One-Year Clinical Outcomes of Depressed Public Sector Outpatients: A Benchmark for Subsequent Studies

A. John Rush, Madhukar Trivedi, Thomas J. Carmody, Melanie M. Biggs, Kathy Shores-Wilson, Hisham Ibrahim, and M. Lynn Crismon

Background: *The symptomatic outcomes of a cohort of public mental health sector depressed outpatients treated for 1 year are described to provide a benchmark for future long-term trials. Baseline moderators of outcome were evaluated.*

Methods: *Outpatients with nonpsychotic major depressive disorder (n = 118) scoring ≥30 on the 30-item Inventory of Depressive Symptomatology–Clinician Rating (IDS-C_{30}) were treated with a medication algorithm and patient/family education package. Response and remission rates were assessed every 3 months with the IDS-C_{30}. Logistic regression analyses evaluated several baseline features in relation to outcome.*

Results: *While response and remission rates increased from 3 to 12 months, the 1-year last observation carried forward (LOCF) response (26.3%) and remission (11.0%) rates were not impressive (sustained response = 14.4%; sustained remission = 5.1%). Younger patients and those with full-time employment (at baseline) were more likely to respond. A shorter length of illness tended to be associated with higher response and remission rates ($p < .10$). Results are generalizable to public sector patients with substantial socioeconomic, general medical, and educational disadvantages who were sufficiently depressed to recommend a change in antidepressant medication.*

Conclusions: *Response and remission rates were modest when compared with outcomes in shorter duration efficacy trials in depressed outpatients with less chronicity, fewer concurrent general medical conditions, and less treatment resistance. Results support the need for more powerful treatments and/or the better delivery of available treatments.*

Key Words: Antidepressant medication, function, long-term outcomes, major depression, symptoms, treatment

Most randomized controlled trials (RCTs) to evaluate medication efficacy for nonpsychotic major depressive disorder (MDD) engage outpatients in 6 to 12 week acute phase trials, although a number of longer-term continuation or maintenance phase trials are available (Prien et al 1973, 1984; Coppen et al 1973; Schou 1979; Kane et al 1982; Bjork 1983; Glen et al 1984; Montgomery et al 1988; Georgotas et al 1989; Frank et al 1990; Rouillon et al 1991; Robinson et al 1991; Doogan and Caillard 1992; Montgomery and Dunbar 1993; Buysse et al 1996; Bauer et al 2000).

Participants in both acute and longer-term efficacy trials are usually recruited by advertising and selected to have minimal concurrent psychiatric (e.g., anxiety disorders, substance abuse) and general medical comorbidities. Those without response to more than one prior medication trial in the current episode are also typically excluded, as are those in a current episode for more than 2 years. In addition, subject samples in these efficacy trials are more often white, better educated, and employed than are patients typically treated in the public sector. Further, participants in efficacy trials are often asked to accept randomization to a placebo, which excludes the most severely ill and those at higher risk for suicide attempts. Consequently, both shorter- and longer-term clinical outcomes of representative outpatients with nonpsychotic MDD treated in daily practice in either the private or public sectors are yet to be well defined (Lepine et al 1997; Tylee et al 1999).

From the Department of Psychiatry (AJR, MT, TJC, MMB, KS-W, HI), University of Texas Southwestern Medical Center, Dallas, Texas; University of Texas College of Pharmacy (MLC), Austin, Texas.

Address reprint requests to A. John Rush, M.D., UT Southwestern Medical Center, Department of Psychiatry, 5323 Harry Hines Blvd., Dallas, Texas, 75390-9086.

-eived September 11, 2003; revised April 13, 2004; accepted April 19, 2004.

The Texas Medication Algorithm Project (TMAP; Rush et al 2003a) included outpatients with psychotic or nonpsychotic MDD diagnosed by participating public sector psychiatrists based on DSM-IV (American Psychiatric Association 1994) criteria who were followed prospectively for at least 12 months from study entry. Participants were treated either with a medication algorithm (Crismon et al 1999) and a patient/family education package (ALGO; Toprac et al 2000) or with treatment as usual (TAU). Primary analyses of the TMAP dataset clearly revealed better clinical and functional outcomes with ALGO than with TAU (Trivedi et al, in press).

Given the dearth of studies describing the clinical outcomes of a large sample of depressed outpatients seen over 12 months in the public sector, we conducted secondary analyses of the TMAP data to provide a benchmark for longer-term efficacy and effectiveness studies. To facilitate comparison with efficacy trials, these analyses focused only on subjects 1) receiving the algorithm package (since they had better outcomes than the TAU group), 2) diagnosed at baseline with nonpsychotic MDD, and 3) having at least a baseline symptom severity that approximates the symptom severity required to enter efficacy trials (i.e., ≥18 on the 17-item Hamilton Rating Scale for Depression [HRS-D_{17}; Hamilton 1960, 1967]; Rush et al 2003b).

Given the descriptive nature of this report, no specific hypotheses were set forth. Rather, we provide these data as potential benchmarks for subsequent studies of longer-term interventions to be used in this type of population. These analyses were aimed at:

1. Defining the overall change in symptom severity based on 30-item Inventory of Depressive Symptomatology–Clinician rating (IDS-C_{30}) (Rush et al 1986, 1996) over the 12-month observation period.
2. Defining the proportion of patients with a response or remission at each quarterly assessment.
3. Defining the proportion of patients with a sustained response or a sustained remission (defined as meeting these thresholds for at least two consecutive quarters).

o-3223/04/$30.00
.oi:10.1016/j.biopsych.2004.04.005

4. Determining whether specific baseline clinical or demographic features (e.g., age, gender, length of illness, or ethnicity) were predictive of response or remission at 12 months or at exit from the study if earlier.

Methods and Materials

The design, rationale, and methods used for the TMAP study have been detailed elsewhere (Rush et al 2003a; Trivedi et al, in press). The study was approved by the University of Texas Southwestern Medical Center at Dallas Institutional Review Board. The following briefly summarizes the study methods.

Participant Selection

All participants completed a written informed consent. Research was conducted in accordance with the ethical principles guiding the conduct of medical research involving human subjects as detailed in the World Medical Association Declaration of Helsinki. Male and female outpatients 18 years of age or older with a clinical diagnosis of nonpsychotic MDD who, in the judgment of their treating physician, required a medication change or were to start a medication for MDD entered ALGO. Patients with schizophrenia, bipolar disorder, schizoaffective disorder, a primary diagnosis of obsessive-compulsive disorder, or an eating disorder (anorexia nervosa or bulimia nervosa) were excluded. Patients requiring hospitalization for detoxification and those receiving mental retardation or Assertive Community Treatment (ACT) program services (Stein and Test 1980) were also excluded at study entry. No other inclusion/exclusion criteria were used.

Treatment

Two different medication management algorithms—a 7-step medication algorithm for nonpsychotic MDD and a 5-step algorithm for psychotic MDD—were used to guide medication management (Crismon et al 1999). Most steps in each algorithm included multiple treatment options, with earlier steps using treatments with the greatest evidence and the best risk/benefit ratios.

To enhance treatment delivery, expert telephone consultation and onsite clinical support from Clinical Coordinators (CCs) were provided. Several tools were used to enhance algorithm implementation, including a detailed treatment manual to guide clinicians in making timely clinical decisions at critical decision points (e.g., at weeks 4 and 8) for each medication step when revisions in treatment strategies or tactics were to be undertaken based on symptomatic response and side-effect burden. The treating physicians along with the CCs implemented the ALGO intervention.

At each clinic visit, routine clinical assessments (and recording on a uniform clinical record) of symptom severity and side-effect burden were used to guide treatment implementation. These assessments included a global assessment of symptoms and associated symptoms, the IDS-C_{30} and the self-report version of the IDS-C_{30} (the IDS-SR_{30}), and a global rating of side-effect burden.

Each patient also received a multistep education package that provided information about the disease, prognosis, treatment options, and medication side effects. The patient education package encouraged patient participation in treatment decisions and treatment adherence (Toprac et al 2000). Eight different educational materials were available. All patients received at least one material during the study, while 72.2% of patients received four, five, or six materials at some point during the study. Only 4.3% of patients received all eight materials.

Study Procedures

Enrollment for the study occurred over 13 months. Research Coordinators (RCs) interviewed patients at baseline and every 3 months thereafter. RCs, while not blind to treatment assignment, were not involved in any treatment. Study participants provided demographic and medical history at baseline.

The IDS-C_{30} collected by the RCs was used to assess symptomatic outcome for this report. The IDS-C_{30} is a 30-item clinician rating that includes all DSM-IV (American Psychiatric Association 1994) diagnostic criterion items for major depressive disorder (MDD; e.g., mood, vegetative, psychomotor, and cognitive symptoms), as well as commonly associated symptoms such as anxiety, irritability, and melancholic and atypical symptom features. The IDS-C_{30} is scored by summing the responses to 28 of 30 items (i.e., either appetite increase or decrease and either weight increase or decrease are scored for a given rating). Each symptom item is scored on a 0 to 3 scale, with higher scores denoting greater symptom severity. The total score range is 0 to 84 (Rush et al 1996).

The patients selected for the present analyses had to receive ALGO treatment and had to score at least 30 at baseline on the IDS-C_{30} to approximate a baseline symptom severity score of $\geq$18 on the HRS-D_{17} that is typically required to enter acute phase randomized efficacy trials (Rush et al 2003b; Trivedi et al 2004).

Measurements were obtained at baseline and at 3, 6, 9, and 12 months following study entry. Altogether, 79.4% of measurements were within $\pm$ 3 weeks of the specified measurement occasion, and 92.4% were within $\pm$ 6 weeks of the specified measurement occasion.

Response was defined a priori as $\geq$ 50% reduction in the baseline IDS-C_{30} total score. Remission was defined as an IDS-C_{30} $\leq$12 based on prior analyses revealing this threshold matches the HRS-D_{17} score of $\leq$7 (Rush et al 2003b). Sustained response (or sustained remission) was defined as achieving the response (or remission) threshold at both 9 and 12 months for the fixed cohort or achieving the relevant thresholds on the last two available consecutive occasions (observed case [OC] and last observation carried forward [LOCF] cohort).

Analytic Methods

Outcomes were reported based on three different samples: 1) OC, 2) LOCF, and 3) fixed membership (subjects who completed all 4 quarterly measurements). Both the OC and LOCF samples required both a baseline and at least one postbaseline measurement. The fixed membership sample required that the IDS-C_{30} total score was available on all subjects at all 5 measurement occasions. All available data within the first year of follow-up were analyzed. Missing data occurred when subjects either discontinued the study or could not be contacted for the requisite measurement occasion.

Each of the following potential baseline moderators were evaluated in regard to their relationship to response and to remission at 12 months (or at exit if earlier) using logistic regression analysis with response and remission as the dependent variables and baseline IDS-C_{30} total score as the covariate in each model. Separate models were formed using each of the following as predictors: age, length of illness, years of education, family size, disposable income, baseline 12-item Short Form Mental Health Summary (SF-12 MHS; Ware et al 1996) score, single versus recurrent MDD, gender, ethnicity (white vs. all others), marital status (married vs. all other categories), presence or absence of concurrent general medical conditions, presence

Table 1. Clinical and Demographic Features of the Sample (n = 118)

Variable	Values
Age (Mean ± SD)	42.1 ± 10.7
Median (range)	41 (19–65)
Female (%)	81.4
Marital Status (%)	
Divorced	32.2
Married	33.9
Single (never married)	16.1
Separated	14.4
Widowed	3.4
Education (Mean ± SD)	11.0 ± 3.0
Median (range)	11 (2–19)
<12 years (%)	50.4
High school diploma/GED (%)	17.1
Partial college (%)	27.4
College degree or greater (%)	5.1
Ethnicity (%)	
African American	8.5
Hispanic	28.0
White	61.0
Other	2.5
Family Size (Mean ± SD)	1.8 ± 1.6
Median (range)	1 (0–7)
Employment (%)	
Full-time	14.8
Part-time	8.7
Unemployed	76.5
Disposable Income[a] (Mean ± SD)	$493 ± 766
Median (range)	$380 ($565–$3,300)
Receiving Public Assistance (%)	32.2
Baseline Scores	
IDS-C_{30} (mean ± SD)	46.0 ± 9.6
Median (range)	45 (30–69)
SF-12 MHS (mean ± SD)	25.2 ± 8.0
Median (range)	24.6 (10.9–46.6)
Current Substance Abuse (%)	
Alcohol abuse only[b]	39.3
Drug abuse only[c]	10.3
Both	6.0
Either	43.6
Concurrent GMCs (%)	
0	42.4
1	35.6
2	7.6
≥3	14.4
Length of Depressive Illness (years)	12.3 ± 11.5
Median (range)	9 (0–52)
(Mean ± SD)	
Recurrent MDD (%)	90.7

DAST, Drug Abuse Screening Test; GED, General Educational Development; GMC, general medical conditions; IDS-C_{30}, 30-item Inventory of Depressive Symptomatology-Clinician Rating; MAST, Michigan Alcohol Screening Test; MDD, major depressive disorder; SF-12 MHS, 12-item Short Form Mental Health Summary.

[a]Total monthly income minus rent or mortgage payment.

[b]Defined by MAST score at baseline ≥5.

[c]Defined by DAST score at baseline >5.

or absence of alcohol- or drug-abuse/dependence based on a drug abuse screening test (DAST; Skinner 1982) score of >5 or a Michigan Alcohol Screening Test (MAST; Selzer 1971) score of ≥5 at baseline, income/public assistance, employment status (full time vs. part time or not employed), and suicidal ideation

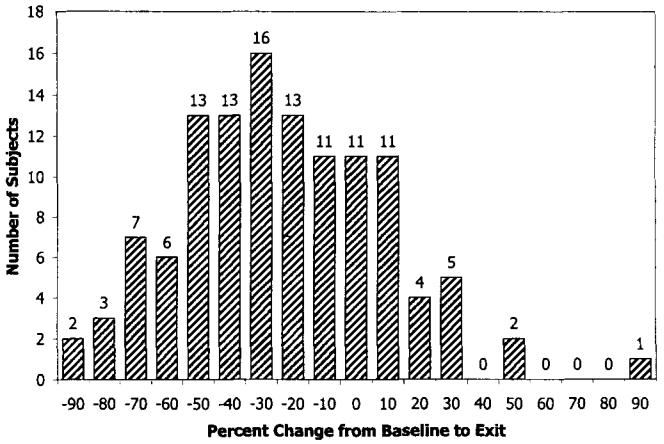

Figure 1. Number of subjects by deciles of percent change in Inventory of Depressive Symptomatology-Clinician Rating (IDS-C_{30}) from baseline to exit (n = 118). Upper end of decile, i.e., categories run from −100 to −90, −89 to −80, −79 to −70, −69 to −60, −59 to −50, etc.

(rated 0–3 on the IDS-C_{30} item). Logistic regression analysis was also used to determine if race (white vs. other), substance abuse (presence/absence), or baseline symptom severity could predict whether or not a patient remained in study at 12 months.

Results

Sample Features

Table 1 summarizes the clinical and demographic features of the total sample (n = 118). The sample had a severe and longstanding illness with an average IDS-C_{30} score of 46, which approximates a HRS-D_{17} total score of 28 (Rush et al 2003b). Overall, subjects contributing data at each measurement occasion were 118 (100%) at baseline, 115 (92.4%) at 3 months, 98 (83.1%) at 6 months, 92 (78.0%) at 9 months, and 78 (66.1%) at 12 months.

Overall, the sample differs sociodemographically from typical treatment samples in that these patients were more likely of ethnic minority status, less educated, less often employed, and more often abusing alcohol or drugs at baseline.

Symptomatic Outcomes

Figure 1 shows the outcomes in terms of categories defined by percentage of reduction in baseline IDS-C_{30} total score at exit (n = 118). Only 31 of 118 (26.3%) achieved at least a 50% reduction in baseline IDS-C_{30} score by exit. If one selects ≥30% reduction to declare at least a clinically meaningful benefit (Figure 1), then 50.8% (60/118) were at least somewhat benefited.

Table 2 shows the response and remission rates for subjects based on the IDS-C_{30} total scores. Response rates, independent of the sample evaluated, were always <30% at any point in this study. Remission rates were even poorer with 10.4% to 12.8% (depending on the sample) achieving remission by 12 months. Note, however, that both response and remission rates rose somewhat over time in each of the three analytic samples (i.e., fixed, OC, LOCF).

To further evaluate the symptomatic outcomes of ALGO treatment, we chose to evaluate the proportion who achieved a sustained response (or sustained remission) by looking at the last 2 available consecutive measurement occasions (Table 3). Only 3.0% to 5.1% achieved a sustained remission, and only 10.5% to 14.4% achieved a sustained response depending on the sample investigated.

Table 2. Response/Remission Rates for Subjects with Nonpsychotic MDD

	Measurement Occasion (Month)				
	BL	3	6	9	12
Fixed Sample (*n*)	(*n* = 67)	(*n* = 67)	(*n* = 67)	(*n* = 67)	(*n* = 67)
IDS-C_{30} response (%)	NA	16.4	25.4	26.9	28.4
IDS-C_{30} remission (%)	NA	4.5	6.0	7.5	10.4
IDS-C_{30} (mean ± SD)	46.5 (9.7)	34.6 (13.8)	30.8 (12.4)	31.5 (14.8)	31.5 (14.1)
SF MHS[a]	25.1 (7.6)	31.9 (10.9)	34.6 (8.5)	34.2 (10.5)	35.8 (10.3)
Observed Case Sample (*n*)	(*n* = 118)	(*n* = 115)	(*n* = 98)	(*n* = 92)	(*n* = 78)
IDS-C_{30} response (%)	NA	19.1	23.5	25.0	29.5
IDS-C_{30} remission (%)	NA	5.2	6.1	6.5	12.8
IDS-C_{30} (mean ± SD)	46.0 (9.6)	34.5 (14.0)	31.8 (12.6)	32.0 (14.9)	31.0 (14.6)
SF MHS	25.2 (8.0)[f]	33.0 (11.5)[b]	34.0 (9.4)[c]	35.9 (11.8)[d]	36.7 (11.0)[e]
LOCF Sample (*n*)	(*n* = 118)	(*n* = 118)	(*n* = 118)	(*n* = 118)	(*n* = 118)
IDS-C_{30} response (%)	NA	18.6	24.6	23.7	26.3
IDS-C_{30} remission (%)	NA	5.1	6.8	7.6	11.0
IDS-C_{30} (mean ± SD)	46.0 (9.6)	34.8 (14.0)	32.2 (13.4)	33.3 (15.3)	32.9 (15.1)
SF MHS[f]	25.2 (8.0)	32.5 (11.4)	33.9 (9.8)	34.7 (11.8)	35.3 (11.7)

Response was defined as an IDS-C_{30} total score of ≤50% of baseline score. Remission was defined as an IDS-C_{30} score ≤12.

BL, baseline; IDS-C_{30}, 30-item Inventory of Depressive Symptoms; LOCF, last observation carried forward; MDD, major depressive disorder; SF MHS, Short Form Mental Health Summary.

[a] *n* = 59
[b] *n* = 108
[c] *n* = 94
[d] *n* = 90
[e] *n* = 77
[f] *n* = 116

To provide a gauge of clinical impact of symptom change/status at exit as compared with status at entry, we calculated the SF-12 MHS total scores for three groups defined by the IDS-C_{30} nonresponders (i.e., <50% reduction in baseline total at exit), remitters (IDS-C_{30} ≤12), and responder-nonremitters. For nonresponders (*n* = 85), the SF-12 MHS went from 25.0 (7.3) to 31.6 (10.2). For remitters, it went from 26.8 (9.4) to 49.4 (7.0) (*n* = 13). For responders/nonremitters, it went from 24.8 (10.2) to 42.4 (9.6) (*n* = 18).

Baseline Predictors of Study Attrition

We conducted a logistic regression analysis to identify potential predictors of staying in the study versus exiting before month 12. Neither substance abuse (yes/no; $p = .9383$) nor baseline severity (IDS-C_{30} total score; $p = .8844$) was predictive of study attrition. Race was a significant predictor (odds ratio [OR] = .37, 95% confidence interval [CI] = [.15–.91] $\chi^2 = 4.70$, $p = .301$). Being white was associated with higher study attrition. Only 64% of white subjects (46/72) were still in the study at month 12, compared with 82% of nonwhites (70% [7/10] for African Americans, 85% [28/33] for Hispanics, and 100% [3/3] for other).

Table 3. Sustained Response/Remission Rates at 12 Months[a]

Sample	Values
Fixed Sample (*n* = 67)	
IDS-C_{30} sustained response	10.5%
IDS-C_{30} sustained remission	3.0%
Observed Case Sample (*n* = 78)	
IDS-C_{30} sustained response	11.5%
IDS-C_{30} sustained remission	3.8%
LOCF Sample (*n* = 118)	
IDS-C_{30} sustained response	14.4%
IDS-C_{30} sustained remission	5.1%

IDS-C_{30}, 30-item inventory of depressive symptoms; LOCF, last observation carried forward; OC, observed case.

[a] Sustained response or remission required that the threshold be met at both 9 and 12 months (fixed cohort) or that the thresholds be met at the last two consecutive available measurement occasions (OC and LOCF samples).

Baseline Predictors of Response and Remission

Significant predictors of response at 12 months (LOCF sample) included both younger age and full-time employment (as opposed to partial or no employment) status at baseline based on the logistic regression analysis. Each additional decade of age reduced the odds of response at 12 months by over one-third (OR = .62; 95% CI = .41–.95; $p = .027$). Response at 12 months was three times more likely for patients with full-time employment than those with either part-time employment or unemployment (OR = 3.05; 95% CI = 1.02–9.09; $p = .046$). The presence of baseline alcohol or drug abuse/dependence tended to be associated with a lower likelihood of response (OR = .46; 95% CI = .19–1.11, $p = .084$).

Turning to remission, younger age was significantly associated with greater chances of remission at 12 months (LOCF sample). Each additional decade of age reduced a patient's odds of remission by about one-half (OR = .46; 95% CI = .24–.87; $p = .016$).

Importantly, in addition to age and employment status, a greater length of illness tended to be associated with a lower likelihood of response and of remission. For each decade increase in the length of illness, the odds of response decreased by about one-third (OR = .66; 95% CI = .43–1.03; $p = .066$), and the odds of remission decreased by about one-half (OR = .52; 95% CI = .25–1.07; $p = .074$).

Finally, of note were the baseline features *not* associated with response or remission. They included education, gender, disposable income, receiving income or food stamp assistance, ethnic-

ity, marital status, and suicidal ideation at baseline. When we divided subjects into white versus nonwhite (including Hispanics and African Americans), nonwhites tended (OR = .48; 95% CI = [.21–1.12] $\chi^2 = 2.91$; $p = .088$) to more likely respond. Nonwhites did not differ from whites in remission rates ($p = .5034$); however, the small sample size could well have precluded finding clinically significant moderators of response or remission.

Discussion

Findings from this study reveal remarkably low response and remission rates, and even lower sustained response and remission rates for public sector outpatients with nonpsychotic MDD. The response rates in this population ranged from 26.3% to 29.5% at 12 months, depending on the sample used, while remission rates ranged from 10.4% to 12.8%. Most clinically relevant, given the chronic nature of MDD in this population, were the very modest sustained response (10.5%–14.4%) and sustained remission (3.0%–5.1%) rates. Importantly, over time there was a slight increase in response (18.6% at 3 months to 26.3% at 12 months LOCF) and remission rates (5.1% at 3 months and 11.1% at 12 months LOCF). Those least likely to respond were older at baseline and lacked full-time employment, while older patients also were less likely to remit. Greater length of illness tended to be associated with lower response ($p = .066$) and lower remission ($p = .074$) rates. Notable for their failure to predict outcome were education, receiving public assistance, ethnicity, and presence/absence of general medical conditions. We did not have a count of number of major depressive episodes (MDEs).

These findings are especially striking given the fact that treatment was delivered under conditions specifically designed to maximize clinical outcomes, which included the use of medication algorithms, additional trained clinical staff support, patient and family education, regular assessment of symptoms and side effects by clinicians at all medication visits, the use of a clinical procedures manual, and expert consultation (via telephone or site visits; Rush et al 2003a; i.e., findings applied to the ALGO group treated in TMAP). Furthermore, the primary analyses of the study clearly showed that this enhanced treatment package (ALGO) produced greater benefits than did treatment as usual, especially for the more severely depressed (Trivedi et al 2004).

Thus, the response/remission rates in this report represent the best case for TMAP subjects in terms of outcomes. In fact, we ran similar analyses on the TAU group (data not shown) defined by the same criteria that we used to isolate the ALGO group in this report. The outcomes in TAU were also poor. At 12 months, the LOCF response rate was 19.4%, and the LOCF remission rate was 7.8%. The LOCF sustained response and remission rates were 9.1% and 3.6%, respectively.

How do these outcomes compare with other reports? In a large, epidemiologically ascertained cohort, Tylee et al (1999) found that 58% were "currently depressed," yet all had seen a health care professional for depression in the previous 6 months, indicating that treatment delivered in representative care settings achieved a 42% response rate in 6 months.

The remission rates in this sample, however, were lower than those reported by others (e.g., Ramana et al 1995; Alexopoulos et al 2000; Uehara et al 1996; Scott et al 1992).

In another study of older patients ($n = 166$), Bosworth et al (2002) found a 45% remission rate in patients ≥60 years of age treated over 1 year with a standard algorithm with remission defined as ≤6 on the Montgomery-Äsberg Depression Rating Scale (MADRS; Montgomery and Äsberg 1979); however, this was not an intent to treat sample. Rather, it was a "completer" sample, which excluded certain subject groups from the analyses (e.g., 17 who died, 30 dropouts, 12 with active substance abuse).

Ezquiaga et al (1998) reported on the 6-month outcome of 90 outpatients receiving pharmacologic treatment at four mental health centers in Madrid. A subsequent report (Ezquiaga et al 1999) revealed the 12-month follow-up results. At 12 months, 59.5% were asymptomatic (HRS-D <8). While this population was drawn from mental health clinics, subjects were excluded if they had drug or alcohol abuse or dependence, antecedents of other psychiatric disease, or if they were on current antidepressant medications before the study. Furthermore, subjects had to be between ages 18 and 65, and none could have been in the current episode of depression for >6 months.

In sum, the available literature has used selected samples rather than broadly representative samples of depressed patients. The samples selected were often less chronically ill with lower rates of concurrent comorbid illness, such as substance abuse.

Why are the response and remission rates in the present study so poor as compared with the usual efficacy RCT findings? There are several possible explanations including a) adherence, b) sociodemographic features of the patient population, c) poor treatment delivery, d) the high rates of concurrent general medical, Axis II, or substance abuse/dependence disorders, e) inclusion of more chronically depressed subjects, and f) inclusion of patients with treatment-resistant depression. Let us consider each potential explanation.

Clearly, treatment adherence can play a major role in the success of a treatment. We did not use blood levels or pill counts to measure adherence in this trial. While we did not measure patient adherence to the medication prescriptions, a patient/family education program was implemented to enhance adherence (Toprac et al 2000). While poor adherence may have reduced the effect of treatment, factors thought to reduce adherence (e.g., baseline substance abuse, minority status) were not predictive of study attrition over the 12-month period.

The enhancement of treatment delivery has been shown to remarkably improve outcome (Katon et al 1995, 1996, 1999, 2002; Thomas et al 2002; Unützer et al 2002). Importantly, some of the studies provided both medication management and patient/family education—as provided in the present report—but they also provided a brief problem-solving therapy. Psychotherapy was largely unavailable in the present trial. This lack of therapy may have contributed to surprisingly poor outcomes.

A second possibility is that the current sample included subjects at substantial socioeconomic disadvantage, thereby reducing the likelihood of response and remission. Specifically, this present sample is less educated, more likely unemployed, and less likely white than efficacy trial samples. Indeed, in this sample, those with full-time employment at baseline were significantly more likely to respond. Supporting this view, Bosworth et al (2002) found that more difficulties in daily living and lower social support were associated with nonremission; however, in a search for baseline moderators of response, we found no effect of education or ethnicity (nor did Bosworth et al 2002); however, full-time employment was associated with a better outcome in this present study.

It is not likely that age per se explains our findings, as the population in this study is only middle-aged, and older age may sometimes reduce the likelihood of response or remission. For example, Ezquiaga et al (1998) found that older age, lower

Global Assessment of Functioning (GAF) at baseline, and the presence of recurrent depression were associated with a lower likelihood of remission (HRS-D ≤ 8; 41/87 [47.1%]) at 6 months. At 1 year, older age, older age at first onset of MDD, lower GAF at baseline, and the presence of recurrent depression were associated with lower remission rates (50/87 [57.5%]; Ezquiaga et al 1999). On the other hand, Hinrichsen and Hernandez (1993) did not find age related to outcome.

Third, what about the quality of treatment? It is true that visit frequency was lower than that typical of efficacy trials. In the first quarter, the average number of clinic visits was 6.5 (SD = 2.2). For the subsequent second, third, and fourth quarters, they were 3.7 (SD = 1.8), 3.2 (SD = 1.6), and 2.6 (SD = 1.3), respectively. On the other hand, as noted above, these subjects were engaged in treatment procedures that exceeded those typical of TAU. Thus, the low rates are not obviously attributable to poor implementation of treatment; however, it should be noted that no patient received electroconvulsive therapy (ECT), even though it was ALGO Step 6 for nonpsychotic MDD. Whether better symptomatic outcomes would have been achieved if ECT were more widely used remains unknown. Whether greater algorithm adherence in general is associated with better outcomes will be the subject of a future report.

The fourth possibility is that the relatively poor long-term outcomes (as compared with those expected based on efficacy trials) were due to the high rates of concurrent comorbid Axis I, II, or III conditions. This population had a substantial rate of concurrent, treated general medical conditions (GMCs; nearly 60%). In addition, about 44% suffered alcohol or drug abuse/dependence based on the DAST and MAST ratings at baseline. We do not have data on the prevalence of Axis II disorders.

It is uncertain whether concurrent GMCs reduce the likelihood of response or remission. For example, Tylee et al (1999) found that 65% of subjects had a concomitant GMC, yet 42% seem to have responded by 6 months. Furthermore, Keitner et al (1992) found that depressed patients with associated concurrent GMCs still had a 48.6% (34/70) response rate at 1 year.

Concurrent active substance abuse or dependence may reduce response and remission rates, since depressed patients without currently active alcohol abuse/dependence were twice as likely to recover from their depression in a 10-year follow-up study than were depressed, currently alcoholic subjects (Mueller et al 1994). It is also likely that those with concurrent substance abuse conditions were more likely than others to comply more poorly with recommended medication treatment; however, antidepressant treatment is effective in those with both depression and alcoholism (Nunes et al 1993).

The fifth potential explanation is that the cohort in this report had longer-standing illness (i.e., more chronic depressions), which could lead to lower response and remission rates. As noted above, efficacy trials often exclude subjects if the current MDE exceeds 2 years. In this sample, the length of the illness (from onset of the first MDE to study entry) was 13.9 years, which is longer than what is seen in practice. For example, Tylee et al (1999) found an average of 45 months between the onset of depression and the research interview; 43% had been ill with depression for >5 years.

In general, longer-standing illness appears to be associated with a worse prognosis. Longer times to remission were associated with greater age, longer episode length, younger age at first onset, and longer length of illness (O'Leary et al 2000). Keller et al (1982) reported that longer illness length predicted a longer time to remit over a 1-year observation period. In a prospective study, Ramana et al (1995) found that if the length of illness was >3 months, it predicted a longer time to remit. Scott et al (1992) also found that a longer length of illness predicted poorer outcome. Longer length of episode was associated with higher recurrence rate after recovery, as was the number of episodes (Mueller et al 1999). In a 1-year study of 166 elderly depressed outpatients, Bosworth et al (2002) found that more MDEs predicted a lower likelihood of remission at 1 year.

A sixth possible explanation is that the current study included a large proportion of subjects who had not responded adequately to one or more prior antidepressant medications. As noted above, efficacy RCTs typically engage symptomatic volunteers with uncomplicated, nontreatment-resistant depressions. Greater treatment resistance, however, is associated with lower response/remission rates (Sackeim et al 2001; Prudic et al 1996). In the present study, we do not have data on the degree of treatment resistance; however, given the nature of the population (substantial disability, unemployment, longstanding illness), it is very likely that many had some degree of treatment resistance.

There remains, as well, an unmeasured variable that could explain our findings. For example, the neurobiology of depression could evolve over time such that those with longer-standing illness have a pathobiology that is simply poorly responsive to currently available treatments. Alternatively, it could be that several of the above potential explanations interact to produce low response and remission rates.

This study has several limitations. First, results are generalizable to adult outpatients with nonpsychotic MDD with substantial social, economic, and educational disadvantages, as well as substantial general medical comorbidity who are treated in the public sector, albeit with a closely managed medication treatment program. Most study participants had not achieved a good result with one or more previous antidepressant medications before entering the study. Secondly, research outcome assessments were not blind, and only quarterly assessments were obtained. On the other hand, these unblinded ratings should, if anything, increase the response/remission rates. Further, adherence by practitioners to the ALGO was adequate, but not extremely high (Trivedi et al, in press); thus, the present results are likely representative of outcomes in these types of patients under somewhat better than average treatment conditions, but they may not represent the efficacy of a more diligent implementation of the ALGO. Finally, the sample size is relatively small so that other predictors of outcome may exist and be clinically significant, yet be undetectable given the current sample size.

These findings suggest that outcomes obtained in short-term efficacy RCTs are not generalizable to these types of patients. Actual patients seen in routine practice typically have more concurrent Axis I, II, and III conditions, and often have longer-standing, more chronic illness course, and/or greater treatment resistance. Given the above-noted studies, results from short-term efficacy RCTs may also not apply to depressed patients treated in representative primary- or specialty-care settings.

The present study cannot determine whether the poor outcomes are due to 1) treatment resistance; 2) concurrent Axis I, II, and III conditions; 3) greater chronicity of illness; 4) socioeconomic disadvantages; 5) less than optimal treatment delivery; or 6) other factors; however, the clinical outcomes obtained in this study clearly suggest that more powerful treatments or the better delivery of available treatments are needed for this patient group.

This research was supported by National Institute of Mental Health Grant No. MH-53799; the Robert Wood Johnson Foundation; the Meadows Foundation; the Lightner-Sams Foundation; the Nanny Hogan Boyd Charitable Trust; the Texas Department of Mental Health and Mental Retardation (MHMR); the Center for Mental Health Services; the Department of Veterans Affairs, Health Services Research and Development Research Career Scientist Award (RCS92-403; T. Michael Kashner, Ph.D., J.D.); the Betty Jo Hay Distinguished Chair in Mental Health and the Rosewood Corporation Chair in Biomedical Science (AJR); the United States Pharmacopoeia Convention, Inc.; Mental Health Connections, a partnership between Dallas County MHMR and the Department of Psychiatry of the University of Texas Southwestern Medical Center, which receives funding from the Texas State Legislature and the Dallas County Hospital District and the University of Texas at Austin College of Pharmacy; and the Southwestern Drug Corporation Centennial Fellowship in Pharmacy (MLC). The following pharmaceutical companies provided unrestricted educational grants: Abbott Laboratories; AstraZeneca; Bristol-Myers Squibb; Eli Lilly & Company; Forest Laboratories; GlaxoSmithKline; Janssen Pharmaceutica; Novartis; Organon; Pfizer, Inc.; and WyethAyerst Laboratories, Inc.

We deeply appreciate the consultations provided by Barbara Burns, Ph.D., Robert Drake, M.D., Susan Essock, Ph.D., William Hargreaves, Ph.D., Teh-wei Hu, Ph.D., Anthony Lehman, M.D., and Greg Teague, Ph.D., without whose expertise this study could not have been accomplished.

We also thank the National and Texas Alliance for the Mentally Ill and the National and Texas Depressive and Bipolar Support Alliance for providing patient and family educational materials. We appreciate greatly the assistance of Gus Sicard, Ph.D., for his translation of the educational, clinical, and research evaluation materials into Spanish.

We gratefully acknowledge the following for their invaluable assistance in the conducting of this study, including Karla Starkweather (Texas Medication Algorithm Project [TMAP] Communications Director, Texas Department of MHMR, Austin). Most important, we wish to express our gratitude to the personnel at each community mental health center for allowing us to conduct this research with their staff and patients, including Andrews Center, Tyler; Center for Health Care Services, San Antonio; Life Management Center, El Paso; Lubbock Regional MHMR Center; the Mental Health and Mental Retardation Authority of Harris County, Houston; Tri-County MHMR Services, Conroe; and Tropical Texas MHMR Services, Edinburgh, Harlingen, and Brownsville. We appreciate the secretarial support of Fast Word, Inc., Dallas, and David Savage, and thank Andrew Sedillo, M.D., for establishing the TMAP website (http://www.mhmr.state.tx.us/centraloffice/medicaldirector/tmap.html).

Alexopoulos GS, Meyers BS, Young RC, Kalayam B, Kakuma T, Gabrielle M, et al (2000): Executive dysfunction and long-term outcomes of geriatric depression. *Arch Gen Psychiatry* 57:285–290.

American Psychiatric Association (1994): Diagnostic and Statistical Manual of Mental Disorders, 4th edition. Washington, DC: American Psychiatric Association Press.

Bauer M, Bschor T, Kunz D, Berghofer A, Strohle A, Muller-Oerlinghausen B (2000): Double-blind, placebo-controlled trial of the use of lithium to augment antidepressant medication in continuation treatment of unipolar major depression. *Am J Psychiatry* 157:1429–1435.

Bjork K (1983): The efficacy of zimelidine in preventing depressive episodes in recurrent major depressive disorders—a double-blind placebo-controlled study. *Acta Psychiatr Scand* 308(suppl):182–189.

Bosworth HB, Hays JC, George LK, Steffens DC (2002): Psychosocial and clinical predictors of unipolar depression outcome in older adults. *Int J Geriatr Psychiatry* 17:238–246.

Buysse DJ, Reynolds CF 3rd, Hoch CC, Houck PR, Kupfer DJ, Mazumdar S, et al (1996): Longitudinal effects of nortriptyline on EEG sleep and the likelihood of recurrence in elderly depressed patients. *Neuropsychopharmacology* 14:243–252.

Coppen A, Peet M, Bailey J, Noguera R, Burns BH, Swani MS, et al (1973): Double-blind and open prospective studies in lithium prophylaxis in affective disorders. *Psychiatria, Neurologia, Neurochirurgia* 76:500–510.

Crismon ML, Trivedi MH, Pigott TA, Rush AJ, Hirschfeld RMA, Kahn DA, et al (1999): The Texas Medication Algorithm Project. Report of the Texas Consensus Conference Panel on medication treatment of major depressive disorder. *J Clin Psychiatry* 60:142–156.

Doogan DP, Caillard V (1992): Sertraline in the prevention of depression. *Br J Psychiatry* 160:217–222.

Ezquiaga E, Garcia A, Bravo F, Pallares T (1998): Factors associated with outcome in major depression: A 6-month prospective study. *Soc Psychiatry Psychiatr Epidemiol* 33:552–557.

Ezquiaga E, Garcia A, Pallares T, Bravo MF (1999): Psychosocial predictors of outcome in major depression: A prospective 12-month study. *J Affect Disord* 52:209–216.

Frank E, Kupfer DJ, Perel JM, Cornes C, Jarrett DB, Mallinger AG, et al (1990): Three-year outcomes for maintenance therapies in recurrent depression. *Arch Gen Psychiatry* 47:1093–1099.

Georgotas A, McCue RE, Cooper TB (1989): A placebo-controlled comparison of nortriptyline and phenelzine in maintenance therapy of elderly depressed patients. *Arch Gen Psychiatry* 46:783–786.

Glen AIM, Johnson AL, Shepherd M (1984): Continuation therapy with lithium and amitriptyline in unipolar depressive illness: A randomized, double-blind, controlled trial. *Psychol Med* 14:37–50.

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

Hamilton M (1967): Development of a rating scale for primary depressive illness. *Br J Soc Clin Psychol* 6:278–296.

Hinrichsen GA, Hernandez NA (1993): Factors associated with recovery from and relapse into major depressive disorder in the elderly. *Am J Psychiatry* 150:1820–1825.

Kane JM, Quitkin FM, Rifkin A, Ramos-Lorenzi JR, Nayak DD, Howard A (1982): Lithium carbonate and imipramine in the prophylaxis of unipolar and bipolar II illness: A prospective, placebo-controlled comparison. *Arch Gen Psychiatry* 39:1065–1069.

Katon W, Robinson P, Von Korff M, Lin E, Bush T, Ludman E, et al (1996): A multifaceted intervention to improve treatment of depression in primary care. *Arch Gen Psychiatry* 53:924–932.

Katon W, Russo J, Von Korff M, Lin E, Simon G, Bush T, et al (2002): Long-term effects of a collaborative care intervention in persistently depressed primary care patients. *J Gen Intern Med* 17:741–748.

Katon W, Von Korff M, Lin E, Simon G, Walker E, Unutzer J, et al (1999): Stepped collaborative care for primary care patients with persistent symptoms of depression: A randomized trial. *Arch Gen Psychiatry* 56:1109–1115.

Katon W, Von Korff M, Lin E, Walker E, Simon GE, Bush T, et al (1995): Collaborative management to achieve treatment guidelines. Impact on depression in primary care. *JAMA* 273:1026–1031.

Keitner GI, Ryan CE, Miller IW, Norman WH (1992): Recovery and major depression: Factors associated with twelve-month outcome. *Am J Psychiatry* 149:93–99.

Keller MB, Shapiro RW, Lavori PW, Wolfe N (1982): Recovery in major depressive disorder: Analysis with the life table and regression models. *Arch Gen Psychiatry* 39:905–910.

Lepine JP, Gastpar M, Mendlewicz J, Tylee A (1997): Depression in the community: The first pan-European study DEPRES (Depression Research in European Society). *Int Clin Psychopharmacol* 12:19–29.

Montgomery SA, Äsberg M (1979): A new depression scale designed to be sensitive to change. *Br J Psychiatry* 134:382–389.

Montgomery SA, Dufour H, Brion S, Gailledreau J, Laqueille X, Ferrey G, et al (1988): The prophylactic efficacy of fluoxetine in unipolar depression. *Br J Psychiatry* 153(suppl 3):69–76.

Montgomery SA, Dunbar GC (1993): Paroxetine is better than placebo in relapse prevention and the prophylaxis of recurrent depression. *Int Clin Psychopharmacol* 8:189–195.

Mueller TI, Lavori PW, Keller MB, Swartz A, Warshaw M, Hasin D (1994): Prognostic effect of the variable course of alcoholism on the 10-year course of depression. *Am J Psychiatry* 151:701–706.

Mueller TI, Leon AC, Keller MB, Solomon DA, Endicott J, Coryell W, et al (1999): Recurrence after recovery from major depressive disorder during 15 years of observational follow-up. *Am J Psychiatry* 156:1000–1006.

Nunes EV, McGrath PJ, Quitkin FM, Stewart JP, Harrison W, Tricamo E (1993): Imipramine treatment of alcoholism with comorbid depression. *Am J Psychiatry* 150:963–965.

O'Leary D, Costello F, Gormley N, Webb M (2000): Remission onset and relapse in depression. An 18-month prospective study of course for 100 first admission patients. *J Affect Disord* 57:159–171.

Prien RF, Klett CJ, Caffey EM (1973): Lithium carbonate and imipramine in prevention of affective episodes. *Arch Gen Psychiatry* 29:420–425.

Prien RF, Kupfer DJ, Mansky PA, Small JG, Tuason VB, Voss CB, et al (1984): Drug therapy in the prevention of recurrences in unipolar and bipolar affective disorders. Report of the NIMH Collaborative Study Group comparing lithium carbonate, imipramine, and a lithium carbonate-imipramine combination. *Arch Gen Psychiatry* 41:1096–1104.

Prudic J, Haskett RF, Mulsant B, Malone KM, Pettinati HM, Stephens S, et al (1996): Resistance to antidepressant medications and short-term clinical response to ECT. *Am J Psychiatry* 153:985–992.

Ramana R, Paykel ES, Cooper Z, Hayhurst H, Saxty M, Surtees PG (1995): Remission and relapse in major depression: A two-year prospective follow-up study. *Psychol Med* 25:1161–1170.

Robinson DS, Lerfald SC, Bennett B, Laux D, Devereaux E, Kayser A, et al (1991): Continuation and maintenance treatment of major depression with the monoamine oxidase inhibitor phenelzine: A double-blind placebo-controlled discontinuation study. *Psychopharmacol Bull* 27:31–39.

Rouillon F, Serrurier D, Miller HD, Gerard MJ (1991): Prophylactic efficacy of maprotiline on unipolar depression relapse. *J Clin Psychiatry* 52:423–431.

Rush AJ, Crismon ML, Kashner TM, Toprac MG, Carmody TJ, Trivedi MH, et al (2003a): Texas Medication Algorithm Project, phase 3 (TMAP-3): Rationale and study design. *J Clin Psychiatry* 64:357–369.

Rush AJ, Giles DE, Schlesser MA, Fulton CL, Weissenburger JE, Burns CT (1986): The Inventory for Depressive Symptomatology (IDS): Preliminary findings. *Psychiatry Res* 18:65–87.

Rush AJ, Gullion CM, Basco MR, Jarrett RB, Trivedi MH (1996): The Inventory of Depressive Symptomatology (IDS): Psychometric properties. *Psychol Med* 26:477–486.

Rush AJ, Trivedi MH, Ibrahim HM, Carmody TJ, Arnow B, Klein DN, et al (2003b): The 16-Item Quick Inventory of Depressive Symptomatology (QIDS), Clinician Rating (QIDS-C) and Self-Report (QIDS-SR): A psychometric evaluation in patients with chronic major depression. *Biol Psychiatry* 54–573–583.

Sackeim HA, Rush AJ, George MS, Marangell LB, Husain MM, Nahas Z, et al (2001): Vagus nerve stimulation (VNS) for treatment-resistant depression: Efficacy, side effects, and predictors of outcome. *Neuropsychopharmacology* 25:713–728.

Schou M (1979): Lithium research at the Psychopharmacology Research Unit, Risskov, Denmark. In: Schou M., Stromgren E (Eds.). *A Historical Account in Origin, Prevention and Treatment of Affective Disorders* (pp. 1–8). London: Academic Press.

Scott J, Eccleston D, Boys R (1992): Can we predict the persistence of depression? *Br J Psychiatry* 161:633–637.

Selzer ML (1971): The Michigan alcoholism screening test: The quest for a new diagnostic instrument. *Am J Psychiatry 1971* 127:1653–1658.

Skinner HA (1982): The drug abuse screening test. *Addict Behav* 7:363–371.

Stein L, Test M (1980): Alternative to mental hospital treatment. I. Conceptual model, treatment program, and clinical evaluation. *Arch Gen Psychiatry* 37:392–397.

Thomas L, Mulsant BH, Solano FX, Black AM, Bensasi S, Flynn T, et al (2002): Response speed and rate of remission in primary and specialty care of elderly patients with depression. *Am J Geriatr Psychiatry* 10:583–591.

Toprac MG, Rush AJ, Conner TM, Crismon ML, Dees M, Hopkins C, et al (2000): The Texas Medication Algorithm Project patient and family education program: A consumer-guided initiative. *J Clin Psychiatry* 61:477–486.

Trivedi MH, Rush AJ, Crismon ML, Kashner TM, Toprac MG, Carmody TJ, et al (in press): Texas Medication Algorithm Project (TMAP): Clinical results for patients with major depressive disorder. *Arch Gen Psychiatry*.

Trivedi MH, Rush AJ, Ibrahim HM, Carmody TJ, Biggs MM, Suppes T, et al (2004): The Inventory of Depressive Symptomatology, Clinician Rating (IDS-C) and Self-Report (IDS-SR), and the Quick Inventory of Depressive Symptomatology, Clinician Rating (QIDS-C) and Self-Report (QIDS-SR) in public sector patients with mood disorders: A psychometric evaluation. *Psychol Med* 34:73–82.

Tylee A, Gastpar M, Lepine JP, Mendlewicz J (1999): Identification of depressed patient types in the community and their treatment needs: Findings from the DEPRES II (Depression Research in European Society II) survey. *Int Clin Psychopharmacol* 14:153–165.

Uehara T, Yokoyama T, Goto M, Ihda S (1996): Expressed emotion and short-term treatment outcome of outpatients with major depression. *Compr Psychiatry* 37:299–304.

Unützer J, Katon W, Callahan CM, Williams JW Jr, Hunkeler E, Harpole L, et al (2002): Collaborative care management of late-life depression in the primary care setting: A randomized controlled trial. *JAMA* 288:2836–2845.

Ware JE, Kosinski M, Keller SD (1996): A 12-item short form health survey (SF-12): Construction of scales and preliminary tests of reliability and validity. *Med Care* 34:220–233.