Placebo Effect in Randomized, Controlled Studies of Acute Bipolar Mania and Depression

Paul E. Keck, Jr., Jeffrey A. Welge, Susan L. McElroy, Lesley M. Arnold, and Stephen M. Strakowski

Randomized, double-blind, placebo-controlled, parallel group clinical trials have been the standard methodology for establishing the efficacy of new treatments for patients with bipolar disorder in manic, mixed, or depressive episodes. We examine the placebo response rate in acute treatment trials of acute mania (and mixed states) and bipolar depression. Also addressed are potential variables associated with placebo response, strategies to minimize placebo response, the optimum duration of placebo-controlled acute treatment trials, possible alternatives to the use of placebo, and the ramifications of these issues with regard to the design of studies in children, adolescents, and older adults with bipolar disorder. Biol Psychiatry 2000;47:748–755 © *2000 Society of Biological Psychiatry*

Key Words: Placebo, bipolar disorder, mania, depression, clinical trials

Introduction

Placebo-controlled, double-blind, randomized clinical trials have been recognized as the standard by which to establish a drug's safety, efficacy, and dose–response relationships in most major psychiatric illnesses (Klerman 1986; Leber 1991, 1996). The justifications for using placebo control groups in such trials are the fluctuating natural course of most psychiatric illnesses, the wide variability in placebo response across patient groups, and the influence of psychosocial factors on treatment response (Klerman 1986; Quitkin 1999; Stanley 1988). To our knowledge, the placebo response in acute treatment clinical trials of patients with bipolar I disorder has not been previously examined. In this article we assess the range of placebo response rates in acute treatment trials of patients with bipolar disorder in manic, mixed, and depressive episodes. In addition, we review potential variables related to placebo response, alternatives in trial designs to minimize placebo response, possible alternatives to the use of placebo control groups, the optimum duration of acute treatment trials involving placebo, and the implications of these issues in studies of children, adolescents, and older adults with bipolar disorder.

From the Biological Psychiatry (PEK, SLM, LMA) and Bipolar and Psychotic Disorders Research (SMS) Programs, Department of Psychiatry, University of Cincinnati College of Medicine, and Center for Biostatistical Services, University of Cincinnati Medical Center (JAW), Cincinnati, Ohio.

Address reprint requests to Paul E. Keck, Jr., M.D., University of Cincinnati College of Medicine, Department of Psychiatry, 231 Bethesda Avenue, PO Box 670559, Cincinnati, OH 45267-0559.

Received October 5, 1999; revised December 1, 1999; accepted December 3, 1999.

Methods and Materials

We searched the scientific literature for all placebo-controlled acute treatment trials in patients with bipolar I disorder in manic, mixed, or depressive episodes using Paperchase. This search was augmented by reviewing all bibliographies of identified papers. Twenty-two studies of acute manic or mixed patients and 13 studies of bipolar depression were identified.

Acute Mania Trials

Of the 22 studies in acute manic or mixed episodes, 13 were excluded from analysis because of the utilization of a crossover design (Ballenger and Post 1978; Chouinard et al 1982; Cohen et al 1982; Dose et al 1986; Dubovsky et al 1986; Emrich et al 1979; Giannini et al 1984; Goodwin et al 1969; Krieg and Berger 1986; Maggs 1963; Schou et al 1954; Stokes et al 1971; Van Berkestijn et al 1990). Crossover studies were excluded from our analysis of placebo response because of their vulnerability to carryover and period effects; abrupt treatment discontinuation effects, which may have artificially lowered placebo response rates; and potential contamination of blindedness (Calabrese and Rapport 1999; Stallone et al 1974). Of note, four lithium crossover studies had additional methodological limitations—two utilized nonrandom assignment (Goodwin et al 1969; Stokes et al 1971) and two others (Chouinard et al 1982; Giannini et al 1984) allowed administration of other antimanic medications. Moreover, many of the early pioneering lithium studies used diagnostic criteria to define bipolar disorder that may not be comparable to those of DSM-III-R (American Psychiatric Association 1987) or DSM-IV (American Psychiatric Association 1994; Bowden et al 1995; Calabrese and Rapport 1999). One parallel group study (Chambers and Naylor 1978) was also excluded from analysis because other antimanic medications were administered during the trials. The eight remaining double-blind, placebo-controlled acute treatment studies of bipolar manic or mixed episodes are displayed in Tables 1–3 (Aldenhoff et al 1986; Bowden et al 1994; Janicak et al 1989, 1998; Klein 1967; Pope et al 1991; Tohen et al 1999a, 1999b).

0006-3223/00/$20.00
PII S0006-3223(99)00311-X

Table 1. Double-Blind, Placebo-Controlled, Parallel Group Studies in Acute Mania: Demographic and Clinical Characteristics

Study	*N*	Criteria	M/F	Age (years)	% mixed	% psychotic	% rapid cycling	Age onset (years)
Klein 1967	13	Clinical	ND	ND	ND	ND	ND	ND
Aldenhoff et al 1986	10	DSM-III	3/7	39 (—)	ND	ND	ND	ND
Janicak et al 1989	21	DSM-III	5/16	32 (11)	10	86	ND	ND
Pope et al 1991	36	DSM-III-R	26/10	37 (12)	ND	67	27	25
Bowden et al 1994	179	RDC	93/86	40 (13)	ND	ND	ND	22
Janicak et al 1998	32	DSM-III-R	19/13	36 (11)	7	ND	ND	ND
Tohen et al 1999b	136	DSM-IV	70/66	40 (11)	17	53	32	ND
Tohen et al 1999a	115	DSM-IV	58/57	39 (10)	43	73	39	ND

ND, no data provided; RDC, Research Diagnostic Criteria.

Acute Depression Trials

Thirteen placebo-controlled studies of acute bipolar depression were identified. Of these, eight utilized a crossover design with short treatment intervals, ranging from 1 to 28 days, of lithium or placebo administration (Baron et al 1975; Donnelly et al 1978; Goodwin et al 1969, 1972; Greenspan et al 1970; Mendels 1976; Noyes et al 1974; Stokes et al 1971). After exclusion of these studies on methodological grounds, five double-blind, placebo-controlled, parallel group acute treatment studies of bipolar I depression remained for analysis (Tables 4–6; Calabrese et al 1999; Cohn et al 1989; Himmelhoch et al 1982; Levine et al 1995; Nemeroff et al, in press).

Placebo Response

The magnitude of placebo response in these acute treatment studies was examined by comparing drug and placebo response rates and by effect size (ES) calculations for drug, placebo, and the drug–placebo difference. Effect size calculations provided an estimation of the magnitude of a difference between treatment groups on the primary measure of outcome in clinical trials (Bowden et al 1997; Cohen 1988). In other words, ES calculations provide a means of assessing whether the change with drug or placebo treatment, and between drug and placebo treatment, is small (ES = 0.2), modest (ES = 0.5), or large (ES = 0.8 or greater). In our analysis of placebo-controlled studies, ES was calculated using the primary measures of mania (e.g., Young Mania Rating Scale) and depression (e.g., Hamilton Depression Rating Scale) and as suggested by Cohen (1988):

$$\text{ES} = \frac{\text{mean at baseline} - \text{mean at endpoint}}{\text{SD}}$$

For computing the ES difference between drug and placebo we used the pooled SD (Cohen 1988), which represented the SD of the change score. When these data were not reported, they were either provided by the authors (Janicak et al 1989) or estimated by using the baseline SD (Aldenhoff et al 1986). For the specific group ES calculation (i.e., ES of drug, ES of placebo) the SD of the corresponding group was used.

Results

The characteristics of the eight acute mania and five acute bipolar depression studies are summarized in Tables 1–6.

Table 2. Double-Blind, Placebo-Controlled, Parallel Group Studies in Acute Mania: Study Methods

Study	Washout duration (days)	Medications (mg/day)	Duration (weeks)	Primary outcome measure	Minimum severity criteria	As-occasion-required medication
Klein 1967	0	Chlorpromazine (to 1200), imipramine (to 300)	6	Global Scale −9 to +9	None	None
Aldenhoff et al 1986	35	4-Methoxyverapamil (75)	2	BRMS	None	CH 2 g/day
Janicak et al 1989	6	Clonidine av. (0.5)	2	9-item MRS	None	None
Pope et al 1991	1	Divalproex	3	YMRS	None	LZPM 4 mg/day to day 9
Bowden et al 1994	3–21	Divalproex, lithium	3	SADS-C (MRS)	MRS ≥14, ≥2 on four items	CH 4 g, LZPM 2 mg/day to day 4, 1 mg/day to day 10
Janicak et al 1998	12	Verapamil (480)	3	YMRS	None	CH 4 g/day, LZPM 2 mg/day tapered by day 10
Tohen et al 1999b	2–4	Olanzapine	3	YMRS	YMRS ≥20	LZPM 4 mg/day to day 7, 2 mg/day to day 10
Tohen et al 1999a	1	Olanzapine	4	YMRS	YMRS ≥20	LZPM 2 mg/day to day 4, 1 mg/day to day 10

BRMS, Bech Rafaelson Mania Scale; CH, chloral hydrate; MRS, Mania Rating Scale; YMRS, Young Mania Rating Scale; LZPM, lorazepam; SADS-C, Schedule for Affective Disorders and Schizophrenia—Change Scale.

Table 3. Double-Blind, Placebo-Controlled, Parallel Group Studies in Acute Mania: Results

Study	Primary outcome (mean change)	% responders		Effect size		Effect size drug–placebo (95% confidence interval)	Significant difference
		Drug	Placebo	Drug	Placebo		
Klein 1967	CPZ +6.1 > P +2.0 CPZ +6.1 > IMI −2.5	ND	ND	ND	ND	ND	ND
Aldenhoff et al 1986	4 MV −7.3 = P −4.0	ND	ND	0.7	0.4	0.5 (−1.1, 2.1)	NA
Janicak et al 1989	ND	8	22	0.4	0.2	−0.1 (−1.0, 0.8)	NA
Pope et al 1991	DVPX −11.4 > P −0.5	53	11	1.9	0.1	1.7 (0.9, 1.8)	Day 7
Bowden et al 1994	DVPX −10 > P −4 LI −10 > P −4	48,49	24	0.7	0.4	0.4 (0.1, 0.8)	Day 10
Janicak et al 1998	V −1.1 = P −1.3	18	13	0.1	0.1	−0.1 (−0.2, 0.2)	NA
Tohen et al 1999b	O −10 > P −5	49	24	1.5	0.4	0.6 (0.1, 0.9)	Day 21
Tohen et al 1999a	O −15 > P −8	65	43	1.2	0.6	0.5 (0.2, 0.9)	Day 7

CPZ, chlorpromazine; P, placebo; IMI, imipramine; ND, no data provided; MV, 4-methoxyverapamil; NA, not applicable; DVPX, divalproex; LI, lithium; V, verapamil; O, olanzapine.

Acute Mania Studies

The demographic and clinical characteristics of the patients in the eight acute mania studies are presented in Table 1. With the exception of the study by Klein (1967) conducted over 3 decades ago, the diagnostic criteria used to define the patient populations in the acute mania studies are comparable. Overall, the gender distribution was nearly 1:1, and patients were on average in their late 30s with an average duration of illness of 14 years. Five studies were conducted at single centers (Aldenhoff et al 1986; Janicak et al 1989, 1998; Klein 1967; Pope et al 1991); the other three were large, multicenter trials. In studies in which clinical features of patients were provided in detail, the majority of patients were manic rather than mixed. Similarly, most patients had psychotic features. Rapid-cycling patients constituted nearly a third of the study sample in the three studies in which the presence of rapid cycling was assessed.

The methods of these eight studies are presented in Table 2. The duration of washout before randomization ranged from 0 to 35 days, but was typically 1–7 days in more recent trials. Studies ranged in duration from 2 to 6 weeks, and specific mania rating scales were used in all but Klein's study. All patients were hospitalized, but only three studies established minimum severity criteria for inclusion. Most studies allowed the use of chloral hydrate and/or lorazepam on a limited, as-occasion-required basis for the first 10 days of randomized treatment.

The results of the eight acute mania studies are summarized in Table 3. Three studies were negative, failing to find a significant difference between drug and placebo (Aldenhoff et al 1986; Janicak et al 1989, 1998). The treatment response was defined as a 50% or more reduction in manic symptoms in four studies (Bowden et al 1994; Pope et al 1991; Tohen et al 1999a, 1999b) and as a 40% or more reduction in two studies (Janicak et al 1989, 1998). The mean (± SD) percentage of placebo responders in all studies was 23% (± 11%; range 11–43%), compared with 40% (± 24%; range 8–65%) of drug responders.

The mean ES for drug-treated patients from all studies was 0.9; for patients receiving placebo it was 0.4. For positive studies (those yielding a significant difference of drug over placebo), the mean ESs were 1.1 for the treatment group and 0.4 for the placebo group. In the three negative studies, the mean ES for both drug and placebo groups was small, 0.2.

There are a number of variables that could contribute to

Table 4. Double-Blind, Placebo-Controlled, Parallel Group Studies in Bipolar Depression: Demographic and Clinical Characteristics

Study	*N*	Criteria	M/F	Age (years) and mean (SD)	Age onset (years) and mean (SD)	Minimum severity criteria
Himmelhoch et al 1982	59	RDC and anergic depression	19/40	40 (12)	ND	RDS ≥ 7
Cohn et al 1989	89	DSM-III	30/59	40 (11)	ND	HDRS ≥ 20 RDS ≥ 20
Levine et al 1995	6	DSM-III-R	2/4	ND	ND	None
Calabrese et al 1999	195	DSM-IV	77/118	42 (10)	21 (14)	HDRS ≥ 15
Nemeroff et al, in press	117	DSM-III-R	52/65	42 (10)	ND	HDRS ≥ 15

RDC, Research Diagnostic Criteria; ND, no data provided; RDS, Raskin Depression Scale; HDRS, Hamilton Depression Rating Scale.

Table 5. Double-Blind, Placebo-Controlled, Parallel Group Studies in Bipolar Depression: Study Methods

Study	Washout duration (days)	Medication(s) (mg/day)	Concomitant mood stabilizers	Duration (weeks)	Primary outcome measures
Himmelhoch et al 1982	ND	Tranylcypromine	None	6	Not specified; RDS, Kupfer–Detre System, SCL-90, New Physicians' Rating List, GAS
Cohn et al 1989	7, placebo	Fluoxetine, imipramine	Lithium (25% of pts)	6	Not specified; HDRS, RDS, Covi Anxiety Scale, CGI
Levine et al 1995	3–7	Inositol 12 g/day	None	4	HDRS
Calabrese et al 1999	5 × t½ previous medications	Lamotrigine 50, 200 mg/day	None	7	HDRS
Nemeroff et al, in press	7, placebo	Paroxetine, imipramine	Lithium (≤0.8, >0.8) +/− CBZ or VPA (9%)	10	HDRS, CGI

ND, no data provided; RDS, Raskin Depression Scale; SCL-90, Symptoms Checklist—90 items; GAS, Global Assessment Scale; HDRS, Hamilton Depression Rating Scale; CGI, Clinical Global Improvement Scale; t½, elimination half-life of prior psychotropic agents; CBZ, carbamazepine; VPA, valproate.

the moderate placebo response in these studies. These include the duration of drug washout prior to randomization; the use, dose and duration of rescue medications; severity of illness; number of sites (variance in interrater reliability and study population); clinical features (e.g., first-episode mania, rapid cycling, psychosis, manic vs. mixed episodes); trial duration; the intensity and therapeutic effect of the hospital milieu; and spontaneous remission or switching. Identification of specific variables associated with placebo response in these studies is limited by the small sample sizes and low placebo response rates in single-site studies. Analyses of variables potentially associated with placebo response have been conducted on two of the three multicenter studies (Bowden et al 1997; Tohen et al 1999b). The highest placebo response rates (24–43%) occurred in the multicenter studies. This suggests that variance in rater reliability or study populations across centers may contribute to placebo response. Bowden et al (1997) calculated the placebo ES on subitem behavioral measures from the mania rating scale results from their earlier study (Bowden et al 1994). Placebo exerted a moderate ES (0.4) on the manic syndrome and elevated mood measures, but a small ES (≤0.2) on measures of increased activity, psychosis, and decreased need for sleep. These latter measures may be more observable and less vulnerable to self-report bias. In addition, the manic syndrome and elevated mood measures may also be more responsive to the psychosocial support provided by study participation. In the first olanzapine study (Tohen et al 1999b), patients with a first episode of mania had a high placebo response. In contrast, in the study with the lowest placebo response (11%; Pope et al 1991) the inclusion of only lithium-refractory patients may have contributed to the low response. Interestingly, other variables such as washout duration, psychosis, manic versus mixed episodes, rapid cycling, and use of rescue medications did not

Table 6. Double-Blind, Placebo-Controlled, Parallel Group Studies in Bipolar Depression: Results

Study	Primary outcome (mean change)	% responders		Effect size		Effect size drug–placebo (95% confidence interval)	Significant difference
		Drug	Placebo	Drug	Placebo		
Himmelhoch et al 1982	NPRL: TCP −16 > P −9 GAS: TCP 19 > P 11	71	13	ND	ND	ND	4 weeks
Cohn et al 1989	FL −18 > IMI −9 FL −18, P −16 ns	FL 85 IMI 57	38	1.5 (FL)	0.5	1.0 (0.6, 1.3)	3 weeks
Levine et al 1995	IN −15, P −14 ns	ND	ND	1.1	1.1	0	NA
Calabrese et al 1999	HDRS: LH −13, LL −13 > P −8 ns (MADRS: LH −13 > LL −11, P −8)	LH 51 LL 45	37	1.2	1.0	0.3 (0.1, 0.4)	5 weeks
Nemeroff et al, in press	Lli: Par −10, IMI −11 > P −6 Hli: Par −10, IMI −9, P −10 Total: Par −10, IMI −10, P −8	P 53, I 3 P 36, I 41 P 46, I 3	32 38 35	1.4	1.1	0.3 (0.1, 0.4)	ND

NPRL, New Physicians' Rating List; TCP, tranylcypromine; P, placebo; GAS, Global Assessment Scale; ND, no data provided; FL, fluoxetine; IMI, imipramine; IN, inositol; NA, not applicable; HDRS, Hamilton Depression Rating Scale; LH, lamotrigine high dose (200 mg/day); LL, lamotrigine low dose (50 mg/day); MADRS, Montgomery–Asberg Depression Rating Scale; Lli, low lithium levels (≤0.8 mEq/L); Par, paroxetine; Hli, high lithium levels (>0.8 mEq/L).

appear to influence placebo response among all studies. In fact, when greater restrictions on the use of as-occasion-required lorazepam were placed on the second Tohen et al olanzapine trial (1999a) in an attempt to reduce placebo response, the placebo response increased.

In positive studies, reductions in manic symptoms from baseline in drug and placebo reached statistical significance by day 7 in two studies (Pope et al 1994; Tohen et al 1999a), by day 10 in one study (Bowden et al 1994), and 21 days in another (Tohen et al 1999b). It is possible that drug separation from placebo may have occurred earlier in the studies that found such a difference by day 7 if ratings had been performed earlier (e.g., at day 3 or day 5). The starting dose and rate of titration may also affect the rate of response. For example, a higher starting dose of olanzapine may have contributed to the earlier separation of drug from placebo (7 vs. 21 days) in the second Tohen et al trial (1999a).

Taken together, these findings suggest that the placebo response in studies of acute mania might be minimized by restricting the number of study sites and patients with first episodes of mania or by including patients who are treatment refractory to at least one mood stabilizer; however, restricting the number of sites would potentially slow the rate of the completion of trials, and excluding first-episode or including only treatment-refractory patients would limit the generalizability of the study results.

The optimum duration of acute mania trials is unclear. The earliest separation of drug from placebo has been at day 7, but an earlier separation may have gone unmeasured. There seems little justification for extending a placebo-controlled study in patients with acute mania beyond 21 days. The optimum trial duration may fall between 7 and 14 days. It should be possible to estimate an optimum duration for such a trial because the ES difference between drug and placebo is a function of the duration of the study.

At least four potential alternatives to the use of placebo control group have been proposed in studies of acute mania (Leber 1991, 1996). These include add-on studies, variable dose designs, establishing *a priori* threshold ES with an active comparison control, and comparisons with historical controls. Add-on designs do not obviate the need for placebo but eliminate placebo monotherapy; however, substantially larger study populations are needed for sufficient power to establish a drug–placebo difference because of the contribution of the primary agent(s) to both drug and placebo effects. This was recently demonstrated in an add-on study of gabapentin (Pande 1999). In that study, gabapentin did not differ significantly in efficacy from placebo. Alternatively, it may be that gabapentin lacks efficacy in the treatment of manic symptoms. The use of add-on designs could also influence the duration of the trial. For example, in patients with more treatment-resistant mania (experiencing symptoms while already receiving at least one agent), 2–3 weeks may not be long enough to gauge time on study drug effects. Variable dose designs allow for the possibility of establishing dose–response relationships; however, it must be clearly specified in the informed consent process that some doses may not exert a therapeutic effect. Data regarding ESs of drug versus placebo reviewed above suggest that establishing a threshold ES that an investigational drug must reach or exceed (e.g., ES 0.9) in a trial with an active control might obviate the need for a placebo control; however, the possibility of a robust placebo effect in both treatment groups still cannot be excluded from such trials. Finally, comparing efficacy results of an investigational agent with historical data from previous trials has been suggested as an alternative to placebo control groups. This approach is limited by the unpredictable course of illness in bipolar disorder, changes in diagnostic criteria, variability in rating scales used, and different patient demographic and clinical characteristics over time.

There is only one placebo-controlled acute treatment study, to our knowledge, in children or adolescents with mania (Geller et al 1998), and there are none in older adults with mania. Geller et al (1998) conducted a 6-week, double-blind, placebo-controlled, parallel group study of lithium in adolescents with bipolar disorder in a manic episode and secondary substance dependence. Response was defined as a score of 65 or more on the Children's Global Assessment Scale in the intent-to-treat sample. By this definition, only 8% (1 of 12) of patients randomized to placebo were responders, compared with 46% (6 of 13) of patients who received lithium. Thus, the placebo response rate in adolescents in this single-site trial was quite low; however, the placebo response in manic adolescents without co-occurring substance dependence is unknown. Substantial differences in the clearance and tolerability of psychotropic agents in these populations need to be addressed in clinical trial designs (Geller et al 1998; Malone and Simpson 1998).

Acute Bipolar Depression Studies

The demographic and clinical characteristics of the patients in the five acute bipolar depression studies are displayed in Table 4. All five studies utilized comparable diagnostic criteria. The study by Himmelhoch et al (1982), which specified that patients also met criteria for anergic depression, is included because nearly half of the patients ($N = 29$) also met criteria for bipolar depression, and another 11 patients were characterized as "pseudounipolar." Overall, the gender distribution was approximately 3:2 (F:M), and patients were on the average in their early

40s. Three studies were conducted at single sites (Cohn et al 1989; Himmelhoch et al 1982; Levine et al 1995) and two were large multicenter trials (Calabrese et al 1999; Nemeroff et al, in press). All but one study (Levine et al 1995), which is limited because of its very small sample size ($N = 6$), specified *a priori* minimum severity criteria based on various depression rating scale scores. The minimum Hamilton Depression Rating Scale (HDRS) severity score in the Nemeroff et al (in press) study was substantially lower than the thresholds used by Calabrese et al (1999) and Cohn et al (1989).

The methods utilized in these five studies are summarized in Table 5. For studies in which the duration of washout was specified, the washout period ranged from 3 to 7 days. Two studies used single-blind placebo run-ins during the washout interval (Cohn et al 1989; Nemeroff et al, in press). Three studies examined the efficacy of putative antidepressants without concomitant mood stabilizers (Calabrese et al 1999; Himmelhoch et al 1982; Levine et al 1995); one study included some (25%) patients already on lithium, introducing significant heterogeneity to the treatment groups (Cohn et al 1989); and one administered study treatment to patients on lithium, stratified into low ($\leq$0.8 mEq/L) and high ($>$0.8) serum concentrations (Nemeroff et al, in press). In this latter study, 9% of patients were also receiving carbamazepine or valproate in addition to lithium. Studies ranged in duration from 4 to 10 weeks, and specific rating scales for depression defining primary outcome measures were used in only the three most recent studies (Calabrese et al 1999; Levine et al 1995; Nemeroff et al, in press). In the study by Himmelhoch et al (1982) the study blind was broken at week 4 for patients who had not responded.

The results of the placebo-controlled acute bipolar depression studies are presented in Table 6. The results of two studies are limited by their analysis of study completers only (Cohn et al 1989; Himmelhoch et al 1982). Cohn et al based their criteria for response on an improvement of 50% or more on the HDRS for patients remaining in the study for at least 3 weeks; however, 29% of patients, including 47% in the placebo group, had dropped out of the study by week 3. Overall, only 49% of all patients enrolled, and only 34% receiving placebo, completed the 6-week trial. By not including study dropouts in an intent-to-treat, last-observation-carried-forward analysis, the placebo response rate in this study is likely to have been substantially inflated (Zornberg and Pope 1993). Three of the five studies were negative, failing to find a significant difference between drug and placebo on the primary measure of outcome (Cohn et al 1989; Levine et al 1995; Nemeroff et al, in press); however, Cohn et al (1989) found fluoxetine significantly more efficacious than imipramine, and Nemeroff et al (in press) found paroxetine and imipramine superior to placebo when they were added to the low–lithium level group, but not when they were added to the high–lithium level group. It is also possible that the therapeutic effects of higher lithium levels in reducing anxiety and psychomotor agitation obscured the impact of paroxetine and imipramine on these behavioral dimensions. This latter finding highlights the difficulty of discerning drug–placebo differences in add-on studies. In these five studies, the mean ($\pm$ SD) percentage of placebo responders (response defined as a 50% or more reduction in depressive symptoms as assessed by the primary depression rating scale) was 29% ($\pm$ 12%; range 13–38), compared with 64% ($\pm$ 24%; range 45–86) of drug responders.

The mean ESs for all studies were 1.3 for drug-treated patients and 0.9 for patients receiving placebo. Variables potentially contributing to the robust placebo ES in these studies include the duration of washout, use of single-blind placebo during washout, severity of illness, number of sites, clinical characteristics of the patient population (e.g., rapid cycling, co-occurring personality disorders), trial duration, spontaneous remission or switching, and type of statistical analysis used. No analyses of the relationship between these variables and placebo response in any study have been conducted, to our knowledge. The earliest placebo-controlled bipolar depression trials (Cohn et al 1989; Himmelhoch et al 1982) used responder analyses on observed data or completer data without distinction between primary and secondary outcome measures. These methodological limitations may have altered the true placebo response rate. The results of the study by Calabrese et al (1999) suggest that the Montgomery–Asberg Depression Rating Scale (MADRS) may be more sensitive in detecting drug–placebo differences, since separation was apparent on the MADRS by week 3 but not until week 5 on the HDRS. Finally, the high placebo response rate in the study by Nemeroff et al (in press) may have been attributable to the lower minimum threshold for severity of depressive symptoms at study entry and the therapeutic effect of lithium, especially at higher serum levels.

The optimum duration of a placebo-controlled trial in patients with acute bipolar depression is unclear. In the only study that revealed a difference in monotherapy between drug and placebo in a homogeneous sample of patient with bipolar I disorder (Calabrese et al 1999), superiority of drug over placebo was evident by the 3rd week of treatment of the MADRS and by the 5th week on the HDRS. The MADRS differs from the HDRS in containing items more closely linked to diagnostic criteria used to define major depression. Furthermore, the optimum duration of acute bipolar depression trials is affected by the time required to reach steady state at a therapeutic dosage. For example, in the study by Calabrese et al

(1999) the time to separation of drug from placebo effect is likely to have been influenced by the conservative dosage titration of lamotrigine.

Similar alternatives to the use of a placebo control group have been proposed for studies of acute bipolar depression as with acute bipolar mania. The hazards of add-on designs are evident in the study by Nemeroff et al (in press), in which the ES for placebo was a robust 1.1. This high ES may have been due, in part, to the therapeutic effect of higher plasma concentrations of lithium. The limitations of variable dose regimens and historical controls were described earlier. Finally, the strategy of establishing minimum ES *a priori* thresholds in active comparator studies of bipolar depression seems especially vulnerable in studies of acute bipolar depression given the high ES associated with placebo in these trials.

Conclusion

The placebo response rate in studies of patients with acute mania is substantial, averaging 23% with an average ES of 0.4. In studies of acute bipolar depression the placebo response rate is even higher, averaging 29% with an average ES of 0.9. These high placebo response rates indicate that alternative trial designs that omit a placebo control group are likely to have limited validity. By taking appropriate measures to ensure adequate informed consent and thorough safeguards to ensure patient safety, placebo-controlled studies for acute mania and bipolar depression remain a viable and necessary standard by which to develop new treatments for bipolar disorder.

Supported in part by a grant from the Theodore and Vada Stanley Foundation, a program of the National Alliance for the Mentally Ill Research Institute (PEK, SLM, LMA, SMS), and National Institute of Mental Health Grant No. MH58170 (SMS).

Aspects of this work were presented at the conference "Clinical Trials in Mood Disorders: The Use of Placebo . . . Past, Present, and Future," September 14–15, 1999, Washington, DC. The conference was sponsored by the National Depressive and Manic-Depressive Association through unrestricted educational grants provided by Abbott Laboratories, Bristol-Myers Squibb Company, Forest Laboratories, Inc., Glaxo Wellcome Inc., Janssen Pharmaceutica Products, L.P., Merck & Company, Pfizer Inc., Pharmacia & Upjohn, SmithKline Beecham Pharmaceuticals, Solvay Pharmaceuticals, Inc., and Wyeth-Ayerst Laboratories.

References

Aldenhoff JB, Schlegal S, Heuser I, Wetzel H (1986): Antimanic effects of the calcium-antagonist D600. A double-blind placebo-controlled study. *Clin Neuropharmacol* 9(suppl 4):553–555.

American Psychiatric Association (1987): *Diagnostic and Statistical Manual of Mental Disorders,* 3rd ed rev. Washington, DC: American Psychiatric Press.

American Psychiatric Association (1994): *Diagnostic and Statistical Manual of Mental Disorders,* 4th ed. Washington, DC: American Psychiatric Press.

Ballenger JC, Post RM (1978): Therapeutic effects of carbamazepine in affective illness: A preliminary report. *Commun Psychopharmacol* 2:159–175.

Baron M, Gershon ES, Rudy V (1975): Lithium carbonate response in depression. *Arch Gen Psychiatry* 32:1107–1111.

Bowden CL, Brugger AM, Swann AC, Calabrese JR, Janicak PG, Petty F, et al (1994): Efficacy of divalproex vs lithium and placebo in the treatment of mania. *JAMA* 271:918–924.

Bowden CL, Calabrese JR, Wallin BA, Swann AC, McElroy SL, Risch SC, et al (1995): Who enters therapeutic trials? Illness characteristics of patients in clinical drug studies of mania. *Psychopharmacol Bull* 31:103–109.

Bowden CL, Davis J, Morris D, Swann A, Calabrese J, Lambert M, Goodnick P (1997): Effect size efficacy measures comparing divalproex, lithium and placebo in acute mania. *Depression Anxiety* 6:26–30.

Calabrese JR, Bowden CL, Sachs GS, Ascher JA, Monaghan E, Rudd GD, et al (1999): A double-blind placebo-controlled study of lamotrigine monotherapy in outpatients with bipolar I depression. *J Clin Psychiatry* 60:79–88.

Calabrese JR, Rapport DJ (1999): Mood stabilizers and the evolution of maintenance study designs in bipolar I disorder. *J Clin Psychiatry* 60(suppl 5):5–13.

Chambers CA, Naylor GJ (1978): A controlled trial of L-tryptophan in mania. *Br J Psychiatry* 132:555–559.

Chouinard G, Young SN, Annable L (1982): A controlled clinical trial of L-tryptophan in acute mania. *Biol Psychiatry* 20:546–557.

Cohen BM, Lipinski JF, Altesman RI (1982): Lecithin in the treatment of mania: double-blind, placebo-controlled trials. *Am J Psychiatry* 139:1162–1164.

Cohen J (1988): *Statistical Power Analysis for the Behavioral Sciences,* 2nd ed. Hillsdale, NJ: Lawrence Erlbaum.

Cohn JB, Collins G, Ashbrook E, Wernicke JF (1989): A comparison of fluoxetine, imipramine and placebo in patients with bipolar depressive disorder. *Int Clin Psychopharmacol* 4:313–322.

Donnelly EF, Goodwin FK, Waldman IN (1978): Prediction of antidepressant responses to lithium. *Psychiatr Res* 18:309–320.

Dose M, Emrich HM, Cording-Tömmel C, von Zerssen D (1986): Use of calcium antagonists in mania. *Psychoneuroendocrinology* 11:241–243.

Dubovsky SL, Franks RD, Allen S, Murphy J (1986): Calcium antagonists in mania: A double-blind study of verapamil. *Psychiatr Res* 18:309–320.

Emrich HM, v Zerssen D, Möller H-J, Kissling W, Cording D, Schietsch HJ, et al (1979): Action of propranolol in mania: comparison of the effects of the d- and l-stereoisomer. *Pharmakopsychiatrie* 12:295–304.

Geller B, Cooper TB, Sun K, Zimerman B, Frazier J, Williams M, et al (1998): Double-blind and placebo-controlled study of lithium for adolescent bipolar disorders with secondary substance dependency. *J Am Acad Child Adolesc Psychiatry* 37:171–178.

Giannini AJ, Houser WL Jr, Loiselle RH, Giannini ML, Price

WA (1984): Antimanic effects of verapamil. *Am J Psychiatry* 141:1602–1603.

Goodwin FK, Murphy DL, Bunney WE Jr (1969): Lithium carbonate treatment in depression and mania. *Arch Gen Psychiatry* 21:486–496.

Goodwin FK, Murphy DL, Dunner DL (1972): Lithium response in unipolar versus bipolar depression. *Am J Psychiatry* 129:76–79.

Greenspan K, Schildkraut JJ, Gordon EK (1970): Catecholamine metabolism in affective disorders-III. *J Psychiatr Res* 7:171–182.

Himmelhoch JM, Fuchs CZ, Symons BJ (1982): A double-blind study of tranylcypromine treatment of major anergic depression. *J Nerv Ment Dis* 170:628–634.

Janicak PG, Sharma RP, Easton M, Comaty JE, Davis JM (1989): A double-blind, placebo-controlled trial of clonidine in the treatment of acute mania. *Psychopharmacol Bull* 25:243–245.

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

Klein DF (1967): Importance of psychiatric diagnoses in prediction of clinical drug effects. *Arch Gen Psychiatry* 16:118–126.

Klerman GL (1986): Scientific and ethical considerations in the use of placebo controls in clinical trials in psychopharmacology. *Psychopharmacol Bull* 22:25–29.

Krieg JC, Berger M (1986): Treatment of mania with the cholinomimetic agent R86 (letter). *Br J Psychiatr* 148:613.

Leber P (1991): The future of controlled clinical trials. Is there an alternative to the randomized controlled trial? *Psychopharmacol Bull* 27:3–8.

Leber P (1996): Challenges to the ethicality of clinical research. *Psychopharmacol Bull* 32:11–13.

Levine J, Barak Y, Gonzalves M, Szor H, Elizur A, Kofman O, et al (1995): Double-blind, controlled trial of inositol treatment of depression. *Am J Psychiatry* 152:792–794.

Maggs R (1963): Treatment of manic illness with lithium carbonate. *Br J Psychiatr* 109:56–65.

Malone RP, Simpson GM (1998): Use of placebos in clinical trials involving children and adolescents. *Psychiatr Serv* 49:1413–1414.

Mendels J (1976): Lithium in the treatment of depression. *Am J Psychiatry* 4:373–378.

Nemeroff CB, Evans DL, Gyulai L, Sachs GS, Bowden CL, Gergel IP, et al (in press): A double-blind, placebo-controlled comparison of imipramine and paroxetine in the treatment of bipolar depression. *Am J Psychiatry.*

Noyes R, Dempsey GM. Blum A (1974): Lithium treatment of depression. *Compr Psychiatry* 15:187–193.

Pande A (1999, June): Adjunctive treatment with gabapentin for bipolar disorder. Paper presented at the Third International Conference on Bipolar Disorder, Pittsburgh.

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

Quitkin FM (1999): Placebos, drug effects, and study design: A clinician's guide. *Am J Psychiatry* 156:829–836.

Schou M, Juel-Nielson N, Strömgren E, Voldby H (1954): The treatment of manic psychoses by the administration of lithium salts. *J Neurol Neurosurg Psychiatry* 17:250–260.

Stallone F, Mendlewicz J, Fieve RR (1974): How blind is the double-blind?: An assessment in a lithium-prophylaxis study (letter). *Lancet* 1:619–620.

Stanley B (1988): An integration of ethical and clinical considerations in the use of placebos. *Psychopharmacol Bull* 24: 18–20.

Stokes PE, Stoll PM, Shamoian CA, Patton MJ (1971): Efficacy of lithium as acute treatment of manic-depressive illness. *Lancet* 1:1319–1325.

Tohen M, Jacobs TG, Grundy SL, Banov MC, McElroy SL, Janicak PG, et al (1999a, May): A double-blind, placebo-controlled study of olanzapine in patients with acute bipolar mania. Paper presented at the annual meeting of the American Psychiatric Association, Washington, DC.

Tohen M, Sanger TM, McElroy SL, Tollefson GD, Chengappa KNR, Daniel DG, et al (1999b): Olanzapine versus placebo in the treatment of acute mania. *Am J Psychiatry* 156:702–709.

Van Berkestijn H, van der Meulen LM-H, Fletge F, Dols L, van den Hoofdakker R (1990): RS 86 in manic disorder. *Biol Psychiatry* 27:109–112.

Zornberg G, Pope HG Jr (1993): Treatment of depression in bipolar disorder: New directions for research. *J Clin Psychopharmacol* 13:397–408.