Published in final edited form as:

J Affect Disord. 2012 September ; 140(1): 57–65. doi:10.1016/j.jad.2012.01.031.

Does differential drop-out explain the influence of study design on antidepressant response? A meta-analysis

Bret R. Rutherford^{a,*}, Joel R. Sneed^b, and Steven P. Roose^a

^aColumbia University College of Physicians and Surgeons, New York State Psychiatric Institute, United States

^bQueens College of the City University of New York, United States

Abstract

Background—Response to antidepressants is higher in active comparator relative to placebo-controlled clinical trials. Increased patient expectancy in comparator trials has been hypothesized to explain this finding, but previous analyses have not accounted for the increased dropout observed in placebo-controlled trials.

Methods—A systematic literature review was conducted to identify published antidepressant clinical trials reporting data on intent-to-treat (ITT) as well as completer patient populations. The influence of participant drop-out on observed antidepressant response was investigated by comparing the ITT and completer data sets in separate multilevel meta-analyses of antidepressant response in placebo-controlled and comparator trials.

Results—18 placebo-controlled and 18 active comparator studies were available for analysis. Using the intent-to-treat data, the odds of responding to medication in comparator trials were 1.9 times the odds in placebo-controlled trials (95% CI=1.3–2.7, $p=0.001$). The same pattern was obtained among study completers, in whom the odds of responding to antidepressant medication were 1.9 times higher in comparator as opposed to placebo-controlled study designs (95% CI=1.2–3.0, $p=0.009$).

Limitations—Publication bias, the use of trial-level summary data, and unreported clinical or demographic differences between the ITT and completer patient populations may have influenced the study results.

Conclusions—Increased drop-out in placebo-controlled vs. active comparator studies of antidepressant medications does not appear to explain the difference in response rates between these study types. Rather, increased patient expectancy resulting from the certainty of receiving active medication in comparator trials may lead to improved response rates.

Keywords

Antidepressant; Clinical trial; Patient expectancy; Placebo effect; Major Depressive Disorder

© 2012 Elsevier B.V. All rights reserved.

*Corresponding author at: 1051 Riverside Drive, Box 98, New York, NY 10032, United States. Tel.: +1 212 543 5746; fax: +1 212 543 6100. brr8@columbia.edu (B.R. Rutherford).

Conflict of interest

Dr. Rutherford and Dr. Sneed have no disclosures to report. Dr. Roose reports serving on a Data and Safety Monitoring Board for Medtronic, Inc. This paper has not been previously presented.

1. Introduction

Methodological features of randomized controlled trials (RCTs) for antidepressant medications influence the treatment outcomes observed. For example, a recent analysis of 52 clinical trials submitted to the Food and Drug Administration (FDA) found that the severity of baseline depressive symptoms, the medication dosing schedule, the number of treatment arms, and the percentage of female patients are associated with finding a significant difference between antidepressant medication and placebo (Khan et al., 2004). Observer-rated outcome measures compared to self-reports have been linked to finding significantly larger effect sizes for placebo response in clinical trials (Rief et al., 2009). Other design characteristics such as the presence and duration of a placebo lead-in period (Trivedi and Rush, 1994) and the administration of other medications during the study (Wernicke et al., 2004) may influence outcome.

A series of meta-analyses performed by our group have reported that response to antidepressant medication is significantly greater in comparator (i.e., one medication vs. a different medication) study designs as opposed to placebo-controlled trials (i.e., one or more medications vs. placebo) (Rutherford et al., 2009; Sneed et al., 2008). In 48 placebo-controlled and 42 comparator trials of antidepressants for MDD in adults aged 18–65, the odds of responding to medication in comparator trials were 1.8 times the odds in placebo-controlled trials (95% CI=1.45–2.17, $p<0.001$) (Rutherford et al., 2009). In a different sample of 16 antidepressant RCTs for patients with late life depression, the odds of medication response in comparator trials were also 1.8 times the odds in placebo-controlled trials (Sneed et al., 2008). In neither meta-analysis were any significant differences identified between the patients enrolled in these different study types or the medications used. Thus, patients appear to respond at higher rates to the same medication if it is administered in a comparator trial rather than a placebo-controlled study. This difference is large (15% response rate differential) and similar in magnitude to the difference observed between antidepressant medication and placebo in placebo-controlled RCTs.

Subsequent reports have been consistent with these findings that the certainty of receiving active medication in an antidepressant clinical trial positively influences patient outcome. Analyzing data from 146 placebo-controlled trials of antidepressant medication, Papakostas and Fava (2009) reported that a higher probability of receiving active medication as compared to placebo was associated with significantly greater medication and placebo response. Along the same lines, Sinyor et al. (2010) found that response to antidepressant medication progressively decreases between drug–drug, drug–drug–placebo, and drug–placebo study designs. These data indicate that the more likely a patient is to receive active medication as opposed to placebo, the greater their response to antidepressants and placebo.

One interpretation of these results is that the higher medication response rates in comparator trials may be explained by increased expectancy among participants in these studies (Rutherford et al., 2010a). Patients in comparator trials know they have a 100% chance of receiving an active medication despite not knowing the exact agent, whereas patients in placebo-controlled trials do not know whether they are receiving an active medication. Patient expectancy is hypothesized to be a primary mechanism of placebo effects in antidepressant RCTs, which tend to be large (Rutherford et al., 2010b; Stewart-Williams and Podd, 2004). In addition, greater patient expectancy has been linked to improved antidepressant treatment outcome in the National Institute of Mental Health (NIMH) Treatment of Depression Collaborative Study (TDCRP) as well as a smaller, single-blind trial of reboxetine for 25 subjects with MDD (Krell et al., 2004; Meyer et al., 2002).

Alternatively, other factors may account for the antidepressant response rate differences observed between placebo-controlled and comparator RCTs. In our meta-analysis of antidepressant response in adults with MDD, drop-out was significantly higher in placebo-controlled vs. comparator studies (34.6 ± 14.4 vs. 22.0 ± 9.7 , respectively, $t=5.54$, $df=145$, $p<0.001$) (Rutherford et al., 2010a). Patients in placebo-controlled trials who remain symptomatic may assume they are receiving placebo and decide to exit the study in order to be assured of receiving active treatment. If drop-out occurs when symptoms remain severe, higher depression scores will tend to be observed at endpoint according to last observation carried forward (LOCF) statistical analyses. Thus, it is possible that decreased attrition relative to placebo-controlled studies rather than increased expectancy-based placebo effects may be responsible for the improved medication response observed in comparator studies.

Determining which among these interpretations is correct has considerable public health import. If enhanced placebo effects cause superior outcome in comparator studies, this would lend support to recent initiatives aimed at identifying the mechanisms of expectancy and placebo effects for the purpose of improving the treatment of MDD (National Institutes of Health, 2011). If, on the other hand, differential attrition explains the differences in antidepressant response observed between study types, then efforts to improve antidepressant outcomes should focus on preventing drop-out rather than enhancing patient expectancy.

In the current study, we re-analyzed data from a prior meta-analysis of antidepressant response in adult outpatients with MDD to compare antidepressant response based on intent-to-treat (ITT) and completer patient populations. If increased patient expectancy mediates the effect of study design on antidepressant response, we reasoned that higher antidepressant response should be observed in comparator vs. placebo-controlled trials in both ITT and completer patient populations. On the other hand, if differential attrition explains the effect of study design, then antidepressant response should be higher in comparator vs. placebo-controlled trials for the ITT but not the completer patient populations. We expected to find similar results for the ITT and completer samples, suggesting that expectancy-based placebo effects rather than differential drop-out mediate the effect of study design on antidepressant outcome.

2. Methods

2.1. Search strategy and selection criteria

The general methods for this study have been described in detail in previous publications (Rutherford et al., 2009, 2011; Sneed et al., 2008). Briefly, a Medline search was conducted using the index terms “depression—drug therapy,” “depressive disorder—drug therapy,” and “antidepressant agents,” in addition to the class and individual generic name of all antidepressants. This returned 19,338 results, which were limited to 1) English language articles, 2) publication year from 1985 to present, 3) age group ≥ 18 (to be inclusive), and 4) publication types including clinical trials, controlled clinical trials, meta-analysis, multi-center study, randomized controlled trial, or review, which yielded 2821 journal articles. The year 1985 was chosen to select trials utilizing more rigorous methods. The first author (BRR) conducted a review of these titles to rule out those which were not clinical trials of antidepressants for depression, resulting in 564 titles.

Three judges (BRR, JRS, and SPR) reviewed the 564 titles, sequentially proceeding from article title, to abstract, and finally full paper text, to determine whether they met inclusion or exclusion criteria (see Fig. 1). These evaluations were pooled, and any differences between judges were resolved by discussion. To further ensure all relevant papers were reviewed, the references of all meta analyses and review articles published since 2000

among the 2821 journal articles were searched for pertinent references. In addition, the Cochrane Database of Systematic Reviews was electronically searched using the topics depression, anxiety, and neurosis. This yielded 24 protocols and completed reviews, each of whose references was reviewed to ensure they were among the reviewed trials.

Inclusion criteria stipulated that articles report RCTs of a Food and Drug Administration (FDA) approved antidepressant medication for Major Depressive Disorder (MDD), enroll outpatients aged 18–65, last between 6 and 12 weeks (inclusive), have comparison group of placebo or another FDA-approved antidepressant medication, be written in English, published 1985 or later, and have response rates specified using a standardized outcome measurement (e.g., Hamilton Rating Scale for Depression (HRSD) (Hamilton, 1960), Beck Depression Inventory (BDI) (Beck et al., 1961), Montgomery-Asberg Depression Rating Scale (MADRS) (Montgomery and Asberg, 1979), Clinical Global Impression (CGI) (Guy, 1976)). The primary change from our earlier publication is that included studies were required to report outcome data for participants completing the study in addition to the ITT data. Trials were excluded for enrolling inpatients, pregnant women, subjects who were psychotic, or those defined to have treatment resistant depression. Also excluded were antidepressant augmentation studies and trials requiring as inclusion criteria a specific subtype of Major Depression, a specific medical illness, or an Axis I disorder other than depression.

2.2. Data extraction

Publication information (year of publication, funding source, type of study, number of groups), demographic characteristics of the included subjects (sample size, age, gender, race, clinical characteristics), details of the treatment condition (medication name, mean dose), and outcome data (pre and post-treatment means, standard deviations, response and remission rates) were extracted from each included RCT. Study quality was also measured by determining whether studies reported critical methodological aspects such as (1) concealment of treatment allocation, (2) blinding of outcome assessment, and (3) use of intent to treat data analyses (Juni et al., 2001). Three judges (BRR, JRS, and SPR) extracted the data, and any differences were resolved by consensus.

2.3. Data analyses

Mixed effects logistic regression models were used to analyze the data on antidepressant response rates, similar to the approach taken by Bryk and Raudenbush (Bryk and Raudenbush, 1992), Hox (Hox, 2002), and Haddock, Rindskopf, and Shadish (Haddock et al., 1998). The multilevel logistic regression model is described by two equations: a within-studies equation and a between-studies equation, which accommodates the hierarchical structure of patients nested within medication conditions nested within studies. The following steps were performed in an identical fashion for both the ITT and completer patient samples.

The initial step was to determine whether there is significant variability in medication response rates across studies. To do this, we ignored the nesting within study (medication group) and fit an unconditional model (Model 1). The within-studies equation for Model 1 is

$$\ln(p/[1-p])=B_0$$

where $\ln(p/[1-p])$ is the log odds of medication response and B_0 is a constant that is assumed initially to be the same for all groups within a study. At the between-studies level, the equation is

$$B_0 = G_{00} + U_0,$$

which describes the true medication response rates as varying around a grand mean (G_{00}) with error (U_0). To determine whether there were genuine differences between the studies (heterogeneity) or whether the variation in findings was compatible with chance alone (homogeneity), we examined the Birge ratio, which is calculated by dividing a chi-square by its degrees of freedom (Birge, 1932; Higgins and Thompson, 2002). The value of the Birge ratio is near 1 when there is only random variation between studies, and as the value exceeds 1, the results of a set of studies lack homogeneity (i.e., they are more varied than expected based on sampling error alone).

If there is significant variability in medication response rates across studies, it is possible to test the hypothesis that medication response rates vary depending on study type (i.e., placebo-controlled vs. comparator). To do this, we included study type as a fixed effect in the between-studies equation (Model 2):

$$B_0 = G_{00} + G_{01}(\text{comparator}) + U_0.$$

‘Comparator’ is a dummy variable coded one for comparator trials and zero for placebo-controlled trials. Using this method, odds ratios and estimated probabilities of response to antidepressant medication in the different study types were computed.

Finally, a full interaction model was constructed by adding study duration to the conditional model of antidepressant response rates. In these analyses, 6 and 7 week duration trials were grouped together under the heading of 6 week trials, 8 and 9 week duration trials were grouped as 8 week trials, and 10 to 12 week duration trials were grouped as 12 week trials. Study duration was modeled despite not being the focus of this analysis, because it was found in our original publication to be a significant predictor of antidepressant response rates.

The regression models were estimated using HLM 6. Differences in participant characteristics between trials were investigated using two-tailed independent samples t tests for continuous variables and chi-square (X^2) tests for categorical variables (SPSS version 18).

3. Results

3.1. Characteristics of included trials and subjects

Eighteen placebo-controlled and 18 comparator trials met the study selection criteria, compared to 48 placebo-controlled and 42 comparator trials included in our original analysis (see Table 1). As shown in Table 2 (summary data provided for ITT samples), there were 40 active treatment conditions enrolling 4620 participants in the 18 placebo-controlled RCTs, since many trials compared more than one medication to placebo. There were 36 active treatment conditions enrolling 3069 subjects in the 18 comparator RCTs. Response rates to medication ranged from 49 to 83% (mean 64.0 ± 9.9) in the placebo-controlled and 42–100% (mean 75.2 ± 14.6) in the comparator trials. Among the placebo controlled trials, 62.5% demonstrated significant differences in depression response rates between medication and placebo, while 2.8% of the comparator trials demonstrated significant differences in depression response rates between medications.

As in our original analysis, placebo-controlled relative to comparator trials enrolled individuals who were significantly younger (40.4 ± 2.9 vs. 42.8 ± 3.8 , $t = -2.911$, $df = 61$, $p = 0.005$) and had higher drop-out rates (37.7 ± 13.1 vs. 20.5 ± 9.5 , $t = 5.937$, $df = 62$, $p < 0.001$). Relative to comparator trials, placebo-controlled trials also more frequently employed fixed dosing strategies (16/40 (40%) vs. 2/36 (5.6%), respectively) and utilized exclusively North American as compared to European sites (17/18 (94.4%) vs. 6/18 (33.3%), respectively). The two study types did not differ in subjects' pre-treatment HRSD scores, mean sample size, medications used, gender, prior depressive episodes, response and remission outcome measures, or quality ratings.

3.2. Analysis of antidepressant response rates: ITT sample

In the unconditional model of the ITT antidepressant response rates, variability between studies was over 13 times that expected by chance alone ($\chi^2/df = 495.4/36 = 13.8$). Therefore, the null hypothesis that antidepressant response rates are homogeneous across studies was rejected, and the analysis proceeded with the conditional model. In the conditional model of the ITT antidepressant response rates, study type accounted for 30.4% of the variability observed (see Table 3, $(0.349 - 0.243)/0.349 = 0.304$). The odds of responding to medication in comparator trials were 1.9 times the odds in placebo-controlled trials (95% CI = 1.3–2.7, $p = 0.001$). Study duration accounted for 15.2% of the variability in anti-depressant response rates once study design was taken into account (see Table 3, $(0.243 - 0.206)/0.243 = 0.152$). There were greater odds of medication response in 8 weeks (OR 1.49, CI 1.06–2.10, $p = 0.024$) compared to 6 week duration clinical trials but not 12 weeks (OR 1.65, CI 0.89–3.1, $p = 0.107$) compared to 6 week duration clinical trials. There was no significant difference found between 8 and 12 week trials, nor were significant interactions found between study duration and study type.

3.3. Analysis of antidepressant response rates: Completer sample

In the unconditional model of antidepressant response rates for study completers, variability between studies was over 12 times that expected by chance alone ($\chi^2/df = 444.5/36 = 12.3$). Therefore, the null hypothesis that antidepressant response rates are homogeneous across studies was rejected, and the analysis proceeded with the conditional model. Study type accounted for 17.8% of the variability observed in the completer data set of antidepressant response rates (see Table 4, $(0.501 - 0.412)/0.501 = 0.178$). Among completers, the odds of responding to medication in comparator trials were 1.9 times the odds in placebo-controlled trials (95% CI = 1.2–3.0, $p = 0.009$). Study duration accounted for 9.7% of the variability in antidepressant response rates once study design was taken into account (see Table 4, $(0.412 - 0.372)/0.412 = 0.097$). There was a trend toward greater odds of medication response in 8 weeks (OR 1.48, CI 0.93–2.34, $p = 0.093$) compared to 6 week duration clinical trials but not 12 weeks (OR 1.95, CI 0.85–4.48, $p = 0.112$) compared to 6 week duration clinical trials. Similar to the ITT data, there was no significant difference found between 8 and 12 week trials, nor were significant interactions found between study duration and study type.

4. Discussion

Results from this re-analysis of a large sample of antidepressant clinical trials do not support differential drop-out rates as an explanation of outcome differences based on study design. Limiting the analysis to study completers, response to antidepressant medication remained significantly higher in comparator studies vs. placebo-controlled trials (OR 1.9, 95% CI = 1.2–3.0, $p = 0.009$). This suggests that patients in comparator trials experience enhanced expectancy-based placebo effects resulting from their certainty of receiving active medication. If increased drop-out among participants in placebo-controlled trials explained

their decreased response rates, then one would not expect to find significant differences among patients completing the two types of studies.

Patient expectancy of improvement is affected by antidepressant study design and in turn influences treatment outcome for patients with MDD (Rutherford et al., 2010a). Given that MDD affects approximately 120 million people worldwide and that many patients will not experience sustained remission of their depression even with maximal treatment, methods of optimizing currently available treatments are urgently needed (Kessler et al., 2003; Rush et al., 2006). Optimizing patient expectancy may be a safe and effective means of improving treatment for MDD. Future research should investigate whether standardized patient education programs may be capable of increasing expectancy by proactively addressing patients' concerns about medication and informing them of the known effectiveness of the agent being prescribed.

Conversely, it would be helpful to minimize patient expectancy when the goal is to find signals of efficacy for new antidepressant treatments. It has been stated that addressing the problem of high placebo response is one of the most important challenges facing the future of psychopharmacologic drug development (Demitrack et al., 1998). The gradual increase of placebo response rates in clinical trials of antidepressants is at least partially responsible for the fact that only 14% of trials for 5 putative antidepressant agents submitted to Food and Drug Administration (FDA) in the 1990s showed active drug to be superior to placebo (Hooper and Amsterdam, 1998). To the extent these numbers indicate the presence of more false negative trials, rising placebo response may delay bringing new treatments to market, increase the costs of drug development, and influence the allocation of private industry resources away from drug development for psychiatric disorders (Fava et al., 2003).

The results of this study should be considered within the context of its limitations. Perhaps the most important limitation is that this analysis utilizes completer data contained in the published journal articles. Only 40% of the articles in our original sample contained this information, and we cannot rule out the possibility that differences obtain between articles publishing completer data and those that did not. Furthermore, the articles we analyzed most often did not provide information about the clinical and demographic characteristics of patients completing the studies, nor did they comment on whether there were significant differences between study drop-outs and completers. Therefore, while we did not discover significant clinical or demographic differences between the ITT patient populations of placebo-controlled and comparator studies in our sample, it is possible that differences between completers of the two study types other than patient expectancy contributed to the results observed.

Second, there were several differences found between the placebo-controlled and comparator trials in our sample that might have affected the results. Placebo-controlled relative to comparator trials enrolled individuals who were significantly younger, more frequently employed fixed dosing strategies, and more often utilized exclusively North American as compared to European sites. Third, publication bias may have affected which studies were included in these analyses, since RCTs failing to demonstrate significant differences between medication and placebo may not have been published. However, publication bias seems unlikely to explain the overall pattern of placebo-controlled trials having lower response rates than comparator trials. In fact, the inclusion of unpublished RCTs not demonstrating a difference between medication and placebo seems likely to strengthen the observed results, since lower antidepressant response rates in placebo-controlled RCTs would increase their relative differences with response rates in comparator trials.

In summary, patient expectancy appears to be an important mediator of treatment response in antidepressant RCTs. If participants in antidepressant clinical trials know they are receiving active medication for depression, this results in significantly higher response rates compared to when participants are aware that they may receive placebo. While increased drop-out rates are observed in placebo-controlled relative to comparator trials, this does not appear to explain the difference in response rates between the two study types. Methods of optimizing patient expectancy should be investigated as means of increasing antidepressant response.

Acknowledgments

Role of the funding source

This work was supported by a National Institute of Mental Health grants K23 MH085236 (BRR), K23 MH075006 (JRS), R21 MH087774 (JRS), T32 MH15144 (SPR), Hope for Depression Research Foundation (BRR), and a NARSAD Young Investigator Award (BRR).

References

- Alves C, Cachola I, Brandao J. Efficacy and tolerability of venlafaxine and fluoxetine in outpatients with major depression. *Primary Care Psychiatry*. 1999; 5:57–63.
- Baldwin DS, Hawley CJ, et al. A multicenter double-blind comparison of nefazodone and paroxetine in the treatment of outpatients with moderate-to-severe depression. *The Journal of Clinical Psychiatry*. 1996; 57 (Suppl 2):46–52. [PubMed: 8626363]
- Beck AT, Ward CH, Mendelson M, et al. An inventory of measuring depression. *Archives of General Psychiatry*. 1961; 4:53–63.
- Benkert O, Szegedi A, et al. Mirtazapine compared with paroxetine in major depression. *The Journal of Clinical Psychiatry*. 2000; 61:656–663. [PubMed: 11030486]
- Birge RT. The calculation of errors by the method of least squares. *Reviews of Modern Physics*. 1932; 40:207–227.
- Bouchard JM, Delaunay J, et al. Citalopram versus maprotiline: a controlled, clinical multicentre trial in depressed patients. *Acta Psychiatrica Scandinavica*. 1987; 76:583–592. [PubMed: 3324648]
- Bryk, AS.; Raudenbush, SW. Hierarchical linear models. Sage Publications; Newbury Park, CA: 1992.
- Byerley WF, Reimherr FW, et al. Fluoxetine, a selective serotonin uptake inhibitor, for the treatment of outpatients with major depression. *Journal of Clinical Psychopharmacology*. 1988; 8:112–115. [PubMed: 3286684]
- Chouinard G, Saxena B, et al. A Canadian multicenter, double-blind study of paroxetine and fluoxetine in major depressive disorder. *Journal of Affective Disorders*. 1999; 54:39–48. [PubMed: 10403145]
- Christiansen PE, Behnke K, et al. Paroxetine and amitriptyline in the treatment of depression in general practice. *Acta Psychiatrica Scandinavica*. 1996; 93:158–163. [PubMed: 8739658]
- Claghorn JL, Kiev A, et al. Paroxetine versus placebo: a double-blind comparison in depressed patients. *The Journal of Clinical Psychiatry*. 1992; 53:434–438. [PubMed: 1487471]
- Cohn CK, Robinson DS, et al. Responders to antidepressant drug treatment: a study comparing nefazodone, imipramine, and placebo in patients with major depression. *The Journal of Clinical Psychiatry*. 1996; 57 (Suppl 2):15–18. [PubMed: 8626358]
- Coleman CC, King BR, et al. A placebo-controlled comparison of the effects on sexual functioning of bupropion sustained release and fluoxetine. *Clinical Therapeutics*. 2001; 23:1040–1058. [PubMed: 11519769]
- Debus JR, Rush AJ, et al. Fluoxetine versus trazodone in the treatment of outpatients with major depression. *The Journal of Clinical Psychiatry*. 1988; 49:422–426. [PubMed: 3053668]
- Demitrack MA, Faries D, Herrera JM, DeBrotta D, Potter WZ. The problem of measurement error in multisite clinical trials. *Psychopharmacology Bulletin*. 1998; 34:19–24. [PubMed: 9564194]

- Detke MJ, Lu Y, et al. Duloxetine, 60 mg once daily, for major depressive disorder: a randomized double-blind placebo-controlled trial. *The Journal of Clinical Psychiatry*. 2002a; 63:308–315. [PubMed: 12000204]
- Detke MJ, Lu Y, et al. Duloxetine 60 mg once daily dosing versus placebo in the acute treatment of major depression. *Journal of Psychiatric Research*. 2002b; 36:383–390. [PubMed: 12393307]
- Dunlop SR, Dornseif BE, et al. Pattern analysis shows beneficial effect of fluoxetine treatment in mild depression. *Psychopharmacology Bulletin*. 1990; 26:173–180. [PubMed: 2236453]
- Fava M, Evins AE, Dorer DJ, Schoenfeld DA. The problem of the placebo response in clinical trials for psychiatric disorders: culprits, possible remedies, and a novel study design approach. *Psychotherapy and Psychosomatics*. 2003; 72:115–127. [PubMed: 12707478]
- Fournier J. A double-blind comparison of sertraline and imipramine in outpatients with major depression: acute 8 weeks and continuation 16 weeks treatment. *Human Psychopharmacology*. 1997; 12:203–215.
- Golden RN, Nemeroff CB, et al. Efficacy and tolerability of controlled-release and immediate-release paroxetine in the treatment of depression. *The Journal of Clinical Psychiatry*. 2002; 63:577–584. [PubMed: 12143913]
- Guy, W. ECDEU Assessment Manual for Psychopharmacology, revised (DHEW Publ No ADM 76–338). National Institute of Mental Health; Rockville, MD: 1976. Clinical global impressions; p. 218–222.
- Haddock CK, Rindskopf D, Shadish WR. Using odds ratios as effect sizes for meta-analysis of dichotomous data: a primer on methods and issues. *Psychological Methods*. 1998; 3:339–353.
- Hamilton M. A rating scale for depression. *Journal of Neurology, Neurosurgery, and Psychiatry*. 1960; 23:56–62.
- Higgins JPT, Thompson SG. Quantifying heterogeneity in a metaanalysis. *Statistics in Medicine*. 2002; 21:1539–1558. [PubMed: 12111919]
- Hooper, M.; Amsterdam, JD. Do clinical trials reflect drug potential? A review of FDA evaluation of new antidepressants. 39th Annual NCDEU Meeting; Boca Raton. June 11–14; 1998.
- Hox, J. *Multilevel Analysis: Techniques and Applications*. Lawrence Erlbaum Publishers; Mahwah, NJ: 2002.
- Juni, P.; Altman, DG.; Egger, M. Assessing the quality of randomized controlled trials. In: Egger, M., editor. *Systematic Reviews in Health Care: Meta-Analysis in Context*. BMJ Publishing Group; New York: 2001. p. 87–108.
- Kasper S, Olivieri L, et al. A comparative, randomised, double-blind study of trazodone prolonged-release and paroxetine in the treatment of patients with major depressive disorder. *Current Medical Research and Opinion*. 2005; 21:1139–1146. [PubMed: 16083521]
- Keegan D, Bowen RC, et al. A comparison of fluoxetine and amitriptyline in the treatment of major depression. *International Clinical Psychopharmacology*. 1991; 6:117–124. [PubMed: 1960381]
- Kessler RC, Berglund P, Demler O, Jin R, Koretz D, Merikangas KR, Rush AJ, Walters EE, Wang PS. The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). *The Journal of the American Medical Association*. 2003; 289:3095–3105.
- Khan A, Kolts RL, Thase ME, Krishnan KRR, Brown W. Research design features and patient characteristic associated with the outcome of antidepressant clinical trials. *The American Journal of Psychiatry*. 2004; 161:2045–2049. [PubMed: 15514405]
- Krell HV, Leuchter AF, Morgan M, Cook IA, Abrams M. Subject expectations of treatment effectiveness and outcome of treatment with an experimental antidepressant. *The Journal of Clinical Psychiatry*. 2004; 65:1174–1179. [PubMed: 15367043]
- Leinonen E, Skarstein J. Nordic Antidepressant Study Group. Efficacy and tolerability of mirtazapine versus citalopram: a double-blind, randomized study in patients with major depressive disorder. *International Clinical Psychopharmacology*. 1999; 14:329–337. [PubMed: 10565799]
- Lepola UM, Loft H, et al. Escitalopram 10–20 mg/day is effective and well tolerated in a placebo-controlled study in depression in primary care. *International Clinical Psychopharmacology*. 2003; 18:211–217. [PubMed: 12817155]

- Mehtonen OP, Sogaard J. Venlafaxine 631 Study Group. Randomized, double-blind comparison of venlafaxine and sertraline in outpatients with major depressive disorder. *The Journal of Clinical Psychiatry*. 2000; 61:95–100. [PubMed: 10732656]
- Mendels J, Johnston R, et al. Efficacy and safety of bid doses of venlafaxine in a dose–response study. *Psychopharmacology Bulletin*. 1993; 29:169–174. [PubMed: 8290661]
- Meyer B, Pilkonis PA, Krupnick JL, Egan MK, Simmens SJ, Sotsky SM. Treatment expectancies, patient alliance, and outcome: further analyses from the national institute of mental health treatment of depression collaborative research program. *Journal of Consulting and Clinical Psychology*. 2002; 70:1051–1055. [PubMed: 12182269]
- Moller HJ, Glaser K, et al. Double-blind, multicenter comparative study of sertraline versus amitriptyline in outpatients with major depression. *Pharmacopsychiatry*. 2000; 33:206–212. [PubMed: 11147927]
- Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. *The British Journal of Psychiatry*. 1979; 134:382–389. [PubMed: 444788]
- Montgomery SA, Huusom AK, et al. A randomised study comparing escitalopram with venlafaxine XR in primary care patients with major depressive disorder. *Neuropsychobiology*. 2004; 50:57–64. [PubMed: 15179022]
- National Institutes of Health. [Accessed on September 7, 2011] The Placebo Effect: Mechanisms and Methodology (R01). at <http://grants.nih.gov/grants/guide/rfa-files/RFA-DA-12-003.html>
- Papakostas GI, Fava M. Does the probability of receiving placebo influence clinical trial outcome? A meta-regression of double-blind, randomized clinical trials in MDD. *European Neuropsychopharmacology*. 2009; 19:34–40. [PubMed: 18823760]
- Patris M, Bouchard JM, et al. Citalopram versus fluoxetine: a double-blind, controlled, multicentre, phase III trial in patients with unipolar major depression treated in general practice. *International Clinical Psychopharmacology*. 1996; 11:129–136. [PubMed: 8803650]
- Perry PJ, Garvey MJ, et al. A comparative trial of fluoxetine versus trazodone in outpatients with major depression. *The Journal of Clinical Psychiatry*. 1989; 50:290–294. [PubMed: 2668259]
- Reimherr FW, Chouinard G, et al. Antidepressant efficacy of sertraline: a double-blind, placebo- and amitriptyline-controlled, multicenter comparison study in outpatients with major depression. *The Journal of Clinical Psychiatry*. 1990; 51 (Suppl B):18–27. [PubMed: 2258378]
- Rickels K, Schweizer E, et al. Nefazodone and imipramine in major depression: a placebo-controlled trial. *The British Journal of Psychiatry*. 1994; 164:802–805. [PubMed: 7952987]
- Rief W, Nestoriuc Y, Weiss S, Welzel E, Barsky AJ, Hofmann SG. Meta-analysis of the placebo response in antidepressant trials. *Journal of Affective Disorders*. 2009; 118:1–8. [PubMed: 19246102]
- Rudolph RL, Fabre LF, et al. A randomized, placebo-controlled, dose–response trial of venlafaxine hydrochloride in the treatment of major depression. *The Journal of Clinical Psychiatry*. 1998; 59:116–122. [PubMed: 9541154]
- Rush AJ, Trivedi MH, Wisniewski SR, Nierenberg AA, Stewart JW, Warden D, Niederehe G, Thase ME, Lavori PW, Lebowitz BD, McGrath PJ, Rosenbaum JF, Sackeim HA, Kupfer DJ, Luther J, Fava M. Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR*D report. *The American Journal of Psychiatry*. 2006; 163:1905–1917. [PubMed: 17074942]
- Rutherford BR, Sneed JR, Roose SP. Does study design affect outcome? The effects of placebo control and treatment duration in antidepressant trials. *Psychotherapy and Psychosomatics*. 2009; 78:172–181. [PubMed: 19321970]
- Rutherford BR, Sneed JR, Eisenstadt R, Roose SP. Antidepressant study design affects patient expectancy: a pilot study. *Psychological Medicine*. 2010a; 40:781–788. [PubMed: 19732481]
- Rutherford BR, Wager TD, Roose SP. Expectancy effects in the treatment of depression: a review of experimental methodology, effects on patient outcome, and neural mechanisms. *Current Reviews in Psychiatry*. 2010b; 6:1–10.
- Rutherford BR, Sneed JR, Tandler J, Peterson BS, Roose SP. Deconstructing pediatric depression trials: an analysis of the effects of expectancy and therapeutic contact. *Journal of the American Academy of Child and Adolescent Psychiatry*. 2011; 50:782–795. [PubMed: 21784298]

- Sauer H, Huppertz-Helmhold S, et al. Efficacy and safety of venlafaxine ER vs amitriptyline ER in patients with major depression of moderate severity. *Pharmacopsychiatry*. 2003; 36:169–175. [PubMed: 14571350]
- Schweizer E, Feighner J, et al. Comparison of venlafaxine and imipramine in the acute treatment of major depression in outpatients. *The Journal of Clinical Psychiatry*. 1994; 55:104–108. [PubMed: 8071246]
- Sinyor M, Levitt AJ, Cheung AH, Schaffer A, Kiss A, Dowlati Y, Lancot KL. Does inclusion of a placebo arm influence response to active antidepressant treatment in randomized controlled trials? Results from pooled and meta-analyses. *The Journal of Clinical Psychiatry*. 2010; 71:270–279. [PubMed: 20122371]
- Smith WT, Glaudin V, et al. Mirtazapine vs amitriptyline vs placebo in the treatment of major depressive disorder. *Psychopharmacology Bulletin*. 1990; 26:191–196. [PubMed: 2236455]
- Sneed JR, Rutherford BR, Rindskopf D, Roose SP. Design makes a difference: antidepressant response rates in placebo-controlled versus comparator trials in late life depression. *The American Journal of Geriatric Psychiatry*. 2008; 16:65–73. [PubMed: 17998306]
- Stewart-Williams S, Podd J. The placebo effect: dissolving the expectancy versus conditioning debate. *Psychological Bulletin*. 2004; 30:324–340. [PubMed: 14979775]
- Swann AC, Bowden CL, et al. Desipramine versus phenelzine in recurrent unipolar depression: clinical characteristics and treatment response. *Journal of Clinical Psychopharmacology*. 1997; 17:78–83. [PubMed: 10950467]
- Trivedi MH, Rush J. Does a placebo run-in or a placebo treatment cell affect the efficacy of antidepressant medications? *Neuropsychopharmacology*. 1994; 11:33. [PubMed: 7945742]
- Walczak DD, Apter JT, et al. The oral dose–effect relationship for fluvoxamine: a fixed-dose comparison against placebo in depressed outpatients. *Annals of Clinical Psychiatry*. 1996; 8:139–151. [PubMed: 8899132]
- Wernicke JF, Dunlop SR, et al. Fixed-dose fluoxetine therapy for depression. *Psychopharmacology Bulletin*. 1987; 23:164–168. [PubMed: 3496625]
- Wernicke JF, Dunlop SR, et al. Low-dose fluoxetine therapy for depression. *Psychopharmacology Bulletin*. 1988; 24:183–188. [PubMed: 3290940]
- Wernicke JF, Saylor ME, Koke SC, Pearson DK, Tollefson GD. Fluoxetine and concomitant centrally acting medication use during clinical trials of depression: the absence of an effect related to agitation and suicidal behavior. *Depression and Anxiety*. 2004; 6:31. [PubMed: 9394873]

Fig 1.
Flow chart for included studies.

Table 1

Placebo-controlled and comparator trials included in meta-analysis of antidepressant response rates.

Placebo-controlled studies					
Author	Medication	Duration	Outcome measure	Response rate	Completer sample
Byerley et al. (1988)	Fluoxetine	6	CGI	.43 ^a	.65
	Imipramine			.41 ^a	.54
	Placebo			.13	.10
Claghorn et al. (1992)	Paroxetine	6	CGI	.42 ^a	.62
	Placebo			.27	.46
Cohn et al. (1996)	Nefazodone	8	HRSD	.64 ^a	.79
	Imipramine			.64 ^a	.78
	Placebo			.36	.45
Coleman et al. (2001)	Bupropion	8	HRSD	.37	.56
	Fluoxetine			.36	.57
	Placebo			.34	.50
Detke et al. (2002a,b)	Duloxetine	9	HRSD	.45 ^a	.62
	Placebo			.23	.29
Detke et al. (2002a,b)	Duloxetine	9	HRSD	.50 ^a	.65
	Placebo			.35	.42
	Fluoxetine 20			.40	.53
Dunlop et al. (1990)	Fluoxetine 40	6	HRSD	.40	.51
	Fluoxetine 60			.35 ^a	.59
	Placebo			.26	.36
Golden et al. (2002)	Paroxetine CR	12	HRSD	.60 ^a	.74
	Paroxetine			.56	.73
	Placebo			.48	.60
Lepola et al. (2003)	Citalopram	8	MADRS	.53	.53
	Escitalopram			.64 ^a	.64
	Placebo			.48	.48
Mendels et al. (1993)	Venlafaxine low	6	CGI	.60	.80

Placebo-controlled studies					
Author	Medication	Duration	Outcome measure	Response rate	Completer sample
Reimherr et al. (1990)	Venlafaxine med			.65	.75
	Venlafaxine high			.68	.81
	Placebo			.50	.59
	Setraline	8	HRSD	.54 ^a	.82
	Amitriptyline			.60 ^a	.83
	Placebo			.35	.52
Rickels et al. (1994)	Nefazodone	8	HRSD	.52	.69
	Imipramine			.36	.66
	Placebo			.31	.45
Rudolph et al. (1998)	Venlafaxine 75	6	HRSD	.42	.58
	Venlafaxine 225			.50 ^a	.66
	Venlafaxine 375			.52 ^a	.67
	Placebo			.30	.38
Schweizer et al. (1994)	Venlafaxine	6	HRSD	.60 ^a	.77
	Imipramine			.37	.60
	Placebo			.35	.47
Smith et al. (1990)	Mirtazipine	6	HRSD	.40 ^a	.63
	Amitriptyline			.51 ^a	.69
	Placebo				.48
Walczak et al. (1996)	Fluvoxamine 25	8	HRSD	.29	.42
	Fluvoxamine 50			.30	.50
	Fluvoxamine 100			.36 ^a	.59
	Fluvoxamine 150			.28 ^a	.58
Wernicke et al. (1987)	Placebo			.23	.38
	Fluoxetine 20	6	HRSD	.39 ^a	.53
	Fluoxetine 40			.44 ^a	.61
	Fluoxetine 60			.30 ^a	.48
	Placebo			.19	.27

Placebo-controlled studies					
Author	Medication	Duration	Outcome measure	Response rate	Completer sample
Wernicke et al. (1988)	Fluoxetine 5	6	HRSD	.46 ^a	.54
	Fluoxetine 20			.50 ^a	.64
	Fluoxetine 40			.48 ^a	.65
	Placebo			.23	.33
Comparator studies					
Author	Medication	Duration	Outcome measure	Response rate	Completer sample
Alves et al. (1999)	Venlafaxine	12	HRSD	.85	.87
	Fluoxetine			.75	.79
Baldwin et al. (1996)	Nefazodone	8	CGI	.55	.72
	Paroxetine			.61	.72
Benkert et al. (2000)	Mirtazipine	6	HRSD	.58	.62
	Paroxetine			.54	.58
Bouchard et al. (1987)	Citalopram	6	MADRS	.78	1.0
	Maprotiline			.73	.94
Chouinard et al. (1999)	Paroxetine	12	HRSD	.67	.86
	Fluoxetine			.68	.88
Christiansen et al. (1996)	Paroxetine	8	CGI	.65	.82
	Amitriptyline			.66	.84
Debus et al. (1988)	Fluoxetine	6	HRSD	.50	.57
	Trazodone			.53	.64
Fournier (1997)	Sertraline	8	HRSD	.71	.76
	Imipramine			.74	.82
Kasper et al. (2005)	Trazodone	6	HRSD	.87	.96
	Paroxetine			.91	.91
Keegan et al. (1991)	Fluoxetine	6	HRSD	.63	.67
	Amitriptyline			.69	.84
Leinonen et al. (1999)	Mirtazipine	8	MADRS	.85	.97
	Citalopram			.88	.94
Mehtonen et al. (2000)	Venlafaxine	8	HRSD	.73 ^a	.83

Comparator studies					
Author	Medication	Duration	Outcome measure	Response rate	Completer sample
Moller et al. (2000)	Sertraline			.59	.68
	Sertraline	6	HRSD	.51	.51
	Amitriptyline			.68	.67
Montgomery et al. (2004)	Escitalopram	8	MADRS	.77	.88
	Venlafaxine XR			.80	.87
Patis et al. (1996)	Citalopram	8	MADRS	.78	.81
	Fluoxetine			.76	.75
Perry et al. (1989)	Fluoxetine	6	HRSD	.71	.71
	Trazodone			.82	.82
Sauer et al. (2003)	Venlafaxine	6	HRSD	.40	.42
	Amitriptyline			.47	.53
Swann et al. (1997)	Phenelzine	6	HRSD	.57	.55
	Desipramine			.57	.60

^aSignificant difference ($p < 0.05$) reported vs. placebo or comparator group.

Table 2

Selected characteristics of included studies.

Characteristic	Placebo controlled trials	Comparator trials
Number of studies	18	18
Number of active treatment groups	40	36
Number of enrolled patients	4620	3069
Mean age	40.4±2.9	42.8±3.8 ^a
Mean pre-treatment HRSD	23.6±3.5	24.9±2.7
Mean dropout rate	37.7±13.1	20.5±9.5 ^a
Mean N completer	71.6±35.6	67.1±43.7

	N active treatment groups	N enrolled patients	N active treatment groups	N enrolled patients
Study duration				
6	22	1796	18	950
7	0	0	0	0
8	14	2001	14	1445
9	2	249	0	0
12	2	417	4	285
Medications used				
SSRI	21	2789	18	1709
SNRI	9	787	4	333
TCA	6	422	7	550
Atypical	4	308	6	454
MAOI	0	0	1	23
Response definition				
HRSD 50%	32	3843	24	1769
MADRS 50%	2	314	8	961
CGI 1 or 2	6	463	4	339

^aSignificant difference (p<0.05) between groups.

Table 3

Multilevel meta-analysis of response rates in medication treatment cells based on ITT data.

Variable	Model 1		Model 2		Model 3	
	Coefficient (SE)	Odds ratio (CI)	Coefficient (SE)	Odds ratio (CI)	Coefficient (SE)	Odds ratio (CI)
Intercept	0.40 (0.10) *	1.49 (1.21–1.83) *	0.061 (0.12)	1.06 (0.83–1.36)	–0.14 (0.14)	0.87 (0.65–1.15)
Comparator	–	–	0.63 (0.17) *	1.87 (1.32–2.66) *	0.63 (0.16) *	1.88 (1.35–2.61) *
Eightwks	–	–	–	–	0.40 (0.17)	1.49 (1.06–2.10)
Twelvewks	–	–	–	–	0.50 (0.30)	1.65 (0.89–3.06)
Variance component	0.349		0.243		0.206	
–2LL	163.0		158.9		156.4	
df	36		35		33	

'Comparator' = dummy variable coded '1' for comparator trials and '0' for placebo-controlled trials. 'Eightwks' = dummy variable coded '1' for 8 week duration trials and '0' for all others. 'Twelvewks' = dummy variable coded '1' for 12 week duration trials and '0' for all others.

* p<0.05.

Table 4

Multilevel meta-analysis of response rates in medication treatment cells based on completer data.

Variable	Model 1			Model 2			Model 3		
	Coefficient (SE)	Odds ratio (CI)		Coefficient (SE)	Odds ratio (CI)		Coefficient (SE)	Odds ratio (CI)	
Intercept	0.94 (0.12) *	2.56 (2.00–3.29) *		0.62 (0.16) *	1.87 (1.35–2.57) *		0.41 (0.19) *	1.51 (1.03–2.20) *	
Comparator	–	–		0.63 (0.22) *	1.88 (1.19–2.97) *		0.62 (0.22) *	1.86 (1.20–2.89) *	
Eightwks	–	–		–	–		0.39 (0.23)	1.48 (0.93–2.34)	
Twelvewks	–	–		–	–		0.67 (0.41)	1.95 (0.85–4.48)	
Variance component	0.501			0.412			0.372		
–2LL	150.9			151.4			147.7		
df	36			35			33		

'Comparator' = dummy variable coded '1' for comparator trials and '0' for placebo-controlled trials. 'Eightwks' = dummy variable coded '1' for 8 week duration trials and '0' for all others. 'Twelvewks' = dummy variable coded '1' for 12 week duration trials and '0' for all others.

* p<0.05.