Duloxetine in Practice-Based Clinical Settings: Assessing Effects on the Emotional and Physical Symptoms of Depression in an Open-Label, Multicenter Study

Madelaine M. Wohlreich, M.D.; Curtis G. Wiltse, Ph.D.; Durisala Desaiah, Ph.D.; Wenyu Ye, Ph.D.; Rebecca L. Robinson, M.S.; Kurt Kroenke, M.D.; Susan G. Kornstein, M.D.; and John H. Greist, M.D.

Objective: In placebo-controlled clinical trials, duloxetine has been shown to be effective and well-tolerated in patients with major depressive disorder (MDD). However, patients in registration trials may not be representative of patients in clinical practice. This study sought to assess the effectiveness, safety, and tolerability of duloxetine in diverse populations of outpatients with MDD.

Method: This open-label study recruited outpatients ≥ 18 years of age with DSM-IV MDD in primary care or psychiatric practice settings and treated them with duloxetine 60 mg q.d. for 7 weeks. Primary outcome measures were (1) the physician-rated Clinical Global Impressions-Severity of Illness scale, (2) the patient-rated 28-item Somatic Symptom Inventory (SSI-28) average, and (3) the patient-rated 16-item Quick Inventory of Depressive Symptomatology-Self Report. Quality of life, disability, and vital signs also were assessed. The first patient visit was August 16, 2004. The last patient visit was January 7, 2005.

Results: Of 3543 outpatients enrolled, 3431 received at least 1 dose of duloxetine, of whom 71.4% completed the study. Most patients were Caucasian (90.8%) and female (75.4%); mean age was 48 years. Duloxetine significantly ($p < .001$) improved all efficacy measures in all treated patients as well as in subgroups based on gender, ethnic origin, age, and patient care setting. Except for the SSI-28 average, all the efficacy measures were in favor of female gender and primary care subgroups. Overall, 10.8% of patients discontinued due to adverse events.

Conclusion: Duloxetine 60 mg q.d. was effective, regardless of gender, ethnic origin, age, and patient care settings, in this 7-week open-label study and was well-tolerated in a diverse population of outpatients with MDD.

Clinical Trials Registration: ClinicalTrials.gov identifier NCT00479726.

(Prim Care Companion J Clin Psychiatry 2007;9:271–279)

Received Dec. 18, 2006; accepted Feb. 28, 2007. *From Lilly Research Laboratories, Eli Lilly and Co., Indianapolis, Ind. (Drs. Wohlreich, Wiltse, Desaiah, and Ye and Ms. Robinson); the Indiana University School of Medicine and Regenstrief Institute, Indianapolis, Ind. (Dr. Kroenke); Virginia Commonwealth University, Richmond, Va. (Dr. Kornstein); and Healthcare Technology Systems, Inc., Madison, Wis. (Dr. Greist).*

This work was sponsored by Eli Lilly and Co., Indianapolis, Ind. The authors accept full responsibility for the conduct of this study, had full access to all data from this study, and participated in the decision to publish the data.

The authors thank all the investigators and patients for their role in the conduct of this study and Douglas E. Faries, Ph.D., for his support of propensity scoring analyses. Dr. Faries is an employee and stock shareholder of Eli Lilly and Co.

Financial disclosure appears at the end of this article.

Corresponding author and reprints: Madelaine M. Wohlreich, M.D., Lilly Research Laboratories; Eli Lilly and Co., DC 4103, Indianapolis, IN 46285 (e-mail: mwmd@lilly.com).

Patients who participate in clinical trials may not be representative of patients in clinical practice. Practice patients often differ from those in controlled clinical trials in terms of demographics and clinical characteristics.[1] Hence, data from controlled clinical trials may not be highly generalizable to the heterogeneous population of patients that may be expected in practice.

Duloxetine hydrochloride is a dual reuptake inhibitor of serotonin and norepinephrine.[2] In placebo-controlled clinical trials, duloxetine has been shown to be effective, safe, and well-tolerated in the treatment of major depressive disorder (MDD).[3–6] A prior practice-based, open-label study of duloxetine in outpatients with depression was too small to allow comparisons in subpopulations based on ethnic origin or other clinical characteristics that might differentiate clinical trial from practice-based populations.[7]

Presenting symptoms and therapeutic response may vary with a patient's ethnic background.[8–10] While most controlled clinical trials tend to report data on Caucasians, there is increasing interest in other ethnic subgroups in psychopharmacology research. A large prospective study for the treatment of depression, Sequenced Treatment Alternatives to Relieve Depression (STAR*D), was conducted in both psychiatric and primary care settings.[11] The

results of this trial found that non-Caucasian race, male gender, unemployment, lower income, less education, poorer functional status, and lower quality of life at baseline were overlapping and independently associated with lower remission rates.[11] Remission status by age or primary care setting did not significantly differ in STAR*D patients.[11] Other studies conducted with selective serotonin reuptake inhibitors (SSRIs) in primary care settings showed that older age was associated with a poor response.[12,13] Escitalopram treatment in an open-label study showed comparable response rates in broadly representative diverse populations of outpatients with depression.[14]

In order to study the effectiveness of duloxetine in diverse outpatient populations with distinct characteristics such as gender, ethnic origin, age, and patient care setting, a large number of patients with major depression were recruited from "real-world" outpatient primary care and psychiatric practice settings. This study of a diverse group of outpatients may provide data about duloxetine in the treatment of emotional and physical symptoms with a degree of generalizability not previously possible for practice-based patients with depression.

METHOD

Study Design

This phase IV multicenter, open-label study was conducted at primary care and psychiatric clinical practices in the United States and Puerto Rico. The study was conducted in accordance with the ethical principles of the Declaration of Helsinki. Data collection at the clinical site was modest in order to allow investigators to recruit and enroll subjects from their practices with as little disturbance to the usual course of outpatient treatment as possible. Qualified patients were assigned to open-label duloxetine 60 mg q.d. for 7 weeks of treatment. Study visits were at baseline and at 2 and 7 weeks. Investigators were instructed to start duloxetine at a dose of 60 mg q.d.; the dose could subsequently be lowered to 30 mg q.d. if needed for tolerability reasons during the first week of treatment. At the investigator's discretion, the dose could be initiated at 30 mg q.d. for up to 7 days. However, following a maximum of 7 days at the lower 30 mg dose, all patients received duloxetine 60 mg q.d. for the remainder of the study.

Sample size calculations were based on data for ethnic distribution and change in CGI-Severity of Illness (CGI-S) scale and the 28-item Somatic Symptom Inventory (SSI-28) average from 6 large, previously published U.S.-based duloxetine studies.[3–6,15] The study was powered such that, if 1% of the 8000 anticipated patients in this study (80 patients) were in the smallest subgroup, then the study would have 90% power to detect mean changes from baseline of 0.54 and 0.21 in CGI-S and SSI-28 average score, respectively, in that ethnic subgroup of this study, assuming standard deviations of 1.46 and 0.57, respectively. Investigators were chosen by geographic location and by their self-reported ability to recruit ethnic subgroups of patients to fulfill the study objectives. The goal of recruiting 8000 patients was not met, however, due to difficulty in obtaining a sufficient number of investigators.

Patients

Study participants were outpatients, at least 18 years of age, who, in the opinion of the investigator, met the *Diagnostic and Statistical Manual of Mental Disorders,* Fourth Edition (DSM-IV)[16] criteria for MDD. Since the focus of the trial was generalizability to outpatient practice, the inclusion and exclusion criteria were not as stringent as they would have been in controlled clinical trials. Female participants had to have a negative urine test for pregnancy. All patients had to be able to communicate with the investigators and complete all self-rated scales. Exclusionary criteria included investigator site personnel and their immediate family members, employees of the sponsor, treatment within the 30 days prior to enrollment in any other investigational drug trials, serious medical or psychiatric illness requiring hospitalization, acute or serious liver disease, current substance abuse or dependence, and treatment with a monoamine oxidase inhibitor within 14 days prior to visit 1. Patients were required to provide written informed consent prior to participation in the study. The first patient visit was August 16, 2004, and the last patient visit was January 7, 2005.

Primary Efficacy Measures

The primary objective of this trial was to assess the effectiveness of duloxetine 60 mg q.d. in diverse populations of outpatients with MDD in practice-based clinical settings as measured by the CGI-S scale,[17] SSI-28 average,[18] and the 16-item Quick Inventory of Depressive Symptomatology-Self Report (QIDS-SR-16) total score.[19] The CGI-S was administered by the investigator at the time of assessment and recorded the severity of illness from 1 (normal, not at all depressed) to 7 (most extremely depressed). The SSI-28 scale assessed the severity of 2 pain symptoms (joint pain and neck pain) in addition to the items contained in the original 26-item scale. Patients recorded each symptom on a rating scale of 1 (not at all) to 5 (a great deal). The QIDS-SR-16 and SSI-28 scales were self-reported by patients. The QIDS-SR-16 scale records severity and change in depressive symptoms on a scale of 0 to 3, with higher scores denoting greater symptom severity.

Secondary Efficacy Measures

The Clinical Global Impressions-Severity of Physical Symptoms (CGI-S-PS) scale[17] is a measure of the severity of overall painful symptoms and was recorded by the physician. The scores range from 1 (none) to 7 (most

extreme). The Patient Global Impressions-Improvement (PGI-I) scales[17] are patient-rated instruments that measure both physical (PGI-IP) and emotional (PGI-IE) symptoms improvement on a scale of 1 (very much better) to 7 (very much worse), with a score of 4 = no change. The Mood And Physical symptoms in Depression (MAP-D) scale is a novel, patient-rated instrument that contains 8 questions to assess specific aspects of depression and painful physical symptoms on a scale of 0 to 10. The MAP-D validation results will be the subject of a future publication. The Quality of Life Enjoyment and Satisfaction Questionnaire-Short Form (Q-LES-Q-SF)[20] is a patient-rated instrument that measures the degree of enjoyment and satisfaction using 16 items on a scale of 1 (very poor) to 5 (very good). The Sheehan Disability Scale (SDS)[21] is a patient-rated instrument; the total score measures the disruption caused by the patient's symptoms on work/school, social life, and family life/home responsibilities. The extent of disruption in each of the areas is rated from 0 (not at all) to 10 (extremely).

The PGI-I, MAP-D, and SDS scales were administered to patients outside scheduled study site visits via an Interactive Voice Response System using a telephone. Other information was obtained during scheduled visits at the study sites.

Safety Measures

Safety was assessed through collection of adverse events during the study period, without regard to the possibility of causal relationships. Study site personnel also were required to report any serious adverse events as well as all discontinuations due to adverse events. Safety measures also included the assessment of blood pressure and heart rate at baseline and end point.

Statistical Analyses

Demographics were assessed for all patients; efficacy and safety analyses were conducted on data from all patients receiving at least 1 dose of study drug. Baseline was defined as the nonmissing visit 1 observation; end point was defined as the last nonmissing value after visit 1 (visit 2 or 3). The change from baseline to end point for the primary, and most of the secondary, end points was analyzed using a paired t test. PGI-I was tested using a 1-sample t test for the difference from 4 (no change). For within-group changes, t tests also were computed for subgroups based on gender, ethnic origin, age, and patient care setting.

Mean changes for the 3 primary efficacy variables were compared between ethnic origin subgroups with at least 80 patients (Caucasians, African descent, Hispanics), although, because of small sample sizes, the African descent and Hispanic subgroups were not compared. Mean changes in efficacy variables also were compared between subgroups based on gender (female vs. male), ethnic origin (Caucasian vs. others), age (< 65 vs. ≥ 65 years), and patient care setting (primary care vs. psychiatric). The subgroups were not randomized, and comparisons between such nonrandomized groups often are imbalanced in baseline patient characteristics, so that statistical adjustment is required.[22] To control for potential bias in comparing groups, the doubly robust inverse propensity score weighting method (DR)[23] was used to obtain an estimated predicted group difference. The inverse weighting based on propensity scoring reduces the imbalance between groups in the observed covariates. The DR method was selected, in particular, because use of the regression-augmented weights makes it relatively robust to model misspecification. As sensitivity analyses, propensity score–adjusted regression and stratified propensity score analyses were also conducted. The results were similar to those from the DR method; thus, only the DR results are reported.

The covariates included in the propensity score model were gender, ethnic origin, age, patient care setting, baseline score for the outcome measure, and use of alcohol and each of 13 classes of medications at enrollment. For specific subgroup analyses (such as comparisons of age groups) the corresponding subgroup variable was considered the dependent variable and removed from the list of independent variables. Missing data for covariates was handled by utilizing nonmissing covariates plus the missing data pattern to compute the propensity score.[22]

All statistical tests were conducted at a 2-sided α level of .05. Patients with missing data were excluded from the denominator in the calculation of percentages. No corrections for multiple outcome measures were made. All analyses were conducted using SAS/STAT software, version 8.2 of the SAS system for the PC (SAS Institute, Inc., Cary, N.C.).

RESULTS

Study Population and Baseline Characteristics

The baseline demographic and clinical characteristics of all 3543 enrolled patients are presented in Table 1. A total of 836 investigators participated in this open-label study and treated patients in primary care (506 centers with 2110 patients) and psychiatric (330 centers with 1433 patients) settings. The mean patient age was 48 years (8.7% $\geq$ 65 years), and most patients were Caucasian (90.8%) and female (75.4%).

Treatment and Concomitant Medications

Of the 3543 enrolled patients, 3431 patients received at least 1 dose of duloxetine, of whom 2419 (71.4%, calculated by omitting 41 patients whose reasons for discontinuation were unknown) completed the study. Most patients (71.7%) started on a dose of 60 mg q.d., while

Table 1. Baseline Demographics and Clinical Characteristics of All Enrolled Patients (N = 3543)

Variable	N	Duloxetine
Age, mean (SD), y	3524	48.0 (12.9)
Sex, N (%)	3533	
Female		2665 (75.4)
Male		868 (24.6)
Ethnic origin, N (%)	3521	
Caucasian		3197 (90.8)
Hispanic		141 (4.0)
African descent		131 (3.7)
East/Southeast Asian		20 (0.6)
West Asian		9 (0.3)
Other		23 (0.7)
Reached menopause, N (%)	2113	
Yes		899 (42.5)
No		1214 (57.5)
Education, mean (SD), y	2817	13.7 (2.6)
Patient care setting, investigators, N (%)	836	
Psychiatric		330 (39.5)
Primary care		506 (60.5)
Patient care setting, patients, N (%)	3543	
Psychiatric		1433 (40.4)
Primary care		2110 (59.6)
CGI-S, mean (SD)	3489	4.29 (0.91)
SSI-28 average, mean (SD)	3194	2.15 (0.73)
QIDS-SR-16 total, mean (SD)	3324	13.97 (5.04)
CGI-S-PS, mean (SD)	3483	3.94 (1.30)
MAP-D total, mean (SD)	2798	46.10 (16.35)
SSI-26 average, mean (SD)	3198	2.10 (0.71)
Q-LES-Q-SF total, mean (SD)	3487	43.78 (16.10)
SDS total, mean (SD)	2795	17.54 (7.46)

Abbreviations: CGI-S = Clinical Global Impressions-Severity of Illness, CGI-S-PS = Clinical Global Impressions-Severity of Physical Symptoms, MAP-D = Mood And Physical symptoms in Depression, QIDS-SR-16 = 16-item Quick Inventory of Depressive Symptomatology-Self Report, Q-LES-Q-SF = Quality of Life Enjoyment and Satisfaction Questionnaire-Short Form, SDS = Sheehan Disability Scale, SSI-26 = 26-Item Somatic Symptom Inventory, SSI-28 = 28-Item Somatic Symptom Inventory.

28.3% started at 30 mg q.d. Of those who started at 60 mg q.d., 7.6% required a dose reduction to 30 mg q.d. prior to visit 2. The mean (SD) number of days on duloxetine therapy was 42.5 (19.6), with a median of 48 days.

Ninety percent of patients who received duloxetine also took at least 1 concomitant medication, with alprazolam (11.9%), levothyroxine (11.8%), and zolpidem (10.2%) being most common.

Efficacy

The mean changes from baseline to end point for all the efficacy measures in patients treated with duloxetine 60 mg are presented in Table 2. Statistically significant ($p < .001$) improvement from baseline to end point was observed for all the efficacy scores.

Propensity scoring methods were used to compare mean changes from baseline to end point between subgroups for protocol-specified efficacy variables. Comparisons were made between ethnic origin subgroups with at least 80 patients treated with duloxetine (Caucasian, African descent, Hispanic), and also between Caucasian and non-Caucasian patients. Results are presented

Table 2. Mean Changes From Baseline to End Point in All the Efficacy Measures in Duloxetine-Treated Patients

Efficacy Measure	N	Baseline, Mean (SD)	Change, Mean[a] (SD)
CGI-S	3234	4.29 (0.90)	–1.43[a] (1.39)
SSI-28 average	2870	2.14 (0.72)	–0.26[a] (0.60)
QIDS-SR-16 total	3023	13.99 (5.01)	–4.69[a] (5.41)
SSI-26 average	2875	2.09 (0.70)	–0.24[a] (0.60)
CGI-S-PS	3224	3.95 (1.29)	–1.03[a] (1.36)
MAP-D total	2423	46.08 (16.32)	–13.71[a] (19.53)
Q-LES-Q-SF total	2963	43.96 (16.01)	12.74[a] (18.27)
SDS total	978	17.46 (7.56)	–5.01[a] (7.83)
SDS work/school	536	4.10 (2.72)	–1.28[a] (2.79)
SDS social	978	6.12 (2.77)	–1.79[a] (2.97)
SDS family/home	978	6.05 (2.70)	–1.71[a] (2.94)
PGI-IP[b]	2744	NA	3.05[a] (1.39)
PGI-IE[b]	2744	NA	2.97[a] (1.36)

[a]Significantly different from the baseline values ($p < .001$).
[b]Actual mean values at end point; p Value for difference from 4 (no change).
Abbreviations: CGI-S = Clinical Global Impressions-Severity of Illness, CGI-S-PS = Clinical Global Impressions-Severity of Physical Symptoms, MAP-D = Mood And Physical symptoms in Depression, NA = not applicable, PGI-IE = Patient Global Impressions-Improvement of Emotional symptoms, PGI-IP = Patient Global Impressions-Improvement of Physical Symptoms, QIDS-SR-16 = 16-item Quick Inventory of Depressive Symptomatology-Self Report, Q-LES-Q-SF = Quality of Life Enjoyment and Satisfaction Questionnaire-Short Form, SDS = Sheehan Disability Scale, SSI-26 = 26-item Somatic Symptom Inventory, SSI-28 = 28-item Somatic Symptom Inventory.

by gender (Table 3), ethnic origin (Table 4), age (Table 5), and patient care setting (Table 6). All of the efficacy measures were improved significantly ($p < .001$) from baseline to end point in all the subgroups. The propensity scoring analyses showed that the estimated predicted group difference in the efficacy measures, except for the SSI-28 average, by gender was significantly ($p \leq .05$) in favor of the female gender (Table 3). There were no statistically significant differences between Caucasian and African descent or Caucasian and Hispanic patients. All of the efficacy measures were in favor of the Caucasian subgroup, but none of the differences between Caucasian and non-Caucasian origin were significant (Table 4). Only CGI-S-PS scores were significantly in favor of the < 65 years age group; other efficacy measures were not significantly different between the 2 age groups (Table 5). All of the efficacy measures, except for the SSI-28 average, were significantly in favor of the primary care as compared with the psychiatric care patient setting (Table 6).

Safety

Serious adverse events among patients treated with duloxetine were reported for a total of 62 patients (1.8%). The most frequently reported serious adverse events were depression (N = 5), anxiety (N = 4), suicidal ideation (N = 4), and suicide attempt (N = 4). No deaths occurred in this study.

Table 3. Mean Change From Baseline to End Point and Estimated Predicted Group Difference for the Efficacy Measures in Duloxetine-Treated Patients by Gender Subgroup

Efficacy Measure	Subgroup	N	Baseline, Mean (SD)	Change, Mean (SD)	EPGD[b] (F – M)
CGI-S	Female	2437	4.32 (0.89)	–1.48[a] (1.40)	–0.18**
	Male	796	4.22 (0.95)	–1.29[a] (1.33)	
SSI-28 average	Female	2164	2.20 (0.73)	–0.29[a] (0.62)	–0.05
	Male	705	1.96 (0.66)	–0.18[a] (0.54)	
QIDS-SR-16 total	Female	2273	14.26 (4.87)	–4.94[a] (5.47)	–0.72**
	Male	749	13.15 (5.32)	–3.93[a] (5.13)	
CGI-S-PS	Female	2429	3.98 (1.27)	–1.06[a] (1.37)	–0.12*
	Male	793	3.83 (1.36)	–0.94[a] (1.32)	
MAP-D total	Female	1843	47.34 (16.06)	–14.72[a] (20.22)	–2.71**
	Male	579	42.14 (16.52)	–10.50[a] (16.80)	
PGI-IP[c]	Female	2084	NA	2.99[a] (1.40)	–0.26***
	Male	659	NA	3.22[a] (1.34)	
PGI-IE[c]	Female	2084	NA	2.92[a] (1.37)	–0.22***
	Male	659	NA	3.13[a] (1.30)	

[a]Significantly different from baseline ($p < .001$)
[b]Significantly different between female and male subgroups favoring females (*$p \le .05$; **$p \le .01$; ***$p \le .001$).
[c]Actual values at end point; p Values for difference from 4 (no change).
Abbreviations: CGI-S = Clinical Global Impressions-Severity of Illness, CGI-S-PS = Clinical Global Impressions-Severity of Physical Symptoms, EPGD = estimated predicted group difference, F = female, M = male, MAP-D = Mood And Physical symptoms in Depression, NA = not applicable, PGI-IE = Patient Global Impressions-Improvement of Emotional Symptoms, PGI-IP = Patient Global Impressions-Improvement of Physical Symptoms, QIDS-SR-16 = 16-item Quick Inventory of Depressive Symptomatology-Self Report, SSI-28 = 28-Item Somatic Symptom Inventory.

Table 4. Mean Change From Baseline to End Point and Estimated Predicted Group Difference for the Primary Efficacy Measures by Ethnic Origin Subgroup

Efficacy Measure	Subgroup	N	Baseline, Mean (SD)	Change, Mean (SD)	EPGD[b]
CGI-S	Caucasian	2950	4.29 (0.89)	–1.43[a] (1.38)	
	African descent	107	4.21 (0.95)	–1.36[a] (1.38)	–0.04[c]
	Hispanic	130	4.53 (0.93)	–1.45[a] (1.44)	0.14[d]
	Non-Caucasian	283	4.36 (0.99)	–1.39[a] (1.41)	0.01[e]
SSI-28 average	Caucasian	2636	2.13 (0.72)	–0.25[a] (0.59)	
	African descent	94	2.19 (0.77)	–0.28[a] (0.51)	0.04[c]
	Hispanic	97	2.41 (0.80)	–0.49[a] (0.80)	0.14[d]
	Non-Caucasian	233	2.26 (0.80)	–0.33[a] (0.66)	0.04[e]
QIDS-SR-16 total	Caucasian	2763	13.90 (4.99)	–4.66[a] (5.35)	
	African descent	100	14.73 (4.89)	–4.63[a] (5.80)	1.07[c]
	Hispanic	115	15.50 (5.24)	–5.56[a] (6.11)	–0.20[d]
	Non-Caucasian	259	14.89 (5.19)	–4.93[a] (5.97)	0.01[e]
CGI-S-PS	Caucasian	2939	3.94 (1.29)	–1.03[a] (1.36)	
	African descent	107	3.84 (1.19)	–0.88[a] (1.20)	0.11[c]
	Hispanic	130	4.29 (1.20)	–1.25[a] (1.42)	0.07[d]
	Non-Caucasian	283	4.04 (1.30)	–1.03[a] (1.31)	0.05[e]
MAP-D total	Caucasian	2234	45.68 (16.10)	–13.65[a] (19.24)	
	African descent	67	49.75 (20.43)	–13.45[a] (21.43)	3.45[c]
	Hispanic	85	53.65 (15.84)	–17.76[a] (23.49)	0.63[d]
	Non-Caucasian	188	50.96 (18.12)	–14.48[a] (22.75)	0.13[e]
PGI-IP[f]	Caucasian	2521	NA	3.04[a] (1.38)	
	African descent	81	NA	3.01[a] (1.26)	–0.02[c]
	Hispanic	103	NA	2.92[a] (1.44)	–0.09[d]
	Non-Caucasian	222	NA	3.08[a] (1.44)	0.01[e]
PGI-IE[f]	Caucasian	2521	NA	2.96[a] (1.36)	
	African descent	81	NA	2.95[a] (1.20)	0.01[c]
	Hispanic	103	NA	2.89[a] (1.41)	0.01[d]
	Non-Caucasian	222	NA	3.01[a] (1.34)	0.05[e]

[a]Significantly different from baseline ($p < .001$).
[b]No significant differences between Caucasians minus other subgroups ($p > .05$).
[c]EPGD between subgroups: Caucasian minus African-descent subgroups.
[d]EPGD between subgroups: Hispanic minus Caucasian subgroups.
[e]EPGD between subgroups: Caucasian minus non-Caucasian subgroups.
[f]Actual values at end point; p Values for difference from 4 (no change).
Abbreviations: CGI-S = Clinical Global Impressions-Severity of Illness, CGI-S-PS = Clinical Global Impressions-Severity of Physical Symptoms, EPGD = estimated predicted group difference, MAP-D = Mood And Physical symptoms in Depression, NA = not applicable, PGI-IE = Patient Global Impressions-Improvement of Emotional Symptoms, PGI-IP = Patient Global Impressions-Improvement of Physical Symptoms, QIDS-SR-16 = 16-item Quick Inventory of Depressive Symptomatology-Self Report, SSI-28 = 28-item Somatic Symptom Inventory.

Table 5. Mean Change From Baseline to End Point and Estimated Predicted Group Difference for the Efficacy Measures in All Duloxetine-Treated Patients by Age Subgroup

Efficacy Measure	Subgroup Age, y	N	Baseline, Mean (SD)	Change, Mean (SD)	EPGD[b] (< 65–≥ 65)
CGI-S	< 65	2940	4.30 (0.90)	–1.44[a] (1.38)	–0.17
	≥ 65	289	4.22 (0.93)	–1.34[a] (1.43)	
SSI-28 average	< 65	2736	2.14 (0.73)	–0.27[a] (0.60)	–0.19
	≥ 65	230	2.08 (0.64)	–0.14[a] (0.54)	
QIDS-SR-16 total	< 65	2766	14.14 (5.01)	–4.79[a] (5.43)	–0.43
	≥ 65	253	12.28 (4.70)	–3.61[a] (5.07)	
CGI-S-PS	< 65	2930	3.93 (1.31)	–1.04[a] (1.36)	–0.21*
	≥ 65	288	4.10 (1.11)	–0.95[a] (1.29)	
MAP-D total	< 65	2205	46.68 (16.12)	–14.02[a] (19.46)	–1.63
	≥ 65	215	40.24 (17.14)	–10.51[a] (20.00)	
PGI-IP[c]	< 65	2502	NA	3.05[a] (1.37)	0.10
	≥ 65	238	NA	2.98[a] (1.50)	
PGI-IE[c]	< 65	2502	NA	2.97[a] (1.36)	0.05
	≥ 65	238	NA	2.97[a] (1.38)	

[a]Significantly different from baseline ($p < .001$).
[b]Significantly different between < 65 years and ≥ 65 years subgroups favoring < 65 years group (*$p \leq .05$).
[c]Actual values at end point; p Values for difference from 4 (no change).
Abbreviations: CGI-S = Clinical Global Impressions-Severity of Illness, CGI-S-PS = Clinical Global Impressions-Severity of Physical Symptoms, EPGD = estimated predicted group difference, MAP-D = Mood And Physical symptoms in Depression, NA = not applicable, PGI-IE = Patient Global Impressions-Improvement of Emotional Symptoms, PGI-IP = Patient Global Impressions-Improvement of Physical Symptoms, QIDS-SR-16 = 16-item Quick Inventory of Depressive Symptomatology-Self Report, SSI-28 = 28-item Somatic Symptom Inventory.

Table 6. Mean Change From Baseline to End Point and Estimated Predicted Group Difference for the Efficacy Measures in Duloxetine-Treated Patients by Clinical Practice Subgroup

Efficacy Measure	Subgroup	N	Baseline, Mean (SD)	Change, Mean (SD)	EPGD[b] (PC–P)
CGI-S	Primary care	1906	4.22 (0.88)	–1.57[a] (1.40)	–0.36***
	Psychiatric	1328	4.41 (0.92)	–1.23[a] (1.34)	
SSI-28 average	Primary care	1683	2.15 (0.72)	–0.31[a] (0.62)	–0.07
	Psychiatric	1187	2.12 (0.73)	–0.19[a] (0.57)	
QIDS–SR-16 total	Primary care	1794	13.45 (4.78)	–5.01[a] (5.31)	–1.08***
	Psychiatric	1229	14.77 (5.23)	–4.21[a] (5.50)	
CGI-S-PS	Primary care	1898	3.96 (1.27)	–1.18[a] (1.39)	–0.31***
	Psychiatric	1326	3.93 (1.33)	–0.81[a] (1.27)	
MAP-D total	Primary care	1410	44.88 (16.21)	–15.59[a] (19.71)	–3.76***
	Psychiatric	1013	47.75 (16.34)	–11.10[a] (18.97)	
PGI-IP[c]	Primary care	1600	NA	2.90[a] (1.34)	–0.31***
	Psychiatric	1144	NA	3.26[a] (1.42)	
PGI-IE[c]	Primary care	1600	NA	2.83[a] (1.29)	–0.28***
	Psychiatric	1144	NA	3.17[a] (1.42)	

[a]Significantly different from baseline ($p < .001$).
[b]Significantly different between primary care and psychiatric practice subgroups favoring primary care (***$p \leq .001$).
[c]Actual values at end point; p Values for difference from 4 (no change).
Abbreviations: CGI-S = Clinical Global Impressions-Severity of Illness, CGI-S-PS = Clinical Global Impressions-Severity of Physical Symptoms, EPGD = estimated predicted group difference, MAP-D = Mood And Physical symptoms in Depression, NA = not applicable, P = psychiatric, PC = primary care, PGI-IE = Patient Global Impressions-Improvement of Emotional Symptoms, PGI-IP = Patient Global Impressions-Improvement of Physical Symptoms, QIDS-SR-16 = 16-item Quick Inventory of Depressive Symptomatology-Self Report, SSI-28 = 28-item Somatic Symptom Inventory.

A total of 367 patients (10.8%) who took at least 1 dose of duloxetine discontinued due to adverse events (Table 7); nausea was the most frequent (2.7%), followed by headache (0.6%), and fatigue and insomnia with 0.5% each. As shown in Table 7, approximately half of the patients (56.4%) who took at least 1 dose of duloxetine experienced at least 1 treatment-emergent adverse event. Treatment-emergent adverse events reported by ≥ 5% of patients were nausea (12.9%), headache (6.3%), insomnia (6.1%), and fatigue (5.1%) (Table 7).

A statistically significant ($p < .001$), but not clinically relevant, mean increase of 1.7 beats per minute in heart rate was observed from baseline to end point. No statistically significant changes were observed for systolic (mean change: 0.27 mm Hg, $p = .313$) or diastolic (mean change: 0.18 mm Hg, $p = .327$) blood pressure in patients treated with duloxetine.

A statistically significant, but not clinically relevant, decrease from baseline to end point in mean weekly alcohol consumption was observed in patients receiving duloxetine treatment, based on the self-reported number of beers or wine coolers/spritzers ($p = .019$) and glasses of wine consumed ($p < .001$), while no statistically significant ($p = .369$) difference from baseline to end-

Table 7. Adverse Events Reported as the Reason for Discontinuation (DCAEs) by at Least 5 Duloxetine-Treated Patients and Treatment-Emergent Adverse Events (TEAEs) Reported by at Least 2.0% of Duloxetine-Treated Patients

Event	DCAEs, N (%) (N = 3373)[a]	TEAEs, N (%) (N = 3431)
Any	367 (10.8)	1934 (56.4)
Nausea	90 (2.7)	443 (12.9)
Headache	19 (0.6)	217 (6.3)
Fatigue	17 (0.5)	176 (5.1)
Insomnia	17 (0.5)	210 (6.1)
Anxiety	11 (0.3)	91 (2.7)
Dizziness	11 (0.3)	119 (3.5)
Depression	10 (0.3)	82 (2.4)
Somnolence	8 (0.2)	74 (2.2)
Agitation	7 (0.2)	19 (0.6)[b]
Vomiting	7 (0.2)	69 (2.0)
Diarrhea	6 (0.2)	101 (2.9)
Dysuria	5 (0.1)	17 (0.5)[b]
Irritability	5 (0.1)	33 (1.0)[b]
Sedation	5 (0.1)	20 (0.6)[b]
Constipation	5 (0.1)	158 (4.6)
Dry mouth	0[b]	109 (3.2)
Hypertension	3 (0.1)[b]	89 (2.6)

[a]Excluding patients with reason for discontinuation unknown (41 patients) or specific adverse event reported as the reason for discontinuation unknown (17 patients); the 41 patients are included in the denominator for the calculation of the percentage of patients discontinuing due to any adverse events.
[b]Included for comparison, although they are below specified number of patients.

point was noted for number of drinks containing distilled spirits.

DISCUSSION

Duloxetine demonstrated significant improvements in the emotional and physical symptoms of depression in this 7-week, open-label study of 3543 patients from over 800 practice-based primary care and psychiatric settings. Significant improvements were seen for both primary and secondary measures regardless of gender, ethnic origin, age, or whether care for major depressive disorder was provided in a primary care or psychiatric setting. The end result of any treatment should be improved function and quality of life; these domains, as measured by the Q-LES-Q-SF and SDS, were significantly improved in this naturalistic study. In addition, duloxetine 60 mg q.d. was found to be safe and well-tolerated in terms of rates of treatment-emergent adverse events and discontinuations due to adverse events. The safety and tolerability profile of duloxetine in this open-label study was similar to that in patients who participated in the placebo-controlled clinical trials.[3–6,15]

In exploring group differences, the efficacy measures, except SSI-28 average, by gender were found to be significantly ($p \leq .05$) in favor of women and by patient care setting in favor of primary care. Conversely, no statistically significant differences in the efficacy measures were observed between ethnic origin subgroups; although the number of patients in the ethnic subgroups other than Caucasian was relatively small, statistical comparisons among the other ethnic origin subgroups would be relatively underpowered. For age subgroup comparisons, only the CGI-S-PS was significantly in favor of those < 65 years of age. Other efficacy measures were not significantly different between the 2 age groups.

Prior duloxetine efficacy studies compared subgroups of patients with MDD by gender,[24] age,[25] and ethnic origin.[26,27] But these were secondary analyses from randomized controlled clinical trials that may not capture a diverse patient population, and they included more stringent participant eligibility criteria.

In this naturalistic study, it cannot be determined whether significant subgroup differences in efficacy measures were due to true differences in treatment responsiveness or to the nonrandomization of subgroup membership. The propensity scoring method was used in order to minimize bias when comparing subgroups; however, some potentially key factors were not obtained in this study that may have enhanced adjustments between groups. Data collection was modest in order to minimally interfere with busy clinical practice sites, which do not ordinarily participate in clinical trials. For example, psychiatric and medical conditions comorbid with depression previously have been shown to be risk factors for poor response to treatment, and they also may be more prevalent in one subgroup over another.[11,28–30] Physical symptoms, especially painful complaints, were shown to predict greater severity of depression,[11,31–34] worse responsiveness to antidepressant treatment,[34] lower rates of remission,[35] and longer time to remission.[36] The propensity scoring analyses included medications as a proxy for comorbid conditions, but further information on comorbid conditions may be necessary to determine differences due to treatment effects.

Another large prospective study of treatment for depression in both psychiatric and primary care settings, the STAR*D trial, found that non-Caucasian race, male gender, unemployment, lower income, less education, poorer functional status, and lower quality of life at baseline were overlapping and independently associated with lower remission rates.[11] Remission status by age or patient care setting did not significantly differ in patients in the STAR*D trial.[11] Other studies conducted in only primary care settings reported mixed results.[12,13] Older age was significantly associated with a poorer response in patients treated with SSRIs[13] and was nonsignificant in patients treated with care deemed appropriate by standard practice guidelines for depression.[12] Ethnic origin was also found to be nonsignificant.[12]

In another study, remission was found to be comparable in primary care and mental health specialty settings when an intervention was implemented to facilitate the

use of depression treatment guidelines.[37] Primary care has been found to negatively differ from mental health specialty care on indicators of quality of depression treatment.[38] Shasha et al.[39] found that psychiatrists were more likely to prescribe antidepressants at an adequate dosage level but nonpsychiatric physicians were more likely to attain adequate duration of treatment. Current evidence suggests that collaborative care models most strongly improve both the likelihood of quality treatment and outcome, especially in depressed patients who are prescribed adequate dosages of antidepressants.[40] A possible explanation for the apparent difference in response between primary care and psychiatric care in the current study of outpatients is increased recognition and treatment of painful physical symptoms by primary care physicians as compared with psychiatrists, who, in general, are more concerned with the emotional symptoms associated with MDD. Also, the patient characteristics of patients seen at these 2 practice settings may differ.

A large percentage (71.4%) of patients who took at least 1 dose of duloxetine completed the study, with only 10.8% of patients discontinuing due to an adverse event. The study design allowed a dose adjustment based on the investigator's concerns regarding tolerability. Only 7.6% of those patients started on a dose of 60 mg q.d. required a dose reduction to 30 mg q.d. prior to visit 2, suggesting few early tolerability issues. Controlled studies of duloxetine that allowed down-titration from an initial dose of 60 mg daily similarly demonstrated that few patients required dose reduction to 30 mg in the first week of therapy. This study differs from typical randomized clinical trials in that it allowed broader inclusion criteria and was carried out in outpatient-based practice settings. Recently, literature has highlighted the importance of conducting research that measures both treatment efficacy and effectiveness across clinically relevant outcomes.[1,41,42] Greater emphasis on the importance of real-world trials has resulted from the recent publications of 2 National Institutes of Health trials, STAR*D and the Clinical Antipsychotic Trials of Intervention Effectiveness.[11,43]

Study Limitations

This study has several limitations. (1) As mentioned earlier, nonrandomization limits the ability to make unbiased comparisons in outcomes between subgroups of patients. For example, the covariates included in the propensity scoring analysis may not have completely removed the bias when comparing the subgroups. (2) The sample size was smaller than anticipated, and, therefore, we are unable to make comparisons involving the Asian population. (3) Some of the data on efficacy measures were incomplete, especially patient-rated data obtained through the Interactive Voice Response System. (4) Recording of spontaneous adverse events may have been incomplete.

Summary

In summary, both primary and secondary efficacy measures demonstrated that duloxetine 60 mg q.d. in this 7-week open-label study was effective in the treatment of major depressive disorder, regardless of gender, ethnic origin, age, or patient care setting, and was well tolerated in a diverse population of outpatients with MDD. This study complements findings from randomized, controlled clinical trials and provides a different dimension of knowledge that allows a broad, inclusive, and more generalizable understanding of treatment for patients with depression.

Drug names: alprazolam (Xanax, Niravam, and others), duloxetine (Cymbalta), escitalopram (Lexapro and others), levothyroxine (Tirosint, Synthroid, and others), zolpidem (Ambien).

Financial disclosure: Drs. Wohlreich, Wiltse, and Ye and Ms. Robinson are employees and stock shareholders of Eli Lilly. Dr. Desaiah is an employee of Eli Lilly. Dr. Kroenke has received grant/research support from Wyeth, Eli Lilly, and Pfizer; has received honoraria from Wyeth; and has served on the speakers/advisory board for Eli Lilly. Dr. Kornstein has received grant/research support from the U.S. Dept. of Health and Human Services, the National Institute of Mental Health, Pfizer, Bristol-Myers Squibb, Eli Lilly, Forest, GlaxoSmithKline, Mitsubishi-Tokyo, Merck, Wyeth, Berlex, Novartis, Sepracor, Boehringer-Ingelheim, Sanofi-Synthelabo, and AstraZeneca; has served on advisory boards and/or received honoraria from Pfizer, Wyeth, Eli Lilly, Bristol-Myers Squibb, Warner-Chilcott, Biovail, Berlex, and Forest; and has received book publishing royalties from Guilford Press. Dr. Greist has received grant/research support from Bristol-Myers Squibb, Forest, GlaxoSmithKline, Janssen, Eli Lilly, Organon, Pfizer, Solvay, UCB, and Wyeth-Ayerst; has served as a consultant for Bristol-Myers Squibb, GlaxoSmithKline, Eli Lilly, Pfizer, and Solvay; has served on the speakers bureaus for Bristol-Myers Squibb, Cyberonics, GlaxoSmithKline, Eli Lilly, Ortho-McNeil, Pfizer, Solvay, and Wyeth-Ayerst; and is a principal in Healthcare Technology Systems, Inc.

REFERENCES

1. Roy-Byrne PP, Sherbourne CD, Craske MG, et al. Moving treatment research from clinical trials to the real world. Psychiatr Serv 2003;54: 327–332
2. Bymaster FP, Dreshfield-Ahmed LJ, Threlkeld PG, et al. Comparative affinity of duloxetine and venlafaxine for serotonin and norepinephrine transporters in vitro and in vivo, human serotonin receptor subtypes, and other neuronal receptors. Neuropsychopharmacology 2001;25:871–880
3. Detke MJ, Lu Y, Goldstein DJ. Duloxetine, 60 mg once daily, for major depressive disorder: a randomized double-blind placebo-controlled trial. J Clin Psychiatry 2002;63:308–315
4. Detke MJ, Lu Y, Goldstein DJ, et al. Duloxetine 60 mg once daily dosing versus placebo in the acute treatment of major depression. J Psychiatr Res 2002;36:383–390
5. Goldstein DJ, Mallinckrodt C, Lu Y, et al. Duloxetine in the treatment of major depressive disorder: a double-blind clinical trial. J Clin Psychiatry 2002;63:225–231
6. Goldstein DJ, Lu Y, Detke MJ, et al. Duloxetine in the treatment of depression: a double-blind placebo-controlled comparison with paroxetine. J Clin Psychopharmacol 2004;24:389–399
7. Wohlreich MM, Martinez JM, Mallinckrodt CH, et al. An open-label study of duloxetine for the treatment of major depressive disorder: comparison of switching versus initiating treatment approaches. J Clin Psychopharmacol 2005;25:552–560
8. Rollman BL, Hanusa BH, Belnap BH, et al. Race, quality of depression care, and recovery from a depression in primary care setting. Gen Hosp Psychiatry 2002;24:381–390
9. Gilmer WS, Trivedi MH, Rush AJ, et al. Factors associated with chronic

depressive episodes: a primary report from the STAR-D project. Acta Psychiatr Scand 2005;112:425–433
10. Kales HC, Neighbors HW, Blow FC, et al. Race, gender, and psychiatrists' diagnosis and treatment of major depression among elderly patients. Psychiatr Serv 2005;56:721–728
11. Trivedi MH, Rush AJ, Wisniewski SR, et al. Evaluation of outcomes with citalopram for depression using measurement-based care in STAR*D: implications for clinical practice. Am J Psychiatry 2006;163: 28–40
12. Sherbourne C, Schoenbaum M, Wells KB, et al. Characteristics, treatment patterns, and outcomes of persistent depression despite treatment in primary care. Gen Hosp Psychiatry 2004;26:106–114
13. Corey-Lisle PK, Nash R, Stang P, et al. Response, partial response, and nonresponse in primary care treatment of depression. Arch Intern Med 2004;164:1197–1204
14. Rush AJ, Bose A. Escitalopram in clinical practice: results of an open-label trial in a naturalistic setting. Depress Anxiety 2005;21:26–32
15. Nemeroff CB, Schatzberg AF, Goldstein DJ, et al. Duloxetine for the treatment of major depressive disorder. Psychopharmacol Bull 2002;36: 106–132
16. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition. Washington, DC: American Psychiatric Association; 1994
17. Guy W. ECDEU Assessment Manual for Psychopharmacology, revised. Rockville, Md: National Institute of Mental Health (NIMH), Psychopharmacology Research Branch; 1976:217–222, 313–331
18. Barsky AJ, Wyshak G, Klerman GL. Hypochondriasis: an evaluation of the DSM-III criteria in medical outpatients. Arch Gen Psychiatry 1986; 43:493–500
19. Rush AJ, Trivedi MH, Ibrahim HM, et al. 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 2003;54:573–583
20. Endicott J, Nee J, Harrison W, et al. Quality of Life Enjoyment and Satisfaction Questionnaire: a new measure. Psychopharmacol Bull 1993; 29:321–326
21. Sheehan DV, Harnett-Sheehan K, Raj BA. The measurement of disability. Int Clin Psychopharmacol 1996;11(suppl 3):89–95
22. Rosenbaum PR, Rubin DB. Reducing bias in observational studies using subclassification on the propensity score. J Am Stat Assoc 1984;79: 516–524
23. Lunceford JK, Davidian M. Stratification and weighting via the propensity score in estimation of causal treatment effects: a comparative study. Stat Med 2004;23:2937–2960
24. Kornstein SG, Wohlreich MM, Mallincrodt CH, et al. Duloxetine efficacy for major depressive disorder in male vs female patients: data from 7 randomized, double-blind, placebo-controlled trials. J Clin Psychiatry 2006;67:761–770
25. Wohlreich MW, Mallinckrodt CH, Watkin JG, et al. Duloxetine for the long-term treatment of major depressive disorder in patients aged 65 and older: an open-label study. BMC Geriatrics 2004;4:11–20
26. Lewis-Fernandez R, Blanco C, Mallincrodt CH, et al. Duloxetine in the treatment of major depressive disorder: comparisons of safety and efficacy in US Hispanic and majority Caucasian patients. J Clin Psychiatry 2006;67:1379–1390
27. Bailey RK, Mallincrodt CH, Wohlreich MM, et al. Duloxetine in the treatment of major depressive disorder: comparisons of safety and efficacy. J Natl Med Assoc 2006;98:437–447
28. Bair MJ, Robinson RL, Eckert GJ, et al. Impact of pain on depression treatment response in primary care. Psychosom Med 2004;66:17–22
29. Cuffel BJ, Azocar F, Tomlin M, et al. Remission, residual symptoms, and nonresponse in the usual treatment of major depression in managed clinical practice. J Clin Psychiatry 2003;64:397–402
30. Swindle RW Jr, Cronkite RC, Moos RH. Risk factors for sustained nonremission of depressive symptoms: a 4-year follow-up. J Nerv Ment Dis 1998;186:462–469
31. Kroenke K, Spitzer RL, Williams JB, et al. Physical symptoms in primary care. Predictors of psychiatric disorders and functional impairment. Arch Fam Med 1994;3:774–779
32. Gerber PD, Barrett JE, Barrett JA, et al. The relationship of presenting physical complaints to depressive symptoms in primary care patients. J Gen Intern Med 1992;7:170–173
33. Kroenke K. Physical symptom disorder: a simpler diagnostic category for somatization-spectrum conditions. J Psychosom Res 2006;60:335–339
34. Bair MJ, Robinson RL, Katon W, et al. Depression and pain comorbidity: a literature review. Arch Intern Med 2003;163:2433–2445
35. Fava M, Mallinckrodt CH, Detke MJ, et al. The effect of duloxetine on painful physical symptoms in depressed patients: do improvements in these symptoms result in higher remission rates? J Clin Psychiatry 2004; 65:521–530
36. Karp JF, Scott J, Houck P, et al. Pain predicts longer time to remission during treatment of recurrent depression. J Clin Psychiatry 2005;66: 591–597
37. Thomas L, Mulsant BH, Solano FX, et al. Response speed and rate of remission in primary and specialty care of elderly patients with depression. Am J Geriatr Psychiatry 2002;10:583–591
38. Robinson RL, Long SR, Chang S, et al. Higher costs and therapeutic factors associated with adherence to NCQA HEDIS antidepressant medication management measures: analysis of administrative claims. J Manag Care Pharm 2006;12:43–54
39. Shasha M, Lyons JS, O'Mahoney MT, et al. Serotonin reuptake inhibitors and the adequacy of antidepressant treatment. Int J Psychiatry Med 1997; 27:83–92
40. Vergouwen ACM, Bakker A, Katon WJ, et al. Improving adherence to antidepressants: a systematic review of interventions. J Clin Psychiatry 2003;64:1415–1420
41. Wells KB. Treatment research at the crossroads: the scientific interface of clinical trials and effectiveness research. Am J Psychiatry 1999;156:5–10
42. Zimmerman M, Chelminski I, Posternak MA. Generalizability of antidepressant efficacy trials: differences between depressed psychiatric outpatients who would or would not qualify for an efficacy trial. Am J Psychiatry 2005;162:1370–1372
43. Keefe RS, Mohs RC, Bilder RM, et al. Neurocognitive assessment in the Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) project schizophrenia trial: development, methodology, and rationale. Schizophr Bull 2003;29:45–55