Efficacy Issues With Antidepressants

Jan Fawcett, M.D., and Robert L. Barkin, M.B.A., Pharm.D.

Clinical trials of antidepressant medications have shown that, overall, these drugs are effective, as measured by a ≥ 50% decrease in Hamilton Rating Scale for Depression (HAM-D) total scores in about two thirds of patients. However, the results of long-term trials under rigorously controlled conditions show that, even with close follow-up and provision of interpersonal psychotherapy, a third or more of the patients will not achieve or maintain a response to medication for depression. Nevertheless, the improved efficacy of some antidepressants for certain features or types of depression has been shown. Factors associated with a better response to a specific agent or class of drugs include severity of symptoms, patient age, and the symptom profile of the depressive episode, as revealed by assessment scales or subscale scores for selected symptoms. Moreover, a number of studies indicate that a patient's early response to a given medication may assist in predicting long-term outcome. However, outcome measures in traditional trials of antidepressant drug efficacy, such as a 50% reduction in scores on one or more depression rating scales, do not necessarily reflect an improvement in the patient's ability to function in the workplace; they only show that a particular patient at a particular time has responded to treatment in a significant manner by measurement of a depression scale.

(J Clin Psychiatry 1997;58[suppl 6]:32–39)

Although several new classes of antidepressant medications have been introduced in recent years, clinical experience with newer medications has not revealed a significant increase in efficacy. Indeed, one third of the patients diagnosed with major depression display symptoms that still remain resistant to pharmacologic intervention. This review will examine the factors that can influence the efficacy of antidepressant medications in an effort to identify which, if any, can be manipulated by the physician to improve the rate of efficacy in the management of depression. Important efficacy issues in the treatment of depression will be highlighted by addressing five questions: (1) How is antidepressant efficacy best measured? (2) How efficacious are antidepressant medications? (3) How early can antidepressant efficacy be detected? (4) Can a patient's response to antidepressant medication be predicted clinically? (5) Can antidepressant efficacy be improved?

From the Departments of Psychiatry (Dr. Fawcett) and Anesthesiology and Family Medicine (Dr. Barkin), Rush-Presbyterian St. Luke's Medical Center, Rush Medical College, Chicago, Ill.

Presented at the symposium "Advances in Antidepressant Pharmacotherapy," February 2, 1996, New Orleans, La., supported by an educational grant from Organon, Inc.

Reprint requests to: Jan Fawcett, MD, Institute for Mental Well-Being, 1725 West Harrison Street, Suite 955, Chicago, IL 60612.

HOW IS ANTIDEPRESSANT EFFICACY BEST MEASURED?

Clinical trials have assessed the response to antidepressant therapy by using a variety of objective scales.

Choice of Assessment Scales

The 17-item and 21-item Hamilton Rating Scales for Depression (HAM-D) are the standard scales for assessing the efficacy of antidepressant medications in clinical trials. However, a number of other scales are often used and have become well accepted for analysis of antidepressant efficacy. For example, van Moffaert et al.[1] found mirtazapine to be statistically significantly better than trazodone in hospitalized patients not only when HAM-D scores were compared, but also when the Brief Psychiatric Rating Scale (BPRS), the General Psychiatric Impression Global Assessment Scale (GAS), and the Beck Self-Rating Depression Scale (Beck) were used. The Montgomery-Asberg Depression Rating Scale (MADRS) can also be used as an objective measure of the severity of major depression and to evaluate the efficacy of antidepressant medications.

Subscale Score Analysis

The results of clinical trials that evaluate antidepressant medication efficacy often analyze specific factors on the HAM-D to identify the particular strengths of each

Table 1. Efficacy of Antidepressant Medications Compared With Placebo in Controlled Trials*

Medication	Response Rate[a] (Percentage of Patients) Drug	Placebo	Difference Drug–Placebo
TCAs (N = 3327)[b]			
Amitriptyline	60	25	35
Amoxapine	67	49	18
Imipramine	68	40	28
SSRIs (N = 2463)[b]			
Paroxetine	45	23	22
Fluoxetine	60	33	27
Fluvoxamine	67	39	28
Sertraline	79	48	31
MAOIs (N = 1944)[b]			
Phenelzine	64	30	34
Moclobemide	64	24	40
NaSSA (N = 277)[b,c]			
Mirtazapine	48	20	28

*Adapted from Davis et al.[3]
[a]Response rate = at least a 50% reduction in baseline HAM-D scores.
[b]N = total number of patients included in placebo-controlled and/or comparison trials of TCAs (tricyclic antidepressants), SSRIs (serotonin selective reuptake inhibitors), MAOIs (monoamine oxidase inhibitors), and NaSSA (noradrenergic and specific serotonergic antidepressant).
[c]Data on file, Organon, Inc., West Orange, NJ.

drug so that therapy might be targeted more closely to each patient's symptom profile.

A meta-analysis by Angst et al.[2] of moclobemide's efficacy in different patient groups exemplifies the type of definition of efficacy that can be achieved using these subscales. These authors examined intent-to-treat data from 40 placebo-controlled and/or comparison trials involving moclobemide, stratifying results separately for patients with low, medium, or high severity of depression. They found moclobemide to be significantly more efficacious than placebo, specifically against symptoms represented by the retarded depression subscale when compared with the agitation/anxiety subscale.

HOW EFFICACIOUS ARE ANTIDEPRESSANT MEDICATIONS?

Efficacy of Antidepressant Drugs Compared With Placebo

A recent meta-analysis[3] provides a good perspective on antidepressant medication efficacy. Davis et al.[3] performed a meta-analysis of approximately 300 double-blind, randomized clinical trials of drug treatments for affective disorders. They determined the number of patients in each study who responded to drug or placebo therapy and pooled the results for individual drugs and for classes of antidepressants.

In the Davis et al. study, response rates for different classes of antidepressant medications were 60% to 68% for the tricyclic antidepressants (TCAs) amitriptyline, amoxapine, and imipramine; 45% to 79% for the serotonin selective reuptake inhibitors (SSRIs) paroxetine, fluoxetine, fluvoxamine, and sertraline; 64% for the monoamine oxidase inhibitors (MAOIs) phenelzine and moclobemide; and 48% for the new noradrenergic and specific serotonergic antidepressant (NaSSA) mirtazapine (Table 1).[3]

A comparison of the response rates to active medication and placebo in the controlled trials shows that both the TCAs and the MAOIs were 18% to 41% more efficacious than placebo in treating the acute stage of depression in the 7103 patients enrolled in 112 controlled trials. Davis et al.[3] found that antidepressant medications were also more efficacious than placebo for maintenance therapy: 23% of patients maintained on an active drug relapsed, while 50% of those receiving placebo relapsed among a total of 2225 patients in 18 studies. Comparisons across studies of drug-placebo response rates are 18% to 34% for TCAs, 22% to 31% for SSRIs, 34% to 41% for MAOIs, and 28% for mirtazapine.

HOW EARLY CAN ANTIDEPRESSANT EFFICACY BE DETECTED?

Another important issue relevant to the pharmacologic management of depression is the onset of therapeutic action and its implications for long-term treatment.

Researchers have long sought evidence in clinical trials of an early response to antidepressant medication, and, indeed, a clinical response within the first 2 weeks of treatment can be detected and may be predictive of therapeutic outcome.[4] Boyer and Feighner[5] examined the issue of the rapidity of onset with antidepressant medication. They performed a meta-analysis of six trials to determine the predictive value of nonresponse to medication early in a clinical trial. These researchers found that patients who failed to achieve at least a 20% reduction in HAM-D score at any point during the first 4 weeks of a study had only a 3.7% chance of being a "responder" as determined by 50% or greater reduction in HAM-D score at the end of the 6-week study. Results were very similar when patients were evaluated using the Clinical Global Impressions (CGI) scale. Boyer and Feighner concluded from this study that a full 6-week trial of an antidepressant is usually not warranted if there is not at least some improvement in the symptoms of depression during the first 4 weeks of a trial.

Nierenberg et al.[6] conducted a double-blind 8-week study of 143 depressed patients who had baseline HAM-D scores (mean ± SD) of 19.5 ± 3.1 to evaluate whether no response (< 20% decrease from baseline HAM-D score) within the first 6 weeks of treatment with fluoxetine was a predictor of a poor outcome. Of patients who showed no improvement at Weeks 2, 4, and 6, the proportion of responders (≥ 50% decrease from baseline HAM-D score)

at Week 8 were 36.4%, 18.9%, and 6.5%, respectively. Of patients who showed at least a partial response to treatment (20% to 40% decrease from baseline HAM-D score) at Week 6, 25% responded to treatment at Week 8. In this study, a nonresponse as early as the second week of treatment predicted a poor outcome, which may be a function of the long half-life of fluoxetine and the active metabolite norfluoxetine.

In a 4-week, double-blind comparison of nomifensine and imipramine, efficacy early in the trial was also predictive of success.[7] The mean decrease in HAM-D scores for patients receiving nomifensine was significantly greater than for those receiving imipramine at the Week 1 assessment, and the total scores continued to favor nomifensine for the rest of the trial, although these later differences were not statistically significant.

Derivan et al.[8] also examined the onset of antidepressant drug effects in outpatients by analyzing two placebo-controlled trials of venlafaxine. Drug doses were escalated rapidly so that most patients received close to 200 mg/day by Day 7 or Day 8 and 340 mg/day by Week 2. In the venlafaxine group, more than half the patients showed both early (Week 1 and/or Week 2) and persistent (not followed by relapse or a CGI score of 3 or above for any subsequent week of the study) improvement in CGI scores, compared with only 15% of those receiving placebo ($p < .001$).

Gasperini et al.[9] studied the clinical psychopharmacology of amitriptyline and fluvoxamine and found that overall response to treatment with either drug could be predicted from the response at the Week 2, but not the Week 1, assessment. In addition, significantly better responses to fluvoxamine and/or amitriptyline could be predicted from scores related to four symptoms: insomnia (improvement predictive of a response to fluvoxamine), psychomotor retardation (failure to improve predictive of no response to either drug), psychic anxiety (improvement predictive of response to either drug), and diurnal variation (improvement predictive of response to amitriptyline). The fact that insomnia and psychic anxiety are usually seen together clinically strengthens the probability that fluvoxamine will prove efficacious in patients with symptoms of anxiety that improve by the second week of treatment.

Several randomized, placebo-controlled studies of mirtazapine also demonstrated an early onset of antidepressant effects.[10–12] The Bremner[10] and Smith et al.[11] studies were double-blind comparisons of mirtazapine, amitriptyline, and placebo. In major depression, the effectiveness of mirtazapine was comparable with that of amitriptyline. In the Bremner study, mean 17-item HAM-D total scores at baseline were > 25, indicative of a large proportion of severely depressed patients. Mirtazapine and amitriptyline showed significantly ($p \leq .05$) greater reductions in the mean HAM-D total scores and Item-1 (depressed mood) scores from Week 1 onward. In the Smith et al. study, mean 17-item HAM-D total scores at baseline were > 23, indicative of a large proportion of moderately to severely depressed patients. Mirtazapine and amitriptyline showed significantly ($p \leq .05$) greater mean HAM-D score reductions compared with placebo at Weeks 1, 2, and 4, and at the endpoint analysis. In the Claghorn and Lesem[12] study, mean 17-item HAM-D total scores at baseline were comparable in the mirtazapine (HAM-D = 21.5) and placebo (HAM-D = 22.7) groups. The mirtazapine group showed a significant ($p < .05$) improvement in HAM-D compared with the placebo group from Week 1 through Week 5, and at the endpoint. The difference between groups at Week 6 may not have been significant owing to the large proportion of premature discontinuations in the placebo group due to lack of efficacy, thus resulting in a small subset of placebo responders at Week 6.

The results of the clinical trials just discussed suggest that a therapeutic effect with antidepressant drugs can be detected within the first 2 weeks of treatment. From these results, Montgomery[13] proposed a novel method for a rapid evaluation of an antidepressant's efficacy. His alternative for predicting efficacy was to design short (2-week), placebo-controlled trials and to define a clinical response as a 25% reduction in the HAM-D scores (half of the 50% reduction required for a 4- to 6-week trial) and a reduction of 4 points on the MADRS. He also proposed twice-weekly assessments to increase the sensitivity of detecting clinical signs of the onset of antidepressant action and follow-up studies to evaluate the long-term efficacy of drugs that showed promise during the brief trial.

CAN A PATIENT'S RESPONSE TO ANTIDEPRESSANT MEDICATION BE PREDICTED CLINICALLY?

The meta-analysis by Davis et al.[3] of the results of trials in which newer antidepressant medications were compared with standard drugs (usually imipramine or amitriptyline) showed no statistically significant difference between the proportion of patients responding to the newer antidepressants and the standard drugs. However, it seems clear that some classes and individual antidepressant medications have a greater chance of eliciting a response depending on the patient's age, severity of the illness, and the type of depressive symptoms present.

Patient Age

An important factor that can affect antidepressant drug efficacy is patient age (adults younger than 65 years versus elderly). Gerson et al.[14] documented differences between older and younger patients in their responses to antidepressant medications administered during clinical trials. More recently, Dunner[15] reviewed the considerations that are made in choosing an antidepressant medication for clinical treatment of depression in elderly patients. Dunner noted that the SSRI class of antidepressants, because of its favorable tolerability, is prescribed more frequently than

the TCAs. In contrast to TCAs, SSRI antidepressants display a minimal tendency to cause side effects, such as postural hypotension and confusion, that increase the elderly patient's risk of an accident. Another alternative for the elderly patient may be an atypical medication such as trazodone, which is safe in overdose and has a better cardiovascular safety profile than the TCAs.[15] However, intolerance to higher doses of trazodone may limit titration to higher doses in an attempt to achieve better efficacy.

Halikas[16] compared the NaSSA mirtazapine with trazodone in a placebo-controlled, 6-week trial in 150 outpatients, 55 years or older, presenting with moderate-to-severe major depressive disorder. The response to mirtazapine was significant for this age group: at every assessment point, patients receiving mirtazapine had greater reductions in HAM-D scores compared with those receiving trazodone, and they had significantly ($p \leq .05$) greater improvement than those receiving placebo at the Week 2, 3, 4, and 6 assessments and at the endpoint of the study. It should also be noted that trazodone induced dose-related adverse experiences, such as anticholinergic symptoms, dizziness, and blood pressure changes that limited the upward dose titration. Clinical improvement for trazodone may require dose titration upwards to 600 mg/day when tolerated by patients.

Table 2. Efficacies of Antidepressant Drugs by Class in Clinical Comparative Trials*

Drug Class	Inpatients†	Outpatients†	Long-Term† (> 12 Weeks)
TCA[a]	50.9 ± 6.2	49.9 ± 6.1	48.3 ± 4.3
Heterocyclics	49.2 ± 4.4	58.9 ± 5.4	38.8 ± 15.8
SSRI[a]	33.3 ± 6.8	59.0 ± 2.2	42.2 ± 16.0
SNRI[a]	62.4 ± 7.4	55.6 ± 3.8	40.6 ± 5.3
NaSSA[a]	71.7[b]	60–70[c]	51.4[d]

*Adapted from Einarson et al[18] and modified to add NaSSA.
†Percentage Responding Overall (mean ± SD) with at least 50% reduction in baseline HAM-D scores.
[a]Abbreviations: TCA = tricyclic antidepressant, SSRI = serotonin selective reuptake inhibitor, SNRI = serotonin-norepinephrine reuptake inhibitor, NaSSA = noradrenergic and specific serotonergic antidepressant.
[b]Data from reference 17.
[c]Data from reference 20.
[d]Data from reference 21.

Severity of Illness

Treatment setting. Clinical trials of antidepressant drug efficacy are usually conducted in either an inpatient or an outpatient setting, which are representative of populations with greater or lesser severity of illness. An example of an inpatient comparison trial is one reported recently by Zivkov and de Jongh.[17] This amitriptyline-controlled trial of mirtazapine involved 251 adults with diagnoses of major depressive episode who were treated at six university psychiatric centers. In this study the efficacy response rate, determined by the proportion of patients who experienced a 50% or greater reduction from baseline HAM-D scores at the end of 6 weeks, was 71.7% for those receiving mirtazapine and 72.1% among those receiving amitriptyline.

A placebo-controlled comparison trial of mirtazapine and amitriptyline was also recently conducted by Bremner in outpatients.[10] At baseline, mean 17-item HAM-D scores were > 25, indicative of a large proportion of severely depressed patients in this study. Mirtazapine or amitriptyline effected a statistically significantly greater reduction in HAM-D scores compared with placebo at each week of this 6-week study. At the end of the study, significantly more patients had responded, as measured by a 50% or greater reduction in their baseline HAM-D scores, to mirtazapine (70%) or amitriptyline (58%) compared with placebo (33%).[10] The results of these studies indicate that amitriptyline and mirtazapine are comparably effective in both outpatients and inpatients.

A meta-analysis by Einarson et al.[18] focused on defining the specific efficacies of classes of antidepressant medications for inpatient, outpatient, and long-term (> 12 weeks) therapy for depression. These investigators reviewed 221 publications and found 34 studies that met their stringent criteria for inclusion in the meta-analysis: double-blind, randomized, prospective trials involving patients with a single or recurrent episode of major depressive disorder meeting DSM-III-R[19] criteria, with HAM-D scores of ≥ 18 at baseline who were treated for a minimum of 4 weeks with antidepressant doses in the therapeutic range. The antidepressants included in the meta-analysis were the TCAs amitriptyline and imipramine; the heterocyclics trazodone and maprotiline; the SSRIs fluoxetine, sertraline, and paroxetine; and the serotonin-norepinephrine reuptake inhibitor (SNRI) venlafaxine. Four to 12 outpatient studies, one to four inpatient studies, and one or two long-term trials evaluating the efficacy of each antidepressant met the inclusion criteria for meta-analysis.

The results for each class of antidepressant studied are shown in Table 2, with studies regarding the efficacy of the NaSSA mirtazapine added. For inpatients (those with more severe depression), efficacy appears to be lowest with an SSRI and highest for the NaSSA mirtazapine and the SNRI venlafaxine. For outpatients and those in long-term treatment, the percentage responding to treatment was similar among the different classes.

Baseline Assessment Scale Scores

Although treatment setting is an indicator of patients' severity of depression, baseline scores on assessment scales are the standard method for classifying severity of illness among patients being recruited for participation in a clinical trial of antidepressant drug efficacy. Baseline scores reveal that inpatients in the Zivkov and de Jongh

Table 3. Baseline HAM-D Scores of Inpatients and Outpatients in Amitriptyline-Controlled Trials of Mirtazapine Efficacy*

Study	Number of Patients[a]	Study Population	Group Mean HAM-D
Bremner[10]	50	Outpatients	≥ 25
Smith et al.[11]	50	Outpatients	23.4
Zivkov and de Jongh[17]	125	Inpatients	28.0
Mullin et al.[22]	79	Inpatients and outpatients	22.5

*Adapted from Zivkov et al.[20] All studies were double-blind, randomized, dose-titration trials lasting 5 to 6 weeks; the outpatient studies were placebo-controlled and employed lower dosages of active medication (5 to 35 mg/day of mirtazapine). Studies including inpatients were multicenter trials and employed higher dosages of medication (20 to 60 mg/day of mirtazapine).
[a]Number of patients = number of patients in the mirtazapine arm of the trial.

Table 4. Percent Change in HAM-D Scores in Patients Treated With Imipramine*

Study	N	Maximal Dose (mg/d)	Treatment Period	Baseline Mean	Baseline SD	Termination Mean	Termination SD	Change
Byerley et al[26]	34	300	6 wk	28.3	4.2	13.7	8.5	52%
Elkin et al.[24]	57	185	16 wk	19.5	4.6	9.8	7.8	50%
Stark and Hardison[27]	185	300	6 wk	28.2	5.8	16.2	10.1	43%

*Adapted from Workman and Short.[25]

study[17] and outpatients in the Bremner study[10] actually had a very similar illness severity, since the HAM-D scores exceeded 25 in both groups. Comparison of the treatment settings and baseline HAM-D scores for patients in these two studies and two others included in a meta-analysis by Zivkov et al.[20] (Table 3) shows that treatment settings do not reliably predict patients' severity of illness, as measured by standard assessment scales.

Patients' baseline scores on standard assessment scales may, however, be good predictors of the efficacy of an antidepressant medication in clinical trials, as shown by the similar efficacy rates for mirtazapine in the Zivkov and de Jongh (71.7%)[17] and the Bremner (70%)[10] trials in patients with similar severity of illness as measured on the 17-item HAM-D.

Mirtazapine (N = 66) was compared with fluoxetine (N = 67) in a recent 6-week, double-blind study, with an optimal extension of up to 6 months.[23] This study enrolled patients with a 17-item HAM-D total score ≥ 21 and a HAM-D depressed mood score of ≥ 2. In this population of more severely depressed patients, successive assessments were made using the HAM-D scores at screening, baseline, at Weeks 1 through 4, and at Week 6. Mirtazapine was rapidly titrated, as follows: Days 1–4, 15 mg; Days 5–7, 30 mg; Days 8–28, 45 mg; and Days 29–42, 45–60 mg. Fluoxetine doses were titrated, as follows: Days 1–28, 20 mg; Days 29–42, 20–40 mg. The mirtazapine group showed greater improvement than the fluoxetine group at Week 3 (HAM-D: 14.5 mirtazapine versus 18.3 fluoxetine; $p < .05$), Week 4 (HAM-D: 12.6 mirtazapine versus 17.0 fluoxetine; $p < .05$), and Week 6 (HAM-D: 11.8 mirtazapine versus 15.8 fluoxetine; $p = .0543$). The four-point differences in HAM-D scores between groups detected in this study are usually seen in trials comparing an antidepressant with placebo.

The NIMH collaborative treatment of depression study[24] failed to demonstrate any statistically significant difference in response rates (HAM-D score of ≤ 6 at study termination) for patients receiving imipramine or interpersonal psychotherapy (IPT) compared with placebo until the data were reanalyzed according to severity of illness at baseline. Reanalysis continued to show little benefit for active therapy compared with placebo for patients with less severe symptomatology (HAM-D < 20 at baseline), but for patients with greater severity of depression, both imipramine and IPT were significantly more effective than placebo.[24] These results demonstrate the importance of evaluating antidepressant drug efficacy in the light of severity of illness, as measured by baseline assessment scales.

A meta-analysis by Workman and Short[25] of comparison trials involving imipramine is interesting in this regard because one of the studies differed from the others in that patients had lower baseline HAM-D scores, clearly representing milder depression (Table 4).

Effectiveness and Functional Capacity

Traditionally, studies of antidepressant drugs have considered a medication effective when it achieves a 50% reduction in HAM-D scores in adults with an episode of major depression. Workman and Short[25] used this criterion in a meta-analysis of double-blind, placebo-controlled clinical trials of the efficacies of various atypical antidepressant medications measured against imipramine. They found close to a mean 50% reduction in HAM-D scores from baseline to termination in imipramine efficacy studies they analyzed (Table 4). Of note, the lowest mean HAM-D score at termination (9.8) is above the HAM-D score of ≤ 7 that is considered representative of remission, with the remaining scores as high as 16.2.

Such HAM-D scores represent significant residual impairment in work function, as confirmed in a study by Mintz et al.[28] According to the results of this study, about one third of patients with observer-rated HAM-D scores of 13 and about half with scores of 16 would have impairment in work function as demonstrated by absenteeism, poor performance, and/or significant interpersonal conflict. Furthermore, impairment in work function (feelings of job inadequacy, distress at work, and lack of interest in the job)

persisted in about a third of patients after remission of affective symptoms.

With the current emphasis in health care on returning patients to work as quickly as possible, a measure of antidepressant drug efficacy that fails to identify patients who have recovered their ability to function on the job may be outdated. Achievement of remission, as indicated by a HAM-D score of 7 or less, and/or absence of work impairment, would provide a more realistic measure of antidepressant drug effectiveness.

CAN ANTIDEPRESSANT DRUG EFFICACY BE IMPROVED?

The final and possibly the most important antidepressant drug efficacy issue is the level of an antidepressant drug response that is attainable. The validity of this issue is largely determined by the answers to the question: Can response rates to antidepressant medication be improved? An examination of the methods and outcomes of studies of imipramine efficacy gives insight into this issue. Davis et al.[3] classified response to active medication in the double-blind, random-assignment studies included in their meta-analysis as moderate improvement or better on a semi-quantitative scale, or a 50% decrease in the HAM-D scores. Some 2% of the studies they included gave no numerical data concerning response but reported a substantial dropout rate for lack of efficacy, and this rate was taken as the nonresponse rate. Using these criteria, the Davis et al.[3] meta-analysis showed an overall response rate to imipramine of 68% in 50 studies, which included a total of 2649 patients who received this medication for the acute treatment of a major depressive episode.

Frank et al.[29] conducted a carefully controlled long-term trial of the efficacy of full therapeutic dosages of imipramine, with or without IPT, versus placebo in preventing the recurrence of depression in patients in their third or greater major depressive episode. To enter the 3-year maintenance phase of this trial, patients had to achieve a HAM-D score of ≤ 7 (indicating remission) during the acute treatment phase and maintain this score for the 17-week continuation phase of treatment.

Of the 230 patients who began the acute treatment phase of this trial, which included imipramine at daily doses averaging slightly more than 200 mg (range 150 to 300 mg) plus weekly to biweekly IPT sessions, 157 (68%) achieved a HAM-D score of ≤ 7 in 12 to 20 weeks, and were thus eligible to enter the continuation phase. Only 128 (56% of the initial group) maintained a HAM-D score of ≤ 7 throughout the continuation phase, during which imipramine was continued at the acute-phase dosage and IPT was provided monthly. During the 3-year maintenance phase, the proportions of patients receiving medication clinic visits and active imipramine (average dose of approximately 200 mg/day) or imipramine plus IPT who survived without a recurrence of major depression were 60.7% and 84.0% at 1 year, 46.4% and 64.0% at 2 years, and 46.4% and 60.0% at 3 years.[29] It was concluded that an average dose of imipramine maintained was an effective method of recurrence prevention and monthly IPT served to lengthen the time intervals between episodes in those patients receiving no pharmacotherapy.

Bremner and Smith[21] demonstrated long-term efficacy with 118 patients in a 20-week extension of two randomized 6-week placebo-controlled short-term studies of mirtazapine (N = 43) versus amitriptyline (N = 48) and placebo (N = 27). Remission (HAM-D score ≤ 7) occurred in 54.7% of the patients in the mirtazapine group who completed the 20-week extension compared with 54.1% of the amitriptyline group and 19.2% of the placebo group. In the mirtazapine group, mean HAM-D scores remained ≤ 7 throughout the 20-week extension, and only 2.4% of the mirtazapine patients had an endpoint HAM-D score ≥ 16. The amitriptyline treatment group had a mean HAM-D score stabilized at ≤ 7 at Week 12, with 10.4% having an endpoint HAM-D score of ≥ 16. The mean HAM-D scores remained ≥ 7 throughout the trial for patients taking placebo, with 23% having an endpoint HAM-D score ≥ 26. Patient representation in the two trials of those who began and completed the extension trials of the three groups was as follows: mirtazapine 43 and 27, amitriptyline 48 and 33, and placebo 27 and 15. Patients dropped out of the trial secondary to lack of efficacy, clinical improvement, and adverse effects. The respective percent dropout for each reason above was as follows: mirtazapine: 2.1%, 11.6%, 9.3%; amitriptyline: 4.6%, 2.1%, 8.3%; placebo: 11.1%, 3.7%, 3.7%. Mirtazapine had a beneficial therapeutic effect and was better tolerated than amitriptyline in continuation treatment of major depression.

The efficacy rate for patients receiving imipramine for treatment of the acute phase of major depression in the studies included in the Davis et al.[3] meta-analysis was 68%, and the remission rate for mirtazapine in the continuation phase of the studies by Bremner and Smith[21] was similar at 76.2%.

These results represent the efficacy rates that can be expected in clinical trials, and they are probably the best rates that can be achieved with standard therapeutic strategies using available medications for acute and long-term treatment of major depression. There is hope, however, that antidepressant efficacy may be improved, if specific classes of antidepressant drugs can effectively treat certain constellations of symptoms. Indeed, a placebo-controlled study that examined the efficacy of mirtazapine, as measured by the Hamilton Rating Scale for Anxiety (HAM-A) and Zung Anxiety Scale, showed a significant benefit for this antidepressant in daily dosages of 15–25 mg for the treatment of outpatients with a primary diagnosis of anxiety.[30] These results led to the evaluation of mirtazapine

Table 5. Group Mean HAM-D Scores*

Scale/Medication	Baseline	Week of Assessment 1	2	3	4	6
HAM-D (total score)						
Paroxetine	26.5	22.5	19.6[a]	18.2[a]	17.2[a]	16.4[a]
Placebo	26.6	23.4	21.8	21.3	20.7	20.9
Imipramine	26.2	22.5	20.0[a]	18.4[a]	17.3[a]	16.4[a]
HAM-D anxiety subscale						
Paroxetine	7.3	6.3	5.6[a,b]	5.2[a]	4.9[a]	4.7[a]
Placebo	7.1	6.4	6.0	5.8	5.6	5.7
Imipramine	7.2	6.5	5.9	5.4[a]	5.2[a]	5.0[a]

*Adapted from Dunbar et al.[31]
[a]$p < .05$ by one-way analysis of variance, drug-placebo.
[b]$p < .05$ by one-way analysis of variance, paroxetine-imipramine.

efficacy on the anxiety/somatization subscale of the HAM-D and the BPRS in the amitriptyline-controlled trial conducted by Zivkov and de Jongh,[17] who found that mirtazapine also has efficacy for anxiety symptoms in depressed patients.

Moreover, in a placebo-controlled trial of paroxetine and imipramine, Dunbar et al.[31] reported subscale scores for each week. From the second week of the study, both active medications were significantly more efficacious than placebo, both in total HAM-D scores and in scores on the cognition, retardation, and sleep subscales. However, early in the study, the two active drugs produced different scores on the anxiety subscale. Paroxetine showed significantly better efficacy in relieving anxiety-related symptoms compared with imipramine (Table 5).

Fawcett et al.[32] analyzed both HAM-A scores and anxiety items on the HAM-D in a meta-analysis of five studies comparing nefazodone (N = 184), imipramine (N = 288), and placebo (N = 345). Patients receiving nefazodone had significantly lower mean agitation item scores on the HAM-D compared with patients receiving either placebo or imipramine at Weeks 1, 3, and 4 ($p \leq .05$), with a trend in this same direction at Week 2 ($p \leq .1$). These studies imply that specific antidepressants may be more effective in certain patients, depending on their most salient symptomatology.

CONCLUSION

Traditional measures of antidepressant drug efficacy show that, overall, these medications are of significant benefit in relieving symptoms of this illness. However, even with intense therapy and close follow-up, as many as one third of patients may fail to respond to antidepressant drug therapy for an acute episode of major depression. To effectively treat all patients with depression, many investigators focus on matching certain subtypes of depression with specific classes of antidepressive drugs. Researchers have also begun to reexamine the validity of outcome measures used to evaluate drug efficacy, particularly with increasing societal pressure to return depressed workers to the workplace as soon as possible. Related issues that have been addressed in clinical trials include how an early patient response to an antidepressant medication may predict long-term outcome to treatment.

Drug names: amitriptyline (Elavil and others), amoxapine (Asendin), fluoxetine (Prozac), fluvoxamine (Luvox), imipramine (Tofranil and others), maprotiline (Ludiomil), mirtazapine (Remeron), nefazodone (Serzone), paroxetine (Paxil), phenelzine (Nardil), sertraline (Zoloft), trazodone (Desyrel), venlafaxine (Effexor).

REFERENCES

1. van Moffaert M, de Wilde J, Vereecken A, et al. Mirtazapine is more effective than trazodone: a double-blind controlled study in hospitalized patients with major depression. Int Clin Psychopharmacol 1995;10:3–9
2. Angst J, Scheidegger P, Stabl M. Efficacy of moclobemide in different patient groups: results of new subscales of the Hamilton Depression Rating Scale. Clin Neuropharmacol 1993;16(suppl 2):S55–S62
3. Davis JM, Wang Z, Janicak PG. A quantitative analysis of clinical drug trials for the treatment of affective disorders. Psychopharmacol Bull 1993;29: 175–181
4. Coryell W, Coppen A, Zeigler VE, et al. Early improvement as a predictor of response to amitriptyline and nortriptyline: a comparison of 2 patient samples. Psychol Med 1982;12:135–139
5. Boyer WF, Feighner JP. Clinical significance of early non-response in depressed patients. Depression 1994;2:32–35
6. Nierenberg AA, McLean NE, Alpert JE, et al. Early nonresponse to fluoxetine as a predictor of poor 8-week outcome. Am J Psychiatry 1995;152: 1500–1503
7. Bremner JD, Abrahams LM, Crupie JE, et al. Multicenter double-blind comparison of nomifensine and imipramine for efficacy and safety in depressed outpatients. J Clin Psychiatry 1984;45(4, sec 2):56–59
8. Derivan A, Entsuah AR, Kikta D. Venlafaxine: measuring the onset of antidepressant action. Psychopharmacol Bull 1995;31:439–447
9. Gasperini M, Gatti F, Bellini L, et al. Perspectives in clinical psychopharmacology of amitriptyline and fluvoxamine: a double-blind study in depressed inpatients. Neuropsychobiology 1992;26:186–192
10. Bremner JD. A double-blind comparison of Org 3770, amitriptyline, and placebo in major depression. J Clin Psychiatry 1995;56:519–525
11. Smith WT, Glaudin V, Panagides J, et al. Mirtazapine vs amitriptyline vs placebo in the treatment of major depressive disorders. Psychopharmacol Bull 1990;26:191–196
12. Claghorn JL, Lesem MD. A double-blind placebo-controlled study of ORG 3770 in depressed outpatients. J Affect Disord 1995;34:165–171
13. Montgomery SA. Are 2-week trials sufficient to indicate efficacy? Psychopharmacol Bull 1995;31:41–44
14. Gerson SC, Plotkin DA, Jarvik LF. Antidepressant drug studies, 1964 to 1986: empirical evidence for aging patients. J Clin Psychopharmacol 1988; 8:311–322
15. Dunner DL. Therapeutic considerations in treating depression in the elderly. J Clin Psychiatry 1994;55(12, suppl):48–58
16. Halikas JA. Org 3770 (mirtazapine) versus trazodone: a placebo controlled trial in depressed elderly patients. Human Psychopharmacology 1995;10: S125–S133
17. Zivkov M, de Jongh. Org 3770 versus amitriptyline: a 6-week randomized double-blind multicentre trial in hospitalized depressed patients. Human Psychopharmacology 1995;10:173–180
18. Einarson TR, Arikian S, Sweeney S, et al. A model to evaluate the cost-effectiveness of oral therapies in the management of patients with major depressive disorders. Clin Ther 1995;17:136–153
19. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, Third Edition, Revised. Washington, DC: American Psychiatric Association; 1987
20. Zivkov M, Roes KCB, Pols AG. Efficacy of Org 3770 (mirtazapine) vs amitriptyline in patients with major depressive disorder: a meta-analysis. Human Psychopharmacology 1995;10:S135–S145
21. Bremner JD, Smith WT. Org 3770 vs amitriptyline in the continuation

treatment of depression: a placebo controlled trial. Eur J Psychiatry 1996; 10:5–15
22. Mullin J, Lodge A, Bennie E, et al. A multicentre, double- blind, amitriptyline-controlled study of mirtazapine in patients with major depression. J Psychopharmacol (Oxf) 1996;10:235–240
23. Wheatley D, Kremer C. A randomized, double-blind comparison of mirtazapine and fluoxetine in patients with major depression. In: New Research Program and Abstracts of the 150th Annual Meeting of the American Psychiatric Association; May 20, 1997; San Diego, Calif. Abstract NR208:124
24. Elkin I, Shea MT, Watkins JT, et al. National Institute of Mental Health treatment of depression collaborative research program: general effectiveness of treatments. Arch Gen Psychiatry 1989;46:971–982
25. Workman EA, Short DD. Atypical antidepressants versus imipramine in the treatment of major depression: a meta-analysis. J Clin Psychiatry 1993; 54:5–12
26. Byerley WF, Reimherr FW, Wood DR, et al. Fluoxetine, a selective serotonin uptake inhibitor, for the treatment of outpatients with major depression. J Clin Psychopharmacol 1988;8:112–115
27. 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(3, sec 2):53–58
28. Mintz J, Mintz LI, Arruda MJ, et al. Treatments of depression and the functional capacity to work. Arch Gen Psychiatry 1992;49:761–768
29. Frank E, Kupfer DJ, Perel JM, et al. Three-year outcomes for maintenance therapies in recurrent depression. Arch Gen Psychiatry 1990;47: 1093–1099
30. Sitsen JMA, Moors J. Mirtazapine, a novel antidepressant, in the treatment of anxiety symptoms: results from a placebo-controlled trial. Drug Invest 1994;8:339–344
31. Dunbar GC, Cohn JB, Fabre LF, et al. A comparison of paroxetine, imipramine and placebo in depressed out-patients. Br J Psychiatry 1991;159: 394–398
32. Fawcett J, Marcus RN, Anton SF, et al. Response of anxiety and agitation symptoms during nefazodone treatment of major depression. J Clin Psychiatry 1995;56(suppl 6):37–42