Efficacy of modern antipsychotics in placebo-controlled trials in bipolar depression: a meta-analysis

Nuria Cruz1,2, Jose Sanchez-Moreno3,4, Ferran Torres5, Jose Manuel Goikolea3, Marc Valentí3 and Eduard Vieta1,2
1 Bipolar Disorders Program, Hospital Clinic, University of Barcelona, IDIBAPS, CIBERSAM, Barcelona, Spain
2 International Consortium for Bipolar Disorders Research, McLean Hospital, Harvard Medical School, Boston, MA, USA
3 Bipolar Disorders Program, Hospital Clinic, University of Barcelona, IDIBAPS, Barcelona, Spain
4 Department of Psychiatry, Universidad Autonoma de Madrid, Spain
5 Laboratory of Biostatistics & Epidemiology (Universitat Autonoma de Barcelona), Clinical Pharmacology Service, IDIBAPS (Hospital Clinic), Barcelona, Spain

Abstract
Randomized, controlled trials have demonstrated efficacy for second-generation antipsychotics in the treatment of acute mania in bipolar disorder. Despite depression being considered the hallmark of bipolar disorder, there are no published systematic reviews or meta-analyses to evaluate the efficacy of modern atypical antipsychotics in bipolar depression. We systematically reviewed published or registered randomized, double-blind, placebo-controlled trials (RCTs) of modern antipsychotics in adult bipolar I and/or II depressive patients (DSM-IV criteria). Efficacy outcomes were assessed based on changes in the Montgomery–Asberg Depression Rating Scale (MADRS) during an 8-wk period. Data were combined through meta-analysis using risk ratio as an effect size with a 95% confidence interval (95% CI) and with a level of statistical significance of 5% ($p < 0.05$). We identified five RCTs; four involved antipsychotic monotherapy and one addressed both monotherapy and combination with an antidepressant. The two quetiapine trials analysed the safety and efficacy of two doses: 300 and 600 mg/d. The only olanzapine trial assessed olanzapine monotherapy within a range of 5–20 mg/d and olanzapine–fluoxetine combination within a range of 5–20 mg/d and 6–12 mg/d, respectively. The two aripiprazole placebo-controlled trials assessed doses of 5–30 mg/d. Quetiapine and olanzapine trials (3/5, 60%) demonstrated superiority over placebo ($p < 0.001$). Only 2/5 (40%) (both aripiprazole trials) failed in the primary efficacy measure after the first 6 wk. Some modern antipsychotics (quetiapine and olanzapine) have demonstrated efficacy in bipolar depressive patients from week 1 onwards. Rapid onset of action seems to be a common feature of atypical antipsychotics in bipolar depression.

Received 28 October 2008; Reviewed 16 December 2008; Revised 15 June 2009; Accepted 26 June 2009; First published online 29 July 2009

Key words: Antipsychotic drugs, bipolar disorder, depression, placebo-controlled trials.

Introduction
Randomized, placebo-controlled trials (RCTs) have demonstrated efficacy for atypical antipsychotics in the treatment of acute mania in bipolar disorder, either as monotherapy or adjunctive treatment (Bowden et al. 2005; Garcia-Amador et al. 2006; Hirschfeld et al. 2004; Keck et al. 2003a, b; Khanna et al. 2005; McIntyre et al. 2005; Potkin et al. 2005; Sachs et al. 2002, 2004, 2006; Smulevich et al. 2005; Tohen et al. 1999, 2000, 2002, 2008; Vieta et al. 2008b, in press; Weisler et al. 2003; Yatham et al. 2003, 2007). Currently available data suggest that combining atypical antipsychotics and mood stabilizers is the most efficacious treatment in acute mania, reinforced after a systematic review and meta-analysis of pooled data conducted by Perlis and...
Scherk (Perlis et al. 2006; Scherk et al. 2007). Atypical antipsychotics have not traditionally been considered as a major option in depression guidelines, unless psychotic features were present during the acute depressive episode (APA, 2002; Grunze et al. 2002, 2003; Yatham et al. 2005b, 2006). Moreover, atypical antipsychotics, are generally classified as a class, despite their marked differences in pharmacodynamic properties (D₂, 5-HT, H₁, α receptor affinities) (Brugue & Vieta, 2007). Based on data from the latest studies and RCTs conducted in bipolar depression, atypical antipsychotics do not seem to induce depressive episodes as anti-manic agents, while evidence suggests some atypical antipsychotics may have antidepressant (Yatham et al. 2005a) and stabilizing effects (Vieta et al. 2008a).

In recent years, a number of placebo-controlled RCTs have been conducted. Meta-analytical procedures allow us to answer questions about overall magnitude of effect and relative effect by week with greater statistical power than individual trials. Despite depression being considered the hallmark of bipolar disorder (Calabrese et al. 2001) and also a leading cause of disability and mortality (Mitchell & Malhi, 2004), meta-analyses addressing the efficacy and effectiveness of atypical antipsychotics in bipolar depression are lacking. Therefore, we conducted the first structured review and meta-analysis of randomized, placebo-controlled trials of atypical antipsychotics as monotherapy for the acute treatment of depression in bipolar I and/or II disorder.

Method

Search strategy

Studies were identified using searches of PubMed/ Medline with the search terms ‘depression’, ‘placebo’, and each of the atypical antipsychotics, limited to randomized, controlled clinical trials; review of abstracts from the 2003 meetings onward of the American College of Neuropsychiatry, American Psychiatric Association, and International Conference on Bipolar Disorder; and consultations with study investigators and representatives of pharmaceutical companies that market atypical antipsychotics. The search included the period 1994–2007.

Study characteristics

We selected for inclusion randomized, controlled atypical antipsychotics approved for any indication by the U.S. Food and Drug Administration since September 2004 (aripiprazole, asenapine, clozapine, paliperidone, quetiapine, risperidone, ziprasidone, olanzapine) and also amisulpride, licensed in other countries, with a placebo control group. To avoid publication bias, we checked the web (www. clinicaltrials.gov) and contacted all industry sponsors for finalized studies. We allowed both monotherapy studies and studies in which the drug was in combination with antidepressants, but only trials dealing with bipolar depression were included.

Data analysis

The primary outcome in all the trials was change from baseline in Montgomery–Asberg Depression Rating Scale (MADRS) score at 8 wk and this was also our criterion. The difference in change scores between each drug and its corresponding placebo arm was computed. That is, how much more improvement was observed in the drug arm compared to the placebo arm. Where standard deviations (s.d.) for change scores were not available, the median s.d. from those trials where s.d. was reported was used.

We also examined outcomes by response rates, defined as the proportion of subjects achieving ≥50% improvement, and remission rates defined as the proportion of subjects achieving MADRS ≤12 at an endpoint. These endpoint definitions were homogeneous in all trials.

The between-treatment comparisons were estimated by means of the odds ratio (OR) and 95% confidence intervals (95% CI) for the binary variables (remission and response) and the mean difference and 95% CI for the MADRS. Since it was considered likely a priori that not all trials would produce exactly equal underlying effect sizes, a random-effects model was considered preferable to a fixed-effects models. The random-effects model incorporates both within-study and between-study variance into the estimate of average treatment effects and is therefore usually more realistic that the fixed-effects model. We also performed a sensitivity analysis to assess the source of heterogeneity by excluding the aripiprazole studies, as they were negative on the primary outcome. In the only olanzapine trial which included an olanzapine–fluoxetine combination (OFC) treatment, we only considered the olanzapine monotherapy arm for the analysis in order to obtain homogeneous comparative results. The analysis was performed using SAS version 9.1.3 software (SAS Institute Inc., USA) and R software version 2.7.0 (R Development Core Team, Austria). The level of significance was established at the 0.05 level (two-sided).
**Results**

The Medline search identified 15 studies. Only three of these met the inclusion criteria. Two additional studies were identified from a review of meeting proceedings or consultation with study investigators. Quetiapine (Calabrese et al. 2005; Thase et al. 2006) and aripiprazole, reporting two studies in one publication (Thase et al. 2008) were each tested in two trials. One trial testing olanzapine included a combination therapy (OFC) in addition to the monotherapy and the placebo arm (Tohen et al. 2003). No placebo-controlled trials of amisulpride, asenapine, clozapine, paliperidone, risperidone or ziprasidone were identified.

### Included studies

A total of five studies were included. Study characteristics are given in Table 1. In general, monotherapy trials were of similar size and design, with the exception of the olanzapine trial, which was larger. Assessment of depressive symptoms was performed using the MADRS as a primary variable in all trials.

The baseline depression scores were similar in all the studies with moderate to severely depressed patients at inclusion (score range from 28.49 to 32.6). The duration of all studies was 8 wk. The analysis included not only the primary variable at endpoint, but also pooled data by week, in order to assess speed of action. The aripiprazole trials (Thase et al. 2008) excluded patients with psychotic features and bipolar II patients were only enrolled in the quetiapine studies (32–34%); no trials excluded patients with rapid cycling, which represented 18–40% of patients in the studies reporting such data (Calabrese et al. 2005; Thase et al. 2006; Tohen et al. 2003).

### Assessment of the MADRS scale

**Pooled data for each antipsychotic and overall magnitude of effect**

Random-effects estimates of each drug effect (pooled across all monotherapy and combination studies that included that drug) and associated 95% CIs are shown in Fig. 1h. All quetiapine and olanzapine trials demonstrated significant superiority over placebo at week 8 (i.e. all CIs in the pooled analysis excluded zero). The overall mean estimate was $-3.91$ (95% CI $-5.55$ to $-2.26$, $p<0.001$). Treatment effects exhibited a high degree of heterogeneity on the global assessment ($p=0.013$). However, when assessing the estimates within each drug, the heterogeneity was substantially reduced (quetiapine: $p=0.803$; olanzapine: $p=0.08$).

---

**Table 1. Atypical antipsychotics in the treatment of bipolar depression: placebo-controlled, randomized, monotherapy and combination therapy studies**

<table>
<thead>
<tr>
<th>Study</th>
<th>Duration (wk)</th>
<th>Comparators</th>
<th>Patients (n)a</th>
<th>Baseline MADRS score</th>
<th>MADRS change from baseline at 8 wk</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Monotherapy trials</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Calabrese et al. (2005)</td>
<td>8</td>
<td>Quetiapine (300 mg)</td>
<td>170</td>
<td>30.3</td>
<td>$-16.7$</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>Quetiapine (600 mg)</td>
<td>172</td>
<td>30.4</td>
<td>$-16.4$</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>Placebo</td>
<td>169</td>
<td>30.6</td>
<td>$-10.3$</td>
</tr>
<tr>
<td>Thase et al. (2006)</td>
<td>8</td>
<td>Quetiapine (300 mg)</td>
<td>155</td>
<td>29.9</td>
<td>$-16.0$</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>Quetiapine (600 mg)</td>
<td>161</td>
<td>31.1</td>
<td>$-16.9$</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>Placebo</td>
<td>151</td>
<td>29.6</td>
<td>$-11.9$</td>
</tr>
<tr>
<td>Thase et al. (2008)</td>
<td>8</td>
<td>Aripiprazoleb</td>
<td>177</td>
<td>29.1</td>
<td>$-12.0$</td>
</tr>
<tr>
<td>(CN138096)</td>
<td>8</td>
<td>Placebo</td>
<td>164</td>
<td>28.5</td>
<td>$-11.4$</td>
</tr>
<tr>
<td>Thase et al. (2008)</td>
<td>8</td>
<td>Aripiprazoleb</td>
<td>176</td>
<td>29.56</td>
<td>$-12.3$</td>
</tr>
<tr>
<td>(CN138146)</td>
<td>8</td>
<td>Placebo</td>
<td>178</td>
<td>29.35</td>
<td>$-11.8$</td>
</tr>
<tr>
<td><strong>Combination trials</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tohen et al. (2003)</td>
<td>8</td>
<td>Olanzapine</td>
<td>377</td>
<td>32.6</td>
<td>$-18.5$</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>OFC</td>
<td>370</td>
<td>30.8</td>
<td>$-11.9$</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>Placebo</td>
<td>86</td>
<td>31.3</td>
<td>$-5.19$</td>
</tr>
</tbody>
</table>

MADRS, Montgomery–Asberg Depression Rating Scale; OFC, Olanzapine–fluoxetine combination.

a Number of patients used in efficacy analyses.
b Range 5–30 mg.
olanzapine: $p$ non-estimable (one study); aripiprazole: $p = 0.813$. Moreover, when performing the sensitivity analysis excluding the aripiprazole group, the heterogeneity was then negligible ($p = 0.302$), and the pooled mean was $-4.90$ (95% CI $-6.21$ to $-3.59$, $p < 0.001$) (see Fig. 2).

Pooled data for each antipsychotic by week

All of the atypical antipsychotics demonstrated significant efficacy from week 1 and throughout the first 6 wk, which was the time with the maximal effect size reported by all of the studies. From week 6 to
endpoint quetiapine and olanzapine maintained their superiority over placebo, but aripiprazole did not (see Fig. 1).

**Assessment of response and remission**

**Response (overall magnitude for all antipsychotics)**

The proportion of patients achieving a clinical response, defined as a 50% reduction in MADRS score from baseline to endpoint, were reported for all trials by active vs. placebo effect. Random-effects analysis were pooled, estimated with an overall effect size (OR 0.66, 95% CI 0.49–0.89) (see Fig. 3a).

Treatment effects regarding response exhibited a high degree of heterogeneity on the global assessment ($p = 0.018$). However, when assessing the estimates within each drug, the heterogeneity was substantially reduced [quetiapine: $p = 0.266$; olanzapine: $p$ non-estimable (one study); aripiprazole: $p = 0.593$]. Further, when performing the sensitivity analysis excluding the aripiprazole group, the heterogeneity was smaller ($p = 0.5930$), and the pooled OR was 0.55 (95% CI 0.41–0.74, $p < 0.001$) (see Fig. 3b).

**Remission (overall magnitude for all antipsychotics)**

Proportion of patients achieving a clinical remission, defined as the proportion of subjects achieving MADRS $\leq 12$ at endpoint, were reported for all trials by active vs. placebo effect. Random-effects analysis were pooled, estimated with an overall effect size of 0.67 (95% CI 0.45–0.98) (see Fig. 4a).

Treatment effects regarding remission also showed a high degree of heterogeneity on global assessment ($p = 0.0010$). However, when assessing the estimates within each drug, the heterogeneity was substantially reduced [quetiapine: $p = 0.119$; olanzapine: $p$ non-estimable (one study); aripiprazole: $p = 0.415$].

Moreover, when performing the sensitivity analysis excluding the aripiprazole group, the heterogeneity was smaller ($p = 0.052$), and the pooled OR was 0.51 (95% CI 0.35–0.74, $p < 0.001$) (see Fig. 4b).

**Discussion**

To our knowledge, this is the first meta-analysis addressing the efficacy of atypical antipsychotics in the...
treatment of bipolar depression. Its results suggest that some second-generation antipsychotics (quetiapine and olanzapine) may additionally represent a monotherapy management option of bipolar I and/or II depression. Bipolar disorders present initially with a depressive episode in > 50% of patients, which is considered the major burden of bipolar disorder in terms of disability and suicide risk (Colom et al. 2006; Daban et al. 2006; Mitchell et al. 2008). Currently, atypical antipsychotics are only considered by treatment guidelines as second- or third-line therapy in the management of bipolar depression (Fountoulakis et al. 2007), despite the dearth of positive placebo-controlled trials with alternative compounds, such as lithium, lamotrigine or various combinations of antidepressant and mood-stabilizing agents, which are still mentioned as first-line treatments for bipolar depression (Yatham et al. 2005a, 2006). Of course, guidelines and clinicians not only look into efficacy, but also safety and tolerability as well when prioritizing treatment options, but it is likely that future updates of the major guidelines may shift upwards atypical antipsychotics in bipolar depression in their suggested algorithms.

We found some atypical antipsychotics (quetiapine and olanzapine) as monotherapy were significantly more efficacious than placebo (Calabrese et al. 2005; Thase et al. 2006; Tohen et al. 2003), as indicated by greater reductions in MADRS scores, in the treatment of acute bipolar depression from week 1 onwards,

---

**Fig. 3.** (a) Random-effects estimates of quetiapine, olanzapine, and aripiprazole and associated 95% confidence intervals for active vs. placebo effect in response rates. (b) Random-effects estimates excluding the aripiprazole group.

**Fig. 4.** (a) Random-effects estimates of quetiapine, olanzapine, and aripiprazole and associated 95% confidence intervals for active vs. placebo effect in remission rates. (b) Random-effects estimates excluding the aripiprazole group.
except for aripiprazole trials which have shown a 6-wk limited superiority compared to placebo, decreasing its effect size at endpoint (Thase et al. 2008).

The very early onset of action of all tested atypical antipsychotics in the treatment of bipolar depression, even as monotherapy, may highlight an overlapping mechanism of action of these drugs as a potential class effect with independence of the monoaminergic pathways.

The relevant question of whether initial combination therapy with a mood stabilizer is superior to monotherapy with an atypical antipsychotic thus cannot be answered with the present available data.

The superiority of quetiapine and olanzapine monotherapy by week is emphasized by higher response and remission rates and, except for aripiprazole, lower drop-out rates due to adverse events. Overall, completion rates were homogeneous with an average of ~60% for patients under quetiapine and OFC and slightly less for olanzapine monotherapy and aripiprazole treatment groups.

We elected to compare drugs based upon their difference from placebo, rather than absolute change in MADRS score, as a means of controlling for study differences.

Remission is often considered the most clinically useful endpoint, defined as an almost complete absence of mood symptoms, which in our meta-analysis represented a proportion of patients achieving a MADRS score of \( \leq 12 \). We obtained significant progressive superiority throughout the first 6 wk for all drugs, which was maintained to an endpoint only by quetiapine and olanzapine.

The homogeneity across the trials’ design and also in their study sample characteristics were the rule. Severity of depression at baseline did not show substantial differences and placebo response between studies was surprisingly highly homogeneous and may be not considered as a source of heterogeneity.

The heterogeneity in this meta-analysis was due to a differential effect in the aripiprazole group compared to the olanzapine and quetiapine groups, but not due to intra-treatment heterogeneity. Therefore, in our results, aripiprazole may be considered an important single source of heterogeneity, therefore there were no relevant differences in design among the bipolar depression studies or in dosage used for the same drugs.

It should be taken into account that we studied higher aripiprazole doses as a monotherapy compared to its use in combination with mood stabilizers. This fact suggests one potential reason for failed results in the aripiprazole trials (tolerability issues, especially akathisia perhaps related to dosing), leading to drop-outs and also related to differences in its mechanism of action such as too high \( D_2 \) affinity and low \( H_1 \) affinity compared to quetiapine and olanzapine.

The studies also differed in whether they included rapid-cycling patients or those in mixed states and also in the proportion of bipolar I or II patients with psychotic features. However, the exploratory analyses reported to date suggest little or no difference in overall efficacy across these subgroups. Overall, bipolar II and rapid-cycling patients highlight lower effect sizes in primary outcome, although with significantly superiority to placebo group in the treatment of bipolar depression. Moreover, the incidence of treatment-emergent mania was low (from 3.6% for higher quetiapine doses to a maximum of 5.7% for olanzapine trials) and not significantly different from placebo for any drug.

Regarding tolerability, quetiapine was generally safe and well tolerated in both tested doses. The most common adverse effects reported in ~30% of patients were not severe, mostly somnolence and sedation leading to withdrawal from the study, with most discontinuations occurring within the first week. Importantly, changes in weight observed in all three groups were relatively small and did not result in withdrawal from the study. In fact, weight gain associated to quetiapine administration was mild and dose-related, with \(< 9\%\) of patients gaining \( \geq 7\% \) from baseline as a clinical meaningfully measurement.

The olanzapine adverse-event profile was consistent with previously reported findings whereas the OFC profile was similar to that of olanzapine, except for higher rates of nausea and diarrhoea. Small but statistically significant mean increases in glucose and cholesterol levels were also seen. Patients under olanzapine treatment reported, as expected, about an 18% significant weight gain compared to the placebo arm. Nevertheless, only 9% of patients dropped out in olanzapine trials due to overall adverse events.

The high rates of drop-outs had been related to aripiprazole trials (study 1, 46.8% vs. 35.1% in placebo arm and study 2, 41.2% vs. 29.8%, respectively), and they were more associated to intolerable adverse events, especially akathisia. No meaningfully clinical changes on weight were reported.

Regarding limitations, there were very few studies available that currently met the inclusion criteria. Moreover, the studies included in that meta-analysis were not adequately powered to detect differences in subpopulations of bipolar depression to allow us to perform subanalysis for bipolar II subtype. Nevertheless, exploratory analysis performed in quetiapine
trials exerts quite qualitative homogeneous results for this subpopulation (Suppes et al. 2008). Further analyses of the pooled data are needed to examine the therapeutic effect with regard to other clinically relevant factors such as chronicity, sex, history of suicide attempts or substance abuse.

Moreover adjunctive studies with mood stabilizers are needed to compare the benefit-risk ratio and also to conduct more placebo-controlled studies of maintenance of antidepressant effect of atypical antipsychotics as monotherapy.

In summary our results suggest some atypical antipsychotics (quetiapine and olanzapine) may be considered as a first-line management option in acute bipolar I and/or II depression, even for poor responder subgroups such as rapid cyclers and patients with psychotic features. The question of their class effect seems to be answered positively regarding its early onset of action but with differences in the magnitude and maintenance of effect. These results raise questions on the current approach by most treatment guidelines and on the potential mechanism of the action involved in the improvement of depressive symptoms in bipolar depression by means of drugs traditionally considered as antipsychotics.

Acknowledgements

We thank the following pharmaceutical companies: AstraZeneca, Bristol–Myers Squibb and Eli-Lilly, for kindly providing us with all the requested data for the analysis. This project was funded by the CIBERSAM (Centro de Investigación Biomédica en Red de Salud Mental, CB07/07/0004) from the Spanish Ministry of Science and Innovation. N.C. is funded by the Spanish Ministry of Science and Innovation, Instituto Carlos III through a ‘Rio Hortega’ contract.

Statement of Interest

E. Vieta has received grant/research support from Almirall, AstraZeneca, Bristol–Myers Squibb, Eli-Lilly, the European 7th Framework Program, GlaxoSmithKline, Janssen-Cilag, Novartis, Organon, Otsuka, Pfizer, Sanofi-Aventis, Seny Foundation, Servier, the Spanish Ministry of Health (CIBERSAM), the Spanish Ministry of Science and Education, and the Stanley Medical Research Institute; has been a member of the speakers’ boards for Almirall, AstraZeneca, Bristol–Myers Squibb, Eli-Lilly, GlaxoSmithKline, Otsuka, and Sanofi-Aventis; and has served as a consultant for AstraZeneca, Bristol–Myers Squibb, Eli-Lilly, Forest Research Institute, GlaxoSmithKline, Janssen, Jazz, Lundbeck, Novartis, Organon, Otsuka, Pfizer, Sanofi-Aventis, Servier, and UBC. J. M. Goikolea has been a member of the speakers’ boards for Bristol–Myers Squibb, Eli-Lilly, GlaxoSmithKline, Otsuka, and Sanofi-Aventis; and has served as a consultant for AstraZeneca and Bristol–Myers Squibb.

References


trial. 4th International Forum on Mood and Anxiety Disorders, Monte Carlo, Monaco.


