Treatments for Acute Bipolar Depression: Meta-analyses of Placebo-controlled, Monotherapy Trials of Anticonvulsants, Lithium and Antipsychotics

Authors
V. Selle1, 2, S. Schalkwijk1, 3, G. H. Vázquez1, 4, R. J. Baldessarini1, 5

Affiliations
1 International Consortium for Bipolar Disorder Research, McLean Hospital, Belmont, Massachusetts, USA
2 Department of Psychiatry, Viareggio Clinic, Lugano, Switzerland
3 Department of Pharmaceutics, Utrecht Institute for Pharmaceutical Sciences (UIPS), Utrecht University, Utrecht, Netherlands
4 Department of Neuroscience, Palermo University, Buenos Aires, Argentina
5 Department of Psychiatry, Harvard Medical School, Boston, Massachusetts, USA

Abstract

Background: Optimal treatments for bipolar depression, and the relative value of specific drugs for that purpose, remain uncertain, including agents other than antidepressants.

Methods: We searched for reports of placebo-controlled, monotherapy trials of mood-stabilizing anticonvulsants, second-generation antipsychotics, or lithium for acute major depressive episodes in patients diagnosed with type I or II bipolar disorder and applied random-effects meta-analysis to evaluate their efficacy, comparing outcomes based on standardized mean drug-placebo differences (SMD) in improvement, relative response rates (RR), and number-needed-to-treat (NNT).

Results: We identified 24 trials of 10 treatments (lasting 7.5 weeks, with ≥50 collaborating sites/trial) that met eligibility criteria: lamotrigine (5 trials), quetiapine (5), valproate (4), 2 each for aripiprazole, olanzapine, ziprasidone, and 1 each for carbamazepine, lithium, lurasidone, and olanzapine-fluoxetine. Overall, pooled drug-over-placebo responder-rate superiority (RR) was moderate (29% [CI: 19–40%]), and NNT was 8.2 (CI: 6.4–11). By SMD, apparent efficacy ranked: olanzapine + fluoxetine ≥ valproate > quetiapine > lurasidone > olanzapine, aripiprazole, and carbamazepine; ziprasidone was ineffective, and lithium remains inadequately studied. Notably, drugs were superior to placebo in only 11/24 trials (5/5 with quetiapine, 2/4 with valproate), and only lamotrigine, quetiapine and valproate had >2 trials. Treatment-associated mania-like reactions were uncommon (drugs: 3.7%; placebo: 4.7%).

Discussion: Controlled trials of non-antidepressant treatments for bipolar depression remain scarce, but findings with olanzapine-fluoxetine, lithium, quetiapine, and perhaps carbamazepine and valproate were encouraging; lithium requires adequate testing.

Introduction

Effective and safe treatment of depressive, dysthymic, and dysphoric or mixed components of bipolar disorders remains among the most challenging problems in modern clinical psychopharmacology [1, 2]. Overall, bipolar disorder patients in mid-course or from onset, treated by current community clinical standards, spend approximately half of the weeks of follow-up in symptomatic morbid states, and fully three-quarters of that morbidity is depressive [3]. Depressive components of bipolar disorder contribute importantly not only to long-term morbidity, but also to co-morbidity, disability, and excess mortality [1, 2]. Despite the pressing need for improved treatments for depressive episodes and the frequent failure of ongoing preventive treatments, remarkably little research has been directed to the problem [2]. Expert treatment recommendations continue to be tentative and inconsistent concerning depressive components of bipolar disorders, but typically ascribe high value to agents usually considered to be mood-stabilizers as well as to some modern antipsychotics [4–7]. Some of the available therapeutic research on these non-antidepressant treatments has been reviewed recently [8–11]. Nevertheless, important uncertainties remain about the relative efficacy and safety of antidepressants, anticonvulsants, lithium salts, second-generation antipsychotics, and several experimental treatments for bipolar depression [2, 8–14]. This uncertainty reflects the striking paucity of well and unambiguously designed, controlled trials specifically for bipolar depression. Given these uncertainties, we collected and analyzed available data concerning the relative efficacy of various anticonvulsants with putative mood-stabilizing properties and second-genera-
tion antipsychotics, as well as lithium salts, specifically for treatment of acute bipolar depression. We hypothesized that these treatments would vary in the amount of information available as well as in apparent efficacy based on data pooled across trials by meta-analytical methods using different outcome measures.

Methods

We performed a comprehensive literature search for reports on treatments for bipolar depression, focusing on randomized, controlled trials (RCTs) of mood-stabilizing anticonvulsants, second-generation antipsychotics, or lithium salts in acute major depressive episodes in patients diagnosed with type I or II BD. We carried out a systematic search [15] of several literature databases (PubMed, PsychInfo, EMBASE, and ClinicalTrials.gov). Search terms included various combinations of “anticonvulsants” and names of individual agents; carbamazepine, lamotrigine, levetiracetam, oxcarbazepine, valproate), “antipsychotics” (second-generation or atypical, and names of individual antipsychotics: amisulpiride, aripiprazole, asenapine, iloperidone, lurasidone, olanzapine, quetiapine, risperidone, ziprasidone); and “lithium” as well as “bipolar”, “controlled”, “depression”, “efficacy”, “randomized”, “treatment”, and “trial”. In addition, we hand-searched citations in identified reports and systematic reviews on this topic. Trial inclusion criteria were: [a] acute phase of major depressive episodes in type I or II BD diagnosed by standard, internationally accepted diagnostic criteria, [b] ≥ 18 patients/trial; [c] randomized treatment; [d] mood-stabilizing anticonvulsants, second-generation antipsychotics, or lithium salts as monotherapy; [e] placebo control (± other comparators); [f] double-blinded; [g] nominal trial duration ≥ 4 weeks. We excluded reports of studies concerning BD patients in episodes other than acute depression, trials involving add-on treatments, special populations (such as geriatric or pediatric patients, or those with known poor treatment responsiveness), and long-term studies of potential prophylactic effects. Abstracts of initially identified reports were screened for possible relevance, and evaluated for meeting our a priori study criteria requirements by independent review of full texts by 2 investigators (VS and SS). As a secondary consideration, several trials of interest that were excluded from the primary meta-analyses due to methodological shortcomings, were considered separately; they included early, small, brief trials of lithium carbonate with crossover designs involving partial placebo controls, or comparisons of bipolar vs. unipolar major depression. We extracted data from included full reports, including the sex distribution and average age of subjects, treatments and doses, subjects per trial arm and number of collaborating sites, mean baseline depression ratings in each trial arm, and approximate average weeks of treatment. Clinical ratings involved changes in scores on a standardized depression symptom rating scale (usually Montgomery-Åsberg depression rating scale [MADRS] or Hamilton depression rating scale [HDRS] with 17 or 21 items). Outcomes were either improvement (change in depression ratings or rates of achieving “response” (usually ≥ 50% reduction of initial depression symptom ratings). We also recorded reported rates of prematurely dropping out of trials in each arm as well as reported rates of mood switching from depression into mania-like (hypomania, mania, or mixed manic depressive) states.

Analyses included random-effects meta-analyses for individual trials and with pooling for overall assessments and for specific agents. Outcomes involved pooled drug/placebo response rate ratios (RR), standardized mean differences (SMD, as Hedges' g statistic), or response rate differences (RD) used to estimate number-needed-to-treat (NNT, as 1/RD), all with 95% confidence intervals (CI). In order to manage variance among outcome measures and its impact on rankings of apparent efficacy, we averaged 3 rankings based on RR, SMD, and NNT for each trial, and noted ranking for all 10 agents included for analysis. Correlations were tested with bivariate linear regression (r) or non-parametric Spearman rank correlation (r_s) methods. Potential covariates with SMD were assessed for at least suggestive associations (p ≤ 0.10) in preliminary bivariate meta-regression analyses in preparation for multivariate meta-regression analysis. Averages are reported as mean ± standard deviation (SD), sometimes weighted by subject number. Changes in depression ratings were standardized by subject counts, and variance is reported as SD calculated or imputed from pooled SD from all trials [16]. Statistical analyses used commercial software [Statview 3.0 (SAS Institute, Cary, NC), and Stata 10 (StataCorp, College Station, TX)].

Results

Trials identified

We identified a total of 4915 potentially relevant report titles at initial screening. Based on review of abstracts, 145 reports met eligibility criteria and were considered likely candidates for inclusion. Subsequent exclusions [121/145; 83.4%] were as follows: [a] 97 (66.9%) trials concerned BD patients in episodes other than acute depression, [b] 13 (8.97%) were long-term studies of potential prophylactic effects; [c] 7 (4.83%) involved add-on treatments, [d] 4 (2.76%) involved special populations. An additional 19 trials did not meet inclusion criteria owing to design limitations but included findings of interest and were considered for comment but not included in primary meta-analyses. In total, 24 trials met all inclusion/exclusion criteria and were included in the primary meta-analytical analyses (Table 1) [17–38]. 2 included studies [18, 27] had 3 arms comparing 2 different drugs to placebo, and 2 others [24, 29] reported on 2 independent trials of the same agents; each of these 4 separate drug-placebo comparisons was included as a separate trial. 5 trials lacked peer-reviewed publications (4 for lamotrigine, 1 for valproate), and their data were extracted from 2 pharmaceutical summary trial reports [32, 33] and 5 reviews [34–38].

Trial characteristics

Overall meta-analysis included a total of 7 307 unique subjects (4 543 randomized to an active agent and 2 764 to placebo, adjusting placebo-treated subjects by 484 used in 2 comparisons [18, 27]); 16/20 trials (80.0%) providing such information involved outpatients, and 4 (20.0%) involved both hospitalized and ambulatory patients. The numbers of collaborating sites/trial varied widely, from 1 to 110, and averaged 55. Proportions of bipolar disorder types varied, from all bipolar I in 86.5% of trials, to all bipolar II participants in 1 trial with lamotrigine (Table 1). 10 treatments were tested, including: aripiprazole (2 trials, at 5–30 mg/day); carbamazepine (1 trial, at an average of 452 mg/day); lamotrigine (5 trials, mean dose 220±48 mg/day); lithium carbonate (1 trial; at 600–1800 mg/day); lurasidone (5 trials; mean dose 220±48 mg/day); olanzapine (3 trials; mean dose 15 mg/day); quetiapine (3 trials; mean dose 300±160 mg/day); risperidone (3 trials; mean dose 1.5 mg/day); ziprasidone (2 trials, at 40–120 mg/day); valproate (2 trials, at 5–30 mg/day); divalproex (2 trials, at 5–30 mg/day); topiramate (1 trial; at 100–200 mg/day); lamotrigine (1 trial; at 50–100 mg/day); lurasidone (1 trial; at 400 mg/day); aripiprazole (1 trial; at 9 mg/day); olanzapine (1 trial; at 5 mg/day); quetiapine (1 trial; at 200 mg/day); risperidone (1 trial; at 1 mg/day); ziprasidone (1 trial; at 2 mg/day).
**Table 1** Placebo-controlled trials of anticonvulsant mood-stabilizers, second-generation antipsychotics, and lithium in acute bipolar depression.

<table>
<thead>
<tr>
<th>Treatment (Study)</th>
<th>Sites</th>
<th>Baseline Score$^a$</th>
<th>Improvement (% ± SD) Drug</th>
<th>Dropout (%) Drug</th>
<th>Switch (%) Drug</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aripiprazole</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thase et al. 2008</td>
<td>&gt;2</td>
<td>29.1</td>
<td>41.4 ± 45.2, 38.2 ± 48.8</td>
<td>44.0</td>
<td>32.4</td>
</tr>
<tr>
<td>Zhang et al. 2007</td>
<td>2</td>
<td>29.7</td>
<td>56.8 ± 43.8, 45.0 ± 41.1</td>
<td>26.5</td>
<td>40.0</td>
</tr>
<tr>
<td>Lamotrigine$^b$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Calabrese et al. 1999</td>
<td>20</td>
<td>28.8</td>
<td>45.7 ± 35.6, 27.5 ± 36.2</td>
<td>29.0</td>
<td>29.0</td>
</tr>
<tr>
<td>SCA2010. 2008, 2009</td>
<td>44</td>
<td>28.4</td>
<td>43.0 ± 35.5, 43.6 ± 35.5</td>
<td>34.0</td>
<td>33.0</td>
</tr>
<tr>
<td>SCA40924. 2008, 2009</td>
<td>2</td>
<td>29.7</td>
<td>45.6 ± 35.4, 40.0 ± 34.0</td>
<td>27.0</td>
<td>33.0</td>
</tr>
<tr>
<td>Tohen et al. 2008</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Calabrese et al. 2005</td>
<td>39</td>
<td>32.0</td>
<td>46.0 ± 40.2, 38.0 ± 48.2</td>
<td>51.6</td>
<td>61.5</td>
</tr>
<tr>
<td>Selle V et al. Treatments for Acute Bipolar ... Pharmacopsychiatry 2014; 47: 43–52</td>
<td>84</td>
<td>32.0</td>
<td>47.2 ± 33.4, 40.7 ± 33.4</td>
<td>22.2</td>
<td>28.6</td>
</tr>
<tr>
<td>Quetiapine</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sachs et al. 2001</td>
<td></td>
<td>31.3</td>
<td>56.8 ± 45.2, 40.6 ± 42.4</td>
<td>34.4</td>
<td>35.1</td>
</tr>
<tr>
<td>Davis et al. 2004</td>
<td>1</td>
<td>30.5</td>
<td>54.3 ± 43.1, 33.3 ± 40.8</td>
<td>39.3</td>
<td>40.9</td>
</tr>
<tr>
<td>Ghaemi et al. 2007</td>
<td>3</td>
<td>29.8</td>
<td>55.0 ± 4.1, 40.2 ± 42.4</td>
<td>44.0</td>
<td>34.5</td>
</tr>
<tr>
<td>Muzina et al. 2011</td>
<td>2</td>
<td>29.9</td>
<td>58.4 ± 45.5, 39.5 ± 45.1</td>
<td>37.9</td>
<td>31.4</td>
</tr>
<tr>
<td>Ziprasidone$^b$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lombardo et al. 2012</td>
<td>56</td>
<td>28.4</td>
<td>51.2 ± 46.0, 46.4 ± 44.0</td>
<td>38.8</td>
<td>31.9</td>
</tr>
<tr>
<td>Lombardo et al. 2012</td>
<td>45</td>
<td>28.6</td>
<td>51.3 ± 4.0, 46.0 ± 44.0</td>
<td>39.5</td>
<td>32.8</td>
</tr>
<tr>
<td>Totals; Weighted-Means ± SD</td>
<td>–55</td>
<td>29.4 ± 17.4</td>
<td>51.0 ± 23.0, 39.6 ± 25.7</td>
<td>35.7 ± 30.1, 39.9 ± 41.4</td>
<td>3.68 ± 2.90, 4.71 ± 5.12</td>
</tr>
</tbody>
</table>

Total subjects = 7 307 (adjusted for repeat use of 484 controls) in 24 drug-placebo comparisons. Averaged values (boldface) are means weighted by subject-numbers (summed [boldface] within drug categories). 3 trials involved extended-release preparations of quetiapine[26] or valproate[22,28] $^*$

$^a$Based on ratings with MADRS or HDRS depression scale that did not differ, and so were averaged between drug and placebo arms

$^b$Switch rates were confirmed by Dr. Muzina’s data manager, P.K. Chan, MS (personal written communication to Dr. Vázquez, 19 March 2013)
done (1 trial, at 20–60 or 80–120 mg/day); olanzapine (2 trials, dosed at 5–20 mg/day), olanzapine/fluoxetine combination (1 trial, dosed at 6/25, 6/50, or 12/50 mg/day); quetiapine (5 trials, at 300 or 600 mg/day), valproate, usually as divalproex (4 trials; mean dose 1225±183 mg/day); and ziprasidone (2 trials dosed at 40–80 or 120–160 mg/day). Baseline depression ratings based on MADRS or HDRS ranged from 27.0 to 32.0 and were very similar in drug and placebo arms, averaging 29.4±17.4 overall. Trial durations ranged from 6 to 10 weeks (weighted mean, 7.5±4.8 weeks; ○ Table 1).

Weighted average rates of trial non-completion (dropout) were similar in drug (35.7%) and placebo arms (39.9%) of the 24 trials, but reasons for dropping out were not provided consistently. With active agents, dropout rates were both highest (53.8%) and lowest (22.2%) in trials with valproate, although only lamotrigine, quetiapine and valproate had more than 2 trials to evaluate (○ Table 1).

Based on available data, risks of mood switching into hypomania, mania, or mixed states with active agents, as defined in each trial, were highest in one trial with valproate (30.8%), moderate with other agents (0.8–7.6%), and ranged from 0.0% to 21.4% in placebo arms, but were not reported for trials involving carbamazepine, lurasidone, or ziprasidone (○ Table 1). Unusually high switch rates appear to be related to particular trials rather than treatments, since switch rates associated with active agents and placebo were highly correlated within trials (r=0.772, p=0.0005). Overall, switch rates were slightly, but not significantly, lower with active drugs (3.68 [CI: 2.07–5.29]) than placebo (4.71% [1.89–7.55]); ○ Table 1).

Improvements in depression symptom ratings

Differences in reported percent improvement in depression symptom ratings were moderate (51.0±23.0% with drugs vs. 39.6±25.7% with placebo), and were superior with drug over placebo in all but 2 of 22 comparisons (including lamotrigine or valproate; ○ Table 1). The overall mean paired difference was small (9.43% [CI: 5.90–13.00], paired t=5.55, p<0.001), and the median was even smaller (7.30%; [IQR: 4.00–15.8%]). Average improvements with antipsychotic drugs were slightly but not significantly greater than with anticonvulsants or lithium (11.5% [7.03–16.00] vs. 6.91% [0.72–13.11]; t=1.38, p=0.18). Average rates of improvement with drugs minus-placebos ranked: olanzapine-fluoxetine (22.1% [1 trial])>quetiapine (16.9% [range: 14.1–21.0%])>carbamazepine (11.8% [1 trial])>olanzapine alone (7.25% [6.50–8.00%])>valproate (7.07% [−9.90 to 16.5%])≥lamotrigine (6.58% [−0.80 to 18.2%])>aripiprazole (3.25 [2.50–4.00%])≥lithium (3.20% [1 trial]).

Meta-analyses

Based on random-effects meta-analyses, there was a highly significant superiority of active agents over placebo, averaging 29% [CI: 19–40%, based on pooled RR value of 1.29 [CI: 1.19–1.40] (z=6.25, p<0.0001; ○ Table 2). The pooled RR value for the 4 most effective agents (olanzapine+fluoxetine, lurasidone, quetiapine, and valproate) was 1.47 [CI: 1.32–1.64]; (z=7.09 p<0.0001; not shown). However, there was considerable heterogeneity across outcomes, such that more than half (54.2% [13/24]) did not significantly differentiate a test agent from placebo. Rates of failed trials, in descending rank-order based on pooled SMD values (○ Table 3) were: aripiprazole (2/2 [100%])>ziprasidone (2/2 [100%])≥lithium (1/1 [100%])>quetiapine (4/5 [80%])>olanzapine alone (1/2 [50%])>valproate (2/4 [50%])>carbamazepine (0/1 [0%])>lurasidone (0.1 [0%])=olanzapine+fluoxetine (0/1 [0%])=quetiapine (0/5 [0%]).

Separate meta-analyses also were made for each treatment, comparing relative rates of responding to drug vs. placebo (RR), estimated number-needed-to-treat (NNT), and standardized mean drug-placebo difference in improvement of depression ratings (SMD), each measure with confidence intervals for each treatment. These comparisons (○ Table 3) indicate similar findings across treatment types with all 3 outcome measures. However, only 5/10 treatments yielded favorable NNT estimates of <10 (ranking: olanzapine-fluoxetine [1.8]<carbamazepine [3.4]<valproate [4.4]<lurasidone [4.6]<quetiapine [5.9]). The 3 outcomes are highly correlated, as expected (Spearman r, values are: RR vs. NNT: 0.946; SMD vs. RR: 0.897; SMD vs. NNT: 0.878; all p<0.008), but do not yield identical rankings by apparent efficacy. In order to deal with this variation among outcomes, we constructed rankings of apparent efficacy (drug-placebo differences) of the 10 treatments for each of the 3 outcome measures and then averaged them. These averaged rankings by apparent efficacy were as follow: olanzapine-fluoxetine (mean rank = 1.67)≠valproate (2.00)>carbamazepine (3.00)>lurasidone (4.00)>quetiapine (4.33)>olanzapine (6.67)≠lithium (7.00)≠lithium (7.33)>ziprasidone (9.00)>aripiprazole (10.00). Of note, 5 of the 10 treatments tested did not show statistical superiority of active drug vs. placebo based on pooled SMD (aripiprazole, carbamazepine, lamotrigine, lithium, and ziprasidone), nor did 3 of these show separation by RR (all but carbamazepine and lamotrigine). In addition, 5/10 treatments (all of the preceding agents as well as olanzapine) yielded relatively unfavorable values of NNT (≥10). Lithium, aripiprazole, and ziprasidone were not superior to placebo based on all 3 outcome measures (○ Table 3). Findings concerning efficacy measures for specific treatments are also illustrated in representative forest plots of SMD (a) and RR (b) values and their confidence intervals (○ Fig. 1, panels a and b). It is important to emphasize that both in ○ Table 3 and ○ Fig. 1, the 95% CIs for most treatments overlap, indicating lack of significant separation. These considerations and the small numbers of trials for most agents, indicate the need for caution in attempting to rank treatments by tested efficacy, based on the available data. Moreover, only lamotrigine, quetiapine, and valproate had more than 2 trials each, and 4 agents had only one (carbamazepine, lithium, lamotrigine, and olanzapine+fluoxetine; ○ Table 3).

Covariates of effect size

The following factors lacked even suggestive covariance with SMD (all p>0.10), as tested by bivariate meta-regression: [a] the proportion of women participants, [b] mean subject-age, [c] proportion of bipolar-I disorder diagnoses, [d] trial size (subject-count) and [e] number of collaborating sites, [f] year of reporting, [g] rating scale employed, [h] trial-duration, [i] assessments/month, [j] dropout rates, and [k] baseline severity score. In addition, [l] the rate of mood switching was suggestively, but again not significantly, associated with SMD (slope = 0.022 [CI: −0.002 to +0.074], t=1.98, p=0.07). Accordingly, none of these factors was tested further in multivariate meta-regression analysis. Of note, however, these comparisons are limited by the small number of trials, lack of multiple trials for several treatments, and limited variance of some measures. Given the small numbers of trials involving highly dissimilar agents, we also did not attempt to test for possible publication bias (as with Egger’s test or funnel plots).

Selle V et al. Treatments for Acute Bipolar... Pharmacopsychiatry 2014; 47: 43–52
Table 2  Meta-analysis of treatment effects in placebo-controlled trials for acute bipolar depression: mood-stabilizing anticonvulsants, antipsychotics, lithium.

<table>
<thead>
<tr>
<th>Trial</th>
<th>Year</th>
<th>Drug</th>
<th>Subjects</th>
<th>Responder/Subjects</th>
<th>RR [95%CI]</th>
<th>SMD [95%CI]</th>
<th>Weight (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calabrese et al. [17]</td>
<td>1999</td>
<td>LTG</td>
<td>129</td>
<td>34/63</td>
<td>1.87 [1.20–2.92]</td>
<td>0.487 [0.301–0.674]</td>
<td>5.83</td>
</tr>
<tr>
<td>Sachs et al. [37, 38]</td>
<td>2001</td>
<td>DVP</td>
<td>43</td>
<td>9/21</td>
<td>1.57 [0.68–3.65]</td>
<td>0.167 [0.432 to 0.766]</td>
<td>1.07</td>
</tr>
<tr>
<td>Tohen et al. [18]</td>
<td>2003</td>
<td>OFC</td>
<td>437</td>
<td>46/82</td>
<td>1.57 [1.19–2.06]</td>
<td>0.453 [0.211–0.695]</td>
<td>4.40</td>
</tr>
<tr>
<td>Davis et al. [20]</td>
<td>2004</td>
<td>DVP</td>
<td>25</td>
<td>6/13</td>
<td>1.85 [0.59–5.79]</td>
<td>0.338 [0.045 to 1.13]</td>
<td>0.64</td>
</tr>
<tr>
<td>Calabrese et al. [19]</td>
<td>2005</td>
<td>QTP</td>
<td>511</td>
<td>198/342</td>
<td>1.60 [0.83–1.34]</td>
<td>0.509 [0.158–0.860]</td>
<td>2.64</td>
</tr>
<tr>
<td>Tohen et al. [18]</td>
<td>2003</td>
<td>OFC</td>
<td>706</td>
<td>137/351</td>
<td>1.28 [1.05–1.57]</td>
<td>0.219 [0.071–0.367]</td>
<td>7.10</td>
</tr>
<tr>
<td>Davis et al. [20]</td>
<td>2004</td>
<td>DVP</td>
<td>25</td>
<td>6/13</td>
<td>1.85 [0.59–5.79]</td>
<td>0.338 [0.045 to 1.13]</td>
<td>0.64</td>
</tr>
<tr>
<td>Calabrese et al. [17]</td>
<td>1999</td>
<td>LTG</td>
<td>129</td>
<td>34/63</td>
<td>1.87 [1.20–2.92]</td>
<td>0.487 [0.301–0.674]</td>
<td>5.83</td>
</tr>
<tr>
<td>Sachs et al. [37, 38]</td>
<td>2001</td>
<td>DVP</td>
<td>43</td>
<td>9/21</td>
<td>1.57 [0.68–3.65]</td>
<td>0.167 [0.432 to 0.766]</td>
<td>1.07</td>
</tr>
<tr>
<td>Tohen et al. [18]</td>
<td>2003</td>
<td>OFC</td>
<td>437</td>
<td>46/82</td>
<td>1.57 [1.19–2.06]</td>
<td>0.453 [0.211–0.695]</td>
<td>4.40</td>
</tr>
<tr>
<td>Davis et al. [20]</td>
<td>2004</td>
<td>DVP</td>
<td>25</td>
<td>6/13</td>
<td>1.85 [0.59–5.79]</td>
<td>0.338 [0.045 to 1.13]</td>
<td>0.64</td>
</tr>
<tr>
<td>Calabrese et al. [19]</td>
<td>2005</td>
<td>QTP</td>
<td>511</td>
<td>198/342</td>
<td>1.60 [0.83–1.34]</td>
<td>0.509 [0.158–0.860]</td>
<td>2.64</td>
</tr>
<tr>
<td>Tohen et al. [18]</td>
<td>2003</td>
<td>OFC</td>
<td>706</td>
<td>137/351</td>
<td>1.28 [1.05–1.57]</td>
<td>0.219 [0.071–0.367]</td>
<td>7.10</td>
</tr>
<tr>
<td>Davis et al. [20]</td>
<td>2004</td>
<td>DVP</td>
<td>25</td>
<td>6/13</td>
<td>1.85 [0.59–5.79]</td>
<td>0.338 [0.045 to 1.13]</td>
<td>0.64</td>
</tr>
</tbody>
</table>

Drugs: APZ, aripiprazole; CBZ, carbamazepine; DVP, divalproex; DVPx, divalproex extended release; Li, lithium carbonate; LTG, lamotrigine; LUR, lurasidone; ONZ, olanzapine; OFC, olanzapine + fluoxetine combined; QTP, quetiapine; QTPx, quetiapine extended-release; ZPS, ziprasidone. Data ranked by reporting year

Non-significant drug-placebo difference in individual trial (13/24 [54.2%] of trials) by RR or SMD

Adjusted for 355 placebo-treated subjects used twice

Overall z-score by random-effects meta-analysis: for RR = 6.25, for SMD = 6.83; both p < 0.0001. Values for SMD vs. RR were strongly correlated (Spearman ρ = 0.838, z = 4.02, p < 0.0001)
### Table 3  Summary of meta-analyses: Effects of specific agents in placebo-controlled trials for acute bipolar depression.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Trials (n)</th>
<th>Subjects (N)</th>
<th>RR [95% CI] (p-value)</th>
<th>NNT [95% CI] (p-value)</th>
<th>SMD [95% CI]</th>
<th>SMD z-score</th>
<th>Favorable Outcomes (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valproate</td>
<td>4</td>
<td>140</td>
<td>2.08 [1.18–3.65] (0.01)</td>
<td>4.4 [2.7–12] (0.002)</td>
<td>0.452 [0.114–0.790]</td>
<td>2.62 (0.009)</td>
<td>2/4</td>
</tr>
<tr>
<td>Carbamazepine</td>
<td>1</td>
<td>70</td>
<td>1.84 [1.01–3.34] (0.05)</td>
<td>3.4 [1.9–19] (0.02)</td>
<td>0.209 [0.291–0.709]</td>
<td>0.82 (0.41)</td>
<td>0/1</td>
</tr>
<tr>
<td>Olanzapine + Fluoxetine</td>
<td>1</td>
<td>437</td>
<td>1.84 [1.44–2.36] (&lt;0.001)</td>
<td>1.8 [2.7–7.2] (&lt;0.001)</td>
<td>0.453 [0.211–0.695]</td>
<td>3.67 (&lt;0.001)</td>
<td>1/1</td>
</tr>
<tr>
<td>Lurasidone</td>
<td>1</td>
<td>485</td>
<td>1.72 [1.32–2.22] (&lt;0.001)</td>
<td>4.6 [3.3–7.8] (&lt;0.001)</td>
<td>0.318 [0.128–0.508]</td>
<td>3.29 (0.001)</td>
<td>1/1</td>
</tr>
<tr>
<td>Quetiapine</td>
<td>5</td>
<td>2485</td>
<td>1.36 [1.24–1.49] (&lt;0.001)</td>
<td>5.9 [4.7–7.8] (&lt;0.001)</td>
<td>0.373 [0.284–0.462]</td>
<td>8.19 (&lt;0.001)</td>
<td>5/5</td>
</tr>
<tr>
<td>Valproate</td>
<td>2</td>
<td>1220</td>
<td>1.25 [1.08–1.44] (0.002)</td>
<td>11 [7.0–30] (0.002)</td>
<td>0.187 [0.072–0.302]</td>
<td>3.18 (0.001)</td>
<td>1/2</td>
</tr>
<tr>
<td>Lamotrigine</td>
<td>1</td>
<td>1071</td>
<td>1.21 [1.07–1.46] (0.005)</td>
<td>10 [6.1–32] (0.004)</td>
<td>0.131 [0.018–0.280]</td>
<td>1.72 (0.03)</td>
<td>1/5</td>
</tr>
<tr>
<td>Lithium</td>
<td>1</td>
<td>265</td>
<td>1.12 [0.92–1.44] (0.27)</td>
<td>11 [7.4–20] (0.27)</td>
<td>0.142 [0.099–0.383]</td>
<td>1.15 (0.25)</td>
<td>0/1</td>
</tr>
<tr>
<td>Aripiprazole</td>
<td>2</td>
<td>690</td>
<td>0.88 [0.74–1.04] (0.49)</td>
<td>&gt;100 [58–200] (0.47)</td>
<td>0.077 [0.072–0.227]</td>
<td>1.07 (0.28)</td>
<td>0/2</td>
</tr>
<tr>
<td>Ziprasidone</td>
<td>1</td>
<td>928</td>
<td>1.02 [0.90–1.17] (0.73)</td>
<td>87 [14–200] (0.74)</td>
<td>0.103 [0.036–0.341]</td>
<td>1.47 (0.14)</td>
<td>0/2</td>
</tr>
<tr>
<td><strong>All Agents Pooled</strong></td>
<td>7</td>
<td>456</td>
<td>1.29 [1.19–1.40] (&lt;0.001)</td>
<td>8.2 [6.4–11] (&lt;0.0001)</td>
<td>0.232 [0.167–0.297]</td>
<td>6.98 (&lt;0.0001)</td>
<td>11/24</td>
</tr>
</tbody>
</table>

Agents are ranked by observed drug/placebo response-rate ratios (RR) based on random-effects meta-analysis of individual trials reported in Table 2. Favorable outcomes include trials with SMD > 0.0. Note that all confidence intervals (CI) for RR, NNT (number-needed-to-treat), and SMD (data adequate for only 6/10 treatments with ≥2 trials) overlap across agents, requiring caution in comparing drugs by efficacy.

*Non-significant separation of drug- and placebo-associated responses or NNT ≥10.

---

**Discussion**

Findings in this review of 24 randomized, placebo-controlled trials of non-antidepressant treatments for acute bipolar depression are consistent with other recent reviews of treatments and confirm the small benefits in small numbers of patients with treatment-resistant bipolar depression [57]. An open-label, uncontrolled trial of carbamazepine [58], an open-label, placebo-controlled trial of olanzapine [59], and a double-blind, placebo-controlled trial of divalproate and levetiracetam [60] found no difference between active medication treatment and placebo in patients with bipolar depression [58]. However, a double-blind, placebo-controlled trial of olanzapine and fluoxetine [59] found significant improvement in the bipolar cases (62 vs. 45%) [61].

Few trials have assessed the efficacy of lithium in the treatment of bipolar depression. In a study of 126 observations [46–50], lithium was found to be significantly superior to placebo (RR = 1.50; 95% CI: 1.15–1.95; p = 0.007). Random-effects meta-analytical summaries of the lithium trials suggest that lithium is more effective than placebo (RR = 1.63; 95% CI: 1.05–2.53; p = 0.03) [41]. These findings are inconclusive regarding the position of lithium in the treatment of bipolar depression.

In several trials involving 124 bipolar disorder subjects (68% type I), after 8 weeks, lamotrigine added to lithium treatment was somewhat more effective than placebo (RR = 1.50; 95% CI: 1.05–2.53; p = 0.03) [41]. These findings are inconclusive regarding the position of lamotrigine in the treatment of bipolar depression.

In earlier reviews, 6/9 uncontrolled trials of topiramate [51] found similarly small improvements (21%; RR = 1.25; 95% CI: 1.00–1.54; p = 0.07) [41]. These findings are inconclusive regarding the position of topiramate in the treatment of bipolar depression.

---

**Additional trials considered secondarily for comment**

A total of 13 trials in acute bipolar depression trials found lamotrigine to be non-significantly less effective than olanzapine and fluoxetine in improving mood stabilization and fewer trials were found to be marginally effective. One trial found no significant difference in response between lithium and placebo, which was inconclusive. This result was similar to the findings of the current meta-analysis.

---

This document was downloaded for personal use only. Unauthorized distribution is strictly prohibited.
few such trials could be identified, and very few treatments (only lamotrigine, quetiapine, and valproate) have been tested in more than one or 2 trials. We also found few additional, relevant trials, including of lithium, that could not be included in primary meta-analyses owing to methodological shortcomings [39–57]. Effects of antidepressants in bipolar depression also have been reviewed extensively [14, 58].

A noteworthy observation arising from this review is that rankings of specific treatments by apparent efficacy varied with the outcome measure employed (raw % improvement, RR, SMD, or NNT; Fig. 1). These outcome measures compare ratings of clinical changes with drugs vs. placebo somewhat differently: RR and NNT pertain to the proportion of persons attaining a criterion level of response, whereas SMD and raw percent improvement pertain to changes in symptomatic ratings of illness severity.

Overall, the evidence reviewed here indicates statistical superiority of active agents over placebo controls in only half of reported trials (Table 2, 3). Outcomes with aripiprazole, lamotrigine, and lithium failed to support efficacy by any reported outcome measure, all but RR for lamotrigine, and results from a single trial of carbamazepine were equivocal by RR and non-significant by SMD (Table 3). Outcomes among other, apparently effective, non-antidepressants varied widely, from relatively high pooled measures of efficacy with lurasidone, olanzapine + fluoxetine, quetiapine and valproate (SMD = 0.318–0.452; RR = 1.36–2.08; NNT < 6) to more moderate values with olanzapine alone (SMD = 0.187; RR = 1.25; NNT = 11; Table 3). However, even treatments with relatively favorable results were: [a] limited to single trials (lurasidone, olanzapine + fluoxetine), [b] not significantly effective by some outcome measures (olanzapine: high NNT; lamotrigine: low SMD), or [c] had a high proportion of trials with negative findings (valproate: 2/4; Table 2, 3). This body of evidence provides some encouraging leads, but does not establish consistent and unambiguous evidence of high levels of efficacy of potential treatments for acute bipolar depression. A possible exception is the atypical antipsychotic agent quetiapine, which has been studied in 5 placebo-controlled trials, with statistical superiority to placebo in all trials (Table 2, 3). However, even this promising treatment had modest effect sizes (e.g., drug vs. placebo average response rate difference with large placebo-associated responses, 16.2% [56.8–40.6%]; Table 1) and it may be risky for long-term use.
owing to its strong association with weight gain and metabolic syndrome [59].

The evidence reviewed was remarkably inconsistent or unfavorable and poorly studied for several treatments for which better effects might have been expected (Table 2, 3). Inconsistency is noteworthy for valproate and contrasts to its wide empirical application in various phases of bipolar disorder, including depression [37,38]. Similarly, lithium, too, is widely employed [2,60,61], despite having virtually no research support of efficacy in acute bipolar depression (Table 3), despite some encouraging findings in trials that did not meet study inclusion criteria. In addition, the performance of lamotrigine was uneven (Table 1, 3), contrasting to its regulatory approval for long-term treatment of bipolar depressive and manic recurrences [2]. Moreover, lamotrigine usually is administered in slowing increasing doses to limit risks of dermatological reactions, making it difficult to employ in acute phases of bipolar disorder [2].

Of other agents considered, carbamazepine and lurasidone, with only one trial each, appear to be promising and require more study, and further study of lithium would be of interest. By comparison with the present study, and further study of lithium would be of interest.

References

1 Baldessarini RJ, Vieta E, Calabrese JR et al. Bipolar depression: overview and commentary. Harv Rev Psychiatry 2010; 18: 143–157
10 Kemp DE, Canooy SJ, Brecher M et al. Clinical value of early partial symptomatic improvement in the prediction of response and remission during short-term treatment trials in 3369 subjects with bipolar I or II depression. J Affect Disord 2011; 130: 171–179
18 Tohen M, Viera E, Calabrese J. Efficacy of olanzapine and olanzapine-fluoxetine combination in the treatment of bipolar I depression. Arch Gen Psychiatry 2003; 60: 1079–1088
46 Goodwin FK, Murphy DL, Bunney WE. Lithium carbonate treatment in depression and mania. Arch Gen Psychiatry 1969; 21: 486–496