0271-0749/02/2203-0309/0
Journal of Clinical Psychopharmacology
Vol. 22, No. 3

Printed in U.S.A.

Multivariate Meta-Analysis of Controlled Drug Studies for Obsessive-Compulsive Disorder

DEBORAH L. ACKERMAN, MS, PHD*, AND SANDER GREENLAND, MS, DRPH*†

**Department of Epidemiology, School of Public Health, and †Department of Statistics, College of Letters and Sciences, University of California, Los Angeles, California*

Meta-analytic reviews of placebo-controlled studies for obsessive-compulsive disorder have found that clomipramine is more effective than drugs with more selective actions on serotonin reuptake, whereas in most direct comparisons, clomipramine's superiority has been less obvious. The authors used metaregression to identify sources of heterogeneity in placebo-controlled trials of clomipramine, fluvoxamine, sertraline, and paroxetine. They evaluated such patient characteristics as age, gender, age of obsessive-compulsive disorder (OCD) onset, and baseline severity of OCD and depression, and such study characteristics as exclusion or inclusion criteria, length of single-blind prerandomization period, length of trial, number of subjects, and publication year. We found considerable heterogeneity across studies that was associated, in part, with publication year, length of single-blind prerandomization period, length of trial, and severity of patients' OCD. The apparent superiority of clomipramine persisted after controlling for these factors. The authors also confirmed previous reports that placebo response is higher in more recent studies. Meta-analyses can help characterize responders and nonresponders. The authors urge investigators to provide summaries of patient characteristics, especially baseline severity, age at onset, and duration of OCD, by patients' response. (J Clin Psychopharmacol 2002;22:309–317)

The effectiveness of the various serotonin reuptake inhibitors (SRIs) to treat obsessive-compulsive disorder (OCD) is well established. Meta-analytic reviews have estimated that response rates for all drugs are between 40% and 60%, whereas placebo response rates are less than 20%.[1–7] These reviews have consistently found that in placebo-controlled studies, clomipramine appears to be more effective (i.e., the clomipramine group shows more improvement relative to the placebo group) than drugs with more selective actions on serotonin reuptake (SSRIs); in most direct comparisons, however, clomipramine's superiority has been less obvious.

Differences in the methodologic characteristics of studies may have contributed to clomipramine's advantage in placebo-controlled studies. Publication year has often been noted: more recent studies have reported smaller effect sizes than earlier ones.[3, 4, 6] Greist and associates[4] and Stein and associates[3] hypothesized that the effect size of more recent clinical trials was affected by inclusion of more treatment-resistant patients and by an increase in placebo response rates over time.

Different patient characteristics have also been identified. Greist and associates[4] noted that patients in the later studies of SSRIs had lower baseline scores on the Yale-Brown Obsessive Compulsive Scale (Y-BOCS) and hypothesized that placebo improvement may be related to less severe OCD. Supplemental analyses of data from the large multicenter trials of clomipramine and fluoxetine found that response rates may be lower with subclinical depression.[8, 9] Poor response to clomipramine, but not fluoxetine, was also associated with early age of OCD onset; however, the subjects in the fluoxetine trial had an earlier average age of onset. Response to fluoxetine varied by previous treatment: subjects with no previous treatment for OCD showed the greatest response.

The aim of this study was to evaluate quantitatively characteristics of the published clinical trials that might account for differences in observed effect size. Previous meta-analyses computed effect size by subtracting the final drug treatment change score from the final placebo change score and dividing the difference by the pooled change standard deviation. We used metaregression, also called effect-size modeling, to identify sources of heterogeneity.[10] Weighted metaregression allowed us to use the actual change score in each treatment arm,

Received December 20, 2000; accepted after revision September 21, 2001.

Address requests for reprints to: Deborah L. Ackerman, MS, PhD, University of California, Los Angeles, School of Public Health, Box 951772, Los Angeles, CA 90095-1772. Address e-mail to: dackerma@ucla.edu.

weighted by the inverse of the variance of that score, while controlling for the change score in the placebo arm of each trial. Protocol information and patient characteristics from each study were included to compare the magnitude of improvement in each study with each drug and the interactions among study, patient, and treatment factors. We evaluated such patient characteristics as age, gender, age of OCD onset, and baseline severity of OCD and depression, and exclusion or inclusion criteria of the protocols, length of single-blind prerandomization period, length of trial, number of subjects, and publication year.

Methods

We searched the computer databases of Current Contents, MEDLINE, and PsychInfo for all entries indexed with the keywords *obsessive*, *compulsive*, and *clinical trial*. We also referred to published reviews and previous meta-analyses. We identified 25 clinical trials that met the following inclusion criteria: randomized, double-blind, parallel trials; 8 weeks or longer; efficacy assessed with Y-BOCS; and point estimates and SD (or SE) provided in or calculable from the report.

Improvement was measured by subtracting baseline Y-BOCS from final Y-BOCS, so a negative change represented improvement. The drug effect was measured by subtracting the improvement in the placebo arms from the improvement in the drug arms, so a negative difference represented greater improvement with active treatment.

When a measure of variability (SD or SE) for the change score was not provided, we estimated it as follows. We first calculated the covariance between baseline and final scores for all studies that included standard deviations for baseline, final, and change scores for each treatment group, using the formula

$$\text{Covariance} = SD^2_{\text{baseline}} + SD^2_{\text{final}} - SD^2_{\text{change}} / 2$$

We then calculated the correlation between baseline and final standard deviations as

$$\text{Correlation} = \text{covariance} / SD_{\text{baseline}} \times SD_{\text{final}}$$

We used the smallest calculated correlation to estimate the missing standard deviations of change score, using the formula

$$SD^2_{\text{change}} = SD^2_{\text{baseline}} + SD^2_{\text{final}} - 2 \text{ orrelation } (SD_{\text{baseline}} \times SD_{\text{final}})$$

We used weighted least-squares regression[10] to estimate average treatment effect within groups of trials (i.e., placebo-controlled trials of clomipramine, fluvoxamine, fluoxetine, sertraline, or paroxetine; placebo-controlled trials of all SSRIs; controlled comparisons of clomipramine with fluvoxamine; comparisons of clomipramine with all SSRIs). Each treatment group in each study was entered as a separate observation. Indicator variables for each study were included to control for changes in Y-BOCS scores in the placebo arm of each study. Weights were computed as the inverse of the variance (SD^2/N) of the change score in each treatment arm:

$$\text{Variance}_{\text{difference}} = SE^2_{\text{baseline}} + SE^2_{\text{final}} - 2 \frac{(\text{correlation}}{SE_{\text{baseline}} \times SE_{\text{final}})}$$

Heterogeneity within groups of studies was evaluated by treating the residual sum of squares as a χ^2 test statistic.[10] Sources of heterogeneity were then explored by entering study factors (i.e., year of publication, inclusion or exclusion criteria, number of weeks of prerandomization and postrandomization) and summary patient characteristics (average baseline Y-BOCS, Hamilton Depression Rating Scale (HAM-D), age, OCD duration, onset age, and gender distribution) into each regression model. If average duration of OCD (or average onset age) was not provided, it was estimated by subtracting average onset age (or duration) from average age. If patient summary data were not provided for each treatment arm separately, the overall averages were used for each arm. Product terms of each factor and each treatment arm indicator were added in order to examine whether drug or placebo effects were modified (predicted) by the factor.

Results

We estimated the effect size as the difference in improvement (decrease in Y-BOCS) between active drug and placebo. The results are summarized in Table 1. A negative effect represents improvement, and a negative difference between groups represents greater effect in the treatment group. For the seven clomipramine trials,[11–16] the net improvement compared with placebo was −8.19 (95% confidence interval [CI], −10.53, −5.85). The multicenter clomipramine study was analyzed as two trials because two different protocols were followed. For four fluvoxamine studies,[17–20] improvement relative to placebo was −4.84 (95% CI, −7.78, −1.83); for three fluoxetine studies[21–23] (two studies[21, 22] compared fixed doses of 20, 40, and 60 mg with one placebo group), net improvement was −1.61 (95% CI, −2.18, −1.04, for 20 mg dose); and for four sertraline studies[24–27] (one study[26] compared fixed doses of 50, 100, and 200 mg with one placebo group), the improvement was −2.47 (95% CI, −6.13, 1.20 for 50 mg). Considerable heterogeneity was found within placebo-controlled comparisons of each drug.

One study compared paroxetine with clomipramine

TABLE 1. Double-blind, placebo-controlled comparisons: Yale-Brown Obsessive-Compulsive Scale change*

Study	Year	N_1	Drug effect	Wt_1	N_2	Placebo effect	Wt_2	Difference	95%	CL
CMI										
CMICollab[11]	1991	118	−11.50	2.39	134	−1.60	3.60	−9.90	−11.59	−8.21
CMICollab[11]	1991	120	−10.10	2.04	129	−0.90	2.65	−9.20	−11.03	−7.37
Jenike et al.[12]	1989	13	−8.90	0.33	14	−2.30	0.49	−6.60	−11.00	−2.20
Greist et al.[13]	1990	16	−9.00	0.50	16	−1.70	0.30	−7.30	−11.84	−2.76
Mavissakalian et al.[14]	1990	13	−16.80	0.16	12	−0.30	0.23	−16.50	−22.94	−10.10
Hoehn-Saric et al.[15]	1993	13	−10.20	0.32	12	−1.70	0.39	8.50	−13.19	−3.81
Zohar and Sudge[16]	1996	99	−8.00	1.47	99	−5.00	1.59	−3.00	−5.24	−0.76
CMI vs. placebo, pooled difference (95% CL): −8.19 (−10.53, −5.85)										
FLV										
Goodman et al.[17]	1989	21	−7.00	0.43	21	2.00	0.58	−9.00	−13.29	−4.71
Jenike et al.[18]	1990	18	−3.80	0.99	20	−0.90	0.32	−2.90	−8.21	2.41
Mallya et al.[19]	1992	14	−6.47	0.40	14	−1.13	0.84	−5.33	−9.09	−1.58
Goodman et al.[20]	1996	78	−4.50	1.97	78	−1.00	3.25	−3.50	−5.27	−1.73
FLV vs. placebo, pooled difference (95% CL): −4.84 (−7.78, −1.83)										
FLX										
Montgomery et al.[21†]	1993	52	−5.13	1.27	57†	−3.70	1.59†	−1.43	−3.76	0.90
Montgomery et al.[21]	1993	52	−4.76	1.09		−3.70		−1.06	−3.49	1.37
Montgomery et al.[21]	1993	55	−6.07	1.26		−3.70		−2.37	−4.76	0.02
Tollefson et al.[22†]	1994	87	−4.70	1.55	89†	−0.70	4.03†	−4.00	−5.85	−2.15
Tollefson et al.[22]	1994	89	−5.40	2.04		−0.70		−4.70	−6.38	−3.02
Tollefson et al.[22]	1994	90	−6.80	1.56		−0.70		−6.10	−7.95	−4.25
Jenike et al.[23]	1997	23	−2.80	0.52	21	−0.20	0.43	−2.60	−6.63	1.43
FLX vs. placebo, pooled difference (95% CL): −1.61 (−2.18, −1.04)†										
SER										
Chouinard et al.[24]	1990	43	−3.79	1.31	44	−1.48	0.51	−2.31	−5.54	0.92
Jenike et al.[25]	1990	10	−2.20	0.12	9	−0.50	0.16	−1.70	−8.25	4.85
Greist et al.[26†]	1995	80	−6.00	0.76	84†	−4.20	1.32[2]	−1.80	−4.62	1.02
Greist et al.[26]	1995	81	−4.50	0.82		−4.20		−0.30	−3.06	2.46
Greist et al.[26]	1995	80	−7.50	1.07		−4.20		−3.30	−5.85	−0.75
Kronig et al.[27]	1999	86	−9.00	1.24	81	−4.00	1.64	−5.00	−7.34	−2.66
SER vs. placebo, pooled difference (95% CL): −2.47 (−6.13, 1.20)†										
PAR										
Zohar and Judge[16]	1996	201	−8.00	3.14	99	−5.00	1.59	−3.00	−4.91	−1.09
TZR										
Pigott et al.[28]	1992	11	−2.70	0.29	6	−2.83	0.70	0.13	−1.34	1.60

*CL, confidence limits; CMI, clomipramine; CMICollab, clomipramine Collaborative Study Group; FLV, fluvoxamine; SSRI, selective serotonin reuptake inhibitor; TRZ, triazolam; Wt, weight (I/variance).

†Fixed dose studies, one placebo group. Pooled result represents mean difference for lowest dose. All SSRIs vs. placebo, pooled difference: −1.85 (−2.43, − 1.27).

and with placebo[16] and found that the net improvement of paroxetine over placebo was −3.00 (95% confidence limit [CL], −4.91, −1.09). Another compared trazodone with placebo and found no difference between groups (difference = 0.13; 95% CL, −1.34, 1.60).[28] The pooled difference for all serotonin reuptake inhibitors (SRIs, clomipramine, and SSRIs combined) and placebo was −2.57 (95% CL, −3.51, −1.63); for all SSRIs, the difference was −1.85 (95% CL, −2.43, −1.27).

As with other meta-analyses of controlled comparisons of clomipramine and an SSRI, we found no difference between treatments in studies that compared clomipramine with fluvoxamine,[29–32] fluoxetine,[33] or paroxetine[16] (Table 2). The pooled difference between clomipramine and fluvoxamine from four direct comparisons was 1.23 (95% CL, −1.11, 3.56). The pooled difference between clomipramine and all SSRIs was 0.15 (95% CL, −8.86, 9.16). There did not appear to be appreciable heterogeneity among direct comparisons.

Desipramine was the control in one fluvoxamine[34] and one sertraline study.[35] The first study found fluvoxamine superior to desipramine (difference, −8.40; 95% CL, −13.26, −3.54). The second found sertraline slightly superior to desipramine (difference, −2.40; 95% CL, −4.90, 0.10).

There were differences across placebo-controlled studies in a number of study characteristics. All of the clomipramine trials and the paroxetine and clomipramine trial followed a 2-week, single-blind, prerandomization period; all fluoxetine trials were conducted after a 1-week period;

TABLE 2. Double-blind direct comparisons: Yale-Brown Obsessive-Compulsive Scale

Study	Year	N_1	CMI effect	Wt_1	N_2	SSRI effect	Wt_2	Difference	95% OL	
CMI vs. FLV										
Smeraldi et al.[29]	1992	5	−10.20	0.03	5	−13.60	0.24	3.40	−8.29	15.09
Freeman et al.[30]	1994	32	−7.80	0.89	34	−8.60	0.94	1.86	−2.55	6.27
Koran et al.[31]	1996	39	−7.30	0.54	34	−7.70	0.55	0.40	−3.35	4.15
Milanfranchi et al.[32]	1997	13	−11.00	0.11	13	−11.30	0.17	0.30	−7.08	7.68
CMI vs. FLV, pooled difference (95% CL): 1.23 (−1.11, 3.56)										
CMI vs. FLX										
Lopez-Ibor et al.[33]	1996	25	−8.90	0.50	30	−7.50	0.35	1.40	−5.74	2.94
CMI vs. PAR										
Zohar and Judge[16]	1996	99	−8.00	1.47	201	−8.00	3.14	0	−1.94	1.94
CMI vs. all SSRIs, pooled difference (95% CL): 0.15 (−8.86, 9.16)										

Study	N_1	SSRI effect	Wt_1	N_2	DESIP effect	Wt_2	Difference	95% CL	
FLV vs. DESIP									
Goodman et al.[34]	21	−8.10	0.27	19	0.30	0.41	−8.40	−13.26	−3.54
SER vs. DESIP									
Hoehn-Saric et al.[35]	79	−8.40	1.23	85	−6.00	1.23	−2.40	−4.90	0.10

*CL, confidence limits; CMI, clomipramine; DESIP, desipramine; FLV, fluvoxamine; FLX, fluoxetine; PAR, paroxetine; SER, sertraline; SSRI, selective serotonin reuptake inhibitor; Wt, Weight (I/variance).

the fluvoxamine trials included 1-week, 2-week, and 3-week single-blind periods; and the sertraline trials followed 1-week or 2-week periods. The length of the studies ranged from 8 weeks to 13 weeks (in one sertraline study). Most studies required a minimum OCD duration of at least 1 year, but three clomipramine trials required 2 years, and the paroxetine and clomipramine trial required only 6 months. Most studies excluded patients with depression, but depressed patients were included in one clomipramine, two fluvoxamine, and two fluoxetine trials. Among studies that reported the minimum Y-BOCS required for inclusion, a Y-BOCS score of 16 or greater was required by five clomipramine, two fluoxetine, and the paroxetine and clomipramine study; a minimum of 20 was required by two sertraline studies.

Patient characteristics also varied across studies. There were differences in OCD severity, duration, onset age, and in the gender distribution. Baseline Y-BOCS scores were highest in the clomipramine studies (mean average, 26.6; range, 24.1–31.0) compared with fluvoxamine (mean, 23.1; range, 19.6–25.6), fluoxetine (mean, 22.9; range, 18.9–25.5), and sertraline (mean, 23.7; range, 22.6–25.2). Among studies that reported the average duration of OCD or age at OCD onset, patients in the sertraline trials had shorter duration of OCD (mean, 11.0 years; range, 5.0–8.0) compared with clomipramine (mean, 14.7; range, 13.0–16.0) and fluvoxamine (mean, 18.0; range, 15.0–22.0). Patients in three fluvoxamine trials also had longer OCD duration (mean, 18.0; range, 15.0–22.0) compared with one fluoxetine trial (mean, 12.9). Age of OCD onset was later in the sertraline studies (mean, 20.4; range, 17.6–23.0) than in the clomipramine (mean, 21.3; range, 20.0–23.0) and fluvoxamine (mean, 20.4; range, 17.6–23.0) studies. The proportion of males in the sertraline studies was 0.63 (range, 0.44–0.80), which was higher than in all other studies (clomipramine, range 0.33–0.57; fluvoxamine, range 0.42–0.55; fluoxetine, range 0.39–0.57; paroxetine, 0.44). Among studies that reported severity of depression, HAM-D baseline scores were higher in two fluoxetine trials (mean, 10.7; range, 9.0–13.0) than in six clomipramine studies (mean, 7.6; range, 4.7–12.2) and in one fluvoxamine study (mean, 7.6).

We used metaregression to examine whether any of the above study differences were sources of heterogeneity within the drug and placebo arms of studies grouped by drug. Regression models included each study factor and the product of the treatment group (drug or placebo) indicator and study factor. We also included baseline Y-BOCS in all models to assess and control for baseline severity. The factors associated with drug response are summarized in Table 3.

Factors associated with drug response

Three clomipramine studies[12–14] were not included because they represented a subset of patients in the large multicenter trials. Within the remaining four clomipramine studies, the study factors associated with improvement in the drug arms were date of publication, length of trial, and required minimum duration of OCD. Since 1991, the difference between clomipramine and placebo has narrowed. Over a period of 5 years, there was less improvement with clomipramine (the average yearly change in Y-BOCS was 1.22 points less; 95% CL, 0.77, 1.66; $p < 0.001$). Longer trials, 12 versus 10 weeks, were also associated with 5.78 points less improvement with clomipramine (95% CL, 3.97, 7.59; $p < 0.001$).

Several characteristics of the patients who participated

TABLE 3. Summary of study design and patient characteristics associated with improvement in drug arms*

	CMI	FLV	FLX	SER
Publication yr (per yr)	1.22 (0.77, 1.66)	1.84 (−3.41, 7.10)	−0.88 (−1.48, −0.28)	−1.93 (−7.98, 4.13)
Prerandomization (per wk)	—‡	−3.07 (−7.43, 1.29)	—§	−0.45 (−32.2, 31.3)
Study Length (per 2 wk)	5.78 (3.97, 7.59)	5.96 (2.45, 9.46)	−0.55 (−0.93, −0.17)	0.96 (−5.57, 7.48)
Minimum required OCD duration (per 1/2 yr)	−1.07 (−3.44, 1.30)	—¶	—¶	—¶
Depression included	1.39 (−4.53, 7.32)‥	−5.96 (−9.46, −2.45)‥	1.06 (−4.30, 6.42)‥	—††
Average age (per 3 yr)	6.80 (1.78, 11.8)	7.10 (−49.7, 63.9)	−0.97 (−3.35, 1.81)	1.61 (−4.24, 7.45)
Average onset age (per 3 yr)	6.41 (2.14, 10.7)	−4.69 (−10.4, 1.03)‡‡	—§§	0.95 (−0.67, 2.57)
Average baseline HAM-D (per 5 points)	−0.40 (−4.94, 4.14)¶¶	—[a]	2.36 (1.16, 3.55)¶¶	—[b]

*CMI, Clomipramine; FLV, Fluvoxamine; FLX, fluoxetine; HAM-D, Hamilton Depression Rating Scale; OCD, obsessive-compulsive disorder; SER, sertraline; Y = BOCS, Yale-Brown Obsessive-Compulsive Scale.

†Data represent change in final Y-BOCS score in the treatment group over levels of each factor (95% confidence limits). Results are from separate regression models that included an indicator variable for treatment arm, each factor, a product term for treatment/factor, baseline Y-BOCS score, and a constant.

‡All CMI studies employed 2-week lead-in.

§All FLX studies employed 1-week lead-in.

¶All FLV and FLX studies required 1-year duration. Two SER studies specified 1-year minimum duration.

‥Depressed patients included in one CMI study, one FLV study, and two FLX studies.

††All SER studies excluded depressed patients.

‡‡Information on average onset age provided in three FLV studies.

§§Only one FLX study provided data on average onset age.

¶¶Information on average baseline HAM-D provided by three CMI studies and two FLX studies.

[a]Information on average baseline HAM-D provided by only one FLV study.

[b]Information on average baseline HAM-D not provided in any SER study.

in the clomipramine studies were also found associated with drug response. The average age of patients ranged from 35 to 38 years. Older patients in the clomipramine arms showed less improvement. Studies with an average patient age of 38 years in the clomipramine group showed 6.79 points less change than studies with an average age of 35 (95% CL, 1.78, 11.9; $p = 0.01$). Later age of OCD onset, which ranged from 20 to 23 years in four studies, was associated with less clomipramine improvement. An average age of onset of 23 years was associated with 6.41 points less improvement than an average onset age of 20 (95% CL, 2.14, 10.7; $p = 0.01$).

Within the fluvoxamine studies, longer trials were also associated with less improvement. Studies that were 10 weeks long yielded less improvement in the fluvoxamine arms (5.96 change points; 95% CL, 2.45, 9.46; $p = 0.01$) than trials that were 8 weeks long. One study[17] included depressed patients and resulted in more improvement than studies that excluded depressed patients: the fluvoxamine change score was 5.96 points greater (95% CL, −9.46, −2.45, $p = 0.01$). A longer prerandomization period (range, 1–3 weeks) was associated with somewhat greater improvement (−3.07 change points per week; 95% CL, −7.43, 1.29; $p = 0.11$). There was also a weak association between improvement and age at OCD onset (range, 18–23 years in three studies). Each 3-year increase in OCD onset age was associated with 4.69 points greater improvement (95% CL, −10.4, 1.07; $p = 0.07$).

Publication year and study length were associated with response in the fluoxetine studies. Unlike the clomipramine studies, for each year between 1993 and 1997, there was more improvement with fluoxetine (change per year, −0.88; 95% CL, −1.48, −0.28; $p = 0.01$). Unlike the clomipramine and fluvoxamine studies, longer trials (8, 10, or 13 weeks) were associated with more improvement with fluoxetine: each 2 weeks resulted in 0.55 points lower Y-BOCS (95% CL, −0.93, −0.17; $p = 0.01$). The only patient characteristic found to be associated with fluoxetine response was average baseline HAM-D score. Two studies[21, 22] included depressed patients and provided information about the HAM-D scores of subjects in each treatment arm. The baseline HAM-D range was 9.0 to 13.0: higher scores were associated with less fluoxetine improvement (2.36 Y-BOCS change points for each 5-point HAM-D increase; 95% CL, 1.16, 3.55; $p = 0.002$).

None of the study design or patient factors was associated with response to sertraline.

Factors associated with placebo response

Many of the factors that were associated with more or less improvement (Y-BOCS decline) in the drug arms were also related to placebo improvement, but in the opposite direction. Within the clomipramine studies, more recent studies and longer trials showed more improvement in the placebo arms since 1991: the average yearly decline in Y-BOCS was 0.71 points (95% CL, −0.96, −0.46; $p < 0.001$). Trials that were 12 weeks long showed 3.56 points greater decline than 10-week trials (95% CL, −4.61, −2.52; $p < 0.001$). Older patients in the placebo arms showed more improvement: studies with an average age of 38 in the placebo group showed 4.14

points more decline than studies with an average age of 35 (95% CL, −6.93, −1.34; $p = 0.01$). Later age of OCD onset was also associated with greater placebo improvement. An average onset age of 23 was associated with 3.95 points greater decline than an onset at 20 (95% CL, −6.76, −1.14; $p = 0.01$).

Within fluvoxamine trials, longer studies were also associated with somewhat greater placebo response. Trials that were 10 weeks long resulted in slightly more improvement (−0.99 points; 95% CL, −2.33, 0.35; $p = 0.10$) than trials that were 8 weeks long.

Unlike the clomipramine studies, more recent fluoxetine trials produced less improvement with placebo: the average yearly change in Y-BOCS was 1.05 points lower (95% CL, 0.21, 1.89, $p = 0.02$). Unlike the clomipramine and fluvoxamine studies, longer fluoxetine trials were associated with less improvement with placebo (1.09 points less change each 2 weeks; 95% CL, 0.43, 1.75; $p = 0.004$). In contrast to the fluvoxamine studies, two fluoxetine studies that included depressed patients[21,22] produced somewhat greater placebo improvement than one that did not (−5.81 points greater change; 95% CL, −12.6, 1.00; $p = 0.09$). Higher baseline HAM-D scores were associated with greater placebo improvement (−4.68 change points for each 5-point increase in HAM-D scores; 95% CL, −6.92,−2.43; $p = 0.001$).

Publication year approximates the temporal relation among studies. Fig. 1 depicts Y-BOCS improvement over time in drug and placebo arms of all placebo-controlled comparisons of SRIs (clomipramine and each SSRI). Placebo response increased somewhat over time. Year of publication, which predicted some of the heterogeneity within clomipramine and fluoxetine studies, was correlated with several study factors and patient characteristics. Studies conducted more recently tended to have shorter prerandomization periods (Pearson correlation, −0.45; $p = 0.002$) and longer lengths (correlation, 0.55; $p < 0.001$). A longer prerandomization period was also associated with lower Y-BOCS scores at randomization (correlation, 0.31; $p = 0.04$). There was a weak trend for more recent studies to include patients with lower Y-BOCS scores at randomization (correlation, −0.29; $p = 0.11$). Among studies that reported patients' HAM-D scores, more recent studies included patients with higher scores (correlation, 0.76; $p < 0.001$).

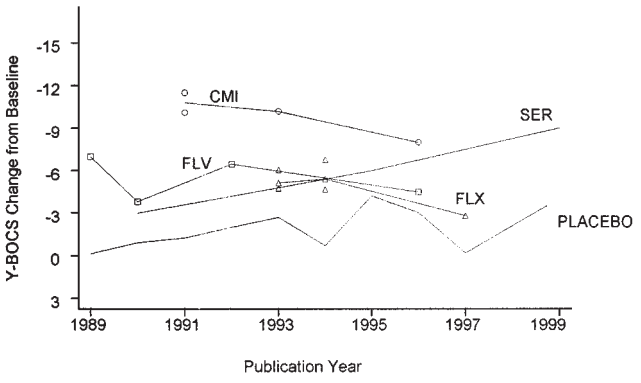

FIG. 1. Yale-Brown Obsessive Compulsive Scale change by publication year, serotonin reuptake inhibitor placebo-controlled studies.

Multiple regression results

All of the study and patient factors associated with either drug or placebo effects among placebo-controlled SRI studies when considered individually were entered into multiple regression models for change in Y-BOCS among treatment groups. The results are summarized in Table 4. The models included active drug indicators (clomipramine, fluvoxamine, fluoxetine, sertraline, and paroxetine), variables for average baseline Y-BOCS score, publication year, prerandomization period, and an indicator for studies that included depressed patients. Minimum required duration was not included because it was not specified in two sertraline studies.[24,26] The full model was generated by data from both the active drug and the placebo arms. The coefficients of each study factor in the full model represent the average effect across treatment arms in all studies.

The factors associated with response were baseline Y-BOCS, publication year, and length of prerandomization period. As expected given regression to the mean, higher baseline Y-BOCS was associated with greater improvement (Y-BOCS change, −2.52 points per baseline Y-BOCS point increase; 95% CL, −4.16, −0.87; $p = 0.004$). More recent studies showed greater improvement (change, −1.15 points per 5 years; 95% CL, −2.20, −0.05; $p = 0.04$). A longer prerandomization period was associated with less improvement (change, 1.20 Y-BOCS points per week; 95% CL, −0.02, 2.38; $p = 0.05$). All active treatments yielded greater improvement relative to placebo. There was more improvement in the clomipramine arms (change, −8.55; 95% CL, −9.90, −7.21) than among all SSRIs, and there were no differences among SSRIs. Required minimum duration of OCD was then added. The resulting model based on data from all but two sertraline studies indicated no association between this factor and improvement.

We then evaluated the same factors in a model generated by data from the placebo arms alone. The results indicated that placebo improvement was associated with baseline Y-BOCS, publication year, and prerandomization period. Higher baseline Y-BOCS was associated with greater placebo improvement (Y-BOCS change, −2.91 points per baseline Y-BOCS point increase; 95% CL, −5.01, −0.81; $p = 0.01$). More recent studies showed greater improvement (change, −0.32 points per year; 95% CL, −0.61, −0.02; $p = 0.04$). A

TABLE 4. Metaregresssion results of relationship between Yale-Brown Obsessive-Compulsive Scale change scores and study factors*†

	Drug and placebo arms				Placebo arms			
Study/patient factors	Coefficient	95% CL		p value	Coefficient	95% CL		p value
Baseline Y-BOCS	−2.52	−4.16	−0.87	< 0.01	−2.91	−5.01	−0.81	0.01
Publication year‡	−1.15	−2.20	−0.05	0.04	−0.32	−0.61	−0.02	0.04
Prerandom weeks	1.20	0.02	2.38	0.05	2.04	0.58	3.50	0.01
Trial length	0.13	−0.18	0.43	0.41	0.42	−0.07	0.90	0.09
Depression included	0.38	−0.73	1.50	0.49	0.85	−0.47	2.18	0.19
CMI arms	−8.55	−9.90	−7.21	< 0.00	—	—	—	—
FLV arms	−4.38	−6.19	−2.57	< 0.00	—	—	—	—
FLX arms	−3.65	−5.01	−2.30	< 0.00	—	—	—	—
SER arms	−3.85	−5.39	−2.31	< 0.00	—	—	—	—
PAR arms	−5.58	−6.48	−3.74	< 0.00	—	—	—	—

*CL, confidence limits; CMI, clomipramine; FLV, fluvoxamine; FLX, fluoxetine; PAR, paroxetine; SER, sertraline; Y-BOCS, Yale-Brown Obsessive-Compulsive Scale.

†Results of multiple regression models that included an indicator variable for each active treatment arm, each study factor, and a constant. Data from all placebo-controlled SRI trials were used.

‡Rescaled to represent change per 5 years.

longer prerandomization period was associated with less placebo improvement (change, 2.04 Y-BOCS points per week; 95% CL, -0.58, 3.50; $p = 0.01$). Longer trials were associated with somewhat less placebo improvement (change, 0.42 points per two weeks; 95% CL, −0.07, 0.90; $p = 0.09$).

We also examined product terms for treatment group indicator and study factors. Product terms were not included for factors that did not vary within groups of drug studies (e.g., prerandomization within clomipramine and fluoxetine studies). The coefficients of such product terms represent the difference in average drug effect associated with a 1-unit increase in the variables. Publication year was associated with less improvement in the clomipramine arms (change, 4.09 points per 5 years; 95% CL, 1.60, 6.58; $p = 0.002$) and fluvoxamine arms (change, 2.26 points per 5 years; 95% CL, −0.46, 4.58; $p = 0.06$). A longer prerandomization period was associated with greater improvement in fluvoxamine arms (change, −2.23 points per week; 95% CL, −4.59, 0.12; $p = 0.06$).

We then included average age, duration of OCD, and age at OCD onset. One clomipramine study,[12] one fluvoxamine study,[19] and two fluoxetine studies[22, 23] were not included because they did not provide this information. The resulting model, generated by data from 12 studies, found no association between duration or onset age and improvement in placebo arms. In clomipramine arms, later onset age was associated with less improvement (change, 3.44 points per 3 years; 95% CL, 1.26, 5.62; $p = 0.003$), whereas in the fluvoxamine arms it was associated with greater improvement (change, −5.58; 95% CL, −9.13, −2.03; $p = 0.003$). Baseline HAM-D score was then added. The resulting model, generated by data from six studies, exhibited almost no association between depression and drug or placebo improvement (change, −0.39 points per 5-point increase in HAM-D; 95% CL, −4.58, 3.81).

Conclusions

Our analyses confirm the findings of Greist and associates,[4] Abramowitz,[5] Kobak and associates,[6] and others that clomipramine has been more effective than all the SSRIs in placebo-controlled trials. We found considerable heterogeneity among studies associated, in part, with publication year, length of single-blind prerandomization period, length of trial, and severity of OCD. Heterogeneity within clomipramine and fluvoxamine studies was also associated with patients' age at OCD onset. The apparent superiority of clomipramine persisted after controlling for these factors. As with other meta-analyses of controlled comparisons of clomipramine and an SSRI, we found no difference between treatments in studies that compared clomipramine directly with fluvoxamine, fluoxetine, or paroxetine.

Several possible explanations for clomipramine's superiority in placebo comparisons have been suggested. The numerous side effects of clomipramine may have contributed to its greater effect size in placebo comparisons. Abramowitz[5] observed that drugs with more side effects had larger effects. In our previous analyses of clomipramine data, we found that response was associated with the number of reported side effects, that certain early side effects predicted good response, and that some of the same side effects also predicted placebo response.[36] In contrast, in a similar analysis of fluoxetine, we did not find that outcome was associated with the number of side effects.[37] A problem with using reported side effects to understand different effect sizes across studies is that each protocol may have a different method for eliciting side effect reports. Even differences across sites within a multicenter study may produce different estimates of the prevalence of adverse reactions.[38]

The inclusion of treatment-resistant patients in later studies may have resulted in lower response rates. Some reports included information about previous treatment history.[17, 20, 25, 34] Unfortunately, they did not evaluate response status and previous treatment. Goodman and associates[39] reexamined data from their fluvoxamine study[20] and found that response rates were lower among subjects who had failed to respond to previous trials of clomipramine or fluoxetine. In our supplemental analysis of data from one multicenter fluoxetine study,[22] we found that improvement was greatest for patients without previous drug treatment.[9]

Another possible explanation for the higher response rates in earlier studies is that more recent trials have been longer. We found that longer trials yielded less improvement. It is possible that in the early 10-week clomipramine and 8-week fluvoxamine trials, some of the subjects in the active treatment arms responded to the nonspecific study effects (such as side effects). In longer studies, that improvement was not maintained.

We have also confirmed previous findings that placebo response has increased over time. Increasing placebo response rates may be the result of several factors. Montgomery and associates[21] hypothesized that higher placebo response may be caused by shorter trials, or by the inclusion of subjects with depression and with less severe and shorter duration OCD. We found that a shorter prerandomization lead-in period and longer trials resulted in greater placebo improvement. More recent studies tended to have shorter (1-week) lead-in periods followed by longer observation periods. We also found that studies with longer lead-in periods had lower Y-BOCS scores at randomization and that less severe OCD resulted in less improvement in the active drug and placebo groups. There was a weak trend toward less improvement in studies with a longer average duration of OCD. Average OCD duration ranged from 13 to 16 years in the clomipramine studies, 15 to 22 years in the fluvoxamine studies, and 5 to 18 years in the sertraline studies.

Finally, among studies that reported patients' HAM-D scores, more recent studies included patients with higher scores. Among studies that included patients with depression, more severe depression was associated with less improvement in the drug arms but more improvement in the placebo arms.

Greist and associates[26] suggested that placebo patients might be benefiting from forms of behavior therapy available through self-help guides and other sources. Behavior therapy may have residual effects that enhance placebo response. Marks and associates[40] found benefits lasting 1 to 5 years after behavioral treatment. In our analysis of fluoxetine data, we found that improvement was also high among the small number of patients with previous behavior therapy.[9]

Later age of OCD onset was associated with less clomipramine improvement but greater improvement with fluvoxamine. Among fluvoxamine studies, the average onset age ranged from 18 to 23 years. This finding from clomipramine studies contrasts with our previous analysis[8] of original data from the multicenter clomipramine trials.[11] In that report, we found that earlier onset was associated with poor response to clomipramine. The difference might have been caused by reliance on an estimated average onset age for all subjects rather than actual patient data. The reported average onset age for clomipramine studies fell within a narrow range from 20 to 23.

Response rates were provided in some reports, although different criteria were sometimes used. For all placebo-controlled comparisons of clomipramine and the SSRIs, response rates in the active drug arms ranged between 0.32 and 0.60 (mean, 0.42). Response rates in placebo arms ranged from 0 to 0.35 (mean, 0.15). The response rate in direct comparisons between clomipramine and SSRIs ranged from 0.32 to 0.85 (mean, 0.53).

Because many patients do not respond to these drugs, additional work is needed to identify characteristics of responders and nonresponders. We found only three reports that provided clinical information about patients by their response status.[17, 26, 34] The baseline Y-BOCS scores of fluvoxamine responders and nonresponders summarized by Goodman and associates[17] indicated that responders had less severe OCD. In another fluvoxamine study, Goodman and associates[34] reported that responders to fluvoxamine had somewhat lower HAM-D baseline scores. Greist and associates[26] reported that responders to both sertraline and placebo had less severe OCD at baseline.

Meta-analyses can help characterize responders and nonresponders to active drug and to placebo if the published results of clinical trials include additional summary information about patients who respond. We urge investigators to include supplemental responder analyses and to provide summaries of patient characteristics, especially baseline severity of OCD and depression, age at onset, and duration of OCD, by their response to treatment. We also recommend that future meta-analysts search clinical trial registries such as the Cochrane Controlled Trial Registry to minimize publication bias. Until we have better information to characterize potential treatment responders, we recommend clinicians follow the Expert Consensus Guidelines for the treatment of OCD.[41]

Acknowledgment

This work was supported by NIMH grant MH01513 to Dr. Ackerman.

We are grateful to Dr. Ken A. Kobak for providing additional information not provided in some published reports.

References

1. Cox BJ, Swinson RP, Morrison B, et al. Clomipramine, fluoxetine, and behavior therapy in the treatment of obsessive-compulsive disorder: a meta-analysis. J Behav Ther Exp Psychiatry 1993;24:149–53.
2. Piccinelli M, Pini S, Bellantuono C, et al. Efficacy of drug treatment in obsessive-compulsive disorder: a meta-analytic review. Br J Psychiatry 1995;166:424–43.
3. Stein DJ, Spadaccini E, Hollander E. Meta-analysis of pharmacotherapy trials for obsessive-compulsive disorder. Int Clin Psychopharmacol 1995;10:11–8.
4. Greist JH, Jefferson JW, Kobak KA, et al. Efficacy and tolerability of serotonin transport inhibitors in obsessive compulsive disorder: a meta-analysis. Arch Gen Psychiatry 1995;52:53–60.
5. Abramowitz J. Effectiveness of psychological and pharmacological treatments for obsessive-compulsive disorder: a quantitative review. J Consult Clin Psychol 1997;65:44–52.
6. Kobak KA, Greist JH, Jefferson JW, et al. Behavioral versus pharmacological treatments of obsessive compulsive disorder: a meta-analysis. Psychopharmacology 1998;136:205–16.
7. Pigott TA, Seay SM. A review of the efficacy of selective serotonin reuptake inhibitors in obsessive-compulsive disorder. J Clin Psychiatry 1999;60:101–6.
8. Ackerman DL, Greenland S, Bystritsky A, et al. Predictors of treatment response in obsessive compulsive disorder: multivariate analyses from a multicenter trial of clomipramine. J Clin Psychopharmacol 1994;14:247–54.
9. Ackerman DL, Greenland S, Bystritsky A. Clinical characteristics of responders to fluoxetine treatment of OCD. J Clin Psychopharmacol 1998;18:185–92.
10. Greenland S. Meta-analysis. In: Rothman K, Greenland S, eds. Modern epidemiology, edn. 2. Philadelphia: Lippincott-Raven, 1998:643–73.
11. Clomipramine Collaborative Study Group. Clomipramine in the treatment of patients with obsessive-compulsive disorder. Arch Gen Psychiatry 1991;48:730–8.
12. Jenike MA, Baer L, Summergrad P, et al. Obsessive-compulsive disorder: a double-blind, placebo-controlled trial of clomipramine in 27 patients. Am J Psychiatry 1989;146:1328–30.
13. Greist JH, Jefferson JW, Rosenfeld R, et al. Clomipramine and obsessive compulsive disorder: a placebo-controlled double-blind study of 32 patients. J Clin Psychiatry 1990;51:292–7.
14. Mavissakalian MR, Jones BA, Olson S, et al. Clomipramine in obsessive-compulsive disorder: clinical response and plasma levels. J Clin Psychopharmacol 1990;10:261–8.
15. Hoehn-Saric R, McLeod DR, Zimmerli WD, et al. Symptoms and physiologic manifestations in obsessive compulsive patients before and after treatment with clomipramine. J Clin Psychiatry 1993;54:272–6.
16. Zohar J, Judge R. Paroxetine versus clomipramine in the treatment of obsessive-compulsive disorder. OCD Paroxetine Study Investigators. Br J Psychiatry 1996;169:468–74.
17. Goodman WK, Price LH, Rasmussen SA, et al. Efficacy of fluvoxamine in obsessive-compulsive disorder: a double-blind comparison with placebo. Arch Gen Psychiatry 1989;46:36–44.
18. Jenike MA, Hyman S, Baer L, et al. A controlled trial of fluvoxamine in obsessive-compulsive disorder: implications for a serotonergic theory. Am J Psychiatry 1990;147:1209–15.
19. Mallya GK, White K, Waternaux C, et al. Short- and long-term treatment of obsessive-compulsive disorder with fluvoxamine. Ann Clin Psychiatry 1992;4:77–80.
20. Goodman WK, Kozak MJ, Liebowitz M, et al. Treatment of obsessive-compulsive disorder with fluvoxamine: a multicentre, double-blind, placebo-controlled trial. Int Clin Psychopharmacol 1996;11:21–9.
21. Montgomery SA, McIntyre A, Osterheider M, et al. A double-blind, placebo-controlled study of fluoxetine in patients with DSM-III-R obsessive-compulsive disorder. Lilly European OCD Study Group. Eur Neuropsychopharmacol 1993;3:143–52.
22. Tollefson GD, Rampey AH Jr, Potvin JH, et al. A multicenter investigation of fixed-dose fluoxetine in the treatment of obsessive-compulsive disorder. Arch Gen Psychiatry 1994;51:559–67.
23. Jenike MA, Baer L, Minichiello WE, et al. Placebo-controlled trial of fluoxetine and phenelzine for obsessive-compulsive disorder. Am J Psychiatry 1997;154:1261–4.
24. Chouinard G, Goodman W, Greist J, et al. Results of a double-blind placebo controlled trial of a new serotonin uptake inhibitor, sertraline, in the treatment of obsessive-compulsive disorder. Psychopharmacol Bull 1990;26:279–84.
25. Jenike MA, Baer L, Summergrad P, et al. Sertraline in obsessive-compulsive disorder: a double-blind comparison with placebo. Am J Psychiatry 1990;147:923–8.
26. Greist J, Chouinard G, DuBoff E, et al. Double-blind parallel comparison of three dosages of sertraline and placebo in outpatients with obsessive-compulsive disorder. Arch Gen Psychiatry 1995; 52:289–95.
27. Kronig MH, Apter J, Asnis G, et al. Placebo-controlled, multicenter study of sertraline treatment for obsessive-compulsive disorder. J Clin Psychopharmacol 1999;19:172–6.
28. Pigott TA, L'Heureux F, Rubenstein CS, et al. A double-blind, placebo controlled study of trazodone in patients with obsessive-compulsive disorder. J Clin Psychopharmacol 1992;12:156–62.
29. Smeraldi E, Erzegovesi S, Bianchi I, et al. Fluvoxamine vs. clomipramine treatment in obsessive-compulsive disorder: a preliminary study. Trends Exp Clin Psychiatry 1992;8:63–5.
30. Freeman CP, Trimble MR, Deakin JF, et al. Fluvoxamine versus clomipramine in the treatment of obsessive compulsive disorder: a multicenter, randomized, double-blind, parallel group comparison. J Clin Psychiatry 1994;55:301–5.
31. Koran LM, McElroy SL, Davidson JR, et al. Fluvoxamine versus clomipramine for obsessive-compulsive disorder: a double-blind comparison. J Clin Psychopharmacol 1996;16:121–9.
32. Milanfranchi A, Ravagli S, Lensi P, et al. A double-blind study of fluvoxamine and clomipramine in the treatment of obsessive-compulsive disorder. Int Clin Psychopharmacol 1997;12:131–6.
33. Lopez-Ibor JJ Jr, Saiz J, Cottraux J, et al. Double-blind comparison of fluoxetine versus clomipramine in the treatment of obsessive compulsive disorder. Eur Neuropsychopharmacol 1996;6:111–8.
34. Goodman WK, Price LH, Delgado PL, et al. Specificity of serotonin reuptake inhibitors in the treatment of obsessive-compulsive disorder: comparison of fluvoxamine and desipramine. Arch Gen Psychiatry 1990;47:577–85.
35. Hoehn-Saric R, Ninan P, Black DW, et al. Multicenter double-blind comparison of sertraline and desipramine for concurrent obsessive-compulsive and major depressive disorders. Arch Gen Psychiatry 2000;57:76–82.
36. Ackerman DL, Greenland S, Bystritsky A, et al. Relationship between early side effects and therapeutic effects of clomipramine therapy in OCD. J Clin Psychopharmacol 1996;16:324–8.
37. Ackerman DL, Greenland S, Bystritsky A. Side effects as predictors of drug response in OCD. J Clin Psychopharmacol 1999;19: 459–65.
38. Tangrea JA, Adrianza ME, Helsel WE. Risk factors for the development of placebo adverse reactions in a multicenter clinical trial. Ann Epidemiol 1994;4:327–31.
39. Goodman WK, Ward H, Kablinger A, et al. Fluvoxamine in the treatment of obsessive-compulsive disorder and related conditions. J Clin Psychiatry 1997;58[suppl 5]:32–49.
40. Marks IM, Stern RS, Mawson D, et al. Clomipramine and exposure for obsessive-compulsive rituals, I. Br J Psychiatry 1980;136:1–25.
41. Treatment of obsessive-compulsive disorder: the Expert Consensus Panel for obsessive-compulsive disorder. J Clin Psychiatry 1997,58[suppl 4]:2–72.