Special Article

Validity of Clinical Trials of Antidepressants

**Frederic M. Quitkin, M.D., Judith G. Rabkin, Ph.D., Jessica Gerald, B.A.,
John M. Davis, M.D., and Donald F. Klein, M.D.**

Objective: Recent reports have criticized the design of antidepressant studies and have questioned their validity. These critics have concluded that antidepressants are no better than placebo treatment and that their illusory superiority depends on methodologically flawed studies and biased clinical evaluations. It has been suggested that the blind in randomized trials is penetrable—since clinician's guesses exceed chance—and that only active placebo can appropriately camouflage the difference between drug and placebo response. Furthermore, evidence has been cited to suggest that psychotherapy is as effective as antidepressants in both the acute and maintenance treatment of depression. These positions are often accepted as valid and have been broadly discussed in both the lay press and scientific literature. The purpose of this review is to reassess the cited data that support these assertions. **Method:** The authors examined the specific studies that were cited in these reports, evaluated their methodology, and conducted aggregate analyses. **Results:** Analyses of the original sources failed to substantiate 1) that standard antidepressants are no more effective than placebo, 2) that active placebo offers an advantage over inactive placebo, or 3) that substantial evidence of a medication bias is suggested by raters' treatment guesses exceeding chance. The authors also note that some researchers have suggested that the interpretation of psychotherapy trials can be complicated by "allegiance effects." **Conclusions:** The issue of bias or allegiance effects for both antidepressant and psychotherapy research is real. Investigators of all orientations must guard against potential bias. However, studies cited as supporting the questionable validity of antidepressant trials fail upon closer examination to support assertions that these trials are invalid.

(Am J Psychiatry 2000; 157:327–337)

A recent series of reports in the professional literature and lay press have critiqued the design of antidepressant studies and have questioned antidepressants' true efficacy (1–10). It has been suggested that antidepressants are only modestly effective, and no more so than psychotherapy or even placebo treatment. While these critics also have challenged phenomenologically based diagnoses and psychopharmacology in general,

only their assessment of antidepressant efficacy for unipolar depression is considered here.

We examine the work of Fisher and Greenberg and their colleagues (1–6), as well as that of Antonuccio et al. (7, 8), Moncrieff et al. (9), and Kirsch and Sapirstein (10), all of whom present basically similar viewpoints. The reports of Fisher and Greenberg and their colleagues (1–6) will be examined more closely, since they are the most extensive contributors. Their work has been broadly referred to in the general media (11–19) and has appeared in various clinical discussions (20–22).

In broad outline, their position has three anchors, the first of which is that placebo treatment alone results in clinically significant improvement. Their sec-

ond point is that by and large antidepressants are no better than placebo and that their illusory superiority hinges on methodologically flawed studies and biased clinical evaluations (1–5). In particular, it is asserted that the integrity of the placebo-controlled, double-blind randomized trial is undermined by a penetrable blind that results in biased ratings by drug-favoring clinicians (4). They assert that clinicians' guesses about treatment type exceed chance because clues from side effects of the active drug are poorly camouflaged by inactive placebo (4). This penetration of the double-blind permits more favorable rating of patients presumed to be on active drug. As a corollary to these assumptions, it is argued that active placebo would reduce rater bias (2). The third point of Fisher and Greenberg et al. is that psychotherapy is as effective as antidepressants in both the acute and maintenance treatment of depression (3). Overall, they conclude that valid evidence for antidepressant efficacy is lacking (5).

This perspective continues to resurface in the media, as seen in a recent *New York Times* op-ed column (16). In it, the columnist argues that even the slight advantage reported for antidepressants over placebo "may be illusory, according to researchers like Roger Greenberg, a psychologist….Because all psychiatric drugs have side effects….both patients and researchers invariably see through the double blind, according to Dr. Greenberg." Further evidence that this position is not confined to fringe faction coverage comes from a recent "News Focus" published in *Science*: "It is suggested that there are challenges to the scientific basis of much of the multibillion dollar market for antidepressant drugs" (23).

We suggest that these arguments rest on inaccurate literature citations and are misleading. We are concerned that these conclusions may discourage depressed people from seeking effective treatment. It has been documented that depressive illness has a low rate of treatment in the general population. As Hirschfeld and colleagues noted in a major policy paper, depressed patients "are being seriously undertreated. The vast majority of patients with chronic major depression receive inappropriate or inadequate treatment, or are given no treatment at all" (24, p. 333). Negative press portrayals will not improve this situation.

While most psychiatrists have a balanced view regarding the advantages and limitations of antidepressants, they may be unfamiliar with the details of the research literature supporting antidepressant efficacy as well as the basis for the criticisms leveled against it. It is likely that patients will ask questions about proposed antidepressant treatment, especially when they read negative reports in authoritative sources like *The New York Times*. This review may be helpful in facilitating informed discussions of the pros and cons of antidepressant treatment and questions about true drug efficacy.

ANTIDEPRESSANT EFFICACY: WHAT IS THE EVIDENCE?

The psychiatric research literature suggests that antidepressants are effective for 50%–60% of patients with DSM-IV unipolar depression and that placebo treatment is effective for 20%–30%, with higher placebo response rates for patients with mild depression (25). In a comprehensive 1999 study prepared for the Agency for Health Care Policy and Research (26), an expert multidisciplinary panel reviewed more than 80 randomized, controlled trials of marketed antidepressants that had lasted at least 6 weeks, including those of the newer antidepressant drugs. Outcome was based on a modified intention-to-treat analysis (number of trial completers rated as improved divided by total number of subjects initially randomized). This intention-to-treat analysis assumes that all dropouts are treatment failures, thus giving a conservative lower limit to estimates of treatment utility. In this review, the reported response rates were 50% for antidepressants and 32% for placebo among patients with major depression. These rates are similar to those cited by Thase (25).

In a meta-analysis of treatment with tricyclics for late-life depression (27), a 7–9-point improvement in score on the Hamilton Depression Rating Scale over that of placebo was observed, which approximates a 30% difference in a categorical global measure. These overviews approximate the global advantage of imipramine versus placebo reported by Klerman and Cole (28) more than 30 years earlier. They rated 65% of 550 drug-treated and 31% of 459 placebo-treated patients as having responded. Indeed, this is approximately the position cited by Fisher and Greenberg (6), who say that "proponents of medication usually conclude that about one-third of patients do not show improvement with drug treatment, one third display improvement with placebos, while the remaining third are believed to show improvements with medications that would not have been achieved with placebo" (pp. 118–119).

The disagreement between proponents of medication and the antidepressant skeptics concerns the validity of these findings. The skeptics suggest that without clinician bias, antidepressants would appear about as effective as placebos. The skeptics cite three bodies of evidence: 1) "blinder" studies in which an experimental drug is compared to both placebo treatment and a so-called "standard" drug, where the standard drug is no more effective than placebo; 2) studies in which clinician and patient ratings of perceived outcomes differ (5); and 3) studies in which inert rather than active placebos are used, compromising the blind—as suggested by raters' treatment guesses exceeding chance—and thus inflating the drug effect (3, 4).

TABLE 1. Antidepressant Efficacy Studies That Compared an Experimental Drug With Both a Standard Drug and Placebo ("Blinder" Studies)

Study[a]	Total N	Standard Drug				New Drug				Placebo			Analysis	
		Drug	N	Responders[b] N	%	Drug	N	Responders[b] N	%	N	Responders[b] N	%	χ^2 (df=2)	p
Dominguez et al.	59	Imipramine	23	18	78	Amoxapine	22	18	82	14	9	64	1.5	0.50
Escobar et al.	40	Imipramine	15	14	93	Trazodone	13	8	62	12	6	50	6.6	0.04
Fabre et al.	28	Imipramine	10	6	60	Trazodone	9	5	56	9	1	11	5.5	0.06
Feighner	45	Imipramine	18	5	28	Trazodone	17	9	53	10	0	0	8.4	0.02
Gershon et al.	263	Imipramine	100	47	47	Trazodone	91	43	47	72	18	25	17.6	<0.01
Goldberg and Finnerty	114	Amitriptyline	34	23	68	Trazodone	42	34	81	38	24	63	3.3	0.19
Goldberg et al. (second study)	272	Amitriptyline	87	59	68	Trazodone	92	57	62	93	34	37	20.4	<0.01
Kiev and Okerson	61	Imipramine	20	16	80	Amoxapine	19	16	84	22	20	91	1.0	0.60
Kellams et al.	28	Imipramine	10	2	20	Trazodone	9	5	56	9	0	0	7.6	0.02
Mann et al.	28	Imipramine	9	2	22	Trazodone	10	7	70	9	1	11	8.2	0.02
Rickels and Case	145	Amitriptyline	46	36	78	Trazodone	46	33	72	53	18	34	24.0	<0.01
Rickels et al.	108	Imipramine	40	32	80	Amoxapine	43	34	79	25	14	56	5.5	0.06
Smith	74	Imipramine	24	14	58	Amoxapine	26	18	69	24	10	42	3.9	0.14
Steinbook et al.	23	Imipramine	9	8	89	Amoxapine	10	9	90	4	4	100	0.5	0.79
Trapp et al.	30	Imipramine	10	6	60	Trazodone	10	6	60	10	4	40	1.1	0.59
Van Der Velde	74	Imipramine	26	16	62	Maprotiline	23	16	70	25	5	20	13.9	<0.01
Total	1,392		481	304	63		482	318	66	429	168	39	85.1	<0.01

[a] Citations listed in the meta-analysis of Greenberg et al. (1).
[b] Defined as those having at least a 50% improvement in score on the Hamilton Depression Rating Scale or, if data not available, a change in CGI score to 1 or 2.

Blinder Studies

This term refers to a three-armed study with an experimental drug, standard drug, and placebo. Standard drugs (e.g., imipramine and amitriptyline) are so identified when they have been widely studied and used, and their efficacy is considered clearly established. It has been suggested that in blinder studies, "there is somewhat less vested interest in establishing the efficacy of the standard drugs" (1, p. 664). Consequently, the response rate of the standard drug in this setting is claimed to be a more valid estimate of its real effect than would be observed in two-armed studies in which investigator enthusiasm may engender bias. In their 1989 book (2), Greenberg and Fisher reviewed 16 blinder studies, while in their 1992 article (1), 22 studies were cited. Our analysis examines the original 16 studies because the six additional blinder studies (29–34) were uninterpretable, since two had no placebo group, one had no outcome data, and three lasted only 21 days (a duration of 6 weeks is needed to see specific drug effects [35, 36]). The 16 studies discussed by Fisher and Greenberg in their meta-analysis are outlined in table 1.

In their meta-analysis (1), Greenberg et al. calculated the effect sizes of the blinder studies and stated that they were "two to four times less than those previously reported" (p. 666), and that 66% of the studies showed no difference between standard drug and placebo. The effect size refers to the difference in the effectiveness of a treatment of interest minus that of a control condition (standard drug or placebo).

Our review of their meta-analysis indicates that in eight of the 16 studies, medication was statistically superior to placebo (p≤0.05) and approached significance (p=0.06) in two other studies (table 1). In 14 of 16 studies, a higher percentage of subjects given the standard medication than of those given placebo responded. In the two aberrant cases, placebo response rates were 91% and 100%, so no treatment could be superior. A sign test yields p=0.004, two-tailed, for the superiority of the standard drug over placebo. Results of a Mantel-Haenszel chi-square analysis (37) that combined standard drug and placebo response data from all studies in table 1 and used a 50% decrease in Hamilton depression scale score (or, if unavailable, a CGI of 1 or 2) as the criterion for treatment response were highly significant (χ^2=68.4, df=1, p≤0.0001). A consistency test did not suggest significant interstudy differences (χ^2=2.9, df=1, n.s.). Contrary to the blinder studies hypothesis, the effect size for the standard drugs (26%, 95% confidence interval [CI]=0.20–0.32) was equivalent to that of the new drugs (27%, 95% CI=0.21–0.33). The effect sizes of the standard and new drugs were not "one-half to one-quarter of those previously reported" as Fisher and Greenberg stated (1), and their equivalency fails to support the suggestion that this paradigm offers a more valid assessment of the standard drug.

Their thesis has another basic flaw. Pharmaceutical companies sponsored the 16 blinder studies to obtain approval from the Food and Drug Administration (FDA) for amoxapine, trazodone, and maprotiline. FDA approval requires evidence that the candidate drug is superior to placebo. Inclusion of a standard drug provides information about the new drug's relative efficacy, while differences between standard drug and placebo outcome help to validate the diagnostic specificity of the sample studied. A study that fails to demonstrate a difference between standard drug and

TABLE 2. Comparison of Patient- and Clinician-Rated Improvements From Baseline to Week 6 in a Placebo-Controlled Study of Phenelzine Efficacy (42)

Measure	Improvement in Rating Scale Score		Analysis		
	Mean	SD	t	df	p
SCL-90 (patient-rated)			−8.195	238	<0.001
Placebo (N=127)	0.18	0.81			
Phenelzine (N=113)	1.08	0.89			
CGI (clinician-rated)			−9.67	254	<0.001
Placebo (N=136)	0.57	1.10			
Phenelzine (N=120)	1.85	1.00			

placebo outcome would raise questions about the sample's diagnostic composition and thus would be considered by the FDA to be a "failed study" (38, 39). If the standard drug actually was ineffective, having approximated placebo in 66% of the studies, sample appropriateness would have been suspect, and the FDA would have withheld drug approval. Leber has referred to "the importance of internal controls to demonstrate what we now call 'assay sensitivity'…to deal with sampling variation" (38, 39). However, the primary point is that the conclusion of Greenberg et al., that standard drugs have minimal benefit, fails to gain support after reanalysis of the blinder studies (table 1).

Doctor-Patient Differences in Perceived Outcomes

The antidepressant skeptics also propose that patient self-ratings, because of a greater freedom from drug bias, are more valid than clinician outcome ratings. We examined the empirical basis for this conclusion and find it flawed.

In their 1992 article (1), Greenberg et al. cited a meta-analysis by Lambert et al. (40) that included studies with both a self-rating scale and a clinician-rated scale in which self-reports were said to provide "a significantly smaller effect size than did clinician rating scales" (p. 665). However, Lambert and colleagues had cautioned that the self- and doctor-rated scales may actually measure different aspects of depression "and therefore should not be expected to yield comparable results." These and other prudent remarks were not cited by Greenberg et al. in their critique.

Concerns about patient-clinician rating scale discrepancies antedate Lambert. Kellner et al. (41) noted that "observer rating scales may be more sensitive in discriminating drug and placebo than self-rating scales."

Do patients who clinicians consider improved consider themselves less depressed? Since self-ratings were rarely included in the 16 blinder studies, we analyzed data from our own trials to contrast magnitude of change derived from patient- and clinician-rated scales. In one data set, patients were randomly assigned to placebo (N=135) or phenelzine (N=120) in a 6-week trial; scores from the clinician-rated Clinical Global Impression and the patient-rated SCL-90 depression subscale were compared (42). Paired t tests

were used to estimate the difference in magnitude of change between baseline and week 6 assessments for the two measures. As shown in table 2, both clinician and patient ratings showed significantly greater improvement for phenelzine-treated patients over placebo. The correlations between doctor and patient ratings at week 6 within each treatment group were quite high (placebo group: r=0.70, df=126, p<0.001; phenelzine group: r=0.78, df=113, p<0.001). With respect to effect size, on the SCL-90 depression subscale, the active drug versus placebo effect size was 0.94. The effect size based on the clinician-rated Clinical Global Impressions change scale was 1.04. Both are extremely robust; anything above 0.80 is considered a strong effect (43). Thus, evidence of a disparity in which drug treatment was only favored according to clinician ratings and not patient self-ratings of improvement was not demonstrated.

Inert Versus Active Placebos

Several groups have suggested that the use of inert placebos is inadequate, asserting that drug side effects enable identification of active drug recipients (3, 9, 44). The rationale for giving an active placebo would be to create drug-mimicking side effects to improve the blind, thus increasing the placebo response rate so that it approximates drug response rate. This would suggest that observed drug benefit is just placebo effect plus effect stemming from patient or clinician attribution. Only a higher response rate with active placebo, compared to rates with inert placebos, would be consistent with this thesis.

Thomson reviewed 68 double-blind studies of tricyclics that used an inert placebo and seven that used an active placebo (44). He found drug efficacy was demonstrated in 59% of studies that employed inert placebo, but only 14% of those that used active placebo (χ^2=5.08, df=1, p=0.02). This appears to demonstrate that in the presence of a side-effect-inducing control condition, placebo cannot be discriminated from drug, thus affirming the null hypothesis. (Thomson's review identified six studies that used active placebo but counted two treatment groups in one study as if independent [45–50].) In addition to the studies Thompson cited, four additional studies are discussed in the more recent "active placebo" reviews (3, 9). These 10 studies are outlined in table 3 (45–54).

Does the use of active placebo increase the placebo response rate? This is not the case. After pooling data from those studies in which a judgment could be made about the proportion of responders, it was found that 22% of patients (N=69 of 308) given active placebos were rated as responders. To adopt a conservative stance, one outlier study (50) with a low placebo response rate of 7% (N=6 of 90) was eliminated because its placebo response rate was unusually low (typical placebo response rates in studies of depressed outpatients are 25%–35%). Even after removing this possibly aberrant placebo group, the aggregate response

TABLE 3. Antidepressant Efficacy Studies With Active Placebo Control Condition

Study	Year	Diagnosis	Treatment and Dose	Duration of Treatment	N	Responders N	Responders %	Design Shortcomings or Findings That Support Antidepressant Efficacy
Friedman et al. (45)	1966	Psychotic depression not otherwise specified	Imipramine (200 mg/day)	3 weeks	36	16	44	No significant difference was found between groups, but study flaws include low dose and short trial duration; high placebo response inexplicable for patients with psychotic depression
			Placebo[a]		26	9	35	
Hollister et al. (46)	1964	Depression not otherwise specified	Imipramine (171 mg/day)	3 weeks	27	—	—	Superior Inpatient Multiphasic Scale scores for imipramine and amitriptyline groups support antidepressant efficacy
			Amitriptyline (157 mg/day)		35	—	—	
			Placebo[a]		31	—	—	
Hussain (47)	1970	Criteria not discussed	Amitriptyline/ perphenazine	Not given	15	11	73	Antidepressant response superior to placebo (χ^2=4.48, df=1, p<0.03)
			Perphenazine		15	11	73	
			Placebo[a]		19	7	37	
Uhlenhuth and Park (48)	1963	Primary depression not otherwise specified	Imipramine (150 mg/day)	4-week crossover	22	12	55	Drug efficacy supported by improvement in morale loss score; treatment response difference in the correct direction (χ^2=2.58, df=1, p<0.11)
			Placebo[a]		20	6	30	
Wilson et al. (49)	1963	Depression not otherwise specified	Imipramine (150 mg/day)	5 weeks	6	5	83	Antidepressant efficacy suggested despite study flaw of small group numbers
			ECT (6 sessions)		6	6	10	
			Placebo[b]		6	2	33	
Daneman (50)	1961	Neurotic and psychotic depressive reaction not otherwise specified	Imipramine (200 mg/day)	At least 4 weeks	78	62	79	Antidepressant response superior to placebo (χ^2=98.97, df=1, p<0.001)
			Placebo[a]		90	6	7	
Fahy et al. (51)	1961	Moderately severe depression not otherwise specified	Imipramine (100 mg/day)	3 weeks	16	5	31	Flawed by low dose, short duration, and low power; poor ECT response makes poor imipramine response uninterpretable
			ECT (6 sessions)		17	6	35	
			Placebo[b]		17	2	12	
McLean and Hakstian (52)	1979	Depression (per Feighner criteria)	Psychotherapy	11 weeks	44	11	25	Results uninterpretable because 37 dropouts were replaced; outcome determined by mailed questionnaire after drug discontinuation
			Relaxation		43	12	28	
			Behavior therapy		42	21	50	
			Amitriptyline (150 mg/day)		49	12	24	
Weintraub and Aronson (53)	1963	Nonschizophrenic depression not otherwise specified	Imipramine (150 mg/day)	4 weeks	37	21	57	Antidepressant response superior to placebo (χ^2=7.10, df=1, p<0.008)
			Placebo[a]		32	8	25	
Friedman (54)	1975	Depression (172 of the patients diagnosed with neurotic or reactive depression)	Amitriptyline (150–200 mg/ day) with minimal or marital therapy	12 weeks (no medication last 2 weeks)	98	45	46	Antidepressant response superior to placebo (χ^2=5.57, df=1, p<0.02)
			Placebo[a] with minimal or marital therapy		98	29	30	

[a] Atropine.
[b] Anesthesia.

rate was 29% (N=63 of 218), typical of an inactive placebo. The active placebo theory gains no support from these data.

Closer scrutiny suggests that the "failure" of these 10 early studies to find typical drug-placebo differences is attributable to design errors that characterize studies done during psychopharmacology's infancy. Eight of the 10 studies had at least one of four types of methodological weaknesses: inadequate sample size, inadequate dose, inadequate duration, and diagnostic heterogeneity. The flaws in medication prescription that characterize these studies are outlined in table 3. In fact, in spite of design measurement and power problems, six of these 10 studies still suggested that antidepressants are more effective than active placebo.

In summary, these reviews failed to note that the active placebo response rate fell easily within the rate observed for inactive placebo, and the reviewers relied on pioneer studies, the historical context of which limits them (3, 9, 44).

Double-Blind Guesses

As further evidence that antidepressant trials are invalid, reference has been made to a study done by our group at Columbia, in which blind raters' guesses exceeded chance (55). In one review of this area, 22 other studies are cited in which it is suggested that treatment guesses collected before breaking the study blind exceed chance (4). This evidence is offered to show that the side effects of the active drug were responsible for this breaking the blind. The possibility that clinical outcome also contributes to guessing correctly is not considered. Clinicians are likely to assume that responders were getting active drug. So if 60% respond on drug and 30% on placebo, merely guessing drug for all responders and placebo for all nonresponders would exceed the chance expectation for correct guesses.

We examined the original sources of the 22 studies cited; only eight were found to include actual clinician guesses (listed in table 4). Others either lacked original data or gave anecdotal impressions. Furthermore, only three of these 22 studies included antidepressant medications, and most of the drugs (e.g., chlorpromazine, meprobamate) today would not be used to treat depression. The relevance of these studies is thus doubtful, particularly when coupled with the aforementioned low response rate noted with active placebo.

ARE PSYCHOSOCIAL TREATMENTS MORE EFFECTIVE THAN DRUGS?

Several reports have suggested that psychotherapy is superior to drugs in treating depression (3, 7, 8, 10). In one review it was asserted that "depression specific psychotherapies more often than not eclipsed the results obtained with drugs" (3). Eight trials that compared a specified type of psychotherapy (for depres-

sion) to antidepressants, reported in seven studies, were included in this review. Three of these trials indicated that psychotherapy was equivalent to medication in fostering improvement (56–58), and five indicated that psychotherapy was superior to the drug in promoting substantial change (56, 58–61). The other reviews of this area (7, 8, 10) used a somewhat different approach, but their conclusions are similar. Space limitations require us to focus on one review, but similar problems are noted in the others and have previously been discussed (62).

Several points should be noted. First is the need for a "psychotherapy placebo." A necessary first step before conducting treatment efficacy comparisons is to demonstrate a treatment's superiority to placebo. Does the treatment produce benefits over and above nonspecific factors such as coming for treatment, having a professional ally, etc.? In medication evaluation, the comparison placebo treatment plays this role. There has been substantial controversy over the appropriate parallel for the evaluation of psychotherapy. Klein (62) has critically assessed the logic of using pill placebo-case management as an adequate control for both medication and psychotherapy.

The relationship between medication and psychotherapy with regard to demonstration of specific benefit is not symmetrical. Standard medications have been shown to be superior to placebo under appropriately controlled circumstances. However, such systematic assessment of specific value is only a recent feature of psychotherapy studies.

Studies that contrast psychotherapy and a standard drug are analogous to studies of a standard drug versus a putatively active new medication. A simple direct comparison of the two medications is not adequate. Even if the two treatments were equivalent, the new treatment's efficacy would not be established. Without calibrating the placebo response rate in any sample, two treatments might appear equally effective when, in fact, neither was having an effect.

An equally relevant criticism is that the data cited do not actually show that psychotherapy was superior to pharmacotherapy. Studies by Blackburn et al. (56) and Rush et al. (63) were cited as examples of psychotherapy's superiority. Blackburn et al. (56) studied two patient groups: hospital-based outpatients and general practice patients, some of whom had been symptomatic for only 2 weeks. Cognitive therapy, drug of choice (amitriptyline or clomipramine), or the combination was administered. All treatment-response differences were observed in the general practice group. Treatment responses in the hospital-based groups were equivalent. Blackburn et al. noted that "poor compliance" or dysphoria associated with adverse "social and economic conditions" seen in general practice may have lowered the medication response in the primary care group (56); however, these cautionary remarks were not mentioned.

Rush et al. (63) contrasted cognitive therapy with imipramine during a 12-visit trial. Before discontinu-

TABLE 4. Studies That Purportedly Support the Penetrability of the Double-Blind Design

Study[a]	Drug(s) Studied	Cohort	New Data Presented?	Guesses Exceed Chance?	Treatment Prediction Rates
Rabkin et al., 1986	Imipramine versus phenelzine and placebo	Subjects with depression (per Research Diagnostic Criteria)	Yes	Yes	Doctor and patient guesses of drug versus placebo exceeded chance, but patients' ability to distinguish phenelzine from imipramine did not
Rickels et al., 1965	Chlordiazepoxide and meprobamate versus placebo, nialamide, and other "inactive drugs"	Anxious and depressed outpatients	Yes	Yes (drug); No (placebo)	After 4 weeks (total N=376)—drug: 68% (N=156 of 228); placebo: 51% (N=75 of 148)
Stallone et al., 1973	Lithium versus placebo	Patients with mood disorders	Yes	Nurse 1: Yes (lithium); no (placebo); Nurse 2: No (lithium or placebo)	Nurse 1—lithium: 90% (N=27 of 30); placebo: 50% (N=8 of 16); Nurse 2—lithium: 60% (N=18 of 30); placebo: 46% (N=6 of 13)
Henker et al., 1979	Methylphenidate versus placebo	Hyperactive boys	Yes	Yes (methylphenidate); No (placebo)	Methylphenidate: 74% (N=81 of 110); placebo: 59% (N=65 of 110)
Weiss et al., 1971	Chlorpromazine, dextroamphetamine, and methylphenidate, each versus placebo	Hyperactive children	Yes	Yes	Numbers not reported; chlorpromazine and methylphenidate predictions 100% correct; dextroamphetamine guesses slightly worse; for the three placebo groups, 80%–90% correctly identified by doctors
Werry and Sprague, 1974	Methylphenidate versus placebo	Hyperactive children	Yes	Yes	Physicians correctly guessed whether drug or placebo 74% of the time (N=31 of 42)
Marini et al., 1976	Lithium versus placebo	Aggressive delinquents	Yes	No	None of predictions by three raters exceeded chance for lithium (40.6%, 53.89%, and 40%) or placebo (72.4%, 60%, and 50%)
Margraf et al., 1991	Alprazolam versus imipramine and placebo	Patients with panic disorder	Yes	Yes	Placebo versus active drug—placebo: 72% (N=13 of 18); active drug: 95% (N=36 of 38); imipramine versus alprazolam—imipramine: 88% (N=15 of 17); alprazolam: 79% (N=15 of 19)
Munjack et al., 1989	Alprazolam versus propranolol and placebo	Patients with panic disorder	Yes	Yes (alprazolam); No (propranolol or placebo)	Alprazolam: 85% (N=11 of 13); propranolol: 38% (N=6 of 16); placebo: 50% (N=7 of 14)
Rickels et al., 1966	Meprobamate versus placebo	Neurotic outpatients	Yes	Yes (meprobamate); No (placebo)	After 6 weeks (total N=91)—meprobamate: 78% (N=36 of 46); placebo: 60% (N=27 of 45)
Rickels et al., 1970	Meprobamate versus placebo	Anxious outpatients	Yes	Yes (meprobamate); No (placebo)	After 6 weeks—meprobamate: 76% (N=55 of 72); placebo: 50% (N=33 of 66)
Engelhardt et al., 1969	Chlorpromazine versus promazine and placebo	Patients with schizophrenia	Yes	Yes (promazine, chlorpromazine); N/A (placebo)	Promazine: 60% (N=66 of 109); chlorpromazine: 45% (N=46 of 103); placebo: 60% (N=59 of 99)
Hurst et al., 1973[b]	Amphetamine, placebo, secobarbital	Normal subjects	Yes	Yes (amphetamine); No (placebo or secobarbital)	Only amphetamine guesses exceeded chance
Kleber et al., 1983	Imipramine versus placebo	Opiate users being treated with methadone	Yes	Nurses: Yes; Patients: No	Nurses—imipramine: 55% (N=17 of 31); placebo: 76% (N=19 of 25); Patients—imipramine: 65% (N=13 of 20); placebo: 63% (N=15 of 24)
Adelman and Compas, 1977	"Stimulants"	Children with learning problems	No	—[c]	—[c]
Braunstein and Moscone—cited by Jospe, 1978	Placebo	Not stated	No	—[c]	—[c]
Baker and Thorpe, 1957	Mepazine versus placebo	Incontinent, psychotic patients	No	—[c]	—[c]
Greenberg et al., 1992	Antidepressants (review of drug trials)	Not stated	No	—[c]	—[c]
Letemendia and Harris, 1959	Nicotinic acid versus placebo	Patients with chronic psychosis	No	—[c]	—[c]
Lipman et al., 1966	Chlordiazepoxide, atropine, placebo	Anxious neurotic patients	No	—[c]	—[c]
Ritvo et al., 1983	Fenfluramine versus placebo	Autistic subjects	Yes	—[c]	—[c]
Wing, 1956	Reserpine versus placebo	Psychotic inpatients	No	—[c]	—[c]

[a] Citations listed in the review of Fisher and Greenberg (4).   [b] Four experiments.   [c] Study did not involve systematic guesses.

ing imipramine, treatment effects were equal. After imipramine discontinuation, a subsequent evaluation noted cognitive therapy's superior outcome. Rush et al. (63) cautioned, "a question may be raised as to whether the level of improvement of some of the pharmacotherapy patients reflected a reduction of drug dosage." In the Blackburn and Rush studies, lack of a placebo group and other design problems noted by the investigators themselves make psychotherapy and pharmacotherapy contrasts less informative.

Antonuccio et al. (7) cited a meta-analysis by Robinson et al. (64) and noted psychotherapy had a statistically significant mean effect size that was 0.13 larger than that for drug therapy. The cautionary statement by Robinson et al. ("after investigator allegiance had been controlled...the advantage of psychotherapy...was no longer statistically significant" [p. 39]) was included only as an afterthought.

The original report of the NIMH collaborative study of depression is cited as another example of the apparent equivalence of medication and two types of psychotherapy: interpersonal therapy and cognitive behavior therapy (65). However, reanalysis of these data indicated imipramine's superiority for severely depressed patients, compared to the psychotherapy and placebo conditions (66, 67). For the most severely ill group, interpersonal therapy was superior to placebo on only one measure, while cognitive behavior therapy was indistinguishable from both interpersonal therapy and placebo.

Post hoc stratification by diagnosis was performed in another reanalysis (68, 69). Patients with atypical depression responded poorly to imipramine (at a rate indistinguishable from placebo). Our group and others have previously documented the mediocre effect of imipramine for this diagnostic group (42). Cognitive behavior therapy was superior to imipramine (interpersonal therapy showed trend superiority to imipramine [p=0.08]). Underlining the relevance of a placebo control for drug-psychotherapy studies is the fact that neither interpersonal therapy nor cognitive behavior therapy were superior to placebo for patients with atypical depression. In the nonatypical patients, imipramine treatment was superior to both placebo and cognitive behavior therapy but not interpersonal therapy (which was superior to placebo). Patients with atypical depression account for 25%–40% of outpatient cases of depression (70). Since previous psychotherapy-drug studies did not consider the poor tricyclic response of patients with atypical depression, this may account for the mediocre outcome of medications in some of the studies.

The issue of "investigator allegiance" to a particular method of psychotherapy is receiving increasing attention in the psychotherapy literature. The concept is equivalent to the problem of "clinician bias" in drug research. A team of investigators led by Luborsky et al. (71) examined 24 psychotherapy studies that included 29 comparisons. Study criteria for this review included clear-cut diagnosis and randomization, but was unique in that a rating of therapeutic allegiance was also included. They concluded that "the results both of past analyses and the present one imply that the researcher's allegiance tends to be strongly associated with the differential outcomes of the treatments...the combination of the allegiance measures shows a very large association with treatment outcomes (r=0.85!)" (p. 103).

In an accompanying editorial, Shaw (72) stated that "their original research leaves me with few questions that, indeed, researchers are biased or at least potentially biased in their impact on the results of their own trials" (p. 131). In a separate commentary, Hollon (73) noted that "allegiance effects are ubiquitous in treatment outcome research" (p. 107). The possibility of bias affecting placebo-controlled medication trials is real, but in psychotherapy research there is not any blind; therefore, allegiance effects (and the resulting potential for bias) may be even a greater problem.

Another consideration concerns drug augmentation. In clinical practice if drug-treated patients have not improved after 6–8 weeks, thyroid hormone, amphetamine, lithium, or an additional antidepressant would be added to the initial regimen (74). This augmented medication trial, which would be completed in the time required for a single psychotherapy trial, is likely to increase the proportion of responders by another 15%–20% (74). There is no equivalent to such augmentation for nonpharmacologic treatment except, of course, the addition of drug treatment.

CONCLUSIONS

Our examination of the original source material cited by antidepressant skeptics suggests that these critiques of the antidepressant literature are largely unsubstantiated. Findings from antidepressant research are usually valid; these medications are often specifically useful.

We do not dismiss the likelihood of biased ratings. It is possible that a reasonable scientist, in evaluating an ambiguous clinical outcome, may use different criteria when assessing patients guessed to be taking drug or placebo. That distinct side effect profiles may permit some piercing of the double blind is probable, but this does not appear to invalidate results. Davis et al. (75) reviewed the effect sizes of a variety of medical treatments. Their meta-analysis suggested that drugs used in psychiatry were two to three times more effective than placebo and as effective as penicillin for pneumococcal pneumonia or streptomycin for tuberculosis.

Investigators of all allegiances should not be complacent about the possibility of bias affecting their studies. However, our review of this antidepressant-critical research suggests that overall, these critiques may be allegiance driven. Cautionary remarks that were included in the original sources frequently were omitted, and doubtful studies with design shortcomings were portrayed as definitive. Misleading informa-

tion about the role of antidepressants and their relative benefit compared to psychotherapy may have a deleterious public health impact. Patients and therapists may assume medication is unnecessary. The potentially serious public health implications of this type of reporting were seen when patients, after the lay press erroneously discussed a possible link between fluoxetine and suicide, generalized their fear and avoidance to other antidepressants (76).

Another issue raised by this review is underlined by the apparent discrepancy between the primary sources we reviewed and the presentation of these data. Peer reviews more easily identify errors of commission than omission of data or absence of a fair representation of alternative hypotheses. The peer-reviewed articles included in this article had errors that were apparently missed by the original journal reviewers. A more balanced view of controversial areas, particularly where allegiance may color interpretation, will occur if articles in these areas are accompanied by critiques and rebuttals by adherents to other positions.

Rather than repeatedly examining old studies for their flaws and strengths, it would be constructive to move to the next stage of research, including the formation of research consortia and the conduct of "mega-trials" with investigators of varied theoretical orientations. In order to clarify relative efficacy in drug-psychotherapy contrasts, minimal requirements should include intrastudy placebo calibration, the participation of expert psychopharmacologists and psychotherapists, and rigorous experimental design and measurement. Multisite studies done by mutually monitored collaborating investigators, some with allegiance to psychotherapy and others to pharmacotherapy, that result in findings repeated at different sites would be most convincing. Such large-scale research programs, as Andrews (77) has noted, are relatively "immune to the methodological and political hazards that can beset small randomized, controlled trials and meta-analyses." The empirical findings thus generated would provide a firm basis for development and implementation of clinical guidelines.

REFERENCES

1. Greenberg RP, Bornstein RF, Greenberg MD, Fisher S: A meta-analysis of antidepressant outcome under "blinder" conditions. J Consult Clin Psychol 1992; 60:664–669
2. Greenberg RP, Fisher S: A second opinion: rethinking the claims of biological psychiatry, in The Limits of Biological Treatments for Psychological Distress: Comparisons With Psychotherapy and Placebo. Edited by Fisher S, Greenberg RP. Hillsdale, NJ, Lawrence Erlbaum Associates, 1989, pp 309–336
3. Greenberg RP, Fisher S: Examining antidepressant effectiveness: findings, ambiguities, and some vexing puzzles. Ibid, pp 1–37
4. Fisher S, Greenberg RP: How sound is the double-blind design for evaluating psychotropic drugs? J Nerv Ment Dis 1993; 181:345–350
5. Greenberg RP, Bornstein RF, Zborowski MJ, Fisher S, Greenberg MD: A meta-analysis of fluoxetine outcome in the treatment of depression. J Nerv Ment Dis 1994; 182:547–551
6. Fisher S, Greenberg R (eds): From Placebo to Panacea: Putting Psychiatric Drugs to the Test. New York, John Wiley & Sons, 1997
7. Antonuccio DO, Danton WG, DeNelsky GY: Psychotherapy versus medication for depression: challenging the conventional wisdom with data. Professional Psychol: Res and Practice 1995; 26:574–585
8. Antonuccio DO, Danton WG, DeNelsky GY, Greenberg RP, Gordon JS: Raising questions about antidepressants. Psychother Psychosom 1999; 68:3–14
9. Moncrieff J, Wessely S, Hardy R: Meta-analysis of trials comparing antidepressants with active placebos. Br J Psychiatry 1998; 172:227–231
10. Kirsch I, Sapirstein G: Listening to Prozac but hearing placebo: a meta-analysis of antidepressant medications, in How Expectancies Shape Experience. Edited by Kirsch I. Washington, DC, American Psychological Association, 1999, pp 303–320
11. Goleman D: Psychologists dispute value of antidepressants. New York Times, Nov 29, 1995, p C10
12. Blakeslee S: Placebos Prove so Powerful Even Experts Are Surprised. New York Times (Science Times), Oct 13, 1998, p C5
13. Schoofs M: Healing head trips: doctors explore the astonishing power of placebos. Village Voice, May 6–12, 1998, p 46
14. Holmes L: Treatment e-journal is reborn. About.com, June 29, 1998 (http://mentalhealth.miningco.com/library/weekly/aa062998.htm)
15. Horgan J: Science triumphant? not so fast. New York Times, Jan 19, 1998, p A17
16. Horgan J: Placebo nation. New York Times, March 21, 1999, section 4, p 15
17. Day M: Mostly in the mind. New Scientist, July 11, 1998 (http://www.newscientist.com/ns/980711/nantidepressant.html)
18. American Psychological Association Press Release: Placebo effect accounts for fifty percent of improvement in depressed patients taking antidepressants. Self-Help & Psychology Magazine, May 21, 1998 (http://www.shpm.com/articles/depress/antidprs.html)
19. American Psychological Association Press Release: Antidepressant medications: miracle drugs or placebos with a buzz? APA Public Communications, July 16, 1998 (http://www.apa.org/releases/debate.html)
20. Munoz RF, Hollon SD, McGrath E, Rehm LP, VandenBos GR: On the AHCPR depression primary care guidelines: further considerations for practitioners. Am Psychol 1994; 49:42–61
21. Young SM: The use of placebos in psychiatry: a response to the draft document prepared by the TRI-Council Working Group. J Psychiatry Neurosci 1996; 21:235–238
22. Healy D: Commentary: meta-analysis of trials comparing antidepressants with active placebos. Br J Psychiatry 1998; 172:232–234
23. Enserink M: Can the placebo be the cure? Science 1999; 284:238–240
24. Hirschfeld R, Keller M, Panico S, Arons B, Barlow D, Davidoff F, Endicott J, Froom J, Goldstein M, Gorman J, Guthrie D, Marek R, Maurer T, Meyer R, Phillips K, Ross J, Schwenk T, Sharfstein S, Thase M, Wyatt R: The National Depressive and Manic-Depressive Association Consensus Statement on the Undertreatment of Depression. JAMA 1997; 277:333–340
25. Thase M: How should efficacy be evaluated in randomized clinical trials of treatments for depression? J Clin Psychiatry 1999; 60(suppl 4):23–31
26. Treatment of depression—newer pharmacotherapies: summary, evidence report/technology assessment number 7. Rockville, Md, Agency for Health Care Policy and Research, March 1999 (http://www.ahcpr.gov/clinic/deprsumm.htm)
27. Klawansky S: Meta-analysis on the treatment of depression in late life, in Diagnosis and Treatment of Depression in Late Life: Results of the NIH Consensus Development Conference. Edited by Schneider LS, Reynolds CF III, Liebowitz BD, Freidhoff AJ. Washington, DC, American Psychiatric Press, 1994, pp 333–352

28. Klerman GL, Cole JO: Clinical pharmacology of imipramine and related compounds. Pharmacol Rev 1965; 17:101–141

29. Ather SA, Ankier SI, Middleton RSW: A double-blind evaluation of trazodone in the treatment of depression in the elderly. Br J Clin Pract 1985; 39:192–199

30. Blacker R, Shanks NJ, Chapman N, Davey A: The drug treatment of depression in general practice: a comparison of nocte administration of trazodone with mianserin, dothiepin, and amitriptyline. Psychopharmacology (Berl) 1988; 95:18–24

31. Klieser E, Lehmann E: Experimental comparison between the effect of standardized trazodone-amitriptyline and placebo treatment in vitalized depressive patients. Psychopharmacology (Berl) 1988; 95:3–5

32. Klieser E, Lehmann E: Experimental examination of trazodone. Clin Neuropharmacol 1989; 12(suppl 1):S18–S24

33. McNair DM, Kahn RJ, Frankenthaler LM, Faldetta LL: Amoxapine and amitriptyline, I: relative speed of antidepressant action. Psychopharmacology (Berl) 1984; 83:129–133

34. McNair DM, Kahn RJ, Frankenthaler LM, Faldetta LL: Amoxapine and amitriptyline, II: specificity of cognitive effects during brief treatment of depression. Psychopharmacology (Berl) 1984; 83:134–139

35. Quitkin FM, Rabkin JG, Ross D, McGrath PJ: Duration of antidepressant drug treatment: what is an adequate trial? Arch Gen Psychiatry 1984; 41:238–245

36. Donovan SJ, Quitkin FM, Stewart JW, Ocepek-Welikson K, Harrison W, McGrath PJ, Nunes EV, Wager S, Tricamo E: Duration of antidepressant trials: clinical and research implications. J Clin Psychopharmacol 1994; 14:64–66

37. Mantel N, Haenszel W: Statistical aspect of the analysis of data from retrospective studies of disease. J Natl Cancer Inst 1959; 22:719–748

38. Leber P: The future of controlled clinical trials. Psychopharmacol Bull 1993; 27:3–8

39. Leber P: There is an alternative to the randomized-controlled trials. Psychiatr Annals 1995; 25:401–403

40. Lambert MJ, Hatch DR, Kingston MD, Edwards BC: Zung, Beck, and Hamilton rating scales as measures of treatment outcome: a meta-analytic comparison. J Consult Clin Psychol 1986; 54:54–59

41. Kellner R, Simpson GM, Sheffield BF: The value of self-rating scales in drug trials (proceedings). Psychopharmacol Bull 1977; 13:42–43

42. Quitkin FM, Stewart JW, McGrath PJ, Tricamo E, Rabkin JG, Ocepek-Welikson K, Nunes E, Harrison W, Klein DF: Columbia atypical depression: a subgroup of depressives with better response to MAOI than to tricyclic antidepressants or placebo. Br J Psychiatry Suppl 1993; 21:30–34

43. Cohen J: Statistical Power Analysis for Behavioral Sciences, revised ed. San Diego, Academic Press, 1977

44. Thomson R: Side effects and placebo amplification. Br J Psychiatry 1982; 140:64–68

45. Friedman AS, Granick S, Cohen HW, Cowitz B: Imipramine (Tofranil) vs placebo in hospitalized psychotic depressives (a comparison of patients' self-ratings, psychiatrists' ratings and psychological test scores). J Psychiatr Res 1966; 4:13–36

46. Hollister L, Overall J, Johnson M, Pennington V, Katz G, Shelton J: Controlled comparison of amitriptyline, imipramine, and placebo in hospitalized depressed patients. J Nerv Ment Dis 1964; 139:370–375

47. Hussain Z: Drugs in depressive illness. Br Med J 1970; 1:482

48. Uhlenhuth E, Park L: The influence of medication (imipramine) and doctor in relieving depressed psychoneurotic outpatients. J Psychiatr Res 1963; 2:101–122

49. Wilson I, Vernon J, Sandiler M: A controlled study of treatment of depression. J Neuropsychiatry 1963; 4:331–337

50. Daneman E: Imipramine in office management of depressive reactions (a double blind clinical study). Dis Nerv Syst 1961; 22:213–217

51. Fahy P, Imlah N, Harrington J: A controlled comparison of electroconvulsive therapy, imipramine and thiopentone sleep in depression. J Neuropsychiatry 1963; 4:310–314

52. McLean PD, Hakstian AR: Clinical depression: comparative efficacy of outpatient treatments. J Consult Clin Psychol 1979; 47:818–836

53. Weintraub W, Aronson H: Clinical judgment in psychopharmacological research. J Neuropsychiatry 1963; 4:65–70

54. Friedman AS: Interaction of drug therapy with marital therapy in depressive patients. Arch Gen Psychiatry 1975; 32:619–637

55. Rabkin JG, Markowitz JS, Stewart JW, McGrath PJ, Harrison W, Quitkin FM, Klein DF: How blind is blind? assessment of patient and doctor medication guesses in a placebo-controlled trial of imipramine and phenelzine. Psychiatry Res 1986; 19:75–86

56. Blackburn IM, Bishop S, Glen AI, Whalley LJ, Christie JE: The efficacy of cognitive therapy in depression: a treatment trial using cognitive therapy and pharmacotherapy, each alone and in combination. Br J Psychiatry 1981; 139:181–189

57. Murphy GE, Simmons AD, Wetzel RD, Lustman PJ: Cognitive therapy and pharmacotherapy: single and together in the treatment of depression. Arch Gen Psychiatry 1983; 41:33–41

58. Weissman MM, Prusoff BA, DiMascio A, Neu C, Goklaney M, Klerman GL: The efficacy of drugs and psychotherapy in the treatment of acute depressive episodes. Am J Psychiatry 1979; 136:555–558

59. Bellack AS, Hersen M, Himmelhoch J: Social skills training compared with pharmacotherapy and psychotherapy in the treatment of unipolar depression. Am J Psychiatry 1981; 138:1562–1567

60. Beutler LE, Scogin F, Kirkish P, Schretlen D, Corbishley A, Hamblin D, Meredith K, Potter R, Levenson A, Bamford CR: Group cognitive therapy and alprazolam in the treatment of depression in older adults. J Consult Clin Psychol 1987; 55:550–556

61. McLean P, Hackstian A: Clinical depression: comparative efficacy of outpatient treatments. J Consult Clin Psychol 1979; 47:818–836

62. Klein DF: Preventing hung juries about therapy studies. J Consult Clin Psychol 1996; 64:81–87

63. Rush AJ, Beck AT, Kovacs M, Hollon S: Comparative efficacy of cognitive therapy and pharmacotherapy in the treatment of depressed outpatients. Cognitive Therapy Res 1977; 1:17–37

64. Robinson LA, Berman JS, Neimeyer RA: Psychotherapy for the treatment of depression. Psychol Bull 1990, 108:30–49

65. Elkin I, Shea MT, Watkins JT, Imber SD, Sotsky SM, Collins JF, Glass DR, Pilkonis PA, Leber WR, Docherty JP, Fiester SJ, Parloff MB: National Institute of Mental Health Treatment of Depression Collaborative Research Program: general effectiveness of treatments. Arch Gen Psychiatry 1989; 46:971–982

66. Elkin I, Gibbons RD, Shea MT, Sotsky SM, Watkins JT, Pilkonis PA, Hedeker D: Initial severity and differential treatment outcome in the National Institute of Mental Health Treatment of Depression Collaborative Research Program. J Consult Clin Psychol 1995; 63:841–847

67. Klein DF, Ross DC: Re-analysis of the National Institute of Mental Health Treatment of Depression Collaborative Research Program General Effectiveness Report. Neuropsychopharmacology 1993, 8:241–252

68. Sotsky SM, Simmens SJ: Pharmacotherapy response in atypical depression: findings from the NIMH Treatment of Depression Collaborative Research Program, in DSM-IV Sourcebook, vol 4. Edited by Widiger TA, Frances AJ, Pincus HA, Ross R, First MB, Davis WW. Washington, DC, American Psychiatric Association, 1996, pp 143–156

69. Stewart JW, Garfinkel R, Nunes EV, Klein DF: Atypical depression and treatment response in the NIMH Treatment of Depression Collaborative Research Program. J Clin Psychopharmacol 1998; 18:429–434

70. Zisook S, Shuchter SR, Gallagher T, Sledge P: Atypical depression in an outpatient population. Depression 1993; 1:268–274

71. Luborsky L, Diguer L, Seligman D, Rosenthal R, Krause E, Johnson S, Halperin G, Bishop M, Berman JS, Schweizer E: The researcher's own therapy allegiances: a "wild card" in comparisons of treatment efficacy. Clin Psychol—Sci and Practice 1999; 6:95–106

72. Shaw B: How to use the allegiance effect to maximize competence and therapeutic outcomes. Clin Psychol— Sci and Practice 1999; 6:131–132

73. Hollon S: Allegiance effects in treatment research: a commentary. Clin Psychol—Sci and Practice 1999; 6:107–112

74. McGrath PJ, Stewart JW, Nunes EV, Ocepek-Welikson K, Rabkin JG, Quitkin FM, Klein DF: A double-blind crossover trial of imipramine and phenelzine for outpatients with treatment-refractory depression. Am J Psychiatry 1993; 150:118–123

75. Davis JM, Wang Z, Janicak PG: Psychiatric medication compares favorably to drugs used in internal medicine: a meta-analysis (abstract). Schizophr Res 1995; 15:148

76. Teicher MH, Glod C, Cole JO: Emergence of intense suicidal preoccupation during fluoxetine treatment. Am J Psychiatry 1990; 147:207–210

77. Andrews G: Randomized controlled trials in psychiatry: important but poorly accepted. Br Med J 1999; 319:562–564