A Double-Blind Comparison of Sertraline and Fluoxetine in Depressed Elderly Outpatients

Paul A. Newhouse, M.D.; K. Ranga Rama Krishnan, M.D.; P. Murali Doraiswamy, M.D.; Ellen M. Richter, Ph.D.; Evan D. Batzar, M.S.; and Cathryn M. Clary, M.D.

Background: There has been a paucity of well-designed studies comparing selective serotonin reuptake inhibitor (SSRI) medications in the treatment of depression in the elderly. This multicenter study was designed to examine the efficacy and safety of sertraline and fluoxetine in depressed elderly outpatients. A secondary objective was to examine the effects of SSRI treatment on quality of life and cognitive function.

Method: Two hundred thirty-six outpatients 60 years of age and older who met DSM-III-R criteria for major depressive disorder received 1 week of single-blind placebo before being randomly assigned to 12 weeks of double-blind, parallel-group treatment with flexible daily doses of either sertraline (range, 50–100 mg) or fluoxetine (range, 20–40 mg). Primary efficacy measures consisted of the 24-item Hamilton Rating Scale for Depression and Clinical Global Impressions scale ratings. Secondary outcome assessments included clinician- and patient-rated measures of depression symptoms and factors, cognitive functioning, and quality of life, as well as plasma drug concentrations, which were correlated with clinical response.

Results: Both drugs produced a similarly positive response on the primary efficacy measures, with 12-week responder rates of 73% for sertraline and 71% for fluoxetine. Sertraline-treated patients showed statistically greater cognitive improvement on several measures. Both drugs were safe and well tolerated.

Conclusion: Data indicate that both drugs are effective antidepressants for the treatment of depressed elderly outpatients. Differences in cognitive performance effects deserve further investigation.

(J Clin Psychiatry 2000;61:559–568)

Received Dec. 1, 1998; accepted March 21, 2000. From the Clinical Neuroscience Research Unit and Geriatric Psychiatry Service, Department of Psychiatry, University of Vermont College of Medicine, Burlington (Dr. Newhouse); the Department of Psychiatry, Duke University Medical Center, Durham, N.C. (Drs. Krishnan and Doraiswamy); and Pfizer Inc, New York, N.Y. (Drs. Richter and Clary and Mr. Batzar).

Supported by a grant from Pfizer Inc.

Financial disclosure: Dr. Newhouse has received grant/research support from and has served as a consultant for Pfizer Inc, Eli Lilly, SmithKline Beecham, Parke-Davis, Merck, Inc., Aventis, and Targacept. Dr. Krishnan has been under contract with Pfizer Inc; has received grant/research support from Pfizer Inc and Eli Lilly; has received honoraria from Pfizer Inc and Eli Lilly; and has served as a consultant for Pfizer Inc. Dr. Doraiswamy has received grant/research support from Pfizer Inc; has served as a speaker for Eli Lilly; has received honoraria from Forest; and has occasionally been a stockholder in SmithKline Beecham. Drs. Richter and Clary and Mr. Batzar are employees of Pfizer Inc.

Partial results of the current study were presented as a poster at the New Clinical Drug Evaluation Unit (NCDEU) annual meeting, Boca Raton, Fla., May 1996.

The authors wish to acknowledge the hard work of all of the investigators, clinical coordinators, and staff who assisted in the successful completion of this study. Investigators included, in addition to the authors: Mildred Farmer, M.D.; James Ferguson, M.D.; Burton Goldstein, M.D.; James Hartford, M.D.; Bharat Raj Nakra, M.D.; Lawrence Lazarus, M.D.; Robert Linden, M.D.; Richard Margolin, M.D.; Robert Riesenberg, M.D.; and Michael Rothman, M.D.

Reprint requests to: Paul A. Newhouse, M.D., Clinical Neuroscience Research Unit and Geriatric Psychiatry Service, Department of Psychiatry, University of Vermont College of Medicine, 1 South Prospect St., Burlington, VT 05401 (e-mail: pnewhous@zoo.uvm.edu).

Surveys of healthy elderly individuals have revealed that 15% of community-dwelling elderly report clinically significant degrees of depressed mood, 4% suffer from a current major depressive disorder, and 6.5% have depression associated with a significant medical illness.[1] The consequences of untreated or inadequately treated depressive disorder in the elderly are serious. Suicide rates rise over the life cycle and increase dramatically in old age.[2] Increased rate of utilization of medical services, increased morbidity and mortality from medical illnesses, polypharmacy, and inappropriate institutionalization are all well-studied consequences of untreated depressive disorders in the elderly.[3–5] A study of medical outcomes[6] revealed that the disability of patients with major depressive illness was comparable to that of most other chronic diseases of the elderly, including diabetes, arthritis, and hypertension. For decades, tricyclic antidepressants (TCAs) have been the standard treatment for

depression in the elderly, although they are often relatively contraindicated in this age group. For example, in an examination of the detection and treatment of depression in hospitalized elderly men, Koenig and coauthors[7] found that over 85% of the patients had a relative or absolute medical contraindication to the use of older TCAs. Although anticholinergic and other adverse events associated with TCAs may pose a significant problem for the elderly, evidence from a recent meta-analysis suggests that secondary amine tricyclics may be better tolerated.[8]

In the past decade, selective serotonin reuptake inhibitors (SSRIs) have increasingly been used in the treatment of the depressed elderly (Scott-Levin Physician Drug and Diagnosis Audit for 1995–1998, Scott-Levin, a division of PMSI Scott-Levin, Inc.). A few randomized clinical studies, mostly non–placebo controlled, suggest that SSRIs are efficacious in geriatric depression, with response rates that are generally comparable to those with traditional agents (reviewed in Newhouse[9]). Side effect profiles are different from those of tricyclics and monoamine oxidase inhibitors (MAOIs) and appear to lead to lower rates of discontinuation, at least in younger populations.[10]

Fluoxetine, the first SSRI available, appears generally efficacious and generally better tolerated than TCAs by the elderly.[11] However, case reports, and indirect comparisons based on clinical experience, suggest that significant side effects, including agitation, insomnia, and weight loss, may occur more frequently with fluoxetine in the elderly than with other SSRIs.[12–15] Further, fluoxetine, at standard clinical doses, is a relatively potent inhibitor of the hepatic isoenzyme cytochrome P450 2D6 (CYP2D6), as is its active metabolite, norfluoxetine, at normal clinical doses.[16] This may produce more potential for drug-drug interactions in the elderly, a population at greater risk for polypharmacy.

Sertraline, like fluoxetine, has demonstrated efficacy in elderly patients.[17] It has a lower potential for certain drug interactions than fluoxetine since it produces less clinically meaningful inhibition of CYP2D6 at commonly used dosages.[16,18] Overall, the tolerability of sertraline in the elderly appears to be similar to what is reported in younger patients.[14] However, no direct comparative studies of sertraline and fluoxetine have been performed in elderly depressed patients. Furthermore, even though depression in the elderly has been found to be associated with impairment in both cognitive function[19–23] and quality of life,[24] these domains of function are not often examined as an outcome in treatment studies in this population. Therefore, this study was conducted with the primary goal of investigating the comparative efficacy and safety of sertraline and fluoxetine in the treatment of elderly outpatients suffering from major depression. A secondary goal was to assess the comparative effects of the 2 compounds on quality of life and cognitive functioning.

METHOD

This 12-week, double-blind, randomized, parallel-group study was conducted at 12 sites in the United States.

Patient Selection

Outpatients 60 years of age or older who met the *Diagnostic and Statistical Manual of Mental Disorders, Third Edition, Revised* (DSM-III-R)[25] criteria for major depressive disorder (single episode or recurrent, without psychotic features) and had a total score of 18 or more on the 24-item Hamilton Rating Scale for Depression (HAM-D) at the end of the washout were eligible for this study. Written informed consent was obtained from all patients prior to study entry.

Complete psychiatric histories were taken, and patients were excluded if they met DSM-III-R diagnostic criteria for any other psychiatric disorders. In addition, patients were required to score at least 24 on the Mini-Mental State Examination (MMSE)[26] to exclude significant cognitive impairment. Subjects were also excluded if they had any medical contraindications to antidepressant therapy; significant hematologic, endocrine, or cardiovascular disease; and conditions that might impair study drug absorption, metabolism, or excretion. Additional exclusions included failure to respond to electroconvulsive therapy in a prior depressive episode or to adequate trials (6 weeks) of 2 or more antidepressants in daily doses equivalent to 150 mg of amitriptyline administered for 3 weeks. Patients were informed that all other psychotropic medications (except temazepam or chloral hydrate used sparingly for sleep) were to be discontinued prior to entry into the study. This criterion was verified by urinalysis performed at the initial visit.

Patients had to be free of any clinically significant medical problems. Complete medical histories were taken, and a thorough physical examination was performed prior to study entry. A 12-lead electrocardiogram was recorded at day 1 of washout (and at the end of weeks 4 and 12 or when a patient discontinued the study). Normal baseline laboratory values were required, but patients with clinically insignificant abnormalities were eligible.

Study Procedures

Following approval by the designated Institutional Review Board, outpatients who met initial selection criteria and provided written informed consent entered a single-blind, placebo run-in phase to allow washout of any prior psychoactive medications and to assess placebo response. The 1-week run-in period could be extended if necessary to ensure fulfillment of all selection criteria.

At the conclusion of the run-in period (baseline), severity of illness was again assessed to ensure continued fulfillment of entry criteria (score of 18 or more on the 24-item HAM-D; Clinical Global Impressions-Severity

scale [CGI-S] score of 3 or more). Eligible patients were then randomly assigned to receive, in a double-blind fashion, either sertraline, 50 mg/day (1 capsule) in the evening, or fluoxetine, 20 mg/day (1 capsule) in the morning. A double-dummy procedure was used to ensure patient and physician blindness to treatment assignment. After 4 weeks of double-blind treatment, doses could be blindly doubled to 100 mg/day for sertraline or 40 mg/day for fluoxetine if, in the investigator's opinion, an adequate clinical response had not been observed and no dose-limiting side effects had occurred. A patient's daily dose could be reduced at any time to 50 mg of sertraline or 20 mg of fluoxetine owing to adverse events and/or if clinically indicated.

Clinical assessments were made at day 1 of washout, the end of washout (baseline), at weekly intervals for the first 4 weeks of double-blind treatment, and at 2-week intervals thereafter. Data analysis for efficacy included all patients for whom a baseline evaluation and at least one subsequent evaluation during drug treatment were available (intent-to-treat analysis). Data collected from patients who had received at least one dose of study medication, regardless of later compliance or protocol violation, were analyzed.

Efficacy Measurements

Primary investigator-rated efficacy measures included the 24-item HAM-D (total and factor scores),[27–29] CGI-S rating, CGI-Improvement scale (CGI-I) rating, and CGI-Efficacy Index rating.[30] Secondary investigator-rated measures of efficacy were the Montgomery-Asberg Depression Rating Scale (MADRS)[31] and the Hamilton Rating Scale for Anxiety (HAM-A).[32] Secondary measurements also included patient-rated measures: the Profile of Mood States (POMS)[33] total and factor scores, the Beck Depression Inventory (BDI),[34] and 8 summary scales of the Quality of Life Enjoyment and Satisfaction Questionnaire (Q-LES-Q).[35] Cognitive assessments included the Shopping List Task (SLT),[36] which is an effortful serial verbal learning task that was administered according to Buschke's selective reminding procedure.[37] Items were chosen that might commonly be seen on a shopping list in order to increase the face validity of the test. Additional cognitive assessments included the Digit Symbol Substitution test[38] and the MMSE.[26]

Analysis of safety was based on reported treatment-related adverse events. The investigators classified adverse events as "due to study drug," "due to uncertain cause," or unrelated. Treatment-related laboratory abnormalities, intercurrent illnesses, vital signs, body weight, and electrocardiogram data were also analyzed. The final safety and efficacy evaluations were made at week 12 or whenever treatment was discontinued.

Plasma samples for pharmacokinetic analysis were drawn at day 1 of washout and at the end of weeks 1, 3, 4, 6, 8, 10, and 12 or at the time of discontinuation if it occurred prior to week 12. Patients were instructed not to take their medication on days when blood was drawn until after the specimen was obtained.

Compliance was assessed by capsule counts of returned medication and by questioning patients at every visit about study medication consumed. Any patient who missed greater than 25% of the prescribed medication on 2 consecutive visits was dropped from the study.

Statistical Method

All efficacy measures were analyzed on the basis of intent-to-treat sample. For continuous measures at baseline, the 2 treatment groups were examined for comparability with 2-way analysis of variance (ANOVA) models that included treatment group and center main effects as well as the treatment-by-center interaction effect. For categorical measures, Cochran-Mantel-Haenszel general association statistics were used, stratifying on center, to test for comparability at baseline where appropriate.

For all continuous efficacy measures, with the exception of CGI-I and CGI-Efficacy Index measures, the mean score and the mean change score from baseline were computed. For CGI-I and CGI-Efficacy Index measures, only mean scores were computed. The significance of mean changes from baseline was determined with respect to between-group differences using least square means from 2-way analysis of covariance (ANCOVA) models with baseline values as covariates. Where only mean scores were computed, their significance was determined with respect to between-group differences using least square means from 2-way ANOVA models. Treatment group, center, and treatment-by-center interaction terms were included in all models. Gender was added into all the main analyses and found not to influence or qualify the results. Significance of within-group changes from baseline was determined using paired t tests. All analyses were performed for patients available at each study visit and at the last observation carried forward for each patient (endpoint).

Responder status was defined using 2 criteria: (1) a 50% or greater reduction from pretreatment baseline in HAM-D total score and (2) a CGI-I score of 1 or 2 ("very much" or "much" improved). Remitter status was defined as an endpoint HAM-D total score ≤ 10. Treatment comparisons for responder and remitter rates based on the full patient cohort were performed using a Cochran-Mantel-Haenszel general association statistic controlling for center. For comparisons of this type based on subgroups, a logistic model containing treatment group and center effects using the standard logit transformation or Fisher exact test (2-tailed) was used where appropriate. Correlation of plasma level with side effect burden was obtained using a Pearson correlation coefficient. Significance of the correlation was determined using a t test. All significance levels were 2-tailed and set at the .05 level.

The issue of post hoc multiple comparisons was addressed in a variety of ways, and only where convergent evidence pointed toward a valid finding was it reported as such. The SAS Proc Multtest (SAS Institute Inc., Cary, N.C.) was used to perform a permutation test to approximate the minimum p value of all tests, stratified by center. To control for an increased type II error that is the result of the conservative nature of this test, a p value cutoff of .10 was chosen. A second independent method used to determine which findings were "real" was to perform a binomial (sign) test of the overall hypothesis of no differences between sertraline and fluoxetine by looking at the results across the 12 study centers on any given measure; if there was no difference between them, approximately half of the sites would favor sertraline and half would favor fluoxetine. Thirdly, graphical analysis provided another filter to determine which significant findings were indeed meaningful.

Plasma Assays for Sertraline and Fluoxetine

Plasma samples were assayed for sertraline and *N*-desmethylsertraline using a proprietary method[39] or were assayed for fluoxetine and norfluoxetine by a validated, high-performance liquid chromatographic separation method with ultraviolet detection,[40] both developed at Hazelton Laboratories (now part of Covance) in Madison, Wisc. The study sample concentrations were assayed in duplicate and calculated on the basis of comparison to standard curves.

RESULTS

A total of 236 patients entered the double-blind phase of the study (of the 292 who entered the single-blind, placebo run-in phase) in either the sertraline (N = 117) or fluoxetine (N = 119) arm (Table 1). As is consistent with the prevalence of major depressive disorder in the general population, women comprised the majority in both treatment groups, making up 63% of the sertraline group and 51% of the fluoxetine group (p = .053). The subject pool in both groups was overwhelmingly white. The mean ± SD age was 68 ± 5.3 years in the sertraline group and 67 ± 5.9 years in the fluoxetine group. Of the patients in the study, 32% were 70 years of age and over. Results in this subgroup are reported separately.[41] The majority of patients (52.5%) suffered from recurrent moderate or severe depression. Slightly more than half of the patients in both treatment groups (52%) reported prior episodes of depression. The mean baseline 24-item HAM-D scores were 25.1 ± 4.2 and 25.0 ± 4.7 for the sertraline and fluoxetine groups, respectively (range, 18–38). No significant differences were observed between treatment groups on other demographic variables at baseline (see Table 1).

One hundred sixteen (99.1%) and 118 patients (99.2%) who took at least one dose of study drug were evaluable for safety in the sertraline and fluoxetine arms, respectively. Completion rates were also nearly identical for both groups. Eighty patients (68.4%) taking sertraline and 80 patients (67.2%) taking fluoxetine completed all 12 weeks of the study. The percentage of randomized patients who withdrew owing to adverse experiences was 18.8% in the sertraline arm versus 24.4% in the fluoxetine arm (NS). Few patients in either group withdrew owing to lack of efficacy: 2.6% of discontinuations in the sertraline group and 4.2% in the fluoxetine group (NS).

The percentage of patients whose dose of study medication was titrated upward was similar in both groups. Among 116 sertraline patients, 66 (56.9%) had a final dose of 50 mg and 50 (43.1%) had a final dose of 100 mg. Among 118 fluoxetine patients, 64 (54.2%) had a final dose of 20 mg and 54 (45.8%) had a final dose of 40 mg. In both treatment groups, about half of the patients who were classified as responders (i.e., at least a 50% reduction in baseline HAM-D score) received the higher doses (50.0% for sertraline and 51.6% for fluoxetine).

Overall Efficacy

Overall, significant improvements in both primary and secondary efficacy measures were observed, indicating that both sertraline and fluoxetine were effective in the relief of depressive symptoms. Significant improvement

Table 1. Clinical and Demographic Characteristics of Study Patients[a]

Characteristic	Sertraline	Fluoxetine	p Value*
Patients randomized, N	117	119	NS
Sex, N (%)			
Men	43 (36.8)	58 (48.7)	NS
Women	74 (63.2)	61 (51.3)	
Age, y, mean ± SD	68 ± 5.3	67 ± 5.9	NS
Race, N (%)			
White	112 (95.7)	119 (100)	NS
Black	4 (3.4)	0 (0)	
Asian American	0 (0)	0 (0)	
Other	1 (0.9)	0 (0)	
Major depression diagnosis, N (%)[b]			
Single episode	56 (47.9)	55 (46.2)	NS
Recurrent episodes	61 (52.1)	63 (52.9)	
Duration since episode was first diagnosed, wk, mean ± SD	437 ± 599.4	513 ± 720.3	
DSM-III-R severity, N (%)			
Mild	9 (7.7)	7 (5.9)	NS
Moderate	88 (75.2)	81 (68.1)	
Severe (nonpsychotic)	20 (17.1)	31 (26.1)	
Severe (psychotic)	0 (0)	0 (0)	
Baseline 24-item HAM-D score			
Mean ± SD	25.1 ± 4.2	25.0 ± 4.7	NS
Range	18–35	18–38	
Baseline CGI-I score			
Mean ± SD	4.0 ± 0.6	3.9 ± 0.4	NS
Range	3–6	3–5	

[a]Abbreviations: CGI-I = Clinical Global Impressions-Improvement scale, HAM-D = Hamilton Rating Scale for Depression, NS = not significant.
[b]Data missing for 1 patient.
*p Value is based on 2-way analysis of variance or Cochran-Mantel-Haenszel test.

Figure 1. Change From Baseline in Hamilton Rating Scale for Depression Total Scores

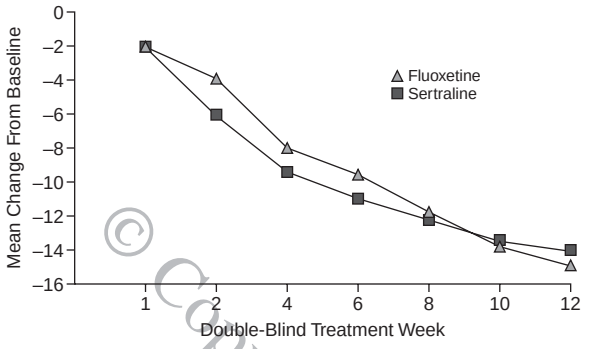
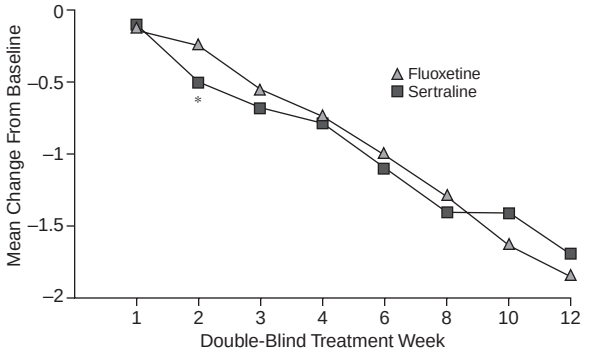


Figure 2. Change From Baseline in Clinical Global Impressions-Severity of Illness Scale Scores



*p < .01 vs. fluoxetine.

(p < .05) from baseline in primary efficacy measures was evident from week 1 onward for both agents.

Primary Efficacy Measures

HAM-D. Significant (p < .01) improvement from baseline in the adjusted mean change in total HAM-D score was seen for both agents by endpoint, with an identical mean change score of –11.3 (Figure 1). The sertraline-treated group showed a trend (p = .089) for a greater decrease in HAM-D score (–6.2 vs. –3.9) by week 2 after making a permutation-based adjustment for multiple comparisons, although this advantage was not sustained. All 7 HAM-D factors (retardation, cognitive disturbance, etc.) tended to favor sertraline over fluoxetine at week 2.

Overall, responder status (≥ 50% reduction from baseline in HAM-D total score) was achieved by 73% of patients treated with sertraline and 71% of patients treated with fluoxetine (NS). When the subgroup of patients suffering from high-severity depression (HAM-D score > 25) was analyzed separately, sertraline-treated patients (N = 47) were found to have a significantly (p = .034) higher proportion of early responders by week 2 (8 patients [17%] were classified as responders) compared with the fluoxetine-treated subgroup (N = 52, with 1 [2%] responding). This was consistent with the difference between the treatment groups in the week-2 HAM-D change score data cited above. This statistically significant early response advantage in favor of sertraline continued through week 4 in this subgroup, with 18 (39%) of the sertraline-treated group considered responders versus 9 (17%) of the fluoxetine-treated group (p = .014). From week 5 on, no significant difference was found between the 2 drugs in the treatment of the high-severity subgroup, with week-12 responder rates for completers of 81% (N = 38) for sertraline-treated patients and 76% (N = 40) for fluoxetine-treated patients. Taking into account all patients in this subgroup who were randomized, 64% of severely depressed sertraline-treated patients were classified as responders compared with 59% of fluoxetine-treated severely-depressed patients (NS).

By week 12, 60% of completers in both the sertraline and fluoxetine groups were classified as remitters (HAM-D score < 10), whereas among the total intent-to-treat sample, 45% of the sertraline-treated group and 46% of the fluoxetine-treated group achieved remitter status. Analysis of subgroups such as those with moderate or severe depression, those whose illnesses had lasted greater than 2 years, and age at onset did not reveal differential efficacy in achieving remission.

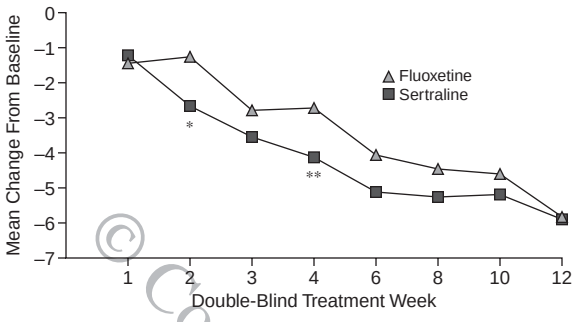
CGI. Results on the CGI were generally similar to those seen on the HAM-D scale. Both treatment groups showed a significant improvement from baseline on all 3 CGI scales (CGI-S, CGI-I, and CGI-Efficacy Index) at all visits. As with the HAM-D, significant (p < .01) between-treatment differences were seen at week 2 in both the CGI-S (p < .01; Figure 2) and CGI-I (p = .011) ratings, with greater improvement in the sertraline-treated group. A treatment responder was also characterized by having a CGI-I score of 1 or 2. Using this definition, the sertraline-treated group showed a significantly (p = .021) higher proportion of responders by week 2 (21%) than the fluoxetine group (10%). An examination of response rates by gender utilizing both the CGI and the HAM-D did not reveal significant evidence of treatment by gender interactions.

Secondary Efficacy Measures: Observer Rated

MADRS. Significant improvement was seen in both treatment groups beginning with week 2. Other than week 2 (p = .037 for sertraline), no statistically significant between-group differences in MADRS scores were observed at any subsequent timepoints.

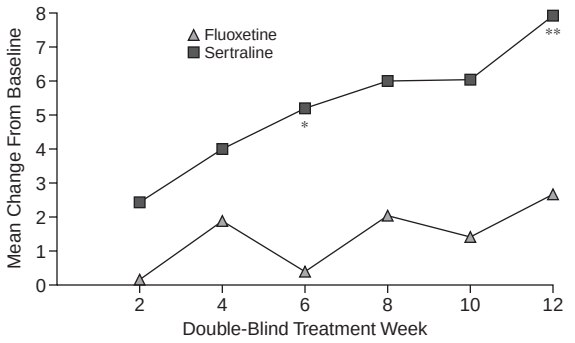
HAM-A. Highly significant (p < .01) improvement from baseline in HAM-A scores was seen for both the sertraline and the fluoxetine groups; no statistically significant between-treatment group differences were seen at any timepoints.

Figure 3. Change From Baseline in Beck Depression Inventory Scores



*p = .019 vs. fluoxetine. **p = .042 vs. fluoxetine.

Figure 4. Change From Baseline in Number of Items Completed on the Digit Symbol Substitution Test



*p = .019 vs. fluoxetine. **p = .037 vs. fluoxetine.

Secondary Efficacy Measures: Patient Rated

BDI. A significant improvement from baseline within both treatment groups was noted for all visits beginning with the first posttreatment week (Figure 3). Consistent with the clinician-rated data, there was a significantly greater magnitude of improvement in the sertraline-treated group at weeks 2 (p = .019) and 4 (p = .042). From week 6 on, no statistically significant between-treatment group differences were observed on the BDI.

POMS. Baseline scores were not significantly different between treatment groups. A significant improvement from baseline in total POMS score was evident in both the sertraline and fluoxetine groups at all visits. Examination of individual POMS subscales revealed some treatment-related differences. The depression-dejection subscale showed a significantly greater improvement in the sertraline-treated group compared with the fluoxetine-treated group at both week 2 (–9.9 vs. –6.6, p = .03) and week 4 (–13.9 vs. –10.0, p = .016). This difference continued to be manifest through week 12. The vigor subscale also showed a significantly greater improvement in the sertraline group at week 4 compared with the fluoxetine group (3.1 vs. 1.2, p = .029).

Q-LES-Q. At week 12 and at endpoint, a highly significant (p < .01) change from baseline within both the sertraline and the fluoxetine groups was noted for the total Q-LES-Q score, as well as for scores on specific Q-LES-Q domains. No significant treatment-related differences were seen. To make sure that this was a direct treatment effect, the effects of depression severity on Q-LES-Q at endpoint were ascertained by including the HAM-D score as a covariate in the endpoint Q-LES-Q analyses. Similarly, the week-12 HAM-D score was included as a covariate in the week-12 Q-LES-Q analyses. The findings remained significant for both drugs.

Cognitive Measures

SLT. The number of items recalled on the SLT was greater at all treatment visits in the sertraline-treated group than in the fluoxetine group. The change from baseline score for items recalled was significantly greater (p = .022) at week 6 on sertraline than fluoxetine treatment with a strong trend toward a significant difference at week 8 (p = .061). Data from week 8 also showed strong trends toward a treatment difference favoring sertraline in the change in the number of items recalled from absolute long-term storage (p = .057) and in long-term storage (p = .075) change from baseline. The size of the learned list was greater at all measurements in the sertraline-treated group and was significantly (p = .051) greater at week 8. No significant treatment-related difference was seen in intrusion error change scores.

Digit Symbol Substitution test. Baseline values for the number of items correctly coded were not significantly different between the treatment groups. Significant improvement (p < .01) from baseline on number correctly coded (4.0) was observed by week 4 and thereafter in the sertraline group, whereas only at endpoint (includes all randomized subjects) was significant improvement (p < .05) observed in the fluoxetine group (3.4). Significant between-treatment effects began by week 6 (p = .019), with a strong trend at weeks 8 (p = .078) and 10 (p = .068), and a significant (p = .037) difference at week 12, all in the direction of a greater improvement in performance on this test in the sertraline-treated group (Figure 4). The magnitude of the differences in treatment group performance was substantial. For example, the mean change for the sertraline group was 5.2 versus 0.4 for fluoxetine at week 6 (p = .019) and 7.9 versus 2.7 at week 12 (p = .037).

Vital Signs

Examination of supine and standing blood pressure (systolic and diastolic) and pulse revealed no clinically significant overall treatment-related changes. However, weight showed a modest but statistically significantly (p = .018) greater decline by endpoint in the fluoxetine-treated group (–3.2 lb) than in the sertraline-treated group (–1.7 lb).

Table 2. Summary of Treatment-Related Adverse Events[a] During Double-Blind Therapy[b]

Patients	Sertraline (N = 116)	Fluoxetine (N = 118)	p Value[c]
Patients with adverse events, N (%)	102 (87.9)	105 (89.0)	.840
Patients with severe adverse events, N (%)	21 (18.1)	14 (11.9)	.203
Patients withdrawn owing to adverse events, N (%)	20 (17.2)	25 (21.2)	.508
Patients requiring dose reductions due to adverse events, N (%)	12 (10.3)	6 (5.1)	.148

[a]Based on World Health Organization dictionary term.[42]
[b]Received at least one dose of double-blind therapy.
[c]Based on Fisher exact test (2-tailed).

Adverse Experiences

Both sertraline and fluoxetine were generally well tolerated. Twenty (17.2%) of the sertraline-treated patients and 25 (21.2%) of the fluoxetine-treated patients withdrew from treatment owing to adverse events (Table 2). The frequency of adverse event reports was consistent with previous experience with these agents.

Headache was the most common central nervous system treatment-related adverse event, affecting 33.6% of sertraline-treated patients and 31.4% of fluoxetine-treated patients. Dizziness, an important consideration in this population, was relatively uncommon, with 7.8% and 10.2% of patients reporting this effect in the sertraline and fluoxetine groups, respectively. The most common autonomic symptom was dry mouth, reported in 15.5% and 7.6% of patients treated with sertraline and fluoxetine, respectively. Nausea and diarrhea were the most common gastrointestinal complaints, with nausea reported in 14.7% of patients taking sertraline versus 18.6% taking fluoxetine; diarrhea affected 22.4% and 16.1% of patients treated with sertraline or fluoxetine, respectively. The vast majority of these reports were mild to moderate in nature, and no statistically significant between-treatment group differences were found.

Insomnia was the most frequently reported psychiatric disturbance, occurring in 13.7% of sertraline-treated patients and 14.4% of those treated with fluoxetine (NS). The peak occurrence for sertraline was between weeks 1 and 2, whereas for fluoxetine-treated patients, weeks 3 through 6 produced the greatest number of reports. Almost as common was anxiety, which was reported by 14.6% of patients in the sertraline group and 12.7% in the fluoxetine group. The occurrence of anxiety reported as an adverse event tended to peak between weeks 3 and 4 for both groups. Appetite was decreased in 9.5% and 11.9% of patients in the sertraline and fluoxetine groups, respectively. The most frequent adverse experiences that resulted in study withdrawal were anxiety, somnolence, and insomnia.

Plasma Drug Concentrations

The mean ± SD trough plasma steady-state sertraline concentration attained by the end of the first week of treatment in patients receiving daily doses of 50 mg was 20.3 ± 1.2 ng/mL. Thereafter, mean steady-state trough concentrations were maintained between 19.7 and 21.6 ng/mL. At sertraline, 50 mg/day, the mean steady-state *N*-desmethylsertraline plasma concentration was 43.9 ± 20.5 ng/mL by the end of the fourth week and was maintained between 42.4 and 45.4 ng/mL.

Patients receiving the 20-mg dose of fluoxetine had mean steady-state fluoxetine concentrations of 91.6 ± 43.9 ng/mL (approximately 3 times the week 1 mean levels), but these were not attained until the eighth week of dosing. The norfluoxetine mean concentrations of 138.9 ± 55.2 ng/mL (approximately 3.5 times the week-1 levels) were also attained during this same time period. Statistically, when dose-normalized mean trough concentrations were compared, steady state was not achieved by fluoxetine and norfluoxetine until weeks 8 and 12, respectively.

In sertraline-treated patients whose dose was doubled from 50 mg to 100 mg at week 4, mean steady-state trough concentrations were achieved by week 6. Mean trough plasma concentrations at week 6 were 46.7 ± 23.9 ng/mL for sertraline and 84.3 ± 37.1 ng/mL for *N*-desmethylsertraline. Thereafter, steady-state concentrations from 44.7 to 48.6 ng/mL for sertraline and from 90.3 to 91.2 ng/mL for *N*-desmethylsertraline were observed from weeks 8 through 12. Dose proportionality occurred with both sertraline and its metabolite.

In the fluoxetine-treated patients at the 40-mg dose, mean plasma concentrations were generally dose proportional, ranging from 183.1 to 243.1 ng/mL, but steady state was not achieved until week 12. Mean trough norfluoxetine concentrations from 186.4 to 239.5 ng/mL (weeks 6–12) did not attain steady state by week 12. These levels were not increased proportional to those occurring at the lower (20 mg) dose of fluoxetine.

Attempts to correlate plasma concentrations of either drug and its metabolite to the primary efficacy measure (HAM-D total score) proved unsuccessful. No linear relationship between plasma drug levels at each week and the primary clinical efficacy variable was observed. However, a significant ($p < .05$) negative correlation was observed between side effect burden (sum of all reported adverse experiences weighted by severity) and plasma fluoxetine ($r = –0.28$) and norfluoxetine ($r = –0.23$) levels at week 12. No significant relationship was observed for sertraline.

DISCUSSION

The results of this study demonstrate that both sertraline and fluoxetine were effective in treating major de-

pression in a large group of elderly outpatients. Consistent and clinically significant improvements were noted for both treatment groups on the 24-item HAM-D total score, as well as individual HAM-D factor scores, the MADRS, and the CGI. Overall, 73% of patients treated with sertraline were classified as responders at week 12, whereas 71% of patients treated with fluoxetine were classified as responders. This overall response rate is higher than was seen in the only placebo-controlled trials in this population.[12] Although subtle differences in the characteristics of the patient population cannot be ruled out, the most likely reason for this higher response rate is the 12-week duration of the study. The majority of previous controlled trials of antidepressants in the elderly have been from 6 to 8 weeks in duration. At 6 weeks in the current study, responder status had been achieved by 50% of patients taking sertraline and 41.5% of patients taking fluoxetine.

It should be noted that patients continued to show improvement in their depressive symptoms throughout the 3 months of the study. This suggests that patience is warranted in the treatment of these older individuals and that full improvement may not manifest before at least 12 weeks of treatment. On the basis of evidence from treatment studies in younger adults, clinicians may be tempted to change treatments if full improvement has not occurred after a relatively short period of time. The results of the current study suggest that this switch of treatment after a short period would be a mistake for the elderly patient suffering from depression.

Time to Response

Although this study was not designed to assess time to response, it appears that sertraline treatment showed a trend toward earlier improvement on several measures (e.g., HAM-D, CGI-S and CGI-I, MADRS, and BDI). This time-to-response effect was somewhat more evident among the more severely depressed subgroup of patients. In this subgroup, a treatment advantage in favor of sertraline continued to be observed through week 4 on some measures. This suggests that any overall earlier antidepressant response observed for sertraline was largely accounted for by patients with higher depression severity.

Even though there is both clinical[43–45] and pharmacokinetic[46] evidence that appears to support the earlier time-to-response advantage for sertraline over fluoxetine, we must consider this finding in the present study to be preliminary. Since the current study had no a priori time-to-response hypothesis, these suggestive findings await confirmation from further research.

Efficacy on Secondary Outcome Measures

Despite similar overall antidepressant response on the primary outcome measures, treatment with sertraline yielded statistically significantly greater improvement than did fluoxetine on a number of secondary efficacy measures that are likely to be clinically important for functioning in the elderly.

Sertraline had significantly greater positive effects on verbal learning and recall as measured by the SLT, as well as on visual tracking, coding, and motor performance as measured by the Digit Symbol Substitution test. These results suggest that sertraline may be more potent in improving the cognitive symptoms often associated with depression in this age group. The reasons for these effects are unclear. Differential cognitive improvement does not appear to be secondary to differential improvement in depression, since the degree of improvement in both groups was similar by study endpoint. It is possible that the slightly more rapid clinical improvement observed with sertraline may have partially contributed to its more favorable effect on cognitive performance. Neither sertraline nor fluoxetine appears to have clinically significant anticholinergic effects, so it is unlikely that a differential central effect on this receptor system was contributory. Since the differences on this task were mostly seen on measures reflecting recall from long-term memory, it may be that sertraline either directly or indirectly through improving depression has a greater impact on consolidation of memory or in improving memory retrieval. Performance on the Digit Symbol Substitution test has been shown to decline with age and to be sensitive to the effects of centrally acting medication and sedation.

The long-term clinical implications of the differential acute effects on cognitive function observed in the current study are uncertain. The topic deserves further research to confirm the result, elucidate what the underlying mechanism might be, and assess the clinical impact on long-term functioning in the elderly.

Dosing and Adverse Events

Both sertraline and fluoxetine were well tolerated by the elderly patients in the study. In general, adverse side effects were similar, both qualitatively and quantitatively, to those observed in other studies of SSRIs in elderly patients.[18,47,48] Clinical experience suggests that side effect rates and treatment dropout rates can be reduced if a gradual titration from low starting doses is employed.[10] Previous studies[49] have suggested that there is a roughly linear relationship between SSRI dosage and side effect rates. The small negative correlation found between fluoxetine and norfluoxetine plasma levels at week 12 and side effect rates is of interest in this regard.

Interestingly, the distribution of final doses suggests that there was no significant difference between the effectiveness of the lower and higher doses for either drug. What is unclear is whether those patients whose dose was raised at week 4 would have improved if simply given more time; the study design does not allow this question to be answered. However, it is possible to conclude that 50 mg of sertraline or 20 mg of fluoxetine is an effective

dose for many patients, and conversely there was no evidence that the lower dose of sertraline was clinically ineffective in any greater proportion of patients than 20 mg of fluoxetine. Rapid dose titration to higher doses may therefore not be justified.

Finally, neither study drug was found to have clinically significant adverse effects on laboratory values or vital signs, with the possible exception of weight loss being greater in the fluoxetine-treated group. Clinically significant weight loss with fluoxetine use in elderly patients has been previously described.[50]

Study Limitations

The most notable limitation of the current study is the lack of a placebo control group. Although no placebo control group was used, the placebo run-in period at the beginning of the study presumably eliminated some early placebo responders. Furthermore, pattern analysis suggests that the clinical effect of placebo is characterized not only by early response, but by response that is not sustained.[51] The 12-week duration of the study, approximately twice the length of most acute treatment studies in younger populations, should also help to minimize non–drug-related treatment effects. Nonetheless, without a placebo group, the proportion of responses due solely to the effect of the medication remains unclear.

CONCLUSION

The efficacy and safety profiles of the SSRIs make them a preferred drug class for treating depression in the elderly. The results of this study indicate that sertraline and fluoxetine have comparable overall efficacy over a 3-month period in the treatment of even long-standing episodes of depression in elderly outpatients. The fact that patients in both treatment groups continued to improve throughout the 3-month treatment period suggests that patience is warranted in the treatment of these individuals, for whom full clinical response may take longer than with younger patients. There was evidence from this direct comparison that sertraline may produce a greater improvement in cognitive performance. Whether this could be due to time-to-response effects or clinical effects in more severely depressed patients remains to be further investigated.

Drug names: amitriptyline (Elavil and others), fluoxetine (Prozac), sertraline (Zoloft), temazepam (Restoril and others).

REFERENCES

1. Blazer D, Williams CD. Epidemiology of dysphoria and depression in an elderly population. Am J Psychiatry 1980;139:439–444
2. Allen A, Blazer DG. Mood disorders. In: Sadavoy J, Lazarus LW, Jarvik LF, eds. Comprehensive Review of Geriatric Psychiatry. Washington, DC: American Psychiatric Press; 1991:337–352
3. Katz IR, Curlik S, Nemetz P. Functional psychiatric disorders in the elderly. In: Lazarus L, Foster J, eds. Essentials of Geriatric Psychiatry: A Guide for Health Professionals. New York, NY: Springer; 1988:113–137
4. Borson S, Barnes RA, Kukull WA. Symptomatic depression in elderly medical outpatients, I: prevalence, demography, and health service utilization. J Am Geriatr Soc 1986;34:341–347
5. Berkman LF, Berkman CS, Kasl S. Depressive symptoms in relation to physical health and functioning in the elderly. Am J Epidemiol 1986;124: 372–388
6. Wells KB, Stewart A, Hays RD, et al. The functioning and well-being of depressed patients: results from the Medical Outcomes Study. JAMA 1989;262:914–919
7. Koenig HG, Meador KG, Cohen HG, et al. Detection and treatment of major depression in older medically ill hospitalized patients. Int J Psychiatry Med 1988;18:17–31
8. Hotopf M, Hardy R, Lewis G. Discontinuation rates of SSRIs and tricyclic antidepressants: a meta-analysis and investigation of heterogeneity. Br J Psychiatry 1997;170:120–127
9. Newhouse PA. Use of serotonin selective reuptake inhibitors in geriatric depression. J Clin Psychiatry 1996;57:12–22
10. Dunbar GV, Cohn JB, Fabre LF. A comparison of paroxetine, imipramine, and placebo in depressed outpatients. Br J Psychiatry 1991;159:394–398
11. Tollefson GD, Bosomworth JC, Heiligenstein JH, et al, for the Fluoxetine Collaborative Study Group. A double-blind, placebo-controlled clinical trial of fluoxetine in geriatric patients with major depression. Int Psychogeriatr 1995;7:89–104
12. Orengo CA, Kunik ME, Molinari V, et al. The use and tolerability of fluoxetine in geropsychiatric inpatients. J Clin Psychiatry 1996;57:12–16
13. Brymer C, Winograd CH. Fluoxetine in elderly patients: is there cause for concern? J Am Geriatr 1992;40:902–905
14. Muir GMM, Gaspic M. Getting a grip on agitation-psychotropic medications. Geriatrics 1994;10:17–22
15. Blazer DG. The treatment of late-life depression. Clin Geriatr 1996;4: 16–24
16. Van Harten J. Clinical pharmacokinetics of selective serotonin reuptake inhibitors. Clin Pharmacokinet 1993;24:203–220
17. Cohn CK, Shrivastava R, Mendels J, et al. Double-blind, multicenter comparison of sertraline and amitriptyline in elderly depressed patients. J Clin Psychiatry 1990;51(12, suppl B):28–33
18. Preskorn SH. Recent pharmacokinetic advances in antidepressant therapy of the elderly. Am J Med 1993;94:2S–12S
19. Weingartner H, Cohen RM, Murphy DL, et al. Cognitive processes in depression. Arch Gen Psychiatry 1981;38:42–47
20. King DA, Caine ED, Conwell Y, et al. The neuropsychology of depression in the elderly: a comparative study of normal aging and Alzheimer's disease. J Neuropsychiatry Clin Neurosci 1991;3:163–168
21. King DA, Caine ED, Cox C. Predicting severity of depression in the elderly at 6 month follow-up: a neuropsychological study. J Neuropsychiatry Clin Neurosci 1991;3:64–67
22. Mangione CM, Seddon JM, Cook EF, et al. Correlates of cognitive function scores in elderly outpatients. J Am Geriatr Soc 1993;41:491–497
23. Young RC, Mattis S, Alexopoulos GS, et al. Verbal memory and plasma drug concentrations in elderly depressives treated with nortriptyline. Psychopharmacol Bull 1991;27:291–294
24. Gurland BJ. The range of quality of life: relevance to the treatment of depression in the elderly. In: Schneider LS, Reynolds CF III, Lebowitz BD, et al., eds. Diagnosis and Treatment of Depression in Late Life. Washington, DC: American Psychiatric Press; 1994
25. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, Third Edition, Revised. Washington, DC: American Psychiatric Association; 1987
26. Folstein MF, Folstein SE, McHugh PR. Mini-Mental State: a practical method for grading the cognitive state of patients for the clinician. J Psychiatry Rev 1975;12:189–198
27. Hamilton M. A rating scale for depression. J Neurol Neurosurg Psychiatry 1960;23:51–56
28. Hamilton M. Development of a rating scale for primary depressive illness. Br J Soc Clin Psychol 1967;6:278–279
29. Bech P. Psychometric developments of the Hamilton scales: the spectrum of depression, dysthymia, and anxiety. Psychopharmacol Ser 1990;9: 72–79
30. Guy W. ECDEU Assessment Manual for Psychopharmacology. US Dept Health, Education and Welfare publication (ADM) 76-338. Rockville, Md: National Institute of Mental Health; 1976:218–222

31. Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. Br J Psychiatry 1979;134:383–389
32. Hamilton M. The assessment of anxiety states by rating. Br J Med Psychol 1959;32:50–55
33. McNair DM, Lorr M, Droppleman LF. Manual for the Profile of Mood States. San Diego, Calif: Educational & Industrial Testing Service; 1971
34. Beck AT, Ward CH, Mendelson M, et al. An inventory for measuring depression. Arch Gen Psychiatry 1961;4:561–571
35. Endicott J, Nee J, Harrison W, et al. Quality of Life Enjoyment and Satisfaction Questionnaire: a new measure. Psychopharmacol Bull 1993;29: 321–326
36. McCarthy M, Ferris SH, Clark E, et al. Acquisition and retention of categorized material in normal aging and senile dementia. Exp Aging Res 1981;7:127–135
37. Buschke H. Short-term retention, learning, and retrieval from long-term memory. In: Deutsch DJ, ed. Short-Term Memory. New York, NY: Academic Press; 1975
38. Wechsler D. Wechsler Adult Intelligence Scale-Revised. New York, NY: Psychological Corp; 1981
39. Tremaine LM, Joerg EA. Automated gas chromatographic-electron capture assay for the selective serotonin reuptake blocker, sertaline. J Chromatogr 1989;496:423–429
40. Wong SH, Dellafera SS, Fernandes R, et al. Determination of fluoxetine and norfluoxetine by high-performance liquid chromatography. J Chromatogr 1990;499:601–608
41. Finkel S, Nakra BR, Richter E, et al. Comparative efficacy of sertraline versus fluoxetine in elderly patients over age 70 with major depression. Am J Geriatr Psychiatry. In press
42. WHO Collaborating Center for International Drug Monitoring. International Monitoring of Adverse Reactions to Drugs: Adverse Reaction Terminology. Sweden: the World Health Organization; 1989
43. Patris M, Bouchard JM, Bougerol T, et al. Citalopram versus fluoxetine: a double-blind controlled, multicentre, phase III clinical trial in patients with unipolar major depression in general practice. Int Clin Psychopharmacol 1996;11:129–136
44. Bougerol T, Scotto JC, Patris M, et al. Citalopram and fluoxetine in major depression: comparison of two clinical trials in a psychiatrist setting and in general practice. Clin Drug Invest 1997;14:77–89
45. Geretsegger C, Bohmer F, Ludwig M. Paroxetine in the elderly depressed patient: randomized comparison with fluoxetine of efficacy, cognitive, and behavioral effects. Int Clin Psychopharmacol 1994;9:25–29
46. Preskorn SH. Clinically relevant pharmacology of selective serotonin reuptake inhibitors: an overview with emphasis on pharmacokinetics and effects on oxidative metabolism. Clin Pharmacokinet 1997;32(suppl 1): 1–21
47. Salzman C, Schneider LS, Alexopolous GS. Pharmacological treatment of depression in late life. In: Bloom FE, Kupfer DJ, eds. Psychopharmacology: The Fourth Generation of Progress. New York, NY: Raven Press; 1995:1471–1477
48. Dunner DL, Cohn JB, Walshe T III, et al. Two combined, multicenter double-blind studies of paroxetine and doxepin in geriatric patients with major depression. J Clin Psychiatry 1992;53:57–60
49. Schone W, Ludwig M. A double-blind study of paroxetine compared with fluoxetine in geriatric patients with major depression. J Clin Psychopharmacol 1993;13(6 suppl 2):34S–39S
50. Brymer C, Winograd CH. Fluoxetine in elderly patients: is there cause for concern? J Am Geriatr 1992;40:902–905
51. Quitkin FM, Rabkin JD, Markowitz JM, et al. Use of pattern analysis to identify true drug response: a replication. Arch Gen Psychiatry 1987;44: 259–264