

Placebo Response in Generalized Anxiety: Its Effect on the Outcome of Clinical Trials

Edward Schweizer, M.D., and Karl Rickels, M.D.

The development of new treatments for generalized anxiety disorder increasingly has been sabotaged by a high placebo-response rate. As a consequence, and in contrast to the surge of approvals for new antidepressants, only one new anxiolytic has been approved by the U.S. Food and Drug Administration in the past 15 years. This article presents a brief review of factors that contribute to the placebo response in treatment studies of generalized anxiety. Since anxiety is a normal emotion that is sensitive to a variety of life stresses, it is particularly difficult to achieve the primary goal of a clinical trial, which is to extract the "signal" of a drug effect from the "noise" of background fluctuations in symptoms. Data from the published literature and from the authors' research unit concerning placebo-response trends are reviewed.

(J Clin Psychiatry 1997;58[suppl 11]:30-38)

In the past 15 years, hundreds of millions of dollars in the United States have been spent on the development of anxiolytics, but with the single exception of buspirone, no new anxiolytic has successfully navigated the regulatory waters to win FDA approval in the American market.

The potential explanations for this failure in drug development are numerous. First, it may be traced to the shift in focus away from the benzodiazepine-GABA-receptor system to the presumably safer candidate anxiolytics that appear to operate through serotonergic mechanisms. Alternatively, it may be that outcome measures such as the Hamilton Rating Scale for Anxiety (HAM-A), which are weighted somatically and designed to demonstrate the efficacy of benzodiazepines, are subtly but crucially biased against nonbenzodiazepine anxiolytics; remember that a loss of only 1-2 points on the HAM-A at the final visit frequently is sufficient to convert a statistically significant drug-placebo difference into a negative overall study. Finally, it may be due to a variety of variables that relate to how anxiety-treatment studies are being conducted now compared with how they were conducted

in previous decades. These variables are thought to act largely through their effect on the placebo response. Understandably, both industry and academia are concerned about this issue of placebo response—how it affects drug approval and how it might be minimized.

This article provides a brief and selective review of some of the major issues of placebo response in anxiolytic trials. Our major focus will be on studies of generalized anxiety disorder (GAD), although panic-type anxiety and depression also will be cited when no data are available for GAD, and when the results may be generalized to GAD.

Elsewhere in this supplement the results of the initial group of studies that tested the anxiolytic efficacy of the partial benzodiazepine-receptor agonist abecarnil are reported. The effect of the rate of placebo response on these studies is both a touchstone and a case study for this article.

The task of a clinical trial is to detect the "signal" of a drug effect amidst the "noise" of background fluctuations in levels (severity) of symptoms. In the case of generalized anxiety, this is an inherently more arduous task than it is in the case of depression. Anxiety is a normal emotion that has both normal and adaptive ranges of intensity. It is triggered by a host of life stresses and situations or even by the anticipation of such situations. Furthermore, anxiety may improve as the result of a variety of psychosocial variables that, though they also influence the outcome in trials of antipsychotics, for example, may not have as much influence in other non-anxious disorders, whether psychiatric or medical.¹

Before examining some of these variables, two preliminary questions deserve attention: What is the placebo response? and How essential is it for the conduct of anxiolytic trials?

From the Mood and Anxiety Disorders Section, Department of Psychiatry, University of Pennsylvania School of Medicine, Philadelphia.

Presented at the 20th Collegium Internationale Neuropsychopharmacologicum symposium titled "A New Concept in the Treatment of Anxiety," held June 23-27, 1996, in Melbourne, Australia. The symposium was made possible by an educational grant from Schering AG.

Reprint requests to: Edward Schweizer, M.D., Mood and Anxiety Disorders Section, Department of Psychiatry, University of Pennsylvania School of Medicine, University Science Center, Suite 803, 3600 Market Street, Philadelphia, PA 19104-2649.

WHAT IS THE PLACEBO RESPONSE?

The simple but arguably incorrect answer to this question is that the placebo response is the improvement that occurs in a patient who is being treated with placebo and not with the active drug. This definition ignores the likelihood that a patient who is treated with active drug may improve for reasons having nothing to do with the drug's mechanism of action. Such improvement might fairly be construed as a placebo response in a patient given active drug, though it might be difficult to "prove" that improvement is occurring for nonspecific reasons. Inferential "proof" of this has been offered, though, by Quitkin and colleagues,² who suggest that early response, especially if it is nonsustained, may be safely attributed to a placebo effect (safely, that is, if one accepts that the mechanism of antidepressant action involves monoaminergic receptor down-regulation that takes 2 or more weeks to occur). Given this criterion, the data from Quitkin and colleagues suggest that 30% of the favorable outcome achieved by antidepressants may be due largely to a placebo effect. In fact, the scientific skepticism enshrined in the null hypothesis asks us to assume that fully 50% of the drug response is actually a placebo effect in, for example, a study in which there is a 30% response to pill placebo and a 60% response to active drug. If this conservative criterion is employed, then almost every published and successful trial of anxiolytics reports a therapeutic effect for active drug in which *only a minority* of the total number of patients in the trial respond *and* do so for reasons referable to a true drug effect that reflects a known mechanism of action.

Before we begin to view our pharmacopoeia as being a high-tech version of snake oil, however, it is helpful to remember that the placebo response is heterogeneous. It may be due not only to the nonspecific and nonpharmacologic effects that are thought of traditionally as the placebo response. It may be due equally to spontaneous remission of illness. Much less is known about course-of-illness patterns in generalized anxiety than in affective illness (see the article by Rickels and Schweizer elsewhere in this supplement). One useful way to identify and control for this remission effect is to use an untreated, naturalistic, parallel control group. Unfortunately, this design is essentially never employed in clinical trials of anxiolytic agents, though spontaneous remissions in generalized or panic-related anxiety may well be higher than in affective or psychotic illness.

In discussing the heterogeneous nature of the placebo response, it is also important to mention that what are characterized as "nonspecific effects" are nonspecific only in terms of causes; they may be operating specifically through the same final pathway of neurochemical mechanisms as the active drug is operating. For example, several (although not all) studies have found that

placebo-evoked reductions in pain sensitivity can be blocked by blinded administration of naloxone, suggesting that the placebo response in this instance is acting through the well-characterized opiate-receptor pathways.³⁻⁵ We have observed (Amsterdam J. April 1987. Oral communication) a more flamboyant instance of faith healing in a patient who had treatment-resistant depression, whose previous, well-documented neuroendocrine abnormalities (dexamethasone-suppression test nonsuppression and blunted response of thyroid-stimulating hormone to thyrotropin-releasing hormone) normalized. This suggests that, in addition to operating through neurochemical mechanisms, nonspecific and placebo-like effects may achieve remissions that are not merely wishful, subjective self reports but that indeed are associated with improvements in objective physiologic measures that have been correlated with illness.¹

ARE PLACEBO-CONTROLLED TRIALS NECESSARY OR EVEN ETHICAL IN STUDYING NEW TREATMENTS FOR GAD?

The simple answer to this question for the past generation is that yes, placebo-controlled trials are the gold standard. In light of the nature of GAD, one would expect that the use of placebo would be even more crucial than it is in treatment studies of other illnesses that are associated with, perhaps, less fluctuation of symptoms. However, in the past few years, ethical concerns have been growing about the use of placebo to treat individuals who suffer from any illness in which proven, effective treatments are available.⁶ Some authors⁷ even have intimated that placebo-controlled trials violate the Declaration of Helsinki, which states, "In any medical study, every patient—including those of a control group, if any—should be assured of the best proven diagnostic and therapeutic method." These authors endorse a narrow, strict constructionist reading of this document, which would appear to preclude the possibility of using not only placebo, but any as-yet-unproven active treatment—which, in effect, would put an end to most clinical-trials research as it is currently conducted. Most clinical researchers, however, support and even mandate the use of a placebo control in clinical trials of anxiolytic agents.⁸

CLINICAL AND DEMOGRAPHIC VARIABLES ASSOCIATED WITH THE PLACEBO RESPONSE

As noted above, the placebo response is a heterogeneous and multifactorial phenomenon. One category of hypothesized determinants of the placebo response consists of clinical and demographic variables.

Personality Variables

It would be convenient if the placebo response could be attributed to a personality trait or traits of the patients being

treated in a clinical trial. This would permit systematic screening and exclusion. Unfortunately, there is little evidence that this is the case—possibly because there has been very little empiric research investigating this possibility. McNair and colleagues⁹ conducted one of the few studies that have examined personality correlates of the placebo response. They found that a high level of placebo response was associated with traits of social acquiescence and traditionalism. It has been hypothesized that similar personality traits, such as openness to experience and suggestibility or a high need for social approval, might also be correlated with an increased likelihood of a placebo response. We are not aware of any published reports of clinical trials that examined the effect of these traits, but this is an area that deserves to be the focus of more empiric research.

The placebo response in medical studies in general has been found to be more likely to occur in patients who are more anxious.¹⁰ This effect of anticipatory anxiety may contribute not only to a positive-valence placebo effect but also to a negative-valence placebo effect that manifests itself as adverse effects. One might hypothesize that placebo-treated patients who have a clear therapeutic response would have a higher rate of adverse effects than placebo-treated patients who have minimal to no therapeutic response.

Illness Variables: Duration, Severity, and Situational Exacerbation

In the treatment of depression, multiple trials have demonstrated that greater chronicity is associated with a more delayed response to treatment, and a lower overall rate of ultimate response. For GAD, the effect of greater chronicity is less well established. In one early study of “neurotic anxiety,” Rickels and colleagues clearly demonstrated that chronic illness was a poor predictor of the response to placebo but not to benzodiazepines.¹¹ The net effect was a modest increase in the difference between drug and placebo.

It is unclear whether and to what extent the effect of chronicity on treatment response is due to failure of previous treatment or other correlated variables. For example, comorbidity has been found to be significantly higher in patients who have GAD of longer duration.¹² The tendency in current clinical trials is to exclude any GAD patient who has any Axis I comorbidity. Such diagnostic fastidiousness, although commendable from the standpoint of making the results of a study easier to interpret, may serve to reduce the generalizability of the results of the GAD trial, since most GAD is associated with comorbidity. It also may bias the outcome of the trial, possibly by increasing the placebo-response rate. In light of this possibility, it is interesting to recall that before the early 1980s, all trials of anxiolytic agents were conducted in a much more diagnostically heterogeneous population, even if GAD was still the principal diagnosis.

The severity of the symptoms of GAD is another relevant illness variable, one that may be a predictor of a low placebo-response rate. Coryell and Noyes¹³ found that a higher HAM-A score at baseline predicted a lower placebo-response rate, but little other evidence concerning the effect of severity has been reported. In the related diagnosis of panic disorder, several studies have *not* found the severity of anxiety at baseline to be predictive of a lower placebo response rate.^{14,15}

GAD, perhaps more than any other psychiatric disorder, exists on a fluctuating continuum of severity that is keyed to life stresses. As discussed in another article in this supplement (Rickels and Schweizer), acute, situationally triggered anxiety may be superimposed on chronic levels of GAD. In fact, some patients likely are motivated to come in for evaluation and treatment only when their chronic anxiety is made transiently worse by situational factors. This makes the pre-randomization assessment period especially important to establish a consistent baseline of the severity of symptoms. The very act of the initial assessment and recruitment into treatment, offering as it does a sense of hope and the promise of improvement, may have the effect of reducing the severity of symptoms, especially in situationally mediated, transient, and superimposed levels of anxiety. The use of a flexible pre-randomization baseline period of 1 to 4 weeks allows the investigator some discretion in ensuring that a stable level of symptoms has been achieved prior to beginning study treatment. For patients who have had a recent exacerbation of chronic anxiety, such an observation period possibly should be extended even further to guard against a partial naturalistic remission.

To our knowledge, no prospective follow-up of outcome has been reported for patients who have been dropped from anxiety studies because of a placebo response, though such a long-term follow-up might provide useful information. Rabkin and colleagues¹⁶ have conducted a 6-month follow-up of 45 patients who were dropped from a *depression* study because of a placebo response during the initial week. At follow-up, 44% remained euthymic, and 56% had suffered a recurrence of their depression. In fact, many patients met Research Diagnostic Criteria (RDC) for “intermittent depression,” which is consistent with a course-of-illness subtype that Rickels and Schweizer hypothesize for anxiety disorders elsewhere in this supplement.

THE EFFECT OF THE SITUATIONAL CONTEXT OF TREATMENT ON THE PLACEBO RESPONSE

Therapeutic-Alliance Effects

Empathy, compassion, a helpful attitude, sympathetic listening, and making patients feel that they are being understood: All of these clinical virtues are subsumed under the rubric of the therapeutic alliance, which is the

framework of shared goals and expectations within which the treatment of the patient is conducted, whether that treatment is pharmacotherapy or psychotherapy. The therapeutic alliance comprises nonspecific effects that cut across various treatment modalities and attempt to enhance the response rates in all of them.¹⁷ Treatment conducted as part of a research study is necessarily fastidious and designed to avoid confounding variables that interfere with the interpretation of the effect of the treatment under study. Manuals have been written¹⁸ in an attempt to preclude the possibility that psychotherapeutic effects may be introduced inadvertently into a drug-treatment condition. Much less attention has been paid to controlling nonspecific variables of treatment such as the therapeutic alliance. Yet Horvath and Symonds¹⁹ have reported that a positive therapeutic alliance has an effect size of .26 on outcome in 24 psychotherapeutic trials. Similar positive correlations between the therapeutic alliance and outcome have been reported for antidepressant drug therapy,²⁰ and Crits-Christoph has found similar positive correlations between the therapeutic alliance and outcome in a pilot treatment study in GAD (Crits-Christoph P. 1996. Unpublished data). In the 1960s, a series of collaborative studies conducted by the NIMH, Johns Hopkins University, and the University of Pennsylvania demonstrated the importance of the empathy and warmth of the physician for treatment outcome.²¹⁻²³

The influence of nonspecific factors and the therapeutic alliance on outcome and the risk that an overly positive alliance will increase the placebo-response rate pose a dilemma for clinical researchers. Some degree of therapeutic alliance appears necessary to ensure compliance with medication and study procedures. The importance of having a good therapeutic rapport with the patient is especially important if the patient has been randomly assigned to the placebo group and is showing no improvement or if the study is of long duration and requires a high level of commitment from the patient.

Our internal quality-assurance monitoring at the University of Pennsylvania Mood and Anxiety Disorder Clinic suggests that simply limiting contact between the patient and the clinician rater to approximately 20 minutes is one useful method of reducing the effect of the therapeutic alliance. Much more research needs to be done, though, to arrive at a better understanding of how to control for these effects on placebo response.

Effects of Cueing, Conditioning, and Expectations

Patients usually enter a study carrying with them a long history of contacts with both medical and mental-health professionals. Getting better has therefore come to be associated over time with seeing such a professional; being evaluated by interview; having laboratory tests, ECGs, and physical examinations; and, finally, getting pills to take—all of which are also part of the ritual of a clinical

trial. Patients have very different perceptions and interpretations of the symbolic and transactional aspects of pill taking.²⁴ For example, some patients may see taking a pill as gaining control over their disturbing symptoms, whereas others see the same act as an admission of defeat and giving up control. In addition, a patient who has been treated successfully several times for anxiety or depression comes to a clinical trial with a very different set of expectations than does a patient who has failed on several previous attempts at drug therapy. Experienced patients who have been treated previously may associate specific side effects with an early response to drugs. Results of antinociceptive studies have suggested that the placebo response can be conditioned.^{25,26} However, the effects of such cueing, conditioning, and expectation are almost completely unstudied in the literature on psychiatric clinical trials. In other medical subspecialties, both gastric contractility²⁷ and airway resistance²⁸ have been shown to either increase or decrease after administration of placebo, depending on the direction of expectation. One might even view the therapeutic alliance and the relationship and attitude of the patient to the clinician as the most potent type of the effect of expectation.

We should emphasize that the estimate of clinical prognosis given to a patient by an experienced clinician does *not* appear to contribute, in and of itself, to a higher placebo-response rate. Downing and Rickels²⁹ examined how well the physician's estimate of the prognosis correlated with outcome in 517 patients treated with either a benzodiazepine or a placebo in several double-blind treatment studies. Patients who were treated with benzodiazepines and given a favorable prognosis had a significantly higher response rate (60%) than those given an unfavorable prognosis (36%). In contrast, the estimation of a favorable vs. an unfavorable prognosis was associated with no difference in outcome for patients treated with placebo (38% vs. 36%).

Side Effect Cueing and the Quality of the Blind

Perhaps the most important threat to the integrity of a double-blind study design is the potential for covert unmasking of the identity of the treatment that results from side effects. The danger of side effect cueing originates unavoidably in the disclosure of drug risks that constitutes an essential part of informed consent. There is evidence that, even with the best of intentions, consent is less "informed" than we might wish it to be.^{30,31} As a consequence, a cynic might feel comfortable dismissing the consent form with its potential for training the patient in what side effects to expect. Nonetheless, every researcher has had occasion to be confronted by a patient who, in the middle of study treatment and feeling no side effects, comments that "I must be on placebo." How a researcher responds to this question should probably be scripted since the response has much potential for doing casual

mischievous to the double-blind, but it is almost always left to the discretion of the investigator. At the University of Pennsylvania Mood and Anxiety Disorder Clinic, we make an effort to standardize our response to this type of statement by study patients.

In treatment studies of GAD, the early and prominent sedative side effects associated with benzodiazepines may serve to cue the patients that they indeed have been randomly assigned to an active drug and not to placebo. This could only serve to accentuate the drug-placebo difference in outcome. One may therefore speculate that the increasing difficulty in obtaining robust results of efficacy with the newer nonbenzodiazepine anxiolytics under investigation might be partially attributable to their lower potential for side effects, which leaves both the patient and the clinician with no way of telling whether random assignment has been to active drug or placebo. This is particularly a problem for those drugs that need weeks to show any effect.

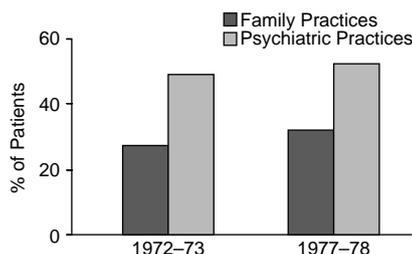
It is interesting that when 5-HT_{1A} partial agonists produce side effects such as dizziness and nausea, they are much more often dysphoric than benzodiazepine-induced side effects and much less likely to be construed by patients as early markers of a therapeutic response.

One of us (K.R.), in collaboration with Johns Hopkins and the NIMH, has previously demonstrated that the side effect of sedation produced by meprobamate was interpreted by many patients positively, whereas the side effect of dry mouth produced by atropine was interpreted by patients negatively and, in fact, was associated with a negative effect on clinical outcome.^{22,23,32} Thus, simple side effect cueing does not necessarily lead to clinical improvement, particularly if the side effects are interpreted as negative instead of positive in their emotional valence.

The only way to avoid, at least partially, side effect cueing in patients would be to use “active” placebo—in other words, substances that have no therapeutic efficacy, but which, in their own right, might cause mild dry mouth or nausea. As Lipman et al.³² have suggested, the use of such active placebos may not always have the desired effects, and active placebos are rarely if ever used in clinical trials of anxiolytic drugs.

Another method of studying the potential unblinding that may occur by side effects or other factors during clinical trials is “medication guesses.” As early as 1965, Rickels et al.³³ were concerned about such possible unblinding in trials of anxiolytic drugs. To the authors’ surprise, however, clinical improvement and not side effects—at least in trials of 4 to 6 weeks’ duration—contributed primarily to “correct” medication guesses. They reported correct medication guesses in 73% of active-drug-treated patients, 56% of inactive-drug-treated patients, and 48% (i.e., random guesses) in placebo-treated patients. In a subsequent report, Rickels et al.³⁴ confirmed these earlier findings. However, they also found that the reporting of side effects had a significant effect on medication guesses after 6 weeks of

Figure 1. Placebo Response in GAD Trials: the Effect of Treatment Setting in Four Studies*



*The percentages of patients who had moderate-to-marked improvement (LOCF analysis). Abbreviations: GAD = generalized anxiety disorder; LOCF = last observation carried forward.

therapy. They speculated that drug-induced side effects are more constant and are maintained over time but that placebo-related side effects are more transient, which, in turn, might contribute to correct medication guesses. Several more recent studies³⁵⁻³⁸ have reported correct patient-guess rates that were high enough to impeach the blind and that were, in some studies, correlated with a more positive therapeutic response.

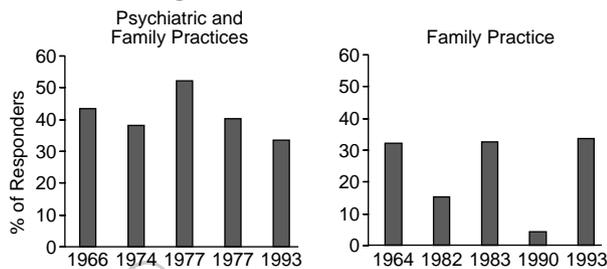
Effects of the Treatment Setting

The treatment setting is another variable that is often neglected when examining correlates and causes of placebo response. The University of Pennsylvania Mood and Anxiety Disorder Clinic has available a multisite Private Practice Research Group network, which often uses the same rater at several sites, affording us an opportunity to examine treatment-setting effects. Figure 1 shows the results of the effect of treatment setting on the placebo response in four treatment studies of generalized anxiety disorder conducted some time ago. As can be seen, a more than 20% higher placebo response was observed when the patient was seen in a private psychiatric practice as opposed to a primary-care setting. We have observed similar effects of setting, not only in anxiety, but also in depression-treatment studies.³⁹ Since the majority of affective and anxiety disorders are treated in primary-care settings, greater generalizability and the potential for a lower placebo-response rate are arguments in favor of including more primary-care sites in drug-development trials.

DOES THE PATTERN OF RESPONSE SIGNIFY A PLACEBO RESPONSE?

Quitkin⁴⁰ has suggested that the abruptness of improvement in antidepressant drug trials is correlated with a lack of sustained response and a greater likelihood, therefore, that a nondrug effect was being observed. This hypothesis finds no parallel empiric support among treatment studies in GAD. In fact, for the benzodiazepines, rapid response (within the first 1 to 2 weeks) appears to be the norm.⁴¹

Figure 2. Placebo Response in GAD Trials: the Effect of Treatment Setting Over Time*



*The percentages of patients who had moderate-to-marked improvement (LOCF analysis). Abbreviations: GAD = generalized anxiety disorder; LOCF = last observation carried forward.

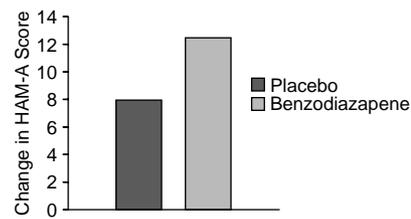
Whether early abrupt placebo response to nonbenzodiazepine anxiolytic agents is not as likely to be sustained and is consistent with a placebo response might be worth examining. Several studies of treatment of panic-related anxiety^{14,42} have found that an early and sustained placebo response is a common pattern, in contrast to what Quitkin has reported in trials of depression. It may be that the results of pattern analysis do not generalize to GAD.

The rationale of using a single-blind placebo-lead-in period is that it will tend to identify and eliminate patients who have a high potential for a placebo response. Empirical evidence to justify the inclusion of this single-blind placebo period, as opposed to a drug-free lead-in period, is lacking. One depression-treatment study⁴³ examined the effect of including “hidden placebo responders” (defined by a 20% improvement on the Hopkins Symptom Checklist depression factor). Exclusion of the 17% of patients who met these hidden-responder criteria resulted in a small *reduction* in placebo difference. Once again, whether this finding generalizes to studies of GAD is uncertain, but it is not clear that a placebo lead-in, especially one that is single blind, is either justified or worth the effort.

**THE PLACEBO RESPONSE:
TRENDS IN CLINICAL TRIALS**

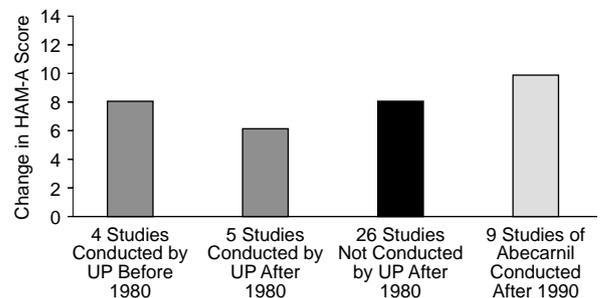
The perception among both investigators and pharmaceutical-company sponsors is that the placebo rate has been rising over the past two decades. We have tested this hypothesis in an admittedly less-than-rigorous and systematic way by examining the placebo-response trends for generalized anxiety trials conducted over the past 30 years at our University of Pennsylvania Mood and Anxiety Disorder Clinic. Figure 2 shows the percentage of moderate-to-marked global improvement (by LOCF) achieved by patients treated with placebo. One panel shows the results for studies in which results from psychiatric and primary-care practices were combined, and the other panel shows the results that were available for primary-care practices alone. As can be seen, there has been no clear trend for a

Figure 3. Combined Results From 26 Placebo-Controlled GAD Studies (Not Conducted by the University of Pennsylvania) Published Since 1980 (4-Week LOCF Data)*



*Abbreviations: GAD = generalized anxiety disorder; LOCF = last observation carried forward.

Figure 4. Placebo-Response Rates Over Time (LOCF Data)*



*Abbreviations: LOCF = last observation carried forward; UP = University of Pennsylvania.

higher placebo-response rate in the past 30 years, despite the fact that all of the trials shown from 1983 onward involved a nonbenzodiazepine anxiolytic agent. More notable, as highlighted above, was the effect of *setting* on treatment response.

Figure 3 shows aggregated results for 26 trials of anxiolytic agents (not conducted by the University of Pennsylvania) published since 1980 that had sufficient LOCF data at Week 4 to be included in a meta-analysis (Pallay A, Wyeth-Ayerst. April, 1996. Oral communication). As can be seen, the mean improvement in HAM-A score for placebo-treated patients in these almost exclusively positive clinical trials was approximately 8 points, whereas the mean improvement for drug-treated patients was in the range of 11.5 to 14 points. In fact, an improvement in the HAM-A score for drug-treated patients consistently falls in this 12- to 14-point range for effective drugs in most trials. What tends to vary, and usually is determinative of a positive or negative study, is the degree of improvement in placebo-treated patients. This can be seen clearly in Figure 4, which combines the mean placebo-improvement scores from the previous figures and contrasts them with the mean change in HAM-A scores (LOCF data) achieved in the abecarnil-development program sponsored by Sandoz/Schering (Data on file, Schering AG, Berlin, Germany). As can be seen, the mean change HAM-A score achieved with placebo treatment in

the nine treatment studies of abecarnil in GAD was 9.8. In fact, seven of the nine studies yielded a change in HAM-A score in the placebo-treatment group that was ≥ 9.8 . In the analysis of completed patients, the placebo response was even higher (change in HAM-A score ≥ 11.4). When the placebo response is this high, it is virtually impossible to obtain a therapeutic signal of drug effect, because the background noise is so loud.

PLACEBO RESPONSE AND THE PROBLEM OF NEGATIVE STUDIES

A negative result for a clinical trial traditionally has been taken to mean that the null hypothesis was confirmed and that no statistically significant difference was demonstrated between the experimental drug and the blinded placebo. This might be because of low efficacy of the experimental drug, a high placebo-response rate, or a combination of the two. Only the presence of an active comparator can help to establish whether the placebo-response rate was high enough to show that the study was, in essence, a "failed" study that could not even find efficacy for a compound of known efficacy, rather than a negative study.

If negative studies have increased in frequency over the past decade (and it is difficult to ascertain this because of the low publication rate of results of negative studies), then it becomes very interesting to understand better what might be contributing to this trend. The first step, reviewed in this article, is to understand better and control the variables correlated with the placebo response. Looking past the placebo response to other factors that might contribute to negative study results, the puzzled investigator must ask several searching questions. First, are we studying different clinical populations today compared with those of yesteryear? There doubtless has been a shift to "symptomatic volunteers" instead of patients seeking treatment. Preliminary studies⁴⁴⁻⁴⁸ have suggested a variety of minor differences between these two populations, but clear and consistent findings that predict differential clinical outcome have not been identified yet. One study of responsivity of GAD found that symptomatic volunteers and walk-in patients had identical placebo-response rates (8.8 vs. 8.1 on the HAM-A), but the symptomatic volunteers had a lower drug-response rate than did the walk-in patients (9.4 vs. 12.8).⁴⁷ This finding has not been confirmed by other studies.

Not only has there been a shift to the greater use of symptomatic volunteers, but also the climate in which studies are conducted has changed. The widespread publicity in the wake of fluoxetine about breakthroughs in the treatment of psychiatric illness has altered the expectation of improvement with which many patients enter a clinical trial. At the University of Pennsylvania, we have observed anecdotally that consumers of mental-health care are more informed and more often voice an interest in receiving

cutting-edge treatment. Many applicants for studies are also opting out of a managed-care health system, which they perceive as not responding to their mental-health needs with high-quality treatment (if any at all is covered).

Possibly, the increase in negative studies might be attributed to problems in the diagnostic identification of appropriate populations of anxious subjects. The irony here is that many pharmaceutical companies, not prompted so much by the FDA as by a meritorious but perhaps misguided urge toward diagnostic rigor and purity, have opted for structured interviews and increasingly restrictive inclusion and exclusion criteria. The paradoxical effect may be to increase the placebo-response rate, as we have already discussed. Before the early 1980s, treatment studies of GAD (then called anxiety neurosis) were routinely conducted on patients who had both panic and social phobia as well as dysthymic disorder. These patients who have higher comorbidity are less likely to respond to placebo or to undergo spontaneous remission than are patients who have lower comorbidity.

A related problem is the possibility of an outcome-measure bias. The most commonly used outcome measure in GAD is the HAM-A scale, which is generally acknowledged to be weighted toward the somatic symptoms of anxiety. The HAM-A is a metric that was introduced to facilitate the testing of benzodiazepine anxiolytics, which have been shown to treat the somatic symptoms of anxiety preferentially, whereas nonbenzodiazepine anxiolytics and antidepressants appear to be more effective in targeting the psychic symptoms of anxiety.⁴⁹⁻⁵¹ Many of the negative studies—both reported and unreported—in the past decade have been conducted on nonbenzodiazepine anxiolytics.

A final explanation for the putative increase in negative studies may lie in changes in the conduct of clinical trials. Clinical trials are now conducted using much larger numbers of patients. Faster recruitment is expected. Multiple sites are often involved, and external contract-research organizations are employed to monitor the progress of the studies. For-profit research sites have proliferated markedly, and both academic and for-profit sites have come to rely much more heavily on professional raters. This enumeration of changes in how clinical trials are conducted is not meant to provide an invidious comparison between today and a bygone golden age of clinical research, but to emphasize how much the conduct of trials has changed and to wonder aloud at what effect some of these changes might be having on the outcome of studies.

WHAT TO DO?

The appropriate remedy, as usual, is to conduct formal research on variables that contribute to the placebo response. As we already stated, it is surprising how little controlled prospective research has been conducted on

this phenomenon despite its crucial importance to the marketing of new therapeutic agents.

We hope that pharmaceutical sponsors will not embark on design strategies that are shaped by reflexive reactions to recent negative studies, which we have seen happen more than once. Although this would be understandable, the remedies proposed in that scenario are not empirically justified and often run the risk of inadvertently contributing to an even higher placebo-response rate. More elaborate and obsessively conducted clinical trials, though they provide the illusion of control, may act like a psychological equivalent of the Heisenberg uncertainty principle and significantly affect what is being observed.

Drug names: buspirone (BuSpar), fluoxetine (Prozac), meprobamate (Equanil and others), naloxone (Narcan).

REFERENCES

- Rickels K, ed. *Non-Specific Factors in Drug Therapy*. Springfield, Ill: Charles C Thomas; 1968
- Quitkin FM, Rabkin JD, Markowitz JM, et al. Use of pattern analysis to identify true drug response: a replication. *Arch Gen Psychiatry* 1987;44: 259–264
- Hersh EV, Ochs H, Quinn P, et al. Narcotic receptor blockade and its effect on the analgesic response to placebo and ibuprofen after oral surgery. *Oral Surg Oral Med Oral Pathol* 1993;75:539–546
- Levine JD, Gordon NC, Fields HL. The mechanism of placebo analgesia. *Lancet* 1978;2:654–657
- Grevert P, Goldstein A. Placebo analgesia, naloxone, and the role of endogenous opioids. In: White L, Tursky B, Schwartz GE, eds. *Placebo: Theory, Research and Mechanisms*. New York, NY: Guilford Press; 1985:332–350
- Streiner DL. The ethics of placebo-controlled trials. *Can J Psychiatry* 1995; 40:165–166
- Rothman KJ, Michels KB. The continuing unethical use of placebo controls. *N Engl J Med* 1994;331:394–398
- Prien RF, Robinson DS, eds. *Clinical Evaluation of Psychotropic Drugs: Principles and Guidelines*. New York, NY: Raven Press; 1994
- McNair DM, Gardos G, Haskell DS, et al. Placebo response, placebo effect, and two attributes. *Psychopharmacology* 1979;63:245–250
- Evans FJ. Expectancy, therapeutic instructions, and the placebo response. In: White L, Tursky B, Schwartz GE, eds. *Placebo: Theory, Research and Mechanisms*. New York, NY: Guilford Press, 1985:215–228
- Rickels K, Lipman R, Raab E. Previous medication, duration of illness and placebo response. *J Nerv Ment Dis* 1966;142:548–554
- Breslau N, Davis GC. DSM-III generalized anxiety disorder: an empirical investigation of more stringent criteria. *Psychiatry Res* 1985;15:231–238
- Coryell W, Noyes R. Placebo response in panic disorder. *Am J Psychiatry* 1988;145:1138–1140
- Dager SR, Khan A, Cowley DS, et al. Characteristics of placebo response during long-term treatment of panic disorder. *Psychopharmacol Bull* 1990; 26:273–278
- Shear MK, Leon AC, Pollack MH, et al. Pattern of placebo response in panic disorder. *Psychopharmacol Bull* 1995;31:273–278
- Rabkin JG, Markowitz JS, Stewart JW, et al. 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
- Shapiro AK, Shapiro E. Patient-provider relationships and the placebo effect. In: Matarazzo JD, Weiss SM, eds. *Behavioral Health: A Handbook of Health Enhancement and Disease Prevention*. New York, NY: Wiley-Interscience; 1984:371–383
- Fawcett J, Epstein P, Fiester SJ, et al. Clinical management—imipramine/placebo administration manual. NIMH Treatment of Depression Collaborative Research Program. *Psychopharmacol Bull* 1987;23:309–324
- Horvath AO, Symonds BD. Relation between working alliance and outcome in psychotherapy: a meta-analysis. *J Counsel Psychol* 1991;38: 139–149
- Marziali E, Marmar C, Krupnick J. Therapeutic alliance scales: development and relationship to psychotherapy outcome. *Am J Psychiatry* 1981; 138:361–364
- Fisher S, Cole JO, Rickels K, et al. Drug-set interaction: the effect of expectations on drug response in outpatients. In: Bradley PB, Fluegel F, Hock P, eds. *Neuropsychopharmacology*. Amsterdam, The Netherlands: Elsevier; 1964:80–84
- Uhlenhuth EH, Rickels K, Fisher S, et al. Drug, doctor's verbal attitude and clinic setting in the symptomatic response to pharmacotherapy. *Psychopharmacologia* 1966;9:392–418
- Rickels K, Lipman RS, Park LC, et al. Drug, doctor warmth and clinic setting in the symptomatic response to minor tranquilizers. *Psychopharmacologia* 1971;20:128–152
- Barsky AJ. Nonpharmacologic aspects of medication. *Arch Intern Med* 1983;143:1544–1548
- Voudouris NJ, Peck CL, Coleman G. Conditioned response models of placebo phenomena: further support. *Pain* 1989;38:109–116
- Voudouris NJ, Peck CL, Coleman G. The role of conditioning and verbal expectancy in the placebo response. *Pain* 1990;43:121–128
- Sternbach RA. The effects of instructional sets on autonomic responsivity. *Psychophysiology* 1964;1:67–72
- Luparello T, Leist N, Lourie CH, et al. The interaction of psychologic stimuli and pharmacologic agents on airway reactivity in asthmatic subjects. *Psychosom Med* 1970;32:509–513
- Downing RW, Rickels K. Physician prognosis in relationship to drug and placebo response in anxious and depressed psychiatric outpatients [Brief Communication]. *J Nerv Ment Dis* 1983;171:182–185
- Mann T. Informed consent for psychological research: do subjects comprehend consent forms and understand their legal rights? *Psychological Sci* 1994;5:140–143
- Grisso T, Appelbaum PS. The MacArthur Treatment Competence Study, III: abilities of patients to consent to psychiatric and medical treatments. *Law Hum Behav* 1995;19:149–174
- Lipman RS, Park LC, Rickels K. Paradoxical influence of a therapeutic side-effect interpretation. *Arch Gen Psychiatry* 1966;15:462–474
- Rickels K, Raab E, Carranza J. Doctor medication guesses: an indicator of clinical improvement in double-blind studies—a contribution to the methodology of psychiatric drug evaluation. *J New Drugs* 1965;5:67–71
- Rickels K, Lipman RS, Fisher S, et al. Is a double-blind clinical trial really double-blind? A report of doctors' medication guesses. *Psychopharmacologia* 1970;16:329–336
- Brownell KD, Stunkard AJ. The double-blind in danger: untoward consequences of informed consent. *Am J Psychiatry* 1982;139:1487–1489
- Rabkin JG, Stewart JW, McGrath PJ, et al. Baseline characteristics of 10-day placebo washout responders in antidepressant trials. *Psychiatry Res* 1987;21:9–22
- Margraf J, Ehlers A, Roth WT, et al. How “blind” are double-blind studies? *J Consult Clin Psychol* 1991;59:184–187
- Morin CM, Colecchi C, Brink D, et al. How “blind” are double-blind placebo-controlled trials of benzodiazepine hypnotics? *Sleep* 1995;18: 240–245
- Rickels K, Schweizer E, Clary C, et al. Nefazodone and imipramine in major depression: a placebo-controlled trial. *Br J Psychiatry* 1994;164: 802–805
- Quitkin FM, Stewart JW, McGrath PJ, et al. Further evidence that a placebo response to antidepressants can be identified. *Am J Psychiatry* 1993;150: 566–570
- Downing RW, Rickels K. Early treatment response in anxious outpatients treated with diazepam. *Acta Psychiatr Scand* 1985;72:522–528
- Mavissakalian M. The placebo effect in agoraphobia, II. *J Nerv Ment Dis* 1988;176:446–448
- Reimherr FW, Ward MF, Byerley WF. The introductory placebo washout: a retrospective evaluation. *Psychiatry Res* 1989;30:191–199
- Hersen M, Bellack AS, Himmelhoch JM. A comparison of solicited and nonsolicited female unipolar depressives for treatment outcome research. *J Consult Clin Psychol* 1981;49:611–613
- Last CG. Treatment outcome for solicited versus nonsolicited unipolar depressed female outpatients. *J Consult Clin Psychol* 1984;52:134–135
- Thase ME. Symptomatic volunteers in depression research: a closer look. *Psychiatry Res* 1984;11:25–33
- Tiller JW, Biddle N. Obtaining outpatient referrals for clinical research. *Aust N Z J Psychiatry* 1991;25:132–133
- Rapaport MH, Frevort T, Babior S, et al. A comparison of demographic

- variables, symptom profiles, and measurements of functioning in symptomatic volunteers and an outpatient clinical population. *Psychopharmacol Bulletin* 1995;31:111-114
49. Rickels K. Buspirone and diazepam in anxiety: a controlled study. *J Clin Psychiatry* 1982;43(12, sect 2):81-86
50. Rickels K, Downing R, Schweizer E, et al. Antidepressants for the treatment of generalized anxiety disorder: a placebo-controlled comparison of imipramine, trazodone, and diazepam. *Arch Gen Psychiatry* 1993;50:884-895
51. McLeod DR, Hoehn-Saric R, Zimmerli WD, et al. Treatment effects of alprazolam and imipramine: physiological versus subjective changes in patients with generalized anxiety disorder. *Biol Psychiatry* 1990;28:849-861

© Copyright 1997 Physicians Postgraduate Press, Inc.
One personal copy may be printed