Use of Pattern Analysis to Predict Differential Relapse of Remitted Patients With Major Depression During 1 Year of Treatment With Fluoxetine or Placebo

Jonathan W. Stewart, MD; Frederic M. Quitkin, MD; Patrick J. McGrath, MD; Jay Amsterdam, MD; Maurizio Fava, MD; Jan Fawcett, MD; Frederick Reimherr, MD; Jerrold Rosenbaum, MD; Charles Beasley, MD; Paul Roback, MS

**Background:** Delayed and persistent (“true drug”) improvement characterizes the response to antidepressant medication. Early or nonpersistent (“placebo”) benefit is typical of a placebo response. The prediction was that patients with a true drug response would sustain their benefit best if they continued to receive the drug and that patients with a placebo response would have an equivalent prognosis whether they continued to receive the drug or were switched to placebo.

**Methods:** Patients with major depression who met the study’s response criteria (a modified Hamilton Depression Rating Scale score ≤7 and failure to meet major depression criteria after each of the last 3 weeks following 12 to 14 weeks of treatment with fluoxetine hydrochloride, 20 mg/d) were enrolled in a 50-week randomized placebo substitution trial during which the return of depressive symptoms defined relapse. The timing and persistence of response during initial treatment defined true drug or placebo response patterns.

**Results:** Patients with a true drug response pattern relapsed significantly more frequently if they were switched to placebo than if they continued to receive fluoxetine (P<.001 for weeks 12–26, P<.005 for weeks 26–50, and P<.41 for weeks 50–62). Patients with a placebo response pattern had an equivalent outcome whether maintained on fluoxetine therapy or placebo (P<.20 for weeks 12–26, test invalid for weeks 26–50, and P<.67 for weeks 50–62). Patients with a placebo response pattern relapsed more often when they continued to receive fluoxetine than patients with a true drug response pattern (P<.01 for weeks 12–26, P<.10 for weeks 26–50, and P<.36 for weeks 50–62).

**Conclusions:** These findings confirm that pattern analysis validly differentiates true drug from nonspecific initial responses and extend its use to the continuation and maintenance phases of treatment for depression. Investigations into the mechanisms of antidepressant activity might best be limited to those that can account for delayed efficacy. Fluoxetine’s efficacy during the continuation and maintenance phases of treatment may be limited to patients with a true drug pattern of initial response.

Arch Gen Psychiatry. 1998;55:334-343
PATIENTS AND METHODS

To test the hypotheses that the timing and persistence of a short-term benefit will predict long-term outcome, we used data from a multisite study designed to determine the optimal duration of fluoxetine treatment. Five sites enrolled patients with major depression into an open, 12-week trial of fluoxetine; responders then received continued therapy for up to 1 year, during which three quarters were switched to placebo at 1 of 3 points (Figure 1). Each site followed an identical protocol.

PATIENT POPULATION

All study subjects were outpatients aged between 18 and 65 years, were physically healthy, met the DSM-III-R22 criteria for major depression, and had a modified 17-item Hamilton Depression Rating Scale23 (mHAMD) score of 16 or more. The mHAMD substituted hypersomnia and hyperphagia for insomnia, anorexia, and weight loss items in patients with reverse vegetative symptoms. Patients were excluded for unstable medical illnesses, pregnancy or lactation, serious suicidal impulses, history of psychosis, organic mental disorder, mania or antisocial personality disorder, substance abuse disorder in the past year, hypothyroidism, previous fluoxetine treatment for longer than 6 months, and the age of 65 years. All study subjects were outpatients aged between 18 and 65 years, were physically healthy, met the DSM-III-R criteria for major depression, and had a modified 17-item Hamilton Depression Rating Scale (mHAMD) score of 16 or more. The mHAMD substituted hypersomnia and hyperphagia for insomnia, anorexia, and weight loss items in patients with reverse vegetative symptoms.

Patients were excluded for unstable medical illnesses, pregnancy or lactation, serious suicidal impulses, history of psychosis, organic mental disorder, mania or antisocial personality disorder, substance abuse disorder in the past year, hypothyroidism, previous fluoxetine treatment for longer than 6 months, and the age of 65 years. Patients were excluded for unstable medical illnesses, pregnancy or lactation, serious suicidal impulses, history of psychosis, organic mental disorder, mania or antisocial personality disorder, substance abuse disorder in the past year, hypothyroidism, previous fluoxetine treatment for longer than 6 months, and the age of 65 years.

CONSENT PROCEDURE

Signed, informed consent was obtained from each patient and approved by each site's institutional review board. Patients were told that the Eli Lilly Corp, Indianapolis, Ind, sponsored the study to determine the long-term efficacy of fluoxetine and that if their condition initially improved, placebo might be substituted for fluoxetine at some time during the ensuing 11 months. If depressive symptoms returned during the double-blind phase, they were instructed to call immediately for an evaluation. They were told that treatment with fluoxetine hydrochloride, 20 mg/d.

INITIAL 12 WEEKS OF TREATMENT

Patients whose symptoms persisted during a drug-free observation week were treated openly with fluoxetine hydrochloride, 20 mg/d, for 12 weeks. The mHAMD and Clinical Global Impression Global Improvement score, which was not obtained in this trial. We substituted the Clinical Global Impression Global Severity score. A score of 1 (“no psychopathology”) or 2 (“minimal psychopathology”) identified an improved week; other scores (ie, 3 “mild psychopathology”) to 7 “extreme psychopathology”) defined unimproved weeks.

Improvement was early if week 1 or 2 was rated improved. Late improvement occurred if the first improved week was after week 2. Persistent response occurred if an improved week was followed by an unimproved week. The pattern was nonpersistent if an unimproved week followed an improved week.

RANDOMIZATION

Responders were randomly assigned within site by a computer-generated code to 1 of 4 groups: 1, immediate placebo substitution; 2, continuation of fluoxetine therapy for 14 weeks followed by placebo substitution; 3, continuation of fluoxetine therapy for 38 weeks followed by placebo substitution; or 4, continuation of fluoxetine therapy for 50 weeks.

Opaque envelopes contained each patient's treatment code, to be opened only in an emergency. Otherwise, the treatment code was known only to Eli Lilly Corp staff not involved with clinical decisions. Placebo substitution was immediate, without tapering.

TREATMENT DURING CONTINUATION AND MAINTENANCE PHASES

After randomization, treatment was double-blind and patients were seen weekly for 2 weeks, then biweekly for 16 weeks, and then monthly. Extra visits were scheduled if depressive symptoms returned. Relapse was defined as either (1) an mHAMD score of 14 or more for 3 consecutive weeks, (2) a continuation of fluoxetine therapy for 14 weeks followed by placebo substitution; 3, continuation of fluoxetine therapy for 38 weeks followed by placebo substitution; or 4, continuation of fluoxetine therapy for 50 weeks.

Three hypotheses were studied:

1. Patients with a true drug response pattern will relapse more often while taking placebo than while taking continued medication.

2. The relapse rates of patients with a placebo pattern of response will not depend on treatment.

3. While continuing to receive fluoxetine, patients with placebo response patterns will relapse more frequently than patients with true drug response patterns.

STATISTICAL ANALYSES

First, randomization groups were compared for distribution of pattern type. Next, each pair of groups to be compared in the analyses was tested for site, demographic, and illness differences. Sites were compared for different rates of relapse and pattern type. Continuous variables were tested by the t test, and categorical variables were tested by the \( \chi^2 \) test with Yates correction in \( 2 \times 2 \) contingency tables. Because 117 comparisons were performed (13 variables times 3 hypotheses each tested at 3 points), we considered \( P < .001 \) to identify nonrandom differences.

The primary analysis for each hypothesis tested whether relapse during weeks 12 to 26 differed between...
delayed and persistent short-term response as characterizing patients who improve while receiving antidepressant medications, whereas early or nonpersistent responses are typical of patients responding to placebo. A demonstration that short-term response patterns alter subsequent prognosis would further validate pattern analysis. This article examines the predictive validity of pattern analysis in a 14-month study of fluoxetine.

comparison groups. The log-rank test assessed Kaplan-Meier survival curves for differences in time to relapse between groups.

Secondary analyses included similar Kaplan-Meier analyses for weeks 26 to 50 and 50 to 62.

Hypothesis 2, predicting no difference between treatments in patients with a placebo response pattern, asserts the null hypothesis, which cannot be directly demonstrated. Therefore, in addition to conventional significance testing, we also provide a power analysis to suggest the smallest relapse difference this sample size could be expected to detect. Only if there is the power to detect a difference too small to be clinically meaningful would it be reasonable to accept the hypothesis of equivalence.

All analyses were performed for the period immediately following each placebo substitution point (ie, 12, 26, and 50 weeks) until the next substitution point or end of the study. Analyses testing hypothesis 1 compared patients with the true drug pattern randomized to group 1 (placebo) (n=63) with those randomized to groups 2 through 4 (continued fluoxetine therapy) (n=191) for weeks 12 to 26 (primary analysis); those in group 2 (placebo) (n=41) with those in groups 3 and 4 (continued fluoxetine therapy) (n=69) for weeks 26 to 50; and those in group 3 (placebo) (n=26) with those in group 4 (continued fluoxetine therapy) (n=19) for weeks 50 to 60 (secondary analyses). For hypothesis 2, analyses compared patients with the placebo pattern randomized to group 1 (placebo) (n=32) with those randomized to groups 2 through 4 (continued fluoxetine therapy) (n=106) for weeks 12 to 26 (primary analysis); those in group 2 (placebo) (n=11) with those in group 3 or 4 (continued fluoxetine therapy) (n=37) for weeks 26 to 50; and those in group 3 (placebo) (n=26) with those in group 4 (continued fluoxetine therapy) (n=19) for weeks 50 to 62 (secondary analyses). Analyses for hypothesis 3 compared fluoxetine-treated patients with a true drug pattern with fluoxetine-treated patients with a placebo pattern. This was groups 2 through 4 (n=191 and n=106) for weeks 12 to 26 (primary analysis) and groups 3 and 4 (n=69 and n=37) for weeks 26 to 50 and group 4 (n=19 and n=9) for weeks 50 to 62 (secondary analyses) (the first n values are for patients with a true drug pattern of initial response, and the second n values are for patients with a placebo pattern of initial response).

Patients who did not complete the study period were censored after they left the study. All tests of hypotheses were 2 tailed and used an α=.05.

The use of short-term response patterns in predicting outcome is examined; the relationship of outcome to fluoxetine blood levels has been reported, while other study aspects will be reported elsewhere, including short-term treatment outcome, general outcome during continuation and maintenance therapy and side effects. Of 839 patients who began to receive medication, 428 met the response criteria. Of these 428 patients, 33 refused treatment or were otherwise determined to be inappropriate candidates for treatment (eg, they had an intercurrent illness or they had moved) and 3 were missing a week 2 rating. Thus, patterns were available for 392 randomized patients who constitute the sample of this report. Table 1 provides the characteristics of these patients.

ADEQUACY OF RANDOMIZATION

The distribution of patterns among the 4 randomization groups did not differ (Table 2). The 9 pairs of comparison groups did not differ in the distribution of response type or demographic or illness variables (Table 3). The response patterns differed by site as follows:

For this analysis, χ²=9.73, df=4, and P<.05. The data are given as the percentage (number) of patients.

Relapse rates also differed among sites (Table 4). The effect of site on outcome was, therefore, assessed separately for each period, assuming a proportional hazard for time to relapse. The hazard for relapse was modeled as a function of pattern, treatment, and site, including all 2- and 3-factor interactions. The Cox proportional hazards analysis did not show a significant site-by-pattern-by-treatment interaction for any period:
for the 12-week analysis, Wald test score = 5.02, df = 1, and P < .29; for the 26-week analysis, Wald test score = 0.04, df = 1, and P < .95; and for the 50-week analysis, Wald test score = 0, df = 1, and P ≤ .99). We also tested whether patients with placebo patterns had different outcomes depending on whether they had an early initial response or a late and nonpersistent initial response. The Kaplan-Meier survival analysis was not significant for any of the periods (log-rank test score = 0.09, df = 1, P < .77 for weeks 19-26; log-rank test score = 0.14, df = 1, P < .71 for weeks 26-50; and log-rank test score = 1.21, df = 1, P < .28 for weeks 50-62). Because sex, site, or placebo pattern subtype were not significant in the analyses described, subsequent analyses did not adjust for these factors.

Finally, we investigated whether patients dropped out of the study equally among comparison groups. Patients with a placebo response pattern who were receiving placebo during weeks 26 to 50 dropped out significantly more often than with patients with the placebo response pattern who continued to receive fluoxetine, invalidating an assumption of the Kaplan-Meier analysis. The other 8 comparison groups did not differ in their likelihood of dropping out during the risk period (Table 5).

**TESTS OF THE HYPOTHESES**

**Table 6** and **Figure 2** through **Figure 10** summarize the survival analyses.

**Do Patients With a True Drug Short-term Response Pattern Relapse More Often While Receiving Placebo Substitution Than While Receiving Continued Fluoxetine Therapy?**

Among patients with a true drug short-term response pattern, the results of the primary analysis weeks 12 to 26 Kaplan-Meier survival analysis were significant (Figure 2). The weeks 26 to 50 Kaplan-Meier survival analysis (Figure 3) also demonstrated a significantly higher relapse in placebo-treated patients. The results of the weeks 50 to 62 Kaplan-Meier survival analysis were not significant (Figure 4), although the power was poor because of the small sample size.

**Do Patients With a Placebo Response Pattern Relapse More Often While Receiving Placebo Than While Receiving Continued Fluoxetine Therapy?**

None of the primary or secondary Kaplan-Meier analyses demonstrated significantly greater relapse with the administration of placebo than with the administration of continued fluoxetine therapy (Figures 5 through 7). However, the sample size was adequate to detect a ratio of relapse with the administration of continued fluoxetine therapy vs placebo of 2.7 in the weeks 12 to 26 analysis and 3.7 in the weeks 26 to 50 analysis. That is, the weeks 12 to 26 analysis had an 80% chance of finding relapse rates of 20% vs 54% to be statistically significant. The weeks 26 to 50 analysis would have required the rates to be at least as different as 20% and 74% to have had an 80% chance of a statistically significant finding. In addition, the high dropout rate in the placebo group available for the weeks 26 to 50 analysis invalidated the use of the Kaplan-Meier analysis.

**While Receiving Continued Fluoxetine Therapy, Is There More Relapse in Patients With Placebo Response Patterns Than in Those With a True Drug Response Pattern?**

The primary (weeks 12-26) analysis demonstrated significantly less relapse with the continued administration of fluoxetine therapy in patients with a true drug response pattern compared with those with a placebo response pattern (Figure 8). The Kaplan-Meier survival analyses for weeks 26 to 50 and 50 to 62 were not significant at P = .05 (Figures 9 and 10).

**COMMENT**

Pattern analysis differentiates true drug response from placebo response during 6- to 12-week initial treatment trials. The results of this study demonstrate the predictive validity of pattern analysis for the continuation (the period from the end of short-term treatment until 6 months) and maintenance (beyond 6 months) phases of
treatment. Hypotheses were largely confirmed because the continued administration of medication was demonstrated to be effective in patients with a true drug short-term treatment response pattern and not in patients with other short-term response patterns. Patients with a placebo short-term response pattern had a similar prognosis whether they continued to receive fluoxetine therapy or not, although the sample size allowed at best modest power to detect differences. In contrast, following the primary switch point and 1 of the 2 secondary switch points, patients who had a true drug short-term response were significantly more likely to relapse while receiving placebo than while continuing to receive fluoxetine (P < .001, P < .005, and P < .41, respectively). After the primary 12-week switch point, fluoxetine less effectively prevented relapse in patients who had a placebo short-term response than in those with a true drug short-term response (P < .01). Thus, all 3 primary analyses, as well as

### Table 3. Adequacy of Randomization

<table>
<thead>
<tr>
<th>Pattern/Treatment Assignment (Group No.)†</th>
<th>No. of Patients</th>
<th>Mean (±SD)</th>
<th>Baseline mHAMD Score</th>
<th>Week 12 mHAMD Score</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Placebo/placebo (1)</strong></td>
<td>32</td>
<td>42±11</td>
<td>27±15</td>
<td>2±4</td>
</tr>
<tr>
<td><strong>Placebo/fluoxetine hydrochloride (2)</strong></td>
<td>106</td>
<td>40±10</td>
<td>24±13</td>
<td>3±10</td>
</tr>
<tr>
<td><strong>“True drug”/placebo (3)</strong></td>
<td>63</td>
<td>39±10</td>
<td>24±12</td>
<td>5±7</td>
</tr>
<tr>
<td><strong>True drug/fluoxetine (4)</strong></td>
<td>191</td>
<td>40±11</td>
<td>23±11</td>
<td>4±8</td>
</tr>
</tbody>
</table>

### Hypothesis 1

1. **Group 3 vs 4**
   - t = 0.47, df = 252, P < .64
   - t = 0.35, df = 250, P < .73
   - t = 0.57, df = 223, P < .58
   - t = 1.50, df = 252, P < .14
   - t = 0.23, df = 252, P < .82

2. **Group 1 vs 2**
   - t = 1.39, df = 136, P < .17
   - t = 1.04, df = 134, P < .30
   - t = 0.53, df = 116, P < .60
   - t = 1.22, df = 44, P < .23
   - t = 0.88, df = 136, P < .39

3. **Group 2 vs 4**
   - t = 0.08, df = 2,95, P < .94
   - t = 0.65, df = 180, P < .59
   - t = 0.54, df = 245, P < .75
   - t = 0.32, df = 294, P < .75
   - t = 0.50, df = 295, P < .62

### Hypothesis 2

1. **Group 3 vs 4**
   - t = 0.79, df = 108, P < .43
   - t = 1.71, df = 107, P < .09
   - t = 0.48, df = 97, P < .64
   - t = 0.97, df = 108, P < .34
   - t = 1.00, df = 65, P < .33

2. **Group 1 vs 2**
   - t = 1.17, df = 46, P < .25
   - t = 1.58, df = 46, P < .13
   - t = 1.36, df = 9, P < .21
   - t = 0.85, df = 46, P < .41
   - t = 0.36, df = 45, P < .73

3. **Group 2 vs 4**
   - t = 1.73, df = 104, P < .09
   - t = 0.07, df = 58, P < .95
   - t = 0.69, df = 91, P < .38
   - t = 0.63, df = 104, P < .54
   - t = 0.46, df = 100, P < .02

### Hypothesis 3

1. **Group 3 vs 4**
   - t = 0.05, df = 43, P < .97
   - t = 0.53, df = 36, P < .60
   - t = 0.21, df = 37, P < .84
   - t = 0.27, df = 43, P < .80
   - t = 0.10, df = 41, P < .92

2. **Group 1 vs 2**
   - t = 0.46, df = 15, P < .66
   - t = 0.06, df = 15, P < .96
   - t = 0.12, df = 14, P < .91
   - t = 0.77, df = 15, P < .46
   - t = 0.39, df = 15, P < .70

3. **Group 2 vs 4**
   - t = 0.21, df = 26, P < .24
   - t = 0.39, df = 25, P < .71
   - t = 0.45, df = 22, P < .66
   - t = 1.56, df = 26, P < .14
   - t = 0.81, df = 11, P < .44

* mHAMD indicates modified Hamilton Depression Rating Scale; SAD, seasonal affective disorder; NA, data are not available; and ellipses, data are not applicable.
† Hypothesis 1 indicates that in patients with a true drug short-term response pattern, relapse will occur more often in those treated with placebo than in those maintained on fluoxetine therapy; hypothesis 2, in patients with a placebo short-term response pattern, there will be no difference in relapse between those switched to placebo and those maintained on fluoxetine therapy; and hypothesis 3, in patients maintained on fluoxetine therapy, those with a placebo short-term response pattern will relapse more often than those with a true drug short-term response pattern.

**N=27.**
§N=86.
¶N=170.
#N=10.
**N=32.**
††N=37.
‡‡N=61.
§§N=8.
||N=23.
¶¶N=16.

©1998 American Medical Association. All rights reserved.
analyses, confirmed that pattern analysis differentiates nonspecific benefit from improvement attributable to specific drug effects.

Because studies validating pattern analysis have included tricyclic antidepressants, monoamine oxidase inhibitors, and selective serotonin reuptake inhibitors, this technique may apply generally to antidepressant medications. Pattern analysis might be applicable to other treatments characterized by the delayed onset of persistent effects, such as lithium carbonate, antipsychotic medications, anticonvulsants, and electroconvulsive therapy. Protein synthesis represents a candidate mechanism for delayed action of treatments because its initiation takes days to weeks, while immediate effects on reuptake or receptors would be candidates if they initiated a cascade of other effects over time.

These data suggest that patients with a placebo pattern of short-term response present the clinician with a dilemma. Regardless of whether fluoxetine therapy is continued, a seemingly high likelihood of relapse can be expected. If fluoxetine therapy was discontinued in a patient with a placebo response pattern, the recurrence of depressive symptoms would lead to the assumption that the medication had been responsible for the initial improvement. These results suggest that this conclusion

<table>
<thead>
<tr>
<th>Site</th>
<th>Weeks</th>
<th>Weeks</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>12-26</td>
<td>26-50</td>
</tr>
<tr>
<td>2</td>
<td>31 (18/58)</td>
<td>26 (6/23)</td>
</tr>
<tr>
<td>3</td>
<td>35 (40/113)</td>
<td>20 (9/44)</td>
</tr>
<tr>
<td>4</td>
<td>46 (27/59)</td>
<td>43 (9/21)</td>
</tr>
<tr>
<td>5</td>
<td>35 (27/77)</td>
<td>21 (7/33)</td>
</tr>
</tbody>
</table>

*All data are given as the percentage of patients relapsing (number of patients relapsing/total number of patients at a site during the particular period indicated). For weeks 12 to 26, χ² = 11.40, df = 4, P < .03; for weeks 26 to 50, χ² = 4.38, df = 4, P < .36; and for weeks 50 to 62, χ² = 11.25, df = 4, P < .03.
might be inaccurate, but it likely would be reached anyway. The discontinuation of an apparently effective agent is, therefore, problematic. Continuation would also be questionable. Although these results call into question the prophylactic efficacy of fluoxetine in patients with a placebo pattern of short-term response, as no treatment has demonstrated efficacy in patients whose short-term response could be considered to be nonspecific, we have no recommendations to make for the patient with a placebo pattern response.

Loss of apparent benefit has clinical and heuristic implications. Some dramatically improved patients later relapse despite the continuation of treatment. Others relapse when the drug is withdrawn, then do not improve when the medication is reinstituted. Pattern analysis suggests that some relapse while continuing to receive medication represents loss of a placebo effect rather than loss of a true drug effect. In this study, for example, of patients relapsing while continuing to receive fluoxetine therapy, almost half (48%) had a nonspecific response pattern. Similarly, some failure to improve on second exposure to the same medication might be due to loss of a placebo effect, rather than acclimatization to pharmacological effects. Attempts to understand the treatment

---

### Table 5. Dropouts by Pattern and Treatment Assignment*

<table>
<thead>
<tr>
<th>Pattern/Treatment Assignment (Group No.)</th>
<th>Weeks</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>12-26</td>
</tr>
<tr>
<td>&quot;Placebo&quot;/placebo (1)</td>
<td>25 (8/32)</td>
</tr>
<tr>
<td>Placebo/fluoxetine hydrochloride (2)</td>
<td>18 (19/106)</td>
</tr>
<tr>
<td>&quot;True drug&quot;/placebo (3)</td>
<td>24 (15/63)</td>
</tr>
<tr>
<td>True drug/fluoxetine (4)</td>
<td>18 (85/191)</td>
</tr>
</tbody>
</table>

Hypothesis†

| 1 (group 3 vs 4)          | $\chi^2=0.90$, df=1, $P<.35$ | $\chi^2=1.62$, df=1, $P<.21$ | $\chi^2=0.0$, df=1, $P<.98$ |
| 2 (group 1 vs 2)          | $\chi^2=0.78$, df=1, $P<.38$ | $\chi^2=6.97$, df=1, $P<.01$ | $\chi^2=2.01$, df=1, $P<.16$ |
| 3 (group 2 vs 4)          | $\chi^2=0.01$, df=1, $P<.94$ | $\chi^2=0.11$, df=1, $P<.74$ | $\chi^2=0.17$, df=1, $P<.68$ |

* Data are given as the percentage of patients who dropped out of the study (number of patients who dropped out/total number of patients in the particular group indicated).
† The hypotheses are defined in the second footnote of Table 3.

### Table 6. Days to Relapse*

<table>
<thead>
<tr>
<th>Pattern/Treatment Assignment (Group No.)</th>
<th>Weeks</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>12-26</td>
</tr>
<tr>
<td>&quot;Placebo&quot;/placebo (1)</td>
<td>69±7 (55-82)</td>
</tr>
<tr>
<td>Placebo/fluoxetine hydrochloride (2)</td>
<td>80±3 (74-85)</td>
</tr>
<tr>
<td>&quot;True drug&quot;/placebo (3)</td>
<td>68±5 (59-77)</td>
</tr>
<tr>
<td>True drug/fluoxetine (4)</td>
<td>83±2 (83-90)</td>
</tr>
</tbody>
</table>

Hypothesis†

| 1 (group 3 vs 4)          | Log-rank test score=22.37, $df=1$, $P<.001$ | Log-rank test score=8.23, $df=1$, $P<.005$ | Log-rank test score=0.71, $df=1$, $P<.41$ |
| 2 (group 1 vs 2)          | Log-rank test score=1.71, $df=1$, $P<.20$ | The log-rank test score was invalid because of a high dropout rate in the placebo group | Log-rank test score=0.19, $df=1$, $P<.67$ |
| 3 (group 2 vs 4)          | Log-rank test score=6.75, $df=1$, $P<.01$ | Log-rank test score=2.74, $df=1$, $P<.10$ | Log-rank test score=0.86, $df=1$, $P<.36$ |

* Data are given as the mean (±SE) days to relapse (95% confidence interval) unless otherwise specified.
† The hypotheses are defined in the second footnote of Table 3.

---

Figure 2. Kaplan-Meier survival curves comparing survival from weeks 12 to 26 among patients with a “true drug” short-term response pattern, the patients were treated with fluoxetine hydrochloride vs placebo (log-rank test score, 22.37; $P<.001$). Vertical bars indicate points that individuals drop out before reaching relapse criteria.
The course of individual patients must consider the possibility that initial response was not due to direct medication effects. Unequal assignment among randomization groups could produce spurious results. For example, if men and women relapsed at different rates and were differently distributed among groups, ignoring sex effects in the analyses could yield results that seem to implicate other factors. Indeed, patterns and relapse were differentially distributed among treatment sites. However, the Cox proportional hazards analysis did not reveal significant site-by-pattern-by-treatment interaction effects on time to relapse ($P > .29$ for weeks 12-26, $P < .95$ for weeks 26-50, and $P < .99$ for weeks 50-62). Other demographic and illness variables were not differentially distributed among response type, site, or comparison groups.

Carryover and withdrawal effects are a concern in any discontinuation trial. Fluoxetine is not associated with withdrawal symptoms. However, the relatively long half-life of fluoxetine and its active metabolite, norfluoxetine, might delay relapse in patients switched to placebo. The resulting underestimate of relapse in placebo-treated patients...
could result in failure to show between-group differences. This is one reason for caution in accepting failure to demonstrate differences in addressing hypothesis 2.

An additional problem with hypothesis 2 is that in predicting no difference, it asserts the null hypothesis. What can be demonstrated is that if a difference exists it is probably smaller than \( x \). If \( x \) is small enough, it might be judged inconsequential. The patients with placebo response patterns in this study had an 80% chance of demonstrating between-treatment differences in the weeks 12 to 26 analysis on the order of 20% vs 54%. A relapse difference of 25% vs 50% might have been missed. The detectable differences in the weeks 26 to 50 and 50 to 62 analyses were even larger. Hypothesis 2 might best be considered not disconfirmed.

These results must be replicated before they should influence clinical practice. While validated for short-term treatment with 3 classes of antidepressant medications, to our knowledge, this study is the first assessing its use beyond short-term treatment. The predictive validity of pattern analysis must be replicated for fluoxetine and extended to other selective serotonin reuptake inhibitor and non–selective serotonin reuptake inhibitor antidepressant medications.
In summary, after the 3- and 6-month switch points, pattern analysis prospectively identified one patient group that showed continued medication effects (ie, different relapse rates while receiving drug vs placebo) and another group that did not. Patients who were more likely to relapse while continuing to receive fluoxetine therapy were also identified. These and previous results suggest that delayed-persistent short-term responses may identify with 80% likelihood patients who had a true drug response, while early or nonpersistent initial responses select patients with a high proportion of nonspecific improvement. In attempting to identify the mechanism of action of antidepressants, limiting inquiries to mechanisms that are delayed and persistent and testing candidate mechanisms only in patients identified as having a true drug response will likely maximize the chances of demonstrating these mechanisms.

Accepted for publication July 30, 1997.

This study was supported by the Eli Lilly Corp, Indianapolis, Ind.

Reprints: Jonathan W. Stewart, MD, 1051 Riverside Dr, Unit 35, New York, NY 10032.

REFERENCES