

A Double-Blind Study of Paroxetine, Fluoxetine, and Placebo in Outpatients with Major Depression

Maurizio Fava, M.D.,^{1,6} Jay D. Amsterdam, M.D.,² Joseph A. Deltito, M.D.,³ Carl Salzman, M.D.,⁴ Michael Schwaller, M.D.,⁵ and David L. Dunner, M.D.⁵

We report results from a multicenter, placebo-controlled, randomized, double-blind comparison of the efficacy and tolerability of paroxetine and fluoxetine in outpatients with major depression. Across five U.S. sites, 128 outpatients (mean age: 41.3 ± 12.6 ; 63 men and 65 women) with moderate to moderately severe major depression without a history of mania or hypomania were recruited between 1993 and 1994. All 128 patients completed a 1-week placebo washout period, and were then randomized to 12 weeks of double-blind treatment with paroxetine up to 50 mg/day ($n = 55$), fluoxetine up to 80 mg/day ($n = 54$), or placebo ($n = 19$). Subjects were evaluated weekly for the first 4 weeks, then at weeks 6, 9, and 12 with the 21-item HAMD and the Covi Anxiety Scale. There were no significant differences among the three treatment groups in baseline and endpoint depression and anxiety severity, as well as in the degree of depression and anxiety improvement. There were no statistically significant differences in rates or mean numbers of adverse events between paroxetine-treated patients and fluoxetine-treated patients. In summary, our results, although limited by the lack of a significant difference from placebo in treatment outcome, suggest that the SSRIs paroxetine and fluoxetine have comparable antidepressant and antianxiety efficacies among depressed outpatients, as well as similar safety and tolerability profiles.

KEY WORDS: paroxetine; fluoxetine; placebo; depression; anxiety.

INTRODUCTION

Selective serotonin reuptake inhibitors have become the most widely prescribed antidepressants in the U.S. Given the relative heterogeneity in chemical structure, norepinephrine and dopamine uptake in-

hibiting properties, and pharmacokinetic characteristics of these antidepressants, one may wonder whether such variability translates into meaningful differences in antidepressant efficacy. In addition to having an indication for the treatment of depression, SSRIs are commonly used in the treatment of anxiety disorders such as panic disorder, social phobia, and obsessive-compulsive disorder (1). However, it is unclear whether all SSRIs have comparable antianxiety effects. Finally, as Finley (1) points out in his review of the literature, the assessment of possible important differences in the prevalence of adverse events awaits the completion of adequate comparative trials.

Unfortunately, very few parallel comparisons of SSRIs have been conducted thus far, often with rather small sample sizes or among rather atypical populations. One multicenter study (2) carried out in Austria and Germany and supported by Smith-

¹Depression Clinical and Research Program, Massachusetts General Hospital, Boston, Massachusetts 02114.

²Depression Research Unit, University of Pennsylvania Medical Center, Philadelphia, Pennsylvania.

³Anxiety and Mood Disorder Program, Cornell University Medical Center, New York Hospital, Westchester Division, White Plains, New York.

⁴Massachusetts Mental Health Center, Boston, Massachusetts.

⁵Center for Anxiety and Depression, University of Washington, Seattle, Washington.

⁶To whom correspondence should be addressed at Depression Clinical and Research Program, Massachusetts General Hospital, ACC 812, 15 Parkman Street, Boston, MA 02114.

Kline Beecham Pharmaceuticals randomized 106 geriatric (age ≥ 65 years) depressed outpatients to six weeks of double-blind treatment with either paroxetine (20–40 mg/day; $n = 54$) or fluoxetine (20 to 60 mg/day; $n = 52$). There was no significant difference in Hamilton Rating Scale for Depression (HAM-D) (3) and Montgomery Asberg Depression Rating Scale (MADRS) (4) change scores between the two treatments at endpoint. Although paroxetine-treated patients showed a statistically significantly higher proportion of responders than fluoxetine-treated patients according to two of four definitions of response, the overall response rates in both groups were quite low, suggesting the possibility that this sample of patients with geriatric depression might have been relatively atypical. In addition, there were no significant differences between the two treatment groups in the overall incidence of adverse events or in the incidence of any specific adverse event.

Another European multicenter, six-week study (5) randomized 178 inpatients with major depression to treatment with either paroxetine (20 mg/day) or fluoxetine (20 mg/day). This study, also supported by SmithKline Beecham Pharmaceuticals, failed to show any significant difference in antidepressant and anti-anxiety efficacy between the two treatments. The incidence of anxiety as an adverse event was 13.5% among paroxetine-treated patients and 11.5% among fluoxetine-treated patients (5). No statistically significant differences in distribution and rates of adverse events were noted between the two groups, with the exception of a significantly higher rate of weight loss on fluoxetine (5).

A relatively smaller European multicenter, 6-week study (6) involved the randomization of 78 evaluable patients to double-blind treatment with either paroxetine or fluoxetine and found no significant differences in efficacy at endpoint between the two treatments. However, patients on paroxetine reported fewer adverse events than those on fluoxetine, with nausea and vomiting being the most commonly reported adverse events in both groups. To our knowledge, no U.S. comparative studies of fluoxetine and paroxetine have as yet been published.

We now report the results derived from two U.S. multicenter studies comparing the efficacy of fluoxetine, paroxetine, and placebo. The results of these two studies, supported by SmithKline Beecham Pharmaceuticals (protocols # 115 and 128) and carried out in 1993–1994, pertain only to five sites out of a higher total of sites participating in the same protocols (the

investigators MF, JAD, and CS participated in protocol 115, and JDA, DLD, MS participated in protocol 128). We decided to pool the data together from these five academic sites as the two protocols shared the same design and methods.

SUBJECTS AND METHODS

Across five U.S. sites, 128 outpatients (mean age: 41.3 ± 12.6 ; 63 men and 65 women) with moderate to moderately severe major depression without a history of mania or hypomania were recruited between 1993 and 1994. For each subject, a Raskin Depression Scale (7) score of at least 8, and larger in value than the Covi Anxiety Scale (8) score was required, as well as a Hamilton Rating Scale for Depression (HAM-D) (3) score of 18 or greater on the first 17 items.

Excluded were patients with the following primary psychiatric diagnoses: schizophrenia, adjustment disorder, bipolar disorder, panic disorder, social phobia, obsessive-compulsive disorder, psychotic depression, and atypical depression. Other exclusions included patients whose HAM-D total score fell by 25% or more between the screen and baseline evaluations, patients with serious concomitant medical illness or significant abnormal laboratory values at screening or baseline, patients with history of seizure disorders (except febrile seizures in childhood), patients judged to have a high suicidal risk by the investigator, patients with a recent history (within 6 months) of alcohol or drug abuse, patients who required therapy with other psychotropic drug within 14 days of baseline (with the exception of chloral hydrate for sleep), patients who received electroconvulsive therapy within 3 months of baseline, were hypersensitive to fluoxetine during a prior treatment period, or used any investigational drug within 30 days of baseline, and patients previously treated with paroxetine. Women of childbearing age were excluded if they had positive pregnancy tests or did not practice medically accepted means of birth control.

Screening included pregnancy testing for women, body weight determination, medical history, physical examination, 12-lead ECG, vital signs (including screen for postural hypotension), laboratory testing, including: hematology (hematocrit, hemoglobin, WBC with differential, platelet count); blood chemistry (alkaline phosphatase, BUN, creatinine, AST, ALT, total bilirubin, total T3, and total T4); and

Table 1. 21-Item HAM-D and Raskin Anxiety Scores Before and After Treatment Among Treatment Groups^a

	Baseline (BL)		Endpoint (EP) (week 12 or earlier)		BL-EP difference	
	Mean	SD	Mean	SD	Mean	SD
21-item HAM-D						
paroxetine	23.1	3.4	12.1	10.0	11.1	9.4
fluoxetine	23.9	3.8	13.1	10.3	10.8	9.5
placebo	23.7	2.7	12.2	9.0	11.6	8.9
Covi Anxiety Scale						
paroxetine	6.2	1.7	5.0	2.5	1.2	2.7
fluoxetine	6.3	1.7	5.0	2.1	1.2	1.8
placebo	5.8	1.2	4.8	1.7	1.1	2.0

^an = 55, 54, 19 for paroxetine, fluoxetine, placebo groups respectively.

urinalysis (if dipstick was positive for blood or protein, full microscopy was performed). Psychiatric screening included: DSM-III-R multi-axial diagnostic evaluations, 21-item HAM-D, Raskin Depression Scale, Covi Anxiety Scale, and a detailed psychiatric history.

Each patient was adequately informed of the methods, anticipated benefits, and possible hazards of the study and signed a consent form before participating. All 128 patients met all criteria for study eligibility, completed a one week placebo washout period, and were then randomized in a double-blind fashion into one of three treatment groups: paroxetine, fluoxetine and placebo. Of these 128 patients, 55 received paroxetine (43%), 54 patients received fluoxetine (42%), and 19 (15%) were given placebo.

Both groups receiving paroxetine and fluoxetine started treatment with 20 mg daily of drug. If, in the investigators' clinical judgment, a greater dose of drug was needed, investigators were permitted to increase the dose at any time during the trial, as long as dose increases were separated by at least 1 week. The protocol for paroxetine allowed 10 mg increments per week to maximum dose of 50 mg daily. Fluoxetine increases were permitted in 20 mg/day increments to a maximum of 80 mg daily. If adverse events occurred, investigators could decrease the dose at any time during the trial period.

After compilation of baseline data, subjects were seen weekly for the first 4 weeks, then at weeks 6, 9, and 12 of the 12-week treatment period. During each visit, the following observations were performed: 21-item HAM-D, Covi Anxiety Scale, vital signs, adverse effect monitoring, concomitant medication records, study medication records, and Activity/Productivity Assessment. During weeks 3, 6, 9,

and 12 laboratory evaluations were completed. At week 12, or upon termination of the study, patients were given physical examinations. A study termination record assessed patients' reasons for completion/termination of the study. Adverse events were grouped *a priori* into 11 categories: anticholinergic, visual-sensory, central nervous system (CNS), nervousness/agitation, urogenital, insomnia, sedation, cardiovascular, respiratory, and gastrointestinal.

Statistical Analyses

We chose to conduct all analyses with an intent to treat approach, and with a Bonferroni correction for multiple comparisons by dividing the level of significance (0.05) by the number of comparisons. Baseline, endpoint, and baseline-endpoint changes in total 21-item HAM-D scores and Covi Anxiety Scale scores were compared across the three groups with the Kruskal Wallis one-way analysis of variance. The same test was used to compare ages at screen across the three treatment groups. The overall chi-square test was used to compare gender distribution, response rates (defined as the percentage of subjects with a $\geq 50\%$ reduction in total 21-item HAM-D scores) and rates of adverse events across the three treatment groups.

RESULTS

There were no significant differences in age or gender distribution across the three groups. As shown in Table 1, there were also no significant differences among the three treatment groups in de-

Table 2. Rates and Mean Number of Spontaneously Reported Adverse Events Across Treatment Groups^a

	No. (%) reporting AEs	Number of AEs	
		Mean	SD
Anticholinergic AEs			
paroxetine	16 (29%)	0.4	0.6
fluoxetine	10 (19%)	0.3	0.7
placebo	2 (11%)	0.1	0.3
Sexual dysfunction AEs			
paroxetine	14 (25%)**	0.3**	0.5
fluoxetine	4 (7%)	0.1	0.3
placebo	0 (0%)	0	0
Visual-sensory AEs			
paroxetine	5 (9%)	0.1	0.3
fluoxetine	4 (7%)	0.1	0.3
placebo	0 (0%)	0	0
CNS AEs			
paroxetine	33 (60%)	0.9	0.9
fluoxetine	28 (52%)	0.8	1.0
placebo	6 (32%)	0.4	0.7
Nervousness/agitation AEs			
paroxetine	14 (25%)	0.5	1.3
fluoxetine	23 (43%)	0.7	1.1
placebo	5 (26%)	0.3	0.6
Urogenital AEs			
paroxetine	7 (13%)	0.1	0.4
fluoxetine	4 (7%)	0.1	0.4
placebo	3 (16%)	0.2	0.4
Insomnia AEs			
paroxetine	16 (29%)	0.3	0.5
fluoxetine	11 (20%)	0.2	0.5
placebo	2 (11%)	0.1	0.3
Sedation AEs			
paroxetine	19 (35%)	0.4	0.6
fluoxetine	14 (26%)	0.3	0.7
placebo	2 (11%)	0.2	0.5
Cardiovascular AEs			
paroxetine	3 (5%)	0.1	0.3
fluoxetine	6 (11%)	0.1	0.3
placebo	2 (11%)	0.2	0.5
Respiratory AEs			
paroxetine	11 (20%)	0.2	0.5
fluoxetine	7 (13%)	0.1	0.4
placebo	5 (26%)	0.3	0.6
Gastrointestinal AEs			
paroxetine	26 (47%)	0.8	1.0
fluoxetine	26 (48%)	0.8	1.1
placebo	3 (16%)*	0.2*	0.4

^aFor n of each group, see Table 1.

* $p < .05$ (nonsignificant after Bonferroni correction for multiple comparisons).

** $p < .005$ (nonsignificant after Bonferroni correction for multiple comparisons).

pression and anxiety severity (as measured by the 21-item HAM-D and the Covi Anxiety Scale) at both baseline and endpoint (week 12 or earlier). Similarly,

there were also no significant differences among the three treatment groups in the degree of depression and anxiety improvement (as measured by the

change score in 21-item HAM-D and Covi Anxiety Scale), nor in response rates (58% for paroxetine, 57% for fluoxetine, and 53% for placebo).

Of all patients randomized to treatment, 16 (29%) dropped out on paroxetine, 16 (31%) dropped out on fluoxetine, and 4 (21%) dropped out on placebo. Of those patients who dropped out ($n = 36$), 15 (42%) were discontinued because of adverse events: 9 on paroxetine and 6 on fluoxetine. Table 2 reports the rates and mean numbers of spontaneously reported adverse events grouped *a priori* into 11 categories. There were two statistically significant findings concerning the overall distribution of adverse events: the fluoxetine-treated and the paroxetine-treated patients reported significantly more gastrointestinal adverse events than the placebo-treated patients, and the paroxetine-treated patients reported significantly more sexual dysfunction adverse events than the fluoxetine- and placebo-treated patients. However, these two differences became nonsignificant after Bonferroni's correction. Although agitation/nervousness were reported more frequently on fluoxetine than on paroxetine or placebo, this difference failed to reach statistical significance. There were no other statistically significant differences in rates or mean numbers of adverse events among the three treatment groups.

DISCUSSION

Although limited by the lack of a significant difference from placebo in treatment outcome, the results of our study suggest that the SSRIs paroxetine and fluoxetine have comparable antidepressant and antianxiety efficacy in a sample of adult outpatients with major depression. Although both antidepressants showed no significant difference from placebo-treated patients in terms of efficacy, the relatively small number of patients randomized to placebo (15% of the entire sample) makes it difficult to interpret our findings and to determine whether placebo treatment is particularly efficacious in the population enrolled into our study. It is possible that the lack of a drug-placebo difference in efficacy may be also related to the overall mild to moderate level of severity of depression of our sample, as the mean 21-item HAM-D score was 23. Elkin *et al.* (9), in fact, had shown that patients with mild to moderate major depression were less likely to show a drug-placebo

difference than those with more severe depression. On the other hand, our findings of comparable efficacy of both fluoxetine and paroxetine are consistent with those of a European multicenter, 6-week study (5) which randomized 178 inpatients with major depression to treatment with either paroxetine or fluoxetine, and failed to show any significant difference in both antidepressant and antianxiety efficacy between the two treatments.

The safety profile of these two SSRIs also appears to be comparable, with perhaps the exception of sexual dysfunction, which we found to occur more frequently among paroxetine-treated patients (although the difference was not statistically significant after Bonferroni's correction for multiple comparisons). This is also consistent with the findings of a multicenter prospective study by Montejo-Gonzalez *et al.* (10) who found that among 344 patients treated with SSRIs paroxetine was associated with significantly higher rates of delayed orgasm and impotence than fluoxetine, sertraline, and fluvoxamine. The fact that there were no statistically significant differences in adverse events such as anxiety, insomnia, and sedation between paroxetine and fluoxetine is in agreement with the findings of a European study by Tignol (5) which found that the incidence of anxiety was 13.5% among paroxetine-treated patients and 11.5% among fluoxetine-treated patients and that no statistically significant differences in distribution and rates of adverse events were present between these two antidepressants, with the exception of a significantly higher rate of weight loss on fluoxetine.

The limitations of this study are primarily related to the small size of the placebo arm, which complicates the interpretation of our results, and to the fact that these data are derived from two large multicenter studies, thereby representing only a selected group of academic sites. As a result of that, the present sample size does not reflect the originally projected statistical power of the entire multicenter studies, and our results may have been different with a larger sample size. The findings from the two larger multicenter studies, sponsored by SmithKline Beecham Pharmaceuticals, have not been published in the literature as of yet. Finally, adverse events were grouped in this study in categories that we had established prior to data entry and analysis, but that were not included in the original protocols.

In summary, our results, although limited by the lack of a significant difference from placebo in

treatment outcome, suggest that the SSRIs paroxetine and fluoxetine have comparable antidepressant and antianxiety efficacy among depressed outpatients, as well as similar safety and tolerability profile. These data challenge the view that there may be significant differences between these two SSRIs in terms of either anxiolytic or sedating vs. activating effects.

ACKNOWLEDGMENTS

This study was supported by SmithKline Beecham Pharmaceuticals.

REFERENCES

1. Finley PR: Selective serotonin reuptake inhibitors: Pharmacologic profiles and potential therapeutic distinctions. *Ann Pharmacother* 1994; 28:1359-1369
2. Schone W, Ludwig M: A double-blind study of paroxetine compared with fluoxetine in geriatric patients with major depression. *J Clin Psychopharm* 1993; 13 (suppl 2):34S-39S
3. Hamilton M: A rating scale for depression. *J Neurol Neurosurg Psychiatry* 1960; 23:56-62
4. Montgomery SA, Asberg M: A new depression scale designed to be sensitive to change. *Br J Psychiat* 1979; 134:382-389
5. Tignol J: A double-blind, randomized, fluoxetine-controlled, multicenter study of paroxetine in the treatment of depression. *J Clin Psychopharm* 1993; 13 (suppl 2):18S-22S
6. De Wilde J, Spiers R, Mertens C, Bartholome F, Schotte G, Leyman S: A double-blind, comparative, multicenter study comparing paroxetine with fluoxetine in depressed patients. *Acta Psychiatr Scand* 1993; 87:141-145
7. Raskin A, Schulterbrandt J, Reatig N, Rice CE: Factors of psychopathology in interview, word behavior, and self-report ratings of hospitalized depressions. *J Consult Psychol* 1967; 31:270-278
8. Covi L, Lipman R, McNair DM, Czerlinsky T: Symptomatic volunteers in multicenter drug trials. *Progr Neuropsychopharmacol* 1979; 3:521-533
9. Elkin I, Shea MT, Watkins JT, Imber SD, Sotsky SM, Collins JF, Glass DR, Pilkonis PA, Leber WR, Docherty JP, Fiester SJ, Parloff MB: National Institute of Mental Health Treatment of Depression Collaborative Research Program: General effectiveness of treatments. *Arch Gen Psychiat* 1989; 46:971-982
10. Montejo-Gonzalez AL, Llorca G, Izquierdo JA, Ledesma A, Bousono M, Calcedo A, Carrasco JL, Ciudad J, Daniel E, De La Gandara J, Derecho J, Franco M, Gomez MJ, Macias JA, Martin T, Perez V, Sanchez JM, Sanchez S, Vicens E: SSRI-induced sexual dysfunction: Fluoxetine, paroxetine, sertraline, and fluvoxamine in a prospective-multicenter, and descriptive clinical study of 344 patients. *J Sex Marital Ther* 1997; 23(3):176-194