HDRS≥50%	40	—	40	—	—	HDRS21	—	—	—	—	—	21	53	No	—
Fabre (1992)	PRX	29	145	6	80	1	62	HDRS %	40	—	40	—	—	HDRS21	—	—	—	—	—	21	53	No	Yes
Feighner (1992)	IMI	113	113	6	79	—	—	HDRS≥50%	40	50	39	12.8	3.9	HDRS21	—	—	—	—	—	41	54	Yes	Yes
Feighner (1992)	PRX	26	130	6	78	—	—	HDRS≥50%	39	28.2	39	12.8	2.2	HDRS21	—	—	—	—	—	60	54	Yes	X
Kiev (1992)	PRX	31	155	6	78	—	45	HDRS≥50%	34	55.9	44	25	2.24	HDRS17	60.4	45	58.3	24	20.7	38	44	No	Yes
Rickels *et al* (1992)	PRX	32	160	6	111	—	64	HDRS≥50%	55	40	56	19.6	2.04	HDRS21	—	—	42.6	—	—	29	21	No	Yes
Feighner *et al* (1993)	IMI	140	140	6	477	6	49	HDRS<10	237	26.6	240	12.9	2.06	HDRS21	42.6	35	44.3	22	13.3	54	53	Yes	Yes
Feighner *et al* (1993)	PRX	30	150	6	480	6	51	HDRS<10	240	24.6	240	12.9	1.9	HDRS21	42.6	38	44.3	22	16.1	42	53	Yes	Yes
Cunningham *et al* (1994)	TZD	297	346	6	153	6	66	CGI	77	59.7	76	55.3	1.08	HDRS21	41	43	39.3	36	7.3	36	36	Yes	Yes
Cunningham *et al* (1994)	VNX	158	142	6	148	6	66	CGI	72	72.2	76	55.3	1.31	HDRS21	41	48	39.3	36	11.1	29	36	Yes	Yes
Doogan and Langdon (1994)	SRT	75	93.8	6	200	—	68	MADRS≥50%	99	50.5	101	39.6	1.28	MADRS	46.6	55	44.5	45	10.4	19	10	Yes	Yes

Table 1 Continued

Trial (reference)	Drug	mg/day	IMI-eq	Weeks	Total *N*	Sites	% Women	Response	*N* Rx	% Resp Rx	*N* Pbo	% Resp Pbo	RR Resp	Ratings	Initial Dep Rx	Change Rx (%)	Initial Dep Pbo	Change Pbo (%)	RD % Change	Dropout Rx (%)	Dropout Pbo (%)	ITT	Washout
Fontaine *et al* (1994)	IMI	214	214	6	90	1	58	HDRS≥50%	45	48.9	45	31.1	1.57	HDRS17	54.2	42	54.2	26	15.6	42	47	Yes	Yes
Rickels *et al* (1994)	IMI	191	191	8	187	12	63	HDRS≥50%	95	65.3	92	44.6	1.46	HDRS17	50	—	50	—	—	49	37	No	Yes
Schweizer *et al* (1994)	IMI	176	176	6	151	2	66	HDRS≥50%	73	61.6	78	47.4	1.3	HDRS21	39.3	43	41	38	5.2	45	27	No	—
Schweizer *et al* (1994)	VNX	182	164	6	151	2	66	HDRS≥50%	73	76.7	78	47.4	1.62	HDRS21	41	55	41	38	16.5	36	27	No	Yes
Silverstone (1994)	IMI	150	150	6	135	13	55	HDRS≥50%	66	50	69	50.7	0.99	HDRS17	52.1	48	50	43	4.4	40	35	No	No
Vartiainen and Leinonen (1994)	MTZ	32.5	140	6	114	8	54	HDRS≥50%	59	—	55	—	—	HDRS21	42.6	59	42.6	48	10.7	37	44	Yes	Yes
Wilcox *et al* (1994)	AMI	122	122	6	99	2	47	HDRS≥50%	50	56	49	24.5	2.29	HDRS21	42.6	59	42.6	45	14.2	44	55	Yes	Yes
Bremner (1995)	AMI	186	186	6	100	—	68	HDRS≥50%	50	58	50	34	1.71	HDRS17	56.3	—	56.2	—	—	20	24	Yes	Yes
Bremner (1995)	CTP	30	129	6	100	—	68	HDRS≥50%	50	70	50	34	2.06	HDRS17	58.3	—	56.2	—	—	18	24	Yes	Yes
Claghorn and Lesem (1995)	MTZ	16	68.8	6	90	—	44	HDRS≥50%	42	50	48	27.1	1.85	HDRS17	45.8	—	47.9	—	—	40	58	No	Yes
Fabre *et al* (1995)	SRT	171	214	6	369	8	53	CGI	278	60.1	91	41.8	1.44	HDRS17	52.1	47	52.1	34	12.3	23	49	Yes	Yes
Guelfi *et al* (1995)	VNX	350	315	4	93	6	85	HDRS≥50%	46	52.2	47	31.9	1.63	HDRS17	—	—	60.4	17	—	24	57	Yes	Yes
Khan (1995)	MTZ	36	155	6	54	1	67	HDRS≥50%	27	55.6	27	37	1.5	HDRS17	47.9	53	45.8	29	24.7	33	41	Yes	Yes
Laakman *et al* (1995)	AMI	102	102	6	146	—	71	HDRS≥50%	72	73.6	74	21.6	3.4	HDRS17	41.7	60	39.6	25	34.9	5	12	No	Yes
Mynors-Wallis *et al* (1995)	AMI	139	139	12	61	15	74	HDRS≤7	31	51.6	30	26.7	1.94	HDRS17	—	—	37.5	36	—	19	60	No	No
Cassano *et al* (1996)	IMI	150	150	6	123	18	52.8	MADRS≥50%	64	—	59	—	—	MADRS	51.8	41	51.7	28	13.1	27	39	Yes	Yes
Claghorn *et al* (1996)	IMI	136	136	6	89	—	64	CGI	44	45.5	45	26.7	1.7	HDRS21	42.6	40	42.6	25	15.2	58	60	Yes	Yes
Cohn *et al* (1996)	IMI	126	126	8	80	—	70	HDRS≥50%	38	60.5	42	35.7	1.69	HDRS17	—	—	47.9	39	—	39	26	Yes	Yes
Feiger (1996)	IMI	224	224	8	81	8	68	CGI	41	61	40	30	2.03	HDRS17	50	46	50	29	16.6	33	55	Yes	Yes
Cunningham (1997)	VNX	128	115	12	278	—	63	HDRS≥50%	179	57.5	99	30.3	1.9	HDRS21	39.3	55	40.8	36	18.8	34	41	Yes	Yes
Lecrubier *et al* (1997)	IMI	114	114	10	151	24	66	MADRS≥50%	75	62.7	76	59.2	1.06	MADRS	40	57	40	54	3.7	31	25	Yes	Yes
Lecrubier *et al* (1997)	VNX	125	112	10	154	24	69	MADRS≥50%	78	82.1	76	59.2	1.39	MADRS	41.7	64	40	54	10.6	29	25	Yes	Yes
Lydiard *et al* (1997)	AMI	91	91	8	260	15	67	HDRS≥50%	131	55.7	129	37.2	1.5	HDRS17	45.8	58	45.8	40	18.1	38	29	Yes	Yes
Lydiard *et al* (1997)	SRT	91	114	8	261	15	67	HDRS≥50%	132	54.6	129	37.2	1.47	HDRS17	45.8	52	45.8	40	11.8	27	29	Yes	Yes
Thase (1997)	VNX	150	135	8	197	12	61	HDRS≥50%	95	57.9	102	29.4	1.98	HDRS21	39.3	48	39.3	30	18.2	27	40	Yes	Yes
Ban *et al* (1998)	DMI	150	112	4	174	6	62	HDRS≥50%	89	48.3	85	35.3	1.37	HDRS17	—	—	—	—	—	10	10	Yes	Yes
Fava *et al* (1998)	FLX	50	250	12	73	5	51	HDRS≥50%	54	57.4	19	52.6	1.09	HDRS21	39.3	45	39.3	48	–3.30	31	21	Yes	Yes
Fava *et al* (1998)	PRX	35	175	12	74	5	51	HDRS≥50%	55	58.2	19	52.6	1.11	HDRS21	37.7	48	39.3	48	–0.50	29	21	Yes	Yes
Khan *et al* (1998)	VNX	142	128	12	382	12	64	HDRS≥50%	286	—	96	—	—	HDRS21	41	45	41	30	15	—	—	Yes	Yes
Massana (1998)	FLX	30	150	8	255	—	—	HDRS≥50%	127	55.9	128	34.4	1.63	HDRS21	—	—	—	—	—	24	41	—	—
Reimherr *et al* (1998)	BUP	218	109	8	362	—	68	HDRS %	241	—	121	—	—	HDRS17	—	—	—	—	—	44	50	Yes	Yes
Rudolph *et al* (1998)	VNX	204	184	6	323	—	33	HDRS≥50%	231	48.9	92	29.3	1.67	HDRS21	—	—	—	—	—	53	41	Yes	Yes
Coleman *et al* (1999)	BUP	290	145	8	235	9	57	HDRS≥50%	118	66.1	117	56.4	1.17	HDRS31	—	59	—	50	9.4	22	32	Yes	Yes
Coleman *et al* (1999)	SRT	106	132	8	226	9	57	HDRS≥50%	109	60.5	117	56.4	1.07	HDRS31	—	57	—	50	6.9	36	32	Yes	Yes
Croft *et al* (1999)	BUP	293	146	8	232	8	51	HDRS≥50%	116	66.4	116	47.4	1.4	HDRS31	—	—	—	—	—	30	34	Yes	Yes
Croft *et al* (1999)	SRT	121	151	8	232	8	50	HDRS≥50%	116	68.1	116	47.4	1.44	HDRS31	—	—	—	—	—	33	34	Yes	Yes
Feighner and Overo (1999)	CTP	33	142	6	650	—	60	MADRS≥50%	521	—	129	—	—	HDRS21	41	46	40.8	38	7.7	—	—	Yes	Yes
Mendels *et al* (1999)	CTP	52	224	4	180	3	62	HDRS≥50%	89	80.9	91	47.3	1.71	HDRS17	50	39	50	29	10.7	48	44	Yes	Yes
Philipp *et al* (1999)	IMI	100	100	8	151	18	78	HDRS≥50%	105	62.9	46	47.8	1.31	HDRS17	45.8	64	47.9	53	10.7	—	—	Yes	Yes
Rudolph and Feiger (1999)	FLX	47	235	8	200	12	70	HDRS≥50%	103	50.5	97	42.3	1.19	HDRS21	42.6	45	41	41	4.6	27	21	—	—
Rudolph and Feiger (1999)	VNX	175	158	8	192	12	70	HDRS≥50%	95	56.8	97	42.3	1.34	HDRS21	41	50	41	41	9.2	19	21	Yes	Yes
Silverstone and Ravindran (1999)	FLX	40	200	12	237	—	60	HDRS≥50%	119	63	118	43.2	1.46	HDRS17	56.3	56	56.2	41	15.1	26	40	Yes	—
Silverstone and Ravindran	VNX	141	127	12	240	—	60	HDRS≥50%	122	64.8	118	43.2	1.5	HDRS17	56.3	58	56.2	41	17.7	29	40	Yes	Yes

Table 1 Continued

Trial (reference)	Drug	mg/day	IMI-eq	Weeks	Total *N*	Sites	% Women	Response	*N* Rx	% Resp Rx	*N* Pbo	% Resp Pbo	RR Resp	Ratings	Initial Dep Rx	Change Rx (%)	Initial Dep Pbo	Change Pbo (%)	RD % Change	Dropout Rx (%)	Dropout Pbo (%)	ITT	Washout
(1999)																							
Corrigan *et al* (2000)	FLX	20	100	8	70	8	—	HDRS≥50%	35	48.6	35	25.7	1.89	HDRS17	45.8	—	43.8	—	—	14	34	Yes	Yes
Stahl (2000)	CTP	57	245	8	215	8	60	HDRS≥50%	107	55.1	108	39.8	1.38	HDRS17	54.2	58	54.2	46	11.5	20	22	Yes	Yes
Stahl (2000)	SRT	143	179	8	216	8	60	HDRS≥50%	108	54.6	108	39.8	1.37	HDRS17	56.3	55	54.2	46	9.1	26	—	Yes	Yes
Coleman *et al* (2001)	BUP	335	168	8	302	15	62	HDRS≥50%	150	56	152	50	1.12	HDRS21	41	65	39.3	55	10.5	37	33	Yes	Yes
Coleman *et al* (2001)	FLX	29	145	8	306	15	63	HDRS≥50%	154	57.1	152	50	1.14	HDRS21	41	63	39.3	55	8.4	37	33	Yes	Yes
Andreoli *et al* (2002)	FLX	40	200	8	255	33	60	HDRS≥50%	127	55.9	128	33.6	1.66	HDRS21	44.3	—	44.3	—	—	8	12	Yes	Yes
Bodkin and Amsterdam (2002)	SLG	6	100	6	176	6	60	HDRS≥50%	88	37.5	88	22.7	1.65	HDRS17	47.7	38	48.5	26	11.8	11	17	Yes	Yes
Burke *et al* (2002)	CTP	40	172	8	244	35	61	MADRS≥50%	125	45.6	119	27.7	1.65	HDRS21	42.6	38	42.6	29	8.8	—	—	Yes	Yes
Burke *et al* (2002)	S-CTP	15	128	8	360	35	66	MADRS≥50%	241	50.6	119	27.7	1.83	HDRS21	41	44	42.6	29	14.3	—	—	Yes	Yes
Davidson (2002)	SRT	75	93.8	8	225	12	66	HDRS≥50%	109	48.6	116	43.1	1.13	HDRS17	47.9	46	47.9	40	5.8	28	28	Yes	Yes
Detke *et al* (2002a)	DLX	60	108	9	267	18	54	HDRS≥50%	139	44.7	128	23	1.95	HDRS17	43.8	52	43.8	29	−11.2	—	—	Yes	Yes
Detke *et al* (2002b)	DLX	60	108	9	245	21	99	HDRS≥50%	123	50.4	122	35.2	1.43	HDRS17	41.7	40	41.7	52	23.1	39	35	Yes	Yes
Golden *et al* (2002)	PRX	43.2	216	12	622	40	65	HDRS≥50%	417	58	205	47.8	1.21	HDRS17	50	53	50	46	7.1	—	—	Yes	Yes
Goldstein *et al* (2002)	DLX	107	193	8	140	8	64	HDRS≥50%	70	64.3	70	48.6	1.32	HDRS17	37.5	53	39.6	34	18.9	34	34	Yes	—
Goldstein *et al* (2002)	FLX	20	100	8	103	8	64	HDRS≥50%	33	51.5	70	48.6	1.06	HDRS17	37.5	44	39.6	34	9.2	36	34	Yes	—
Wade *et al* (2002)	S-CTP	10	85	8	380	40	76	MADRS≥50%	191	55	189	41.8	1.32	MADRS	48.3	51	48.3	42	9.2	16	15	Yes	Yes
Amsterdam (2003)	SLG	6	100	8	289	16	64	HDRS≥50%	145	32.4	144	27.8	1.17	HDRS17	47.5	36	47.9	29	6.7	28	28	Yes	Yes
Lepola *et al* (2003)	CTP	28	120	8	313	69	72	MADRS≥50%	159	52.8	154	48.1	1.1	MADRS	—	—	48.3	42	—	5	10	Yes	Yes
Lepola *et al* (2003)	S-CTP	14	119	8	309	69	72	MADRS≥50%	155	63.9	154	48.1	1.33	MADRS	—	—	48.3	42	—	6	—	—	—
Detke *et al* (2004)	DLX	100	180	8	281	—	70	HDRS≥50%	188	68.1	93	44.1	1.54	HDRS17	41.7	58	41.7	44	13.8	11	19	Yes	Yes
Detke *et al* (2004)	PRX	20	100	8	179	—	71	HDRS≥50%	86	74.4	93	44.1	1.69	HDRS17	41.7	58	41.7	44	14.5	12	19	Yes	Yes
Goldstein *et al* (2004)	DLX	60	108	8	266	19	60	HDRS≥50%	177	47.5	89	31.5	1.51	HDRS17	37.5	44	35.4	29	15	39	42	Yes	Yes
Goldstein *et al* (2004)	PRX	20	100	8	176	19	64	HDRS≥50%	87	40.2	89	31.5	1.28	HDRS17	37.5	36	35.4	29	6.7	44	42	Yes	Yes
Trivedi *et al* (2004)	PRX	19	95	8	447	40	58	HDRS≥50%	301	—	146	—	—	HDRS17	47.9	52	50	42	9.5	21	23	Yes	Yes
Bjerkenstedt *et al* (2005)	FLX	20	100	4	109	15	79	HDRS≥50%	54	37	55	—	—	HDRS21	39.3	37	41	38	−1.10	11	5	Yes	Yes
Brannan *et al* (2005)	DLX	60	108	7	280	25	65	HDRS≥50%	141	42	141	39.7	1.06	HDRS17	47.9	46	45.8	46	—	—	—	Yes	Yes
Fava *et al* (2005)	FLX	20	100	12	90	2	59	HDRS<8	47	29.8	43	20.9	1.42	HDRS17	41.7	32	41.7	37	−4.60	49	51	Yes	Yes
Clayton *et al* (2006)	S-CTP	13	158	8	549	—	59	HDRS≥50%	276	61.2	273	48.7	1.26	HDRS17	47.9	58	47.9	52	4.4	25	24	Yes	Yes
Clayton *et al* (2006)	BUP	316	110	8	554	—	59	HDRS≥50%	281	59.1	273	48.7	1.21	HDRS17	50	56	47.9	52	6.9	25	24	Yes	Yes
Feiger *et al* (2006)	SLG	9	150	8	265	3	57	HDRS≥50%	132	40.2	133	30.1	1.34	HDRS17	48.8	37	49.4	32	5.6	24	20	Yes	Yes
Gastpar *et al* (2006)	CTP	20	86	6	257	21	69	HDRS≥50%	127	55.9	130	39.2	1.43	HDRS17	45.8	53	45.8	41	11.9	18	19	Yes	—
Jefferson *et al* (2006)	BUP	352	176	8	274	24	68	CGI	135	53.3	139	38.1	1.4	IDSIVR30	54.8	46	54.8	38	8.1	24	21	Yes	No
Moreno *et al* (2006)	FLX	20	100	8	46	1	83	HDRS≥50%	20	55	26	42.3	1.3	HDRS21	24.6	53	26.2	31	22.1	20	27	Yes	Yes
Perahia *et al* (2006)	DLX	100	180	8	295	22	70	HDRS≥50%	196	66.3	99	51.5	1.29	HDRS17	43.8	59	43.8	52	6.2	12	10	Yes	No
Perahia *et al* (2006)	PRX	20	100	8	196	22	70	HDRS≥50%	97	60.8	99	51.5	1.18	HDRS17	43.8	57	43.8	52	4.3	9	10	Yes	—
DeMartinis *et al* (2007)	*d*VNX	233	291	8	470	25	60	HDRS≥50%	350	—	120	—	—	HDRS17	47.9	45	47.9	33	11.7	25	18	Yes	Yes
Liebowitz *et al* (2007)	*d*VNX	187	234	8	238	—	60	HDRS≥50%	121	43	117	34.2	1.26	HDRS17	50	40	50	36	4.2	18	25	Yes	Yes
Nemeroff and Thase (2007)	FLX	41	205	6	206	13	62	HDRS≥50%	104	45.2	102	37.3	1.21	HDRS21	39.3	—	—	—	—	17	24	Yes	—
Nemeroff and Thase (2007)	VNX	142	128	6	204	13	62	HDRS≥50%	102	52.9	102	37.3	1.42	HDRS21	39.3	—	—	—	—	24	24	Yes	—
Nierenberg *et al* (2007)	S-CTP	10	108	8	410	36	66	HDRS≥50%	273	45.3	137	37.2	1.22	HDRS17	37.5	41	37.5	34	9.5	24	29	Yes	Yes
Nierenberg *et al* (2007)	DLX	60	85	8	411	36	63	HDRS≥50%	274	48.7	137	37.2	1.31	HDRS17	37.5	43	37.5	34	6.9	31	29	Yes	Yes
Septien-Velez *et al* (2007)	*d*VNX	300	375	8	369	35	66	HDRS≥50%	245	58	124	37.9	1.53	HDRS17	52.1	49	52.1	37	12.6	27	22	Yes	No
Boyer *et al* (2008)	*d*VNX	75	93.8	8	485	44	70	HDRS≥50%	324	63.9	161	50.3	1.27	HDRS17	50	56	50	45	11.4	15	9	Yes	Yes
Lieberman *et al* (2008)	VNX	162	399	8	471	—	68	HDRS≥50%	226	60.3	245	46.9	1.29	HDRS17	54.2	57	54.2	47	9.2	18	14	Yes	Yes

Table 1 Continued

Trial (reference)	Drug	mg/day	IMI-eq	Weeks	Total *N*	Sites	% Women	Response	*N* Rx	% Resp Rx	*N* Pbo	% Resp Pbo	RR Resp	Ratings	Initial Dep Rx	Change Rx (%)	Initial Dep Pbo	Change Pbo (%)	RD % Change	Dropout Rx (%)	Dropout Pbo (%)	ITT	Washout
Lieberman *et al* (2008)	dVNX	319	146	8	487	—	66	HDRS ≥ 50%	242	54.9	245	46.9	1.17	HDRS17	52.1	56	54.2	47	10	26	14	Yes	Yes
Liebowitz *et al* (2008)	dVNX	75	93.8	8	447	25	60	HDRS ≥ 50%	297	32.7	150	24	1.36	HDRS17	47.9	49	47.9	41	7.6	21	16	Yes	Yes
Cutler *et al* (2009)	DLX	60	108	6	308	38	60	MADRS ≥ 50%	151	49.7	157	36.3	1.37	HDRS17	52.1	49	52.1	40	8.3	30	21	Yes	Yes
Feiger *et al* (2009)	dVNX	349	436	8	230	12	65	HDRS ≥ 50%	117	41.9	118	29.7	1.41	HDRS17	47.9	46	47.9	32	14.1	25	13	Yes	Yes
Sheehan *et al* (2009)	FLX	58	290	6	194	22	66	HDRS ≥ 50%	99	35.3	95	36.7	0.96	HDRS21	49.2	39	47.5	37	1.2	23	29	Yes	No
Sheehan *et al* (2009)	VNX	235	212	6	186	22	56	HDRS ≥ 50%	91	51.6	95	36.7	1.41	HDRS21	49.2	48	47.5	37	10.4	29	29	Yes	No
Tourian *et al* (2009)	dVNX	75	108	8	317	21	64	HDRS ≥ 50%	157	44	160	38.1	1.15	HDRS17	47.9	44	50	36	8.5	25	27	Yes	Yes
Tourian *et al* (2009)	DLX	60	93.8	8	458	21	62	HDRS ≥ 50%	298	47.1	160	38.1	1.24	HDRS17	47.9	45	50	36	7.8	26	27	Yes	Yes
Hewett *et al* (2010)	BUP	180	90	8	390	65	66	MADRS	203	57.1	187	49.2	1.16	MADRS	51.7	48	51.7	43	4.9	22	22	No	No
Hewett *et al* (2010)	VNX	85	76.5	8	385	65	66	MADRS	198	66.2	187	49.2	1.34	MADRS	50	56	51.7	43	13.4	23	22	No	—
Means/sums	**142 Trials**	—	**156 ± 67**	**7.2 ± 1.8**	**27 127[a]**	**16 ± 15**	**62 ± 9.9**	—	**17 059**	**53.8 ± 10.9**	**9925[a]**	**36.6 ± 10.9**	**1.57 ± 0.49**	—	**45.3 ± 6.4**	**48.6 ± 8.4**	**45.1 ± 36.0**	**36 ± 9.5**	**12.6 ± 8.2**	**29.8 ± 12.3**	**33.3 ± 15.7**	**78.90%**	**81.7%**

Abbreviations: AMI, amitriptyline; BUP, bupropion; CMI, clomipramine; CTP, citalopram; S-CTP, escitalopram; Dep, depression rating; DLX, duloxetine; DMI, desipramine; FLX, fluoxetine; IMI, imipramine; ITT, findings based on intent-to-treat: at least one dose and one assessment; MTZ, mirtazapine; *N*, responders or cases treated; NRT, (nortriptyline;); Pbo, placebo; PNZ, phenelzine; PRX, paroxetine; RD, response difference, %-Improvement, drug—placebo; RR, response rate ratio, drug–placebo; Rx, drug; SLG, selegiline; SRT, sertraline; TCP, tranylcypromine; TZD, trazodone; VNX, venlafaxine; dVNX, desvenlafaxine.
[a]Corrected for 3677 repeated placebo cases in 35 trials.
There are a total of 142 drug–placebo comparisons from 107 studies, involving a total of 26 948 patients.

Figure 1 Summary of meta-analytically computed relative rates (RR) of response after randomization to drug *vs* placebo) with 95% confidence intervals (CI, horizontal bars when $n \geqslant 2$ trials per drug) for controlled trials of each of 19 antidepressants (with numbers of trials on the left axis, and numerical values on the right). Drugs are listed by descending apparent efficacy, with symbol-size approximately proportional to weighting by trials per drug. The vertical solid line = null (1.0); vertical dotted line and solid diamond (width = CI) = pooled RR for all agents tested (*$p < 0.05$; **$p \leqslant 0.01$; ***$p \leqslant 0.001$). Overall pooled RR = 1.42 (CI: 1.38–1.48), indicating an average of 42% superiority of antidepressants over placebos. Note that phenelzine, clomipramine, tranylcypromine and trazodone ($n = 1$ trial each) appear to be outliers.

subjects per trial, on responses to drugs and placebos and their ratio (Figure 2). Both sites and subjects per trial increased between 1983 and 2010 (Figures 2a and b). Responses in placebo-arms of trials increased across the same era, but responses to antidepressant drugs decreased slightly (Figures 2c and d) to yield highly significant decreases in the drug–placebo responder rate-ratio (RR) across the same years (Figure 2e); drug–placebo differences in rates of response or of percentage improvement also declined (not shown). Responder RR values also declined significantly as the number of subjects per trial (Figure 2f) as well as sites per trial (not shown) increased. Trial-duration also increased significantly across the years sampled ($r_s = 0.603$, $p < 0.0001$), and longer-trials led selectively to larger responses with placebos (slope, 1.90 (CI: 0.85–2.95), $p = 0.0005$) than with drugs (0.92 (–0.15 to 1.99)).

Multivariate linear regression modeling indicated that the following factors were associated significantly and independently with more recent reporting-years, as follows: (a) more drugs other than TCAs, (b) larger numbers of subjects per trial (or sites per trial), (c) lower response to drugs, (d) greater responses to placebo, (e) higher proportions of depressed women, and (f) longer trials. However, there was no evidence of secular changes in ratings of depression-severity at intake, IMI-eq drug-dose, or dropout rates (Table 3).

Table 2 Comparisons Among Antidepressant Types and Reporting Years

Measures	All drugs	TCAs	MAO inhibitors	SRIs	SNRIs	Atypicals	Early (1983–1997)	Late (1998–2010)
Trials (*n*):	124	31	5	47	30	11	57	67
Responder RR								
Pooled RR	1.42	1.62	1.39	1.37	1.40	1.25	1.63	1.32
95% CI	1.38–1.48	1.47–1.78	1.11–1.48	1.27–1.48	1.30–1.51	1.15–1.35	1.49–1.78	1.26–1.38
z-Score	15.7	9.86	2.88	8.28	8.60	5.50	10.9	12.6
p-Value	<0.0001	0.0001	0.004	0.0001	0.0001	0.0001	0.0001	0.0001
Responder RD								
Pooled RD	16.3%	21.4%	12.1%	14.6%	16.4%	11.9%	20.7%	13.4%
95% CI	14.4–18.2	17.7–25.1	3.58–20.5	11.5–17.7	12.3–20.5	8.15–15.7	17.5–23.8	11.1–15.6
z-Score	16.6	11.3	2.79	9.21	7.81	6.19	12.9	11.7
p-Value	<0.0001	0.0001	0.005	0.0001	0.0001	0.0001	0.0001	0.0001
Improvement RD								
Pooled RD	12.5%	16.2%	16.0%	11.5%	9.80%	12.8%	16.8%	9.80%
95% CI	11.0–14.1	13.3–19.1	0.98–33.0	8.70–14.2	7.14–12.5	8.19–17.4	14.5–19.2	7.17–10.2
Paired-*t*	16.1	11.4	2.62	8.40	7.54	6.18	14.5	11.3
p-Value	<0.0001	<0.0001	0.05	<0.0001	<0.0001	0.0001	<0.0001	<0.0001
NNT	8.0	6.2	6.2	8.7	10.2	7.8	6.0	10.2
95%CI	7.1–9.1	5.2–7.5	3.0–102	7.0–11.5	8.0–14.0	5.7–12.2	5.2–6.9	9.8–13.9

Abbreviations: MAO, monoamine oxidase; NNT, number-needed-to-treat (reciprocal of RD); SNRI, serotonin-norepinephrine reuptake inhibitors; SRI, serotonin-reuptake inhibitor; TCA, tricyclic antidepressants.
Based on meta-analytic computation of ratios of responder rates with antidepressants/placebos (RR) or their differences (RD), and on differences in percentage-improvement in initial depression ratings with drug—placebo for 19 antidepressants tested for efficacy in 124 trials summarized in Table 1. Note that most CIs overlap between agents, and that ranking by apparent potency varies among the three outcome measures, but that TCAs appear to be consistently more effective than other types of antidepressants, including SNRIs, SRIs, MAO inhibitors, or atypical agents (bupropion, mirtazapine, and trazodone). Also, early trials (reported in 1983–1997 *vs* 1998–2010) yield consistently greater drug–placebo differences.

Finally, we carried out a preliminary, hypothesis-generating *post hoc* analysis of deciles of meta-analytically determined drug–placebo responder RR values as well as depression-improvement RD values *vs* trial-sizes (not shown). By both outcome measures, the apparently optimal number was 2–10 sites per trial, and 30–75 subjects per trial, with lower efficacy found at both lower and higher counts.

DISCUSSION

The present findings are congruent with reviews discussed above indicating that antidepressant drug-*vs*-placebo differences in published reports of controlled trials are generally moderate (Baldessarini, 2005; Gartlehner *et al*, 2008; Kirsch *et al*, 2008; Tsapakis *et al*, 2008; Bridge *et al*, 2009; Wooley *et al*, 2009; Masi *et al*, 2010; Pigott *et al*, 2010; Khin *et al*, 2011). This conclusion was reached in the previous literature despite typical reliance on initial improvement on scale ratings rather than less readily achieved clinical remission, and despite growing evidence of publication bias toward underreporting of studies without significant drug–placebo differences (Ioannidis, 2008; Turner *et al,* 2008). Following nearly identical mid-range, initial depression ratings across drug and placebo arms and reporting-years, the crude response rates in the reports reviewed here averaged 54% with FDA-approved antidepressants that are employed clinically to treat major depression in the United States, compared with 37% with placebo. These differences consistently favor active drugs, but by only 17%.

The present findings also support the broad consensus that drug–placebo differences have been declining for a variety of psychotropic drugs in recent decades, making it increasingly difficult to demonstrate efficacy (Khin *et al*, 2011; Yildiz *et al*, 2011a, b). This trend probably has encouraged increased reliance on larger trials (more subjects and collaborating sites) in order to maintain statistical power. Moreover, increasing reliance on complex trials carried out in varied geographic locations and cultures may tend to limit the reliability of research findings (Vázquez *et al*, 2011).

It is evidently widely held that differences in efficacy among specific drugs or types of antidepressants in the treatment of acute episodes of major depressive disorder are generally minor (Healy, 1997; Baldessarini, 2005, 2012; Cipriani *et al*, 2007; Gartlehner *et al*, 2008; Ghaemi, 2008; Pigott *et al*, 2010; Khin *et al*, 2011). The present findings support the conclusion that pooling of data from placebo-controlled trials does not yield clear rankings of specific drugs or drug-types by apparent efficacy (Figure 1). Unexpectedly, however, there were significant differences in reported apparent efficacy between TCAs and newer anti-

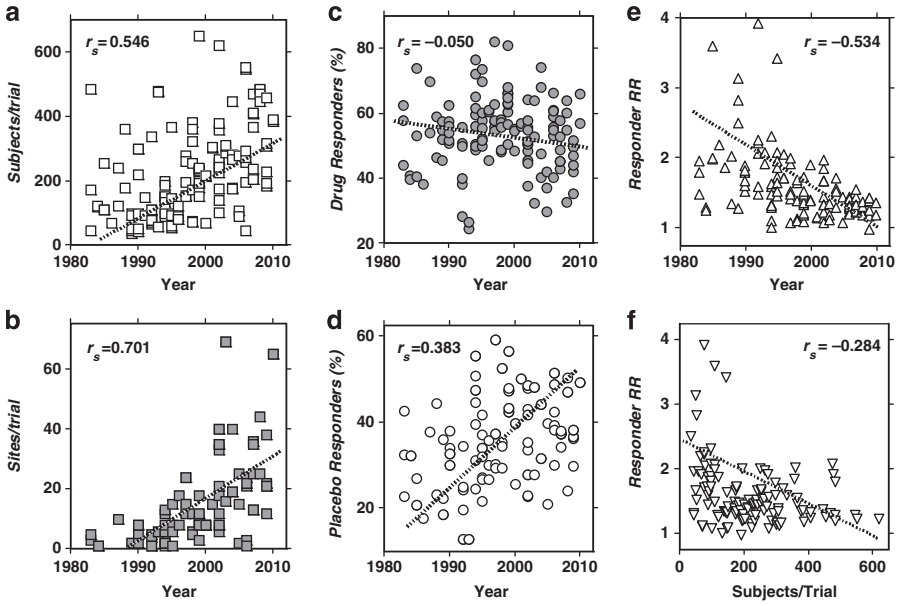

Figure 2 Correlations with Spearman nonparametric correlation coefficients (r_s): (a) Subjects per trial *vs* year of trial reporting ($p<0.0001$); (b) Collaborating sites per trial *vs* year ($p<0.0001$); (c) Meta-analytic responder rate (% of subjects) for antidepressants *vs* year ($p=0.58$); (d) Responder rate (%) for placebo *vs* year ($p<0.0001$); (e) Responder rate ratio (RR: drug–placebo) *vs* year ($p<0.0001$); (f) RR *vs* subjects per trial ($p=0.002$). In addition RR decreased significantly with more sites per trial ($r_s=-0.302$, $p=0.004$). Note that response after randomization to placebo but not antidepressant drugs selectively increased over years, as subject and site counts per trial increased, with corresponding decreases in drug–placebo relative response rate-ratio (RR).

Table 3 Multivariate Linear Regression Model: Factors Associated with Year of Publication of Trial Reports

Factors	Slope function (β) (95% CI)	t-Score	*p*-Value
More subjects per trial	+0.013 (+0.006 to +0.021)	3.40	0.001
More placebo response	+0.183 (+0.067 to +0.300)	3.13	0.002
Less drug response	–0.179 (–0.288 to –0.071)	3.27	0.001
Less use of TCAs	–5.73 (–7.97 to –3.50)	3.13	0.002
More women subjects	+0.155 (+0.053 to +0.258)	3.00	0.003
Longer trials	+0.807 (+0.268 to +1.345)	2.97	0.004

Factors listed are independently and significantly associated with more recent reports among reporting-years (1983–2010) as the continuous outcome measure. Factors *not* associated with more recent trials included: initial depression severity rating, IMI-eq drug dose, and dropout-rate.

depressants (Table 2). We propose that this outcome may reflect important changes in characteristics of clinical trials for depression over the past three decades. These include increasing size and complexity, with selective increases in response rates with placebos and somewhat decreasing responses with antidepressants (Figure 2). It is particularly noteworthy that when placebo-response data from the generally older TCA trials were substituted for those in more recent trials of modern drugs, both types of agents yielded identical meta-analytically pooled RR values. In contrast, we did not find evidence of significant changes over the years in initial ratings of depression-severity (adjusted for variance among rating scales), in approximate IMI-eq antidepressant doses, or in several other measured characteristics of trials (Table 3).

It is increasingly clear that drug–placebo differences in trials of antidepressants and other psychotropic agents have been declining (Gartlehner *et al*, 2008; Ioannidis, 2008; Kirsch *et al*, 2008; Tsapakis *et al*, 2008; Turner *et al*, 2008; Bridge *et al*, 2009; Masi *et al*, 2010; Khin *et al*, 2011; Vázquez *et al*, 2011; Yildiz *et al*, 2011a, b). In accord with recent findings in controlled treatment trials for mania (Yildiz *et al*, 2011a, b), a secular increase in sites and participants per trial was associated, selectively, with rising placebo-associated response rates, resulting in declining drug–placebo contrasts or effect-size (Figure 2; Table 3). We propose that this tendency may, at least in part, reflect declining quality-control and greater heterogeneity of diagnostic and clinical assessments in large, complex, multi-site trials, particularly when dissimilar cultures are involved and local standardization of assessment methods is limited (Yildiz *et al*, 2011a, b; Vázquez *et al*, 2011). We propose that selective increases in response rates associated with randomized placebo-treatment might reflect 'regression-to-mean' effects (Anderson, 1990; Bland and Altman, 1994) or random outcomes. Placebo-associated responses have increased from former levels of 20 to 30% to current levels of 30 to 50%, and to as high as 59.2% in a 1997 trial involving paroxetine (Lecrubier *et al*, 1997).

Alternative factors that may contribute to the observed secular trends include changes in the types of patients recruited into antidepressant trials, including less severely ill patients willing to accept potential randomization to a placebo, and even partially treated subjects. Levels of training and expertise of personnel providing diagnostic and symptom-rating assessments may also have declined. In addition, trials have become longer over the years sampled (Table 2), requiring more clinical assessments with greater risk of measurement-variance, and providing more clinical

contact and more time for spontaneous improvement—all of which may favor responses associated with placebo treatment. Additional technical factors may include less reliance on expert raters, with greater risk of less stable assessments in a very heterogeneous disorder (Healy, 1997).

If the preceding interpretation of the present findings is correct, it suggests several practical considerations for the design and conduct of therapeutic trials for major depression and perhaps other disorders. These include seeking an optimal range of trial-sizes, with redoubled efforts to maximize quality-control, limit placebo-associated responses, and maximize drug–placebo differences. Preliminary analyses of the present data suggest that an optimal range of collaborating sites per trial may be 2–10, and of subjects per trial, about 30–75. Such conservative considerations for the design of future trials may improve outcomes. Additional potential benefits may include reduced time, complexity, and costs, as well as limiting exposure of as many acutely depressed patient-subjects to placebo-treatment as possible.

Limitations of this study include a lack of relevant details in many reports of controlled trials, sometimes including inconsistent reporting of definitions and outcomes for responder rate and percentage improvement, of the number of rating-scale items and of their maximum attainable scores in a few trials. Also, in most trials, exposure times are estimated from nominal protocol requirements since precise, subject-based actual weeks of treatment usually are not stated. Also, numbers of patients with defined outcomes are usually, but not always, based on prevalent intention-to-treat methods, which can limit responses owing to early dropout. Routine reporting of such details would greatly benefit future meta-analyses. Additional limitations to generalization arise from our requirements of peer-review and publication of findings in placebo-controlled trials concerning antidepressants approved and marketed in the United States for acute adult, major depression.

In conclusion, the present meta-analytic review of outcomes of placebo-controlled trials of antidepressants for acute episodes of major depressive disorder found evidence that older antidepressants, particularly TCAs, yielded somewhat superior apparent efficacy to some modern, second-generation agents. However, such nominal differences appear to have been influenced by secular changes in the nature of such trials over the past three decades. These include rising subject- and site-numbers and increasing placebo-associated responses, leading to falling drug–placebo differences or effect-size. We hypothesize that more conservative numbers of subjects and sites, with improved quality-control of trial methods, may paradoxically yield superior results in controlled trials of some psychotropic drugs, and do so more economically. Finally, the lack of major and compelling differences in apparent efficacy among specific antidepressants, and moderate differences among drug-types, suggest that meta-analyses of controlled trials may have limited value in efforts to develop an evidence-basis (Sackett *et al*, 1996) for identifying superior treatments.

ACKNOWLEDGEMENTS

This work was supported by a grant from the Bruce J Anderson Foundation and by the McLean Private Donors Psychopharmacology Research Fund (to RJB), and by a Traveling Research Fellowship from the University of Barcelona and the Fundación Española de Psiquiatría y Salud Mental (to JU). Leonardo Tondo, MD provided helpful technical advice.

DISCLOSURE

The authors declare no conflict of interest.

REFERENCES

Amsterdam JD (2003). Double-blind, placebo-controlled trial of the safety and efficacy of selegiline transdermal system without dietary restrictions in patients with major depressive disorder. *J Clin Psychiatry* **64**: 208–214.

Anderson B (1990). *Methodological Errors in Medical Research: Incomplete Catalog*. Blackwell: New York.

Anderson IM (2001). Meta-analytical studies on new antidepressants. *Br Med Bull* **57**: 161–178.

Andreoli V, Caillard V, Deo RS, Rybakowski JK, Versiani M (2002). Reboxetine, new noradrenaline selective antidepressant, is at least as effective as fluoxetine in the treatment of depression. *J Clin Psychopharmacol* **22**: 393–399.

Baldessarini RJ (2005). Drug therapy of depression and anxiety disorders. Chapter 17. In: Brunton LL, Lazo JS, Parker KL (eds). *Goodman and Gilman's The Pharmacological Basis of Therapeutics*, 11th edn. McGraw-Hill: New York, pp 429–459.

Baldessarini RJ (2012). *Chemotherapy in Psychiatry*, 3rd edn. New York: Springer-Verlag (in press).

Baldessarini RJ, Tondo L, Ghiani C, Lepri B (2010). Discontinuation rate *vs* recurrence risk following long-term antidepressant treatment in major depressive disorder patients. *Am J Psychiatry* **167**: 934–941.

Bakish D, Bradwejn J, Nair N, McClure J, Remick R, Bulger L (1992). Comparison of moclobemide, amitriptyline and placebo in depression: Canadian multicenter study. *Psychopharmacology* **106**(Suppl): S98–S101.

Ban TA, Aguglia E, Batista R, Castillo A, Lipcsey A, Macher JP *et al* (1998). Clinical efficacy of reboxetine: comparative study with desipramine, with methodological considerations. *Hum Psychopharmacol* **13**(Suppl): S29–S39.

Bjerkenstedt L, Edman GV, Alken RG, Mannel M (2005). Hypericum extract LI-160 and fluoxetine in mild to moderate depression: randomized, placebo-controlled multi-center study in outpatients. *Eur Arch Psychiatry Clin Neurosci* **255**: 40–47.

Bland JM, Altman DG (1994). Regression towards the mean. *BMJ* **308**: 1499–1500.

Bodkin A, Amsterdam JD (2002). Transdermal selegiline in major depression: a double-blind, placebo-controlled, parallel-group study in outpatients. *Am J Psychiatry* **159**: 1869–1875.

Boyer P, Montgomery S, Lepola U, Germain JM, Brisard C, Ganguly R *et al* (2008). Efficacy, safety, and tolerability of fixed-dose desvenlafaxine 50 and 100 mg/day for major depressive disorder in a placebo-controlled trial. *Int Clin Psychopharmacol* **23**: 243–253.

Brannan SK, Mallinckrodt CH, Brown EB, Wohlreich MM, Watkin JG, Schatzberg AF (2005). Duloxetine 60 mg once-daily in the treatment of painful physical symptoms in patients with major depressive disorder. *J Psychiatr Res* **39**: 43–53.

Bremner JD (1995). Double-blind comparison of Org-3770, amitriptyline, and placebo in major depression. *J Clin Psychiatry* **56**: 519–525.

Bridge JA, Birmaher B, Iyengar S, Barbe RP, Brent DA (2009). Placebo response in randomized controlled trials of

antidepressants for pediatric major depressive disorder. *Am J Psychiatry* **166**: 42–49.
Burke WJ, Gergel I, Bose A (2002). Fixed-dose trial of the single isomer SSRI escitalopram in depressed outpatients. *J Clin Psychiatry* **63**: 331–336.
Carman JS, Ahdieh H, Wyatt-Knowles E, Warga E, Panagides J (1991). Controlled study of mianserin in moderately to severely depressed outpatients. *Psychopharmacol Bull* **27**: 135–139.
Cassano GB, Heinze G, Lôo H, Mendlewicz J, Sousa MP (1996). Double-blind comparison of tianeptine, imipramine and placebo in the treatment of major depressive episodes. *Eur Psychiatry* **11**: 254–259.
Cipriani A, Geddes JR, Furukawa TA, Barbui C (2007). Metareview on short-term effectiveness and safety of antidepressants for depression: evidence-based approach to inform clinical practice. *Can J Psychiatry* **52**: 553–562.
Claghorn J, Gershon S, Goldstein BJ, Behrnetz S, Bush DF, Huitfeldt B (1983). Double-blind evaluation of zimelidine in comparison to placebo and amitriptyline in patients with major depressive disorder. *Prog Neuropsychopharmacol Biol Psychiatry* **7**: 367–382.
Claghorn JL, Earl CQ, Walczak DD, Stoner KA, Wong LF, Kanter D *et al.* (1996). Fluvoxamine maleate in the treatment of depression: single-center, double-blind, placebo-controlled comparison with imipramine in outpatients. *J Clin Psychopharmacol* **16**: 113–120.
Claghorn JL, Kiev A, Rickels K, Smith WT, Dunbar GC (1992). Paroxetine *vs* placebo: double-blind comparison in depressed patients. *J Clin Psychiatry* **53**: 434–438.
Claghorn JL, Lesem MD (1995). Double-blind placebo-controlled study of Org-3770 in depressed outpatients. *J Affect Disord* **34**: 165–171.
Clayton AH, Croft HA, Horrigan JP, Wightman DS, Krishen A, Richard NE *et al* (2006). Bupropion extended release compared with escitalopram: effects on sexual functioning and antidepressant efficacy in two randomized, double-blind, placebo-controlled studies. *J Clin Psychiatry* **67**: 736–746.
Cohn CK, Robinson DS, Roberts DL, Schwiderski UE, O'Brien K, Ieni JR (1996). Responders to antidepressant drug treatment: study comparing nefazodone, imipramine, and placebo in patients with major depression. *J Clin Psychiatry* **57**(Suppl 2): 15–18.
Cohn JB, Wilcox C (1985). Comparison of fluoxetine, imipramine, and placebo in patients with major depressive disorder. *J Clin Psychiatry* **46**: 26–31.
Cohn JB, Wilcox CS (1992). Paroxetine in major depression: double-blind trial with imipramine and placebo. *J Clin Psychiatry* **53**(Suppl): 52–56.
Coleman CC, Cunningham LA, Foster VJ, Batey SR, Donahue RM, Houser TL *et al* (1999). Sexual dysfunction associated with the treatment of depression: placebo-controlled comparison of bupropion sustained release and sertraline treatment. *Ann Clin Psychiatry* **11**: 205–215.
Coleman CC, King BR, Bolden-Watson C, Book MJ, Segraves RT, Richard N *et al* (2001). Placebo-controlled comparison of the effects on sexual functioning of bupropion sustained release and fluoxetine. *Clin Ther* **23**: 1040–1058.
Corrigan MH, Denahan AQ, Wright CE, Ragual RJ, Evans DL (2000). Comparison of pramipexole, fluoxetine, and placebo in patients with major depression. *Depress Anxiety* **11**: 58–65.
Croft H, Settle Jr E, Houser T, Batey SR, Donahue RM, Ascher JA (1999). Placebo-controlled comparison of the antidepressant efficacy and effects on sexual functioning of sustained-release bupropion and sertraline. *Clin Ther* **21**: 643–658.
Cunningham LA (1997). Once-daily venlafaxine extended release (XR) and venlafaxine immediate release (IR) in outpatients with major depression. *Ann Clin Psychiatry* **9**: 157–164.
Cunningham LA, Borison RL, Carman JS, Chouinard G, Crowder JE, Diamond BI *et al* (1994). Comparison of venlafaxine, trazodone, and placebo in major depression. *J Clin Psychopharmacol* **14**: 99–106.
Cutler AJ, Montgomery SA, Feifel D, Lazarus A, Astrom M, Brecher M (2009). Extended release quetiapine fumarate monotherapy in major depressive disorder: placebo- and duloxetine-controlled study. *J Clin Psychiatry* **70**: 526–539.
Davidson JR (2002). Effect of Hypericum perforatum (St John's wort) in major depressive disorder: randomized controlled trial. *JAMA* **287**: 1807–1814.
DeMartinis NA, Yeung PP, Entsuah R, Manley AL (2007). Double-blind, placebo-controlled study of the efficacy and safety of desvenlafaxine succinate in the treatment of major depressive disorder. *J Clin Psychiatry* **68**: 677–688.
Detke MJ, Lu Y, Goldstein DJ, Hayes JR, Demitrack MA (2002a). Duloxetine, 60 mg once daily, for major depressive disorder: a randomized double-blind placebo-controlled trial. *J Clin Psychiatry* **63**: 308–315.
Detke MJ, Lu Y, Goldstein DJ, McNamara RK, Demitrack MA (2002b). Duloxetine 60 mg once daily dosing versus placebo in the acute treatment of major depression. *J Psychiatr Res* **36**: 383–390.
Detke MJ, Wiltse CG, Mallinckrodt CH, McNamara RK, Demitrack MA, Bitter I (2004). Duloxetine in the acute and long-term treatment of major depressive disorder: a placebo- and paroxetine-controlled trial. *Eur Neuropsychopharmacol* **14**: 457–470.
Doogan DP, Langdon CJ (1994). Double-blind, placebo-controlled comparison of sertraline and dothiepin in the treatment of major depression in general practice. *Int Clin Psychopharmacol* **9**: 95–100.
Fabre LF (1992). Six-week, double-blind trial of paroxetine, imipramine, and placebo in depressed outpatients. *J Clin Psychiatry* **53**(Suppl): 40–43.
Fabre LF, Abuzzahab FS, Amin M, Claghorn JL, Mendels J, Petrie WM *et al* (1995). Sertraline safety and efficacy in major depression: a double-blind fixed-dose comparison with placebo. *Biol Psychiatry* **38**: 592–602.
Fava M, Alpert J, Nierenberg AA, Mischoulon D, Otto MW, Zajecka J *et al* (2005). Double-blind, randomized trial of St John's wort, fluoxetine, and placebo in major depressive disorder. *J Clin Psychopharmacol* **25**: 441–447.
Fava M, Amsterdam JD, Deltito JA, Salzman C, Schwaller M, Dunner DL (1998). Double-blind study of paroxetine, fluoxetine, and placebo in outpatients with major depression. *Ann Clin Psychiatry* **10**: 145–150.
Feiger AD (1996). Double-blind comparison of gepirone extended release, imipramine, and placebo in the treatment of outpatient major depression. *Psychopharmacol Bull* **32**: 659–665.
Feiger AD, Rickels K, Rynn MA, Zimbroff DS, Robinson DS (2006). Selegiline transdermal system for the treatment of major depressive disorder: an 8-week, double-blind, placebo-controlled, flexible-dose titration trial. *J Clin Psychiatry* **67**: 1354–1361.
Feiger AD, Tourian KA, Rosas GR, Padmanabhan SK (2009). Placebo-controlled study evaluating the efficacy and safety of flexible-dose desvenlafaxine treatment in outpatients with major depressive disorder. *CNS Spectrums* **14**: 41–50.
Feighner JP (1992). Double-blind comparison of paroxetine, imipramine and placebo in depressed outpatients. *Int Clin Psychopharmacol* **6**(Suppl 4): 31–35.
Feighner JP, Aden GC, Fabre LF, Rickels K, Smith WT (1983). Comparison of alprazolam, imipramine, and placebo in the treatment of depression. *JAMA* **249**: 3057–3064.
Feighner JP, Boyer WF, Merideth CH, Hendrickson GG (1989a). Double-blind comparison of fluoxetine, imipramine and placebo in outpatients with major depression. *Int Clin Psychopharmacol* **4**: 127–134.
Feighner JP, Boyer WF, Meredith CH, Hendrickson GG (1989b). Placebo-controlled inpatient comparison of fluvoxamine maleate and imipramine in major depression. *Int Clin Psychopharmacol* **4**: 239–244.

Feighner JP, Cohn JB, Fabre Jr LF, Fieve RR, Mendels J, Shrivastava RK *et al* (1993). Study comparing paroxetine, placebo, and imipramine in depressed patients. *J Affect Disord* **28**: 71–79.

Feighner JP, Overo K (1999). Multicenter, placebo-controlled, fixed-dose study of citalopram in moderate-to-severe depression. *J Clin Psychiatry* **60**: 824–830.

Fontaine R, Ontiveros A, Elie R, Kensler TT, Roberts DL, Kaplita S *et al* (1994). Double-blind comparison of nefazodone, imipramine, and placebo in major depression. *J Clin Psychiatry* **55**: 234–241.

Gartlehner G, Gaynes BN, Hansen RA, Thieda P, DeVaugh-Geiss A, Krebs EE *et al* (2008). Comparative benefits and harms of second-generation antidepressants. *Ann Intern Med* **149**: 734–750.

Gastpar M, Singer A, Zeller K (2006). Comparative efficacy and safety of a once-daily dosage of hypericum extract STW3-VI and citalopram in patients with moderate depression: double-blind, randomized, multicentre, placebo-controlled study. *Pharmacopsychiatry* **39**: 66–75.

Gelenberg AJ, Wojcik JD, Falk WE, Baldessarini RJ, Zeisel SH, Schoenfeld D *et al.* (1990). Tyrosine for depression: double-blind trial. *J Affect Disord* **19**: 125–132.

Ghaemi SN (2008). Why antidepressants are not antidepressants. *Bipolar Disord* **10**: 957–968.

Golden RN, Nemeroff CB, McSorley P, Pitts CD, Dube EM (2002). Efficacy and tolerability of controlled-release and immediate-release paroxetine in the treatment of depression. *J Clin Psychiatry* **63**: 577–584.

Goldstein DJ, Lu Y, Detke MJ, Wiltse C, Mallinckrodt C, Demitrack MA (2004). Duloxetine in the treatment of depression: double-blind placebo-controlled comparison with paroxetine. *J Clin Psychopharmacol* **24**: 389–399.

Goldstein DJ, Mallinckrodt C, Lu Y, Demitrack MA (2002). Duloxetine in the treatment of major depressive disorder: double-blind clinical trial. *J Clin Psychiatry* **63**: 225–231.

Guelfi JD, White C, Hackett D, Guichoux JY, Magni G (1995). Effectiveness of venlafaxine in patients hospitalized for major depression and melancholia. *J Clin Psychiatry* **56**: 450–458.

Guy W (1976). *ECDEU Assessment Manual for Psychopharmacology*. US Department HEW publication (ADM) 76-338: Rockville, MD. National Institute of Mental Health, pp 218–222.

Hamilton M (1960). Rating scale for depression. *J Neurol Neurosurg Psychiatry* **23**: 56–62.

Healy D (1997). *The Antidepressant Era*. Harvard University: Cambridge, MA.

Hewett K, Gee MD, Krishen A, Wunderlich HP, Le Clus A, Evoniuk G *et al* (2010). Double-blind, placebo-controlled comparison of antidepressant efficacy and tolerability of bupropion-XR and venlafaxine-XR. *J Psychopharmacol* **24**: 1209–1216.

Hollyman JA, Freeling P, Paykel ES, Bhat A, Sedgwick P (1988). Double-blind placebo-controlled trial of amitriptyline among depressed patients in general practice. *J R Coll Gen Pract* **38**: 393–397.

Ioannidis JP (2008). Effectiveness of antidepressants: an evidence myth constructed from a thousand randomized trials? *Philos Ethics Humanit Med* **3**: 14–22.

Itil TM, Shrivastava RK, Mukherjee S, Coleman BS, Michael ST (1983). Double-blind placebo-controlled study of fluvoxamine and imipramine in out-patients with primary depression. *Br J Clin Pharmacol* **15**(Suppl 3): S433S–S4338.

Jefferson JW, Rush AJ, Nelson JC, VanMeter SA, Krishen A, Hampton KD *et al* (2006). Extended-release bupropion for patients with major depressive disorder presenting with symptoms of reduced energy, pleasure, and interest: findings from a randomized, double-blind, placebo-controlled study. *J Clin Psychiatry* **67**: 865–873.

Khan A, Fabre LF, Rudolph R (1991). Venlafaxine in depressed outpatients. *Psychopharmacol Bull* **27**: 141–144.

Khan A, Upton GV, Rudolph RL, Entsuah R, Leventer SM (1998). Use of venlafaxine in the treatment of major depression and major depression associated with anxiety: dose-response study. *J Clin Psychopharmacol* **18**: 19–25.

Khan M (1995). Randomized, double-blind, placebo-controlled, 5-weeks study of Org-3770 (mirtazapine) in major depression. *Hum Psychopharmacol* **10**(Suppl): S119–S124.

Khin NA, Chen YF, Yang Y, Yang P, Laughren TP (2011). Exploratory analyses of efficacy data from major depressive disorder trials submitted to the US Food and Drug Administration in support of new drug applicatons. *J Clin Psychiatry* **72**: 464–472.

Kiev A (1992). Double-blind, placebo-controlled study of paroxetine in depressed outpatients. *J Clin Psychiatry* **53**(Suppl): S27–S29.

Kirsch I, Deacon B, Huedo-Medina TB, Moore TJ, Johnson BT (2008). Initial severity and antidepressant benefits: meta-analysis of data submitted to the Food and Drug Administration. *PLoS Med* **5**: e45–e53.

Laakman G, Faltermaier-Temizel M, Bossert-Zaudig S, Baghai T, Lorkowski G (1995). Treatment of depressive outpatients with lorazepam, alprazolam, amitriptyline and placebo. *Psychopharmacology* **120**: 109–115.

Larsen JK, Holm P, Høyer E, Mejlhede A, Mikkelsen PL, Olesen A *et al* (1989). Moclobemide and clomipramine in reactive depression: placebo-controlled randomized clinical trial. *Acta Psychiatr Scand* **79**: 530–536.

Lecrubier Y, Bourin M, Moon CA, Schifano F, Blanchard C, Danjou P *et al* (1997). Efficacy of venlafaxine in depressive illness in general practice. *Acta Psychiatr Scand* **95**: 485–493.

Lepola UM, Loft H, Reines EH (2003). Escitalopram (10–20 mg/day) is effective and well tolerated in a placebo-controlled study in depression in primary care. *Int Clin Psychopharmacol* **18**: 211–217.

Lieberman DZ, Montgomery SA, Tourian KA, Brisard C, Rosas G, Padmanabhan K *et al* (2008). A pooled analysis of two placebo-controlled trials of desvenlafaxine in major depressive disorder. *Int Clin Psychopharmacol* **23**: 188–197.

Liebowitz MR, Manley AL, Padmanabhan SK, Ganguly R, Tummala R, Tourian KA (2008). Efficacy, safety, and tolerability of desvenlafaxine 50 mg/day and 100 mg/day in outpatients with major depressive disorder. *Curr Med Res Opin* **24**: 1877–1890.

Liebowitz MR, Yeung PP, Entsuah R (2007). Randomized, double-blind, placebo-controlled trial of desvenlafaxine succinate in adult outpatients with major depressive disorder. *J Clin Psychiatry* **68**: 1663–1672.

Lineberry CG, Johnston JA, Raymond RN, Samara B, Feighner JP, Harto NE *et al* (1990). Fixed-dose (300 mg) efficacy study of bupropion and placebo in depressed outpatients. *J Clin Psychiatry* **51**: 194–199.

Lydiard RB, Stahl SM, Hertzman M, Harrison WM (1997). Double-blind, placebo-controlled study comparing the effects of sertraline *vs* amitriptyline in the treatment of major depression. *J Clin Psychiatry* **58**: 484–491.

Masi G, Liboni F, Brovedani P (2010). Pharmacotherapy of major depressive disorder in adolescents. *Expert Opin Pharmacother* **11**: 375–386.

Massana J (1998). Reboxetine vs fluoxetine: overview of efficacy and tolerability. *J Clin Psychiatry* **59**(Suppl 14): 8–10.

Mendels J, Kiev A, Fabre LF (1999). Double-blind comparison of citalopram and placebo in depressed outpatients with melancholia. *Depress Anxiety* **9**: 54–60.

Mendels J, Schless AP (1986). Comparative efficacy of alprazolam, imipramine, and placebo administered once a day in treating depressed patients. *J Clin Psychiatry* **47**: 357–361.

Miller SM, Naylor GJ, Murtagh M, Winslow G (1989). Double-blind comparison of paroxetine and placebo in the treatment of depressed patients in a psychiatric outpatient clinic. *Acta Psychiatr Scand Suppl* **350**: 143–144.

Montgomery SA, Åsberg M (1979). New depression scale designed to be sensitive to change. *Br J Psychiatry* **134**: 382–389.

Moreno RA, Teng CT, Almeida KM, Tavares JH (2006). Hypericum perforatum *vs* fluoxetine in the treatment of mild to moderate depression: randomized double-blind trial in a Brazilian sample. *Rev Bras Psiquiatr* **28**: 29–32.

Mynors-Wallis LM, Gath DH, Lloyd-Thomas AR, Tomlinson D (1995). Randomized controlled trial comparing problem-solving treatment with amitriptyline and placebo for major depression in primary care. *BMJ* **310**: 441–445.

Nemeroff CB, Thase ME (2007). Double-blind, placebo-controlled comparison of venlafaxine and fluoxetine treatment in depressed outpatients. *J Psychiatr Res* **41**: 351–359.

Nierenberg AA, Greist JH, Mallinckrodt CH, Prakash A, Sambunaris A, Tollefson GD *et al* (2007). Duloxetine *vs* escitalopram and placebo in the treatment of patients with major depressive disorder: onset of antidepressant action, non-inferiority study. *Curr Med Res Opin* **23**: 401–416.

Papakostas GI, Thase ME, Fava M, Nelson JC, Shelton RC (2007). Are antidepressant drugs that combine serotonergic and noradrenergic mechanisms of action more effective than the selective serotonin reuptake inhibitors in treating major depressive disorder? Meta-analysis of studies of newer agents. *Biol Psychiatry* **62**: 1217–1227.

Perahia DG, Wang F, Mallinckrodt CH, Walker DJ, Detke MJ (2006). Duloxetine in the treatment of major depressive disorder: placebo- and paroxetine-controlled trial. *Eur Psychiatry* **21**: 367–378.

Philipp M, Kohnen R, Hiller KO (1999). Hypericum extract *vs* imipramine or placebo in patients with moderate depression: randomized multicenter study of treatment for eight weeks. *BMJ* **319**: 1534–1538.

Pigott HE, Leventhal AM, Alter GS, Boren JJ (2010). Efficacy and effectiveness of antidepressants: current status of research. *Psychother Psychosom* **79**: 267–279.

Pitts WM, Fann WE, Halaris AE, Dressler DM, Sajadi C, Snyder S *et al* (1983). Bupropion in depression: tri-center placebo-controlled study. *J Clin Psychiatry* **44**(Part 2): 95–100.

Quitkin FM, McGrath PJ, Stewart JW, Harrison W, Wager SG, Nunes E *et al* (1989). Phenelzine and imipramine in mood reactive depressives: further delineation of the syndrome of atypical depression. *Arch Gen Psychiatry* **46**: 787–793.

Reimherr FW, Chouinard G, Cohn CK, Cole JO, Itil TM, LaPierre YD *et al* (1990). Antidepressant efficacy of sertraline: double-blind, placebo- and amitriptyline-controlled, multicenter comparison study in outpatients with major depression. *J Clin Psychiatry* **51**(Suppl B): S18–S27.

Reimherr FW, Cunningham LA, Batey SR, Johnston JA, Ascher JA (1998). Multicenter evaluation of the efficacy and safety of 150 and 300 mg/day sustained-release bupropion tablets *vs* placebo in depressed outpatients. *Clin Ther* **20**: 505–516.

Rickels K, Amsterdam J, Clary C, Fox I, Schweizer E, Weise C (1992). Efficacy and safety of paroxetine compared with placebo in outpatients with major depression. *J Clin Psychiatry* **53**(Suppl): S30–S32.

Rickels K, Chung HR, Csanalosi IB, Hurowitz AM, London J, Wiseman K *et al* (1987). Alprazolam, diazepam, imipramine, and placebo in outpatients with major depression. *Arch Gen Psychiatry* **44**: 862–866.

Rickels K, Feighner JP, Smith WT (1985). Alprazolam, amitriptyline, doxepin, and placebo in the treatment of depression. *Arch Gen Psychiatry* **42**: 134–141.

Rickels K, Schweizer E, Clary C, Fox I, Weise C (1994). Nefazodone and imipramine in major depression: a placebo-controlled trial. *Br J Psychiatry* **164**: 802–805.

Roth D, Mattes J, Sheehan KH, Sheehan DV (1990). Double-blind comparison of fluvoxamine, desipramine and placebo in outpatients with depression. *Prog Neuropsychopharmacol Biol Psychiatry* **14**: 929–939.

Rudolph RL, Fabre LF, Feighner JP, Rickels K, Entsuah R, Derivan AT (1998). Randomized, placebo-controlled, dose-response trial of venlafaxine hydrochloride in the treatment of major depression. *J Clin Psychiatry* **59**: 116–122.

Rudolph RL, Feiger AD (1999). Double-blind, randomized, placebo-controlled trial of once-daily venlafaxine extended release (XR) and fluoxetine for the treatment of depression. *J Affect Disord* **56**: 171–181.

Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS (1996). Evidence based medicine: what it is and what it isn't. *BMJ* **12**: 71–72.

Schweizer E, Feighner J, Mandos LA, Rickels K (1994). Comparison of venlafaxine and imipramine in the acute treatment of major depression in outpatients. *J Clin Psychiatry* **55**: 104–108.

Septien-Velez L, Pitrosky B, Padmanabhan SK, Germain JM, Tourian KA (2007). Randomized, double-blind, placebo-controlled trial of desvenlafaxine succinate in the treatment of major depressive disorder. *Int Clin Psychopharmacol* **22**: 338–347.

Sheehan DV, Nemeroff CB, Thase ME, Entsuah R (2009). Placebo-controlled inpatient comparison of venlafaxine and fluoxetine for the treatment of major depression with melancholic features. *Int Clin Psychopharmacol* **24**: 61–86.

Silverstone PH (1994). Multicenter comparative trial of moclobemide, imipramine and placebo in major depressive disorder. *Int Clin Psychopharmacol* **9**: 109–113.

Silverstone PH, Ravindran A (1999). Once-daily venlafaxine extended release (XR) compared with fluoxetine in outpatients with depression and anxiety. *J Clin Psychiatry* **60**: 22–28.

Smith WT, Glaudin V, Panagides J, Gilvary E (1990). Mirtazapine *vs* amitriptyline *vs* placebo in the treatment of major depressive disorder. *Psychopharmacol Bull* **26**: 191–196.

Stahl SM (2000). Placebo-controlled comparison of the selective serotonin reuptake inhibitors citalopram and sertraline. *Biol Psychiatry* **48**: 894–901.

Thase ME (1997). Efficacy and tolerability of once-daily venlafaxine extended release (XR) in outpatients with major depression. *J Clin Psychiatry* **58**: 393–398.

Tourian KA, Padmanabhan SK, Groark J, Brisard C, Farrington D (2009). Desvenlafaxine 50 and 100 mg/day in the treatment of major depressive disorder: an 8-week, phase III, multicenter, randomized, double-blind, placebo-controlled, parallel-group trial and a *post hoc* pooled analysis of three studies. *Clin Ther* **31**: 1405–1423.

Trivedi MH, Pigotti TA, Perera P, Dillingham KE, Carfagno ML, Pitts CD (2004). Effectiveness of low doses of paroxetine controlled release in the treatment of major depressive disorder. *J Clin Psychiatry* **65**: 1356–1364.

Tsapakis EM, Soldani F, Tondo L, Baldessarini RJ (2008). Efficacy of antidepressants in juvenile depression: meta-analysis. *Br J Psychiatry* **193**: 10–17.

Turner EH, Matthews AM, Linardatos E, Tell RA, Rosenthal R (2008). Selective publication of antidepressant trials and its influence on apparent efficacy. *N Engl J Med* **358**: 252–260.

Vartiainen H, Leinonen E (1994). Double-blind study of mirtazapine and placebo in hospitalized patients with major depression. *Eur Neuropsychopharmacol* **4**: 145–150.

Vázquez G, Baldessarini RJ, Yildiz A, Tamayo J, Tondo L, Salvatore P (2011). Multi-site international collaborative clinical trials in mania: commentary. *Int J Neuropsychopharmacol* **14**: 1013–1016.

Wade A, Michael-Lemming MO, Bang-Hedegaard K (2002). Escitalopram 10 mg/day is effective and well tolerated in a placebo-controlled study in depression in primary care. *Int Clin Psychopharmacol* **17**: 95–102.

Walsh BT, Seidman SN, Sysko R, Gould M (2002). Placebo response in studies of major depression: variable, substantial, and growing. *JAMA* **287**: 1840–1847.

Wernicke JF, Dunlop SR, Dornseif BE, Bosomworth JC, Humbert M (1988). Low-dose fluoxetine therapy for depression. *Psychopharmacol Bull* **24**: 183–188.

Wernicke JF, Dunlop SR, Dornseif BE, Zerbe RL (1987). Fixed-dose fluoxetine therapy for depression. *Psychopharmacol Bull* **23**: 164–168.

White K, Razani J, Cadow B, Gelfand R, Palmer R, Simpson G *et al* (1984). Tranylcypromine vs nortriptyline vs placebo in depressed outpatients: a controlled trial. *Psychopharmacology* **82**: 258–262.

Wilcox CS, Cohn JB, Katz BB, Mijares CP, Guarino JJ, Panagides J *et al* (1994). Double-blind, placebo-controlled study comparing mianserin and amitriptyline in moderately depressed outpatients. *Int Clin Psychopharmacol* **9**: 271–279.

Wooley SB, Cardoni AA, Goethe JW (2009). Last-observation carried forward imputation method in clnical efficacy trials: review of 352 antidepressant studies. *Pharmacotherapy* **29**: 1408–1416.

Yildiz A, Vieta E, Leucht S, Baldessarini RJ (2011a). Efficacy of antimanic treatments: meta-analysis of randomized, controlled trials. *Neuropsychopharmacology* **36**: 375–389.

Yildiz A, Vieta E, Tohen M, Baldessarini RJ (2011b). Placebo effects in treatment trials for acute mania. *Int J Neuropsychopharmacol* **14**: 863–875.