International Journal of Neuropsychopharmacology (2011), **14**, 863–875. © CINP 2011
doi:10.1017/S1461145710001641

ARTICLE

Factors modifying drug and placebo responses in randomized trials for bipolar mania

Ayşegül Yildiz[1,2], Eduard Vieta[2,3], Mauricio Tohen[2,4] and Ross J. Baldessarini[2]

[1] *Department of Psychiatry, Dokuz Eylül University, Izmir, Turkey*
[2] *Department of Psychiatry, Harvard Medical School and International Consortium for Bipolar Disorder Research, McLean Division of Massachusetts General Hospital, Boston, MA, USA*
[3] *Bipolar Disorders Program, Institute of Clinical Neuroscience, Hospital Clinic, University of Barcelona, IDIBAPS, CIBERSAM, Barcelona, Spain*
[4] *University of Texas, Health Science Center at San Antonio, TX, USA*

Abstract

Randomized placebo-controlled trials (RCTs) are standard for assessing efficacy and safety of treatments. We pursued preliminary indications that some factors are associated differentially with responses to placebo or drugs in RCTs for bipolar mania. We meta-analysed data from RCTs to assess influences of study-site count, subjects' age, sex distribution, diagnostic subgroups, clinical features, trial-completion rates, and publication year on mean difference (MD) in mania ratings between intake and final assessments. In 38 RCTs involving 3812 placebo-treated and 6988 drug-treated patients, symptomatic improvement was similar in placebo arms of trials of effective (6.77, 95% CI 5.77–7.76) and ineffective (7.61, 95% CI 5.47–8.75) drugs. Lesser placebo responses (MD) and greater drug–placebo differences (Hedges' *g*) were associated with fewer study sites, younger patients' age, and male sex. More patients with initial psychotic features and more trial completion in drug arms were associated with greater drug-associated improvement (MD) and drug–placebo contrast (Hedges' *g*), whereas more mixed-state diagnoses decreased both measures. Identifying modifying factors can support more efficient and cost-effective designs of therapeutic trials. In trials for mania, fewer sites may limit placebo response and enhance drug–placebo contrasts.

Received 21 October 2010; Reviewed 28 November 2010; Revised 4 December 2010; Accepted 14 December 2010; First published online 7 February 2011

Key words: Bipolar I disorder, mania, meta-analysis, placebo response, randomized controlled trials.

Introduction

Randomized, placebo-controlled, trials (RCTs) are standard for testing efficacy and safety of medical treatments. However, ethical, clinical, and practical concerns arise in the use of placebos, particularly when they interrupt ongoing active treatment or make subject recruitment into experimental trials difficult (Charney *et al*. 2002; Ehni & Wiesing, 2008; Kotzalidis *et al*. 2008; Vieta & Carne, 2005). Nevertheless, use of placebos as controls in experimental therapeutic trials continues to be accepted and even required by some regulatory agencies including the US Food and Drug Administration (FDA). Alternative trial designs, such as comparisons of experimental to established treatments usually require large numbers of subjects in attempts to detect small differences. Such comparison designs also risk random and potentially misleading findings of 'non-significant differences'. Spurious results are especially likely in trials with limited quality control in subject selection and clinical ratings, particularly in disorders in which non-specific or 'placebo' response rates typically are moderately high, including mania (Charney *et al*. 2002; Kotzalidis *et al*. 2008). With 'non-inferiority' trials, there is no expected difference in efficacy between test treatments (null hypothesis), and sample size cannot be guided by statistical power analysis, so that non-rejection of the null hypothesis may be inconclusive, even with large samples.

Address for correspondence: Dr A. Yildiz, Department of Psychiatry, Dokuz Eylül University Inciralti, Izmir, Turkey.
Tel.: +90 232 412 4160 *Fax*: +90 232 412 2198
Email: agul_yildiz@hotmail.com
Preliminary findings of this paper were presented at the Breaking News Symposia, 23rd Congress of the European College of Neuropsychopharmacology, Amsterdam, Netherlands, 30 August 2010.

As therapeutic innovation in psychiatry increases, more controlled treatment trials are needed, resulting in greater pressure to recruit patient-subjects quickly and efficiently in large samples that may vary in quality control. Accordingly, the costs of conducting such trials, as well as the major challenges of maintaining scientific quality control are increasing, with the risk of loss of sensitivity and a paradoxical increase in costs as well as increased assessment variance by requiring very large samples. Quality control is particularly difficult in increasingly common international trials carried out in geographically separate and heterogeneous sites. These circumstances provide increasing challenges to limiting variance in subject recruitment, diagnosis, and ongoing assessment, as well as limiting typically high dropout rates with loss of data. Such problems appear to be reflected in indications that, even in severe psychiatric disorders including schizophrenia and mania, clinical change during placebo treatment has been rising in recent years (Kemp *et al*. 2010; Sysko & Walsh, 2007; Vieta & Cruz, 2008). A recent collaborative conference involving the International Society for CNS Clinical Trials and Methodology (ISCTM) and the International Society for CNS Drug Development (ISCDD) reviewed factors associated with placebo-associated responses in trials (Kemp *et al*. 2010). Factors considered included: longer duration of current acute illness episodes, effects of 'rescue' medications such as sedatives, variability in assessments, lack of optimized drug-dosing, as well as inconsistent medication compliance and protocol adherence (Kemp *et al*. 2010). Pardoxically, imperfect control of variance and typically modest effect sizes (drug–placebo contrasts) in the conduct of trials encourages larger samples derived from more collaborating sites, but at greater costs and with greater risk of increasing variance with the risk of regression-to-the-mean in both placebo and drug arms of RCTs (Kemp *et al*. 2010). Such circumstances have uncertain impact on the quest for larger or more stable drug–placebo contrasts. It is also an ethical truism that larger trials require exposing more patients to inactive placebo treatment.

We hypothesize that identifying characteristics of research subjects and of trial designs associated with smaller placebo responses and larger drug–placebo contrasts, as well as identifying subgroups that may be particularly treatment-responsive or -tolerant (or the opposite), will have value in optimizing trial design. Such considerations should limit the complexity of trials and decrease their expense, duration, and the number of persons exposed to potentially inactive treatments or placebos, as well as facilitating subject recruitment and retention (Charney, 2000; Charney *et al*. 2002; Kent & Hayward, 2007; Kotzalidis *et al*. 2008). Moreover, placebo-response rates can influence observed drug–placebo differences in trials, and inferences about efficacy of particular treatments relevant to rational or evidence-based clinical treatment (Kent & Hayward, 2007).

In response to the many challenges of including placebo-controls in treatment trials, we conducted a meta-analysis of RCTs of single treatments in acute mania. In this analysis, we identified clinical and trial-design factors that may be associated with responses to randomized placebo or drug treatments. In particular, we sought factors that might enhance drug–placebo contrasts, or at least limit placebo response, as a contribution to improving the design of experimental therapeutic trials.

Methods

Data sources

We conducted meta-analyses of all available randomized, placebo-controlled monotherapy trials for acute mania in patients with manic or mixed phases of type I bipolar disorder (BPD). We identified reports through computerized searches of the PubMed/Medline; ClinicalTrials.gov; Cochrane Central Register of Controlled Trials; Controlled-trials.com; and EMBASE/Excerpta Medica databases (search terms: 'bipolar', 'mania', 'placebo', 'trial', and 'names' of individual anticonvulsant, antipsychotic, or other drugs tested in mania, up to 12 January 2010). We extended the search by reviewing bibliographies of reports identified, as well as reports of presentations at meetings of the American Psychiatric Association (APA), American and European Colleges of Neuropsychopharmacology, and the International Conference on Bipolar Disorder. We also consulted study investigators and representatives and scientists of pharmaceutical companies that produce antimanic agents for leads to other trial reports and for access to data missing from identified reports.

Study selection and data extraction

We included only trials involving a placebo and one or more active treatments or doses, assigned randomly, to test efficacy among consenting adults diagnosed with acute mania or mixed states of type I BPD by definitions in the APA Diagnostic and Statistical Manual (DSM-III to DSM-IV-TR) or Research Diagnostic Criteria (RDC) (APA, 2000; Spitzer *et al*. 1978). We excluded trials involving patients diagnosed with

unspecified BPD (NOS), BPD II, or schizoaffective disorders, and trials permitting use of psychotropic agents other than modest doses of sedatives or hypnotics (benzodiazepines or chloral hydrate). For quality control, required trial data were extracted by two scientists (A. Yildiz and S. Özer).

Information extracted and evaluated included study-site counts, sample sizes in each trial arm, based on intent-to-treat (ITT) principles (i.e. patients with at least one post-randomization assessment), mean age of subjects, sex distribution, initial diagnosis (mania *vs.* manic-depressive mixed state), presence of psychotic features, baseline illness severity ratings (percent of maximum attainable mania scale scores), nominal trial duration, rates of trial completion, source of support, and year of reporting. The primary outcome of interest was mean change in mania ratings between intake and last observation, expressed as mean difference (MD), considered separately for drug and placebo arms of each trial. We also considered drug–placebo differences in outcomes, expressed as Hedges' *g*, based on standardized MD between changes in mania ratings with drug *vs.* placebo. The main secondary outcome measure was rate of response (proportion of subjects with $\geqslant 50\%$ improvement in mania ratings) with drug and placebo (Tohen *et al*. 2009). Ratings of clinical change were based on the 11-item Young Mania Rating Scale (YMRS; maximum score = 60) used in 80.4% of trials, or the 11-item Mania Rating Scale (MRS; maximum score = 52), which are similar in scoring characteristics and ability to detect treatment effects (Endicott & Spitzer, 1978; Vieta, 2003; Young *et al*. 1978). We examined the primary outcome measure (MD for placebo or drugs; Hedges' *g* for drug–placebo contrasts) for association with site counts, mean age, percent men, initial diagnosis (% manic *vs.* mixed), proportion (%) with psychotic features, trial completion rates (%), and publication year.

Meta-analyses

We combined data across trials by meta-analytial modelling of mean changes in mania ratings during randomized treatment with placebo or drugs. When standard deviations (S.D.) for changes in score were not reported, we estimated them by a standard formula (Whitley & Ball, 2002). Computed values are shown with 95% confidence intervals (CIs). Since the studies analysed involved different investigators and varied environments, we used random-effects meta-analytical modelling, with or without evidence of inter-study heterogeneity as evaluated with preliminary *Q* tests (Borenstein *et al*. 2009).

We examined potential effects of the previously stated parameters on the primary, continuous outcome measures of MD for drug arms and placebo arms; and Hedges' *g* for drug–placebo differences by using unrestricted maximum-likelihood, mixed-effects meta-regression modelling to generate a slope function (β coefficient) that reflects the influence of each factor on the corresponding outcome measure (Berkey *et al*. 1995). MD for change in mania ratings between baseline and end-point in placebo arms and drug arms of trials are considered placebo effect and drug effect, respectively; Hedges' *g* indicates adjusted mean drug–placebo difference in change in mania ratings (treatment effect). To limit risk of false-positive (type I) errors, we corrected $\alpha = 0.05$ by dividing it with the number of moderator variables ($n = 7$), requiring $p \leqslant 0.007$ (0.05/7) to establish statistical significance.

Trials with negative results (statistical non-separation of drug *vs.* placebo) are less likely to be published than those with statistically significant drug–placebo contrasts. To limit such reporting bias, we performed a comprehensive literature search and included data from unpublished reports. Then, we applied the funnel-plot method (pooled MD *vs.* its standard error) to evaluate presence of reporting bias (Sterne & Egger, 2001). We also estimated Orwin's fail-safe *N* values (number of additional hypothetical studies with zero-effect required to make the pooled effects derived from meta-analysis trivial, i.e. effect size $\leqslant 0.10$); larger fail-safe *N* indicates less likelihood of publication bias (Orwin, 1983). Finally, we employed Duval & Tweedie's trim-and-fill approach to calculate a best-estimate of unbiased effect size by removing extremely small studies from the funnel plot individually until the plot became symmetrical about the (new) pooled effect size (Duval & Tweedie, 2000). We used Biostat commercial software for meta-analyses (Comprehensive Meta-Analysis, version 2.2; BioStat Inc., USA).

Results

Characteristics of trials and subjects

We identified 38 studies yielding 56 comparisons (13 with negative results) of 17 active drugs *vs.* placebo, involving a total of 5929 patients randomized to a placebo [Table 1]. Corrected for duplicate counting of placebo-arm patients in multi-arm trials, a total of $n = 3812$ patient-subjects were randomized to placebo arms of the trials considered, and had at least one post-randomization assessment during protocol-guided treatment (ITT samples). Most studies (34/38 = 89.5%)

Table 1. Characteristics of randomized, placebo-controlled monotherapy studies in mania

	Randomized		ITT sample						Completers (%)		Baseline mania (% max)		Mania improvement (%)		
Sites (n^a)	Total	PBO	Total	PBO	Men (%)	Age, yr (±S.D.)	Psychotic (%)	Mixed (%)	Rx	PBO	Rx	PBO	Rx	PBO	Source
Aripiprazole															
56	270	134	259	130	47.0	41.0±11.7	–	38.7	39.7	40.3	46.4	47.1	38.8	35.8	El Mallakh *et al*. (2010)
56	265	134	257	130	49.1	40.0±11.5	–	39.2	42.7	40.3	46.6	47.1	35.8	35.8	El Mallakh *et al*. (2010)
38	262	132	243	120	43.5	40.5±12.2	25.7	32.8	41.5	21.2	47.0	49.5	29.1	11.4	Keck *et al*. (2003*a*)
42	330	165	317	163	51.6	39.7±11.0	21.8	38.2	44.2	47.3	47.5	48.2	44.4	31.2	Keck *et al*. (2009)
29	272	135	268	132	48.5	38.8±11.5	13.1	41.5	54.7	51.9	48.0	47.4	43.4	25.3	Sachs *et al*. (2006)
59	320	153	318	152	44.7	40.4±11.9	10.6	19.4	75.4	71.2	46.7	47.2	42.8	34.3	Young *et al*. (2009)
Asenapine															
70	283	98	277	94	49.5	38.6±12.4	36.8	31.1	67.0	58.2	49.0	47.2	48.3	38.2	McIntyre *et al*. (2009*a*)
64	298	104	292	103	55.7	39.1±12.0	29.1	30.2	67.0	58.2	47.2	48.3	46.3	25.5	McIntyre *et al*. (2009*b*)
Carbamazepine															
24	204	103	192	98	52.5	38.0±11.0	–	52.9	49.5	44.7	44.3	45.5	32.8	18.9	Weisler *et al*. (2004)
25	239	117	235	115	70.3	37.0±11.0	–	20.9	65.6	54.7	47.4	46.6	53.0	25.5	Weisler *et al*. (2005)
Cariprazine															
29	238	120	236	118	63.3	38.5	–	–	63.6	61.9	51.0	50.3	43.5	23.8	Knesevich *et al*. (2009)
Haloperidol															
49	200	101	198	100	36.9	42.9	41.9	0.00	77.8	60.4	53.8	55.2	48.6	25.1	McIntyre *et al*. (2005)
20	284	140	282	138	52.5	39.0	32.0	0.00	89.0	85.0	53.5	52.5	47.0	29.8	Smulevich *et al*. (2005)
33	260	88	258	88	57.5	37.9±13.0	32.7	15.4	71.3	50.0	59.0	60.2	51.9	19.5	Vieta *et al*. (2010*b*)
59	318	153	313	152	45.0	40.9±12.0	−9.40	17.9	73.3	71.2	46.0	47.2	46.5	34.3	Young *et al*. (2009)
16	–	–	117	97	–	–	–	–	–	–	–	–	–	6.8	Katagiri *et al*. (2010)
Lamotrigine															
47	180	95	179	95	52.8	37.2±10.7	41.7	22.2	62.4	64.2	50.8	49.8	35.2	36.7	Goldsmith *et al*. (2003)
Licarbazepine															
28	324	163	313	158	53.4	40.0±11.6	36.7	36.7	63.4	68.7	45.8	45.7	33.5	30.3	Novartis (2007)
Lithium															
9	110	74	107	72	61.8	39.1±10.6	22.7	56.0	38.9	36.5	52.0	54.1	34.3	14.4	Bowden *et al*. (1994)
38	195	97	193	95	58.5	40.1	28.5	0.00	85.7	69.1	55.5	56.7	45.6	19.7	Bowden *et al*. (2005)
47	131	95	131	95	51.9	39.4±12.3	39.7	22.9	44.4	64.2	50.4	49.8	40.8	36.7	Goldsmith *et al*. (2003)
42	325	165	318	163	52.3	39.7±10.9	24.6	37.5	48.8	47.3	49.0	48.2	40.9	31.2	Keck *et al*. (2009)
40	224	111	224	111	46.0	42.5±13.5	29.5	14.7	74.3	73.9	50.2	50.0	42.9	25.7	Kushner *et al*. (2006)
40	226	112	226	112	38.1	41.5±11.5	19.5	10.2	81.6	86.6	51.2	52.8	45.0	26.5	Kushner *et al*. (2006)
Olanzapine															
70	303	98	297	94	54.5	38.3±11.7	33.7	31.7	78.5	58.2	49.5	47.2	54.2	38.2	McIntyre *et al*. (2009*a*)
64	295	104	291	103	56.3	39.8±11.7	29.2	32.3	79.6	61.5	47.7	48.3	48.6	25.5	McIntyre *et al*. (2009*b*)
16	139	69	136	66	51.8	39.5±11.0	53.2	17.3	61.4	34.8	47.8	46.1	35.8	17.6	Tohen *et al*. (1999)
24	115	60	110	56	49.6	38.7±10.4	55.7	42.6	61.8	41.7	47.9	49.1	51.4	27.6	Tohen *et al*. (2000)
42	320	105	300	99	48.3	40.1±12.4	0.00	29.0	74.0	73.3	39.7	39.2	39.5	35.5	Tohen *et al*. (2008)
16	–	–	201	97	–	–	–	–	–	–	–	–	12.6	6.80	Katagiri *et al*. (2010)

Paliperidone														
44	236	121	235	121	55.7	39.8 ± 11.2	22.1	33.2	65.2	58.7	47.0	48.0	49.3	9.90	Berwaerts *et al*. (2009)
44	233	121	233	121	53.6	39.5 ± 11.3	25.3	37.3	63.4	58.7	47.7	48.0	33.6	9.90	Berwaerts *et al*. (2009)
44	240	121	239	121	54.4	39.4 ± 11.2	23.0	32.6	58.0	58.7	46.5	48.0	41.9	9.90	Berwaerts *et al*. (2009)
52	300	105	294	104	56.1	39.4 ± 10.6	18.0	34.0	82.1	61.9	45.5	44.2	48.4	27.9	Vieta *et al*. (2010*a*)
Quetiapine														
38	204	97	202	95	56.9	36.7	29.7	0.00	90.7	69.1	54.5	56.7	44.7	19.7	Bowden *et al*. (2005)
48	316	161	308	159	–	–	–	–	–	–	–	–	–	–	AstraZeneca (2010)
49	203	101	201	100	36.8	41.7	42.8	0.00	64.7	60.4	56.7	55.2	36.1	25.1	McIntyre *et al*. (2005)
52	298	105	296	104	57.4	38.5 ± 10.5	22.6	38.5	78.8	61.9	46.0	44.2	42.4	27.9	Vieta *et al*. (2010*a*)
Risperidone														
20	259	125	246	119	56.8	38.8 ± 12.1	42.5	0.00	56.0	41.6	48.5	48.7	36.4	16.4	Hirschfeld *et al*. (2004)
8	291	145	286	142	62.1	35.1 ± 17.1	59.3	4.50	89.0	70.8	61.8	62.5	61.2	28.0	Khanna *et al*. (2005)
20	294	140	291	138	52.4	40.4 ± 13.1	32.7	0.00	90.3	85.0	52.2	52.5	44.4	29.8	Smulevich *et al*. (2005)
Tamoxifen														
1	66	31	58	26	48.5	34.8 ± 12.3	66.7	6.10	82.9	67.7	64.3	62.0	43.0	−12.9	Yildiz *et al*. (2008)
1	16	8	16	8	87.5	35.4 ± 7.80	50.0	31.3	50.0	62.5	50.5	40.5	60.4	−19.2	Zarate *et al*. (2007)
Topiramate														
2	314	100	308	99	46.8	38.7 ± 11.0	27.3	53.2	58.9	72.0	48.7	47.2	27.6	27.2	Kushner *et al*. (2006)
2	215	106	213	106	63.4	40.5 ± 11.5	33.3	34.3	56.0	73.6	50.7	49.2	16.8	21.7	Kushner *et al*. (2006)
40	331	111	326	111	50.3	42.3 ± 13.7	29.4	15.0	70.0	73.9	50.8	50.0	19.7	25.7	Kushner *et al*. (2006)
40	228	112	227	112	37.9	40.5 ± 12.0	25.1	11.5	87.1	86.6	51.3	52.8	26.6	26.5	Kushner *et al*. (2006)
Valproate														
9	143	74	139	72	54.5	39.7 ± 11.4	24.5	62.4	52.2	36.5	52.2	54.1	34.0	14.4	Bowden *et al*. (1994)
99	377	185	364	177	57.4	37.6 ± 10.5	20.6	43.7	57.8	51.9	51.2	51.2	44.7	33.8	Bowden *et al*. (2006)
29	225	78	222	78	51.4	39.1 ± 11.4	40.1	31.5	17.0	17.9	63.3	63.5	30.7	25.8	Hirschfeld *et al*. (2010)[b]
9	36	19	36	19	72.2	37.2 ± 13.3	–	0.00	23.5	21.1	47.0	47.7	40.5	−0.60	Pope *et al*. (1991)
42	306	105	285	99	49.1	39.9 ± 12.3	0.00	29.5	75.1	73.3	39.8	39.2	34.3	31.5	Tohen *et al*. (2008)
Verapamil														
1	32	15	20	12	59.4	36.2 ± 10.6	71.9	6.30	17.6	40.0	55.8	50.0	3.80	5.00	Janicak *et al*. (1998)
Ziprasidone														
24	210	70	197	66	54.3	38.0 ± 10.5	45.7	35.5	53.6	44.3	51.9	51.3	45.9	29.2	Keck *et al*. (2003*b*)
23	206	66	202	65	50.7	39.0 ± 11.6	31.2	40.8	60.7	54.5	50.4	50.8	42.4	21.3	Potkin *et al*. (2005)
33	266	88	264	88	59.4	37.7 ± 12.7	33.1	18.4	66.9	50.0	56.9	60.2	35.2	19.5	Vieta *et al*. (2010*b*)

ITT, Intent-to-treat; MRS, Mania Rating Scale; PBO, placebo; Rx, study drug; S.D., standard deviation; YMRS, Young Mania Rating Scale.

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

[b] A negative trial of divalproex extended release (ER) against placebo.

were multi-centre collaborations with a large average (±S.D.) site count [29.7±18.9 (range 1–70) sites/trial], and the same proportion were sponsored by manufacturers of drugs tested.

Outcomes with effective vs. *ineffective drugs*

Improvements in mania ratings in placebo arms of the 56 trials were 6.92 (95% CI 6.02–7.82) scale points. Effect-sizes in placebo arms were similar in the 48 trials of 13 drugs with significant overall drug–placebo contrasts (aripiprazole, asenapine, carbamazepine, cariprazine, haloperidol, lithium, olanzapine, paliperidone, quetiapine, risperidone, tamoxifen, valproate, ziprasidone: pooled MD = 6.77, 95% CI 5.77–7.76), and seven trials of four drugs with non-significant overall drug–placebo contrasts (lamotrigine, licarbazepine, topiramate, verapamil: pooled MD = 7.61; 95% CI 5.47–9.75; Table 1). Pooled effects in drug arms of the trials involving effective drugs (MD = 12.7, 95% CI 12.0–13.4) were nearly twice as large as the placebo effects noted above. Improvements in mania ratings in drug arms of the ineffective treatments were 7.31 (95% CI 5.35–9.27), indicating 1.7-fold superior improvement in successful *vs.* unsuccessful trials.

Also based on meta-analytical pooling, the overall, secondary outcome measure of responder rate (improving ⩾50%) was substantial over 3 wk, at 30.8% (95% CI 28.6–33.2) for all placebo-treated patients. For trials involving effective drugs, pooled responder rates were 48.3% (95% CI 46.2–50.4) in drug arms and 30.7% (95% CI 28.1–33.3) in placebo arms – a 1.6-fold difference. The pooled responder rate in drug arms of ineffective agents was much lower, as expected, at 30.8% (95% CI 26.2–35.9), and close to the corresponding placebo-associated responder rate of 31.6% (95% CI 26.6–37.2). Notably, however, placebo responder rates in trials of effective and ineffective drugs (agents with non-significant overall drug–placebo contrasts) were nearly identical (30.7%, 95% CI 28.1–33.3 *vs.* 31.6%, 95% CI 26.6–37.2, respectively).

These findings indicate that most of the contrast in outcomes of trials of apparently effective *vs.* ineffective agents was due to marked differences in drug response, and that differences in placebo responses were minor.

Factors associated with placebo responses

Factors of interest were tested for association with placebo effect based on changes in mania ratings (MD) in all placebo-treatment trial arms for which data were available, and included only once in the meta-regression modelling (Tables 2 and 3). For comparison, we considered the same factors for association with drug effect (also based on MD for change in mania ratings) as well as treatment effect (based on Hedges' g for effects of drug *vs.* placebo). For testing those associations of candidate factors with drug effect and treatment effect we considered only the 13 drugs found significantly more effective than placebo in a meta-analysis of individual drugs for bipolar mania (Yildiz *et al*. 2010), in order to avoid potential confounding by drug ineffectiveness, which would influence the drug-associated benefit as well as observed drug and placebo differences. Regression modelling of drug effect involved only patients in drug arms, whereas modelling of treatment effect involved patients in both drug and placebo arms. However, reported proportions of patients given active drugs who completed each trial were also tested for association with treatment effect. Since most of the studies did not report subjects' ages in each treatment arm separately, we used average age for entire study samples in corresponding regression models. Since seven preselected covariates were tested by meta-regression, statistical significance was set at two-tailed $\alpha = 0.007$.

Meta-regression modelling suggested that a lower number of collaborating study sites was strongly associated with lesser placebo-induced improvement of mania ratings (MD), and greater drug–placebo contrasts (Hedges' g) (see Fig. 1). Of note, more study sites corresponded to larger patient samples, and with larger placebo effects (in 38 trials; $\beta = +0.06$, 95% CI 0.04–0.08, $z = 6.47$, $p < 0.0001$). Notably, however, site counts had no apparent association with drug-associated improvements in mania ratings (MD). Male sex was associated with lesser placebo effects, and greater treatment effects. On the contrary, younger mean age was associated with lesser placebo effects and greater drug-associated improvements, and correspondingly greater drug–placebo contrasts. Diagnostic subgroups (mania with or without psychotic features or manic *vs.* mixed states) had no apparent influence on placebo effects overall, although outcomes were rarely reported separately for such subgroups. However, presence of psychotic symptoms increased both observed effects of treatment over placebo and drug-associated benefit, whereas the proportion of cases with mixed state decreased both measures (Tables 2 and 3).

More trial completers (lower dropout rates) in placebo arms was not related to improvements with placebo, but higher proportion of trial completion in drug arms was associated with both greater drug-associated benefit and drug–placebo contrast or treatment effect (Tables 2 and 3). No outcome was associated with publication year, and there were too few

Table 2. Association of factors with changes in mania ratings with placebo or drug (MD), and drug *vs.* placebo contrasts (Hedges' *g*), in randomized, monotherapy trials for antimanic effects

	Placebo effect (MD)					Drug effect (MD)					Drug-placebo contrast (Hedges's *g*)				
Factors	*N*	Slope (%, 95% CI)	*z*	*Q* model	*p*	*N*	Slope (%, 95% CI)	*z*	*Q* model	*p*	*N*	Slope (%, 95% CI)	*z*	*Q* model	*p*
More collaborating sites (*n*[a])	38	+**0.11** (+0.06 to +0.15)	4.67	21.8	<0.0001	48	−**0.02** (−0.06 to +0.03)	−0.80	0.64	0.425	48	−**0.007** (−0.01 to −0.003)	−3.79	14.4	0.0002
Age (yr[b])	36	+**0.92** (+0.33 to +1.15)	3.07	9.44	0.002	46	−**0.69** (−1.15 to −0.24)	−2.99	8.96	0.003	46	−**0.09** (−0.13 to −0.04)	−4.03	16.2	0.00006
More men (%)	35	−**0.18** (−0.29 to −0.06)	−3.04	9.23	0.002	46	+**0.08** (+0.01 to +0.16)	2.28	5.19	0.023[d]	46	+**0.02** (+0.007 to +0.03)	3.49	12.2	0.0005
More with psychotic features (%)	31	−**0.08** (−0.16 to −0.009)	−2.20	4.83	0.028[d]	40	+**0.10** (+0.06 to +0.15)	4.22	17.8	0.00002	40	+**0.85** (+0.43 to +1.28)	3.91	15.2	0.00009
More mixed state (%)	35	+**0.03** (−0.04 to +0.10)	0.90	0.81	0.368	45	−**0.07** (−0.12 to −0.03)	−3.26	10.6	0.001	45	−**0.59** (−0.99 to −0.19)	−2.92	8.51	0.004
Trial completers (%)	36	+**0.05** (−0.02 to +0.11)	1.45	2.10	0.147	46	+**0.08** (+0.04 to +0.13)	3.99	15.9	0.00007	46	+**0.41** (−0.01 to +0.83)	1.89	3.58	0.059
Trial completers: drug arms (%)	–	–	–	–	–	46	+**0.08** (+0.04 to +0.13)	3.99	15.9	0.00007	46	+**0.006** (+0.002 to +0.009)	2.86	8.20	0.004
Publication year (without outliers)[c]	36	+**0.37** (+0.22 to +0.52)	4.86	23.6	<0.0001	46	+**0.12** (−0.09 to +0.32)	1.13	1.28	0.258	46	−**0.02** (−0.03 to +0.001)	−1.82	3.32	0.068

MD, Mean difference; *N*, number of comparisons included in corresponding regression models. *Q* model indicates the amount variance or heterogeneity in effect sizes predicted by the model. Effect-size measure is MD for placebo effects and drug effects; and Hedges' *g* for drug–placebo contrast. Only effective drugs were considered for drug effect and drug–placebo contrast, but all available placebo arms were included for placebo effect.

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

[b] Mean age for the study sample was the moderator variable used for all three outcomes.

[c] Meta-regression results after exclusion of outliers (two single-site tamoxifen trials with worsening of mania ratings with placebo).

[d] Indicates non-significance after correction for multiple comparisons (adjusted $p = 0.05/7 = 0.007$).

Table 3. Factors associated with change in mania ratings in placebo arms and drug arms, and drug–placebo contrasts

Factors	Mania improvement with treatment		Drug–placebo contrasts
	Placebo	Drug	
More collaborating sites (n[a])	Increased	No effect	Decreased
More with psychotic features (%)	No effect	Increased	Increased
Age (yr[b])	Increased	Decreased	Decreased
More men (%)	Decreased	No effect	Increased
More mixed-state diagnoses (%)	No effect	Decreased	Decreased
Trial completers (%)	No effect	Increased	No effect
Trial completers: drug arms (%)	–	Increased	Increased
Publication year (without outliers)[c]	Increased	No effect	No effect

Only effective drugs (>placebo) were considered for drug effects and drug–placebo differences (in 40–48 randomized, monotherapy trials), but all available placebo arms were included for placebo effect (31–38 trials; unmatched numbers owing to multiple active treatments in some trials). Improvement is based on mean difference for placebo effects and drug effects, and Hedges' g for drug–placebo contrasts.

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

[b] Mean age for the study sample was the moderator variable used for all three outcomes.

[c] There were no significant associations between publication year and any effect measure. However, exclusion of outliers (two single-site tamoxifen trials with worsening of mania ratings with placebo) indicated rising placebo-related improvements in recent years.

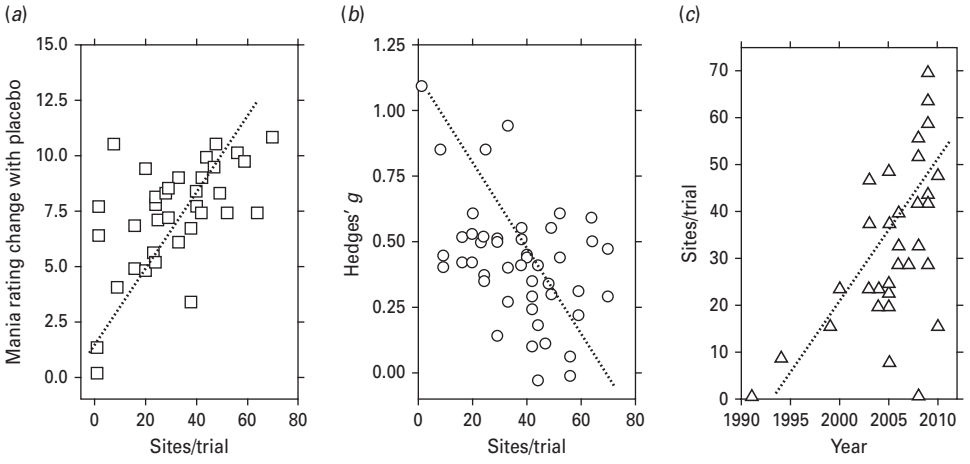

Fig. 1. Improvement in mania ratings *vs.* the number of collaborating sites in randomized, placebo-controlled trials in manic or mixed bipolar disorder patients. (*a*) Improvement with placebo treatment (mean difference between intake and end-point) for all 38 studies (Q model = 21.8, $p < 0.0001$). (*b*) Efficacy (Hedges' g for difference between drug- and placebo-associated changes) for 48 comparisons involving effective drugs only (Q model = 14.4, $p = 0.0002$). (*c*) For illustration, there was an expected, strong secular trend (correlation) of sites/trial and year of reporting ($r = 0.587$, $p < 0.0001$); two single-site tamoxifen trials reporting worsening with placebo treatment were excluded.

trials not supported financially by pharmaceutical manufacturers to test effectively for potential effects of sources of study support. However, exclusion of two small, academic, non-industrial, single-site trials with unusually low placebo effects led to a secular association of rising placebo effects in more recent years, as well as a secular effect of rising numbers of collaborating sites/trials (Fig. 1).

Publication bias

Examination of funnel plots of the primary outcome (MD in placebo arms of trials *vs.* its standard error)

indicated publication bias (Kendall's $\tau = -0.26$, $z = 2.31$, $p = 0.02$). Orwin's fail-safe N was 2793, indicating that a very large number of studies with zero effect would need to be added to produce a trivial outcome, and that is very unlikely to have arisen by omission of overlooked reports. Finally, Duval & Tweedie's trim-and-fill method did not identify any aberrant studies, and the summary effect remained unchanged (MD = 6.92, 95% CI 6.03–7.81).

Discussion

This is the first meta-analytical evaluation of patient-characteristics or other trial-design factors for association with placebo and drug responses and drug–placebo contrasts (effect sizes) in RCTs for mania. After diligent searching to identify all available trials, we considered RCTs involving DSM-III or DSM-IV BPD patients in acute manic or mixed states, randomized to the placebo arms or drug arms of 38 studies with 56 drug–placebo contrasts, and reported on drug efficacy previously (Yildiz *et al*. 2010). These RCTs indicated 13 effective (with statistically superior overall outcomes over placebo: 49 contrasts) and four ineffective (seven contrasts) candidate drugs. Factors associated with lesser placebo effects included: (*a*) fewer collaborating sites, (*b*) younger patients, and (*c*) higher proportion of men in placebo arms (Tables 2 and 3). Since some of these factors are potentially modifiable, their consideration in the design of treatment trials may help to limit sample sizes, placebo exposure, trial failures, financial costs, and generally yield trials that are more efficient and attractive to patients, their families, and physicians.

To help with power and sample-size estimates for future treatment trials in mania, we considered responses following randomization to placebo or drug in 49 trials of the 13 effective antimanic agents. In these successful trials, the meta-analytically pooled MD between initial and final mania ratings was 6.77 scale points for placebo arms ($n = 3199$), and 12.7 for drug arms ($n = 6075$), indicating nearly 2-fold superiority of effective drugs over placebo. In secondary analyses, pooled responder rates (% of patients showing $\geqslant 50\%$ improvement in mania ratings) were 30.7% ($n = 2944$) with placebo and 48.3% ($n = 5827$) with effective drugs, a somewhat lesser (1.6-fold) contrast. This moderate absolute difference of 17.6% in the responder rates with drug *vs.* placebo is similar to an estimate of 18% (41% to drug *vs.* 24% to placebo) in 38 RCTs involving 7323 participants for schizophrenia (Leucht *et al*. 2009). In earlier reviews of eight (Keck *et al*. 2000) and 21 (Sysko & Walsh, 2007) RCTs in acute mania, reported placebo responder rates averaged 23.0% and 31.2%, respectively, or close to our meta-analytical estimate of 30.8% (95% CI 28.6–33.2) across 38 studies with 56 drug–placebo contrasts (Yildiz *et al*. 2010).

An initially observed association between presence of psychotic symptoms and lesser placebo-induced improvements in mania ratings was not significant after correction for multiple comparisons. Yet, concurrent psychotic symptoms in mania seemed to enhance drug-associated benefit as well as drug–placebo contrasts (Tables 2 and 3). Moreover, a higher proportion of cases diagnosed with manic-depressive mixed states resulted in lesser treatment and drug effects, with no apparent effect on placebo-induced improvements in mania ratings (Tables 2 and 3). These observations regarding mixed states suggest a subgroup that is particularly difficult to treat, especially with an antimanic monotherapy, and a subgroup that may not appropriately be combined with mania (Baldessarini *et al*. 2010). Paradoxically, although mixed states often are more severe or clinically challenging than relatively pure mania, their mania rating scores are often lower. This impression was sustained in the present findings, in that baseline mania ratings were lower with higher proportions of mixed-state patients in 52 trials with available data ($\beta = -0.10$, 95% CI -0.13 to -0.06, $z = -5.19$, $p < 0.0001$), with lesser corresponding improvements with drugs in 46 of the trials ($\beta = +0.26$, 95% CI 0.13–0.40, $z = 3.80$, $p = 0.0002$). The observed impact of mixed states on trial outcomes suggests that such patients should be excluded from mania trials, or at least considered separately, ideally with assessments of depressive as well as manic symptoms (Baldessarini *et al*. 2010). In general, we favour consideration of subgroups within trials in order to identify patients who are particularly responsive or tolerant to particular treatments (or not), as well as trials that continue to syndromal remission, and so can better differentiate speed of effect from actual efficacy.

Some observers have challenged the findings of available clinical trials on mania because of variable or low baseline severity of mania ratings among patients included. Initial mania scores in identified trials considered in this study were 39.2–65.4% of maximum attainable scores, and fewer than half of the trials involved baseline mania scores of $\geqslant 50\%$ of maximum attainable scores. As we reported previously, greater initial manic symptom severity did not enhance observed drug–placebo contrasts, but instead, amplified benefit from the drugs selectively (Yildiz *et al*. 2010). This association may reflect a more treatment-sensitive clinical subtype, or, more likely, the technical

effect of greater apparent improvement with higher initial severity scores. That is, the 'law of initial values' (more deviant initial assessments tend to yield greater change with interventions) may well apply to experimental therapeutics (Benjamin, 1963).

When the present findings are considered together, they suggest a profile of manic patients less likely to improve with a placebo, or with greater drug–placebo contrasts, i.e. a greater likelihood of a successful trial outcome. Such characteristics included: (*a*) trials with fewer study sites (a particularly strong effect, presumably reflecting better quality control, and with fewer subjects), (*b*) younger age, (*c*) male sex, and (*d*) psychotic features. In addition, more completion of drug treatments was associated with greater benefit from drug and enhanced drug–placebo contrast.

A particularly notable finding of this study is that improvement during placebo treatment was substantially and selectively smaller with fewer collaborating sites. It challenges the recent trend towards more sites/trials (Fig. 1). Indeed, some single-site trials in mania have reported very slight improvement, or even worsening with placebo treatment (Pope *et al*. 1991; Yildiz *et al*. 2008; Zarate *et al*. 2007). With such rare small trials omitted (Yildiz *et al*. 2008; Zarate *et al*. 2007), we found evidence of a secular trend for increasing placebo effects over time, probably as a manifestation of rising numbers of study sites and patient-subjects. Further meta-regression after exclusion of the rare single-site studies with particularly small placebo effects (Yildiz *et al*. 2008; Zarate *et al*. 2007), verified the observed associations between higher numbers of collaborating study sites and greater placebo-induced improvement in mania ratings (36 trials; $\beta = +0.06$, 95% CI 0.03–0.10, $z = 3.66$, $p = 0.00025$), as well as smaller drug–placebo contrasts (46 trials; $\beta = -0.05$, 95% CI -0.009 to -0.002, $z = -2.86$, $p = 0.004$). Improvements observed during treatment with a placebo may reflect the natural course of a time-limited acute illness like mania, variance in relatively non-specific interventions related to clinical management, regression to mean outcome ratings, as well as an hypothesized 'placebo effect' itself (Klosterhalfen & Enck, 2006). Yet, given the large variance in placebo responses in the trials considered here, despite nominally similar patients and diagnostic and assessment methods, it is likely that other factors are also involved. We propose that that 'noise', arising from local and individual variation in the application of diagnostic and assessment methods and difficulties in maintaining high levels of inter-rater and inter-site reliability across geographically and culturally diverse sites tends to promote regression to average outcome ratings, as a major determinant of placebo response. We suggest that a simple means of mitigating such effects may be to employ fewer and more homogenous collaborating sites with better control over diagnostic and symptom assessments.

Although, the available data were not adequate to support a conclusive assessment, it is also possible that the sources of financial support for trials may affect their outcome. Most of the acute treatment trials in mania or mixed states identified for this study were sponsored by the manufacturers of the drugs tested, and only four were supported by non-profit-making organizations. Meta-analytically pooled improvement in mania ratings with placebo treatment in 34 industry-sponsored trials was 7.71 (95% CI 7.08–8.35), compared to a far lower value of -2.92 (95% CI -5.68 to -0.16) in only four trials with non-industry sponsorship. For drugs with significant overall drug–placebo contrasts, effect sizes in drug arms were comparable in 45 industry-sponsored trials (pooled MD = 12.5, 95% CI 11.8–13.3), and three non-industry-sponsored trials (pooled MD = 15.8, 95% CI 12.1–19.6). The small number of non-industry-sponsored trials obviates a meaningful statistical comparison, the findings suggest that industrial funding was associated with relatively large placebo-related, but not with drug-related benefits in trials in acute mania. In turn, this effect may reflect the relatively larger numbers of sites and subjects in the pharmaceutical-sponsored trials.

Several limitations of this study should be noted. Randomization can protect against many subject selection biases in individual studies but not in meta-regressions (Borenstein *et al*. 2009). As such, identified associations between study-level moderator variables and effect sizes are observational results, and cannot be presumed to be definitive. Further meta-analytical evaluation of identified categorical factors (young/old, male/female, manic/mixed, psychotic/non-psychotic, completer/non-completer) on the outcome measures was not possible since individual studies did not separately report on these potentially important subgroups. Similarly, dosing of rescue medications, exact durations of pre-randomization washout of previous treatments, and days in hospital during trials, as well as such clinical details as the numbers, types, durations, and severity of episodes of previous illness per study arm were not reported in most trials. We therefore, strongly encourage the inclusion of such information, with responses to drug and placebo in subgroups in future studies of acute mania and other major disorders.

In conclusion, meta-analysis and meta-regression modeling based on 56 comparisons of candidate drugs

with placebo controls in 38 RCTs involving diverse BPD I patients in acute episodes of mania or mixed manic-depressive states indicated that a smaller number of collaborating sites was strongly associated with lesser responses to placebo treatment, with little influence on responses to effective drugs. We propose that this aspect of trial design is readily modified, and should improve the efficiency of RCTs, at least for mania. Further research on this, and other indicated factors may lead to more cost-effective and feasible trials, exposure of fewer persons to placebo or potentially ineffective treatments, and facilitate subject recruitment and retention.

Acknowledgements

This work was supported, in part, by the Spanish Ministry of Science and Innovation (PR 2007-0358), Centro de Investigacion Biomedica en Red de Salud Mental (CIBERSAM) (to E.V.); by a grant from the Bruce J. Anderson Foundation and the McLean Private Donors' Research Fund (to R.J.B.). The sponsors had no influence on conduct, analysis, or report. Dr Ayşegül Yildiz made meta-analytical calculations. Seren Özer, B.Sc. made valuable contributions to study identification and selection, and data extraction. For providing unpublished information, we thank Drs Charles Bowden, Paul Keck Jr., Roger McIntyre, Harrison G. Pope Jr., Steven Potkin, Richard Weisler, and their colleagues, as well as clinical investigators or other representatives of Abbott, Bristol–Myers Squibb, Eli Lilly, Forest, GlaxoSmithKline, Johnson & Johnson, Novartis, and Pfizer Pharmaceutical Corporations.

Statement of Interest

Dr Yildiz has received research grants from and/or served as a consultant or speaker for Actavis, Ali-Raif, AstraZeneca, Bristol–Myers Squibb, GlaxoSmithKline, Janssen-Cilag, Pfizer, Sanofi-Aventis, and Servier Corporations. Dr Vieta has received grant support, speakers' honoraria, or has been an advisor to Almirall, AstraZeneca, Bristol–Meyer Squibb, Ei Lilly, Forest, GlaxoSmithKline, Janssen-Cilag, Jazz, Johnson & Johnson, Merck, Novartis, Lundbeck, Otsuka, Pfizer, Sanofi-Aventis, Schering-Plough, Servier, Solvay, Takeda, United BioSource, and Wyeth Corporations, but neither he nor family members hold equity positions in pharmaceutical or biomedical corporations. Dr Tohen is a former employee of Eli Lilly (2008). He has acted as consultant or speaker for the following companies: AstraZeneca, Bristol–Myers Squibb, Eli Lilly, GlaxoSmithKline, Otzuka, Sepracor and Wyeth; his spouse is a current Eli Lilly employee and minor stockholder.

References

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

AstraZeneca (2010). Multicenter, double-blind, randomized, parallel-group, placebo controlled, phase III study of the efficacy and safety of quetiapine fumarate (Seroquel®) sustained-release as monotherapy in adult patients with acute bipolar mania. Study ID number: D144CC00004. ClinicalTrials.gov: NCT00422123. AstraZeneca (data on file).

Baldessarini RJ, Salvatore P, Khalsa H-MK, Tohen M (2010). Dissimilar morbidity following initial mania vs. mixed-states in type-I bipolar disorder. *Journal of Affective Disorders* **126**, 299–302.

Benjamin LS (1963). Statistical treatment of the law of initial values (LIV) in autonomic research: a review and recommendation. *Psychosomatic Medicine* **25**, 556–566.

Berkey CS, Hoaglin DC, Mosteller F, Colditz GA (1995). Random-effects regression model for meta-analysis. *Statistics in Medicine* **14**, 395–411.

Berwaerts J, Xu H, Nuamah I, Lim P, *et al.* (2009). Randomized, double-blind, placebo-controlled, dose-response study of paliperidone ER for acute manic and mixed episodes in bipolar I disorder. ClinicalTrials.gov: NCT00299715. American Psychiatric Association 162nd Annual Meeting, San Francisco, CA, May 16–21.

Borenstein M, Hedges LV, Higgins JPT, Rothstein HR (2009). *Introduction to Meta-analysis*, 1st edn. West Sussex, UK: John Wiley & Sons.

Bowden CL, Brugger AM, Swann AC, Calabrese JR, *et al.* (1994). Efficacy of divalproex *vs.* lithium and placebo in the treatment of mania. *Journal of American Medical Association* **271**, 918–924.

Bowden CL, Grunze H, Mullen J, Brecher M, *et al.* (2005). A randomized, double-blind, placebo-controlled efficacy and safety study of quetiapine or lithium as monotherapy for mania in bipolar disorder. *Journal of Clinical Psychiatry* **66**, 111–121.

Bowden CL, Swann AC, Calabrese JR, Rubenfaer LM, *et al.* (2006). Randomized, placebo-controlled, multicenter study of divalproex sodium extended release in the treatment of acute mania. *Journal of Clinical Psychiatry* **67**, 1501–1510.

Charney DS (2000). Use of placebos in randomized clinical trials of mood disorders: well justified, but improvements in design are indicated. *Biological Psychiatry* **47**, 687–688.

Charney DS, Nemeroff CB, Lewis L, Laden SK, *et al.* (2002). National depressive and manic-depressive association consensus statement on use of placebo in clinical trials of mood disorders. *Archives of General Psychiatry* **59**, 262–270.

Duval S, Tweedie R (2000). Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics* **56**, 455–463.

Ehni HJ, Wiesing U (2008). International ethical regulations on placebo-use in clinical trials: a comparative analysis. *Bioethics* **22**, 64–74.

El Mallakh RS, Vieta E, Rollin L, Marcus R, *et al.* (2010). A comparison of two fixed doses of aripiprazole with placebo in acutely relapsed, hospitalized patients with bipolar disorder I (manic or mixed) in subpopulations. Study ID Number: CN138-007. *European Neuropsychopharmacology* **20**, 776–783.

Endicott J, Spitzer RL (1978). A diagnostic interview: the Schedule for Affective Disorders and Schizophrenia. *Archives of General Psychiatry* **35**, 837–844.

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

Hirschfeld RM, Bowden CL, Vigna NV, Wozniak P, *et al.* (2010). A randomized, placebo-controlled, multicenter study of divalproex sodium extended release in the acute treatment of mania. ClinicalTrials.gov: NCT00060905. *Journal of Clinical Psychiatry* **71**, 426–432.

Hirschfeld RM, Keck Jr. PE, Kramer M, Karcher K, *et al*. (2004). Rapid antimanic effect of risperidone monotherapy: 3-week multicenter, double-blind, placebo-controlled trial. *American Journal of Psychiatry* **161**, 1057–1065.

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.

Katagiri H, Takita Y, Tohen M, Takahashi M (2010). Efficacy and safety of olanzapine in the treatment of Japanese patients with manic or mixed episode of bipolar I disorder: a multicenter, randomized, double-blind, parallel, placebo- and haloperidol-controlled study. ClinicalTrials.gov: NCT00129220. International College of Neuropsychopharmacology 27th Annual Congress, Hong Kong, China, June 6–10.

Keck Jr. PE, Marcus R, Tourkodimitris S, Ali M, *et al.* (2003*a*). A placebo-controlled, double-blind study of the efficacy and safety of aripiprazole in patients with acute bipolar mania. *American Journal of Psychiatry* **160**, 1651–1658.

Keck PE, Orsulak PJ, Cutler AJ, Sanchez R, *et al.* (2009). Aripiprazole monotherapy in the treatment of acute bipolar I mania: a randomized, double-blind, placebo- and lithium-controlled study. *Journal of Affective Disorders* **112**, 36–49.

Keck Jr. PE, Versiani M, Potkin S, West SA, *et al.* (2003*b*). Ziprasidone in the treatment of acute bipolar mania: a three-week, placebo-controlled, double-blind, randomized trial. *American Journal of Psychiatry* **160**, 741–748.

Keck Jr. PE, Welge JA, McElroy SL, Arnold LM, *et al*. (2000). Placebo effect in randomized, controlled studies of acute bipolar mania and depression. *Biological Psychiatry* **47**, 748–755.

Kemp AS, Schooler NR, Kalali AH, Alphs L, *et al.* (2010). What is causing the reduced drug-placebo difference in recent schizophrenia clinical trials and what can be done about it? *Schizophrenia Bulletin* **36**, 504–509.

Kent DM, Hayward RA (2007). Limitations of applying summary results of clinical trials to individual patients. *Journal of the American Medical Association* **298**, 1209–1212.

Khanna S, Vieta E, Lyons B, Grossman F, *et al.* (2005). Risperidone in the treatment of acute mania: double-blind, placebo-controlled study. *British Journal of Psychiatry* **187**, 229–234.

Klosterhalfen S, Enck P (2006). Psychobiology of the placebo response. *Autonomic Neuroscience: Basic and Clinical* **125**, 94–99.

Knesevich MA, Wang O, Papadakis K, Korotzer A, *et al.* (2009). The efficacy and tolerability of cariprazine in acute mania associated with bipolar I disorder: a Phase II Trial. ClinicalTrials.gov: NCT00488618. American Psychiatric Association 162nd Annual Meeting, San Francisco, CA, 16–21 May 2009.

Kotzalidis G, Pacchiarotti I, Manfredi G, Vavioja V, *et al.* (2008). Ethical questions in human clinical psychopharmacology: should the focus be on placebo administration? *Journal of Psychopharmacology* **22**, 590–597.

Kushner SF, Khan A, Lane R, Olson WH (2006). Topiramate monotherapy in the management of acute mania: results of four double-blind placebo-controlled trials. ClinicalTrials.gov: NCT00037674; NCT00240721; NCT00035230. *Bipolar Disorders* **8**, 15–27.

Leucht S, Arbter D, Engel RR, Kissling W, *et al.* (2009). How effective are second-generation antipsychoticdrugs? A meta-analysis of placebo-controlled trials. *Molecular Psychiatry* **14**, 429–447.

McIntyre RS, Brecher M, Paulsson B, Huizar K, *et al.* (2005). Quetiapine or haloperidol as monotherapy for bipolar mania: 12-week, double-blind, randomized, parallel-group, placebo-controlled trial. *European Neuropsychopharmacology* **15**, 573–585.

McIntyre RS, Cohen M, Zhao J, Alphs L, *et al.* (2009*b*). A 3-week, randomized, placebo-controlled trial of asenapine in the treatment of acute mania in bipolar mania and mixed states. *Bipolar Disorders* **11**, 673–686.

McIntyre RS, Cohen M, Zhao J, Alphs L, *et al.* (2009*a*). Asenapine *vs.* olanzapine in acute mania: a double-blind extension study. *Bipolar Disorders* **11**, 815–826.

Novartis (2007). Clinical Trial Results Database. A 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. Novartis (data on file).

Orwin RG (1983). A fail-safe N for effect-size in meta-analysis. *Journal of Educational Statistics* **8**, 157–159.

Pope Jr. HG, McElroy SL, Keck Jr. PE, Hudson JI (1991). Valproate in the treatment of acute mania: placebo-controlled study. *Archives of General Psychiatry* **48**, 62–68.

Potkin SG, Keck Jr. PE, Segal S, Ice K, *et al.* (2005). Ziprasidone in acute bipolar mania: 21-day randomized, double-blind, placebo-controlled replication trial. *Journal of Clinical Psychopharmacology* **25**, 301–310.

Sachs G, Sanchez R, Marcus R, Stock E, *et al.* (2006). Aripiprazole in the treatment of acute manic or mixed episodes in patients with bipolar I disorder: a 3-week placebo-controlled study. *Journal of Psychopharmacology* **20**, 536–546.

Smulevich AB, Khanna S, Eerdekens M, Karcher K, *et al.* (2005). Acute and continuation risperidone monotherapy in bipolar mania: 3-week placebo-controlled trial followed by a 9-week double-blind trial of risperidone and haloperidol. *European Neuropsychopharmacology* **15**, 75–84.

Spitzer RL, Endicott J, Robins E (1978). Research diagnostic criteria: rationale and reliability. *Archives of General Psychiatry* **35**, 773–882.

Sterne JAC, Egger M (2001). Funnel plots for detecting bias in meta-analysis: Guidelines on choice of axis. *Journal of Clinical Epidemiology* **54**, 1046–1055.

Sysko R, Walsh BT (2007). Systematic review of placebo response in studies of bipolar mania. *Journal of Clinical Psychiatry* **68**, 1213–1217.

Tohen M, Frank E, Bowden CL, Colom F, *et al.* (2009). International Society for Bipolar Disorders (ISBD) Task Force report on the nomenclature of course and outcome in bipolar disorders. *Bipolar Disorders* **11**, 453–473.

Tohen M, Jacobs TG, Grundy SL, McElroy SL, *et al.* (2000). Efficacy of olanzapine in acute bipolar mania: a double-blind, placebo-controlled study. *Archives of General Psychiatry* **57**, 841–849.

Tohen M, Sanger TM, McElroy SL, Tollefson GD, *et al.* (1999). Olanzapine *vs.* placebo in the treatment of acute mania. *American Journal of Psychiatry* **156**, 702–709.

Tohen M, Vieta E, Goodwin G, Sun B, *et al.* (2008). Olanzapine *vs.* divalproex *vs.* placebo in the treatment of mild to moderate mania: randomized, 12-week, double-blind study. *Journal of Clinical Psychiatry* **69**, 1776–1789.

Vieta E (2003). Divalproex *vs.* olanzapine in mania. *Journal of Clinical Psychiatry* **64**, 1266.

Vieta E, Carne X (2005). Use of placebo in clinical trials on bipolar disorder: new approach for an old debate. *Psychotherapy and Psychosomatics* **74**, 10–16.

Vieta E, Cruz N (2008). Increasing rates of placebo response over time in mania studies. *Journal of Clinical Psychiatry* **69**, 681–682.

Vieta E, Nuamah IF, Lim P, Yuen EC, *et al.* (2010*a*). A randomized, placebo- and active-controlled study of paliperidone extended release for the treatment of acute manic and mixed episodes of bipolar I disorder. ClinicalTrials.gov: NCT00309699. *Bipolar Disorders* **12**, 230–243.

Vieta E, Ramey T, Keller D, English PA, *et al.* (2010*b*). Ziprasidone in the treatment of acute mania: 12-week placebo-controlled, haloperidol-referenced study. *Journal of Psychopharmacology* **24**, 547–558.

Weisler RH, Kalali AH, Ketter TA, SPD417 Study Group (2004). Multicenter, randomized, double-blind, placebo-controlled trial of extended-release carbamazepine capsules as monotherapy for bipolar disorder patients with manic or mixed episodes. *Journal of Clinical Psychiatry* **65**, 478–484.

Weisler RH, Keck Jr. PE, Swann AC, Cutler AJ, *et al.* (2005). Extended-release carbamazepine capsules as monotherapy for acute mania in bipolar disorder: a multicenter, randomized, double-blind, placebo-controlled trial. *Journal of Clinical Psychiatry* **66**, 323–330.

Whitley E, Ball J (2002). Statistics review 5: comparison of means. *Critical Care* **6**, 424–428.

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

Yildiz A, Vieta E, Leucht S, Baldessarini RJ (2010). Efficacy of anti-manic treatments: meta-analysis of randomized controlled trials. *Neuropsychopharmacology*. Published online: 27 October 2010. doi:10.1038/npp.2010.192.

Young AH, Oren DA, Lowy A, McQuade RD, *et al.* (2009). Aripiprazole monotherapy in acute mania: a 12-week, randomized placebo- and haloperidol-controlled study. *British Journal of Psychiatry* **194**, 40–48.

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

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