

Response, Partial Response, and Nonresponse in Primary Care Treatment of Depression

Patricia K. Corey-Lisle, PhD, RN-CSP; Rowena Nash, BA, BS; Paul Stang, PhD; Ralph Swindle, PhD

Background: Depressive disorders are one of the most common reasons for visits to primary care physicians. This study identifies factors related to poor response to depression treatment with selective serotonin reuptake inhibitors (SSRIs) in primary care settings by (1) examining clinical response taking into account treatment, (2) comparing baseline characteristics and outcomes between patients classified by response, and (3) examining characteristics predicting poor response.

Methods: A Randomized Trial Investigating SSRI Treatment (ARTIST) was a prospective naturalistic trial comparing effectiveness of SSRI therapy. Eligible patients were randomized to treatment (N=601) and followed up for 9 months. Treatment patterns were classified as “adequate” (6-month continuous medication), “aggressive” (defined by a treatment algorithm), or “inadequate” (discontinuations) by patient-reported medication use. Clinical response was determined by use of the Symptom Checklist–20 (SCL-20), with patients classified as remitters (score ≤ 6), partial remitters (50% decrease in symptoms), or nonresponders. Groups were compared

on baseline characteristics, functioning, and treatment patterns. Multinomial logistic regression was used to determine predictors of response.

Results: Of patients completing 6-month evaluations (n=482), 46% were classified as nonresponders. Additionally, 53% (n=256) received adequate therapy but did not achieve remission and 13% (n=61) had aggressive therapy associated with treatment resistance. Significant predictors of nonresponse included older age, diagnosis, worse physical functioning, and lower energy level.

Conclusions: A substantial number of adequately treated patients did not respond to antidepressant therapy. Some of these patients may be considered undertreated or treatment-resistant according to current treatment guidelines recommending dose increases or medication switches for less than adequate clinical response.

Arch Intern Med. 2004;164:1197-1204

From Eli Lilly & Company, Lilly Corporate Center, Indianapolis, Ind (Drs Corey-Lisle and Swindle); Department of Psychology, Indiana University, Bloomington (Dr Swindle); Department of Pharmacy, Florida A & M University, Tallahassee (Ms Nash); and Primary Care Network and Galt Associates, Inc, Blue Bell, Pa (Dr Stang). Dr Corey-Lisle is now with Bristol-Myers Squibb, Wallingford, Conn. Drs Corey-Lisle and Swindle are employees and stockholders of Eli Lilly & Company; Ms Nash was an intern for and holds stock in Eli Lilly & Company; and Dr Stang has been a paid consultant for Eli Lilly & Company.

THE ADVENT OF SELECTIVE serotonin reuptake inhibitors (SSRIs) with simpler dosing, more favorable adverse event profiles, and fewer toxic effects in the event of overdoses^{1,2} has resulted in the majority of depressed patients now being seen in primary care settings.³ In fact, depressive disorders are one of the most common reasons for visits to a primary care physician (PCP), with PCPs accounting for 50% to 60% of all antidepressant therapy prescriptions.⁴⁻⁶

There are numerous studies demonstrating the efficacy of antidepressants in randomized clinical trials; however, there remains a substantial group of individuals who do not fully respond to therapy in typical clinical settings.^{4,7,8} A number of factors related to less than optimal treatment outcomes in primary care settings have been identified. These factors include treatment

noncompliance,^{9,10} inability to tolerate medication,¹¹ side effects,¹² social stigma,⁷ inappropriate diagnosis,^{3,13,14} inadequate dosing and early medication discontinuation,⁸ and nonadherence to published treatment guidelines.^{4,15-17}

While there are many studies looking at depression treatment outcomes, few studies have focused on patient-reported treatment patterns and clinical response in naturalistic studies within primary care settings. An examination of medication treatment without assessment of clinical response does not provide information about the adequacy of treatment. The purpose of this study was to (1) look at clinical response taking into account treatment patterns, (2) compare baseline characteristics and treatment outcomes between patients classified by response, and (3) determine if there are characteristics that significantly predict poor response.

STUDY DESIGN

The data for this study were obtained from A Randomized Trial Investigating SSRI Treatment (ARTIST): Comparable Effectiveness of Paroxetine, Fluoxetine, and Sertraline. The study methods are described in detail elsewhere¹⁸ and are summarized here. ARTIST was a prospective, naturalistic evaluation of patients from 2 primary care research networks. Patients enrolled into the study were deemed appropriate for SSRI therapy by their PCP and were followed up through structured telephone interviews for 9 months after enrollment. The differences from “usual care” in this study included randomization to the initial therapy and completion of periodic computer-assisted telephone questionnaires. The goal was to impact usual care as delivered by a PCP in a general practice setting as little as possible, while evaluating clinical response to therapy.

The decision to initiate SSRI therapy was based on the PCP's clinical judgment. At the start of the study, patients were randomized to 1 of 3 SSRI therapies (fluoxetine hydrochloride, paroxetine, or sertraline hydrochloride). After randomization, neither the patient nor the physician was blinded to the assigned treatment. The patient and the PCP made all decisions about dose titration, medication discontinuation, or switches to a different antidepressant. To assess outcomes, telephone interviews were conducted at baseline, and at 1, 3, 6, and 9 months after enrollment. Patients and physicians were explicitly instructed that telephone assessments were strictly for research purposes and not a substitute for regular medical care. Patients were instructed to contact their physician for any concerns with treatment or questions about their care. While the original ARTIST analyses found similar effectiveness for depressive symptoms between the 3 antidepressants,¹⁸ the data provided valuable insights into depression treatment and outcomes in PCP settings.

STUDY SETTING

Patients were enrolled over an 8-month period (April through November 1999). Overall, there were 77 practitioners from a total of 37 sites from clinical practices in 2 primary care research networks. Sites within the networks were heterogeneous in both geographic location and established practice patterns. The Primary Care Network (N=51 study practitioners) is a not-for-profit voluntary organization of more than 10 000 family practitioners, internists, and pediatricians throughout the country. The Duke Primary Care Research Consortium (N=26 study practitioners) is an academic site management organization within the Duke University Health System consisting of over 150 family physicians, internists, and pediatricians who collaborate in adult and pediatric clinical outcomes trials. Although this network was academically owned, all physician participants were practicing clinicians in the community. The median number of patients enrolled per practitioner was 6 (range, 1-30). The study's naturalistic design was intended to reproduce real-world care; therefore, physicians did not receive any additional training or treatment guidelines that would potentially change their usual practice patterns with regard to recognition, decision to treat, or treatment of depressive symptoms.

PATIENT SELECTION

Patients were screened for depressive symptoms by their PCP and were eligible for enrollment into the study if they were 18 years or older, considered appropriate candidates for SSRI antidepressant therapy, and had access to a home telephone. Exclusion criteria included (1) being actively suicidal; (2) current treatment or treatment within past 2 months with an SSRI;

(3) taking a non-SSRI antidepressant either for depression or for a nondepressive disorder at more than low doses; (4) history of bipolar disorder; (5) active substance abuse; and (6) pregnancy or breastfeeding. Additionally, patients with cognitive impairments (eg, dementia or psychosis); inability to read, speak, or write in English; or a terminal illness were excluded from participation.¹⁸ This study was approved by the institutional review boards of Indiana University, Research Triangle Institute, and Duke University and by a central institutional review board for the Primary Care Network.

BASELINE CHARACTERISTICS

Demographic characteristics were recorded upon entry into the study. Background information included age, sex, and ethnicity. History of treatment characteristics and resource utilization included past history of treatment for depression; reported suicidal ideation within 2 weeks of entering study; alcohol use; and use of outpatient services, emergency department care, and inpatient hospitalizations.

OUTCOMES ASSESSMENT

Interviewers used a structured computer-assisted telephone survey to complete outcomes assessments. Information obtained by the telephone interviewers was not disclosed to the PCP, except if the patient was suicidal.

Depression was assessed using the Primary Care Evaluation of Mental Disorders (PRIME-MD) and severity was assessed by use of the Symptom Checklist-20 (SCL-20). The PRIME-MD is a validated measure for determining diagnostic subgroups of depressive disorders (eg, major depression, dysthymia, and minor depression) and has also been widely used in primary care research settings.^{19,20} For this study, the depression module was used to determine the number of *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV)* depressive symptoms and the diagnostic subgroups of depressive disorders. Patients who met criteria for both major depression and dysthymia were recorded as having double depression. The SCL-20 (a modified subscale of the Hopkins Symptom Checklist and Brief Symptom Inventory) has been successfully used in a number of primary care depression trials in which it has demonstrated ability to detect changes in depression severity between treatment groups.^{5,8,16}

Secondary outcomes included additional depressive symptoms, health-related quality of life, and psychological measures. The 36-Item Short-Form Health Survey (SF-36) component summary scores and selected domain scores were used to assess health-related quality of life. The SF-36 and the subscales are well validated and widely used to assess components of health-related quality of life.^{21,22} The mental and physical component summary scores (MCS and PCS) provide measures of general physical or mental health outcomes. The role-emotional subscale specifically measures problems with work or other activities of daily living associated with emotional problems, and the social functioning subscale measures interference with normal social activities due to physical or emotional problems. Additionally, a scale from the RAND Medical Outcomes Study (MOS) was used to assess positive well-being.²³ Fatigue or lack of energy, a core component of depression, was assessed by use of an energy scale, derived from the SCL-20.²⁴ Two other measures consistently correlated with depression, self-reported disability and total number of symptoms, were also used as general measures of impairment.

Medication status was obtained by a 38-item structured computer-assisted interview. The questionnaire was administered by trained interviewers at each of the assessments (baseline and 1, 3, 6, and 9 months). The medication status questions elicited patient reports of medication use including

initiation of medication, continuation of previous therapy, current medication(s), dose of medication, number of times medication was taken per day, reasons for discontinuation, adherence to therapy, and use of vitamin or herbal remedies. Interviewers had no latitude to divert from the interview format. They were required to use preselected response options they matched to patients' answers to each question.

TREATMENT PATTERN CLASSIFICATION

Treatment guidelines and algorithms are meant to assist medical personnel in the determination of optimal therapeutic pathways based on clear evidence of clinical efficacy. Current depression treatment guidelines²⁵ and published treatment algorithms^{26,27} recommend an adequate trial of antidepressant medication as an initial primary treatment for major depressive disorder. An adequate initial trial length for an antidepressant agent is generally considered 6 to 8 weeks^{25,28,29} and includes titration to a therapeutic dose depending on the development of side effects, the patient's age, and the presence of comorbidities.²⁵ If a moderate improvement is not achieved, the guidelines recommend an upward titration or a change in therapy.²⁷ If symptoms remain, subsequent treatment may include concomitant administration of multiple antidepressants, mood stabilizers, or atypical antipsychotics.^{25,26} Once remission is attained, maintenance therapy is recommended for an additional 4 to 9 months.^{25,28,29}

In ARTIST, patients were initially, randomized to 1 of 3 SSRI therapies: fluoxetine (20 mg/d), paroxetine (20 mg/d), or sertraline hydrochloride (50 mg/d). After the initial randomization and enrollment the study was entirely naturalistic and observational, with treatment changes determined by the decisions of physicians and patients. For the purposes of this study, a treatment pattern algorithm was developed, then applied to individual responses in the medication status interview at each assessment to determine initiation of medication therapy, continuation of therapy, discontinuation of therapy, switches in medication, and attempts at dose optimization. Patients were classified as being inadequately treated, adequately treated, or aggressively treated at the 6-month evaluation.

To be classified as aggressively treated, a patient was required to be taking medication for each of the evaluation periods beginning with the 1-month evaluation. Moreover, they were required to have a minimum of one medication switch and one medication increase or more than one switch in therapy or have later stage therapies such as augmentation with a mood stabilizer or atypical antipsychotic agent for at least one evaluation. These treatment patterns are consistent with recommended treatments for more treatment-resistant depressions (TRD).

With TRD, a depressed individual fails to achieve or sustain remission despite receiving appropriate medication(s), at proper dosages, for a suitable length of time.³⁰ Management of individuals considered to be treatment resistant generally follows a "staged approach."^{26,27,30} If a patient does not respond or exhibits a partial response to the initial antidepressant, treatment strategies typically begin with upward titration of the initial antidepressant followed, if necessary, by a switch to another antidepressant. If symptoms remain, subsequent treatment may include concomitant administration of multiple antidepressants, mood stabilizers, or atypical antipsychotics.^{25,26} For this study, the 6-month evaluation was selected, as it would allow sufficient time for a patient to fail 2 trials of different antidepressants.

Patients were classified as adequately treated if they continued taking medication from the first-month assessment through the sixth-month assessment. Six months of therapy are required to meet the minimum practice guidelines for acute and maintenance therapy.²⁵ Patients who were off medication at the

sixth-month assessment or were intermittently on medication throughout the study were classified as inadequately treated.^{25,27}

CLINICAL RESPONSE

To account for imprecision in classification due to measurement error, clinical response was defined as change from the baseline scores on the SCL-20 greater than the reliable change index (RCI).³¹ The RCI limits the misclassification of individual patients due to measurement error, by defining a range in which an individual score may fluctuate due to the imprecision of the measurement tool. The RCI forms a confidence interval around a cutoff point that provides a criterion for improvement or deterioration that is psychometrically sound. The standard error of all patients' baseline scores is calculated based on the standard deviation of all patient baseline scores and the reliability coefficient of the measure. Using this calculation, the standard error of differences is computed. This formula is

$$SED = [2(SE)^2]^{1/2}.$$

The SE for this patient population for the SCL-20 was 4.45, resulting in an SED of 6.29. Using the SED, the significance of an individual's clinical change or RCI can be calculated as

$$RCI = (x_L - x_A) / SED,$$

where x_L is the patient's last score, x_A is the patient's baseline score, and SED is the standard error of the differences. The RCI is computed by simply subtracting the baseline score from the last score and dividing by the SED. If the RCI is greater than 1.96, the patient has reliably improved; if the RCI is less than -1.96, the patient has reliably deteriorated. If the RCI is between 1.96 and -1.96, the patient has not significantly changed. Therefore, an increase or decrease greater than 12.3 was required before a change was considered clinically meaningful.

The SCL-20 scores at the 6-month evaluation were determined to be clinically meaningful based on the RCI calculation. Patients who did not meet the minimal RCI criteria were classified as nonresponders ($n=221$), including those whose SCL-20 baseline scores were low enough at baseline that they could attain a score of 6 or less on the SCL-20 and not meet the minimal RCI criteria ($n=28$). For scores meeting the RCI criteria, remission was defined as a score of 6 or less on the SCL-20. This score was selected because patients with a score of 6 or less generally reported no symptoms of depression on the PRIME-MD. This number was selected prior to the data analysis. Partial response was defined as patients not meeting remission criteria but experiencing a 50% decrease in the SCL-20 from baseline. Patients that did not meet these criteria for remission, or partial response were classified as nonresponders.

STATISTICAL ANALYSIS

Descriptive analyses were undertaken using demographic and baseline characteristics of all participants. One-way analyses of variance (ANOVAs) and χ^2 tests were used to analyze differences (at the 95% significance level) in demographic, treatment history, and characteristics between patients classified as remitters, partial responders, and nonresponders at the 6-month evaluation period. One-way ANOVA was used to evaluate differences between the groups in baseline and 9-month scores of depression severity, quality-of-life, and social functioning measures. The Tukey test for multiple comparisons was used for post hoc determination of differences between groups. Bonferroni corrections were used to adjust for multiple comparisons ($\sigma = .05/[3 \times \text{number of tests with related outcome variables}]$). Related outcome variables included quality-of-life outcomes (SF-36 MCS, PCS, role-emotional subscale, and social functioning subscale) and symptom severity/disability outcomes (SCL-total, energy scale, disability days, depressive symp-

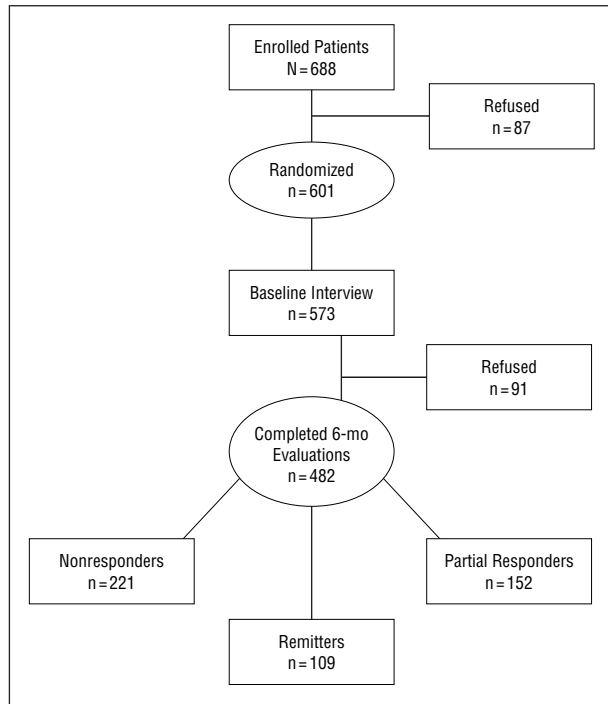


Figure 1. Patient disposition enrollment through 6 months.

tom count, and positive well-being). Multinomial logistic regression models were used to determine significant baseline predictors for remission, partial response, or nonresponse.

RESULTS

PATIENT CHARACTERISTICS

The total sample ($n=573$) was predominantly white (84%) and female (79%).²⁰ Disposition of patients is presented in **Figure 1**. Of the 601 patients who provided informed consent and were randomized into treatment, 573 completed the baseline telephone assessments. The 28 pre-baseline dropouts had similar demographic characteristics as those patients completing the baseline examination.¹⁸

Baseline characteristics between patients who were classified at 6 months as nonresponders, partial responders, and remitters are presented in **Table 1**. The mean age ranged from 43.7 years (for partial responders) to 50.4 years (for nonresponders). Age significantly differed across response groups ($F_2=11.15$, $P=.000$), as patients classified as nonresponders were significantly older than either partial responders or remitters. Statistically significant baseline differences were also found between groups for depressive disorder diagnosis ($\chi^2_8=47.2$, $P=.001$) and suicidal ideation ($\chi^2_4=10.3$, $P\leq.05$). Remitters were less likely to have baseline double depression diagnoses or suicidal ideation. There were no statistically significant differences between groups based on sex, race, past history of depression treatment, alcohol use in the past month, use of health care services, or initial drug randomization.

Classification of patients by treatment and response following 6 months of treatment is shown in **Table 2**. Forty-six percent of all patients completing the 6-month evaluation ($N=482$) were classified as nonresponders ($n=221$), 32% were partial responders ($n=152$), and only 23% met criteria for complete remission ($n=109$).

Table 1. Baseline Characteristics of Patients on 6-Month Response Rate*

Characteristic	Nonresponders (n = 221)	Partial Responders (n = 152)	Remitters (n = 109)
Age, mean (SD), y†	50.4 (15.4)	43.7 (14.9)	43.9 (15.7)
Female	82	77	83
Race			
White	83	88	84
Black	15	9	14
Other	2	3	2
Depressive disorder diagnosis†			
Major depression	31	35	34
Dysthymia	6	3	15
Double depression†	40	54	22
Minor depression	9	5	10
No depression diagnosis	14	3	19
Past history of depression treatment			
Any suicidal ideation past 2 wk†	11	16	5
Any alcohol use in past month	39	47	50
Outpatient visits in past 3 mo	43	39	45
Emergency department visits past 3 mo	14	14	7
Hospitalizations in past 12 mo	20	16	11
Randomization treatment			
Fluoxetine hydrochloride	33	38	31
Paroxetine	34	34	30
Sertraline hydrochloride	34	29	39

*Data are percentage of patients, except for age.

†Significant at $P\leq.05$.

Most patients (77%) received adequate (continuous) or aggressive treatment. Adequate or aggressive therapy did not guarantee remission. A total of 239 patients, or 50% of all patients receiving 6 months of adequate therapy, were classified as either nonresponders or partial responders. Thirteen percent of all patients had aggressive therapy consistent with a TRD ($n=61$).

Additional examinations of the predictors of treatment response and treatment classification (inadequate, adequate, aggressive) were conducted. Patient baseline demographic, clinical severity, and physical functional status had no relationship to treatment classification. There were also no differences between networks in the accrual of patients into the study or outcomes and classification of treatment.

CHANGES IN DEPRESSION SEVERITY OVER TIME BY TREATMENT PATTERNS AND RESPONSE STATUS

Mean change in depression severity by patient treatment and response groups is presented in **Figure 2**. Patients were classified into 1 of 7 groups by treatment patterns and clinical response. Classifications included nonresponse for aggressive therapy, adequate therapy, and inadequate; partial response for aggressive therapy, adequate therapy, and inadequate therapy; and all patients who achieved remission. Between group scores were significantly different at baseline ($F_6=8.8$, $P<.001$), 3 months

($F_6=26.2, P<.001$), 6 months ($F_6=100.7, P<.001$), and 9 months ($F_6=38.5, P<.001$). The greatest response to therapy was early in treatment with the largest changes in mean scores by the 3-month evaluation. As demonstrated in Figure 2, groups clustered more by response than by treatment. At the 6-month evaluation, remitters were not only significantly better than nonresponders in depression severity, but also significantly better than partial responders regardless of their treatment classification.

BASELINE DEPRESSION SEVERITY AND PSYCHOLOGICAL FUNCTIONING

Baseline depression and other psychological measures are presented in **Table 3**. Depression severity, health-related quality of life, psychological well-being, depressive symptom counts, and self-reported disability days are compared between groups of patients defined by response. Statistically significant differences were found between all groups except for self-reported disability days. Overall, remitters were the least impaired at baseline while partial responders had the least favorable health scores on each of the baseline measures (except the SF-36 PCS), indicating that these groups of patients not only have the most severe depression symptoms but were the most functionally impaired group overall.

DEPRESSION SEVERITY AND PSYCHOLOGICAL FUNCTIONING AT THE 9-MONTH EVALUATION

Nine-month depression and other psychological measures are presented in **Table 4**. Response groups were compared on depression severity, psychological functioning, and disability at the end of the study evaluation. All measures were statistically significant between groups at this time, including self-reported disability days. Patients who remitted at 6 months continued to have the best average scores on each of the depression and psychological measures. Nonresponders have average scores on each measure indicating the poorer health states. Post hoc analyses indicated that patients achieving remission were not only statistically better than nonresponders in all measures but remitters were also statistically better than partial responders in all measures except for self-reported disability days. These findings support that the overall treatment goal for depression should be complete remission of symptoms.

MULTINOMIAL LOGISTIC MODEL OF BASELINE PREDICTORS OF RESPONSE

The results of the multinomial logistic regression model, estimating the probability of response based on baseline characteristics and 6-month treatment patterns are presented in **Table 5**. Our initial logistic model considered using all demographic variables as predictors for 3 levels of clinical response: remission, partial response, and non-response. Significant demographic predictors of response included patient age and initial diagnostic category (major depressive disorder, dysthymia, minor depression, double depression). Next, we augmented the patient characteristics model with baseline quality-of-life variables (SF-36 MCS, PCS, role-emotional subscale, social func-

Table 2. Classification of Patients by Treatment and Response at 6-Month Evaluation*

Treatment Classification	No. (%) of Patients			
	Overall (n = 482)	Nonresponders (n = 221)	Partial Responders (n = 152)	Remitters (n = 109)
Aggressive therapy	61 (13)	32 (14)	22 (14)	7 (6)
Adequate therapy	311 (64)	137 (62)	102 (67)	72 (66)
Inadequate therapy	110 (23)	52 (24)	28 (18)	30 (28)

*Overall χ^2 treatment by response was not significant.

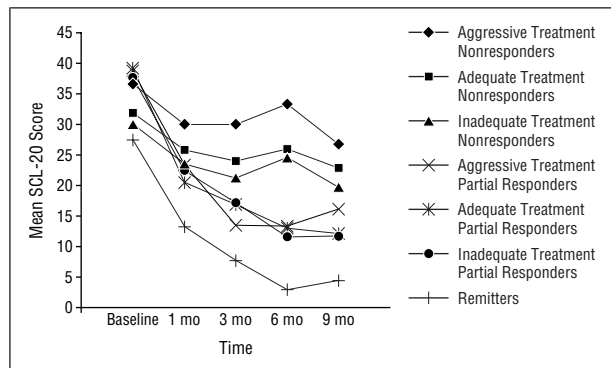


Figure 2. Mean change in Symptom Checklist–20 (SCL-20) score from baseline to end of the study for patients defined by treatment and response.

tioning scale), baseline depression severity/disability variables (SCL-20, energy scale, MOS positive well-being subscale, depressive symptom count, and disability days).

Our final model included all predictors that were found to be statistically significant in the initial models ($P<.10$) (patient age, diagnosis, the SF-36 PCS, and the SCL-20 energy subscale) plus the treatment classification (inadequate therapy, adequate therapy, aggressive therapy) factor that we wished to study. Patients classified as nonresponders tended to be older, report lower energy levels, and have worse health status on the SF-36 PCS. The diagnosis of dysthymia (and also no diagnosis) predicted that the response to therapy would not be partial. We found no evidence that treatment classification significantly influenced response status. While the overall χ^2 statistic for our final model was highly significant ($\chi^2_{18}=128.71, P<.001$), the corresponding R^2 statistic of 0.126 is somewhat low, indicating that much of the observed variation in response is not captured by the variables included in our model.

COMMENT

Overall response to pharmacotherapy in this group of primary care patients was less than optimal, with 46% of patients being classified as nonresponders, and 32% more classified as partial responders. The 23% remission rate for this study is slightly lower than remission rates reported in clinical trial data. Reported remission rates for SSRI therapy in standard randomized clinical trials range from 23%³² to 45%,³³ with pooled rates in a meta-analysis reported at 35%.³⁴ Although the average sever-

Table 3. Baseline Depression and Other Psychological Measures by Response Classification

Measure	Score*	F	Mean (SD) Score		
			Nonresponders (n = 221)	Partial Responders (n = 152)	Remitters (n = 109)
SF-36 mental component	0-100	18.5†	33.0 (11.9)	26.6 (9.8)	34.0 (12.3)
SF-36 role-emotional	0-100	15.2†	57.7 (24.5)	47.8 (20.4)	62.5 (21.8)
SF-36 social functioning	0-100	14.3†	57.7 (28.4)	49.6 (25.8)	67.8 (26.1)
SF-36 physical component	0-100	26.2†	47.1 (10.1)	51.1 (9.5)	55.0 (8.6)
SCL-20 symptom severity	0-100	24.6†	31.9 (14.5)	38.8 (10.9)	27.4 (14.1)
SCL-20 energy subscale	0-25	28.9†	14.3 (4.6)	16.2 (3.8)	12.6 (4.3)
MOS positive well-being	0-100	14.4†	42.3 (20.0)	34.8 (16.0)	47.2 (20.5)
Depressive symptom count	0-9	21.1†	5.7 (2.2)	6.6 (1.7)	4.9 (2.3)
Disability days	0-28	1.9	1.39 (2.2)	1.6 (2.7)	1.0 (2.0)

Abbreviations: MOS, Medical Outcomes Study scale; SCL-20, Symptom Checklist-20; SF-36, 36-Item Short-Form Health Survey.

*Range of possible scores. Shown in boldface is the best score (ie, best health).

†Significant at $P \leq .001$.

Table 4. Nine-Month Depression and Other Psychological Measures by Response Classification

Measure	Score*	F	Mean (SD) Score		
			Nonresponders (n = 201)	Partial Responders (n = 136)	Remitters (n = 99)
SF-36 mental component	0-100	58.1†	43.8 (11.3)	50.41 (7.8)	55.7 (6.2)
SF-36 role-emotional	0-100	42.4†	74.0 (24.2)	85.5 (16.4)	95.5 (11.4)
SF-36 social functioning	0-100	42.1†	71.3 (27.4)	85.5 (19.6)	95.3 (11.8)
SF-36 physical component	0-100	32.0†	44.5 (10.2)	49.5 (0.3)	53.4 (7.8)
SCL-20 symptom severity	0-100	112.1†	22.7 (13.1)	12.4 (8.3)	4.5 (4.9)
SCL-20 energy subscale	0-25	79.8†	11.5 (4.4)	8.9 (3.0)	6.1 (1.6)
MOS positive well-being	0-100	70.2†	55.0 (20.8)	66.3 (16.4)	80.9 (13.1)
Depressive symptom count	0-9	77.1†	4.2 (2.6)	2.5 (2.1)	0.9 (1.2)
Disability days	0-28	7.0†	1.2 (2.5)	0.6 (1.7)	0.3 (1.5)

Abbreviations: MOS, Medical Outcomes Study scale; SCL, Symptom Checklist-20; SF-36, 36-Item Short-Form Health Survey.

*Range of possible scores. Shown in boldface is the best score (ie, best health).

†Significant at $P \leq .001$.

Table 5. Multinomial Regression of Baseline Predictors of Response Classification (SCL-20)

Variable	Wald χ^2	P Value
Patient age	9.00	.01
Diagnostic category	17.50	.02
SF-36 PCS baseline	29.77	<.001
Medication classification	5.37	.25

Abbreviations: SCL-20, Symptom Checklist-20; SF-36 PCS, 36-Item Short-Form Health Survey physical component summary score.

ity of depressive symptoms in this study decreases with treatment,¹⁸ most patients do not achieve complete remission. Research indicates that primary care patients have less than optimal response to therapy because they often receive lower doses of medication, do not receive adequate follow-up, have less than adequate monitoring of symptoms, and prematurely discontinue therapy.⁶ All patients in this study were randomized to medication doses that are typically considered therapeutic. We found that 64% of all patients (n=311) were classified as adequately treated with an adequate dose of medication for an adequate duration; however, only 23% (n=72) achieved remission. Guideline-based care suggests that most of these patients required more aggressive treatment than

they obtained to achieve better outcomes. This highlights the need to redefine adequacy of treatment within the context of measured clinical improvement. If a patient does not improve clinically and no attempts have been made to optimize treatment, they should not be considered adequately treated.^{25,28,29}

However, the treatment classification guidelines used in this study were more conservative than those usually found in standard clinical practice. In this study, aggressive treatment required patients to continuously receive medication for 6 months and have either one switch and one titration, more than one titration or switch, or augmentation therapy. To our knowledge, more than one titration and switch is infrequent in usual primary care. Patients were classified as receiving adequate care only if they reported 6 months of continuous medication therapy; this is in contrast to 90 days or less commonly reported in primary care practice.^{15,16} Therefore, despite using more rigorous treatment strategies than generally reported in primary care studies, there continues to be less than optimal remission and partial response rates.

These findings suggest that it may be beneficial to continue to develop and disseminate educational programs and organizational systems intended for primary care to help PCPs recognize depression with more certainty, to evaluate patients' response to treatment over

time, and to adhere to published treatment guidelines in response to observed treatment responses.^{5,35} Providing methods for monitoring patient outcomes (eg, automated systems to collect patient reported symptom severity) may provide an efficient mechanism for PCPs to receive needed outcomes data in order to provide quality of care without increasing the length of a patient visit.

This study collected a limited amount of information on medical and psychiatric comorbidities and substance abuse. Screening and diagnosis is important for adequate treatment of depression. Individuals with personality disorders (eg, obsessive-compulsive, dependent, narcissistic, borderline disorders) are more prone to depressive disorders and exhibit less than satisfactory response to antidepressant medication treatment.²⁵ Additionally, depression is frequently comorbid with substance abuse and anxiety disorders. These conditions must be carefully screened for as these patients generally have greater impairment and are more clinically challenging. Published treatment guidelines provide additional information to aid in treatment.²⁵

This study also found that 13% of patients reported receiving aggressive treatment strategies generally necessary for addressing treatment resistance. Thus, there appears to be substantial numbers of patients with TRD being treated in primary care settings. This finding is not surprising given that between 10% and 20% of all patients with major depressive disorder are treatment resistant³⁶ and that the majority of patients with major depressive disorder are treated in primary care settings. These and Rush²⁶ have described subcategories of treatment resistance. Relative resistance is considered failure to respond to an average dose of an antidepressant for a minimally acceptable period of time, while absolute treatment resistance refers to patients who fail to respond to a maximal dose of an antidepressant over an extended treatment period. In this study, 7 patients who received more aggressive therapy ultimately were classified as responders and might be considered to have relative treatment resistance. These patients also are an indication that patients with TRD can be successfully treated in primary care. Twenty-three patients who received aggressive therapy and did not have a remission of symptoms may be considered exemplars of absolute treatment resistance and could potentially benefit from development of new therapies.

The most common response criteria in depression trials are prespecified levels of improvement on a depression symptom rating scale.²⁶ Many studies and clinical trials classify patients achieving a 50% decrease in depression severity scores (eg, Hamilton Depression Rating Scale) as having a positive clinical response. In general, decreases in depression severity are correlated with improvement in patient functioning and health-related quality of life.^{6,18} In this study we differentiated between partial responders and remitters to allow comparisons between depression severity, health-related quality of life, and disability for these groups.

In general, a 50% reduction of depression severity corresponds to a moderate improvement in the clinician's global clinical impression. Often, patients meeting these criteria continue to have considerable residual

symptoms.²⁶ Residual symptoms have been associated with less than optimal levels of functioning and higher levels of disability.²⁶ Our study supported these findings in that patients who were in remission had significantly better symptoms and levels of functioning than patients who had experienced a 50% symptom improvement and were classified as partial responders. This difference highlights the need to strive for symptom remission as the desired outcome for acute treatment.²⁶

These results need to be viewed within the context of several limitations. This was a naturalistic, open-label study designed to compare outcomes associated with SSRI therapy (fluoxetine, paroxetine, or sertraline) in usual primary care practices. It is possible that there were patient or clinician biases that may have influenced treatment strategies and could have resulted in less than optimal outcomes. Patients were not randomly assigned to receive continuous therapy or more aggressive therapies associated with TRD (titrations, switches, or augmentations) and physicians were not given algorithms for treating resistant depression. It is possible that overall outcomes would have been improved if there were designated strategies for dealing with less than adequate clinical response.

A key criterion for patient selection was the physician's belief that the patient was appropriate for SSRI therapy. After this criterion was met, only patients willing to take an SSRI were eligible for this study. Therefore, the results do not generalize to patients who the physician did not believe warranted SSRI therapy, who refused any treatment, who refused SSRIs, or who preferred psychotherapy.

Additionally, the outcomes evaluations were obtained independent of the clinical care of the patients. These evaluations were not shared with the physicians. Patients may have continued antidepressant therapy without scheduled physician visits, or minimized symptoms at the time of regularly scheduled visits. Physicians may not have fully evaluated symptoms due to practice or educational limitations. If information about clinical response had been made available to the clinicians,³³ they may have modified the treatment of patients with suboptimal response. However, these changes would have converted this study from one focused on usual care, to a more tightly controlled clinical efficacy trial.

For this study, medication status was determined by patient self-report. While there is the potential for a discrepancy between patient self-reported medication history and adherence to therapy, the self-report module was very thorough in eliciting information about continuation, dose of medication, number of times per day medication was taken, and changes in medication. Medication information was collected by telephone interview, by a trained interviewer not connected with the treating physician, to minimize reporting bias due to social acceptability.

The design of observational studies provides an evaluation of treatment in "real-world" settings and exhibit superior external generalizability; however, there is reduced confidence in the reliability of findings due to the lack of rigorous controls found in randomized clinical trials. Finally, since this was an observational study

after initial randomization, treatment pattern was confounded with clinical status. Patients and clinicians determined whether to terminate medication prematurely, maintain doses despite inadequate response, or collaborate for aggressive medication. Patients and clinicians were also free to seek mental health specialty care. These findings did not assess the impact of variations in mental health intensity or type of outcome.

Numerous published studies describe deficiencies in primary care antidepressant treatment, yet few studies have applied the same rigor in the determination of outcomes of psychiatric care or compared outcomes between primary care and specialty care. One study comparing the two groups found that there were only modest differences between outcomes for psychiatrists and PCPs.⁴ Future naturalistic studies investigating depression outcomes would benefit from adding specialty mental health sites for comparison; however, such a study must assess the impact of refractory disease, comorbidity, and referral bias to be credible.

The results of this study demonstrate the complex interactions between response and treatment patterns for patients with depressive disorders. It also highlights the documented differences between expert recommendations for depression treatment and actual treatment in primary care practices. One important recommendation is that outcomes for depressed primary care patients have the potential for providing more aggressive treatment based on patient-reported outcomes. Efforts should be made to increase the intensity of treatment for patients in whom initial treatment fails but also for patients who only achieve a partial response.

Accepted for publication May 29, 2003.

This project was supported by Eli Lilly & Company.

Portions of this article were presented at the 155th Annual Meeting of the American Psychiatric Association, May 18-23, 2002, Philadelphia, Pa; and the Seventh Annual Meeting of the International Society for Pharmacoeconomic and Outcomes Research, May 19-23, 2002, Arlington, Va.

We would like to thank Donald P. Hay, MD, for review and editorial comments on the manuscript and Robert L. Obenchain, PhD, for statistical guidance and review.

Corresponding author and reprints: Ralph Swindle, PhD, Eli Lilly & Company, Lilly Corporate Center, Indianapolis, IN 46285 (e-mail: swindle@lilly.com).

REFERENCES

- Mason J, Freemantle N, Eccles M. Fatal toxicity associated with antidepressant use in primary care. *Br J Gen Pract.* 2000;50:366-370.
- Nelson JC. Safety and tolerability of the new antidepressants. *J Clin Psychiatry.* 1997;58(suppl 6):26-31.
- Coyne JC, Schwenk TL, Fechner-Bates S. Nondetection of depression by primary care physicians reconsidered. *Gen Hosp Psychiatry.* 1995;17:3-12.
- Simon GE, Von Korff M, Rutter CM, Peterson DA. Treatment process and outcomes for managed care patients receiving new antidepressant prescriptions from psychiatrists and primary care physicians. *Arch Gen Psychiatry.* 2001;58:395-401.
- Katon W, Robinson P, Von Korff M, et al. A multifaceted intervention to improve treatment of depression in primary care. *Arch Gen Psychiatry.* 1996;53:924-932.
- Revicki DA, Simon GE, Chan K, Katon W, Heiligenstein J. Depression, health-related quality of life, and medical cost outcomes of receiving recommended levels of antidepressant treatment. *J Fam Pract.* 1998;47:446-452.
- Arean PA, Alvidrez J. Treating depressive disorders: who responds, who does not respond, and who do we need to study? *J Fam Pract.* 2001;50:E2.
- Simon GE, Lin EH, Katon W, et al. Outcomes of "inadequate" antidepressant treatment. *J Gen Intern Med.* 1995;10:663-670.
- Souery D, Amsterdam J, de Montigny C, et al. Treatment resistant depression: methodological overview and operational criteria. *Eur Neuropsychopharmacol.* 1999;9:83-91.
- Lin EH, Von Korff M, Katon W, et al. The role of the primary care physician in patients' adherence to antidepressant therapy. *Med Care.* 1995;33:67-74.
- Janicak PG, Martis B. Strategies for treatment-resistant depression. *Clin Cornerstone.* 1999;1:58-71.
- Warneke L. Management of resistant depression. *Can Fam Physician.* 1996;42:1973-1980.
- McQuaid JR, Stein MB, Laffaye C, McCahill ME. Depression in a primary care clinic: the prevalence and impact of an unrecognized disorder. *J Affect Disord.* 1999;55:1-10.
- Cass AR, Volk RJ, Nease DE Jr. Health-related quality of life in primary care patients with recognized and unrecognized mood and anxiety disorders. *Int J Psychiatry Med.* 1999;29:293-309.
- Crown WH, Treglia M, Meneades L, White A. Long-term costs of treatment for depression: impact of drug selection and guideline adherence. *Value Health.* 2001;4:295-307.
- Katon W, Von Korff M, Lin E, et al. Collaborative management to achieve treatment guidelines: impact on depression in primary care. *JAMA.* 1995;273:1026-1031.
- Sturm R, Wells KB. How can care for depression become more cost-effective? *JAMA.* 1995;273:51-58.
- Kroenke K, West SL, Swindle R, et al. Similar effectiveness of paroxetine, fluoxetine, and sertraline in primary care: a randomized trial. *JAMA.* 2001;286:2947-2955.
- Mulrow CD, Williams JW Jr, Gerety MB, Ramirez G, Montiel OM, Kerber C. Case-finding instruments for depression in primary care settings. *Ann Intern Med.* 1995;122:913-921.
- Spitzer RL, Williams JB, Kroenke K, et al. Utility of a new procedure for diagnosing mental disorders in primary care: the PRIME-MD 1000 Study. *JAMA.* 1994;272:1749-1756.
- Ware JE, Kosinski M, Keller SD. *SF-36 Physical and Mental Health Summary Scales: A User's Manual.* Boston, Mass: Health Institute, New England Medical Center; 1994.
- Ware JE. *SF-36 Health Survey: Manual and Interpretation Guide.* Boston, Mass: Health Institute, New England Medical Center; 1993.
- Stewart AL, Ware JE. *Measuring Functioning and Well-being: The Medical Outcomes Study Approach.* Durham, NC: Duke University Press; 1992.
- Swindle R, Kroenke K, Braun L. Energy and improved workplace productivity in depression. In: Farquhar I, Summers K, Sorkin A, eds. *Investing in Health: The Social and Economic Benefits of Health Care Innovation.* Vol 14. Amsterdam, the Netherlands: JAI-Elsevier Science; 2001:323-341.
- Practice guideline for the treatment of patients with major depressive disorder (revision). *Am J Psychiatry.* 2000;157(4 suppl):1-45.
- Thase ME, Rush AJ. Treatment-resistant depression. In: Bloom FE, Kupfer DJ, eds. *Psychopharmacology: The Fourth Generation of Progress.* New York, NY: Raven Press Ltd; 1995:1081-1097.
- O'Reardon JP, Amsterdam JD. Treatment-resistant depression: progress and limitations. *Psychiatr Ann.* 1998;28:633-640.
- Depression Guidelines Panel. *Depression in Primary Care: Detection and Diagnosis.* Vol 1. Rockville, Md: US Dept of Health and Human Services, Public Health Service, Agency for Health Care Policy and Research; 1993.
- Depression Guidelines Panel. *Depression in Primary Care: Detection and Diagnosis.* Vol 2. Rockville, Md: US Dept of Health and Human Services, Public Health Service, Agency for Health Care Policy and Research; 1993.
- Rush AJ, Crismon ML, Toprac MG, et al. Consensus guidelines in the treatment of major depressive disorder. *J Clin Psychiatry.* 1998;59(suppl 20):73-84.
- Jacobson NS, Truax P. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *J Consult Clin Psychol.* 1991;59:12-19.
- Rudolph RL, Feiger AD. A double-blind, randomized, placebo-controlled trial of once-daily venlafaxine extended release (XR) and fluoxetine for the treatment of depression. *J Affect Disord.* 1999;56:171-181.
- Dierick M, Ravizza L, Realini R, et al. A double-blind comparison of venlafaxine and fluoxetine for treatment of major depression in outpatients. *Prog Neuropsychopharmacol Biol Psychiatry.* 1996;20:57-71.
- Thase ML, Entsuah AR, Rudolph RL. Remission rates during treatment with venlafaxine or selective serotonin reuptake inhibitors. *Br J Psychiatry.* 2001;178:234-241.
- Wells KB, Sherbourne C, Schoenbaum M, et al. Impact of disseminating quality improvement programs for depression in managed primary care: a randomized controlled trial. *JAMA.* 2000;283:212-220.
- Dyck MJ. Treatment-resistant depression: a critique of current approaches. *Aust N Z J Psychiatry.* 1994;28:34-41.