Onset and Early Behavioral Effects of Pharmacologically Different Antidepressants and Placebo in Depression

Martin M Katz[1], Janet L Tekell[2], Charles L Bowden[1], Steve Brannan[3], John P Houston[4], Nancy Berman[5] and Alan Frazer*[6,7]

[1]Department of Psychiatry, The University of Texas Health Science Center, San Antonio, TX, USA; [2]Department of Psychiatry, Dallas VA Medical Center, Southwest Medical Center, Dallas, TX, USA; [3]Cyberonics, Houston, TX, USA; [4]Eli Lilly and Company, Indianapolis, IN, USA; [5]Harbor UCLA Medical Center, Torrance, CA, USA; [6]South Texas Veterans Health Care System, San Antonio, TX, USA; [7]Department of Pharmacology, The University of Texas Health Science Center, San Antonio, TX, USA

This study was aimed at resolving the time course of clinical action of antidepressants (ADs) and the type of early behavioral changes that precede recovery in treatment-responsive depressed patients. The first goal was to identify, during the first 2 weeks of treatment, the onset of clinical actions of the selective serotonin reuptake inhibitor (SSRI), paroxetine, and the selective noradrenergic reuptake inhibitor, desipramine (DMI). The second aim was to test the hypothesis that the two pharmacologic subtypes would induce different early behavioral changes in treatment-responsive patients. The design was a randomized, parallel group, placebo-controlled, double-blind study for 6 weeks of treatment following a 1-week washout period. The study utilized measures of the major behavioral components of the depressive disorder as well as overall severity. The results indicated that the onset of clinical actions of DMI ranged from 3 to 13 days, averaged 13 days for paroxetine, and was 16–42 days for placebo. Furthermore, as hypothesized, the different types of ADs initially impacted different behavioral aspects of the disorder. After 1 week of treatment, DMI produced greater reductions in motor retardation and depressed mood than did paroxetine and placebo, and this difference persisted at the second week of treatment. Early improvement in depressed mood–motor retardation differentiated patients who responded to DMI after 6 weeks of treatment from those that did not. Paroxetine initially reduced anxiety more in responders than in nonresponders, and by the second week, significantly improved depressed mood and distressed expression in responders to a greater extent. Depressed patients who responded to placebo showed no consistent early pattern of behavior improvement. Early drug-specific behavioral changes were highly predictive of ultimate clinical response to the different ADs, results that could eventually be applied directly to clinical practice.
Neuropsychopharmacology (2004) **29,** 566–579, advance online publication, 19 November 2003; doi:10.1038/sj.npp.1300341

Keywords: onset of antidepressant effects; antidepressants; desipramine; paroxetine; unipolar depression; norepinephrine; serotonin

INTRODUCTION

Despite several decades of research on the mechanisms of action of antidepressants (ADs), the specific therapeutic actions that bring about recovery in depressed patients remain unresolved. In addition, the time of onset of drug-induced clinical actions, the types of behavioral change that precede recovery in treatment-responsive patients, and whether different types of ADs produce similar patterns of early behavioral changes in treatment-responsive patients are not well characterized. We previously reported that the onset of improvement in several symptom domains in hospitalized patients with major depressive disorder who

*Correspondence: A Frazer, Department of Pharmacology, University of Texas Health Science Center, 7703 Floyd Curl Drive, MSC 7764, San Antonio, TX, USA, Tel: +1-210-567-4205,
E-mail: frazer@uthscsa.edu

were treated with tricyclic antidepressants (TCAs) occurred within the first 10 days of treatment (Katz *et al*, 1987). This finding was at variance with the then current view that despite the almost immediate impact of ADs on the functioning of central monoamine transmitter systems, the clinical actions of the drugs 'lag', that is, were not thought to occur until after 2–3 weeks of treatment (see Gelenberg and Chesen, 2000). These early clinical effects of TCAs, found in patients who would eventually respond to the drugs, were also contrary to the results of much of the research conducted during the 1980s (eg Quitkin *et al*, 1984). In our earlier study (Katz *et al*, 1987), the clinical effects detected within the first 2 weeks were large and were predictive of eventual positive clinical outcome. Findings from a number of other studies (Small *et al*, 1981; Coryell *et al*, 1982; Khan *et al*, 1989; Nagayama *et al*, 1991; Stassen *et al*, 1993; Boyer and Feighner, 1994) have also confirmed the positive role of early clinical actions in predicting treatment outcome. The capacity to predict outcome early in treatment would not only shorten the length of a treatment trial but should also decrease morbidity and the

risk of suicide from prolonged depressive symptoms (Kiloh et al, 1988).

Our earlier study did not include a placebo control group. Consequently, questions remained about whether these early clinical effects in treatment responders were due to drug or to a placebo response. The lack of a placebo control group in several of the other prediction studies, for example, Nagayama et al (1991), also brought into question the etiology of these reported early clinical actions. Recent reviews of relevant research on the issue of the onset of action of AD drugs (Stassen et al, 1993; Katz et al, 1996/1997; Gelenberg and Chesen, 2000) concluded that none of these studies, including ours, had fully met the requirements of a definitive study on response onset. Thus, the question as to whether ADs do or do not induce specific predictive behavioral actions within the first 2 weeks of drug treatment is still unresolved. Here, we report a randomized, parallel group, placebo-controlled study designed to address this question. The two primary aims of this study were first to determine, through intensive behavioral examination during the first 2 weeks, the time course and onset of action of the selective serotonin reuptake inhibitor (SSRI), paroxetine, and the selective norepinephrine reuptake inhibitor (NRI), desipramine (DMI). The second aim was to compare early behavioral changes elicited by the two types of ADs to those associated with treatment with placebo.

In contrast to our earlier study that used TCAs, imipramine, and amitriptyline, which have similar pharmacologic effects, the current study evaluated paroxetine and DMI, two ADs with distinct initial pharmacologic profiles (Frazer, 1997). The study focuses on the behavioral impact of the different drug types during the first 2 weeks of treatment and applies advanced statistical methods to determine rate of improvement, timing, and course of early clinical actions. This study is not a clinical trial aimed at evaluating the efficacy of the two active drugs. Rather, the drugs were selected as they have been shown in numerous studies to be efficacious and they are representatives of the SSRI and selective NRI classes of ADs (see Frazer, 1997). The design allowed us to test hypotheses based on earlier findings (Katz et al, 1994) that early behavioral change to the SSRI, paroxetine, would be on anxiety, whereas those to the selective NRI, DMI, would be on motor activity and that such early improvement would be predictive of a positive outcome at the end of treatment.

METHODS AND PROCEDURES

Research Design

The design entailed randomized, parallel group, double-blind assignment of patients to DMI, paroxetine, or placebo for a period of 6 weeks.

Patient Sample

Patients with a diagnosis of primary major depression, unipolar type, single, or recurrent episode were identified from newly admitted inpatients. All subjects provided written informed consent and the study was conducted as approved by the University of Texas Health Science Center at San Antonio's Institutional Review Board (IRB) and the Dallas VA Medical Center's IRB. Diagnostic interviews were conducted using the Structured Clinical Interview for DSM-IIIR (SCID) (Spitzer and Williams, 1983) and medical chart review. Diagnoses were confirmed at weekly research conferences. The exclusion criteria included a history of neurologic disease, recent head trauma, current psychotic symptoms, active substance dependence, unstable medical problems, and suicidal thoughts with intent or previous electroconvulsive therapy. Patients were required to score $\geqslant 18$ on the Hamilton Depression Rating Scale (Ham-D, 21-item version). (Hamilton, 1960).

A total of 82 patients were initially enrolled and randomly assigned to treatment. Of these, 12 patients dropped out prior to receiving the minimum 3 weeks of treatment to qualify for the evaluable sample. The reasons for their dropping out are indicated below. A total of 70 patients completed the requirements of the protocol. The patients, 58 male and 12 female, averaged 46 years, with an age range between 20 and 69 years. All were diagnosed with major unipolar depressive disorder and most patients were moderately (56%) to severely (38%) ill (Clinical Global Improvement (CGI)) (Guy, 1976). In all, 78% had had at least one prior episode, with 18 months as the average length of the current episode. The average patient score at baseline on the Ham-D for the sample was 23.5 ± 4.8 ($X + SD$) and these scores were essentially equivalent for the patients at the two sites, being 23.8 ± 6.4 in San Antonio and 22.2 ± 3.7 in Dallas.

Baseline Period

Study subjects were treated on an in-patient research unit for 6 weeks at two Veterans Administration Hospitals. All patients were maintained on placebo during an initial 1-week drug 'washout' period prior to beginning a 6-week treatment period. Use of a placebo run-in period in this study was primarily carried out to eliminate pre-existing drug that could confound identification of early behavioral changes related to the study medications. Clinical response at the end of the pretreatment period was measured using the Ham-D and CGI (Guy, 1976) scales. Regardless of the initial level of severity, a patient was not included in the treatment study if by day 7 the patient no longer scored $\geqslant 18$ on the Ham-D. During this baseline period, 20 patients were excluded from the randomization stage primarily because they changed their mind about participating in the study or did not meet the inclusion criteria.

Treatment Period

After the 7-day pretreatment period, patients were randomly assigned to either DMI, paroxetine, or placebo. The dosage of DMI was started at 50 mg and was raised as necessary to a maximum of 350 mg/day to reach a blood level of $\geqslant 125$ ng/ml by 7 days. Steady-state concentrations of DMI were reached by day 13 for 80% of the patients. Of the initial patient sample entered into the study, 29 were assigned to treatment with DMI. Of these, three dropped out within 2 weeks due to side effects. In all, 28 patients were assigned to paroxetine. Of these, four did not complete at least 3 weeks of the protocol, two of which were due to

personal choice and two because of protocol violation. The dosage for paroxetine ranged from 20 to 60 mg daily. The dosage was adjusted to achieve a minimum steady-state serum concentration of drug of at least 10 ng/ml. Although there is no current evidence of a relationship between blood level and clinical response to paroxetine, measurement of levels assured adequate absorption and metabolism of paroxetine. Steady-state concentrations of paroxetine were reached by day 6 for 94% of these patients. A total of 25 patients were assigned to placebo, and five did not complete 3 weeks of treatment; for four, this was due to personal choice (changed mind because duration of hospitalization was too long) and one had a protocol violation. Thus, of the 102 patients enrolled in the study originally, 32 did not complete the study, 20 of whom dropped out during the initial 1-week placebo period. Drug plasma levels were monitored daily during the first week of the study, twice weekly for the next 2 weeks, then once a week until the end of the study. Thus, we were able to assure equivalent treatment in the active medication groups.

As the focus of the study was on early behavioral change, it was decided *a priori* that patients who did not complete at least 3 weeks of treatment would not generate useful data for the analyses. As just indicated, 12 patients were, therefore, not included in the analyses. However, of the 70 patients included in the data analyses, five completed 5 weeks of treatment, one completed 4 weeks, and three completed 3 weeks. For such patients, their last behavioral scores were considered to be their final score and they were assigned a behavioral outcome based on these scores.

Measurement of Behavioral Components and Severity of the Depressed State

To address the questions and hypotheses raised, measures of specific behavioral components as well as the severity of the disorder are required. In this multivantaged approach (Katz *et al*, 1984), doctor ratings included the Ham-D, CGI, Schedule for Affective Disorders, and Schizophrenia-Change Version (SADS-C, Endicott and Spitzer, 1978), Global Assessment Scale (GAS, Endicott *et al*, 1976), Video Interview Behavioral Evaluation Scales (VIBES) (Katz and Itil, 1974); nurse ratings: the Affective Disorder Rating Scale (ADRS) (Murphy *et al*, 1980), NIMH Mood Scale (Raskin *et al*, 1969), Global Ward Behavior Scale (Raskin *et al*, 1969); patient self-reports included the Symptoms Checklist (SCL-90; Derogatis *et al*, 1974); NIMH Mood Scale (Raskin *et al*, 1969); and the Buss–Durkee Hostility Inventory (Buss and Durkee, 1957). Other measures of the components of depression, for example, psychomotor performance, de-rived from our earlier work, are described below. These have been described previously (Katz *et al*, 1982, 1984, 1989).

Behavioral components (state constructs). Following administration of the rating scales at baseline to a broad sample of patients with affective disorders and normal subjects, (Katz *et al*, 1984), psychometric analyses of scores obtained from the various instruments led to a set of 11 state constructs that measure the major components of depression. These studies included tests of discriminant and construct validity (Katz *et al*, 1984, 1987). The 11 include

disturbed emotions and abnormalities of motor movement, cognition, social behavior, and somatic functioning, specifically, (1) depressed mood, (2) anxiety, (3) retardation of movement and speech, (4) agitation, (5) hostility, (6) somatization, (7) distressed physical, primarily facial, expression, (8) interpersonal sensitivity, (9) positive adaptation that describes assertive, open social behavior in the interview, (10) cognitive impairment, which includes both thinking and concentration problems, and (11) sleep disorder. It is these constructs that are used for statistical analyses and not scores directly obtained on the various behavioral instruments. Based on the hypotheses being tested, construct numbers 1–4 and number 7 were used in the primary statistical analyses.

Severity dimensions. In a principal components analysis of the intercorrelations of the 11 state constructs, three orthogonal dimensions were identified, accounting for approximately 75% of the variance (Katz *et al*, 1982, 1984). The dimensions, which consist of combined constructs, represent three independent components of severity of the disorder and help to explain the major areas of functioning measured: (I) Depressed mood–retardation of movement and speech (DM-MR dimension). (II) Anxiety–agitation–somatization–sleep disorder, apparently reflecting the many aspects of anxiety, psychological and physical. (III) Hostility–interpersonal sensitivity. Based on the hypotheses being tested, the severity dimension, DM-MR, was used in the primary statistical analyses.

Categorical outcome (multicomponent). The procedure applied in this study for identifying 'therapeutic actions' required classifying patients as responders (essentially recovered) or nonresponders (essentially unresponsive to treatment) after 6 weeks of treatment. It is based on the measurement of two major dimensions of overall outcome, general psychopathology, and severity of the depressed state. Patients in the recovered group have to achieve an illness level of no greater than 'mild' severity, demonstrate markedly reduced severity scores, and be reported as having a level of functioning on the GAS that would indicate no further evidence of psychiatric disorder (GAS > 60). Non-recovered patients show minimal to no decrease on the severity of illness scale, that is, are still judged to be moderately ill or worse, show a zero to minimal increase on the global improvement scale, continue to have a Ham-D score above 16, and show a level of functioning on the GAS that reflects the continued presence of severe symptomatology or impairment. The specific operational criteria can be found in Appendix 1 of Katz *et al*, (1984). Even after 6 weeks of treatment, it is not possible to classify some patients as responders or nonresponders. Previously, we classified such partially responsive patients as 'indeterminate'. In this study, which focuses on the identification of clear responders, we have treated the nonresponders and indeterminates as a single group for analysis, subsequently to be called 'nonresponders'.

A second categorical outcome based on the Hamilton Scale, where 'marked response' is defined as a 50% decrease in the total (21 item) score, has also been defined.

Frequency of Measurement

The behavioral measurements were administered twice weekly during the first 3 weeks of treatment and weekly from weeks 4 to 6. The patients were interviewed in each assessment period utilizing the semistructured question format (20–25 min) that accompanies the SADS-C method (Endicott and Spitzer, 1978). On the basis of that interview, the clinician completes the SADS-C ratings, the HAM-D (it is incorporated into the SADS-C interview), and the brief VIBES (25 mood and behavior items). Since the SADS-C and HAM-D are rated during the course of the interview, the entire interview-rating procedure takes about 30 min. The nurse-rating procedures on the ward, that is, the ADRS, Mood Scale, and GWBS, require 30–45 min. The patient completes the full NIMH Mood Scale and SCL-90 and the psychomotor performance tests at baseline and at 42 days (outcome). These procedures were shortened at the interim assessment points, days 7, 10, 13, 17, 21, 28, and 35, through the use of brief forms of the SCL (60 items) and the Mood Scale (33 items). The assessment procedure at baseline took about 60 min. Owing to the relative brevity of the interim assessments (30–35 min), the patients did not find the tests to be laborious or fatiguing.

Biochemical Assays

Plasma concentrations of drugs were measured by high-performance liquid chromatography exactly as described in Javors *et al* (2000).

Statistical Methods

Descriptive statistics, such as means and standard deviations, medians, or frequencies were produced for all variables. Unless stated otherwise, the following analyses were run separately for each of the state constructs, severity dimensions, and the Hamilton total score.

Analysis of treatment efficacy. χ^2 analysis and Fisher's exact test were used to compare the treatment groups on two binary outcome measures, the categorical outcome and the Hamilton Scale outcome of at least a 50% reduction in the total score. Analysis of covariance (ANCOVA) on the final value of each measure (day 42), with treatment as the independent group and covarying for the baseline value, was used to assess the absolute difference between drug groups after 6 weeks of treatment. Post-test contrasts were used for pairwise comparisons. Survival analysis was used to determine if the overall course of change over the 6-week period was different between different drug groups. For these models, the outcome was a 20% or better improvement in the measure that was sustained for the remainder of the study (Stassen *et al*, 1993).

Analysis of early behavioral actions and onset comparing treatment results in all subjects.

1. To determine if the two drugs and placebo differed in terms of specific behavioral changes at days 7, 10, and 13 of treatment, ANCOVA using the baseline value as a covariate as described above was run for each of these days.
2. To determine if one drug acts faster than the other and/ or placebo in reducing the overall severity or specific symptomatic behaviors during this 2-week treatment period, the linear mixed model (Laird and Ware, 1982) was used to model the longitudinal change in each construct. The mixed model is similar to ordinary least-squares regression, but has the advantage of using all the data. Treatment was included in the model and the interaction between treatment and time effect (slopes test) was tested for significance. This model measures the differences in the rate of change or reduction of a given variable among treatments during the first 2 weeks.
3. The 'median time of onset' was defined as the earliest time point at which $\geqslant 50\%$ of patients changed a minimum of 20% on a given behavioral construct, a change that was then sustained throughout the course of treatment. The treatment groups were compared on the construct measures of behavior and the dimensional measures of severity of the disorder.

Analysis of onset of 'therapeutic' action within each treatment group. Using the categorical outcome described above, clear responders were compared with nonresponders. These comparisons were conducted within each of the treatment groups separately and were designed to identify the initial behavioral actions that precede positive therapeutic action for each drug. To identify the times of onset of improvement in those behavioral indices that distinguish treatment responders, ANCOVA, covarying for the baseline value with response as the group effect was used. Effect size statistics (Cohen, 1969) were calculated to ascertain the degree to which the improvements that resulted were clinically recognizable.

Association of early change with outcome at 6 weeks. To determine whether early drug-induced changes in specific behaviors are correlated with the amount of improvement on a primary outcome measure, namely the Ham-D Total Score, percent improvement in a specific behavior at weeks 1 and/or 2 was correlated with % improvement at the end of 6 weeks on the Ham-D.

Prediction of outcome. Logistic regression (Hosmer and Lemeshow, 1989) was used to develop an algorithm for estimating the probability that a patient would recover by 6 weeks of treatment based on values on the behavioral constructs after 1 or 2 weeks of treatment. Different models of individual prediction were tested for each drug independently. We used the results of the earlier analyses to select variables for inclusion in the models, which could include just a single construct or a linear combination of several constructs as independent variables. We did not test models including variables that did not discriminate between recovered and nonrecovered subjects at any of the early time points. For each model, a classification table was developed that assumed different values of the probability of recovery as a threshold to predict that a subject would recover. The predicted outcomes were then crosstabulated with the actual outcomes to determine the

'sensitivity' (percentage of true positives) and 'specificity' (percent of true negatives), at each proposed threshold. The model and threshold that provided the best combination of sensitivity and specificity was then selected as the prediction model for recovery.

RESULTS

Response to Treatment with DMI, Paroxetine, and Placebo at 6 Weeks

DMI was significantly superior to placebo on the $\geqslant 50\%$ improvement Ham-D rating scale criterion, with 62% of patients treated with DMI meeting this criterion compared with 30% of placebo-treated patients achieving it ($p < 0.05$). The percentage of patients treated with paroxetine who had this level of improvement (46%) was not significantly different from the proportions for either DMI or placebo. Essentially identical proportions of responders were found using this criterion when applying the HAM-17 scoring rather than the HAM-21, with the same statistical significance (data not shown). Similar trends were seen with the multicomponent categorical index of response with 62, 46, and 45% of patients treated, respectively, with DMI, paroxetine, or placebo categorized as responders. However, the difference in response between DMI and placebo on this index was not significant because of the higher percentage of patients responding to placebo.

Early Behavioral Actions of DMI, Paroxetine, and Placebo

At 1 week of treatment, DMI significantly improved depressed mood and motor retardation when compared with placebo and when these factors were combined to form the DM-MR-dimensional measure (Table 1, Figure 1a and b). The size of the change in motor retardation was greater than that in depressed mood and the change was more durable (Table 1, Figure 1a) These changes in motor

retardation and in the severity dimension were sustained at 13 days (Table 1, Figure 1a). These effects of DMI were not only significantly greater than those seen in patients treated with placebo but also of those treated with paroxetine as well. In addition, by 13 days, DMI significantly reduced hostility (Figure 2) and distressed expression compared with placebo-treated patients (Table 1). Such effects were also detectable using the more global severity scale, namely the HAM-21, in that DMI caused a significantly greater reduction in the score at 1 week than did either paroxetine or placebo. At 2 weeks, improvement due to DMI was superior to that caused by paroxetine, but not to that due to placebo.

The slopes test indicated a significantly more rapid paroxetine-induced reduction in hostility during the first 2 weeks of treatment than that due to placebo (see Table 1a and Figure 2). Treatment with paroxetine, however, shows no other significant differences from placebo during this period. DMI had greater effects on distressed expression than either paroxetine or placebo had at 2 weeks, and greater effects on cognition than paroxetine had at 2 weeks. No other differences in drug effects on behavioral constructs were found (Table 1).

Therapeutic Actions: Relationship of Early Behavioral Changes to Recovery Status at 6 Weeks

These analyses compared the early behavioral changes in patients who responded with a clear positive outcome to treatment at 6 weeks, and in patients treated with the same drug who were nonresponders at 6 weeks. For DMI-treated patients at 1 week, responders showed significantly greater reductions in motor retardation ($p < 0.001$) and in the dimension, DM-MR ($p < 0.001$), than the nonresponder group (see Table 2a) These differences, particularly in motor retardation and in the severity dimension, DM-MR, constitute 'large'-effect sizes (Cohen, 1969) (Table 2a). At days 10 and 13, the significant improvements in motor retardation and DM-MR for responders were sustained. The

Table 1 DMI vs Paroxetine vs Placebo: Comparison of Behavioral Changes at 1 and 2 Weeks of Treatment

	DMI (n = 26)[a]			Paroxetine (n = 24)			Placebo (n = 20)			Difference in change		
	Baseline	1 week	2 week	Baseline	1 week	2 week	Baseline	1 week	2 week	1 week	2 week	Slopes test
State constructs												
Depressed mood	4.95 ± .27[b]	3.55 ± .34	2.83 ± .33	5.76 ± .27	4.64 ± .32	4.37 ± .34	5.90 ± .23	5.26 ± .34	4.31 ± .35	0.042[c]		
Anxiety	3.63 ± .29	2.55 ± .28	2.11 ± .24	3.74 ± .28	2.90 ± .24	2.55 ± .27	3.65 ± .19	3.01 ± .26	2.61 ± .25			
Motor retardation	2.97 ± .24	2.09 ± .25	1.34 ± .22	2.94 ± .22	3.13 ± .28	2.61 ± .27	3.36 ± .29	3.14 ± .30	2.33 ± .26	0.004[c,d]	<0.001[c,d]	<0.001[c,d]
Hostility	1.37 ± .14	1.11 ± .18	0.87 ± .13	1.54 ± .14	1.23 ± .18	1.12 ± .18	1.55 ± .20	1.60 ± .19	1.55 ± .15		0.011[c]	<0.001[c,e]
Distressed expression	2.75 ± .24	2.17 ± .21	1.54 ± .20	2.76 ± .28	2.27 ± .18	2.12 ± .22	2.71 ± .27	2.50 ± .29	1.97 ± .19		0.024[c,d]	
Severity												
Hamilton Scale	23.4 ± 1.0	15.5 ± 1.4	12.9 ± 1.3	24.0 ± .96	19.7 ± .98	18.2 ± 1.1	25.9 ± .92	20.6 ± 1.1	16.6 ± 1.2	0.023[c,d]	0.023[c]	
DM-MR dimension	3.96 ± .22	2.82 ± .27	2.08 ± .24	4.35 ± .23	3.88 ± .28	3.49 ± .28	4.63 ± .23	4.20 ± .29	3.32 ± .28	0.011[c,d]	<0.001[c,d]	

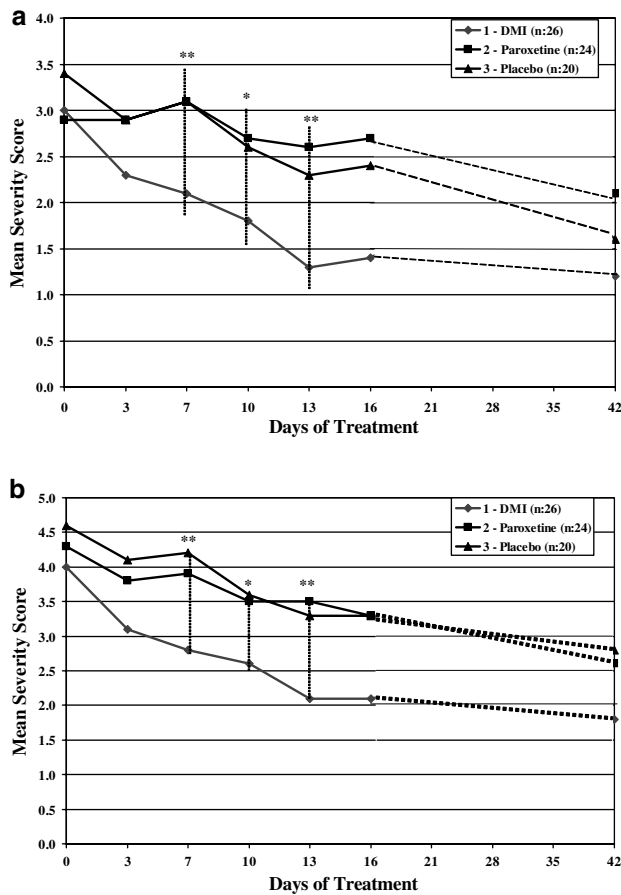[a]Number of subjects.
[b]Mean ± SEM.
[c]DMI > placebo.
[d]DMI > paroxetine.
[e]Paroxetine > placebo.

Figure 2 *Early drug actions*: Comparing treatments across time periods on hostility. DMI and paroxetine reduce hostility at a significantly faster rate during the first 2 weeks than placebo (slopes test, $p < 0.05$). At day 13 of treatment, DMI improved hostility significantly more than placebo (*$p < 0.05$, ANCOVA).
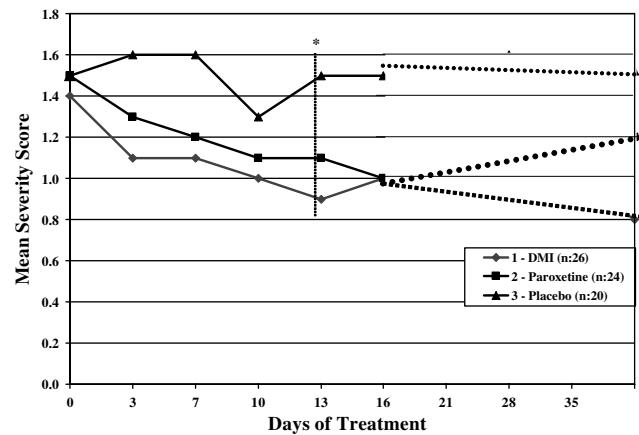
Figure 1 *Early drug actions*: Comparing DMI, paroxetine, and placebo across time periods. (a) *Motor retardation*: At days 7 and 13 of treatment, DMI improved motor retardation significantly more than that caused by either paroxetine or placebo (**< 0.01, ANCOVA) and at day 10, the same effect was observed (*$p < 0.05$, ANCOVA). (b) *Severity dimension*: DM-MR: At days 7 and 13 of treatment, DMI improved the severity dimension significantly more than that due to either paroxetine or placebo (**< 0.01, ANCOVA) and at day 10, the improvement caused by DMI was significantly greater than that caused by paroxetine (*$p < 0.05$, ANCOVA). Although behavioral assessments were carried out at days 21, 28, and 35 of treatment, these time points were not essential for the hypothesis being tested. Consequently, values obtained at these time points are omitted from Figures 1a, b and 2 for the clarity of data presentation.

HAM-21 also showed eventual responders to have a drop in overall severity after 1 week that was significantly greater than that in nonresponders. This difference, though, was not sustained after 2 weeks of treatment. Thus, the early DMI-induced behavioral improvement in motor retardation and depressed mood was large and was associated with its therapeutic action.

For paroxetine, no specific behavioral changes after week 1 were related to response. Anxiety was the earliest behavior influenced early in treatment in paroxetine responders, in that the anxiety construct improved to a significantly greater extent than in nonresponders at 10 days and this effect appeared to be sustained (Table 2b and Figure 3). Consistent with this effect, the dimension of anxiety/agitation/somatization improved to a significantly greater extent at both 10 and 13 days of treatment in responders than in nonresponders (data not shown). At 2 weeks, reductions were observed in depressed mood, motor

retardation, and distressed expression, as well as in the severity dimension, DM-MR, in responders to paroxetine with large effect sizes on depressed mood and DM-MR (Table 2b). This specific pattern of behavioral improvement produced by paroxetine in eventual responders differed from that caused by DMI, in that it appeared later and not only involved anxiety but also depressed mood, motor retardation, and distressed expression. Interestingly, the global severity measure (HAM-21) detected differences between paroxetine responders and nonresponders as early as 1 week, and this difference was sustained at 2 weeks (Table 2b).

For placebo, except for a reversal on motor retardation, that is, the responders were worse than nonresponders at week 1, no statistically significant differences in behavioral change that could distinguish responders from nonresponders were found during the first 2 weeks of treatment (data not shown). Global severity, that is, the HAM-21, also did not change significantly more early in treatment in placebo responders than in nonresponders. Therefore, early improvement among placebo-treated patients was unrelated to eventual responder status.

These results were obtained using the categorical index described earlier that takes into account both improvement and final status on the HAM-D, to divide patients into responders and nonresponders. If the more traditional 50% improvement index in HAM-21 scores is used to create these categories, the population of patients in the two groups is slightly different. It is relevant to ask, then, if the results just presented would be similar if response categorization was based solely on the HAM-21, 50% index. They are; for example, DMI responders still showed significant improvement in motor retardation at 1 and 2 weeks of treatment ($p < 0.005$ and < 0.05, respectively) as compared with that measured in nonresponders. This was true for the DM/MR dimension as well ($p < 0.005$ and < 0.025, respectively, at these times). The HAM-21 total score also dropped significantly more at 1 week ($p < 0.05$), but not at 2 weeks, of treatment in DMI responders than in nonresponders. Patterns of response to paroxetine in

Table 2 Recovered *vs* Nonrecovered Patient Groups: Changes at 1 and 2 Weeks of Treatment

(a) DMI

	Recovered (n = 16)				Nonrecovered (n = 10)				Difference in means (P-values)			Effect size		
	Baseline	1 week	10 days	2 weeks	Baseline	1 week	10 days	2 weeks	1 week	10 days	2 weeks	1 week	10 days	2 weeks
State constructs														
Depressed mood	4.34 ± .26[a]	2.68 ± .34	2.54 ± .35	2.18 ± .38	5.93 ± .42	4.93 ± .39	4.52 ± .48	3.86 ± .44	0.048					
Anxiety	3.12 ± .33	1.94 ± .35	1.62 ± .30	1.62 ± .29	4.44 ± .43	3.52 ± .25	3.18 ± .28	2.90 ± .30		0.033			0.26	
Motor retardation	2.74 ± .32	1.46 ± .23	1.30 ± .22	.83 ± .16	3.33 ± .37	3.10 ± .35	2.63 ± .42	2.14 ± .40	0.001	0.012	0.004	0.87	0.58	0.51
Distressed expression	2.46 ± .27	1.87 ± .26	1.48 ± .30	1.42 ± .26	3.21 ± .41	2.67 ± .31	2.80 ± .32	1.75 ± .30		0.039				
Severity														
Hamilton Scale	21.0 ± .0.8	11.6 ± 1.3	10.9 ± 1.7	10.3 ± 1.5	27.2 ± 1.8	21.8 ± 1.8	19.1 ± 1.7	17.0 ± 1.8	0.016	0.004				
DM-MR dimension	3.54 ± .23	2.07 ± .25	1.92 ± .02	1.50 ± .24	4.63 ± .36	4.02 ± .31	3.57 ± .39	3.00 ± .35	0.002	0.027	0.026	0.87	0.53	0.38

(b) Paroxetine

	Recovered (n = 11)				Nonrecovered (n = 13)				Difference in means (P-values)			Effect size		
	Baseline	1 week	10 days	2 weeks	Baseline	1 week	10 days	2 weeks	1 week	10 days	2 weeks	1 week	10 days	2 weeks
State constructs														
Depressed mood	5.28 ± .38[a]	3.88 ± .43	3.55 ± .28	3.15 ± .30	6.16 ± .36	5.28 ± .38	5.03 ± .44	5.41 ± .39			<0.001			1.64
Anxiety	3.57 ± .41	2.67 ± .39	2.11 ± .21	2.08 ± .31	3.88 ± .39	3.09 ± .31	2.88 ± .31	2.94 ± .40		0.019			0.64	
Motor retardation	2.82 ± .35	2.67 ± .32	2.36 ± .31	2.02 ± .34	3.04 ± .30	3.51 ± .42	2.99 ± .40	3.10 ± .36			0.040			0.67
Distressed expression	2.28 ± .41	1.93 ± .19	1.67 ± .17	1.40 ± .12	3.17 ± .37	2.56 ± .26	2.56 ± .30	2.72 ± .30			0.004			0.44
Severity														
Hamilton Scale	23.4 ± 1.1	16.9 ± .93	17.4 ± 1.1	14.8 ± 101	24.54 ± 1.5	22.0 ± 1.3	19.5 ± 1.4	21.1 ± 1.4	0.004		0.003			
DM-MR dimension	4.05 ± .33	3.28 ± .37	2.96 ± .25	2.58 ± .30	4.60 ± .30	4.40 ± .36	4.01 ± .38	4.26 ± .31			<0.001			1.34

[a]Mean ± SEM.
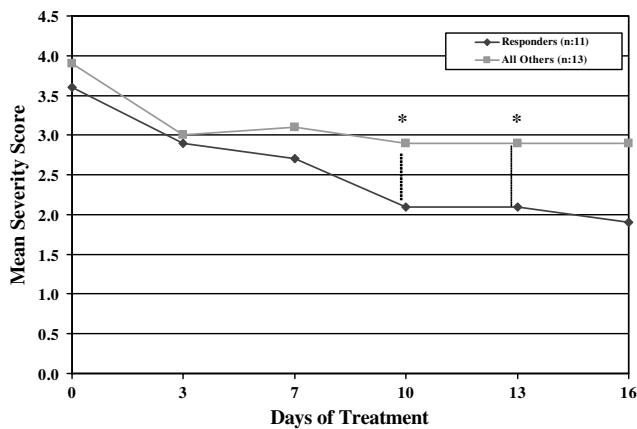
Figure 3 *Paroxetine responders vs nonresponders*: Course of change on anxiety during first 2 weeks of treatment. Paroxetine improved anxiety in eventual responders by day 10 of treatment significantly more than in nonresponders (*$p < 0.05$, ANCOVA).

responders and nonresponders (categorized according to the HAM-21 criterion) were somewhat different from those shown in Table 2b, in that the latter did not detect the early change in anxiety, but showed depressed mood to improve significantly more in responders at 1 week ($p < 0.025$), as well as at 2 weeks. Similarly, motor retardation and the severity dimension, DM/MR, improved significantly more in responders at 2 weeks than in nonresponders. Categorization of placebo responders using the 50% decrease in HAM-21 criterion also did not result in any early improvement in specific behaviors or global severity as compared with that measured in eventual nonresponders. Such data reinforce those presented in Table 2a and b and indicate that early improvement can be used to predict eventual response, irrespective of the criterion used to categorize responders.

Estimating Onset and Time Course for Treatment Responders Through Survival Analyses

Table 3 summarizes the survival analysis results with treatment responders. Responders to DMI showed consistently more rapid improvement on both overall severity (as represented by the dimension of DM-MR (see Figure 4) and several behavioral indices. The time of the onset estimates for paroxetine responders was more delayed, being 13 days for these indices. The time for the first sustained response was even later for the placebo responders (Table 3). Patients who responded to either DMI or paroxetine had earlier improvement on some dimensions than responders to placebo. The initial changes elicited by DMI are different from those elicited in patients treated with the SSRI, the former showing more rapid and greater changes in the motor activity and depressed mood aspects of the disorder.

Predicting Outcome from Early Behavioral Changes

As shown in Table 4 for DMI-treated patients, there are significant relationships between the amount of improvement on several behaviors at the end of week 1 with percent improvement on the Ham-D total score at the end of 6 weeks of treatment. These relationships are also found with

Table 3 Time of Onset[a] of Improvement in Treatment Responders (Survival Analysis)

	Day of onset		
	DMI	**Paroxetine**	**Placebo**
State constructs			
Depressed mood	3	13	16
Anxiety	7	10	16
Motor retardation	7	13	21
Distressed expression	13	13	42
Severity			
Hamilton Scale	7	10	10
DM-MR dimension	3	13	21

[a]*Onset*: Time point (days of treatment) at which $\geq 50\%$ of patients show $\geq 20\%$ of improvement that is sustained through 6 weeks of treatment.
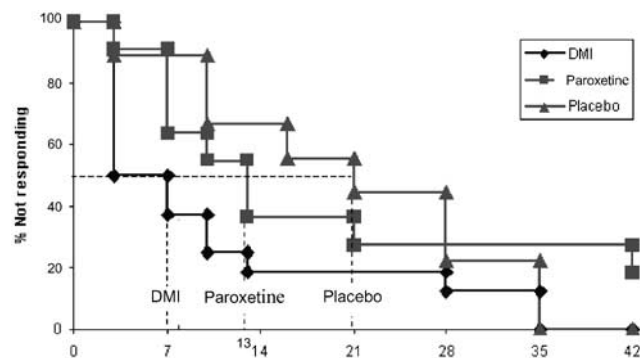
Figure 4 *Median time to onset and time course for treatment responders*: DM-MR (survival analysis). There is a significant difference among these curves ($p < 0.05$, Wilcoxon).

Table 4 Product Moment Correlations of % Change on Behaviors at 1 and 2 Weeks With % Improvement on Ham-D Total Score at 6 Weeks

	DMI (n = 26)		Paroxetine (n = 24)		Placebo (n = 20)	
	1 week	2 week	1 week	2 week	1 week	2 week
State constructs						
Depressed mood	0.56**	0.55**		0.62**		
Anxiety	0.49*	0.50*				
Motor retardation	0.43*	0.45*				
Distressed expression	0.43*					
Severity						
DM-MR dimension	0.56**	0.59**		0.48*		

*$p < 0.05$.
**$p < 0.01$.

the severity indices DM-MR and Hostility–Interpersonal Sensitivity. In general, similar relationships were found for behavioral changes at the end of week 2.

These relationships were stronger for DMI than for paroxetine. At 1 week, only a significant improvement in social behavior (positive adaptation) was associated with response at outcome. At 2 weeks, however, improvement in depressed mood, and the severity indices, anxiety–agitation–somatization and DM-MR, were significantly associated with outcome among paroxetine-treated patients. For placebo, only one significant association was found at week 2 between a behavioral characteristic, interpersonal sensitivity, and outcome, which was reduced more in placebo nonresponders than in responders.

Applying Logistic Regression to Develop Algorithms for Predicting Outcome from Early Behavioral Changes

The best predictive model for the patients treated with DMI was the simple model that had only the dimensional construct DM-MR after 1 week as the independent variable. This model achieved a combination of sensitivity and specificity of 0.90 and 0.88, respectively. The same model at 2 weeks showed less balance, but could achieve 0.70 and 0.81 for a balance in favor of specificity or 0.90 and 0.69 for a balance in favor of sensitivity. There were no variables with any predictive value at 1 week for paroxetine. The best model was based on a linear combination of the distressed expression, cognitive impairment, and depressed mood constructs from the 2-week data that achieved a sensitivity and specificity of 0.85 and 0.91. The model was slightly improved with the addition of the dimensional construct DM-MR, to give 0.92 sensitivity and 0.91 specificity.

DISCUSSION

The most important finding in this study is that it is possible to detect, within the first week or two of treatment, clinically significant improvement in those depressed patients who will eventually respond to pharmacotherapy after a 6-week trial. Furthermore, ADs that target either noradrenergic or serotonergic neurons induce initial improvement on different facets of depressive symptomatology. The selective NRI, DMI, initiated improvement through effects on depressed mood and motor retardation. The SSRI, paroxetine, on the other hand, produced initial improvement somewhat more slowly than DMI, initiating recovery by improving anxiety and somewhat later, depressed mood, distressed expression, and in cognitive functioning. In contrast to these consistent early behavioral effects in patients who eventually responded to these drugs, depressed patients who responded to 6 weeks of treatment with placebo showed no consistent early pattern of behavioral improvement. This finding could mean that the onset of placebo response is not initiated by the same neuronal systems affected by most AD drugs.

When considering these results, it is important to consider the patient population studied. As the study was carried out in two Veterans Administration hospitals, the majority (83%) of patients were male and had illnesses of sufficient severity to warrant hospitalization. Over three-quarters had at least one prior episode of depression and the average duration of the current episode was quite long, about 1.5 years. As in-patients, they were carefully monitored, both with respect to frequent behavioral assessment as well as to the development of side effects. This latter factor probably contributes to relatively few (15%) dropping out after randomization, although about 20% dropped out during the initial 1-week placebo run-in period, primarily because they changed their mind about participating due to the length of hospitalization required.

Another factor to consider is that we are measuring a number of behavioral variables in our patients, which raises the statistical issue of multiple comparisons. Although the behavioral methodology used has 11 state constructs and three severity dimensions, the results of our earlier study (Katz *et al*, 1987) resulted in specific hypotheses (as mentioned in the Introduction) that focused on five constructs—depressed mood, anxiety, distressed expression (the physical expression of depressed mood and anxiety), motor retardation, and hostility—as well as the severity dimension of depressed mood/motor retardation. Thus, there were a limited number of variables involved in the primary analyses. Further, many of the primary findings have *p*-values below the 0.05 level and they are consistent across several methods, for example, ANCOVA, linear mixed model, and survival analyses. We think such considerations strongly reinforce the validity of the major findings.

Efficacy, Onset, and Timing of Clinical Actions

On a traditional efficacy measure such as the 50% drop in the HAM-D score, treatment with DMI caused significantly greater improvement than placebo treatment did, but paroxetine did not. The categorical index resulted in a greater percentage of responders to placebo (45%) than that found using the HAM-D criterion. The likely reason for this difference is that the HAM-D criterion is an absolute one based on a single measure, such that patients who may be just one percentage point apart (49 *vs* 50% decrease) will be given different outcomes. By contrast, the categorical index criterion uses the HAM-D, the CGI, and GAS so as not only to measure improvement and final score but also general psychopathology and social functioning. Thus, the categorical index criterion is less black and white and perhaps more akin to what a clinician might use when evaluating a patient. The difference in the placebo response rates is attributed to three patients whose HAM-21 scores dropped by 44, 46, and 48%, but were classified as responders using the categorical index.

It is likely that the lower response rate found with paroxetine is due to the fact that most of the depressed patients in this study were male. Both Korstein *et al* (2000) and Joyce *et al* (2003) reported that men responded less well to SSRIs such as sertraline and fluoxetine than to TCAs. In these studies, the response rate of male patients to these drugs was about 40–45%, quite similar to the response rate of 46% in this study

In spite of the fact that the different criteria for assigning outcome generated slightly different populations of patients in the two groups, the results obtained using either method of classification were similar. For example, in patients who

responded to DMI after 6 weeks of treatment (using either the HAM-D or categorical index criteria), improvement in motor retardation and depressed mood was detected early in treatment, with such behavioral improvement not observed at these times in nonresponders. Similarly, the ability of paroxetine to initiate improvement in certain behaviors early in treatment in eventual responders was independent of the criterion used to categorize response, although there were some differences in the specific behaviors affected. Importantly, no specific pattern of behavioral improvement was detected early in treatment in placebo responders, irrespective of the criteria used to assign patients to this category.

The results on the onset and timing of response agree generally with those found by Stassen et al (1993, 1997), in that significant improvement occurs as early as 13 days for most ADs. However, our results do not agree with Stassen et al (1993, 1997) on other issues. The results with DMI, for example, indicate that the drug acted more rapidly than 13 days (DMI was not included in the Stassen et al, 1997 analysis). The time pattern with placebo was also different from that found by these authors, in that the placebo responders in this study did not respond similarly to the AD responders. Stassen et al (1993, 1997) conducted their analyses using the total Hamilton Scale score, and that confined their generalizations to global severity only and did not extend to specific behavioral components of the disorders.

Central to the study and to our analysis, however, is the controversy over whether ADs take several weeks to initiate clinical action (eg Quitkin et al, 1984) or whether the behavioral effects begin to act sooner, within the first week, as originally observed (Kuhn, 1958; Kielholz and Poldinger, 1968; Angst, 1970), and as found in a number of subsequent clinical studies (Katz et al, 1987; Khan et al, 1989; Stassen et al, 1993). Large meta-analytic studies analyzing data post hoc from pharmaceutical trials additionally found evidence of early behavioral changes in the first 2 weeks of treatment (Dunbar and Fuell, 1992; Tollefson and Holman, 1994). While such large sample sizes may allow minimal clinical differences to reach statistical significance, these studies also suggest the onset of AD action early in treatment.

The disagreement about the length of time between the effect of the ADs on the functioning of neurotransmitter systems and clinical response has been in great part a function of the ambiguity of the concept of 'onset of clinical response'. The fact that marked or full response, conventionally defined as equal to or greater than a 50% reduction on the Ham-D total score, can require several weeks or even months to occur is not disputed. If recovery is used as the definition of onset of clinical response, there is no question that a significant lag time exists between it and initial neurochemical effects of ADs. However, if the onset of clinical response is defined as 'improvement' (eg, a 20% sustained reduction of the Ham-D total score; Stassen et al, 1993) rather than recovery, the average time for the onset of clinical response to ADs is 13 days, significantly less than the 20 days required for 'full response' (the 50% criterion). Although the figures from similar studies vary somewhat, the results of the Stassen et al (1997) analysis are based on sufficient data to provide reliable estimates of these different time points.

The current study assured equivalent treatment in the medication arms by monitoring drug in plasma and examined improvement in a more detailed manner, that is, replacing the Ham-D total score with multiple measures of the major behavioral components of the disorder. This behavioral methodology results in the detection of specific behavioral change even earlier than 13 days. For DMI, motor activity and depressed mood began to change in 3–7 days; for paroxetine, the reduction in anxiety began by 10 days. Thus, by further refining the measurement of clinical response, it was possible to detect effects on behavior even earlier than previously reported. Further, these behaviors represent critical facets of the depressive disorder and were shown to be correlated with eventual treatment response. They, therefore, can provide the links to later full response and can be tracked alongside further changes in the functioning of the neurotransmitter systems. We conclude that the lag in time between neurochemical and the initial behavioral actions of effective ADs is a matter of days, not many weeks.

Drug-Induced Effects on Monoamines and Behavioral Changes

It is interesting to consider the early behavioral effects produced by DMI and paroxetine from the perspective of their actions on noradrenergic or serotonergic neurons. DMI appeared to produce an initial stimulatory motor effect that was closely, but secondarily, associated with elevation of mood, and this was followed by improvement in anxiety. Although its effect in reducing motor retardation has been noted frequently since the early descriptions by Kielholz and Poldinger (1968) and Carlsson et al (1969), there has been little hard evidence until now that this is a specific clinical action of DMI, that is, an effect greater than that which occurs with other ADs and/or placebo, and furthermore, which is associated with the recovery process. Tonic noradrenergic activity is closely related to the overall state of behavioral arousal, which may be conceptualized as the degree to which an organism is 'engaged' with or responsive to its external environment (Aston-Jones et al, 1991; Jacobs et al, 1991). Brain noradrenergic activity is closely associated with behavioral and physiological indices of arousal, such as locomotion, eye movements, EEG activity, heart rate, and overt behavioral activity (Rasmussen et al, 1986; Berridge and Foote, 1991; Page et al, 1993; Rosario and Abercrombie, 1999). Given the close relationship that has been demonstrated between tonic noradrenergic activity and alertness, behavioral activation and arousal, the tonic elevation of extracellular NE levels produced by selective NRIs such as DMI likely contribute to an alleviation of the 'inhibitory' symptoms of depression such as psychomotor retardation, fatigue/languor, and depressed mood. Our data show that DMI exerts prominent and clinically significant initial effects on such components of depression, consistent with enhancement of noradrenergic transmission.

It is more difficult to obtain consensus about what general behavioral roles to ascribe to serotonin. Some consider serotonin to function mainly in a tonic steady-state mode of activity also related primarily to behavioral arousal (Rueter and Jacobs, 1996), although this relationship can be

modified (Chaouloff et al, 1999). Others, though, consider that even if there is some constant tonic level of serotonergic activity, there can still be local modulation of serotonin release that may be both brain region and stimulus specific (Kirby et al, 1995, 1997). From a behavioral perspective, the activation of serotonergic neurons has been hypothesized to produce behavioral inhibition (Soubrié, 1986; Spoont, 1992; Handley, 1995), that is, serotonin may promote response suppression, inhibition, passivity, and waiting. More specifically at the behavioral level, it has been associated with the regulation of impulsivity (Soubrié, 1986), aggression (Linnoila et al, 1983), and anxiety (Briley and Chopin, 1991). Clinical research reaffirms the association of the functioning of the serotonin system and impulsive, aggressive behavior (Coccaro et al, 1989), and also indicates a specific and strong association with anxiety (Dunbar and Fuell, 1992; Montgomery, 1992; Gorman, 2002). In view of this, treatment with SSRIs might be expected to improve, at least initially, components of depression reflecting arousal, such as anxiety, agitation, and hostility. Consistent with this view, hostility was the earliest behavioral component in all patients that was shown to be reduced. And the earliest behavior that improved in responders to paroxetine compared to paroxetine nonresponders was anxiety. Improvement in depressed mood and cognitive impairment in responders occurred somewhat later. This later behavioral effect of paroxetine may not be due to serotonin exclusively, but to sequential or direct interactions between serotonin and noradrenergic or even dopaminergic systems (Willner, 1997; Bonhomme and Esposito, 1998; Bourin et al, 2001).

Our data show that paroxetine and DMI, prototypes for SSRIs and selective NRIs, initiate behavioral improvement somewhat differently. This leads to the possibility that ADs with different mechanisms of action may initiate their own unique pattern of response. Such data are consistent with those generated in the innovative work of Delgado and Moreno (2000), who demonstrated that abruptly reducing the availability of 5-HT through rapid depletion of dietary tryptophan interrupted the therapeutic effect in patients responsive to an SSRI. When the same procedure was applied to patients responsive to a selective NRI, it had no effect. By contrast, catecholamine depletion transiently reversed responses produced by selective NRIs without affecting the response to SSRIs. This work reinforces the idea that the neurobiological mechanisms underlying responsiveness to different classes of ADs involve initial actions on different neurotransmitter systems. Nevertheless, it is evident that different types of ADs can achieve the same therapeutic efficacy.

Drug Actions on the Dimensional Structure of the Depressive Disorder

To determine how changes in discrete behaviors brought about by drug-induced effects on neuroamine systems resolve the complex, clinical state of depression, it is useful to view the impact of the drugs on the three major severity dimensions of the disorder, that is, DM-MR, anxiety–agitation–somatization ('arousal'), and hostility–interpersonal sensitivity (Katz et al, 1984; Katz and Maas, 1994). The

sequence of behavioral actions would indicate that the selective NRI acted first on the DM-retardation dimension, followed by impact on the anxiety arousal dimension. The secondary effects in the sequence, reduction of anxiety, and hostility might then be associated with the somewhat later effect that DMI has on the serotonin system (Frazer, 2000). The therapeutic pattern for paroxetine as an AD, ie the initial effect on anxiety, has been found in many other studies, so there seems little question that reducing anxiety is a major focus of paroxetine's clinical activity. The pattern is further supported by paroxetine's use as an 'anxiolytic', now FDA approved for the treatment of a range of anxiety disorders (Dunbar and Fuell, 1992; Lydiard and Bobes, 2000; Pollack et al, 2001). The sequence of behavioral changes for paroxetine appears to reflect an initial impact on the serotonergic system, that is, the 'calming' of the negative arousal and hostility dimensions, followed by elevation of mood and improvement in other aspects of the disorder. The results, therefore, support a clinical analysis of AD actions on depression, which views the SSRI and selective NRI drugs as impacting the three major dimensions of the depressed state in contrasting sequences. The NRI first improves motor retardation-depressed mood and, secondly, arousal components of the illness. By contrast, the SSRI reduces arousal first and then depressed mood.

Placebo Response

In this study, responders to placebo treatment responded significantly more slowly than responders to drugs. The placebo responders also showed no differences in early behavioral changes from those who did not respond to placebo. Thus, placebo responders were 'late' in response and showed no pattern of early improvement that was distinctive, as was found with responders to ADs. This result is in contrast to those of Stassen et al (1997) who found the time course of response to placebo to be virtually identical to the time course for responders to a range of active ADs, that is, initial improvement was evident by day 13 for all treatments. The differences in results from the two studies may be due to differences in the clinical populations sampled, although both were primarily in-patient studies. Our data are also in contrast to those of Quitkin et al (1984, 1987), who described a pattern of early response to placebo that was not sustained. The limitations of these studies have been discussed in detail previously (Katz et al, 1996/1997). It is clear that the basis for placebo response is still not well understood. This issue and our results will be explored in greater detail in a subsequent paper.

Prediction of Drug Response

The finding that the onset of therapeutic changes for DMI and paroxetine occurs within the first 2 weeks of treatment suggests that it may be possible in the future to use these early changes to predict outcome at 6 weeks of treatment. As a first step in testing the clinical applicability of these results, we determined whether the amount of improvement produced by each of the drugs on the various behavioral and severity indices at the end of weeks 1 and 2 was correlated with the amount of improvement on a primary measure of outcome, the Hamilton depression total score, at

6 weeks. If so, formulas could be generated utilizing the findings at weeks 1 and 2, to identify the critical behavioral factors to be used in predicting outcome. Decisions might then be made at 2 weeks of treatment to determine whether a patient should be maintained on the initial drug or shifted to a new treatment. The findings in this study indicate that by the end of the first week of treatment, changes in several behavioral variables produced by DMI are associated with clinical response at 6 weeks. They include the severity dimension, DM-MR, and the behavioral facets, depressed mood, anxiety, and somatization (Table 4). These behavioral associations are sustained at 2 weeks. A reduction in severity dimensions is also associated with outcome for paroxetine, but these associations do not appear until the end of 2 weeks of treatment.

These promising relationships were reinforced by the sensitivity and specificity analyses. It appears that status on the DM-MR severity dimension can be used to predict outcome for DMI at 1 week resulting in 0.90 sensitivity and 0.88 specificity. A combination of behavioral facets can be used for paroxetine, as well as the DM-MR dimension, to also achieve high sensitivity and specificity at 2 weeks. A study would be required to test whether these high estimates are equally valid when applied in the outpatient situation, but it is clear that high-level prediction is quite possible using a limited set of behavioral and severity indices. The generalizability of our findings may be limited by the fact that most of our patients were men of lower socioeconomic status with recurrent depression. This further substantiates the need for replication of this work in an outpatient population.

There has been great difficulty in prior studies trying to identify predictors of AD response based on pretreatment patient characteristics (Joyce and Paykel, 1989). There have been, however, a number of studies over the years (Small *et al*, 1981; Coryell *et al*, 1982; Katz *et al*, 1987; Khan *et al*, 1989; Nagayama *et al*, 1991; Boyer and Feighner, 1994) that have demonstrated a relationship between early response to ADs, mainly tricyclics, and treatment outcome. It is clear from the current study, however, that drug-specific types of behavioral responses in the first 1 or 2 weeks of treatment with DMI or paroxetine are highly predictive of later outcome. Consistent with these results, Szegedi *et al* (2003) found that improvement in the first 2 weeks in depressed patients treated with either mirtazapine or paroxetine was highly predictive of a positive response after 6 weeks of treatment. Lack of early improvement was also highly predictive of lack of improvement after 6 weeks. Thus, an increasing body of data is emerging that refutes the concept that ADs initiate behavioral improvement quite slowly.

Earlier studies may have insufficiently assessed the nature and sequence of the specific behavioral changes that might accompany early drug-induced changes in the functioning of the monoamine systems. The difficulty in describing this neurochemical-behavioral process in earlier studies was influenced by the use of 'full clinical response' as the criterion of drug action, rather than treating recovery as a process that begins with changes in parts of the disorder and evolves, sometimes rapidly, into resolution of the full disorder. Our results imply that despite the current emphasis in AD development on selective 5-HT receptor

subtypes in order to increase the rate of clinical action (Artigas *et al*, 1996), more rapid improvement is seen with drugs that target the noradrenergic system than serotonergic neurons. It is important to recognize this in future drug development.

ACKNOWLEDGEMENTS

REFERENCES

Angst J (1970). Mechanism of action of imipramine and its therapeutics use. *Tofranil J Angst Bern Stampfli Cie*, BERN, p36.

Artigas F, Romero L, de Montigny C, Blier P (1996). Acceleration of the effect of selected antidepressant drugs in major depression by 5-HT1A antagonists. *Trends Neurosci* **19**: 378–383.

Aston-Jones G, Akaoka H, Charlety P, Chouvet G (1991). Serotonin selectively attenuates glutamate-evoked activation of noradrenergic locus coeruleus neurons. *J Neurosci* **11**: 760–769.

Berridge CW, Foote SL (1991). Effects of locus coeruleus activation on electroencephalographic activity in neocortex and hippocampus. *J Neurosci* **11**: 3135–3145.

Bonhomme N, Esposito E (1998). Involvement of serotonin and dopamine in the mechanism of action of novel antidepressant drugs: a review. *J Clin Psychopharmacol* **18**: 447–454.

Bourin M, Chue P, Guillon Y (2001). Paroxetine: a review. *CNS Drug Rev* **7**: 25–47.

Boyer WF, Feighner JP (1994). Clinical significance of early non-response in depressed patients. *Depression* **2**: 32–35.

Briley M, Chopin P (1991). Serotonin in anxiety: evidence from animal models. In: Sandler M, Coppen A, Hartnett S (eds) *5-Hydroxytryptamine in Psychiatry*. Oxford University Press: Oxford.

Buss AH, Durkee A (1957). An inventory for assessing different kinds of hostility. *J Consult Psychology* **21**: 343–347.

Carlsson A, Corrodi H, Fuxe K, Hokfelt T (1969). Effect of antidepressant drugs on the depletion of intraneuronal brain 5-hydroxytryptamine stores caused by 4-methyl-a-ethyl-meta-tyramine. *Eur J Pharmacol* **5**: 357–366.

Chaouloff F, Berton O, Mormède P (1999). Serotonin and stress. *Neuropsychopharmacol* **21**: 28S–32S.

Coccaro EF, Siever LJ, Klar HM, Maurer G, Cochrane K, Cooper TB *et al* (1989). Serotonergic studies in patients with affective and personality disorders. Correlates with suicidal and impulsive aggressive behavior. *Arch Gen Psychiatry* **46**: 587–599.

Cohen J (1969). *Statistical Power Analysis for the Behavioral Sciences*. Academic Press: New York.

Coryell W, Coppen A, Zeigler VE, Biggs JT (1982). Early improvement as a predictor of response to amitriptyline and nortriptyline: a comparison of two patient samples. *Psychol Med* **12**: 135–139.

Delgado PL, Moreno FA (2000). Role of norepinephrine in depression. *J Clin Psychiatry* **61**(Suppl 1): 5–12.

Derogatis LR, Lipman RS, Rickels K, Uhlenhuth EH, Covi L (1974). The Hopkins Symptom Checklist (HSCL): a measure of primary symptom dimensions. In: *Psychological Measurements in Psychopharmacology: Modern Problems in Pharmacopsychiatry*. Karger: Basel, 7: 79–110.

Dunbar GC, Fuell DL (1992). The anti-anxiety and anti-agitation effects of paroxetine in depressed patients. *Int Clin Psychopharmacol* 6: 81–90.

Endicott J, Spitzer R (1978). A diagnostic interview. The schedule for affective disorders and schizophrenia. *Arch Gen Psychiatry* 34: 837–844.

Endicott J, Spitzer RL, Fleiss JL, Cohen J (1976). The global assessment scale. A procedure for measuring overall severity of psychiatric disturbance. *Arch Gen Psychiatry* 33: 766–771.

Frazer A (1997). Antidepressants. *J Clin Psychiatry* 58(Suppl 6): 9–25.

Frazer A (2000). Norepinephrine involvement in antidepressant action. *J Clin Psychiatry* 61(Suppl 10): 25–30.

Gelenberg AJ, Chesen CL (2000). How fast are antidepressants? *J Clin Psychiatry* 61: 712–721.

Gorman JM (2002). Treatment of generalized anxiety disorder. *J Clin Psychiatry* 63: 17–23.

Guy W (1976). Early clinical drug evaluation program (ECDEU). In: *Assessment Manual of Psychopharmacology*. Department of Health, Education and Welfare: Rockville MD.

Hamilton M (1960). A rating scale for depression. *J Neurol Neurosurg Psychiatry* 23: 56–62.

Handley SL (1995). 5-Hydroxytryptamine pathways in anxiety and its treatment. *Pharmacol Ther* 66: 103–148.

Hosmer DW, Lemeshow S (1989). *Applied Logistic Regression*. Wiley: New York.

Jacobs BL, Abercrombie ED, Fornal CA, Levine ES, Morilak DA, Stafford IL (1991). Single-unit and physiological analyses of brain norepinephrine function in behaving animals. *Prog Brain Res* 88: 159–165.

Javors MA, Houston JP, Tekell JL, Brannan SK, Frazer A (2000). Reduction of platelet serotonin content in depressed patients treated with either paroxetine or desipramine. *Int J Neuropsychopharmacol* 3: 229–235.

Joyce PR, Paykel ES (1989). Predictors of drug response in depression. *Arch Gen Psychiatry* 46: 89–99.

Joyce PR, Mulder RT, Luty SE, McKenzie JM, Rae AM (2003). A differential response to nortriptyline and fluoxetine in melancholic depression: the importance of age and gender. *Acta Psychiatr Scand* 108: 20–23.

Katz MM, Itil TM (1974). Video methodology for research in psychopathology and psychopharmacology. *Arch Gen Psychiatry* 31: 204–210.

Katz MM, Koslow SH, Berman N, Secunda S, Maas JW, Casper R et al (1984). A multi-vantaged approach to measurement of behavioral and affect states for clinical and psychobiological research. *Psychol Rep* 55: 619–671.

Katz MM, Koslow SH, Frazer A (1996/1997). Onset of antidepressant activity: reexamining the structure of depression and multiple actions of drugs. *Depression Anxiety* 4: 257–267.

Katz MM, Koslow SH, Maas JW, Frazer A, Bowden CL, Casper R et al (1987). The timing, specificity and clinical prediction of tricyclic drug effects in depression. *Psychol Med* 17: 297–309.

Katz MM, Maas JW (1994). Psychopharmacology and the etiology of psychopathological states: are we looking in the right way? *Neuropsychopharmacology* 10: 139–144.

Katz MM, Maas JW, Frazer A, Koslow SH, Bowden CL, Berman N et al (1994). Drug-induced actions on brain neurotransmitter systems and changes in the behaviors and emotions of depressed patients. *Neuropsychopharmacology* 11: 89–100.

Katz MM, Robins E, Croughan J, Secunda A, Swann A (1982). Behavioral measurement and drug response characteristics of unipolar and bipolar depression. *Psychol Med* 17: 297–309.

Katz MM, Wetzler S, Koslow SH, Secunda S (1989). Video methodology in the study of the psychopathology and treatment of depression. *Psychiatr Ann* 19: 372–381.

Khan A, Cohen S, Dager S, Avery DH, Dunner DL (1989). Onset of response in relation to outcome in depressed outpatients with placebo and imipramine. *J Affect Disord* 17: 33–38.

Kieholz P, Poldinger W (1968). Die behandlung endogener depressionen mit psychopharmaka. *Dt med Wschr* 93: 701–704.

Kiloh LG, Andrews G, Neilson M (1988). The long-term outcome of depressive illness. *Br J Psychiatry* 153: 752–757.

Kirby LG, Allen AR, Lucki I (1995). Regional differences in the effects of forced swimming on extracellular levels of 5-hydroxytryptamine and 5-hydroxyindoleacetic acid. *Brain Res* 682: 189–196.

Kirby LG, Chou-Green JM, Davis K, Lucki I (1997). The effects of different stressors on extracellular 5-hydroxytryptamine and 5-hydroxyindoleacetic acid. *Brain Res* 760: 218–230.

Kornstein SG, Schatzberg AF, Thase ME, Yonkers KA, McCullough JP, Keitner GI et al (2000). Gender differences in treatment response to sertraline versus imipramine in chronic depression. *Am J Psychiatry* 157: 1445–1452.

Kuhn R (1958). The treatment of depressive states with G 22355 (imipramine hydrochloride). *Am J Psychiatry* 115: 459–464.

Laird NM, Ware JH (1982). Random effects model for longitudinal data. *Biometrics* 38: 963–974.

Linnoila M, Virkkunen M, Scheinin M, Nuutila A, Rimon R, Goodwin FK (1983). Low cerebrospinal fluid 5-hydroxyindoleacetic acid concentration differentiates impulsive from nonimpulsive violent behavior. *Life Sci* 33: 2609–2614.

Lydiard RB, Bobes J (2000). Therapeutic advances: paroxetine for the treatment of social anxiety disorder. *Depress Anxiety* 11: 99–104.

Montgomery SA (1992). The advantages of paroxetine in different subgroups of depression. *Int Clin Psychopharmacol* 6: 91–110.

Murphy DL, Pickar D, Alterman JS (1980). Methods for the quantitative assessment of depressive and manic behavior. In: Burdock EJ, Sudilovsky A, Gershon S (eds) *Quantitative Techniques for the Evaluation of the Behavior of Psychiatric Patients*. Marcel-Decker: New York, 355–392.

Nagayama H, Nagano H, Ikezaki A, Tashiro T (1991). Prediction of efficacy by 1-week test therapy in depression. *J Affect Disord* 23: 213–216.

Page ME, Berridge CW, Foote SL, Valentino RJ (1993). Corticotropin-releasing factor in the locus coeruleus mediates EEG activation associated with hypotensive stress. *Neurosci Lett* 164: 81–84.

Pollack MH, Zaninelli R, Goddard A, McCafferty JP, Bellew KM, Burnham DB et al (2001). Paroxetine in the treatment of generalized anxiety disorder: results of a placebo-controlled, flexible-dosage trial. *J Clin Psychiatry* 62: 350–357.

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

Quitkin FM, Rabkin JG, Ross D, Stewart JW (1984). Identification of true drug response to antidepressants. *Arch Gen Psychiatry* 41: 782–786.

Raskin A, Schulterbrandt J, Reatig N, McKeon JJ (1969). Replication of factors of psychopathology in interview, ward behavior, and self-ratings of hospitalized depressives. *J Nerv Ment Dis* 148: 87–98.

Rasmussen K, Morilak DA, Jacobs BL (1986). Single unit activity of locus coeruleus neurons in the freely moving cat. I. During naturalistic behaviors and in response to simple and complex stimuli. *Brain Res* 371: 324–334.

Rosario LA, Abercrombie ED (1999). Individual differences in behavioral reactivity: correlation with stress-induced norepinephrine efflux in the hippocampus of Sprague–Dawley rats. *Brain Res Bull* 48: 595–602.

Rueter LE, Jacobs BL (1996). A microdialysis examination of serotonin release in the rat forebrain induced by behavioral/ environmental manipulations. *Brain Res* **739**: 57–69.

Small JG, Milstein V, Kellams JF, Small IF (1981). Comparative onset on improvement in depressive symptomatology with drug treatment electroconvulsive therapy, and placebo. *J Clin Psychopharmacol* **1**(Suppl 6): 62S–69S.

Soubrié P (1986). Reconciling the role of central serotonin neurons in human and animal behavior. *Behav Brain Sci* **9**: 319–363.

Spitzer RL, Williams JBW (1983). *Instruction manual For the Structured Clinical Interview for DSM III (SCID), Revision.* Biometrics Research Department, M.Y. State Psychiatric Institute: New York.

Spoont MR (1992). Modulatory role of serotonin in neural information processing: implications for human psychopathology. *Psychol Bull* **112**: 330–350.

Stassen HH, Angst J, Delini-Stula A (1997). Delayed onset of action of antidepressant drugs? Survey of recent results. *Eur Psychiatry* **12**: 166–176.

Stassen HH, Delini-Stula A, Angst J (1993). Time course of improvement under antidepressant treatment: a survival-analytical approach. *Eur Neuropsychopharmacol* **3**: 127–135.

Szegedi A, Muller MJ, Anghelescu I, Klawe C, Kohnen R, Benkert O (2003). Early improvement under mirtazapine and paroxetine predicts later stable response and remission with high sensitivity in patients with major depression. *J Clin Psychiatry* **64**: 413–420.

Tollefson GD, Holman SL (1994). How long to onset of antidepressant action: a meta-analysis of patients treated with fluoxetine or placebo. *Int Clin Psychopharmacol* **9**: 245–250.

Willner P (1997). Validity, reliability and utility of the chronic mild stress model of depression: a 10-year review and evaluation. *Psychopharmacol* **134**: 319–329.