Psychological Medicine, Page 1 of 19. © Cambridge University Press 2014
doi:10.1017/S0033291714001305

ORIGINAL ARTICLE

A network meta-analysis on comparative efficacy and all-cause discontinuation of antimanic treatments in acute bipolar mania

A. Yildiz[1,2]*, M. Nikodem[3], E. Vieta[2,4], C. U. Correll[5] and R. J. Baldessarini[2,6]

[1] *Department of Psychiatry, Dokuz Eylül University, Izmir, Turkey*
[2] *International Consortium for Bipolar Disorder Research & Psychopharmacology Program, McLean Division of Massachusetts General Hospital, Boston, MA, USA*
[3] *Faculty of Applied Mathematics, AGH University of Science and Technology, Krakow, Poland*
[4] *Bipolar Disorders Program, Institute of Clinical Neuroscience, Hospital Clinic, University of Barcelona, IDIBAPS, CIBERSAM, Barcelona, Catalonia, Spain*
[5] *Division of Psychiatry Research, Department of Psychiatry, Zucker Hillside Hospital, New York, NY, USA*
[6] *Department of Psychiatry, Harvard Medical School Boston, MA, USA*

Background. Evidence synthesis methods enabling direct and indirect comparisons over the entire set of relevant clinical data produce quantitative point estimates for the treatments contrasts between competing interventions, and provide a hierarchical rank ordering between them. We aimed to provide evidence-based guidance on the efficacy and all-cause discontinuation of antimanic treatments.

Method. We conducted a network meta-analysis within a Bayesian framework. We searched all standard literature databases without language restrictions up to 15 January 2014 to identify reports of short-term, randomized, blinded trials of putative antimanic drugs as monotherapy for adults with bipolar-I mania.

Results. Altogether, 14256 manic patients randomized to one of 18 active treatments or placebo provided 95 direct comparisons on 128 data points. For the primary outcome, standardized mean difference as Hedges' *g* (standardized mean difference; SMD), the hierarchies indicated by surface under the cumulative ranking (SUCRA) probabilities were in agreement with the point estimates for all antimanic drugs identified as effective. For the 12 effective antimanic drugs on clinical use, SMDs against placebo ranged from 0.32 to 0.66 without superiority of one over another, except for risperidone *v.* aripiprazole and valproate. Aripiprazole, olanzapine, quetiapine, risperidone, and valproate had less all-cause discontinuation rates than placebo. Sensitivity analysis by drug class indicated similar efficacy profiles for haloperidol, second-generation antipsychotics, and mood stabilizers.

Conclusions. Hierarchical rank ordering by comparative efficacy and risk of all-cause discontinuations should help to guide antimanic treatment choices by clinicians, healthcare policy makers, and guideline developers.

Received 3 April 2014; Revised 11 May 2014; Accepted 12 May 2014

Key words: All-cause discontinuations, antimanic, bipolar, mania, multiple treatments meta-analysis.

Introduction

Bipolar disorder leads to extreme and erratic shifts of mood, thinking, and behavior and has a lifetime prevalence of 1–4.5% (Merikangas *et al*. 2007). This prevalent, complex, and clinically challenging disorder is associated with increased risks of dying not only by suicide, which is roughly 20 times more common in this population, but also from prevalent medical disorders with a consequent shortening in life expectancy of 10–13.6 years relative to the general population of same age (Ösby *et al*. 2001; Tondo *et al*. 2007; Chang *et al*. 2011; Laursen, 2011; Leboyer *et al*. 2012; Cipriani *et al*. 2013). Excited or manic states are characterized by extremes of energy and activity, as well as risk-taking, impulsive actions, substance abuse, and rapid or irrational thinking, sometimes with psychotic delusions (Goodwin & Jamison, 2007). Poor judgment and risk-taking often lead to devastating economic, occupational, and interpersonal problems. At the stage of treatment engagement, the manic syndrome is generally already escalated and the expectation is prompt alleviation, which can only be managed by selecting an antimanic treatment with high potency. Since bipolar disorder is a recurrent and life-long illness, the

* Address for correspondence: Dr A. Yildiz, Depression & Bipolar Disorders Foundation/Duygudurum Vakfi-DUVAK, Yenikale Mah. Sutculer Sok. No.7/3 Narlidere, Izmir, Turkey.
(Email: agul_yildiz@hotmail.com)

treatment of acute manic or mixed episodes also provides a 'portal of entry' to establish a foundation of treatment aimed at subsequent maintenance and relapse prevention. Accordingly, besides efficacy, acceptability of selected treatment is critical, not only for prompt alleviation of manic syndrome but also for the long-term prophylaxis against emerging episodes, disability, and premature death by associated medical causes as well as suicide (Ahrens *et al*. 1995; Leboyer *et al*. 2012; Cipriani *et al*. 2013). Available antimanic treatments include lithium, valproate and carbamazepine, which have been considered 'mood stabilizers' (MSs); and several 'atypical' or 'second-generation' antipsychotic drugs (SGAs) (Yildiz *et al*. 2011*a*; Baldessarini, 2013).

Evidence-based comparisons among these and other treatments are not readily accessible, since most of the accumulated randomized controlled trial (RCT) evidence is against placebo (Yildiz *et al*. 2011*a*). In situations where an active comparator is used, the sample size is often not calculated for a test of superiority or non-inferiority between the two active treatments, but rather for superiority of an experimental treatment to placebo, with the active comparator serving as 'assay sensitivity' for the trial's integrity (Vieta & Cruz, 2012). Standard pair-wise meta-analyses (SPM) can overcome the problem associated with power for such active treatment comparisons. However, since SPM cannot integrate all evidence from several comparators and many available drugs have not been compared head-to-head, a hierarchy on comparative efficacy cannot be obtained. A meta-analytic technique known as network analysis or multiple-treatments meta-analysis (MTM), enables indirect comparisons among competing interventions by combining findings from separate trials involving a common comparator (Salanti *et al*. 2008). Though relying on strong assumptions, MTM allows integration of all available evidence, maximizes power, increases precision, and enables evidence-based hierarchies of comparative efficacy, acceptability and safety of competing interventions (Caldwell *et al*. 2005; Glenny *et al*. 2005; Salanti *et al*. 2009).

In an attempt to update and recapture hierarchies on efficacy and acceptability of antimanic treatments, we constructed a comprehensive network involving all candidate antimanic treatments, and performed MTM on continuous and categorical measures of efficacy as well as risk of all-cause discontinuations.

Methods

Search strategy and selection criteria

We searched the PubMed/Medline; EMBASE; PsycINFO; ClinicalTrials.gov; Cochrane Central Register of Controlled Trials; and Controlled-trials.com databases to identify RCTs of orally administered, putative antimanic drugs for adults with acute bipolar-I mania available up to 15 January 2014. Search-terms were: 'bipolar AND mania', and names of individual antipsychotics, anticonvulsants, lithium, or other drugs. Search results were supplemented by reports identified through previous reviews, conference proceedings, and by data requested from pharmaceutical companies and authors. A flow diagram is provided in Appendix 1 (the Appendices are available as supplementary online material).

We included published and unpublished RCTs that compared a test-agent with placebo or an active comparator as monotherapy, with blinded outcome assessment. The requirement for blindness led to exclusion of three randomized but open, head-to-head trials of lithium *v*. valproate (Bowden *et al*. 2008, 2010; Mosolov *et al*. 2009). We further excluded trials involving participants with hypomania or mild mania, or with type II or unspecified (NOS) bipolar, or schizoaffective disorders (McElroy *et al*. 2010*a*, *b*). We also excluded trials involving add-on or combination treatments and those permitting adjunctive use of psychotropic agents other than hypnotic sedatives for sleep or antiparkinsonism drugs for extrapyramidal side effects. Through that requirement, besides numerous trials with a combination or add-on design, four apparently single-agent trials permitting use haloperidol or chlorpromazine, as needed were excluded (Brown *et al*. 1989; Garza-Treviño *et al*. 1992; Ortega Soto, 1993; Shafti & Shahveisi, 2008). Since similarity of trial circumstances and patient characteristics is fundamental for MTM, we only included trials which evaluated manic symptom improvement via a scale specifically developed for mania, providing quantifiable data. This requirement led to exclusion of one head-to-head trial of lithium *v*. carbamazepine (Lerer *et al*. 1987); and another on ziprasidone *v*. olanzapine, terminated early without efficacy analysis (Pfizer, 2008). We further excluded trials comparing different dosing schedules or immediate- *v*. slow-release formulations of the same drug, or trials without a control group.

As we aimed to evaluate antimanic drugs' efficacy for acute treatment, we targeted outcome assessments at 3 or 4 weeks. If 3 or 4 weeks' data were not available, we used data at the points closest to 3 weeks over 1–8 weeks.

We considered 18 orally administered antimanic drugs used as monotherapies, including all flexible-dose studies since in these studies dose of test drug(s) was titrated up to adequate dose according to individual patients' needs. For fixed-dose studies, we included target doses up to maximum doses as indicated in the international consensus study of antipsychotic dosing (Gardner *et al*. 2010). Bearing in

mind FDA recommendations on placebo equivalence or non-equivalence we did not set a minimum dose for any drug tested against placebo. Comparability of doses in studies involving two active agents was evaluated on the basis of international agreements on antipsychotic and mood stabilizer dosing (Gardner *et al*. 2010; Baldessarini, 2013). In situations where an unfairly low dosing strategy may potentially affect ranking by efficacy among proven antimanic treatments these comparisons were to be excluded from the MTM. However, no study was excluded for this reason since a few identified low dosing schedules involved both of the compared drugs or a drug with previously proven inefficacy (Goldsmith *et al*. 2003; Yildiz *et al*. 2011*a*).

Outcome measures and data extraction

The primary outcome was improvement in manic symptoms, which was assessed by the Young Mania Rating Scale (YMRS; 11 items, scoring range: 0–60) or the Mania Rating Scale developed from the Schedule for Affective Disorders and Schizophrenia, Change Version (MRS; 11 items, range: 0–53), as change in total score from baseline to endpoint (Endicott & Spitzer, 1978; Young *et al*. 1978). If data from these scales were not available, we used change in total score by another scale developed for mania, and then ratings with these scales at study endpoint. In the present network, in only one study the Beigel–Murphy Manic State Rating Scale (MSRS; Beigel & Murphy, 1971) was used; data was presented as endpoint scores (Shafti, 2010). Except for this and one additional study providing endpoint scores with the MRS (Freeman *et al*. 1992), all included RCTs provided data on the primary outcome as the mean change in YMRS or MRS total scores.

Intention-to-treat (ITT) datasets were used whenever available. Secondary outcomes were the rate of attaining *response* (usually defined as $\geqslant$50% reduction of initial mania scores), and all-cause discontinuation rates. For all but two comparisons, standard deviation (S.D.) for change in scores was available. In these two occasions, missing S.D.s were estimated through the mean S.D. of the other included studies.

Meta-analytic calculations

MTM or network analysis integrates evidence by using both direct comparisons available in the conducted RCTs, and also indirect comparisons hypothetically made between competing interventions by using a third, common, comparator (Salanti *et al*. 2008; Kriston & Yildiz, in press). Since the main principle enabling indirect comparisons is *transitivity*, or the assumption that trials are virtually *exchangeable* in subject samples, design, conduct, and outcome measures, we first aimed for utmost similarity of the trials to be included in the network (Lu & Ades, 2004; Glenny *et al*. 2005).

For MTM, we applied random-effects modeling for multi-arm studies with a Bayesian approach, using WinBUGS 1.4.3 software (Biostatistics Unit, Cambridge University, UK) based on Markov-chain Monte Carlo methods (Lu & Ades, 2004; Caldwell *et al*. 2005; Ades *et al*. 2006). WinBUGS codes are provided in Appendix 2. Bayesian hierarchical modeling enables utilization of all available information provided in the entire data network for computing each parameter without any duplication (Ades *et al*. 2006). Thus, for evidence structures comparing multiple treatments, including three-arm studies, a Bayesian approach is particularly suitable for estimating effect-size parameters. Our primary outcome was the standardized mean difference (SMD) as Hedges' adjusted-*g* between competing interventions. For the computations of SMD we used a modified version of the mixed treatment comparisons model originally defined for dichotomous data by Ades *et al*. (2014; https://www.bris.ac.uk/cobm/research/mpes/mtc.html). We established the prior distribution for random-effects S.D., using a formula to convert log-odds ratio (OR), for response with drug *v*. placebo or active comparator, to SMD (Chinn, 2000). For categorical measures of response, we used OR instead of relative risk (RR), which in the case of MTM could lead to reversal of treatment effects (Eckermann *et al*. 2009). For this aim, we used a mixed treatment comparisons model for dichotomous data by Ades *et al*. (2014). Both types of effect sizes are reported with their 95% credible intervals (CrI). To rank the treatments we used the surface under the cumulative ranking (SUCRA) probabilities (Salanti *et al*. 2011). Ranking probabilities provide simple and appealing statements, such as 'treatment A has 30% chance to be the best' (Salanti *et al*. 2011). However, since SUCRA probabilities can be sensitive to small changes in the posterior distribution of effect sizes (either the mean or the variance), and high probabilities in low ranks do not necessarily indicate the most effective treatments (Salanti *et al*. 2011), we provided effect sizes of pairwise comparisons, along with measures on probabilities.

We also calculated absolute differences in responder rates (RD), with associated number-needed-to-treat (NNT), the estimated number of patients to be treated for one additional patient to benefit from a particular drug *v*. placebo or an active comparator (Altman, 1998). For the computations of RD and NNT within the context of MTM we employed an approach described by Woo *et al*. (2010), and selected the most extensively tested antimanic agent: lithium, as a

standard for an auxiliary estimate of average responder rate. Of note, choice of reference treatment, unless it is an outlier, has no influence on the computed values of RD and NNT. To avoid misinterpretation, NNT values are reported only if significant with the entire 95% CrI being above or below zero.

The underlying assumption of transitivity suggests that there is no imbalance in distribution of potential effect modifiers across the pairwise comparisons in the network (Yildiz *et al*. 2014*b*). If the transitivity assumption held, there would not be major disagreements or inconsistency between direct and indirect evidence. We assessed network consistency by testing agreement between direct and indirect evidence for each pair of drugs with both direct and indirect evidence (Dias *et al*. 2010). If direct evidence as denoted by μ exist for comparisons between drug A *v*. drug B, drug B *v*. drug C, and drug C *v*. drug A, a consistency parameter $\mathrm{Cons}_{\mathrm{ABC}}$ of the loop A,B,C is defined as follows:

$$\mathrm{Cons}_{\mathrm{ABC}} = \mu_{\mathrm{AB}} + \mu_{\mathrm{BC}} + \mu_{\mathrm{CA}}$$
$$\mathrm{Var}(\mathrm{Cons}_{\mathrm{ABC}}) = \mathrm{Var}(\mu_{\mathrm{AB}}) + \mathrm{Var}(\mu_{\mathrm{BC}}) + \mathrm{Var}(\mu_{\mathrm{CA}}).$$

If the computed 95% CrI of $\mathrm{Cons}_{\mathrm{ABC}}$ does not include zero, an inconsistency on the loop A,B,C is significant. This approach can be extended to loops consisting of more than three drugs. However, in the dense network structures as considered here, with employment of longer loops, it is unlikely to detect any additional signs of inconsistency. As such, we examined three-treatment loops. For further assessment of network consistency, we carried out an auxiliary, pairwise meta-analysis using the same random-effects Bayesian approach with their 95% CrI for estimates as in the overall MTM, but without transitivity of effect allowed (i.e. no influence of indirect evidence on results of direct comparisons), to permit a sensitive comparison with the MTM results (Spiegelhalter *et al*. 2002; Lu & Ades, 2004; Ades *et al*. 2006). For further assessment of network consistency, parameters of Deviance Information Criterion (DIC) and residual deviance (D_{res}) were computed in order to measure model fit (Dias *et al*. 2010; Jonas *et al*. 2013).

As employed in the MTM of antipsychotics in patients with schizophrenia (Leucht *et al*. 2013), we performed several *post-hoc* sensitivity analyses on the primary outcome. These were conducted by (1) excluding studies with unusually high or low treatment effects; (2) excluding haloperidol to rule out differences in its dosing as a potential source of bias; (3) excluding placebo to examine if hypothetical increases in placebo-associated responses over trial years might have reduced the apparent efficacy of newer drugs; and (4) excluding candidate drugs which are not considered as first-generation antipsychotic (FGA) or SGA or MS to test validity of class effect via a group-based comparison. Since all included studies but one (Kakkar *et al*. 2009) employed an ITT-based last observation carried forward (LOCF) approach, we did not require sensitivity analysis by exclusion of studies not analyzed on an ITT basis. Likewise, since all studies but one (Shafti, 2010), used a similar rating methods for the primary outcome assessment, we did not require sensitivity analysis for the rating scale used. Unfair dose comparisons, dissimilar diagnostic criteria or trial durations, incomparable methods for concealing treatment allocations or to provide blinding were additional potential effect modifiers considered. However, the network was quite homogeneous, with a trial duration of 3 or 4 weeks in 94%, and employment of recent DSM diagnostic criteria in 98% of trials; unreported allocation concealment or unfair dose comparisons being identified in only 9% and 4% of the included RCTs, respectively (Spitzer *et al*. 1978; APA, 2000; Li *et al*. 2008).

We planned several multiple-treatment meta-regressions (MTR) *a priori* on the primary outcome for: (1) presence of psychotic features, (2) presence of mixed-state diagnosis, (3) all-cause discontinuation rates, (4) publication year, and (5) funding source. For the meta-regression analysis we employed the random-effects model as described in Jonas *et al*. (2013), using WinBUGS 1.4.3 software (Appendix 2). The continuous covariates were centered at their mean values to improve model convergence. Finally, we explored presence of publication bias or small study effects with a funnel plot technique expanded to MTM (Moreno *et al*. 2009; Chaimani & Salanti, 2012). Statistical significance required two-tailed $p<0.05$.

Results

This updated network by differential exclusion of all add-on or combination trials (N=15 RCTs, 32 study arms, 13 treatments), and also four apparently single-agent trials (N=4, eight study arms, Lerer *et al*. 1987; Brown *et al*. 1989; Ortega Soto, 1993; Pfizer, 2008); along with further consideration of three recent cariprazine (Knesevich *et al*. 2009; Bose *et al*. 2012; Calabrese *et al*. 2013), one licarbazepine (Novartis, 2007), two tamoxifen (Zarate *et al*. 2007; Yildiz *et al*. 2008), and one verapamil (Janicak *et al*. 1998) trials (N=7 RCTs, 13 study arms, four new treatments), is different than the previous one in 53 data points deriving from 29 comparisons (Cipriani *et al*. 2011).

The present total of 57 studies, reported up to 15 January 2014, involved 95 direct comparisons with 14256 randomized participants (Fig. 1). Study and patient characteristics of the included RCTs are

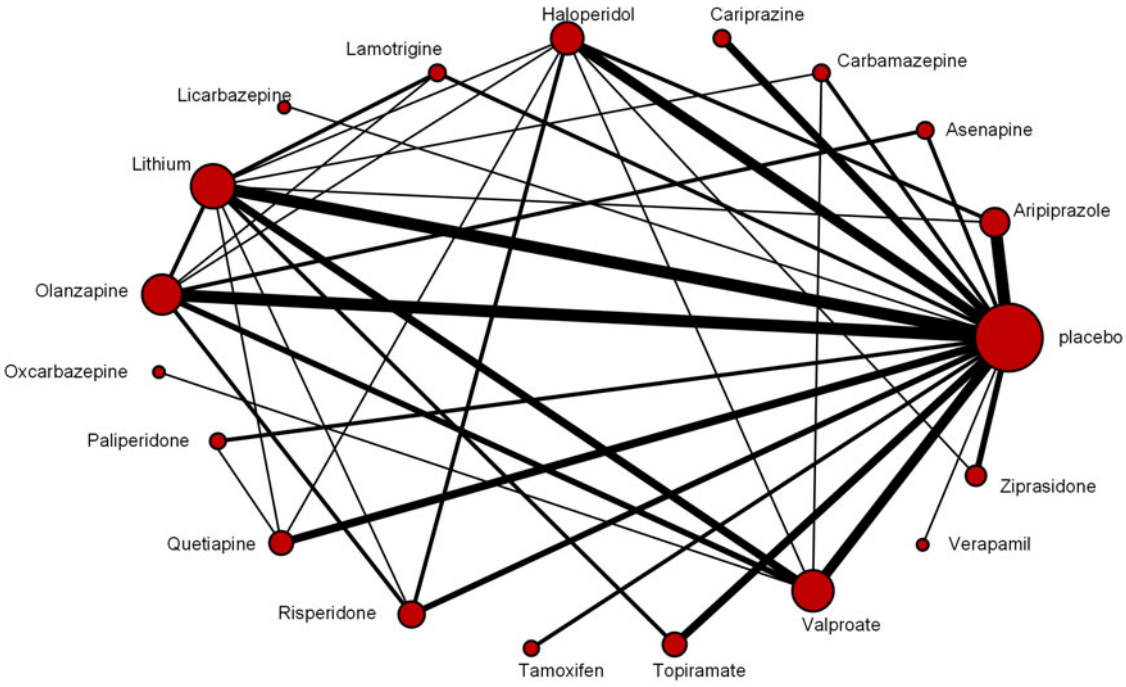

Fig. 1. Network of antimanic treatment comparisons. The size of the nodes corresponds to the number of trials involving that treatment. Directly comparable treatments are linked with a line, the thickness of which corresponds to the number of trials assessing that comparison.

reported in Tables 1 and 2. Bipolar I manic or mixed state was diagnosed with DSM-IV or DSM-IV-TR in 91%, DSM-III-R in 6%, Research Diagnostic Criteria (RDC) in 2%, and a Chinese version of DSM-IV (excluding patients with hypomania or first mania) in 1% of the comparisons (Spitzer *et al*. 1978; APA, 2000; Li *et al*. 2008). Except for one head-to-head comparison of lithium *v*. olanzapine (Shafti, 2010), all studies used YMRS or MRS for dimensional measurement of manic syndrome severity. One study was single-blinded with allocation concealed from the assessors; the rest (99%) were double-blinded. For all placebo-controlled studies, and all but seven head-to-head studies some explanatory information on the method of allocation concealment was provided. While 64% of the comparisons in the network were against placebo, in 87%, a pharmaceutical company provided funding. In all but four included RCTs, study duration for the primary endpoint was 3 or 4 weeks. Mean (±S.D.) age of trial participants was 38.9±12.3 years. The proportion of male patients was similar in placebo controlled (53%) and head-to-head (48%) trials. Trial completion rates per study arm could be obtained for all comparisons but one (Kakkar *et al*. 2009); and averaged 64% in drug *v*. placebo, and 72% in head-to-head comparisons. With reported and tabulated study and patient characteristics this reasonably homogenous network consisted of 19 joints of candidate antimanic treatments or placebo, as monotherapy in an ITT based analytic sample of 13920 manic patients with 128 data points (number of study arms) (Tables 1 and 2; Fig. 1).

Assessment of antimanic drugs efficacy

MTM indicated superior improvement in manic symptom ratings (based on SMD and CrI) over placebo for 13 drugs: aripiprazole (0.37, 0.2–0.55), asenapine (0.36, 0.08–0.63), carbamazepine (0.44, 0.15–0.71), cariprazine (0.47, 0.22–0.73), haloperidol (0.54, 0.38–0.7), lithium (0.45, 0.3–0.61), olanzapine (0.48, 0.34–0.62), paliperdone (0.37, 0.08–0.66), quetiapine (0.35, 0.14–0.56), risperidone (0.65, 044–0.85), tamoxifen (2.92, 2.38–3.48), valproate (0.32, 0.15–0.5), and ziprasidone (0.33, 0.08–0.59); and lack of efficacy for 5 others: oxcarbazepine (0.3, −0.35 to 0.95) lamotrigine (0.13, −0.16 to 0.44), licarbazepine (0.09, −0.34 to 0.53), topiramate (0.07, −0.15 to 0.28), and verapamil (−0.02, −0.96 to 0.94) (Figs 2 and 3). We created hierarchies of effect sizes on the basis of SUCRA rankings for all outcomes, and also provided point estimates as SMD or OR with their CrI. Among clinically used antimanic treatments, improvements in manic symptom ratings were comparable for all drugs except risperidone *v*. aripiprazole (0.27, 0.01–0.54) and *v*. valproate (0.33, 0.06–0.58), both significantly favoring risperidone.

MTM based on the categorical responder rates (OR) validated effectiveness of the same 13 agents and lack

Table 1. *Characteristics of included randomized, placebo-controlled single-agent trials of acute mania (61 comparisons from 41 studies*

Sites (n)[a]	Dose (mg/d, mEq/l)	Patients (n, ITT)	Mania rating scale	Completers (%)		Baseline mania (mean±S.D.)		Severity (% of max)		Mania score change (mean±S.D.)		Change (%)		Response (%)		Source (references)
				Rx	PBO	Rx	PBO	Rx	PBO	Rx	PBO	Rx	PBO	Rx	PBO	
Aripiprazole																
38	15–30	243	YMRS	41.5	21.2	28.2±5.0	29.7±5.0	47.0	49.5	8.2±12.0	3.4±12.0	29.1	11.4	39.8	19.2	Keck *et al.* (2003*a*)
29	15–30	268	YMRS	54.7	51.9	28.8±4.9	28.5±4.9	48.0	47.4	12.5±11.0	7.2±11.0	43.4	25.3	52.9	32.6	Sachs *et al.* (2006)
42	15–30	317	YMRS	44.2	47.3	28.5±5.6	28.9±5.9	47.5	48.2	12.6±10.0	9.0±10.0	44.4	31.2	46.8	34.4	Keck *et al.* (2009)
59	15–30	318	YMRS	75.4	71.2	28.0±5.8	28.3±5.8	46.7	47.2	12.0±10.0	9.7±10.0	42.8	34.3	47.0	38.2	Young *et al.* (2009)
56	15	257	YMRS	42.7	40.3	27.9±5.4	28.3±5.4	46.6	47.1	10.0±11.0	10.1±11.0	35.8	35.8	40.9	37.7	El Mallakh *et al.* (2010)
56	30	259	YMRS	39.7	40.3	27.8±5.5	28.3±5.4	46.4	47.1	10.8±11.0	10.1±11.0	38.8	35.8	45.0	37.7	El Mallakh *et al.* (2010)
Asenapine																
55	18.2±3.1	292	YMRS	62.9	61.5	28.3±5.5	29.0±6.1	47.2	48.3	13.1±11.3	7.4±11.6	46.3	25.5	42.3	25.2	McIntyre *et al.* (2009)
61	18.4±2.7	277	YMRS	67.0	58.2	29.4±6.7	28.3±6.3	49.0	47.2	11.5±10.8	7.8±10.7	39.1	27.6	42.6	34.0	McIntyre *et al.* (2010)
Carbamazepine																
24	756	192	YMRS	49.5	44.7	26.6±5.5	27.3±5.3	44.3	45.5	8.7±11.0	5.2±9.4	32.8	18.9	40.4	21.4	Weisler *et al.* (2004)
25	643	235	YMRS	65.6	54.7	28.5±4.4	27.9±4.9	47.4	46.6	15.1±9.6	7.1±9.2	53.0	25.5	60.8	28.7	Weisler *et al.* (2005)
Cariprazine																
29	3–12	236	YMRS	63.6	61.9	30.6±5.4	30.2±5.4	51.0	50.3	13.3±12.0	7.2±11.9	43.5	23.8	48.3	24.8	Knesevich *et al.* (2009)
28	3–12	310	YMRS	68.4	69.7	32.3±5.8	32.1±5.6	53.8	53.5	17.3±12.2	12.8±11.7	53.6	39.9	58.9	44.1	Bose *et al.* (2012)
66	3–6	327	YMRS	78.2	76.3	33.2±5.6	32.6±5.8	55.3	54.3	17.5±11.2	11.3±11.7	52.7	34.7	60.6	37.5	Calabrese *et al.* (2013)
66	6–12	325	YMRS	71.3	76.3	32.9±4.7	32.6±5.8	54.8	54.3	16.8±10.4	11.3±11.7	51.1	34.7	59.3	37.5	Calabrese *et al.* (2013)
Haloperidol																
49	2–8	198	YMRS	77.8	60.4	32.3±6.0	33.1±6.6	53.8	55.2	15.7±13.0	8.3±13.0	48.6	25.1	56.1	35.0	McIntyre *et al.* (2005)
20	2–12	282	YMRS	89.0	85.0	32.1±6.9	31.5±6.7	53.5	52.5	15.1±10.0	9.4±11.0	47.0	29.8	47.7	33.3	Smulevich *et al.* (2005)
59	5–15	313	YMRS	73.3	71.2	27.6±5.6	28.3±5.8	46.0	47.2	12.8±10.0	9.7±10.0	46.5	34.3	49.7	38.2	Young *et al.* (2009)
33	8–30	258	MRS	71.3	50.0	30.7±7.4	31.3±7.7	57.9	59.1	15.9±10.6	6.1±9.9	51.9	19.5	54.7	20.5	Vieta *et al.* (2010*b*)
56	2.5–10	117	YMRS	40.0	52.5	26.6±4.5	26.9±5.6	44.3	44.8	14.3±11.4	6.8±14	53.8	25.3	65.0	44.3	Katagiri *et al.* (2012)
Lamotrigine																
38	25–200	151	MRS	50.0	60.0	29.6±7.8	29.5±7.0	55.9	55.7	11.6±14.0	11.4±12.3	39.2	38.6	55.4	39.0	Bowden *et al.* (2000)
47	50	179	MRS	62.4	64.2	26.4±6.5	25.9±6.1	49.8	48.9	9.3±11.0	9.5±11.0	35.2	36.7	44.0	46.3	Goldsmith *et al.* (2003)
Licarbazepine																
28	1000–2000	313	YMRS	63.4	68.7	27.5±5.2	27.4±5.3	45.8	45.7	9.2±10.3	8.3±9.3	33.5	30.3	35.5	34.8	|||CLIC477D2301 2007

Lithium															
9	⩾1200	107	MRS	38.9	36.5	27.1±7.4	28.1±6.3	51.1	53.1	9.3±16.0	4.1±11.0	34.3	14.4	48.6	25.0	Bowden *et al.* (1994)
38	0.7–1.3	154	MRS	73.0	60.0	28.7±6.9	29.5±7.0	54.2	55.7	15.6±13.0	11.4±12.3	54.4	38.6	62.3	39.0	Bowden *et al.* (2000)
47	0.8–1.3	131	MRS	44.4	64.2	26.2±5.9	25.9±6.1	49.4	48.9	10.7±12.0	9.5±11.0	40.8	36.7	41.7	46.3	Goldsmith *et al.* (2003)
38	900	193	YMRS	85.7	69.1	33.3±7.1	34.0±6.9	55.5	56.7	15.2±15.0	6.7±15.0	45.6	19.7	53.1	27.4	Bowden *et al.* (2005)
40	1500	224	YMRS	74.3	73.9	30.1±7.4	30.0±6.3	50.2	50.0	12.9±12.0	7.7±12.0	42.9	25.7	46.0	28.0	Kushner *et al.* (2006)
40	1500	226	YMRS	81.6	86.6	30.7±7.5	31.7±7.3	51.2	52.8	13.8±12.0	8.4±12.0	45.0	26.5	46.0	28.0	Kushner *et al.* (2006)
42	900–1500	318	YMRS	48.8	47.3	29.4±5.9	28.9±5.9	49.0	48.2	12.0±10.0	9.0±10.0	40.9	31.2	45.8	34.4	Keck *et al.* (2009)
Olanzapine															
16	5–20	136	YMRS	61.4	34.8	28.7±6.7	27.7±6.5	47.8	46.1	10.3±13.0	4.9±12.0	35.8	17.6	48.6	24.2	Tohen *et al.* (1999)
24	5–20	110	YMRS	61.8	41.7	28.8±6.7	29.4±6.8	47.9	49.1	14.8±13.0	8.1±13.0	51.4	27.6	64.8	42.9	Tohen *et al.* (2000)
42	11.4±2.5	300	YMRS	74.0	73.3	23.8±2.8	23.5±2.5	39.7	39.2	9.4±8.5	7.4±8.0	39.5	31.5	40.8	31.3	Tohen *et al.* (2008)
55	15.8±2.3	291	YMRS	79.6	61.5	28.6±5.9	29.0±6.1	47.7	48.3	13.9±10.7	7.4±11.6	48.6	25.5	50.0	25.2	McIntyre *et al.* (2009)
61	15.9±2.5	297	YMRS	78.5	58.2	29.7±6.6	28.3±6.3	49.5	47.2	14.6±11.4	7.8±10.7	49.2	27.6	54.7	34.0	McIntyre *et al.* (2010)
16	5–20	201	YMRS	68.6	52.5	27.7±5.9	26.9±5.6	46.2	44.8	12.6±10.0	6.8±14.0	45.5	25.3	51.0	44.3	Katagiri *et al.* (2012)
Paliperidone															
52	6–12	294	YMRS	82.0	62	27.3±5.0	26.5±5.0	45.5	44.2	13.2±8.7	7.4±10.7	48.4	27.9	44.2	34.6	Vieta *et al.* (2010*a*)
45	12	224	YMRS	65.2	59	28.2±5.0	28.7±5.2	47	47.8	13.5±9.2	10.1±10.2	47.9	35.2	51.0	43.0	Berwaerts *et al.* (2012)
45	6	227	YMRS	57.5	59	28.0±5.6	28.7±5.2	46.7	47.8	11.4±10	10.1±10.2	40.7	35.2	43.0	43.0	Berwaerts *et al.* (2012)
45	3	222	YMRS	63.4	59	28.7±6.3	28.7±5.2	47.8	47.8	9.1±11.2	10.1±10.2	31.7	35.2	36.0	43.0	Berwaerts *et al.* (2012)
Quetiapine															
38	600–800	202	YMRS	90.7	69.1	32.7±6.5	34.0±6.9	54.5	56.7	14.6±16.0	6.7±15.0	44.7	19.7	53.3	27.4	Bowden *et al.* (2005)
49	600–800	201	YMRS	64.7	60.4	34.0±6.1	33.1±6.6	56.7	55.2	12.3±14.0	8.3±13.0	36.1	25.1	42.6	35.0	McIntyre *et al.* (2005)
52	100–800	296	YMRS	79.0	62.0	27.6±5.1	26.5±5.0	46.0	44.2	11.7±9.3	7.4±10.7	42.4	27.9	49.0	34.6	Vieta *et al.* (2010*a*)
48	400–800	308	YMRS	71.6	72	28.8±5.4	28.4±5.1	48.0	47.3	14.3±11.0	10.5±11.0	49.8	37.0	55.0	33.3	Cutler *et al.* (2011)
Risperidone															
30	1–6	246	YMRS	56.0	41.6	29.1±5.1	29.2±5.5	48.5	48.7	10.6±9.5	4.8±9.5	36.4	16.4	43.3	24.4	Hirschfeld *et al.* (2004)
8	1–6	286	YMRS	89.0	70.8	37.1±8.5	37.5±8.4	61.8	62.5	22.7±13.0	10.5±16.0	61.2	28.0	74.3	36.6	Khanna *et al.* (2005)
20	1–6	291	YMRS	90.3	85.0	31.3±6.5	31.5±6.7	52.2	52.5	13.9±10.0	9.4±11.0	44.4	29.8	47.2	33.3	Smulevich *et al.* (2005)
Tamoxifen															
1	20–140	16	YMRS	50.0	62.5	30.3±7.0	24.3±5.3	50.5	40.5	18.3±4.3	−4.7±4.1	60.4	−19.2	62.5	12.5	Zarate *et al.* (2007)
1	80	58	YMRS	82.9	67.7	38.6±5.0	37.2±6.6	64.3	62.0	16.6±12.0	−4.8±9.1	43.0	−12.9	43.8	3.8	Yildiz *et al.* (2008)
Topiramate															
40	200+400	326	YMRS	70.0	73.9	30.5±7.5	30.0±6.3	50.8	50.0	6.0±12	7.7±12.0	19.7	25.7	27.0	28.0	Kushner *et al.* (2006)
2	400+600	308	YMRS	58.9	72.0	29.2±5.7	28.3±5.8	48.7	47.2	8.1±11	7.7±10.0	27.6	27.2	27.0	28.0	Kushner *et al.* (2006)
2	400	213	YMRS	56.0	73.6	30.4±7.3	29.5±5.7	50.7	49.2	5.1±10	6.4±10.0	16.8	21.7	27.0	28.0	Kushner *et al.* (2006)
40	400	227	YMRS	87.1	86.6	30.8±6.8	31.7±7.3	51.3	52.8	8.2±12	8.4±12.0	26.6	26.5	27.0	28.0	Kushner *et al.* (2006)

Table 1 (*cont.*)

Sites (*n*)[a]	Dose (mg/d, mEq/l)	Patients (*n*, ITT)	Mania rating scale	Completers (%)		Baseline mania (mean±S.D.)		Severity (% of max)		Mania score change (mean±S.D.)		Change (%)		Response (%)		Source (references)
				Rx	PBO	Rx	PBO	Rx	PBO	Rx	PBO	Rx	PBO	Rx	PBO	
Valproate																
1	⩾750	36	YMRS	23.5	21.1	28.2±5.8	28.6±6.9	47.0	47.7	11.4±10.0	0.2±9.9	40.5	0.6	52.9	10.5	Pope *et al.* (1991)
9	⩾1000	139	MRS	52.2	36.5	27.2±7.6	28.1±6.3	51.3	53.1	9.2±12.0	4.1±11.0	34.0	14.4	47.8	25.0	Bowden *et al.* (1994)
33	3057	364	MRS	57.8	51.9	26.6±5.6	26.6±5.6	50.2	50.2	11.9±11.0	9.0±11.0	44.7	33.8	48.1	33.9	Bowden *et al.* (2006)
42	848±136	285	YMRS	75.1	73.3	23.9±2.8	23.5±2.5	39.8	39.2	8.2±8.5	7.4±8.0	34.3	31.5	40.3	31.3	Tohen *et al.* (2008)
29	500–2500	222	MRS	17.0	18.0	32.9±5.8	33.0±6.7	62.1	62.3	10.1±10.8	8.5±12	30.7	25.8	–	–	Hirschfeld *et al.* (2010)
Verapamil																
1	480	20	MRS	17.6	40.0	29.0±9.0	26.0±7.0	54.7	49.1	1.1±11	1.3±13.0	3.8	5.0	37.5	16.7	Janicak *et al.* (1998)
Ziprasidone																
24	130±34	197	MRS	53.6	44.3	27.0±3.8	26.7±7.0	50.9	50.4	12.4±12	7.8±13.0	45.9	29.2	50.4	34.8	Keck *et al.* (2003*b*)
23	126	202	MRS	60.7	54.5	26.2±7.2	26.4±7.5	49.4	49.8	11.1±12	5.6±9.6	42.4	21.3	46.7	29.2	Potkin *et al.* (2005)
33	116	264	MRS	66.9	50.0	29.6±8.0	31.3±7.7	55.9	59.1	10.4±11.1	6.1±9.9	35.2	19.5	36.9	20.5	Vieta *et al.* (2010*b*)

S.D., Standard deviation; Rx, treated with study drug; PBO, placebo. Test drugs are listed alphabetically.

Ratings and changes are based on mania ratings by: YMRS (Young Mania Rating Scale: 11 items, range 0–60) or MRS (Mania Rating Scale: 11 items, range 0–53) from the Schedule for Affective Disorders and Schizophrenia–Change Version. Response is defined as ⩾50% improvement in the mania rating scale used relative to baseline.

References for included studies are provided in Appendix 4.

[a] In studies where actual site numbers are not reported, they are estimated as twice the reported number of countries.

of effect in the five others (Fig. 3). However, when effect size estimates for the categorical responder rates and their corresponding NNT values were considered, no significant differences between any of the clinically available effective drugs could be detected. Hierarchies by SUCRA rankings as well as OR-based point estimates between competing interventions and their corresponding NNT values are available in Fig. 1 of Appendix 3.

Assessment of all-cause discontinuations

All-cause discontinuation (dropout) was used as a measure of acceptability and Fig. 2 (upper triangle) provides ORs with their CrI based on this measure. Aripiprazole (0.68, 0.48–0.96), olanzapine (0.47, 0.35–0.63), quetiapine (0.63, 0.41–0.94), risperidone (0.60, 0.38–0.93), and valproate (0.67, 0.47–0.97) indicated significantly less all-cause, short-term discontinuation rates compared to placebo. Among drugs in clinical use, olanzapine was associated with significantly lower short-term discontinuation rates than haloperidol (0.63, 0.42–0.94), asenapine (0.53, 0.31–0.91), carbamazepine (0.53, 0.28–0.97), lithium (0.49, 0.33–0.73), and cariprazine (0.45, 0.25–0.81). Other drugs did not differ by acceptability. Fig. 4 presents a forest plot for all drugs listed from the most to least acceptable as compared to placebo according to their OR-based point estimates.

Sensitivity analyses

Considering that the unusually large effect size observed with tamoxifen, and the small effect size with topiramate might have artificially affected and skewed the computations made in the context of MTM, we performed a sensitivity analysis by excluding these outliers (Appendix 3: Fig. 2). This network without outliers yielded the exact same SUCRA-based as well as point estimate-based rank orderings by drug efficacy as SMD with the whole network (Figs 3 and 5; Appendix 3: Fig. 2). This analysis also confirmed the SMD-based superior effect estimates for risperidone against aripiprazole and valproate as detected in the complete network, and further found risperidone to be superior to ziprasidone (0.32, 0.03–0.6) and quetiapine (0.30, 0.04–0.56), as well as haloperidol over valproate (0.23, 0.02–0.43). Both networks were in full agreement in regard to significance or insignificance of the pooled drug effects against placebo.

In the present network 15 data points involved haloperidol. For ruling out differences in administered haloperidol doses as a potential source of bias, we repeated MTM for the primary outcome by excluding the haloperidol node from the network. This analysis also confirmed the original results by the entire network on the effectiveness and ineffectiveness of the individual drugs against placebo, as well as the superiority of risperidone over aripiprazole and valproate (Appendix 3: Fig. 3). The network without haloperidol node further indicated a minor but significant difference for olanzapine over valproate (0.2, 0.01–0.37).

The next sensitivity analysis considered if claimed increase in placebo responses, which itself was tested in a meta-regression analysis by publication year, might be related to reduced efficacy of newer drugs. For this analysis we ran MTM on the primary outcome after excluding 41 data points involving placebo treatments. Given that two thirds of the available RCTs in acute bipolar mania involved a placebo treatment, exclusion of the placebo node from the network resulted in 14 remaining joints. While in this analysis, risperidone was no longer significantly superior against aripiprazole and valproate, several new significant contrasts between antimanic treatments emerged. However, neither our previous standard meta-regression analysis (Yildiz *et al*. 2011*b*), nor the present one by a MTM approach, validated a steady increase in the placebo-associated responses or reduction in the treatment contrasts over time (see the section on meta-regressions). As such, while meaningfulness of this sensitivity analysis for assessing efficacy of antimanic treatments has been called into question, it may contribute by illustrating how ignorance of the major part of the evidence by placebo treatment may potentially change the results of evidence synthesis (Cipriani *et al*. 2009; Yildiz *et al*. 2014*b*). SUCRA-based rank orderings over the network without a placebo node as well as the point estimates between competing interventions are provided in Appendix 3: Fig. 4.

Comparative efficacy assessments of FGA v. SGAs v. MSs

In a sensitivity analysis, considering drugs in groups of FGA, SGAs, and MSs, we tested superiority of one group to another. This analysis by class effect confirmed antimanic effectiveness of the sole FGA haloperidol with SMD and CrI of 0.54 (0.39–0.69); SGAs 0.44 (0.36–0.51); and MSs 0.39 (0.28–0.49) compared to placebo, with no indication of significant differences among each other. With a perspective for future development of clinically usable brain penetrable protein kinase C inhibitors, we repeated the same analysis considering tamoxifen within the group of MSs. This analysis did not change the conclusion on group-based comparisons, but the point estimate-based rank ordering with SMD values and CrI of 0.60 (0.41–0.81) for MSs; 0.56 (0.27–0.84) for FGA; and 0.48 (0.32–0.63) for SGAs.

Table 2. *Characteristics of randomized, single-agent trials comparing two active drugs for treatment of acute mania (34 comparisons; 16 from three-armed placebo controlled trials)*

		Drugs (Rx)		Patients (*n*, ITT)			Completers (%)		Baseline mania (mean±S.D.)		Severity (% of max)		Mania score change (mean±S.D.)		Change (%)		Response (%)		
Design	Sites[a]	Rx1	Rx2	Rx1	Rx2	Mania rating/ trial weeks	Rx1	Rx2	Rx1	Rx2	Rx1	Rx2	Rx1	Rx2	Rx1	Rx2	Rx1	Rx2	Source (references)
RDB	1	Li	LTG	15	15	MRS/4	80.0	86.7	31.6	34.4	59.6	64.9	18.4	20.1	58.2	58.4	60.0	53.3	Ichim *et al.* (2000)[b]
RDB	38	Li	LTG	77	74	MRS/6	73.0	50.0	28.7±6.9	29.6±7.8	54.2	55.9	15.6±13.0	11.6±14.0	54.4	39.2	62.3	55.4	Bowden *et al.* (2000)[c]
RDB	47	Li	LTG	36	84	MRS/3	44.4	62.4	26.2±5.9	26.4±6.5	49.4	49.8	10.7±11.6	9.3±10.9	40.8	35.2	41.7	44.1	Goldsmith *et al.* (2003)[c]
RDB	40	Li	TPM	113	215	YMRS/3	74.0	70.0	30.1±7.4	30.5±7.5	50.2	50.8	12.9±11.8	6±12.1	42.9	19.7	46.0[d]	27.0[d]	Kushner *et al.* (2006)[c]
RDB	40	Li	TPM	114	115	YMRS/3	82.0	87.0	30.7±7.5	30.8±6.8	51.2	51.3	13.8±11.9	8.2±11.8	45.0	26.6	46.0[d]	27.0[d]	Kushner *et al.* (2006)[c]
RDB	1	Li	CBZ	24	24	YMRS/8	45.8	70.8	30.3	30.9	50.5	51.5	9.7	8.5	32.0	27.5	33.3[e]	33.3[e]	Small *et al.* (1991)
RDB	1	Li	VPA	13	14	MRS[f]/3	76.9	85.7	43.4±20.3	52.9±12.3	–	–	10.2±8.8[g]	27.1±20.9[g]	76.5	48.8	92.3	64.3	Freeman *et al.* (1992)[g]
RDB	9	Li	VPA	35	67	MRS/3	39.0	52.0	27.1±7.4	27.2±7.6	51.1	51.3	9.3±15.8	9.2±12	34.3	34.0	48.6	47.8	Bowden *et al.* (1994)[c]
RDB	42	Li	APZ	155	154	YMRS/3	49.0	47.0	29.4±5.9	28.5±5.6	49.0	46.7	12±10.3	12.6±10.4	40.9	44.4	45.8	46.8	Keck *et al.* (2009)[c]
RDB	1	Li	HAL	15	15	YMRS/4	93.3	80.0	28.4	24.8	47.3	41.3	15.7	10.2	55.3	41.1	–	–	Segal *et al.* (1998)[h]
RDB	1	Li	OLZ	15	15	MRS/4	80.0	93.3	31.6	31.7	59.6	59.8	18.4	21.5	58.2	67.8	–	–	Berk *et al.* (1999)[b]
RDB	7	Li	OLZ	71	69	YMRS/4	78.9	91.3	32.4±7.2	34±6.8	54.0	56.7	20.2±11.4	24.6±11.3	62.2	72.4	73.2	87.0	Niufan *et al.* (2008)
RDB	1	Li	OLZ	20	20	MSRS/3	85.0	90.0	79.2±9.6	80.3±8.7	60.9	61.8	53.7±9.5[g]	62.6±8.9[g]	17.3	6.8	25	15	Shafti (2010)[g]
RDB	38	Li	QTP	98	107	YMRS/3	85.7	90.6	33.3±7.1	32.7±6.5	55.5	62.9	15.2±15.4	14.6±15.7	45.7	44.7	53.1	53.3	Bowden *et al.* (2005)[c]
RDB	2	Li	QTP	77	77	YMRS/4	80.5	94.8	29.8±5.7	29.3±5.8	49.7	48.8	–	–	–	–	59.7	77.9	Li *et al.* (2008)
RDB	1	Li	RSP	15	15	YMRS/4	93.3	86.7	28.4	28.6	47.3	47.7	15.7	12.4	55.3	43.4	–	–	Segal *et al.* (1998)[h]
RDB	1	VPA	CBZ	15	15	YMRS/4	80.0	80.0	43.5±8.1	41.3±6.5	72.5	68.8	32.8±11.4	20.8±15.7	75.4	50.4	73.3	53.3	Vasudev *et al.* (2000)
RDB	2	VPA	OxCBZ	30	30	YMRS/3 (12)[i]	–	–	34.6±6.5	33.8±5.4	57.7	56.3	10±6.6	9.9±5	28.8	29.4	90.0[i]	80.0[i]	Kakkar *et al.* (2009)
RSB	1	VPA	HAL	21	15	YMRS/1	100	100	36.1±11	37.2±8.8	60.2	62.0	15.4±11	12.9±10.7	42.7	34.7	47.6	33.3	McElroy *et al.* (1996)
RDB	48	VPA	OLZ	123	125	YMRS/3	64.3	68.8	27.9±6.6	27.4±5.2	46.5	45.7	10.4±10.4	13.4±8.8	37.3	48.9	42.3	54.4	Tohen *et al.* (2002)
RDB	21	VPA	OLZ	60	55	MRS/3	62.0	68.0	30.8	32.3	58.1	60.9	14.8	17.2	48.1	53.3	–	–	Zajecka *et al.* (2002)
RDB	42	VPA	OLZ	186	201	YMRS/3	75.1	74.0	23.9±2.8	23.8±2.8	39.8	39.7	8.2±8.5	9.4±8.5	34.3	39.5	40.3	40.8	Tohen *et al.* (2008)[c]
RDB	76	APZ	HAL	173	164	YMRS/3 (12)	76.6	55.2	31.1±6.6	31.5±7.9	51.8	52.5	15.7	15.7	50.5	49.8	50.9	42.6	Vieta *et al.* (2005)
RDB	59	APZ	HAL	166	161	YMRS/3	75.0	73.0	28±5.8	27.6±5.6	46.7	46.0	12±10.3	12.8±10.2	42.8	46.5	47.0	49.7	Young *et al.* (2009)[c]
RDB	55	OLZ	ASN	188	189	YMRS/3	79.6	62.9	28.6±5.9	28.3±5.5	47.7	47.2	13.9±10.7	13.1±11.3	48.6	46.3	50.0	42.3	McIntyre *et al.* (2009)[c]
RDB	61	OLZ	ASN	203	183	YMRS/3	78.5	67.0	29.7±6.6	29.4±6.7	49.5	49.0	14.6±11.4	11.5±10.8	49.2	39.1	54.7	42.6	McIntyre *et al.* (2010)[c]
RDB	58	OLZ	HAL	231	213	YMRS/6 (12)	70.9	64.4	31.1±7.6	30.6±7.7	51.8	51.0	21.3±10.9[j]	23.5±10.7[j]	–	–	55.0[j]	62.0[j]	Tohen *et al.* (2003)
RDB	56	OLZ	HAL	104	20	YMRS/3	68.6	40.0	27.7±5.9	26.6±4.5	46.2	44.3	12.6±10	14.3±11.4	45.5	53.8	51.0	65.0	Katagiri *et al.* (2012)[c]
RDB	30	OLZ	RSP	164	164	YMRS/3	78.7	67.0	26.6±5.0	26.7±5	44.3	44.5	15	16.6	56.5	62.3	62.1	59.5	Perlis *et al.* (2006)
RDB	49	QTP	HAL	101	98	YMRS/3	64.7	77.8	34±6.1	32.3±6	56.7	53.8	12.3±13.5	15.7±13.4	36.2	48.6	42.6	56.1	McIntyre *et al.* (2005)[c]
RDB	52	QTP	PPD	192	190	YMRS/3	79.0	82.0	27.6±5.1	27.3±5	46.0	45.5	11.7±9.3	13.2±8.7	42.4	48.4	49.0	44.2	Vieta *et al.* (2010*a*)[c]
RDB	1	RSP	HAL	15	15	YMRS/4	86.7	80.0	28.6	24.8	47.7	41.3	12.4	10.2	43.4	41.1	–	–	Segal *et al.* (1998)[h]

RDB	20	RSP	HAL	153	144	YMRS/3	89.0	90.0	32.1±6.9	31.3±6.5	53.5	52.2	15.1±10.3	13.9±10.3	47.0	44.4	47.7	47.2	Smulevich *et al.* (2005)[c]
RDB	33	ZPS	HAL	176	170	MRS/3	66.9	71.3	29.6±8.0	30.7±7.4	55.9	57.9	10.4±11.1	15.9±10.6	35.2	51.9	36.9	54.7	Vieta *et al.* (2010*b*)[c]

Studies listed as based on comparisons.

S.D., Standard deviation; Rx, treated with study drug; RDB, Randomized double blind. RO, Randomized open. Drugs: APZ, aripiprazole; ASN, asenapine; CBZ, carbamazepine; HAL, haloperidol; Li, lithium; OLZ, olanzapine; OxCBZ, oxcarbazepine; PPD, paliperidone; QTP, quetiapine; RSP, risperidone; TPM, topiramate; VPA, valproate; ZPS, ziprasidone.

Ratings and changes are based on mania ratings by: YMRS (Young Mania Rating Scale: 11 items, range 0–60) or MRS (Mania Rating Scale: 11 items, range 0–53) from the Schedule for Affective Disorders and Schizophrenia–Change Version or MSRS (Beigel-Murphy Manic State Rating Scale: 26 items, range 0–130).

Unless otherwise specified response is defined as $\geqslant$50% improvement in the mania rating scale scores administered as relative to baseline.

References for included studies are provided in Appendix 4.

[a] In studies where actual site numbers are not reported, they are estimated as twice the reported number of countries.

[b] Lithium arm from the same study with two active controls published separately.

[c] Indicates results from placebo-controlled studies with two active treatment arms.

[d] Indicates pooled results.

[e] Response is defined as moderate improvement with CGI (Clinical Global Impressions Scale).

[f] Freeman *et al.* (1992) used a software system that unit normalized the MRS scores.

[g] Actual score at the end point; please note that smaller score at the end point reflects larger improvement.

[h] Same study tabulated three times as it has three comparisons between three active drugs.

[i] Mania score change results are at 3 weeks; and response indicates rate of remission (defined as YMRS of $\leqslant$12) at 12 weeks.

[j] Mania score change is defined as least squares mean and response is defined as $\geqslant$70% improvement with the YMRS at 6 weeks end point.

Meta-regression analysis

Informed by our previous standard meta-analysis indicating greater drug–placebo contrasts in the presence of psychosis (Yildiz *et al.* 2011*b*), we examined if such a moderator effect survives when head-to-head RCTs are involved in the context of MTR. This analysis revealed that psychosis was no longer acting as a potential moderator. Subgroup analysis considering antipsychotics and MSs separately also did not distinguish any significant effect of psychosis on the computed treatment contrasts (Appendix 2: Table 1). Various other MTR models considering mixed-state diagnosis, all-cause discontinuation rates, publication year, and sponsorship, as potential effect-modifiers did not reveal significance (Appendix 2: Table 1).

Publication bias

Finally, we explored presence of publication bias or small study effects with a funnel plot technique expanded to MTM. As visible in the depicted funnel plot, the observed asymmetry was primarily caused by one small study (Zarate *et al.* 2007), suggesting a small study effect rather than publication bias (Fig. 6).

Consistency checking

By considering each of the primary and secondary outcomes we checked consistency in the three treatment loops. For the primary outcome as SMD all of the 32 loops, but one was found consistent. In regard to the secondary outcomes three of the 22 loops on responder rates, and two out of 30 loops on all-cause discontinuation rates were inconsistent; overall, indicating a good fit for our model. Likewise, the model fit as assessed by the consistency parameters of DIC and D_{res} was good for all of the networks, examined (Appendix 2: Table 2). Pairwise comparisons by the random-effects Bayesian approach, also confirmed compatibility between the direct and indirect evidence (Fig. 3).

Discussion

This updated and homogenous network confirmed effectiveness of 13 antimanic treatments and provided evidence-based hierarchies on their efficacy and acceptability. For the primary outcome as SMD, the hierarchies indicated by the SUCRA probabilities were in full agreement with the point estimates on all effective antimanic drugs (Figs. 2 and 3). Point estimates for the SMDs against placebo for the 12 effective antimanic drugs in clinical use ranged from 0.32 to 0.66 with no indication of superiority of one to another, but for risperidone over aripiprazole and valproate. MTM

Tamoxifen	1.07 (0.3 to 3.59)	0.86 (0.25 to 2.8)	1.36 (0.41 to 4.4)	0.61 (0.17 to 2.09)	0.67 (0.2 to 2.16)	0.72 (0.2 to 2.53)	1.06 (0.29 to 3.79)	0.93 (0.27 to 3.07)	0.72 (0.2 to 2.54)	1.02 (0.29 to 3.39)	—	0.89 (0.25 to 3.07)	0.94 (0.27 to 3.11)	0.18 (0.02 to 1.55)	0.55 (0.15 to 1.97)	0.5 (0.12 to 2.06)	0.63 (0.2 to 1.97)	0.47 (0.13 to 1.57)
2.27 (1.7 to 2.88)	**Risperidone**	0.81 (0.48 to 1.37)	1.28 (0.78 to 2.1)	0.57 (0.29 to 1.14)	0.63 (0.37 to 1.08)	0.67 (0.32 to 1.37)	1 (0.48 to 2.11)	0.88 (0.5 to 1.54)	0.68 (0.34 to 1.36)	0.95 (0.52 to 1.76)	—	0.84 (0.43 to 1.64)	0.89 (0.5 to 1.56)	0.17 (0.02 to 1.1)	0.51 (0.24 to 1.08)	0.47 (0.17 to 1.24)	**0.6 (0.38 to 0.93)**	**0.44 (0.23 to 0.83)**
2.38 (1.82 to 2.97)	0.11 (−0.13 to 0.35)	**Haloperidol**	**1.58 (1.06 to 2.38)**	0.71 (0.39 to 1.31)	0.78 (0.5 to 1.21)	0.83 (0.43 to 1.58)	1.24 (0.64 to 2.45)	1.09 (0.72 to 1.66)	0.84 (0.45 to 1.58)	1.18 (0.71 to 1.98)	—	1.04 (0.59 to 1.82)	1.1 (0.68 to 1.78)	0.21 (0.03 to 1.32)	0.63 (0.32 to 1.26)	0.58 (0.23 to 1.47)	0.74 (0.53 to 1.04)	**0.54 (0.32 to 0.95)**
2.44 (1.88 to 3.02)	0.17 (−0.06 to 0.39)	0.06 (−0.14 to 0.25)	**Olanzapine**	**0.45 (0.25 to 0.81)**	**0.49 (0.33 to 0.73)**	**0.53 (0.28 to 0.97)**	0.78 (0.41 to 1.51)	0.69 (0.44 to 1.07)	**0.53 (0.31 to 0.91)**	0.74 (0.45 to 1.23)	—	0.65 (0.37 to 1.17)	0.69 (0.47 to 1.02)	**0.13 (0.02 to 0.83)**	**0.4 (0.21 to 0.78)**	**0.37 (0.14 to 0.91)**	**0.47 (0.35 to 0.63)**	**0.34 (0.2 to 0.58)**
2.44 (1.86 to 3.07)	0.18 (−0.15 to 0.5)	0.07 (−0.24 to 0.36)	0.01 (−0.28 to 0.3)	**Cariprazine**	1.1 (0.61 to 1.99)	1.17 (0.55 to 2.47)	1.74 (0.81 to 3.81)	1.53 (0.83 to 2.84)	1.18 (0.56 to 2.51)	1.66 (0.87 to 3.23)	—	1.46 (0.71 to 2.99)	1.55 (0.83 to 2.89)	0.3 (0.04 to 1.93)	0.9 (0.41 to 1.97)	0.82 (0.3 to 2.23)	1.04 (0.63 to 1.73)	0.77 (0.39 to 1.51)
2.46 (1.91 to 3.05)	0.2 (−0.06 to 0.44)	0.09 (−0.13 to 0.29)	0.03 (−0.17 to 0.22)	0.02 (−0.28 to 0.31)	**Lithium**	1.06 (0.57 to 1.96)	1.58 (0.83 to 3.06)	1.39 (0.9 to 2.17)	1.07 (0.58 to 2.01)	1.51 (0.94 to 2.45)	—	1.33 (0.73 to 2.39)	1.4 (0.89 to 2.2)	0.27 (0.04 to 1.68)	0.81 (0.45 to 1.49)	0.74 (0.29 to 1.86)	0.94 (0.69 to 1.29)	0.7 (0.43 to 1.13)
2.48 (1.88 to 3.13)	0.21 (−0.13 to 0.56)	0.1 (−0.21 to 0.43)	0.04 (−0.26 to 0.36)	0.03 (−0.33 to 0.42)	0.02 (−0.28 to 0.34)	**Carbamazepine**	1.48 (0.67 to 3.41)	1.31 (0.68 to 2.54)	1.01 (0.46 to 2.23)	1.41 (0.72 to 2.87)	—	1.24 (0.59 to 2.67)	1.32 (0.7 to 2.53)	0.26 (0.03 to 1.68)	0.76 (0.34 to 1.75)	0.69 (0.25 to 1.97)	0.88 (0.51 to 1.56)	0.65 (0.33 to 1.34)
2.54 (1.94 to 3.19)	0.28 (−0.08 to 0.63)	0.17 (−0.17 to 0.49)	0.11 (−0.22 to 0.43)	0.1 (−0.28 to 0.49)	0.08 (−0.24 to 0.41)	0.07 (−0.35 to 0.46)	**Paliperidone**	0.88 (0.44 to 1.72)	0.68 (0.3 to 1.51)	0.95 (0.5 to 1.81)	—	0.84 (0.39 to 1.8)	0.89 (0.44 to 1.75)	0.17 (0.02 to 1.13)	0.51 (0.22 to 1.18)	0.47 (0.16 to 1.32)	0.6 (0.33 to 1.06)	**0.44 (0.21 to 0.91)**
2.54 (1.98 to 3.14)	**0.27 (0.01 to 0.54)**	0.17 (−0.04 to 0.37)	0.11 (−011 to 0.33)	0.1 (−0.21 to 0.41)	0.08 (−0.14 to 0.31)	0.06 (−0.27 to 0.39)	0 (−0.34 to 0.34)	**Aripiprazole**	0.77 (0.4 to 1.48)	1.08 (0.64 to 1.86)	—	0.95 (0.52 to 1.74)	1.01 (0.61 to 1.66)	0.2 (0.03 to 1.22)	0.58 (0.3 to 1.16)	0.53 (0.21 to 1.35)	**0.68 (0.48 to 0.96)**	**0.5 (0.29 to 0.87)**
2.56 (1.96 to 3.19)	0.29 (−0.05 to 0.62)	0.18 (−0.13 to 0.49)	0.12 (−0.15 to 0.39)	0.12 (−0.26 to 0.49)	0.09 (−0.21 to 0.41)	0.08 (−0.32 to 0.46)	0.01 (−0.39 to 0.42)	0.02 (−0.31 to 0.34)	**Asenapine**	1.4 (0.71 to 2.8)	—	1.23 (0.58 to 2.61)	1.31 (0.69 to 2.46)	0.25 (0.03 to 1.68)	0.76 (0.34 to 1.71)	0.69 (0.25 to 1.92)	0.88 (0.51 to 1.53)	0.65 (0.32 to 1.31)
2.57 (1.99 to 3.17)	0.3 (0 to 0.58)	0.19 (−0.07 to 0.43)	0.13 (−0.12 to 0.37)	0.12 (−0.2 to 0.45)	0.1 (−0.14 to 0.35)	0.09 (−0.27 to 0.42)	0.02 (−0.3 to 0.34)	0.02 (−0.25 to 0.29)	0.01 (−0.34 to 0.35)	**Quetiapine**	—	0.88 (0.46 to 1.67)	0.93 (0.53 to 1.59)	0.18 (0.02 to 1.14)	0.54 (0.26 to 1.1)	0.49 (0.19 to 1.28)	**0.63 (0.41 to 0.94)**	**0.46 (0.25 to 0.83)**
2.62 (1.78 to 3.48)	0.35 (−0.33 to 1.02)	0.24 (−0.43 to 0.9)	0.18 (−0.48 to 0.83)	0.17 (−0.53 to 0.87)	0.15 (−0.51 to 0.82)	0.14 (−0.57 to 0.83)	0.07 (−0.64 to 0.78)	0.07 (−0.6 to 0.74)	0.06 (−0.64 to 0.75)	0.05 (−0.63 to 0.73)	**Oxcarbazepine**	—	—	—	—	—	—	—
2.58 (1.99 to 3.2)	0.31 (−0.02 to 0.63)	0.21 (−0.08 to 0.48)	0.15 (−0.14 to 0.43)	0.14 (−0.22 to 0.5)	0.12 (−0.17 to 0.42)	0.1 (−0.28 to 0.47)	0.04 (−0.35 to 0.42)	0.04 (−0.27 to 0.34)	0.02 (−0.35 to 0.4)	0.02 (−0.31 to 0.34)	−0.03 (−0.73 to 0.66)	**Ziprasidone**	1.06 (0.57 to 1.98)	0.21 (0.03 to 1.33)	0.61 (0.28 to 1.35)	0.56 (0.21 to 1.52)	0.71 (0.43 to 1.18)	0.53 (0.27 to 1.03)
2.59 (2.04 to 3.18)	**0.33 (0.06 to 0.58)**	0.22 (−0.02 to 0.43)	0.16 (−0.04 to 0.34)	0.15 (−0.16 to 0.45)	0.13 (−0.09 to 0.34)	0.12 (−0.22 to 0.42)	0.05 (−0.29 to 0.38)	0.05 (−0.2 to 0.29)	0.04 (−0.28 to 0.34)	0.03 (−0.24 to 0.29)	−0.02 (−0.65 to 0.6)	0.01 (−0.3 to 0.31)	**Valproate**	0.19 (0.03 to 1.21)	0.58 (0.29 to 1.16)	0.53 (0.2 to 1.36)	**0.67 (0.47 to 0.97)**	**0.5 (0.28 to 0.88)**
2.93 (1.84 to 4.04)	0.66 (−0.32 to 1.63)	0.56 (−0.42 to 1.51)	0.5 (−0.47 to 1.45)	0.49 (−0.5 to 1.47)	0.47 (−0.5 to 1.42)	0.45 (−0.55 to 1.43)	0.39 (−0.62 to 1.37)	0.39 (−0.59 to 1.35)	0.37 (−0.63 to 1.35)	0.37 (−0.61 to 1.33)	0.31 (−0.84 to 1.46)	0.35 (−0.64 to 1.32)	0.34 (−0.64 to 1.3)	**Verapamil**	2.99 (0.45 to 22.8)	2.72 (0.37 to 22.5)	3.46 (0.58 to 24.3)	2.55 (0.41 to 18.6)
2.78 (2.18 to 3.41)	**0.51 (0.14 to 0.87)**	**0.4 (0.06 to 0.73)**	**0.35 (0.02 to 0.66)**	0.34 (−0.06 to 0.72)	0.32 (0 to 0.63)	0.3 (−0.12 to 0.7)	0.24 (−0.18 to 0.64)	0.24 (−0.11 to 0.57)	0.22 (−0.19 to 0.61)	0.22 (−0.15 to 0.57)	0.16 (−0.55 to 0.87)	0.2 (−0.2 to 0.58)	0.19 (−0.15 to 0.52)	−0.15 (−1.15 to 0.85)	**Lamotrigine**	0.91 (0.31 to 2.61)	1.16 (0.64 to 2.1)	0.86 (0.41 to 1.76)
2.82 (2.15 to 3.55)	**0.56 (0.07 to 1.04)**	0.45 (−0.02 to 0.91)	0.39 (−0.07 to 0.85)	0.38 (−0.12 to 0.88)	0.36 (−0.1 to 0.83)	0.35 (−0.18 to 0.85)	0.28 (−0.24 to 0.81)	0.28 (−0.18 to 0.75)	0.27 (−0.25 to 0.78)	0.26 (−0.22 to 0.74)	0.21 (−0.57 to 0.99)	0.24 (−0.26 to 0.75)	0.23 (−0.23 to 0.7)	−0.11 (−1.14 to 0.94)	0.04 (−0.48 to 0.58)	**Licarbazepine**	1.27 (0.53 to 3.05)	0.94 (0.36 to 2.48)
2.92 (2.38 to 3.48)	**0.65 (0.44 to 0.85)**	**0.54 (0.38 to 0.7)**	**0.48 (0.34 to 0.62)**	**0.47 (0.22 to 0.73)**	**0.45 (0.3 to 0.61)**	**0.44 (0.15 to 0.71)**	**0.37 (0.08 to 0.66)**	**0.37 (0.2 to 0.55)**	**0.36 (0.08 to 0.63)**	**0.35 (0.14 to 0.56)**	0.3 (−0.35 to 0.95)	**0.33 (0.08 to 0.59)**	**0.32 (0.15 to 0.5)**	−0.02 (−0.96 to 0.94)	0.13 (−0.16 to 0.44)	0.09 (−0.34 to 0.53)	**Placebo**	0.74 (0.48 to 1.15)
2.98 (2.41 to 3.59)	**0.71 (0.41 to 1)**	**0.6 (0.34 to 0.86)**	**0.55 (0.29 to 0.79)**	**0.54 (0.21 to 0.87)**	**0.52 (0.29 to 0.75)**	**0.5 (0.14 to 0.84)**	**0.44 (0.08 to 0.8)**	**0.44 (0.16 to 0.71)**	**0.42 (0.07 to 0.77)**	**0.42 (0.12 to 0.71)**	0.36 (−0.32 to 1.05)	**0.4 (0.07 to 0.73)**	**0.39 (0.12 to 0.66)**	0.05 (−0.92 to 1.03)	0.2 (−0.15 to 0.56)	0.16 (−0.33 to 0.64)	0.07 (−0.15 to 0.28)	**Topiramate**

Fig. 2. Efficacy and all-cause discontinuation of antimanic drugs. Drugs are reported in the order of SUCRA–based efficacy ranking. Comparisons between treatments should be read from left to right and the estimate is in the cell in common between the column-defining treatment and the row-defining treatment. For efficacy, standard mean differences (SMDs) with 95% credible intervals (CrI) higher than 0 favor the column-defining treatment. For all-cause discontinuation, odds ratios (ORs) with CrI lower than 1 favor the raw-defining treatment and ORs with CrI higher than 1 favor the column-defining treatment. To obtain SMDs for comparisons in the opposite direction, negative values should be converted into positive values, and vice versa. To obtain ORs for comparisons in the opposite direction, reciprocals should be taken. Significant results are in bold and underlined.

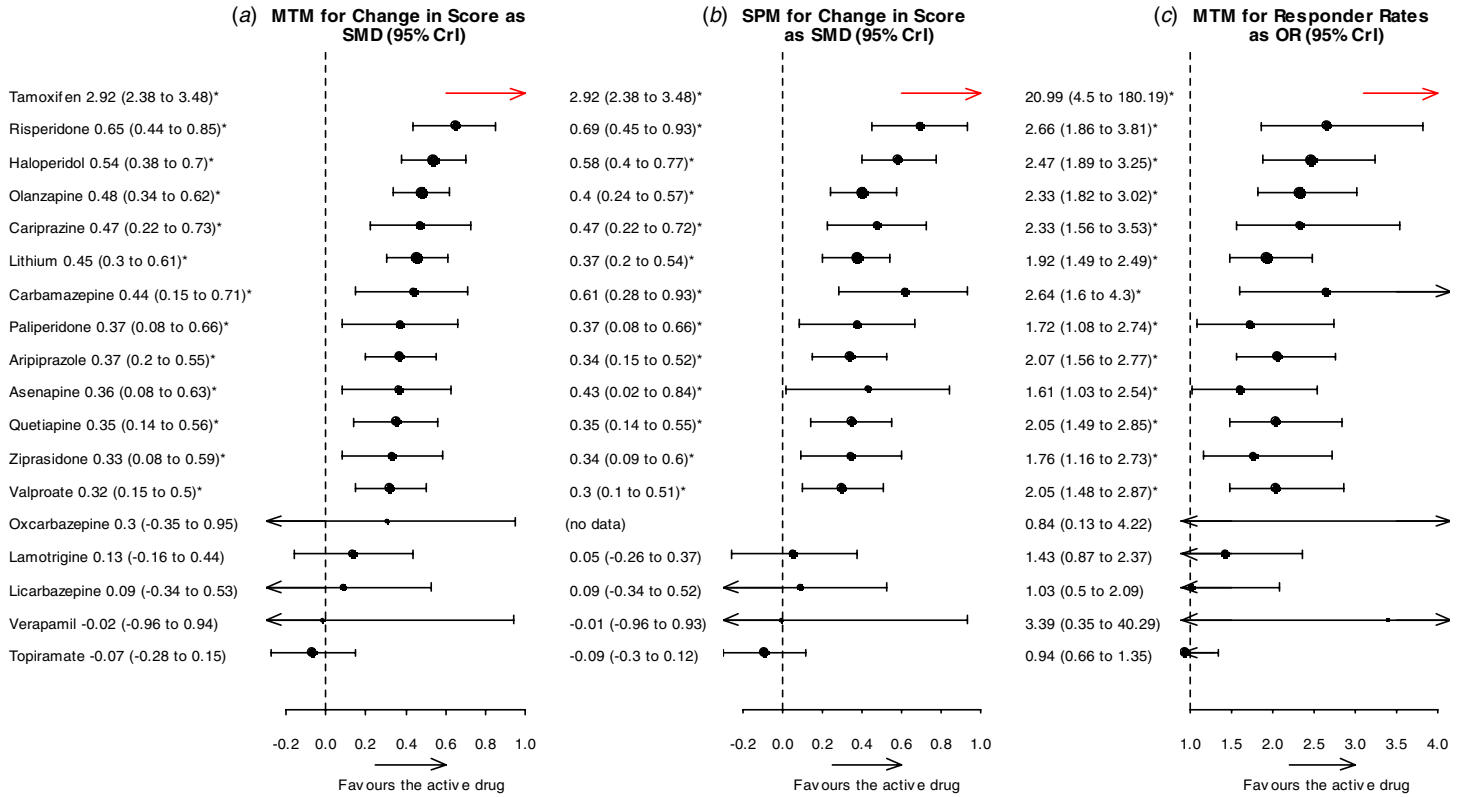

Fig. 3. Forest plots for efficacy of antimanic drugs compared with placebo. Treatments are ranked by the point estimates on the primary outcome. (*a*) Standardized mean differences [SMD, with 95% CrI (horizontal bars)] in reduction of mania symptom ratings between 18 candidate antimanic agents *v*. placebo, based on the multiple treatment meta-analysis (MTM); (*b*) the same measures analyzed by the standard pair-wise meta-analysis (SPM) with a Bayesian approach; (*c*) differences in drug–placebo response rates for the same antimanic agents as odds ratio (OR) with 95% CrI (horizontal bars), based on MTM. Vertical lines at SMD=0.0 and OR=1.0 indicates no drug *v*. placebo difference. (*) Indicates significantly superior effect compared to placebo.

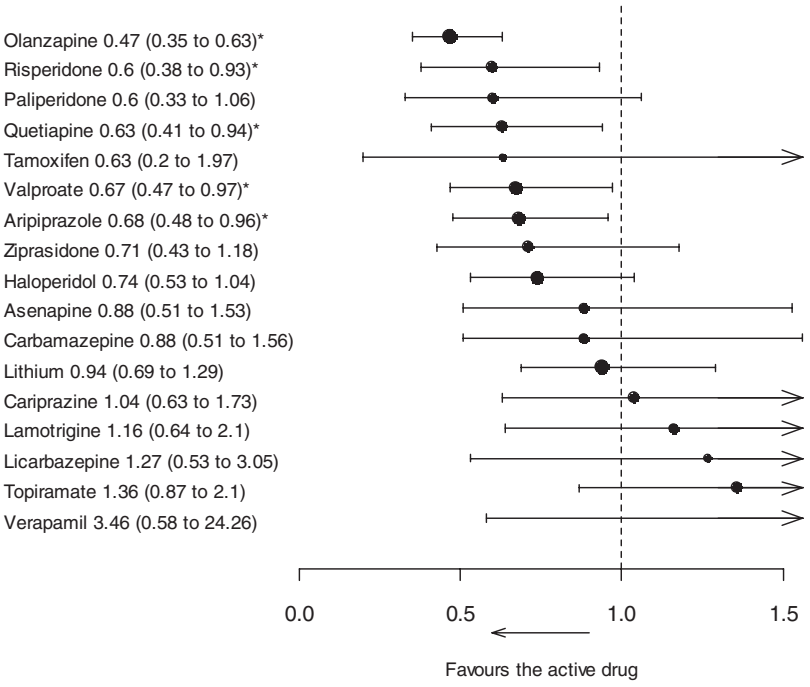

Fig. 4. Forest plot for all-cause discontinuation rates of antimanic drugs compared with placebo. Treatments are ranked by the point estimates as odds ratio (OR) with 95% credible interval (horizontal bars), based on MTM. Vertical line at OR=1.0 indicates no drug *v.* placebo difference. (*) Indicates significantly less all-cause discontinuation rates compared to placebo.

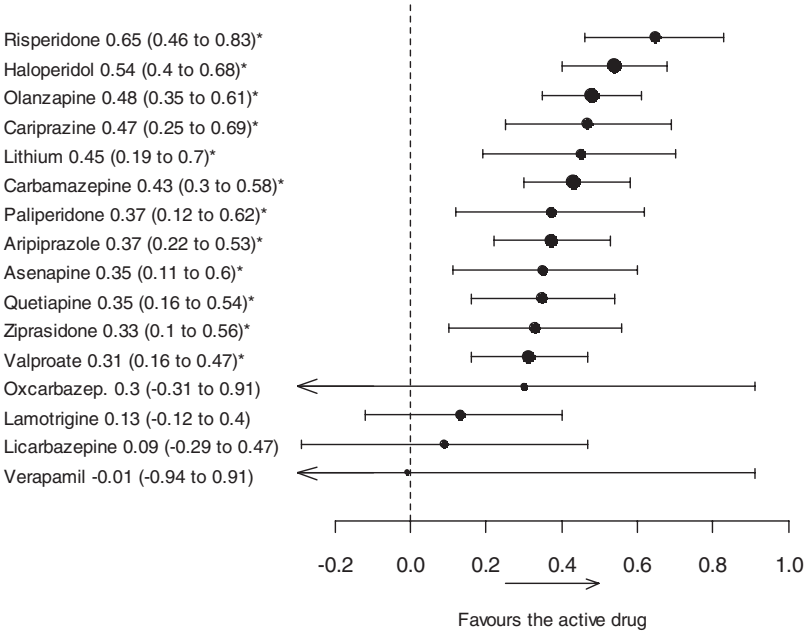

Fig. 5. Forest plot for efficacy of antimanic drugs compared with placebo without outliers. Treatments are ranked by the point estimates as SMD with 95% credible interval [horizontal bars], based on the MTM. Vertical line at SMD=0.0 indicates no drug *v.* placebo difference. (*) Indicates significantly superior effect compared to placebo.

based point estimates computed relative to placebo were compatible with the present Bayesian analysis and a recent frequentist SPM (Yildiz *et al.* 2011*a*). However, numerous disagreements between the present and a previous MTM of antimanic drugs (Cipriani *et al.* 2011) were detected on both SUCRA-based and point estimate-based ranking by efficacy, as well as on the reported pair-wise effect-size

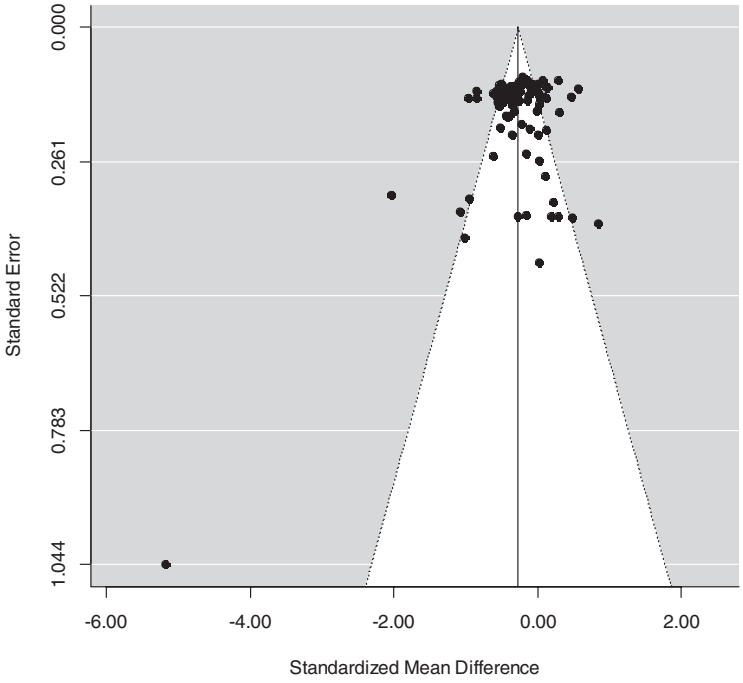

Fig. 6. The funnel plot for assessing presence of publication bias or small study effects. Asymmetry observed in the funnel plot primarily caused by one small study.

estimates. In the earlier MTM, haloperidol was found superior to seven other agents (aripiprazole, asenapine, carbamazepine, lithium, quetipaine, valproate, ziprasidone); and both olanzapine and risperidone over two other treatments (valproate, ziprasidone) (Cipriani *et al*. 2011). Disagreements between the present and previous network were indeed unsurprising given that 41% of the evidence structure at 53 data points was different. Besides, the initial network including add-on and combination treatments combined their findings with data attained in monotherapy trials (e.g. haloperidol node included data derived not only from trials of haloperidol but also haloperidol plus lithium or divalproex). The previous network also differentially considered chemically similar drugs jointly in the same node (e.g. data from RCTs involving paliperidone as monotherapy or add-on were incorporated in the joint for risperidone) (Cipriani *et al*. 2011). Consequently, while the effect estimate for risperidone in the present network derived solely from the trials involving risperidone as monotherapy, in the previous network it involved risperidone plus lithium or divalproex or carbamazepine and paliperidone plus lithium or divalproex as well as RCTs of risperidone and paliperidone, as monotherapy (Cipriani *et al*. 2011). Indeed, where similar trial circumstances and objective outcome measures operate, MTM allows consideration of add-on and monotherapy trials in the same network, so long as justified by a relevant clinical research question and identity of each treatment is respected (Mauri *et al*. 2008). However, considering that add-on trials of acute mania have special characteristics, often involving patients unresponsive to previous treatments (Yildiz *et al*. 2014*b*), we did not consider them jointly with the monotherapy trials.

The present MTM identified five drugs: aripiprazole, olanzapine, quetiapine, risperidone, and valproate with lower all-cause discontinuation rates compared to placebo. It is likely that acceptability of a drug, especially in the short term, reflects a fine-tuning between the drug' efficacy, speed of action, and tolerability based on favorably limited adverse effects. Yet, short-term acceptability may not generalize to longer treatment. Notably, olanzapine had significantly less all-cause discontinuation rates than asenapine, carbamazepine, cariprazine, haloperidol, and lithium. However, olanzapine risks weight gain and metabolic syndrome in the longer term (Centorrino *et al*. 2012; Grande *et al*. 2014; Mizuno *et al*. 2014). Given that short-term choices often end up as long-term treatments (Vieta *et al*. 2013), our findings on acute discontinuation rates should be interpreted with caution by considering undesirable drug effects also for the long term.

The analysis by drug class indicated comparative efficacy profiles for SGAs, MSs, and haloperidol. Group-based effect-size measures computed in the context of MTM were comparable with the previously reported SMDs and 95% confidence intervals (CIs) of 0.38 (0.26–0.50) for MSs; 0.40 (0.32–0.47) for SGAs;

and 0.54 (0.34–0.74) for FGA, by frequentist SPM (Yildiz *et al*. 2011*a*).

Sensitivity analysis conducted by excluding outliers yielded consistent SUCRA-based and point estimate-based rankings, and between-drug effect sizes, in support of internal validity of the reported findings (Appendix 3: Fig. 2). Likewise, exclusion of the haloperidol node yielded effect sizes compatible with the entire network (Appendix 3: Fig. 3). However, removal of the placebo node, resulting in loss of one third of the data points and five nodes from the entire network, skewed the findings on drugs with small treatment effects and less data points (e.g. ziprasidone; Appendix 3: Fig. 4).

The present network is meticulously structured with utmost care for similarity and comprehensiveness. Consistency of the network and internal validity of the present results was supported by several tests of consistency checking and sensitivity analyses. A number of clinical parameters were examined in several MTR models, but none was identified as an effect modifier (Appendix 2: Table 1), indicating a proper base for the MTM assumptions and indirect comparisons.

Robust analytic information obtained through advanced meta-analytic approaches can contribute to guide clinical decisions and reimbursement policies. However, their findings are based on RCT-based evidence; and RCT settings may not adequately represent real-world clinical samples (Yildiz *et al*. 2014*a*, *b*). Accordingly, this critical information can most efficiently be used by concurrent consideration of other accumulated data on a given drug's receptor or post-receptor level effects with consequent clinical reflections in light of patient's individual needs. Looking at the example of haloperidol and lithium or valproate, one may visualize how ranking by efficacy may be challenged by other sources of evidence as for detrimental effects of haloperidol in brain in contrast to lithium or valproate (Drevets *et al*. 1997; Bachmann *et al*. 2009; Machado-Vieira *et al*. 2009; Vernon *et al*. 2012). In conclusion, integrated information processing, supported by quantitative evidence-based point estimates for efficacy and safety, enriched by other data pertaining to drug-associated risks and benefits, and consideration of the needs and responses of individual patients should guide decisions for a skilful clinical care.

Supplementary material

For supplementary material accompanying this paper please visit http://dx.doi.org/10.1017/S0033291714001305.

Acknowledgments

We thank Drs Joris Berwaerts, Charles Bowden, Paul Keck Jr., Bernard Lerer, Roger McIntyre, Harrison G. Pope Jr., Steven Potkin, Gary S. Sachs, Alan Swann, Richard Weisler, and their colleagues, as well as representatives of Abbott, Bristol-Myers Squibb, Eli Lilly, Forest, Lundbeck, GlaxoSmithKline, Johnson & Johnson, Novartis, and Pfizer Corporations for sending us missing information on their published or unpublished trials.

No specific funding was received for this work.

Declaration of Interest

A. Yildiz has received research grants from and/or served as a consultant or speaker for Abdi Ibrahim, Actavis, Ali Raif, AstraZeneca, Bristol-Myers-Squibb, Janssen-Cilag, Pfizer, Sanofi-Aventis, and Servier Corporations. C. U. Correll has been a consultant or advisor to or has received honoraria from: Actelion, Alexza, AstraZeneca, Bristol-Myers-Squibb, Cephalon, Eli Lilly, Gerson Lehrman Group, Genentech, IntraCellular Therapies, Lundbeck, Medavante, Medscape, Merck, Novartis, Ortho-McNeill/Janssen/Johnson & Johnson, Otsuka, Pfizer, ProPhase, Roche, Sunovion, Takeda, Teva and Vanda. He has received grant support from Bristol-Myers-Squibb, Janssen/Johnson & Johnson, and Otsuka. E. Vieta has received grants and honoraria related to consultancy or advisory boards from several pharmaceutical companies, including Adamed, Alexza, Almirall, AstraZeneca, Bristol-Myers-Squibb, Cephalon, Eli Lilly, Ferrer, Forest Research Institute, Gedeon Richter, Glaxo-Smith-Kline, Janssen, Jazz, Johnson & Johnson, Lundbeck, Merck Sharpe & Dohme, Novartis, Otsuka, Pfizer, Roche, Sanofi, Senovel, Servier, Shering-Plough, Shire, Sunovion, Takeda, Teva, and United Biosource Corporation. M. Nikodem, R. J. Baldessarini, and the immediate family members of all authors have no current financial relationships with commercial entities that might represent potential conflicts of interest in this work.

References

Ades AE, Sculpher M, Sutton A, Abrams K, Cooper N, Welton N, Lu G (2006). Bayesian methods for evidence aynthesis in cost-effectiveness analysis. *Pharmacoeconomics* **24**, 1–19.

Ades AE, Welton N, Lu G. Introduction to Mixed Treatment Comparisons. (https://www.bris.ac.uk/cobm/research/mpes/mtc.html). Accessed 20 January 2014.

Ahrens B, Müller-Oerlinghausen B, Schou M, Wolf T, Alda M, Grof E, Grof P, Lenz G, Simhandl C, Thau K,

Volk J, Wolf R, Wolf T (1995). Excess cardiovascular and suicide mortality of affective disorders may be reduced by lithium prophylaxis. *Journal of Affective Disorders* **33**, 67–75.

Altman DG (1998). Confidence intervals for the number needed to treat. *British Medical Journal* **317**, 1309–1312.

American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders*, 4th edn. text revision. APA: Washington, DC.

Bachmann RF, Wang Y, Yuan P, Zhou R, Li X, Alesci S, Du J, Manji HK (2009). Common effects of lithium and valproate on mitochondrial functions: protection against methamphetamine-induced mitochondrial damage. *International Journal of Neuropsychopharmacology* **12**, 805–822.

Baldessarini RJ (2013). *Chemotherapy in Psychiatry*, 3rd edn. Springer Press: New York, NY.

Beigel A, Murphy DL (1971). Assessing clinical characteristics of the manic state. *American Journal of Psychiatry* **128**, 688–694.

Bose A, Starace A, Wang Q, Diaz E, Goodman J, Ruth A, Németh G, Laszlovszky I (2012). Cariprazine in the treatment of acute mania in bipolar disorder: a double-blind, placebo-controlled, phase III trial. ClinicalTrials.gov: NCT01058096. *European Neuropsychopharmacology* **22** (Suppl. 2), S285.

Bowden CL, Göğüş A, Grunze H, Häggström L, Rybakowski J, Vieta E (2008). 12-week, open, randomized trial comparing sodium valproate to lithium in patients with bipolar I disorder suffering from a manic episode. *International Journal of Clinical Psychopharmacology* **23**, 254–262.

Bowden CL, Mosolov S, Hranov L, Chen E, Habil H, Kongsakon R, Manfredi R, Lin HN (2010). Efficacy of valproate vs. lithium in mania or mixed mania: randomized, open 12-week trial. *International Journal of Clinical Psychopharmacology* **25**, 60–67.

Brown D, Silverstone T, Cookson J (1989). Carbamazepine compared to haloperidol in acute mania. *International Journal of Clinical Psychopharmacology* **4**, 229–238.

Calabrese JR, Keck PE Jr., Starace A, Lu K, Ruth A, Laszlovszky I, Nemeth G, Durgam S (2013). Efficacy and safety of low- and high-dose cariprazine in patients with acute mania associated with bipolar I disorder. ClinicalTrials.gov: NCT01058668. *European Neuropsychopharmacology* **23** (Suppl. 2), S378–S379.

Caldwell DM, Ades AE, Higgins JP (2005). Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *British Medical Journal* **331**, 897–900.

Centorrino F, Masters GA, Talamo A, Baldessarini RJ, Öngür D (2012). Metabolic syndrome in psychiatrically hospitalized patients treated with antipsychotics and other psychotropics. *Human Psychopharmacology* **27**, 521–526.

Chaimani C, Salanti G (2012). Using network meta-analysis to evaluate the existence of small-study effects in a network of interventions. *Research Synthesis Methods* **3**, 161–176.

Chang CK, Hayes RD, Perera G, Broadbent MT, Fernandes AC, Lee WE, Hotopf M, Stewart R (2011). Life expectancy at birth for people with serious mental illness and other major disorders from a secondary mental health care case register in London. *PLoS ONE* **6**, e19590.

Chinn S (2000). Simple method for converting an odds ratio to effect size for use in meta-analysis. *Statistics in Medicine* **19**, 3127–3131.

Cipriani A, Barbui C, Salanti G, Rendell J, Brown R, Stockton S, Purgato M, Spineli LM, Goodwin GM, Geddes JR (2011). Comparative efficacy and acceptability of antimanic drugs in acute mania: multiple-treatments meta-analysis. *Lancet* **378**, 1306–1315.

Cipriani A, Furukawa TA, Salanti G, Geddes JR, Higgins JP, Churchill R, Watanabe N, Nakagawa A, Omori IM, McGuire H, Tansella M, Barbui C (2009). Comparative efficacy and acceptability of 12 new-generation antidepressants: a multiple-treatments meta-analysis. *Lancet* **373**, 746–758.

Cipriani A, Hawton K, Stockton S, Geddes JR (2013). Lithium in the prevention of suicide in mood disorders: updated systematic review and meta-analysis. *British Medical Journal* **346**, f3646.

Dias S, Welton NJ, Caldwell DM, Ades AE (2010). Checking consistency in mixed treatment comparison meta-analysis. *Statistics in Medicine* **29**, 932–944.

Drevets WC, Price JL, Simpson JR Jr., Todd RD, Reich T, Vannier M, Raichle ME (1997). Subgenual prefrontal cortex abnormalities in mood disorders. *Nature* **386**, 824–827.

Eckermann S, Coory M, Willan AR (2009). Indirect comparison: relative risk fallacies and odds solution. *Journal of Clinical Epidemiology* **62**, 1031–1036.

Endicott J, Spitzer RL (1978). Diagnostic interview: the schedule for affective disorders and schizophrenia. *Achieves of General Psychiatry* **35**, 837–844.

Freeman TW, Clothier JL, Pazzaglia P, Lesem MD, Swann AC (1992). Double-blind comparison of valproate and lithium in the treatment of acute mania. *American Journal of Psychiatry* **149**, 108–111.

Gardner DM, Murphy AL, O'Donnell H, Centorrino F, Baldessarini RJ (2010). International consensus study of antipsychotic dosing. *American Journal of Psychiatry* **167**, 686–693.

Garza-Treviño ES, Overall JE, Hollister LE (1992). Verapamil vs. lithium in acute mania. *American Journal of Psychiatry* **149**, 121–122.

Glenny AM, Altman DG, Song F, Sakarovitch C, Deeks JJ, D'Amico R, Bradburn M, Eastwood AJ (2005). Indirect comparisons of competing interventions. *Health Technology Assessment* **9**, 1–134.

Goldsmith DR, Wagstaff AJ, Ibbotson T, Perry CM (2003). Lamotrigine: review of its use in bipolar disorder. Study ID Number: SCAA2008. *Drugs* **63**, 2029–2050.

Goodwin RK, Jamison KR (2007). *Manic-Depressive Illness*, 2nd edn. Oxford University Press: New York, NY.

Grande I, Bernardo M, Bobes J, Saiz-Ruiz J, Alamo C, Vieta E (2014). Antipsychotic switching in bipolar disorders: a systematic review. *International Journal of Neuropsychopharmacology* **17**, 497–507.

Janicak PG, Sharma RP, Pandey G, Davis JM (1998). Verapamil for the treatment of acute mania: double-blind, placebo-controlled trial. *American Journal of Psychiatry* **155**, 972–973.

Jonas D, Wilkins T, Bangdiwala S, Bann C, Morgan L, Thaler K, Amick H, Gartlehner G (2013). *Findings of Bayesian Mixed Treatment Comparison Meta-Analyses, Comparison and Exploration using Real-World Trial data and Simulation*. Agency for Healthcare Research and Quality (US), Methods Research Report No. 13-EHC039-EF.

Kakkar AK, Rehan HS, Unni KES, Gupta NK, Chopra D, Kataria D (2009). Comparative efficacy and safety of oxcarbazepine versus divalproex sodium in the treatment of acute mania: pilot study. *European Psychiatry* **24**, 178–182.

Knesevich M, Papadakis K, Bose A, Andor G, Laszlovszky I (2009). The efficacy and tolerability of cariprazine in acute mania associated with bipolar I disorder: a phase II trial. ClinicalTrials.gov: NCT00488618. *European Neuropsychopharmacology* **19** (Suppl. 3), S469.

Kriston L, Yildiz A (in press). Traditional and novel research synthesis approaches to support evidence-based treatment decisions In *Bipolar Book: History, Neurobiology, and Treatment* (ed. A. Yildiz, C. Nemeroff and P. Ruiz). Oxford University Press: New York.

Laursen TM (2011). Life expectancy among persons with schizophrenia or bipolar affective disorder. *Schizophrenia Research* **131**, 101–104.

Leboyer M, Soreca I, Scott J, Frye M, Henry C, Tamouza R, Kupfer DJ (2012). Can bipolar disorder be viewed as a multi-system inflammatory disease? *Journal of Affective Disorders* **141**, 1–10.

Lerer B, Moore N, Meyendorff E, Cho SR, Gershon S (1987). Carbamazepine vs. lithium in mania: double-blind study. *Journal of Clinical Psychiatry* **48**, 89–93.

Leucht S, Cipriani A, Spineli L, Mavridis D, Orey D, Richter F, Samara M, Barbui C, Engel RR, Geddes JR, Kissling W, Stapf MP, Lässig B, Salanti G, Davis JM (2013). Comparative efficacy and tolerability of 15 antipsychotic drugs in schizophrenia: a multiple-treatments meta-analysis. *Lancet* **382**, 951–962.

Li H, Ma C, Wang G, Zhu X, Peng M, Gu N (2008). Response and remission rates in Chinese patients with bipolar mania treated for 4 weeks with either quetiapine or lithium: randomized and double-blind study. ClinicalTrials.gov: NCT00448578. *Current Medical Research and Opinion* **24**, 1–10.

Lu G, Ades AE (2004). Combination of direct and indirect evidence in mixed treatment comparisons. *Statistics in Medicine* **23**, 3105–3124.

Machado-Vieira R, Manji HK, Zarate Jr. CA (2009). The role of lithium in the treatment of bipolar disorder: convergent evidence for neurotrophic effects as a unifying hypothesis. *Bipolar Disorders* **11** (Suppl. 2), 92–109.

Mauri D, Polyzos NP, Salanti G, Pavlidis N, Ioannidis JP (2008). Multiple-treatments meta-analysis of chemotherapy and targeted therapies in advanced breast cancer. *Journal of the National Cancer Institute* **100**, 1780–1791.

McElroy SL, Martens BE, Creech RS, Welge JA, Jefferson L, Guerdjikova AI, Keck PE Jr. (2010*a*). Randomized, double-blind, placebo-controlled study of divalproex extended release loading monotherapy in ambulatory bipolar spectrum disorder patients with moderate-to-severe hypomania or mild mania. *Journal of Clinical Psychiatry* **71**, 557–565.

McElroy SL, Martens BE, Winstanley EL, Creech R, Malhotra S, Keck Jr. PE (2010*b*). Placebo-controlled study of quetiapine monotherapy in ambulatory bipolar spectrum disorder with moderate-to-severe hypomania or mild mania. *Journal of Affective Disorders* **124**, 157–163.

Merikangas KR, Akiskal HS, Angst J, Greenberg PE, Hirschfeld RM, Petukhova M, Kessler RC (2007). Lifetime and 12-month prevalence of bipolar spectrum disorder in the national comorbidity survey replication. *Achieves of General Psychiatry* **64**, 543–552.

Moreno SG, Sutton AJ, Ades AE, Stanley TD, Abrams KR, Peters JL, Cooper NJ (2009). Assessment of regression-based methods to adjust for publication bias through a comprehensive simulation study. *BMC Medical Research Methodology* **9**, 2.

Mosolov SN, Kostiukova EG, Kapiletti SG, Belen'kaia DV, Shafarenko AA, Fedotov DD (2009). Open randomized comparative twelve-week study of lithium and valproate in manic episode [in Russian]. *Zhurnal nevrologii i psikhiatrii imeni S.S. Korsakova* **109**, 47–52.

Mizuno Y, Suzuki T, Nakagawa A, Yoshida K, Mimura M, Fleischhacker WW, Uchida H (2014). Pharmacological strategies to counteract antipsychotic-induced weight gain and metabolic adverse effects in schizophrenia: systematic review and meta-analysis. *Schizophrenia Bulletin*. Published online: 17 March 2014. PMID: 24636967.

Novartis Clinical Trial Results Database (2007). Randomized, double-blind, placebo-controlled multicenter study to evaluate efficacy and tolerability of licabazepine 1000–2000 mg/d in the treatment of manic episodes of bipolar I disorder over 3 weeks. Study ID Number: CLIC477D2301. ClinicalTrials.gov: NCT00099229 (data on file).

Ortega Soto HA (1993). Carbamazepine vs haloperidol in treatment of manic episodes: controlled clinical trial. *(Spanish)* Salud Mental **16**, 44–50.

Ösby U, Brandt L, Correia N, Ekbom A, Sparén P (2001). Excess mortality in bipolar and unipolar disorder in Sweden. *Achieves of General Psychiatry* **58**, 844–850.

Pfizer Clinical Trial Results Database (2008). Multicenter, randomized, double-blind, parallel group study, comparing the efficacy and tolerability of ziprasidone (Zeldox, Geodon) vs olanzapine (Zyprexa) in the treatment and maintenance of response in patients with acute mania. Study ID Number: A1281147. ClinicalTrials.gov: NCT00329108 (data on file).

Salanti G, Ades AE, Ioannidis JP (2011). Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. *Journal of Clinical Epidemiology* **64**, 163–171.

Salanti G, Higgins JPT, Ades AE, Ioannidis JPA (2008). Evaluation of networks of randomized trials. *Statistical Methods in Medical Research* **17**, 279–301.

Salanti G, Marinho V, Higgins JPT (2009). Case study of multiple-treatments meta-analysis demonstrates that covariates should be considered. *Journal of Clinical Epidemiology* **62**, 857–864.

Shafti SS (2010). Olanzapine vs. lithium in management of acute mania. *Journal of Affective Disorders* **122**, 273–276.

Shafti SS, Shahveisi B (2008). Comparison between lithium and valproate in the treatment of acute mania. *Journal of Clinical Psychopharmacology* **28**, 718–720.

Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society* **64**, 583–639.

Spitzer RL, Endicott J, Robins E (1978). Research diagnostic criteria: rationale and reliability. *Archives of General Psychiatry* **35**, 773–782.

Tondo L, Lepri B, Baldessarini RJ (2007). Risks of suicidal ideation, attempts and suicides among 2826 men and women with types I and II bipolar, and recurrent major depressive disorders. *Acta Psychiatrica Scandinavica* **116**, 419–428.

Vasudev K, Goswami U, Kohli K (2000). Carbamazepine and valproate monotherapy: feasibility, relative safety and efficacy, and therapeutic drug monitoring in manic disorder. *Psychopharmacology* **150**, 15–23.

Vernon AC, Natesan S, Crum WR, Cooper JD, Modo M, Williams SC, Kapur S (2012). Contrasting effects of haloperidol and lithium on rodent brain structure: magnetic resonance imaging study with postmortem confirmation. *Biological Psychiatry* **71**, 855–863.

Vieta E, Cruz N (2012). Head to head comparisons as an alternative to placebo-controlled trials. *European Neuropsychopharmacology* **22**, 800–803.

Vieta E, Langosch JM, Figueira ML, Souery D, Blasco-Colmenares E, Medina E, Moreno-Manzanaro M, Gonzalez MA, Bellivier F (2013). Clinical management and burden of bipolar disorder: results from a multinational longitudinal study (WAVE-bd). *International Journal of Neuropsychopharmacology* **16**, 1719–1732.

Woo G, Tomlinson G, Nishikawa Y, Kowgier M, Sherman M, Wong DK, Pham B, Ungar WJ, Einarson TR, Heathcote EJ, Krahn M (2010). Tenofovir and entecavir are the most effective antiviral agents for chronic hepatitis B: systematic review and Bayesian meta-analysis. *Gastroenterology* **139**, 1218–1229.

Yildiz A, Guleryuz S, Ankerst DP, Ongür D, Renshaw PF (2008). Protein kinase C inhibition in the treatment of mania: double-blind, placebo-controlled trial of tamoxifen. *Archives of General Psychiatry* **65**, 255–263.

Yildiz A, Undurraga J, Vieta E, Popovic D, Wooderson S, Young A (in press *a*). Strategies for improving randomized trial evidence in treatment of bipolar disorder. In *Bipolar Book: History, Neurobiology, and Treatment* (ed. A. Yildiz, C. Nemeroff and P. Ruiz). Oxford University Press: New York.

Yildiz A, Vieta E, Correll CU, Nikodem M, Baldessarini RJ (in press *b*). Critical issues on use of network meta-analysis in psychiatry. *Harvard Review of Psychiatry*.

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

Yildiz A, Vieta E, Tohen M, Baldessarini RJ (2011*b*). Factors modifying drug and placebo responses in randomized trials for acute mania. *International Journal of Neuropsychopharmacology* **14**, 863–875.

Young RC, Biggs JT, Ziegler VE, Meyer DA (1978). Rating scale for mania: reliability, validity and sensitivity. *British Journal of Psychiatry* **133**, 429–435.

Zarate CA Jr., Singh JB, Carlson PJ, Quiroz J, Jolkovsky L, Luckenbaugh DA, Manji HK (2007). Efficacy of a protein kinase C inhibitor (tamoxifen) in the treatment of acute mania: a pilot study. *Bipolar Disorders* **9**, 561–570.