

Antidepressant Pharmacotherapy in the Treatment of Depression in the Very Old: A Randomized, Placebo-Controlled Trial

Steven P. Roose, M.D.

Harold A. Sackeim, Ph.D.

K. Ranga Rama Krishnan, M.B.,
Ch.B.

Bruce G. Pollock, M.D., Ph.D.

George Alexopoulos, M.D.

Helen Lavretsky, M.D.

Ira R. Katz, M.D., Ph.D.

Heikki Hakkarainen, M.D.

Old-Old Depression Study Group

Objective: This study determined the efficacy of antidepressant medication for the treatment of depression in the "old-old."

Method: This randomized 8-week medication trial compared citalopram, 10–40 mg/day, to placebo in the treatment of patients 75 and older with unipolar depression.

Results: A total of 174 patients who were 58% women with a mean age of 79.6 years (SD=4.4) and a mean baseline Hamilton Depression Rating Scale score of 24.3 (SD=4.1) were randomly assigned to treatment at 15 sites. There was a main effect for site but not for treatment condition. The remission rate, defined as a final

Hamilton depression scale score <10, was 35% for the citalopram and 33% for the placebo groups. However, patients with severe depression (baseline Hamilton depression scale score >24) tended to have a higher remission rate with medication than with placebo (35% versus 19%).

Conclusions: In the oldest group of community-dwelling patients to be studied to date, medication was not more effective than placebo for the treatment of depression. However, given the considerable psychosocial support received by all patients, the placebo condition represents more than the ingestion of an inactive pill. Across sites, there was considerable range in response to medication, 18% to 82%, and to placebo, 16% to 80%.

(*Am J Psychiatry* 2004; 161:2050–2059)

In the next few decades, there will be a dramatic increase in the population over the age of 65, and people over age 85 constitute the most rapidly growing segment of the population (1). Thus, illnesses prevalent in older individuals will increasingly challenge the health care system. Major depression is of special concern; if undiagnosed or inadequately treated, the illness presents direct risks, e.g., suicide and indirect risks through a negative impact on the course of other major diseases prevalent in older patients. For example, depressed patients after a myocardial infarction have a higher cardiac mortality than nondepressed patients with comparable cardiovascular illness (2).

Antidepressant medications and specific psychotherapies are effective in the treatment of major depression in nongeriatric adults (i.e., ages 18–60). However, data on efficacy and safety in younger patients may not generalize to older groups. For example, the efficacy, safety, and tolerability of antidepressant medications can be altered by age-associated changes in pharmacokinetics and pharmacodynamics (3–5). Thus, the use of antidepressant medications or psychotherapies in patients with late-life depression should be based on the results of randomized controlled trials in geriatric groups.

Although there have been many placebo or comparator-controlled trials of antidepressant medications in late-life major depression, most have had important methodolog-

ical limitations, including insufficient statistical power, use of designs not fully masked or randomized, problematic medication dosing strategies, and inadequate statistical analysis and presentation of findings (e.g., few studies have reported remission rates). A review of late-life antidepressant treatment trials found that all but three placebo-controlled and five comparator trials had more than one of these problems.

The three placebo-controlled trials used as active treatments fluoxetine (6-week trial duration) (6), sertraline (8 weeks) (7), and both venlafaxine and fluoxetine (8 weeks) (8). In a study of 771 patients, the remission rate, defined as a final Hamilton Depression Rating Scale score ≤ 7 , was significantly greater in the fluoxetine (28%) than placebo (18%) groups. In a study of 716 patients, the remission rate, defined as a final Hamilton depression scale score ≤ 10 , was significantly greater in the patients treated with sertraline (29%) than the patients given placebo (23%). In the third study, which randomly assigned 300 patients, there were no significant differences in the rates of remission, defined as a final Hamilton depression scale score of ≤ 8 , across the venlafaxine (42%), fluoxetine (29%), and placebo (38%) groups. Thus, in two of the three studies, the active medication was more effective than placebo. However, the remission rates in these studies were low, and the differences between active medication and placebo were limited.

The five active comparator trials compared sertraline and fluoxetine (12 weeks) (9), sertraline and nortriptyline (12 weeks) (10), mirtazapine and paroxetine (8 weeks) (11), and two trials, paroxetine and nortriptyline (6 weeks and 12 weeks) (12, 13). No trial found a significant difference in the remission rates between active treatments. Remission rates ranged from 29% to 63%, with most remission rates clustering around 60%.

Differences in trial duration, dosing strategies, and remission criteria make it problematic to do comparisons across studies, even within the placebo-controlled and comparator groupings. Furthermore, as in younger adults, comparator antidepressant trials in late-life depression samples consistently report higher remission rates than placebo-controlled trials, even when the same active medication is studied (14). What is most notable is that given the widespread use of antidepressant medications in geriatric patients, the data indicating the efficacy of these treatments are strikingly limited.

A major limitation of the available data is that the average age of patients in geriatric trials to date ranges from 60 to 72 years, and only a small number of patients over the age of 75 were included in any study. Late-life spans a broad age range and is often divided into the “young-old” (ages 60–75) and the “old-old” (ages ≥75). Among the “old-old,” antidepressant treatment may be especially complicated because of the high frequency and severity of comorbid conditions, such as cognitive impairment or heart disease. To date, the few studies of pharmacological treatment in this age group have primarily or exclusively included patients in residential settings (15, 16). However, most depressed patients ages 75 and above—indeed, most people ages 75 and above—are self-sufficient, live in the community, and do not have cognitive impairment comparable to nursing home residents.

To address the need for evidence-based depression treatment in this vulnerable population, a group of clinical investigators designed a clinical trial for the treatment of major depression in the old-old. Given the absence of data in this age group, it was felt that placebo control was essential. To further the goal of generalizability, the inclusion and exclusion criteria allowed enrollment of “real-world” patients, e.g., patients with mild to moderate cognitive impairment and serious medical illness, and did not exclude most concomitant medications. The investigators approached Forest Laboratories, which agreed to provide financial support. A condition of this agreement was that the investigators would have possession of the database and perform all statistical analyses.

Method

The study was a multicenter, double-blind, randomized 8-week trial comparing citalopram to placebo in depressed patients 75 and older. All sites recruited patients by radio and newspaper advertisements and/or through referral from other physicians. At the initial visit, a comprehensive psychiatric evaluation, includ-

ing a Structured Clinical Interview for DSM-IV, a Hamilton depression scale, a Mini-Mental Status Examination (MMSE), and a medical history, were performed to confirm diagnosis, severity and preliminary eligibility. If the patient met inclusion criteria and signed informed consent, a physical examination, an ECG, and routine blood work were performed.

Inclusion criteria were 1) 75 years old or older and not living in a residential setting; 2) with unipolar depression, single or recurrent, nonpsychotic, by DSM-IV criteria with the modification that the current episode must be at least 4 weeks in duration; 3) with a Hamilton depression scale score of ≥20 on the 24-item Hamilton depression scale at the initial visit and at the end of 1 week of placebo; and 4) willing and able to give informed consent. Exclusion criteria were 1) having bipolar disorder, obsessive-compulsive disorder, psychotic disorder, or current substance abuse or substance dependence within the past year (other than nicotine) by DSM-IV criteria; 2) having current suicide intent or serious suicide attempt within the past year; 3) meeting National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association criteria for probable Alzheimer's disease or probable vascular dementia; 4) having an MMSE score ≤18; 5) having Parkinson's disease; 6) having an acute, severe, or unstable medical illness, 7) in the current episode of major depression, failing to respond to either a trial of a selective serotonin reuptake inhibitor (SSRI) (fluoxetine, paroxetine, or citalopram at 20 mg/day or sertraline 50 mg/day for at least 4 weeks) or trials of two or more different classes of antidepressants other than SSRIs.

Patients who met inclusion and exclusion criteria entered a 1-week single-blind placebo lead-in. The assessments performed at the end of the placebo week (baseline visit) and every visit thereafter included the 24-item Hamilton depression scale, the Montgomery-Åsberg Depression Rating Scale (17), the Beck Depression Inventory, and the Clinical Global Impression (CGI) of severity and improvement at all visits after baseline. The Hamilton Anxiety Rating Scale and the Center for Epidemiologic Studies Depression Scale (CES-D Scale) (18) were performed at baseline and at the end of weeks 2, 4, and 8 of treatment; the Positive Affects Scale (19), the Medical Outcomes Study 36-Item Short-Form Health Survey (20), the Instrumental Activities of Daily Living (21), and the MMSE were performed at baseline and at the end of week 8 or upon early termination. At baseline and at the end of 8 weeks or upon early termination, an ECG, routine blood work, a coagulation profile, and a serum anticholinergic assay were performed. At baseline, the patient received a magnetic resonance imaging (MRI) scan and at baseline and week 8, an extensive neuropsychological battery.

At the end of the placebo week, if a patient continued to meet inclusion and exclusion criteria, he or she was randomly assigned to citalopram, 20 mg/day, or placebo. Assignment to treatment group was performed by a computer-generated randomization schedule, and for each site, patient randomization numbers were assigned in ascending sequence. If intolerable adverse events emerged, the dose could be reduced to 10 mg/day (half a tablet), and the patient could subsequently be rechallenged with a higher dose. At the end of week 4, patients with a Hamilton depression scale score >10 had the dose increased to two pills per day, i.e., 40 mg/day of either citalopram or placebo.

After random assignment, patients returned for study visits at weeks 1, 2, 3, 4, 5, 6, and 8 (final week). At each study visit, the patient met with the treating physician to review progress and side effects and with the study raters and research assistants for structured ratings, blood pressure measurements, etc., and completed self-report measures. Compliance was determined by weekly pill count, and if a patient returned more than 20% of the tablets that were to have been taken during that treatment week, they were considered to be noncompliant for that week.

To establish interrater reliability for the Hamilton depression scale, the primary outcome measure, all raters for this study independently viewed five taped interviews of older patients with major depression of varying severity with mild, moderate, or no cognitive impairment. Each taped interview lasted 30 to 45 minutes, and after viewing the tape, the raters completed the 24-item Hamilton depression scale. To allow for calculation of intraclass correlation coefficients, raters were expected to rate all items for all patients based on the information available on the tape. If a rater felt that there was no information pertinent to a specific item, the corresponding symptom was to be considered absent and the item scored as 0. The study was approved by the institutional review board at each participating site.

Statistical Methods

The database for this study is in the possession of the investigators. All the data analyses reported in this article were performed by two of the investigators (H.A.S. and S.P.R.). The intraclass correlation for the 24-item Hamilton depression scale score was 0.92.

The intent-to-treat group for the efficacy analyses was defined as the patients who completed the placebo lead-in, started randomized treatment, and had at least one subsequent assessment of clinical status. The completer group was composed of all patients who received 8 weeks of randomized treatment and were compliant. To enhance the power of detecting site differences, the sites that enrolled eight or fewer patients were combined into one "virtual" site. Seven of the 15 study sites constituted the virtual site, contributing 38 patients to the intent-to-treat group and 33 patients to the completer group. Demographic and clinical features as a function of treatment group and study site were compared by using analyses of variance (ANOVAs) and logistic regression for continuous and dichotomous measures, respectively.

Primary and secondary clinical outcome measures were specified a priori. Baseline was defined as the assessment conducted at the visit following the 1-week placebo lead-in, just before randomization (week 0), and the final score was that obtained at the last patient assessment (last observation carried forward, week 8 for completers). The primary outcome measures were serial changes in Hamilton depression scale scores, final Hamilton depression scale score adjusted for baseline values and covariates, and the rates of response and remission. Patients who showed a reduction in Hamilton depression scale scores of 50% or greater at the final assessment were classified as responders; remitters were defined as patients who had a final Hamilton depression scale score of 10 or less.

The secondary clinical outcome measures included final scores on the Montgomery-Åsberg Depression Rating Scale and the CES-D Scale, each adjusted for baseline values and covariates. A secondary categorical outcome measure was the rate of response, as defined by a final CGI improvement score of 1 or 2.

The primary analyses of clinical outcome were conducted in the intent-to-treat group. The first analysis applied a mixed-effects model to the Hamilton depression scale scores at each visit (22–24). The fixed effects included treatment group (two levels) and site (nine levels) and their interaction as between-subject terms, study week (eight levels) as a repeated-measures factor, and the classifications of depression severity (two levels) and age at onset (two levels) as covariates. The model was fully factorial with respect to the between-subject and repeated-measures factors. The main effects of the two covariates and their first- and second-order interactions with treatment condition and time point of Hamilton depression scale assessment (study week) were also modeled. Analyses of covariance (ANCOVAs) and logistic regression analyses were conducted on the remaining primary and the secondary outcome measures. In these analyses, the between-subject factors (treatment group and study site) and covariates (depression severity and age at onset) were identical to

those used in the random regression model. In the ANCOVAs and logistic regressions, the main effects of each covariate and the first-order interaction between treatment group and each covariate were modeled.

The covariates, depression severity and age at onset, were selected a priori based on their status as patient-level factors previously associated with outcomes in antidepressant trials. In both mid- and late-life depression, there is some evidence that the greatest separation of placebo and active treatments occurs in patients with more severe baseline symptom profiles (7, 25, 26). Since there is no consensually held definition of severe depression, the group was separated into "high" and "low" severity groups by using the mean baseline Hamilton depression scale score as the cutoff. There is also little consensus on the cutoff for defining late-onset depression, although onset after age 50 or 60 is commonly used (27). Particularly since relations between age at onset and clinical outcome are not expected to be linear throughout the age range, the use of binary classification is helpful. In light of the minimum age of the group (75 years) and the desire to maximize homogeneity within the late-onset group, a relatively high cutoff of 60 years was used to define the late-onset group. Other variables that were considered for use as covariates included age, sex, cumulative medical comorbidities, duration of the current episode, and MMSE scores. Screening ANCOVAs showed that none of these five variables had significant associations with any of the clinical outcome measures, either across or within the treatment groups. Thus, these variables were not used as covariates in the final analyses.

Symptomatic improvement and changes in quality of life can be dissociated, especially in the long-term treatment of patients with late-life major depression. Changes in function comprised another set of primary outcomes. Raw scores on each of the eight Medical Outcomes Study 36-Item Short-Form Health Survey subscales were transformed into standardized scores, with a potential range of 0 to 100 (28), and the change from baseline to post-treatment was computed for each subscale. A repeated-measures ANCOVA was conducted on these change scores, with treatment group and site as between-subject factors, subscale as the repeated-measures factor, and depression severity and age at onset groupings as covariates.

The final set of analyses concerned the frequency and type of adverse effects experienced after random assignment to citalopram or placebo. Fisher's exact test was used to compare the treatment groups in the rates of any side effect that was reported in at least 5% of the patients randomly assigned to either treatment group. The treatment groups were also compared in the rates and causes given for dropout after random assignment. Because of the low rate of early termination during the acute treatment trial, these comparisons were also made with Fisher's exact test.

The analyses conducted in the intent-to-treat group were repeated in the completer group. Throughout, an alpha level of 0.05 was considered significant, and all statistical tests were two-tailed.

Results

Of the 253 patients screened for the study, 184 (72.7%) met inclusion/exclusion criteria and agreed to participate (Figure 1). After the single-masked placebo lead-in, 178 (97.8%) of the 184 patients continued to meet inclusion/exclusion criteria and were randomly assigned. Of these 178 patients, 174 (97.8%) participated in at least one clinical outcome evaluation after random assignment and comprised the intent-to-treat group. Of these 174 patients,

29 (16.7%) did not complete the 8-week trial, leaving 145 patients in the completer group.

Intent-to-Treat Group

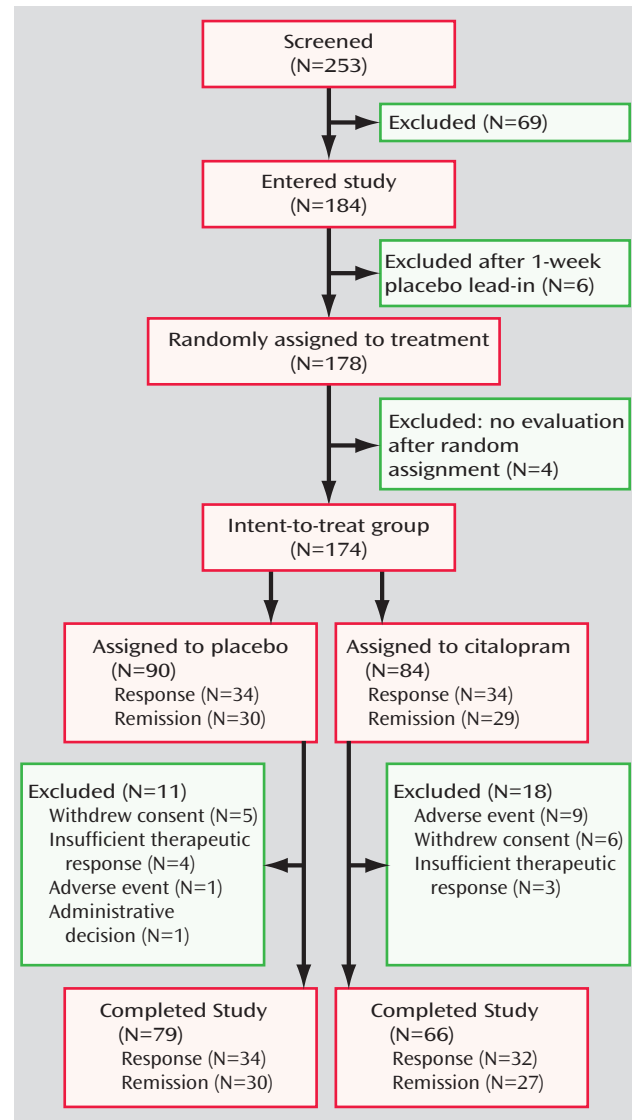
Demographic and clinical features. The sites and treatment groups and sites were compared by clinical and demographic features (Table 1). The treatment groups differed only in baseline MMSE score ($F=4.04$, $df=1$, 156, $p<0.05$), with the citalopram group averaging a less than 1 point higher score than the placebo group. The sites differed in education ($F=5.08$, $df=8$, 155, $p<0.0001$), duration of the current depressive episode ($F=4.05$, $df=8$, 156, $p=0.0002$), baseline score on the posttreatment Montgomery-Åsberg Depression Rating Scale ($F=5.20$, $df=8$, 156, $p<0.0001$), and cumulative medical comorbidities ($F=6.41$, $df=8$, 156, $p<0.0001$). These analyses indicated that there were minimal differences in patient characteristics between the treatment conditions but significant study site differences.

Clinical outcome. In the mixed-model analysis on serial Hamilton depression scale scores, the main effect of treatment group ($F=0.63$, $df=1$, 152, $p=0.43$) and the interaction between treatment group and time point of assessment ($F=0.72$, $df=7$, 1064, $p=0.65$) did not approach significance. Similarly, in the ANCOVAs and logistic regression analyses, there were no differences between the two treatment groups in the extent of improvement on the Hamilton depression scale score or in the rates of response or remission (Table 2). With the exception of the self-report CES-D Scale scores, there were also no main effects of treatment condition in the analyses of the secondary clinical outcome measures (Table 2). There was an advantage for citalopram on the self-report CES-D Scale ($F=7.45$, $df=1$, 148, $p=0.007$), but the average degree of improvement was modest.

The mixed-model analysis yielded a main effect of study site ($F=2.78$, $df=8$, 152, $p<0.007$) and an interaction between study site and the time point of assessment ($F=2.62$, $df=56$, 1064, $p<0.0001$). The site differences in clinical outcome pertained to both the placebo and citalopram conditions, and across all the analyses of the primary and secondary clinical outcome measures, there were no significant interactions involving site and treatment condition. Significant site differences were found for every clinical outcome measure, except for the response rate based on the CGI improvement scale (Table 2). When we combined the two treatment conditions across the study sites, the rates of response ranged from 19% to 82%, and rates of remission ranged from 14% to 73% (Figure 2). Thus, in the intent-to-treat group, with the exception of the self-report CES-D Scale, citalopram did not differ from placebo in antidepressant effects. In contrast, the sites differed markedly in the extent of clinical improvement, and these differences were independent of treatment condition.

In the mixed-model analysis, there was also a main effect of the classification of baseline depression severity ($F=45.32$, $df=1$, 152, $p<0.0001$), an interaction between this

FIGURE 1. Flowchart of a Randomized 8-Week Medication Trial Comparing Citalopram, 10–40 mg/day, to Placebo in the Treatment of Patients Age 75 and Older With Unipolar Depression



factor and treatment condition ($F=3.91$, $df=1$, 152, $p<0.05$), and an interaction among this factor, treatment condition, and time point of assessment ($F=1.95$, $df=14$, 1064, $p<0.02$). A difference between the treatment groups in therapeutic outcome as a function of depression severity classification was also significant in the logistic regression analysis on response rate and was of marginal significance in the analyses of endpoint Hamilton depression scale and CES-D Scale scores and remission rates (Table 3). Post hoc *t* tests on least squares adjusted means indicated that endpoint Hamilton depression scale scores in the high severity group were lower in the patients treated with citalopram compared to placebo ($t=2.37$, $df=66$, $p=0.02$), with no other differences among the four subgroups. The patients treated with placebo in the high severity group had higher endpoint CES-D Scale scores than the patients in

TABLE 1. Demographic and Clinical Characteristics of an Intent-to-Treat Group of Patients Age 75 and Older With Unipolar Depression Receiving Citalopram or Placebo

Variable	Total Study Group (N=174)		Citalopram Group (N=84)		Placebo Group (N=90)	
	Mean	SD	Mean	SD	Mean	SD
Age (years)	79.6	4.4	79.8	4.0	79.3	4.7
Education (years)	13.8	3.3	13.8	2.8	13.7	3.7
Age at onset of mood disorder (years)	68.3	18.4	67.5	20.2	69.1	16.6
Hamilton Depression Rating Scale score	24.3	4.1	24.4	4.3	24.2	3.9
Montgomery-Åsberg Depression Rating Scale score	24.7	5.8	24.4	5.9	25.0	5.9
Clinical Global Impression severity score	4.2	0.5	4.2	0.4	4.3	0.5
Mini-Mental State Examination score	28.0	2.1	28.4	1.6	27.6	2.5
Cumulative Illness Rating Scale total score	7.2	3.9	7.4	3.7	7.1	4.0
	N	%	N	%	N	%
Women	101	58.1	45	53.6	56	62.2
Late-onset depression	133	76.4	64	76.2	69	76.7
Hamilton depression scale severity grouping: severe depression	68	39.1	37	44.1	31	34.4
	Median		Median		Median	
Duration of current episode (months)	13.7		17.7		12.2	

TABLE 2. Clinical Outcome as a Function of Randomized Treatment Conditions for Patients Age 75 and Older With Unipolar Depression Receiving Citalopram or Placebo

Outcome Variable	Citalopram Group (N=84)		Placebo Group (N=90)		Analysis of Treatment Group		Analysis of Study Site	
	Mean	SD	Mean	SD	F	p	F	p
Continuous outcome variables ^a								
Percent change in Hamilton Depression Rating Scale score	39.6	32.9	38.3	31.1	2.92	0.09	2.98	0.004
Percent change in Montgomery-Åsberg Depression Rating Scale score	39.2	34.5	34.0	35.9	2.32	0.13	2.57	0.01
Percent change in Center for Epidemiologic Studies Depression (CES-D) Scale score	25.8	34.8	16.2	48.3	7.45	0.007	1.97	0.05
Categorical outcome variables ^b	N	%	N	%	χ^2	p	χ^2	p
Responders	34	40.5	34	37.8	0.97	0.32	30.4	0.0002
Remitters	29	34.5	30	33.3	0.29	0.59	18.5	0.02
Clinical Global Impression improvement score of 1 or 2	37	44.1	39	43.3	1.53	0.22	8.7	0.36

^a F and p values for the main effect of treatment group from the analyses of covariance of endpoint depression symptom scores (df=1, 151) for the Hamilton depression scale and Montgomery-Åsberg Depression Rating Scale (df=1, 148) for the CES-D Scale. Percent change values for the treatment conditions are not adjusted for covariate or study site effects. F and p values for the main effect of study site from the analyses of covariance on posttreatment scores (df=8, 151) for the Hamilton depression scale and Montgomery-Åsberg Depression Rating Scale and (df=1, 148) for the CES-D Scale.

^b Likelihood ratio chi-square and p values for the main effect of treatment group (df=1) from the logistic regression analyses. Response and remitter rates for the two conditions are not adjusted for covariate or study site effects. Likelihood ratio chi-square and p values for the main effect of study site (df=8) from the logistic regression analyses.

each of the three other subgroups (all $p < 0.05$), with no other pairwise differences. Within the high severity group, the citalopram condition had a higher rate of response relative to placebo ($\chi^2 = 4.03$, $df = 1$, $p = 0.04$), while there was no effect of treatment condition among the patients with low baseline depression severity ($\chi^2 = 0.99$, $df = 1$, $p = 0.32$). Thus, there were indications that citalopram was more effective than placebo in patients with high, but not low, baseline depression severity.

The mixed-model analysis also yielded a significant interaction among the classification of age at onset, treatment condition, and time point of assessment ($F = 1.74$, $df = 14$, 1064, $p = 0.04$). The ANCOVAs revealed differences be-

tween the age-at-onset groups in the extent of improvement with citalopram compared to placebo for the endpoint Hamilton depression scale and CES-D Scale scores (Table 4). Patients with onset of major depression before 60 years of age had poorer outcome when treated with placebo than each of the other three subgroups (all $p < 0.05$).

The mixed-model analysis on the serial Hamilton depression scale scores was repeated, adding as covariates those variables for which there had been site (education, Montgomery-Åsberg Depression Rating Scale score) or treatment group (MMSE score) differences in baseline patient characteristics. Addition of these three covariates did not result in a main effect of treatment group or a first-or-

FIGURE 2. Rates of Response and Remission Among Depressed Patients Age 75 Years and Older Receiving 8 Weeks of Citalopram, 10 to 40 mg/day, and Placebo Across Study Sites

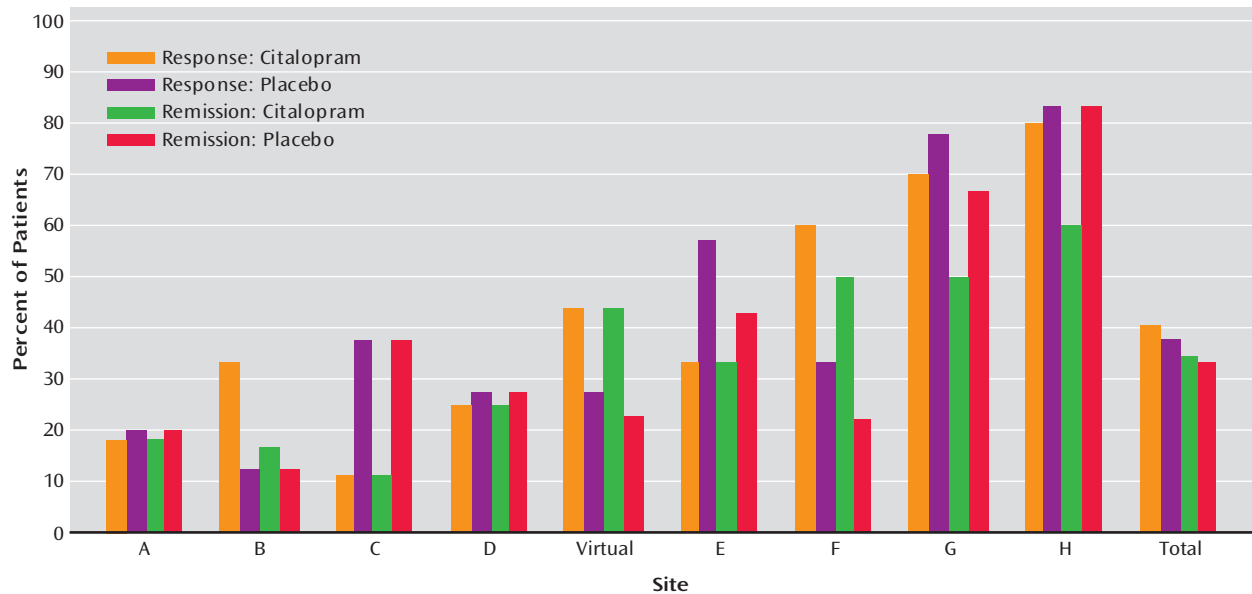


TABLE 3. Clinical Outcome Measures as a Function of Treatment Condition and Depression Severity Classification for Patients Age 75 and Older With Unipolar Depression Receiving Citalopram or Placebo

Outcome Variable	Citalopram Group, Low Severity (N=47)		Placebo Group, Low Severity (N=59)		Citalopram Group, High Severity (N=37)		Placebo Group, High Severity (N=31)		Treatment-by-Severity Grouping	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	F	p
Continuous outcome variables^a										
Percent change in Hamilton Depression Rating Scale score	34.7	34.7	41.4	32.2	45.8	29.7	32.4	28.4	3.62	<0.06
Percent change in Montgomery-Åsberg Depression Rating Scale score	37.0	36.3	36.3	38.4	42.1	32.2	29.8	30.8	0.54	0.46
Percent change in Center for Epidemiologic Studies Depression (CES-D) Scale score	20.9	37.9	18.3	48.5	32.1	29.9	12.0	48.4	2.84	0.09
	N	%	N	%	N	%	N	%	χ^2	p
Categorical outcome variables^b										
Responders	17	36.2	27	45.8	17	46.0	7	22.6	7.05	0.008
Remitters	16	34.0	24	40.7	13	35.1	6	19.4	3.22	0.07
Clinical Global Impression improvement score of 1 or 2	19	40.4	30	50.8	18	48.7	9	29.0	1.98	0.16

^a F and p values for the interaction term of treatment group and depression severity classification from the analyses of covariance of endpoint depression symptom ratings (df=1, 151) for the Hamilton depression scale and Montgomery-Åsberg Depression Rating Scale and (df=1, 148) for the CES-D Scale.

^b Likelihood ratio chi-square and p values for the interaction term of treatment group and depression severity grouping from logistic regression analyses.

der interaction between treatment condition and time point of assessment. Despite statistical control for the variables that showed site differences at baseline, the main effect of site ($F=3.39$, $df=8$, 146 , $p=0.001$) and the interaction between site and time point ($F=2.32$, $df=56$, 1022 , $p<0.0001$) continued to be significant. The effects involving depression severity and age-at-onset classifications were also unaltered.

Functional outcomes. The repeated-measures ANCOVA on the change in Medical Outcomes Study 36-Item Short-Form Health Survey subscale scores yielded no significant effects. Thus, there was no indication that treatment

groups or sites differed with respect to change in functional outcomes. Across the group, significant improvement was observed at posttreatment for all subscales other than physical functioning (all $p<0.05$). The only change that produced an effect size that would be considered moderate or greater occurred with the mental health subscale (mean change=12.3, $SD=20.9$, effect size=0.59) ($t=7.4$, $df=159$, $p<0.0001$). The change in the mental health subscale had moderate associations with the change in the Hamilton depression scale score ($r=0.55$, $df=158$, $p<0.0001$) and CES-D Scale score ($r=0.65$, $df=158$, $p<0.0001$).

TABLE 4. Clinical Outcome Measures as a Function of Treatment Condition and Classification of Age at Onset of Major Depression for Patients Age 75 and Older With Unipolar Depression Receiving Citalopram or Placebo

Outcome Variable	Citalopram Group, Not Late Onset (N=20)		Placebo Group, Not Late Onset (N=21)		Citalopram Group, Late Onset (N=64)		Placebo Group, Late Onset (N=69)		Treatment-by- Onset Grouping	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	F	p
Continuous outcome variables^a										
Percent change in Hamilton Depression Rating Scale score	40.8	34.6	22.5	30.9	39.2	32.6	43.1	29.7	5.64	0.02
Percent Change in Montgomery-Åsberg Depression Rating Scale score	40.9	33.1	24.6	33.6	38.7	35.1	36.9	36.3	2.55	0.11
Percent Change in Center for Epidemiologic Studies Depression (CES-D) Scale score	31.1	37.0	-0.9	50.1	24.1	34.3	21.5	46.8	8.06	0.005
Categorical outcome variables^b										
	N	%	N	%	N	%	N	%	χ^2	p
Responders	8	40.0	5	23.8	26	40.6	29	42.0	0.66	0.41
Remitters	6	30.0	4	19.1	23	35.9	26	37.7	0.36	0.55
Clinical Global Impression improvement score of 1 or 2	10	50.0	5	23.8	27	42.2	34	49.3	2.81	0.09

^a F and p values for the interaction term of treatment group and age at onset classification from the analyses of covariance on endpoint depression symptom ratings (df=1, 151) for the Hamilton depression scale and Montgomery-Åsberg Depression Rating Scale and (df=1, 148) for the CES-D Scale.

^b Likelihood ratio chi-square and p values for the interaction term of treatment group and depression severity grouping from logistic regression analyses.

TABLE 5. Side Effects Reported in at Least 5% of Either Treatment Group for Patients Age 75 and Older With Unipolar Depression Receiving Citalopram or Placebo

Side Effect	Citalopram Group		Placebo Group	
	N	%	N	%
Asthenia	8	9.2	6	6.7
Headache	10	11.5	4	4.4
Dizziness	8	9.2	7	7.8
Constipation	10	11.5	4	4.4
Diarrhea	13	14.9	6	6.7
Dry mouth	6	6.9	12	13.3
Dyspepsia	9	10.3	3	3.3
Nausea	6	6.9	6	6.7
Somnolence	5	5.7	4	4.4
Insomnia	3	3.4	5	5.6
Rash	3	3.4	6	6.7

Side effects and dropout. Side effects reported in at least 5% of the group randomly assigned to either citalopram (N=87) or placebo (N=90) are listed in Table 5. Treatment groups did not differ in the rates of any treatment-emergent side effect.

Of the 174 patients in the intent-to-treat efficacy group, 29 (16.7%) patients left the trial prematurely. The rate of dropout did not differ significantly for the citalopram (21.4%) and placebo (12.2%) groups ($\chi^2=2.65$, df=1, $p=0.10$) or by site. Treatment groups differed in only one source of dropout; despite no difference in rates of individual side effects, early termination due to an adverse event was more common in the citalopram (10.7%) than placebo (1.1%) group ($p<0.008$, Fisher's exact test).

Completer Group

The final mixed model applied in the analysis of serial Hamilton depression scale scores in the intent-to-treat group was reapplied to the completer group. There were no changes in the results.

Discussion

This randomized, placebo-controlled trial of antidepressant medication was conducted in the oldest group (mean age 80 years, SD=4) and with the highest rating of medical illness (mean Cumulative Illness Rating Scale score=7.2, compared to scores of 4.2 and 4.8 in late-life depression studies with a mean age of 68–70 years) studied to date, as far as we know. In this comparison of placebo to citalopram, there were no significant drug-placebo differences in outcome with respect to change in Hamilton depression scale score, response, or remission rates. This was an 8-week trial, and it could be reasonably argued that if the trial had been extended to 12 weeks, the remission rates might be considerably higher. However, there is no reason to presume that a significant difference between medication and placebo would become manifest if the trial were of longer duration. Of note is that the medication was well tolerated in this group of older patients with significant medical comorbidity. The treatment groups did not differ in the frequency of any treatment-emergent side effect, and there was no difference in the rate of dropout. However, early termination due to an adverse event was more common in citalopram- than placebo-treated patients.

The planned secondary analyses with respect to severity and late-onset depression produced significant and intriguing results. With respect to severity, a comparison of citalopram versus placebo in patients classified as having severe (Hamilton depression scale score ≥ 24) versus not severe (Hamilton depression scale score < 24) depression, showed a significant drug-placebo difference favoring medication in the patients with severe depression. This result is consistent with other antidepressant treatment tri-

als in late-life depression that have also reported that medication is significantly more effective than placebo in the severe depression group, with "severe" defined differently in each study (7, 11). However, the statistically significant difference between drug and placebo did not result from an increased response to medication in the severe versus nonsevere group but rather from a decreased response to placebo in the severe group. Therefore, it is more accurate to state that patients with severe depression do not respond as well to placebo as patients with less severe depression rather than that patients with severe depression respond better to medication than patients with nonsevere depression. Although the theoretical implications of the difference in response rates in the severe versus nonsevere group depend on whether the difference in treatment response results from increased response to medication versus a decreased response to placebo, in the clinical situation, the conclusion is simply that an antidepressant medication is a more efficacious treatment than placebo for patients with severe depression.

The other planned secondary analysis compared response rates of patients with late-onset (over the age of 60) to early-onset depression. There have been a number of reports that patients with late-onset depression have a lower response rate to antidepressant medications than patients with early-onset depression, perhaps reflecting the presumed vascular etiology of late-onset depressive illness (27, 29). The unanticipated finding in this study was that there was no significant difference in treatment response to either medication or placebo in the late-onset versus early-onset groups. This result suggests that it is premature to draw conclusions about differences in medication response depending on age at onset of depressive illness and that resolution of this issue awaits prospective clinical trials.

Perhaps the most striking finding in this study was the marked site variability in response rates of both drug and placebo. Across sites, the medication response ranged from 18% to 82% and placebo response from 16% to 80%. In most sites, there was no drug-placebo difference; whether a patient responded or not depended on the site where they were treated rather than what treatment they received. A possible reason for site variability is the significant differences in the patient cohorts treated at the various sites despite the standardization of diagnostic criteria, the methods of assessment, and the establishment of interrater reliability. There were some minimal differences in the patient characteristics across sites but nothing in either character or magnitude that would reasonably explain the marked site variability in response. Although the primary outcome measure was the Hamilton depression scale, intriguingly, the outcome evaluated by self-report using the CES-D Scale did not demonstrate significant differences between sites. This may suggest that in spite of ongoing efforts to attain and maintain reliability across sites, a component of the differences in outcomes be-

tween centers may be related to the process of assessment. This finding, together with the greater significance of the differences between treatment groups, suggests the utility of self-reports as measures of outcome in this group and the need for further research comparing self-reports and clinician-rated outcomes. Although perhaps disconcerting from a methodological perspective, the finding of significant site variability is not surprising. Site variability has been the rule rather than the exception in other multicenter trials of medication treatment for late-life depression (30). The frequency and impact of site variability and the lack of a cogent explanation for this phenomenon underscores that we are not aware of all the moderators and mediators that significantly affect outcome in treatment trials in late-life depression.

In this trial and in the other placebo-controlled trials in late-life depression, the placebo condition is not remotely close to a "no-treatment" condition. In addition to receiving the medication or placebo in the context of medication management, as described by Fawcett et al. (31), other components of this clinical trial received by patients included 1) paid transportation to and from appointments (most sites); 2) a free medical workup, including a physical examination, an ECG, measures of serum chemistries and electrolytes, a CBC, a thyroid profile, and measures of folate and B₁₂ levels; 3) a free MRI; 4) a free neuropsychological evaluation; 5) free medications; 6) weekly visits with raters, nurses, and research assistants (averaging about 30 minutes); and 7) at most sites, free treatment for 3 months after completion of the study. Thus, it would be incorrect to interpret the results of a placebo-controlled medication trial in which no difference was found between the two treatments as indicating that treatment with medication was equivalent to no intervention. There are many aspects of the treatment protocol in a placebo-controlled trial that may have therapeutic effects and are obviously different from "no intervention" or standard clinical care, e.g., the frequency and duration of visits, free medical evaluation, free medication. Furthermore, it may be that patients willing to enter a placebo-controlled trial at this point in time, when there are many antidepressants available that are approved by the Food and Drug Administration, are a nonrepresentative group.

Nonetheless, the results of the study further support the necessity of doing placebo-controlled trials. Ironically, despite the absence of data on the "old-old," many investigators in this study had a difficult time convincing their internal review boards that it was not only ethical but necessary to do placebo-controlled trials in this population. The results of this study and the features of placebo-controlled trials that limit the generalizability of the results lead to the conclusion that it is not whether placebo-controlled trials need to be done but whether they can be designed so that the results are more clinically relevant.

How should the results of this study, to our knowledge, the only randomized clinical trial, placebo or comparator,

in depressed patients over 75, affect clinical treatment? There are some treatments proven effective for depression in the old-old, notably some specific psychotherapies, ECT, and tricyclic antidepressants. There is considerable evidence from clinical trials supporting the efficacy of specific psychotherapies—cognitive behavior therapy (32), interpersonal psychotherapy (33), and problem solving (34)—for the acute treatment of late-life depression. Furthermore, interpersonal psychotherapy has been shown to enhance the effectiveness of maintenance medication in the prevention of recurrence in depressed patients over the age of 70 years (35). ECT remains one of the safest and most effective treatments for depression, and the response rate is not adversely affected by increasing age (36). New methods for the administration of ECT have minimized cognitive side effects, and this treatment is probably underutilized in older patients (37). However, even effective ECT treatment requires continuation medication to sustain response (38). Thus, there is a compelling need to establish a safe and effective antidepressant medication treatment in the old-old. The side effect profile of tricyclics, particularly with respect to anticholinergic effects and their deleterious cardiovascular properties, limits their use in an older population (39).

Are there sufficient data to support the efficacy of other classes of antidepressants in the older-age population? To date, two of the four placebo-controlled trials of SSRIs in late-life depression do not distinguish drug from placebo, and in the other two, remission rates were below 30%. Although not robust, nonetheless, these results are comparable to outcomes in antidepressant trials in younger depressed patients. This underscores the necessity for more studies rather than the premature conclusion that the SSRI medications are not clinically useful in this patient population. With respect to other classes of medications, there has been one placebo-controlled trial of venlafaxine, in which neither fluoxetine nor venlafaxine were superior to placebo, and one comparator study of mirtazapine and paroxetine, in which the medications were comparably effective. The distressing fact is that despite the prevalence and terrible deleterious effect of depression in older patients, the evidence on which to base treatment recommendations is appallingly limited. Perhaps the single most important conclusion to be drawn from the results of this study is that there is a compelling need to do a more systematic study of antidepressant treatments, including brain stimulation, medications, and psychotherapies, in this understudied population.

Received Oct. 9, 2003; revision received Jan. 24, 2004; accepted Feb. 12, 2004. From the New York State Psychiatric Institute, Departments of Psychiatry and Radiology, College of Physicians and Surgeons, Columbia University; the Department of Psychiatry and Behavioral Sciences, Duke University, Durham, N.C.; the Academic Division of Geriatric Psychiatry, University of Pittsburgh, Pittsburgh; Cornell Institute of Geriatric Psychiatry, Weill Medical College of Cornell University, White Plains, N.Y.; UCLA-Neuropsychiatric Institute & Hospital, David Geffen School of Medicine at UCLA, Los Angeles; the

Department of Psychiatry, University of Pennsylvania, Philadelphia; and the Mental Illness Research Education and Clinical Center, Forest Research Institute, Philadelphia VA Medical Center, Philadelphia. Address reprint requests to Dr. Roose, New York State Psychiatric Institute, 1051 Riverside Dr., New York, NY 10032; spr2@columbia.edu.

Supported by a grant from Forest Laboratories.

The members of the Old-Old Depression Study Group include William Burke, M.D., George T. Grossberg, M.D., Howard Hassman, D.O., Alan F. Jacobson, Ph.D., Louis C. Kirby II, M.D., Benoit H. Mulsant, M.D., J. Craig Nelson, M.D., Charles F. Reynolds III, M.D., Robert Riesenber, M.D., Stephen Salloway, M.D., Carl Salzman, M.D.

References

1. United States Census Report. Washington, DC, US Bureau of the Census, 2002
2. Frasure-Smith N, Lesperance F, Talajic M: Depression and 18-month prognosis after myocardial infarction. *Circulation* 1995; 91:999–1005
3. DeVane CL, Pollock BG: Pharmacokinetic considerations of antidepressant use in the elderly. *J Clin Psychiatry* 1999; 60(suppl 20):38–44
4. Pollock BG: Drug interactions, in *Geriatric Psychopharmacology*. Edited by Nelson JC. New York, Marcel Dekker, 1998, pp 43–60
5. Pollock BG: Adverse reactions of antidepressants in elderly patients. *J Clin Psychiatry* 1999; 60(suppl 20):4–8
6. Tollefson GD, Bosomworth JC, Heiligenstein JH, Potvin JH, Holman S (Fluoxetine Collaborative Study Group): A double-blind, placebo-controlled clinical trial of fluoxetine in geriatric patients with major depression. *Int Psychogeriatr* 1995; 7:89–104
7. Schneider LS, Nelson JC, Clary CM, Newhouse P, Krishnan KRR, Shioyitz T, Weihs K (Sertraline Elderly Depression Study Group): An 8-week multicenter, parallel-group, double-blind, placebo-controlled study of sertraline in elderly outpatients with major depression. *Am J Psychiatry* 2003; 160:1277–1285
8. Schatzberg AF, Cantillon M: Antidepressant early response and remission of venlafaxine and fluoxetine in geriatric outpatients. *Eur Neuropsychopharmacol* 2000; 10:S225–S226
9. Bondareff W, Alpert M, Friedhoff AJ, Richter EM, Clary CM, Batzar E: Comparison of sertraline and nortriptyline in the treatment of major depressive disorder in late life. *Am J Psychiatry* 2000; 157:729–736
10. Newhouse PA, Krishnan KRR, Doraiswami PM, Richter EM, Batzar ED, Clary CM: A double blind comparison of sertraline and fluoxetine in depressed elderly outpatients. *J Clin Psychiatry* 2000; 61:559–568
11. Schatzberg AF, Kremer C, Rodrigues HE, Murphy GM Jr (Mirtazapine vs Paroxetine Study Group): Double-blind randomized comparison of mirtazapine and paroxetine in elderly depressed patients. *Am J Geriatr Psychiatry* 2002; 10:541–550
12. Mulsant BH, Pollock BG, Nebes R, Miller MD, Sweet RA, Stack J, Houch PR, Bensasi S, Muzumdar S, Reynolds CF: A twelve-week, double-blind randomized comparison of nortriptyline and paroxetine in older depressed patients and outpatients. *Am J Geriatr Psychiatry* 2001; 9:406–414
13. Nelson JC, Kennedy JS, Pollock BG, Laghrissi-Thode F, Narayan M, Nobler MS, Robin DW, Gergel I, McCafferty J, Roose S: Treatment of major depression with nortriptyline and paroxetine in patients with ischemic heart disease. *Am J Psychiatry* 1999; 156:1024–1028
14. Roose S, Schatzberg A: The efficacy of antidepressants in the treatment of late-life depression: evidence-based conclusions. *J Clin Psychopharmacol* (in press)
15. Streim JE, Oslin DW, Katz IR, Smith BD, DiFilippo S, Cooper TB, Ten Have T: Drug treatment of depression in frail elderly nursing home residents. *Am J Geriatr Psychiatry* 2000; 8:150–159

16. Trappler B, Cohen CI: Use of SSRIs in "very old" depressed nursing home residents. *Am J Geriatr Psychiatry* 1998; 6:83–89
17. Montgomery SA, Åsberg M: A new depression scale designed to be sensitive to change. *Br J Psychiatry* 1979; 134:382–389
18. Radloff LS: The CES-D scale: a self-report depression scale for research in the general population. *Appl Psychol Measurement* 1977; 1:385–401
19. Watson D, Clark LA, Tellegen A: Development and validation of brief measures of positive and negative affect: the PANAS scales. *J Per Soc Psychol* 1988; 54:1063–1070
20. Ware JE, Sherbourne CD: The MOS 36-item short-form health survey (SF-36): I. Conceptual framework and item selection. *Med Care* 1992; 30:473–483
21. Katz S, Ford AB, Moskowitz RW, Jackson BA, Jaffe MW: Studies of illness in the aged; the index of ADL: a standardized measure of biological and psychosocial function. *JAMA* 1963; 185: 914–919
22. Brown H, Prescott R: *Applied Mixed Models in Medicine*. New York, John Wiley & Sons, 1999
23. Gibbons RD, Hedeker D, Elkin I, Waternaux C, Kramer HC, Greenhouse JB, Shea T, Imber SD, Sotsky SM, Watkins JT: Some conceptual and statistical issues in analysis of longitudinal psychiatric data: application to the NIMH Treatment of Depression Collaborative Research Program dataset. *Arch Gen Psychiatry* 1993; 50:739–750
24. Lavorie P: Clinical trials in psychopharmacology: should protocol deviation truncate patient data? *Neuropsychopharmacology* 1992; 6:39–48
25. Klein DF, Ross DC: Reanalysis of the National Institute of Mental Health Treatment of Depression Collaborative Research Program general effectiveness report. *Neuropsychopharmacology* 1993; 8:241–251
26. Glassman AH, O'Connor CM, Califf RM, Swedberg J, Schwartz P, Bigger JT Jr, Krishnan KRR, van Zyl LT, Swenson JR, Finkel MS, Landau C, Shapiro PA, Pepine CJ, Mardekian J, Harrison WM, Barton D, McIvor M (Sertraline Antidepressant Heart Attack Randomized Trial [SADHART] Group): Sertraline treatment of major depression in patients with acute MI or unstable angina. *JAMA* 2000; 288:701–709; correction, 288:1720
27. Figiel GS, Krishnan KRR, Doraiswamy PM, Rao VP, Nemeroff CB, Boyko OB: Subcortical hyperintensities on brain magnetic resonance imaging: a comparison between late age onset and early onset elderly depressed subjects. *Neurobiol Aging* 1991; 12:245–247
28. Ware JE Jr: *SF-36 Health Survey: Manual and Interpretation Guide*. Boston, New England Medical Center, Health Institute, 1993
29. Hickie I, Scott E, Mitchell P, Wilhelm K, Austin MP, Bennett B: Subcortical hyperintensities on magnetic resonance imaging: clinical correlates and prognostic significance in patients with severe depression. *Biol Psychiatry* 1995; 37:151–160
30. Small GW, Schneider LS, Hamilton SH, Bystritsky A, Meyers BS, Nemeroff CS (Fluoxetine Collaborative Study Group): Site variability in a multisite geriatric depression trial. *Int J Geriatr Psychiatry* 1996; 11:1089–1095
31. Fawcett J, Epstein P, Fiester SJ, Elkin I, Autry JH: Clinical management—imipramine/placebo administration manual: NIMH Treatment of Depression Collaborative Research Program. *Psychopharmacol Bull* 1987; 23:309–324
32. Arean PA, Cook BL: Psychotherapy and combined psychotherapy/pharmacotherapy for late-life depression. *Biol Psychiatry* 2002; 52:292–303
33. Reynolds CF III, Frank E, Perel JM, Imber SD, Cornes C, Miller MD, Mazumdar S, Houck PR, Dew MA, Stack JA, Pollock BG, Kupfer DJ: Nortriptyline and interpersonal psychotherapy as maintenance therapies for recurrent major depression: a randomized controlled trial in patients older than 59 years. *JAMA* 1999; 281:39–45
34. Williams JW, Barrett J, Oxman T, Frank E, Katon W, Sullivan M, Cornell J, Sengupta A: Treatment of dysthymia and minor depression in primary care: a randomized controlled trial in older adults. *JAMA* 2000; 284:1519–1526
35. Reynolds CF, Frank E, Dew MA, Houck PR, Miller MD, Mazumdar S, Perel JM, Kupfer DJ: Treatment in 70+ year olds with major depression: excellent short-term but brittle long-term response. *Am J Geriatr Psychiatry* 1999; 7:64–69
36. Tew JD Jr, Mulsant BH, Haskett RF, Prudic J, Thase ME, Crowe RR, Dolata D, Begley AE, Reynolds CF III, Sackeim HA: Acute efficacy of ECT in the treatment of major depression in the old-old. *Am J Psychiatry* 1999; 156:1865–1870
37. Sackeim HA, Prudic J, Devanand DP, Nobler MS, Lisanby SH, Peyser S, Fitzsimons L, Moody BJ, Clark J: A prospective, randomized, double-blind comparison of bilateral and right unilateral electroconvulsive therapy at different stimulus intensities. *Arch Gen Psychiatry* 2000; 57:425–434
38. Sackeim HA, Haskett RF, Mulsant BH, Thase ME, Mann JJ, Pettinati HM, Greenberg RM, Crowe RR, Cooper TB, Prudic J: Continuation pharmacotherapy in the prevention of relapse following electroconvulsive therapy: a randomized controlled trial. *JAMA* 2001; 285:1299–1307
39. Roose S: Treatment of depression in patients with heart disease. *Biol Psychiatry* 2003; 54:262–268