Placebo Response in Studies of Major Depression
Variable, Substantial, and Growing

B. Timothy Walsh, MD

Stuart N. Seidman, MD

Robyn Sysko, BA

Madelyn Gould, PhD

CONTENTIOUS DEBATE CONTINUES about the ethics of placebo administration in clinical trials in general and in psychiatry in particular.[1-4] Supporters of the use of placebo maintain that, in the study of many illnesses, the assignment of individuals to receive placebo is scientifically critical and that such assignment is ethical if it is not expected to contribute to increased mortality or irreversible morbidity.[5-8] Critics of the use of placebo assert that after the efficacy of 1 or more treatments for an illness has been established, it is not ethical to assign subjects in a research study to an intervention that is expected to be less efficacious.[9-11] Such critics suggest that the scientific aims of placebo use (eg, determination of the utility of new interventions) may be achieved with alternative designs, including the use of active comparators (interventions with established efficacy) or previously collected information on the rate of placebo response.[1,12-14]

The study of pharmacologic treatments for major depressive disorder (MDD) has traditionally used assignment to placebo in randomized controlled trials. Based on trials using this design, the Food and Drug Administration (FDA) has approved for marketing antidepressant agents that are among the most frequently prescribed medications in the world and are widely viewed as effective for MDD. Paradoxically, the clinical success of these medications has contributed to ethical concerns about the continued use of placebo in trials to assess the efficacy of new antidepressant medications.[11] It has been argued that the use of placebo assignment in such trials remains justified because the demonstra-

Context Intense debate persists about the need for placebo-controlled groups in clinical trials of medications for major depressive disorder (MDD). There is continuing interest in the development of new medications, but because effective antidepressants are already available, ethical concerns have been raised about the need for placebo groups in new trials.

Objective To determine whether the characteristics of placebo control groups in antidepressant trials have changed over time.

Data Sources and Study Selection We searched MEDLINE and PsychLit for all controlled trials published in English between January 1981 and December 2000 in which adult outpatients with MDD were randomly assigned to receive medication or placebo. Seventy-five trials met our criteria for inclusion.

Data Extraction Data were extracted from the articles by 2 of the authors and discrepancies were resolved via discussion and additional review by a third author.

Data Synthesis The mean (SD) proportion of patients in the placebo group who responded was 29.7% (8.3%) (range, 12.5%-51.8%). Most studies examined more than a single active medication, and, in the active medication group with the greatest response, the mean (SD) proportion of patients responding was 50.1% (9.0%) (range, 31.6%-70.4%). Both the proportion of patients responding to placebo and the proportion responding to medication were significantly positively correlated with the year of publication (for placebo: n=75; r=0.45; 95% confidence interval [CI], 0.25-0.61; P<.001; for medication: n=75; r=0.26; 95% CI, 0.03-0.46; P=.02). The association between year of publication and response rate was more statistically robust for placebo than medication.

Conclusions The response to placebo in published trials of antidepressant medication for MDD is highly variable and often substantial and has increased significantly in recent years, as has the response to medication. These observations support the view that the inclusion of a placebo group has major scientific importance in trials of new antidepressant medications and indicate that efforts should continue to minimize the risks of such studies so that they may be conducted in an ethically acceptable manner.

JAMA. 2002;287:1840-1847   www.jama.com

Author Affiliations and Financial Disclosures are listed at the end of this article.

Corresponding Author and Reprints: B. Timothy Walsh, MD, Department of Psychiatry, College of Physicians and Surgeons of Columbia University, 1051 Riverside Dr, Unit 98, New York, NY 10032.

tion of efficacy compared with placebo eliminates the risk to public health of approving ineffective medications and because there is no evidence that treatment delay or assignment to placebo results in permanent harm.[8,15] Along with these ethical arguments, investigators have suggested that, in recent years, it has become more difficult to demonstrate specific antidepressant effects compared with placebo, suggesting that the response to placebo may have increased.[16-18] In this context of persistent ethical and scientific concerns about the use of placebo in studies of MDD, we examined data on the response to placebo and medication in controlled trials of adult outpatients with MDD published between 1981 and 2000 to determine whether the characteristics of such studies and the response to placebo have changed.

METHODS
Sources of Data and Criteria for Review

We collected peer-reviewed articles describing randomized, placebo-controlled trials of medication for outpatients with unipolar MDD published between 1981 and 2000. Studies published between 1981 and 1990 were taken primarily from an extensive, previously published literature review.[19] To create a list of articles published between 1991 and 2000, we used similar methods. We specified key words, including *placebo* and generic names of all putative antidepressant medications, and conducted multiple computer searches using MEDLINE and PsycLit.

To be included in this review, articles were required to meet the following criteria: (1) reported in English; (2) published between January 1981 and December 2000; (3) primarily composed of outpatients with MDD according to research diagnostic criteria (RDC), *Diagnostic and Statistical Manual of Mental Disorders, Third Edition*, *Diagnostic and Statistical Manual of Mental Disorders, Revised Third Edition*, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*, *International Classification of Diseases, Ninth Revision (ICD-9)*,

or *International Classification of Diseases, 10th Revision (ICD-10)* criteria for major depression (we excluded trials that focused on bipolar disorder, but did include several early studies[20-22] that contained a small number of bipolar depressed patients); (4) had at least 20 patients in the placebo group; (5) lasted at least 4 weeks; (6) randomly assigned patients to receive a putative antidepressant drug or drugs and placebo and assessed patients under double-blind conditions; and (7) reported the total number of patients assigned to placebo and medication group(s) and the number who had responded as determined by a reduction of at least 50% in their score on the Hamilton Rating Scale for Depression (HRSD)[23] and/or a Clinical Global Impression (CGI)[24] rating of markedly or moderately improved (CGI score of 1 or 2).

Data were extracted from the articles by 2 of the authors (S.S., R.S.) and discrepancies were resolved via discussion and additional review by a third author (B.T.W.). For each study, we calculated the response rate for the placebo and medication group(s) by dividing the number of patients in each group who had responded according to these criteria by the number of patients assigned to the group. In a few studies,[25-29] only the number of patients randomized who returned for at least 1 visit was given, and this number was used as the denominator. If response data were provided only on study completers, we assumed that noncompleters had not responded.

Statistical Methods

To examine the strength of the linear relationship between continuous variables, we used Pearson correlation coefficients (2-tailed) and linear regression. To examine differences between groups in continuous variables, we used the *t* statistic. To assess whether the proportion of studies that used certain design features (eg, placebo lead-in before randomization) had changed linearly with time, we determined the number of studies that used the feature during each of the four 5-year pe-

riods between 1981 and 2000 and calculated a χ^2 statistic for linearity.[30] Tests of statistical significance regarding fractions of patients responding were also performed using an arcsin transformation of the fraction, specifically, $2 \times$ arcsin. The results of the analyses using the untransformed and transformed data were virtually identical, and only results of analyses of untransformed response rates are reported herein. Effect size was calculated as the difference between the arcsin-transformed response rate on active medication and the arcsin-transformed response rate on placebo.[31] Statistical calculations were performed using SPSS for Windows, version 10 (SPSS Inc, Chicago, Ill).

RESULTS
Study Characteristics

Seventy-five trials of medication for adult outpatients with MDD meeting our criteria for review were identified.[20-22,25-29,32-99] Data are presented as mean (SD). The trial length averaged 45.4 (17.9) days (range, 28-168 days) and increased significantly with year of publication (n=74; r=0.45; 95% confidence interval [CI], 0.25-0.61; P<.001). The average number of patients per group was 83.8 (59.7) (range, 20-336) for placebo and 85.0 (60.0) (range, 20-335) for medication; neither changed significantly with year of publication (placebo: n=75; r=0.12; 95% CI, –0.11 to 0.34; P=.30; medication: n=75; r=0.10; 95% CI, –0.13 to 0.32; P=.41). The diagnostic systems used have changed over time. For example, the proportion of studies using RDC declined from 28.6% between 1981 and 1985 to 0% between 1996 and 2000 (χ^2_1=8.5, P=.004). The diagnostic systems generally require a minimum duration of 2 weeks of symptoms. Seventeen trials (22.7%) required a longer duration of illness, usually 4 weeks, to merit a diagnosis of MDD. Most trials (74.7%) used a 1- to 2-week initial placebo lead-in period; patients who responded during this time were not randomized. Neither the number of studies that required more than 2 weeks of symptoms nor the number

using placebo lead in changed significantly with time ($\chi^2_1 = 1.8$ and 1.7, and $P = .18$ and .19, respectively).

Seventy-four studies used the HRSD to measure severity of depression, and 67 reported which version was used to assess severity of depression at study entry: the 17-item version was used in 33 studies (49%), the 21-item version in 32 (48%), and 24-item version in 2 (2.9%). The proportion of studies using the 17-item version did not change significantly with time ($\chi^2_1 = 0.31$, $P = .58$). Sixty-eight studies required a minimum score on the HRSD for study entry. To compare the HRSD scores across studies using different versions of the HRSD, we prorated the HRSD score by dividing the score by the number of items in the HRSD version and multiplied by 17. Among the 63 studies that provided sufficient information to perform this calculation, the prorated minimum entry score averaged 16.7 (2.1) (range, 12.9-22.0) and increased with year of publication ($n = 63$; $r = 0.37$; 95% CI, 0.13-0.57; $P = .003$). Forty-five of the studies provided explicit information on whether the use of concomitant medications (usually chloral hydrate or a benzodiazepine) was permitted during the trial for treatment of anxiety or insomnia. In 38 (84%) of these 45 trials, concomitant medications were permitted. The proportion of trials permitting concomitant medication did not change significantly with time ($\chi^2_1 = 0.72$, $P = .39$).

Patient Characteristics

On average, 60.1% (9.7%) (range, 36.2%-85.0%) of participants were women, and this proportion did not change significantly with year of publication ($n = 63$; $r = 0.00$; 95% CI, –0.25 to 0.25; $P = .99$). The average age was 42.6 (7.3) years (range, 35.5-72.0 years) and increased significantly with year of publication ($n = 61$; $r = 0.30$; 95% CI, 0.05-0.51; $P = .02$). The average observed HRSD score at study entry, prorated for the number of items in the HRSD version, was 22.5 (2.7) (range, 13.8-30.2) and did not change significantly with year of publication ($n = 61$; $r = –0.10$; 95% CI, –0.34 to 0.16; $P = .43$).

Placebo Response

Fifty-three studies provided information on the number of patients responding to placebo as indicated by a 50% or greater reduction in HRSD score. By this measure, the average proportion of patients per study who responded while receiving placebo was 30.0% (8.0%), and this proportion increased significantly with year of publication ($n = 53$; $r = 0.43$; 95% CI, 0.18-0.63; $P = .001$). Forty-seven studies provided information on the number of patients responding to placebo as indicated by a CGI rating of 1 or 2. By this measure, the average proportion of patients per study who responded while receiving placebo was 32.8% (11.1%), and this proportion increased significantly with year of publication ($n = 47$; $r = 0.53$; 95% CI, 0.29-0.71; $P < .001$).

We were interested in combining data from all studies to examine the relationship between study characteristics and placebo response rate. Twenty-two studies presented placebo response information only on the basis of CGI rating. To combine information from these studies with data from the other 53 studies that provided placebo response information on the basis of a 50% or greater reduction in HRSD score, we estimated the proportion responding according to HRSD as follows. Twenty-five studies reported both the number of patients receiving placebo who responded according to HRSD criteria and the number who responded according to CGI criteria. These proportions were highly correlated ($n = 25$; $r = 0.79$; 95% CI, 0.57-0.90; $P < .001$). We estimated the proportion responding according to HRSD criteria using the equation for the best-fit straight line obtained from the 25 studies that provided data on response according to both measures (HRSD response rate $= 0.658 \times$ CGI response rate $+ 0.0832$). According to this combined HRSD measure, 29.7% (8.3%) (range, 12.5%-51.8%) of patients assigned to placebo responded. The proportion of patients taking placebo responding by this measure was significantly correlated with year of publication ($n = 75$; $r = 0.45$; 95% CI, 0.25-

0.61; $P < .001$) (**FIGURE**), length of the trial ($n = 74$; $r = 0.27$; 95% CI, 0.04-0.47; $P = .02$), and prorated minimum HRSD score required ($n = 63$; $r = 0.27$; 95% CI, 0.02-0.49; $P = .03$). In a multiple linear regression model, with proportion of patients taking placebo responding according to the combined HRSD measure as the dependent variable and year of publication, length of trial, and prorated minimum HRSD score as independent variables, only year of publication remained significant ($\beta = 0.366$, $t_{58} = 2.7$, $P = .009$). There were no statistically significant associations between the proportion of patients taking placebo who responded and prorated average initial HRSD score ($n = 61$; $r = 0.10$; 95% CI, –0.16 to 0.34; $P = .46$), average patient age ($n = 61$; $r = 0.11$; 95% CI, –0.15 to 0.35; $P = .42$), fraction of patients who were women ($n = 63$; $r = 0.09$; 95% CI, –0.16 to 0.33; $P = .49$), presence of a placebo lead-in phase ($t_{73} = 0.35$, $P = .73$), minimum duration of symptoms required being greater than 2 weeks ($t_{73} = 0.23$, $P = .82$), or concomitant medication use ($t_{35} = 0.86$, $P = .39$).

Medication Response

To put the response to placebo in context, we examined the response to active medication in the same 75 studies. The 25 studies that provided information regarding placebo response on the basis of both a 50% reduction in HRSD score and CGI response contained similar information on medication response. Since most studies examined more than 1 active medication, we determined the best-fit straight line describing the relationship between CGI response and HRSD response to the first medication described in each of these studies: HRSD response rate $= 0.744 \times$ CGI response rate $+ 0.0805$. The slope for this equation does not differ significantly from that obtained for placebo response rates ($t_{46} = 0.56$, $P = .72$). For studies that provided response rates only using CGIs, we estimated the proportion responding according to HRSD criteria using this equation.

Sixty studies (80%) assigned more than 1 group of patients to active medi-

cations (average number of active medication groups, 1.97 per study; range, 1-4), and many assigned different groups of patients to receive different doses or preparations of the same medication. For each study with more than 1 active medication group, the greatest actual or estimated response rate using HRSD criteria in one medication group was designated the maximum medication response rate for that study. The 2 classes of medication most frequently studied were the tricyclic antidepressants (TCAs, 43 studies) and the selective serotonin reuptake inhibitors (SSRIs, 33 studies). For each study that examined a TCA and/or an SSRI, we determined the TCA and/or the SSRI response rates; for studies that assigned different groups of patients to different preparations or doses within the same class, we used data from the group with the greatest response rate.

The average maximum proportion of patients responding to medication in a study was 50.1% (9.0%) (range, 31.6%-70.4%). The proportion of patients responding to TCAs was 46.9% (10.6%) (range, 27.5%-65.6%) and to SSRIs was 48.9% (10.3%) (range, 25.0%-70.4%). The proportion of patients responding to medication was correlated with year of publication (for maximum proportion of patients responding to a medication, n=75; $r=0.26$; 95% CI, 0.03-0.46; $P=.02$; for TCA response: n=43; $r=0.29$; 95% CI, −0.01 to 0.54; $P=.06$; for SSRI response: n=33; $r=0.47$; 95% CI, 0.15-0.70; $P=.006$) (Figure). However, when significant covariates were entered into the regression models (trial length for maximum proportion responding and SSRI response; prorated initial HRSD score for TCA response), neither the relationship of medication response with year of publication nor with the covariate remained significant.

We calculated effect size, the difference between the response to medication and the response to placebo, using an arcsin transformation. The average effect size across studies for the maximum proportion of patients responding to a medication was 0.43 (0.22) (for TCA response, 0.38 [0.22];

for SSRI response, 0.40 [0.24]). There were no statistically significant correlations between any of these effect sizes and year of publication (for maximum proportion of patients responding to a medication, n=75; $r=−0.17$; 95% CI, −0.38 to 0.06; $P=.14$; for TCA response: n=43; $r=−0.09$; 95% CI, −0.38 to 0.22; $P=.56$; for SSRI response: n=33; $r=−0.16$; 95% CI, −0.48 to 0.19; $P=.38$).

Time Course

We were interested in determining the week at which a significant difference in the mean HRSD score between medication and placebo groups was first noted. Fifty-three reports found a putative antidepressant to be superior to placebo, and 40 of these provided information, usually a graph, that allowed determination of the first week at which a statistically significant difference ($P<.05$) between mean HRSD scores of medica-

tion and placebo groups was noted. If data on several medications or medication doses were included in one report, we used the week of first detectable medication vs placebo difference for the medication or dose with the greatest overall response rate. A significant drug effect was first noted, on average, 2.40 (1.48) weeks (range, 1-7 weeks) following randomization, and, in 92% (37 of 40 studies), by 4 weeks after randomization. In 19 studies that provided data on the time course of response to TCAs, a medication-placebo difference was noted in 2.53 (1.54) weeks (range, 1-6 weeks), and, in 19 studies with data on SSRIs, a medication-placebo difference was noted in 2.74 (1.63) weeks (range, 1-7 weeks).

COMMENT

This review of 75 placebo-controlled trials of medication for adult outpatients

Figure. Proportion of Patients Assigned to Placebo, Tricyclic Antidepressants (TCAs), and Selective Serotonin Reuptake Inhibitors (SSRIs) Who Showed a 50% or Greater Improvement in Hamilton Rating Scale For Depression Score by Year of Publication
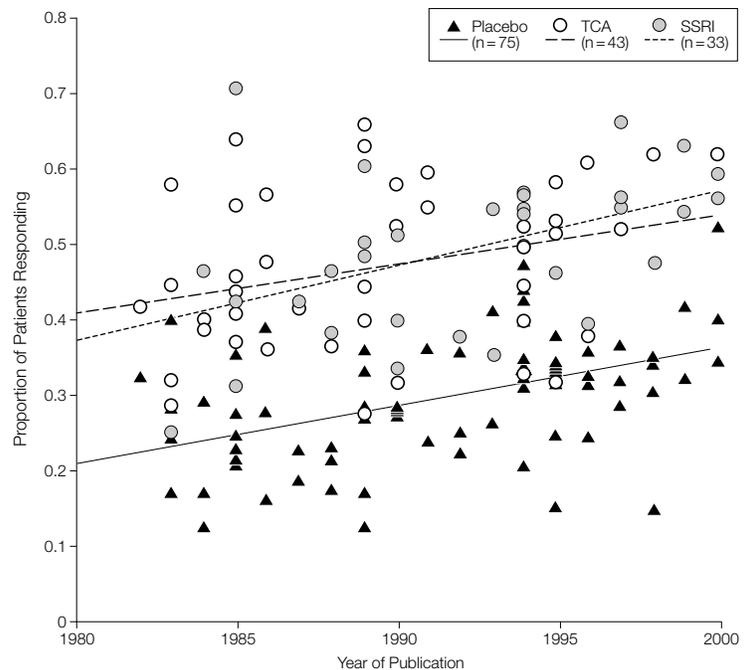
For studies that did not provide data on response using this criterion, an estimated proportion was calculated based on the proportion of patients rated moderately or markedly improved on the Clinical Global Impression (see "Results" section). The lines are the best-fit straight lines describing the relationship between proportion of patients responding and year of publication for placebo ($r=0.45$, n=75, $P<.001$), TCAs ($r=0.29$, n=43, $P=.06$), and SSRIs ($r=0.47$, n=33, $P=.006$).

with MDD published between 1981 and 2000 leads to several conclusions relevant to considering the inclusion of placebo groups in future studies of antidepressant medication. First, the response to placebo across the trials varied, ranging from approximately 10% to more than 50%, and frequently was substantial; in approximately half of the studies, 30% or more of the patients assigned to placebo exhibited a clinically significant improvement. Second, in the last 2 decades, the proportion of patients responding to placebo has clearly increased, at the rate of approximately 7% per decade, and a similar increase has occurred in the fraction of patients responding to active medication. The effect size in published studies has not changed significantly, emphasizing the importance of examining unpublished data to determine whether the increased rate of placebo response may have contributed to the difficulty of establishing the efficacy of new antidepressants.[18,100]

The change in placebo response rate does not appear to be directly explained by changes in study characteristics. Although the diagnostic systems used have shifted with time as psychiatry has adopted the American Psychiatric Association's editions of the *Diagnostic and Statistical Manual of Mental Disorders*, all the systems are close descendants of the RDC, and the specific criteria are extremely similar in the nature and duration of the symptoms required for diagnosis of MDD. Thus, there is little reason to think that the shift from RDC to *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition* during the last 20 years is responsible for changes in response rates. The average age of patients has increased, but the relationship between age and placebo response rate was not significant. The length of randomized controlled trials has increased, and we found, as have others, that the proportion of patients responding to placebo increases with trial length.[101-103] Presumably, this association reflects both the cumulative effects of the nonspecific interventions in-

herent in clinical trials and a longer period during which spontaneous recovery could occur. We also found that, after prorating for the number of items in the HRSDs used in different studies, more recent studies required a higher minimum HRSD score for study entry and that there was a significant positive correlation between minimum entry score and placebo response rate. However, in a multiple regression model that included year of publication, length of trial, and minimum required HRSD score, only year of publication was a significant predictor of placebo response. It is of note that the average placebo response rate of studies that used a placebo lead in did not differ significantly from that of studies that did not use a placebo lead in.[104] Data on other potentially relevant characteristics, such as duration of illness or number of prior episodes of depression, were not reported in a sufficiently similar form across studies to permit evaluation of their impact on placebo response rate.

A recently published meta-analysis[105] found little evidence that the course of patients receiving a placebo can be distinguished from the course of patients receiving no treatment, suggesting that the effect of placebo is usually no greater than the effect of the passage of time. This conclusion was not as clear-cut for studies examining more subjective outcomes such as pain, and the meta-analysis located only a few studies of the short-term treatment of MDD. Therefore, it is unclear whether placebo, when used in studies of MDD, has a greater effect than no treatment. In either case, the current review indicates that the impact of placebo in trials of MDD, even if it reflects no more than the natural course of untreated MDD, has changed in the last 2 decades. In addition, most studies permit the use of sedatives and anxiolytics, so that what is reported as response to placebo typically encompasses the impact of such agents.

Our review found that the response rates to both placebo and active medication in trials of MDD have increased

significantly in the last 20 years and that only year of publication is a significant predictor of placebo response when other potentially relevant factors are taken into account. Some factor or factors associated with the level of placebo response must therefore have changed significantly during this period. Unfortunately, we were not able to determine the identity of these factors. The only characteristic of MDD available across studies was the prorated average initial HRSD score, which, despite an increase in the minimum HRSD score required for study entry, did not change significantly in the last 20 years. Because the HRSD was the sole measure of illness severity, was used in 1 of 3 different forms, and can be difficult to compare across studies,[106] it is possible that the HRSD was not sufficiently sensitive to capture relevant changes in the characteristics of MDD.

In recent years, as effective treatments for depression have become more widely available and socially acceptable, there has been a marked increase in the proportion of the population receiving treatment for depression.[107] The factors responsible for this trend may also have contributed to the participation in clinical trials of individuals with milder, briefer, and more responsive forms of depression. In addition, the methods by which patients are recruited into therapeutic trials have changed. Twenty years ago, most patients were referred by other clinicians or learned of research studies while trying to obtain treatment from established centers. In more recent years, it has become commonplace for investigators to advertise the availability of research trials to the public. It is likely that clinically important characteristics of patients participating in treatment studies have changed as a result of these practices[18] and have contributed to changes in placebo and medication response rates.

The variability of the response of patients with MDD to placebo across studies and the fact that the average rate of response has changed over time strongly argue against the use of a his-

torical standard as a valid benchmark against which to measure the efficacy of new antidepressants. Since the response to placebo is variable, often substantial, and increasing, it is not surprising that in many randomized controlled trials the response associated with placebo is similar to that associated with an established antidepressant.[18,100,108] In addition, the association between medication response rate and year of publication was not as robust as that of placebo. These facts, coupled with the inability to predict accurately the response rate to placebo or medication on the basis of study characteristics, raise serious doubts about the scientific validity of trials of new medications in which there is no placebo group and the only criterion for antidepressant efficacy is a response statistically indistinguishable from that obtained with an accepted antidepressant.[5]

Among trials that ultimately detected a difference between the active medication group and the placebo group, a statistically significant difference in mean HRSD score was apparent soon, usually by 3 weeks and almost always by 4 weeks after randomization. This observation was true even for trials of SSRIs, which may require longer treatment to be fully effective.[109] Thus, although the magnitude of the difference in response between medication and placebo generally increases with length of study,[60,103] the current review found, as have others,[110] that a statistical difference in HRSD score was detectable early. Because an early drug-placebo difference in the HRSD score may only reflect improvement in depressive symptoms, such as sleep disturbance and agitation, a statistically significant difference in mean HRSD scores is likely to be well short of a clinically significant difference. In addition, sufficient data were not generally available to assess the change in effect size during a study, and none of these studies was specifically designed to determine first onset of antidepressant activity.[111] Nonetheless, the observation that a difference between the effects of anti-

depressant medication and placebo can be detected in less than 6 weeks suggests that it may be possible to limit the length of placebo-controlled trials of new antidepressant medications. Such changes would both reduce the duration of time patients receive placebo and help minimize the risks of such treatment.

Our study has several limitations. We examined only published data. For a number of reasons, including the fact that most placebo-controlled trials of antidepressant medications are supported by the pharmaceutical industry, studies in which a significant drug-placebo difference is detected are presumably much more likely to be published than ones in which the responses to placebo and active medication are statistically indistinguishable. Therefore, it seems probable that the placebo response rates in the published studies we examined are lower, and the medication response rates higher, than those in unpublished studies. Several authors[100,101,112] have recently begun to access data submitted to the FDA, much of which is unpublished. Until these and other unpublished studies are examined, uncertainty will persist about changes in the response to placebo and to medication over time. In addition, the variable delays in publication of results could bias the results regarding date of publication. Another limitation is that the results presented herein are based on somewhat arbitrary criteria for clinical response. We used a 50% decrease in HRSD score and/or a CGI improvement score of 1 or 2 as the criteria for response because they are routinely reported and generally accepted, for example, by the FDA. It would be of interest to know whether the results we describe are reflected in other measures of clinical response. In addition, because the focus of this review is on changes in the characteristics of studies of MDD during the last 2 decades, the unit of analysis is the study, not the patient. As a consequence, we have very limited ability to examine the relationship between patient characteristics and

the likelihood of response. For example, we could examine only the relationship between average patient age and average response rate across studies, not the relationship between an individual's age and the likelihood that he or she would respond.

In summary, our review of recent placebo-controlled trials of medication for MDD indicates that the response to placebo is variable but often substantial and has increased in recent years. These data support the view[5,8,15,112] that the inclusion of a placebo group has major scientific importance in trials of new antidepressant medications and indicate that efforts should continue to minimize the risks of such studies so that they may be conducted in an ethically acceptable manner.

REFERENCES

1. Rothman KJ, Michels KB. The continuing unethical use of placebo controls. *N Engl J Med.* 1994;331: 394-398.
2. Rothman KJ. Declaration of Helsinki should be strengthened. *BMJ.* 2000;321:442-445.
3. Enserink M. Psychiatry: are placebo-controlled drug trials ethical? *Science.* 2000;288:416.
4. Al-Khatib SM, Califf RM, Hasselblad V, Alexander JH, McCrory DC, Sugarman J. Medicine: pla-

cebo-controls in short-term clinical trials of hypertension. *Science.* 2001;292:2013-2015.

5. Temple R, Ellenberg SS. Placebo-controlled trials and active-control trials in the evaluation of new treatments, part 1: ethical and scientific issues. *Ann Intern Med.* 2000;19:455-463.

6. Tramer MR, Reynolds DJ, Moore RA, McQuay HJ. When placebo controlled trials are essential and equivalence trials are inadequate. *BMJ.* 1998;317:875-880.

7. Leber P. The use of placebo control groups in the assessment of psychiatric drugs: an historical context. *Biol Psychiatry.* 2000;47:699-706.

8. Quitkin FM. Placebos, drug effects, and study design: a clinician's guide. *Am J Psychiatry.* 1999;156:829-836.

9. Kraemer HC. Statistical analysis to settle ethical issues? *Arch Gen Psychiatry.* 2000;57:327-328.

10. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA.* 2000;284:3043-3045.

11. Michels KB. The placebo problem remains. *Arch Gen Psychiatry.* 2000;57:321-322.

12. Rosner F. The ethics of randomized clinical trials. *Am J Med.* 1987;82:283-290.

13. Freireich EJ, Gehan EA. Historical controls reincarnated: an examination of techniques for clinical therapeutic research. *Cancer Chemother Rep.* 1974;58:623-626.

14. Gehan EA. Comparative clinical trials with historical controls: a statistician's view. *Biomedicine.* 1978;28:13-19.

15. Leber P. Placebo controls: no news is good news. *Arch Gen Psychiatry.* 2000;57:319-320.

16. Klein DF. Improvement of phase III psychotropic drug trials by intensive phase II work. *Neuropsychopharmacology.* 1991;4:251-258.

17. Montgomery SA. The failure of placebo-controlled studies: ECNP Consensus Meeting, September 13, 1997, Vienna, European College of Neuropsychopharmacology. *Eur Neuropsychopharmacol.* 1999;9:271-276.

18. Robinson DS, Rickels K. Concerns about clinical drug trials. *J Clin Psychopharmacol.* 2000;20:593-596.

19. Rush AJ, Crismon ML, Toprac MG, et al. Consensus guidelines in the treatment of major depressive disorder. *J Clin Psychiatry.* 1998;59:73-84.

20. Muijen M, Roy D, Silverstone T, Mehmet A, Christie M. A comparative clinical trial of fluoxetine, mianserin and placebo in depressed outpatients. *Acta Psychiatr Scand.* 1988;78:384-390.

21. Itil TM, Shrivastava RK, Mukherjee S, Coleman BS, Michael ST. A double-blind placebo-controlled study of fluvoxamine and imipramine in out-patients with primary depression. *Br J Clin Pharmacol.* 1983;15(suppl 3):433S-438S.

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

23. Williams JB. A structured interview guide for the Hamilton Depression Rating Scale. *Arch Gen Psychiatry.* 1988;45:742-747.

24. Guy W. *ECDEU Assessment Manual for Psychopharmacology.* Rockville, Md: National Institute of Mental Health; 1976:76-338.

25. Byerley WF, Reimherr FW, Wood DR, Grosser BI. Fluoxetine, a selective serotonin uptake inhibitor, for the treatment of outpatients with major depression. *J Clin Psychopharmacol.* 1988;8:112-115.

26. Bremner JD. A double-blind comparison of Org 3770, amitriptyline, and placebo in major depression. *J Clin Psychiatry.* 1995;56:519-525.

27. Rickels K, Robinson DS, Schweizer E, Marcus RN, Roberts DL. Nefazodone: aspects of efficacy. *J Clin Psychiatry.* 1995;56(suppl 6):43-46.

28. Cohn CK, Robinson DS, Roberts DL, Schwiderski UE, O'Brien K, Ieni JR. Responders to antidepressant drug treatment: a study comparing nefazodone, imipramine, and placebo in patients with major depression. *J Clin Psychiatry.* 1996;57(suppl 2):15-18.

29. Schatzberg AF. Clinical efficacy of reboxetine in major depression. *J Clin Psychiatry.* 2000;61(suppl 10):31-38.

30. Fleiss JL. *Statistical Methods for Rates and Proportions.* New York, NY: John Wiley & Sons; 1981.

31. Cohen J. *Statistical Power Analysis for the Behavioral Sciences.* Hilldale, NJ: Lawrence Erlbaum Associates; 1988.

32. Rickels K, Weise CC, Sandler K, Schless A, Zal M, Norstad N. Nomifensine, imipramine and placebo in depressed outpatients. *Int Pharmacopsychiatry.* 1982;17:73-88.

33. Merideth CH, Feighner JP. A double-blind, controlled evaluation of zimeldine, imipramine and placebo in patients with primary affective disorders. *Acta Psychiatr Scand Suppl.* 1983;308:70-79.

34. Claghorn J, Gershon S, Goldstein BJ, Behrnetz S, Bush DF, Huitfeldt B. A double-blind evaluation of zimeldine in comparison to placebo and amitriptyline in patients with major depressive disorder. *Prog Neuropsychopharmacol Biol Psychiatry.* 1983;7:367-382.

35. Feighner JP, Aden GC, Fabre LF, Rickels K, Smith WT. Comparison of alprazolam, imipramine, and placebo in the treatment of depression. *JAMA.* 1983;249:3057-3064.

36. Giller EJ, Bialos D, Harkness L, Riddle M. Assessing treatment response to the monoamine oxidase inhibitor isocarboxazid. *J Clin Psychiatry.* 1984;45:44-48.

37. Reimherr FW, Wood DR, Byerley B, Brainard J, Grosser BI. Characteristics of responders to fluoxetine. *Psychopharmacol Bull.* 1984;20:70-72.

38. White K, Razani J, Cadow B, et al. Tranylcypromine vs nortriptyline vs placebo in depressed outpatients: a controlled trial. *Psychopharmacology (Berl).* 1984;82:258-262.

39. Cohn JB, Wilcox C. A comparison of fluoxetine, imipramine, and placebo in patients with major depressive disorder. *J Clin Psychiatry.* 1985;46:26-31.

40. Dominguez RA, Goldstein BJ, Jacobson AF, Steinbook RM. A double-blind placebo-controlled study of fluvoxamine and imipramine in depression. *J Clin Psychiatry.* 1985;46:84-87.

41. Hormazabal L, Omer LM, Ismail S. Cianopramine and amitriptyline in the treatment of depressed patients: a placebo-controlled study. *Psychopharmacology (Berl).* 1985;86:205-208.

42. Stark P, Hardison CD. A review of multicenter controlled studies of fluoxetine vs imipramine and placebo in outpatients with major depressive disorder. *J Clin Psychiatry.* 1985;46:53-58.

43. Stewart JW, McGrath PJ, Liebowitz MR, Harrison W, Quitkin F, Rabkin JG. Treatment outcome validation of DSM-III depressive subtypes: clinical usefulness in outpatients with mild to moderate depression. *Arch Gen Psychiatry.* 1985;42:1148-1153.

44. Rickels K, Feighner JP, Smith WT. Alprazolam, amitriptyline, doxepin, and placebo in the treatment of depression. *Arch Gen Psychiatry.* 1985;42:134-141.

45. Lipman RS, Covi L, Rickels K, et al. Imipramine and chlordiazepoxide in depressive and anxiety disorders, I: efficacy in depressed outpatients. *Arch Gen Psychiatry.* 1986;43:68-77.

46. Amsterdam JD, Case WG, Csanalosi E, Singer M, Rickels K. A double-blind comparative trial of zimelidine, amitriptyline, and placebo in patients with mixed anxiety and depression. *Pharmacopsychiatry.* 1986;19:115-119.

47. Mendels J, Schless AP. Comparative efficacy of alprazolam, imipramine, and placebo administered once a day in treating depressed patients. *J Clin Psychiatry.* 1986;47:357-361.

48. Rickels K, Chung HR, Csanalosi IB, et al. Alprazolam, diazepam, imipramine, and placebo in outpatients with major depression. *Arch Gen Psychiatry.* 1987;44:862-866.

49. Wernicke JF, Dunlop SR, Dornseif BE, Zerbe RL. Fixed-dose fluoxetine therapy for depression. *Psychopharmacol Bull.* 1987;23:164-168.

50. Davidson JR, Giller EL, Zisook S, Overall JE. An efficacy study of isocarboxazid and placebo in depression, and its relationship to depressive nosology. *Arch Gen Psychiatry.* 1988;45:120-127.

51. Feighner JP, Boyer WF. Paroxetine in the treatment of depression: a comparison with imipramine and placebo. *Acta Psychiatr Scand Suppl.* 1989;350:125-129.

52. Peselow ED, Stanley M, Filippi AM, Barouche F, Goodnick P, Fieve RR. The predictive value of the dexamethasone suppression test: a placebo controlled study. *Br J Psychiatry.* 1989;155:667-672.

53. Stewart JW, McGrath PJ, Quitkin FM, et al. Relevance of *DSM-III* depressive subtype and chronicity of antidepressant efficacy in atypical depression: differential response to phenelzine, imipramine, and placebo. *Arch Gen Psychiatry.* 1989;46:1080-1087.

54. Versiani M, Oggero U, Alterwain P, et al. A double-blind comparative trial of moclobemide v. imipramine and placebo in major depressive episodes. *Br J Psychiatry Suppl.* October 1989;(6):72-77.

55. Amsterdam JD, Dunner DL, Fabre LF, Kiev A, Rush AJ, Goodman LI. Double-blind, placebo-controlled, fixed dose trial of minaprine in patients with major depression. *Pharmacopsychiatry.* 1989;22:137-143.

56. Conti L, Dell'Osso L. Clinical predictors of response to fluvoxamine, imipramine, and placebo. *New Trends Exp Clin Psychiatry.* 1989;5:221-229.

57. Dunlop SR, Dornseif BE, Wernicke JF, Potvin JH. Pattern analysis shows beneficial effect of fluoxetine treatment in mild depression. *Psychopharmacol Bull.* 1990;26:173-180.

58. Robinson DS, Rickels K, Feighner J, et al. Clinical effects of the 5-HT1A partial agonists in depression: a composite analysis of buspirone in the treatment of depression. *J Clin Psychopharmacol.* 1990;10(suppl 3):67S-76S.

59. Lineberry CG, Johnston JA, Raymond RN, et al. A fixed-dose (300 mg) efficacy study of bupropion and placebo in depressed outpatients. *J Clin Psychiatry.* 1990;51:194-199.

60. Reimherr FW, Chouinard G, Cohn CK, et al. Antidepressant efficacy of sertraline: a double-blind, placebo- and amitriptyline-controlled, multicenter comparison study in outpatients with major depression. *J Clin Psychiatry.* 1990;51(suppl B):18-27.

61. Smith WT, Glaudin V, Panagides J, Gilvary E. Mirtazapine vs amitriptyline vs placebo in the treatment of major depressive disorder. *Psychopharmacol Bull.* 1990;26:191-196.

62. Gelenberg AJ, Wojcik JD, Falk WE, Spring B, Brotman AW, Galvin-Nadeau M. Clovoxamine in the treatment of depressed outpatients: a double-blind, parallel-group comparison against amitriptyline and placebo. *Compr Psychiatry.* 1990;31:307-314.

63. Khan A, Dager SR, Cohen S, Avery DH, Scherzo B, Dunner DL. Chronicity of depressive episode in relation to antidepressant-placebo response. *Neuropsychopharmacology.* 1991;4:125-130.

64. Rickels K, London J, Fox I, Hassman H, Csanalosi I, Weise C. Adinazolam, diazepam, imipramine, and placebo in major depressive disorder: a controlled study. *Pharmacopsychiatry.* 1991;24:127-131.

65. Bakish D, Bradwejn J, Nair N, McClure J, Remick R, Bulger L. A comparison of moclobemide, amitriptyline and placebo in depression: a Canadian multicentre study. *Psychopharmacology (Berl).* 1992;106(suppl):S98-S101.

66. Claghorn JL, Kiev A, Rickels K, Smith WT, Dunbar GC. Paroxetine versus placebo: a double-blind comparison in depressed patients. *J Clin Psychiatry.* 1992;53:434-438.

67. Ose E, Holm P. Moclobemide and placebo in mild major depression: a double-blind randomized trial. *Psychopharmacology.* 1992;106:S114-S115.

68. Mendels J, Johnston R, Mattes J, Riesenberg R. Efficacy and safety of b.i.d. doses of venlafaxine in a dose-response study. *Psychopharmacol Bull.* 1993; 29:169-174.

69. Tollefson GD, Holman SL. Analysis of the Hamilton Depression Rating Scale factors from a double-blind, placebo-controlled trial of fluoxetine in geriatric major depression. *Int Clin Psychopharmacol.* 1993; 8:253-299.

70. Ferguson JM, Mendels J, Manowitz NR, for the Prothiaden Collaborative Study Group. Dothiepin versus doxepin in major depression: results of a multicenter, placebo-controlled trial. *J Clin Psychiatry*. 1994; 55:258-263.

71. Wilcox CS, Cohn JB, Katz BB, et al. A double-blind, placebo-controlled study comparing mianserin and amitriptyline in moderately depressed outpatients. *Int Clin Psychopharmacol.* 1994;9:271-279.

72. Cunningham LA, Borison RL, Carman JS, et al. A comparison of venlafaxine, trazodone, and placebo in major depression. *J Clin Psychopharmacol.* 1994; 14:99-106.

73. Schweizer E, Feighner J, Mandos LA, Rickels K. Comparison of venlafaxine and imipramine in the acute treatment of major depression in outpatients. *J Clin Psychiatry.* 1994;55:104-108.

74. Fontaine R, Ontiveros A, Elie R, et al. A double-blind comparison of nefazodone, imipramine, and placebo in major depression. *J Clin Psychiatry.* 1994;55: 234-241.

75. Thompson C. Management of depression in real-life settings: knowledge gained from large-scale clinical trials. *Int Clin Psychopharmacol.* 1994;9(suppl 3): 21-25.

76. Heiligenstein JH, Tollefson GD, Faries DE. Response patterns of depressed outpatients with and without melancholia: a double-blind, placebo-controlled trial of fluoxetine versus placebo. *J Affect Disord.* 1994;30:163-173.

77. UK Moclobemide Study Group. A multicentre comparative trial of moclobemide, imipramine and placebo in major depressive disorder. *Int Clin Psychopharmacol.* 1994;9:109-113.

78. Chouinard G, Saxena BM, Nair NP, et al. A Canadian multicentre placebo-controlled study of a fixed dose of brofaromine, a reversible selective MAO-A inhibitor, in the treatment of major depression. *J Affect Disord.* 1994;32:105-114.

79. Fabre LF, Abuzzahab FS, Amin M, et al. Sertraline safety and efficacy in major depression: a double-blind fixed-dose comparison with placebo. *Biol Psychiatry.* 1995;38:592-602.

80. Nair NP, Amin M, Holm P, et al. Moclobemide and nortriptyline in elderly depressed patients: a randomized, multicentre trial against placebo. *J Affect Disord.* 1995;33:1-9.

81. Claghorn JL, Lesem MD. A double-blind placebo-controlled study of Org 3770 in depressed outpatients. *J Affect Disord.* 1995;34:165-171.

82. Laakman G, Faltermaier-Temizel M, Bossert-Zaudig S, Baghai T, Lorkowski G. Treatment of depressive outpatients with lorazepam, alprazolam, amytriptyline and placebo. *Psychopharmacology (Berl).* 1995;120:109-115.

83. Khan MC. A randomised, double-blind, placebo-controlled, 5-weeks' study of Org 3770 (mirtazapine) in major depression. *Hum Psychopharmacol.* 1995; 10:S119-S124.

84. Halikas JA. Org 3770 (mirtazapine) versus trazodone: a placebo controlled trial in depressed elderly patients. *Hum Psychopharmacol.* 1995;10:S125-S133.

85. Wilcox CS, Ferguson JM, Dale RN, Heiser JF. A double-blind trial of low- and high-dose ranges of gepirone-ER compared with placebo in the treatment of depressed outpatients. *Psychopharmacol Bull.* 1996;32:335-342.

86. Claghorn JL, Earl CQ, Walczak DD, et al. Fluvoxamine maleate in the treatment of depression: a single-center, double-blind, placebo-controlled comparison with imipramine in outpatients. *J Clin Psychopharmacol.* 1996;16:113-120.

87. Rickels K, Derivan A, Kunz N, Pallay A, Schweizer E. Zalospirone in major depression: a placebo-controlled multicenter study. *J Clin Psychopharmacol.* 1996;16:212-217.

88. Stassen HH, Angst J, Delini-Stula A. Delayed onset of action of antidepressant drugs? survey of results of Zurich meta-analyses. *Pharmacopsychiatry.* 1996;29:87-96.

89. Lydiard RB, Stahl SM, Hertzman M, Harrison WM. A double-blind, placebo-controlled study comparing the effects of sertraline versus amitriptyline in the treatment of major depression. *J Clin Psychiatry.* 1997; 58:484-491.

90. Cunningham LA, for the Venlafaxine XR 208 Study Group. Once-daily venlafaxine extended release (XR) and venlafaxine immediate release (IR) in outpatients with major depression. *Ann Clin Psychiatry.* 1997;9:157-164.

91. Thase ME, for the Venlafaxine XR 209 Study Group. Efficacy and tolerability of once-daily venlafaxine extended release (XR) in outpatients with major depression. *J Clin Psychiatry.* 1997;58:393-398.

92. Schweizer E, Rickels K, Hassman H, Garcia-Espana F. Buspirone and imipramine for the treatment of major depression in the elderly. *J Clin Psychiatry.* 1998;59:175-183.

93. Rudolph RL, Fabre LF, Feighner JP, Rickels K, Entsuah R, Derivan AT. A randomized, placebo-controlled, dose-response trial of venlafaxine hydrochloride in the treatment of major depression. *J Clin Psychiatry*. 1998;59:116-122.

94. Lapierre YD, Silverstone P, Reesal RT, et al. A Canadian multicenter study of three fixed doses of controlled-release ipsapirone in outpatients with moderate to severe major depression. *J Clin Psychopharmacol.* 1998;18:268-273.

95. Schrader E, Meier B, Brattstroem A. Hypericum treatment of mild-moderate depression in a placebo-controlled study: a prospective, double-blind, randomized, placebo-controlled, multicentre study. *Hum Psychopharmacol.* 1998;13:163-169.

96. Rudolph RL, Feiger AD. A double-blind, random-ized, placebo-controlled trial of once-daily venlafaxine extended release (XR) and fluoxetine for the treatment of depression. *J Affect Disord.* 1999;56:171-181.

97. Silverstone PH, Ravindran A, for the Venlafaxine XR 360 Study Group. Once-daily venlafaxine extended release (XR) compared with fluoxetine in outpatients with depression and anxiety. *J Clin Psychiatry.* 1999;60:22-28.

98. Ruether E, Degner D, Munzel U, et al. Antidepressant action of sulpride: results of a placebo-controlled double-blind trial. *Pharmacopsychiatry.* 1999;32:127-135.

99. Stahl SM. Placebo-controlled comparison of the selective serotonin reuptake inhibitors citalopram and sertraline. *Biol Psychiatry.* 2000;48:894-901.

100. Ghosh K, Kramer MS. A review of placebo-controlled efficacy trials for antidepressants as evaluated by the FDA. *New Clin Drug Eval Unit Program Abstracts.* 1999;poster No. 36.

101. Khan A, Warner HA, Brown WA. Symptom reduction and suicide risk in patients treated with placebo in antidepressant clinical trials: an analysis of the Food and Drug Administration database. *Arch Gen Psychiatry.* 2000;57:311-317.

102. Quitkin FM, Rabkin JD, Markowitz JM, Stewart JW, McGrath PJ, Harrison W. Use of pattern analysis to identify true drug response: a replication. *Arch Gen Psychiatry.* 1987;44:259-264.

103. Quitkin FM, McGrath PJ, Stewart JW, et al. Chronological milestones to guide drug change: when should clinicians switch antidepressants? *Arch Gen Psychiatry.* 1996;53:785-792.

104. Trivedi MH, Rush H. Does a placebo run-in or a placebo treatment cell affect the efficacy of antidepressant medications? *Neuropsychopharmacology.* 1994;11:33-43.

105. Hrobjartsson A, Gotzsche PC. Is the placebo powerless? an analysis of clinical trials comparing placebo with no treatment. *N Engl J Med.* 2001;344:1594-1602.

106. Zitman FG, Mennen MF, Griez E, Hooijer C. The different versions of the Hamilton Depression Rating Scale. *Psychopharmacol Ser.* 1990;9:28-34.

107. Olfson M, Marcus SCF, Druss B, Elinson L, Tanielian T, Pincus HA. National trends in the outpatient treatment of depression. *JAMA.* 2002;287:203-209.

108. Leber PD. Hazards of inference: the active control investigation. *Epilepsia.* 1989;30(suppl 1):S57-S63.

109. Stewart A. Choosing an antidepressant: effectiveness based pharmacoeconomics. *J Affect Disord.* 1998;48:125-133.

110. Stahl SM, Nierenberg AA, Gorman JM. Evidence of early onset of antidepressant effect in randomized controlled trials. *J Clin Psychiatry.* 2001;62: 17-23.

111. Leon AC. Measuring onset of antidepressant action in clinical trials: an overview of definitions and methodology. *J Clin Psychiatry.* 2001;62:12-16.

112. Charney DS. The use of placebos in randomized clinical trials of mood disorders: well justified, but improvements in design are indicated. *Biol Psychiatry.* 2000;47:687-688.