Citalopram for Acute and Preventive Efficacy in Bipolar Depression (CAPE-BD):

A Randomized, Double-Blind, Placebo-Controlled Trial

S. Nassir Ghaemi, MD, MPH[a,b,*]; Elizabeth A. Whitham, MD[c]; Paul A. Vohringer, MD, MPH[a,d]; Sergio A. Barroilhet, MD[a,d]; Andrea Amerio, MD[a,e,f]; Oleksandr Sverdlov, PhD[g]; and Ashwin A. Patkar, MD[h]

ABSTRACT

Objective: To assess the efficacy and safety of citalopram in the acute and maintenance phases of bipolar depression in a randomized, double-blind, placebo-controlled trial.

Methods: Between 2007 and 2014, 119 subjects with acute major depressive episodes diagnosed with *DSM-IV* bipolar disorder, type I or type II, were randomized blindly to citalopram or placebo, added to standard mood stabilizers. They were followed for 6 weeks for acute efficacy (primary outcome) and up to 1 year for maintenance efficacy (secondary outcome) using scores on the Montgomery-Asberg Depression Rating Scale (MADRS) and the Mania Rating Scale of the Schedule for Affective Disorders and Schizophrenia (MRS-SADS). The study was powered for a clinically meaningful effect size.

Results: Mean ± SD MADRS scores changed from a baseline value of 27.4 ± 9.1 to 13.1 ± 8.4 at the end of the acute phase for citalopram versus a change from 27.4 ± 7.3 to 15.2 ± 9.9 for placebo, a clinically and statistically nonsignificant difference. Maintenance efficacy also was not better with citalopram than with placebo. Acute manic/hypomanic episodes were similar in both groups, and subjects with type II illness did not have better outcomes than subjects with type I illness. In maintenance treatment, MRS-SADS scores were greater overall, especially in subjects with a rapid-cycling illness course, with citalopram versus placebo.

Conclusions: Citalopram, added to standard mood stabilizers, did not have clinically meaningful benefit versus placebo for either acute or maintenance treatment of bipolar depression. Acute mania did not worsen with citalopram, but maintenance treatment led to worsened manic symptoms, especially in subjects with a rapid-cycling course.

Trial Registration: ClinicalTrials.gov identifier: NCT00562861

J Clin Psychiatry 2021;82(1):19m13136

To cite: Ghaemi SN, Whitham EA, Vohringer PA, et al. Citalopram for Acute and Preventive Efficacy in Bipolar Depression (CAPE-BD): a randomized, double-blind, placebo-controlled trial. *J Clin Psychiatry*. 2021;82(1):19m13136.

To share: https://doi.org/10.4088/JCP.19m13136

[a]Department of Psychiatry, Tufts University School of Medicine, Boston, Massachusetts

[b]Department of Psychiatry, Harvard Medical School, Boston, Massachusetts

[c]Department of Psychiatry, Louisiana State University, New Orleans, Louisiana

[d]Department of Psychiatry, Universidad of Chile, Santiago, Chile

[e]Department of Neuroscience, Rehabilitation, Ophthalmology, Genetics, Maternal and Child Health, Section of Psychiatry, University of Genoa, Italy

[f]IRCCS Ospedale Policlinico San Martino, Genoa, Italy

[g]Novartis Institutes for Biomedical Research, Cambridge, Massachusetts

[h]Department of Psychiatry, Duke University Medical Center, Durham, North Carolina

**Corresponding author:* S. Nassir Ghaemi, MD, MPH, Department of Psychiatry, Tufts Medical Center, 800 Washington St, Boston, MA 02111 (Nassir.Ghaemi@tufts.edu).

Bipolar depression is recognized to be common, deadly, and difficult to treat.[1] Antidepressants are the most widely used class of psychotropic medications for bipolar illness.[2] Not only clinicians, but also bipolar experts often recommend their use for bipolar depression.[3] However, the most recent meta-analysis of randomized clinical trials (RCTs) of antidepressants for bipolar depression[4] found them to be ineffective, and some bipolar experts have urged caution.[5] In addition, meta-analyses[6] of maintenance treatment of depressive episodes also failed to find evidence of preventive efficacy, though most studies involved older agents.

Besides the question of efficacy for acute depressive episodes in bipolar illness, the risk of induction of acute mania has been much discussed,[7] with some evidence that such harm can happen, more with older tricyclic antidepressants[8] than with newer serotonin reuptake inhibitors (SRIs).[9,10] There is evidence of long-term worsening of bipolar illness, with more mood episodes over time, related to maintenance treatment with antidepressants, primarily in patients with a rapid-cycling course of illness.[11,12]

The CAPE-BD (Citalopram for Acute and Preventive Efficacy in Bipolar Depression) study represents the first placebo-controlled RCT of citalopram in acute bipolar depression at 6 weeks after treatment initiation and the first placebo-controlled RCT of any SRI in maintenance prevention of depressive episodes in bipolar illness at 1-year follow-up.

METHODS

The methods, including those described in Supplementary Appendix 1, are reported following the guidelines of the Consolidated Standards of Reporting Trials (CONSORT; Figure 1). Supplementary Appendix 1 provides details on implementation of the trial. More study details, including exclusion criteria, can be found on www.clinicaltrials.gov (identifier NCT00562861).

Study Population

Male and female subjects aged 18–64 years were recruited between 2007 and 2014 if they had (*a*) a diagnosis of bipolar disorder (type I, or type II) and (*b*) a current major depressive episode lasting 8 weeks or longer (all diagnosis were based on *DSM-IV* criteria

Clinical Points

- Clinicians and patients often choose antidepressants, especially serotonin reuptake inhibitors (SRIs), to treat bipolar depression, but the evidence for benefit and safety of these medications has been poorly proven or controversial.
- SRIs like citalopram are not helpful to treat bipolar depression or to prevent it, and they may worsen manic symptoms if used long-term, especially in patients with a rapid-cycling course.
- Antidepressants should be avoided in bipolar depression.

using the Structured Clinical Interview for *DSM-IV* mood modules[13]).

Subjects had to either agree to start standard mood stabilizers (lithium, divalproex, carbamazepine, or lamotrigine) or already be taking mood stabilizers for at least 4 weeks prior to study entry. Therapeutic serum levels were assessed. Mood stabilizers were continued unchanged if already given or were chosen and dosed by trial investigators—based on clinical judgment and patient preference—blinded to study treatment allocation. Other ongoing psychotropic treatments were allowed to be continued unchanged. Any change in any non-study medication except benzodiazepines was not allowed, and if such a change was needed, it was a reason for study termination.

Outcomes

Outcomes were assessed in weekly intervals for 6 weeks from randomization (acute phase) and in monthly intervals for 1 year from completion of the acute phase (maintenance phase). All outcomes listed in this section were prespecified before analysis.

The primary acute endpoint was Montgomery-Asberg Depression Rating Scale (MADRS)[14] scores over 6 weeks. Secondary acute outcomes were (*a*) 50% decline in the MADRS score (clinical response) and (*b*) MADRS score ≤7 (clinical remission). The main outcomes assessed at the maintenance endpoint were scores over 12 months on the MADRS and the Mania Rating Scale from the Schedule for Affective Disorders and Schizophrenia[15] (MRS-SADS). Clinical response and remission were defined in the maintenance phase as they were in the acute phase, using repeated-measures analysis as defined in the Statistical Analysis section.

Other a priori secondary outcomes were (*a*) the aforementioned outcomes stratified by bipolar disorder type I versus type II diagnostic status, (*b*) occurrence of *DSM-IV*–defined manic/hypomanic episodes during follow-up stratified by randomization status or changes in MRS-SADS scores, and (*c*) the aforementioned outcomes stratified by rapid-cycling status. These subgroups were expected to be underpowered, and thus descriptive statistics with relative risks and confidence intervals were reported primarily.

Patients were randomized on day 1 to the entire study of 1-year duration. The 6-week endpoint was only for analysis purposes. Patients were not re-randomized at 6 weeks, nor was there a second decision point about continuation in the study at 6 weeks. The rationale for not re-randomizing patients at 6 weeks of treatment is that the current design is the most valid and conservative assessment of maintenance efficacy in prevention of mood episodes.[16] The commonly used randomized discontinuation design involves preselecting only acute treatment responders for assessment of prevention. That design has been criticized for being invalid due to bias toward treatment responders, as reflected in acute withdrawal effects.[16] Non-enriched maintenance designs, as are standard in most clinical medicine studies (eg, cardiology), typically involve simple continuation after the acute phase, which does not preselect or bias results in favor of acute treatment responders only.[16]

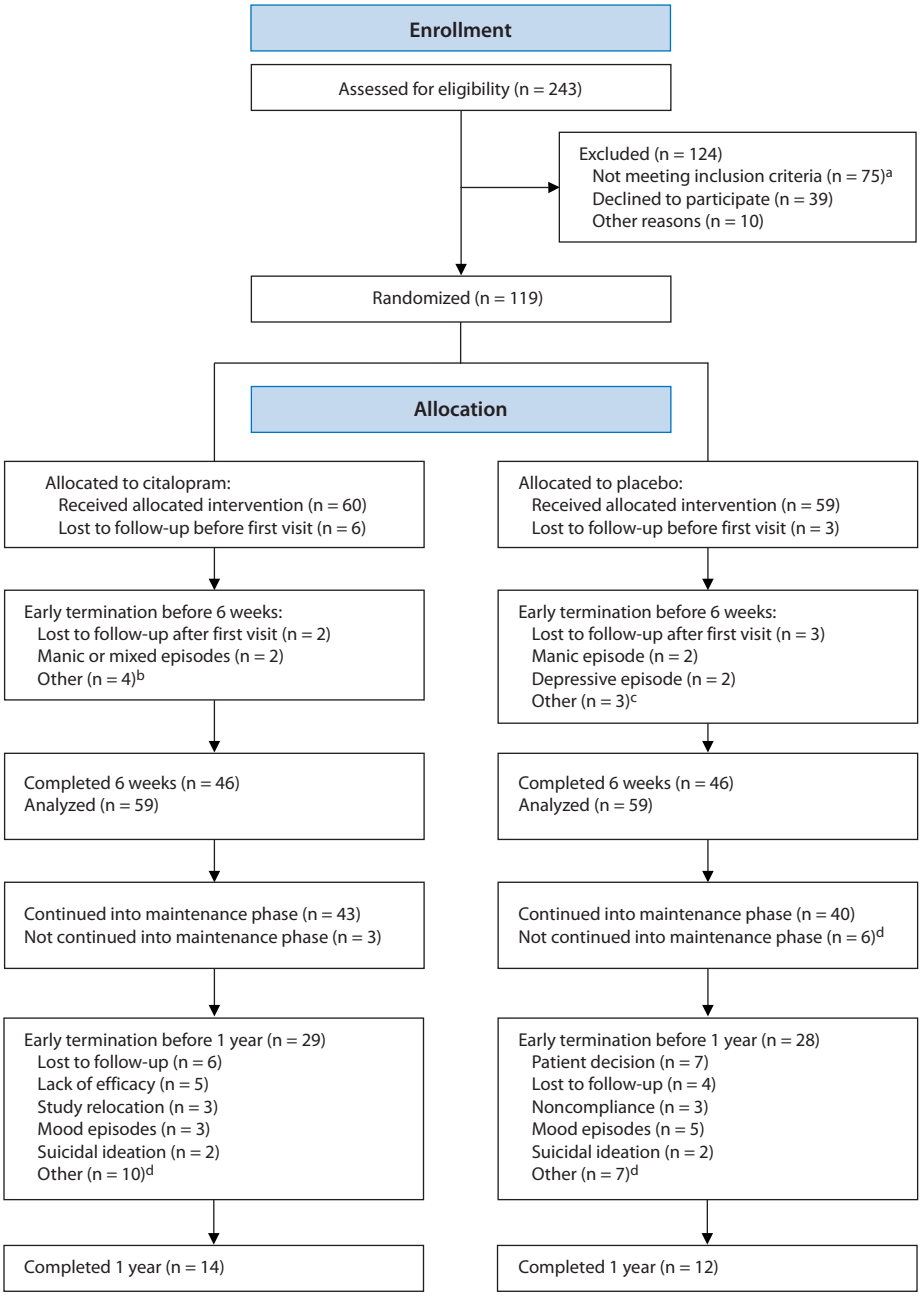
Clinical Assessment

Randomized subjects were followed up with a goal of being seen on a routine basis weekly in the acute phase and monthly in the maintenance phase. There was visit flexibility so patients could be seen more or less frequently depending on clinical needs. Baseline demographic and clinical characteristics were assessed using a standardized form. Diagnosis was based on the Structured Clinical Interview for *DSM-IV* Axis I Disorders Patients Edition (SCID),[13] and the MADRS and the MRS-SADS were administered at each visit.

Sample Size

Sample size estimation was based on standard minimal clinically meaningful differences, ie, moderate effect sizes, seen for acute bipolar depression treatment. Power analysis indicated that data from 148 subjects (74 per treatment group) would provide 80% power to detect statistically significant difference between 2 groups at 2-sided α of .05 using a 2-sample *t* test. This calculation assumed the true mean group difference in MADRS scores of 4 points and the common standard deviation of 9 points (ie, standardized mean group difference Cohen $d = 0.44$). This MADRS difference was based on effect sizes in the range of standard randomized clinical trials of agents that received registration approval for bipolar depression.[17,18] It also is slightly lower than the standard Cohen *d* effect size of 0.5, which is generally seen as moderate and corresponding to a minimal clinically meaningful effect.[19] It further corresponds to accepted definitions of clinically meaningful effect sizes in absolute change with standard depression rating scales, such as in the guidelines of the National Institute for Clinical Excellence (NICE) in the United Kingdom.[20] The total sample goal of 148 subjects was not met fully in over 5 years of recruitment, but with 119 subjects the observed effect size was too small to be clinically meaningful or statistically significant even had the full sample been attained.

Figure 1. CONSORT Diagram



[a]Not in current depressive episode or euthymic: n = 22, current mixed episode: n = 10, current manic/hypomanic episode: n = 6, non-bipolar diagnosis: n = 8, age > 70 years: n = 2, failed citalopram in past or already taking citalopram: n = 12, pregnant: n = 1, current substance abuse: n = 2, unable to give consent due to cognitive impairment: n = 1, other: n = 11.
[b]Each n = 1: completed suicide, unable to tolerate the required concomitant mood stabilizer, hospitalized with PTSD-related flashbacks unrelated to current mood state, severe nausea.
[c]Each n = 1: need for surgery unrelated to the study, severe nausea, patient decision.
[d]See Supplementary Appendix 1 for details.
Abbreviation: PTSD = posttraumatic stress disorder.

Statistical Analysis

Clinical and demographic measures between treatment groups were examined to show comparability at baseline. The intent-to-treat (ITT) sample includes all randomized subjects with at least 1 observation of the primary endpoint (MADRS, n = 118). One subject, randomized to citalopram, had no MADRS data and so was not part of the ITT sample. Effect sizes along with their 95% confidence intervals (CIs) are reported. Proportions of subjects who achieved response or remission between treatment arms were compared using Fisher exact test statistics at endpoint. The main statistical analysis involved linear mixed-effects repeated-measures

models. Further detail is provided in Supplementary Appendix 2.

RESULTS

Clinical and demographic characteristics of the sample (Table 1) indicate that patients were middle-aged, mostly female, about half college-educated or higher, and mostly Caucasian (European-American). Almost two-thirds of the sample had a diagnosis of bipolar type I illness. The most commonly used mood stabilizer was lithium. A total of 28% of patients had rapid-cycling features, and 53% reported past substance abuse. In most cases, no notable differences between the 2 arms were present. The mean ± SD dose in the acute phase was 26.9 ± 15.5 mg/d and in the maintenance phase was 28.4 ± 17.1 mg/d.

Primary Outcome: Acute Phase Efficacy

At the end of the acute phase, raw mean ± SD MADRS scores changed from a baseline value of 27.4 ± 9.1 to 13.1 ± 8.4 for citalopram versus from 27.4 ± 7.3 to 15.2 ± 9.9 for placebo, indicating a 2.1-point difference in MADRS score in favor of citalopram. With adjustment for time and baseline severity in a mixed-effects regression model, citalopram decreased MADRS scores by a mean of 1.7 points overall compared with placebo.

As seen in Figure 2 (and Supplementary Appendix 2), the main acute outcome (MADRS score for 6 weeks) was compared between citalopram and placebo groups over time using a linear mixed-effects model with repeated measures. As noted in Figure 2, both groups improved notably, with a mean ± SD overall improvement in raw MADRS scores of 10.3 ± 12.5 points in the total sample. In this primary outcome, MADRS scores exhibited significant changes over time in both groups ($P < .001$), and rapid-cycling status also predicted outcome ($P = .04$), but the small effect size of difference between the groups was not statistically significant ($P = .17$). Further, there were no significant differential temporal changes between drug and placebo groups when interacted with time (treatment-by-visit interaction, $P = .12$).

Acute treatment response (50% reduction from baseline MADRS scores) rate was 48.3% (29/60) for citalopram versus 45.8% (27/59) for placebo ($P = .85$). Treatment remission (final MADRS scores below 7) rate was 31.7% (19/60) for citalopram versus 27.1% (16/59) for placebo ($P = .69$).

Secondary Outcome: Maintenance Efficacy

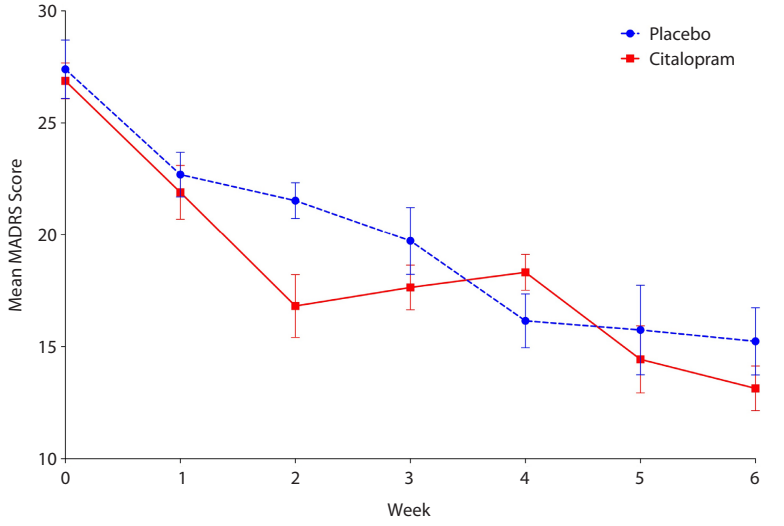
After the acute phase (6 weeks), 40 placebo-treated and 43 citalopram-treated subjects were

Table 1. Baseline Demographic and Clinical Characteristics of the Study Population

	Citalopram (n = 60)		Placebo (n = 59)		Overall (N = 119)	
Variable	n	%	n	%	n	%
Sex						
Male (n = 49)	22	36.7	27	45.8	49	41.2
Female (n = 70)	38	63.3	32	54.2	70	58.8
Race						
White (n = 67)	34	56.7	33	55.9	67	56.3
Black (n = 41)	21	35.0	20	33.9	41	34.5
Other (n = 11)[a]	5	8.3	6	10.2	11	9.2
Education						
High school or below (n = 19)	12	20.0	7	11.9	19	16.0
College (n = 55)	25	41.7	30	50.8	55	46.2
Graduate school (n = 7)	5	8.3	2	3.4	7	5.9
Missing (n = 38)	18	30.0	20	33.9	38	31.9
Site of enrollment						
Emory (n = 10)	5	8.3	5	8.5	10	8.4
Tufts (n = 74)	38	63.3	36	61.0	74	62.2
Duke (n = 35)	17	28.3	18	30.5	35	29.4
Bipolar disorder diagnosis						
Type I (n = 75)	34	56.7	41	69.5	75	63.0
Type II (n = 44)	26	43.3	18	30.5	44	37.0
Prior rapid cycling						
Present (n = 33)	14	23.3	19	32.2	33	27.7
Absent (n = 86)	46	76.7	40	67.8	86	72.3
Past substance abuse						
Present (n = 63)	29	48.3	34	57.6	63	52.9
Absent (n = 56)	31	51.7	25	42.4	56	47.1
Mood stabilizers						
Lithium (n = 61)	24	40.0	37	62.7	61	51.3
Divalproex (n = 17)	9	15.0	8	13.6	17	14.3
Carbamazepine (n = 21)	14	23.3	7	11.9	21	17.6
Lamotrigine (n = 20)	13	21.7	7	11.9	20	16.8
Other psychiatric medications						
Antipsychotics (n = 23)	10	16.7	13	22.0	23	19.3
Benzodiazepines (n = 18)	12	20.0	6	10.2	18	15.1
Other baseline antidepressants (n = 11)	6	10.0	5	8.5	11	9.2
Antianxiety (non-benzodiazepine) (n = 1)	0	0	1	1.7	1	0.8
Anticonvulsants (not listed otherwise) (n = 6)	4	6.7	2	3.4	6	5.0
None (n = 60)	28	46.7	32	54.2	60	50.4
Other medications						
Present (n = 27)	14	23.3	13	22.0	27	22.7
Absent (n = 92)	46	76.7	46	78.0	92	77.3
	Mean (SD) [n]		Mean (SD) [n]		Mean (SD) [n]	
Age, y	40.9 (12.6) [59]		42.1 (11.3) [59]		41.5 (11.9) [118]	
Mood stabilizer dose, mg						
Lithium	896 (235) [21]		784 (305) [28]		832 (280) [49]	
Divalproex	1156 (352) [8]		1208 (246) [6]		1179 (301) [14]	
Carbamazepine	510 (251) [10]		433 (151) [6]		481 (217) [16]	
Lamotrigine	125 (74) [7]		92 (72) [6]		110 (72) [13]	
Scores on clinical measures at assessment						
Montgomery-Asberg Depression Rating Scale	27.4 (9.1) [58]		27.4 (7.3) [58]		27.4 (8.2) [116]	
MRS-SADS	7.2 (5.3) [58]		8.2 (6.5) [58]		7.7 (5.9) [116]	
Global Assessment of Functioning	52.7 (7.4) [57]		53.3 (5.9) [56]		53.0 (6.7) [113]	
CGI-Mania	2.3 (1.0) [56]		2.2 (0.9) [56]		2.2 (1.0) [112]	
CGI-Depression	4.5 (0.7) [56]		4.5 (0.7) [56]		4.5 (0.7) [112]	
CGI-Overall	4.5 (0.8) [55]		4.5 (0.7) [56]		4.5 (0.7) [111]	
SF-36	1768 (512) [33]		1899 (597) [38]		1838 (559) [71]	
Sheehan Disability Scale	20.7 (5.6) [43]		19.4 (7.0) [44]		20.0 (6.4) [87]	

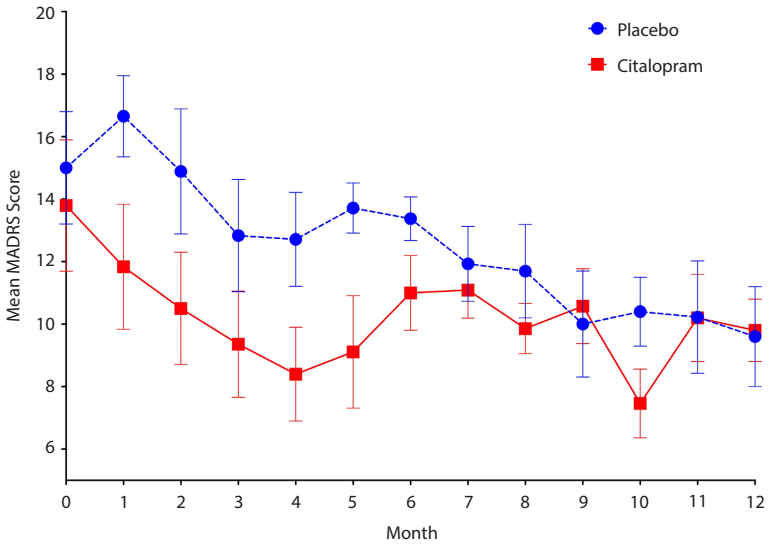
[a]Other = Hispanic, Asian, or mixed race.
Abbreviations: CGI = Clinical Global Impression scale, MRS-SADS = Mania Rating Scale from the Schedule for Affective Disorders and Schizophrenia, SF-36 = 36-Item Short Form Health Survey.

Figure 2. Primary Outcome: Acute Phase (6-Week) Efficacy of Citalopram Versus Placebo as Indicated by MADRS Scores[a]



[a]Bars indicate standard error.
Abbreviation: MADRS = Montgomery-Asberg Depression Rating Scale.

Figure 3. Secondary Outcome: Long-Term (12-Month) Efficacy of Citalopram Versus Placebo as Indicated by MADRS Scores[a]



[a]Bars indicate standard error.
Abbreviation: MADRS = Montgomery-Asberg Depression Rating Scale.

continued in the same randomized assignment into the long-term phase (up to 12 months).

Seventeen citalopram-treated subjects did not enter the maintenance phase, of whom 2 had completed the acute phase. Of those 2 patients, 1 patient was not continued into the maintenance phase because the treating clinical researcher felt the patient's clinical depression was too severe to continue current treatment unchanged; the other patient did not continue into the maintenance phase for unknown reasons. Nineteen placebo-treated subjects did not enter the maintenance phase, of whom 6 had completed the acute phase. Of those 6 patients 3 were not continued into the maintenance phase because the treating clinical researcher felt the patient's clinical depression was too severe to continue current treatment unchanged, 1 experienced a worsened mixed episode, and 2 had improved but chose not to continue the study for unknown reasons.

The main long-term outcome (MADRS score over time) was compared between the citalopram and placebo groups using a linear mixed-effects model with repeated measures (Figure 3 and Supplementary Appendix 2). The model indicated no statistically significant efficacy for citalopram versus placebo at 12 months ($P = .43$). There was some evidence of the temporal effect in both groups ($P = .08$); however the temporal changes were not significantly different between drug and placebo groups (treatment-by-visit interaction, $P = .58$). Visit-by-visit effect estimates in the mixed-effects regression model indicated that citalopram decreased MADRS scores by a mean of only 0.4 points overall compared with placebo.

Maintenance treatment response rate including all results up to the 12-month endpoint (50% reduction from baseline MADRS scores) was 31.8% (14/44) for citalopram versus 41.5% (17/41) for placebo ($P = .38$). Treatment remission (final MADRS scores below 7 at all endpoints up to 12 months) rate was 34.1% (15/44 for citalopram, 14/41 for placebo) for both groups ($P = 1$).

Other Secondary Outcomes

Acute treatment efficacy was assessed stratified by type II versus type I bipolar illness. Acute response was seen in 53.8% of type II subjects (14/26) with citalopram vs 50% (9/18) with placebo. Acute response was seen in 44.1% (15/34) of type I subjects versus 43.9% (18/41) with placebo. Comparing acute response in type II versus type I subjects, the numerical benefit in type II illness (53.8% in type II versus 44.1% in type I) was not statistically significant ($P = .60$; odds ratio [OR] = 1.47; 95% CI, 0.47 to 4.67).

Acute remission was seen in 26.9% of type II subjects (7/26) with citalopram versus 27.8% (5/18) with placebo. Acute response was seen in 35.3% of type I subjects (12/34) with citalopram versus 26.8% (11/41) with placebo. Comparing acute remission in type II versus type I subjects, the numerical benefit in type I illness (remission rate of 35.3% in type I versus 26.9% in type II) was not statistically significant (OR = 1.47; 95% CI, 0.43 to 5.38).

Manic episodes, defined as new onset of manic symptoms meeting *DSM-IV* criteria for manic or hypomanic episodes at any time during follow-up, were compared between groups. There were 9 full *DSM-IV*–defined manic/hypomanic episodes in the whole sample during acute and long-term follow-up assessment, without a higher rate in the citalopram group (3/60 with citalopram, 6/59 with placebo; OR = 2.14 for placebo; 95% CI, 0.43 to 13.87). However, manic symptoms as measured by the MRS-SADS were similar in both groups at the end of the acute phase, but were greater with citalopram in the maintenance phase (see summary statistics in Supplementary Appendix 2). The model-adjusted mean (95% CI) change in MRS-SADS from baseline to final visit was 0.1 (−1.5 to 1.6) for citalopram versus −1.8 (−3.2 to −0.4) for placebo ($P = .07$ for the treatment difference). There was evidence of significant effects due to prior rapid-cycling features ($P = .001$) and baseline MRS-SADS score ($P < .001$) as well (Supplementary Appendix 3).

Per analyses planned a priori, rapid-cycling status demonstrated notable influence on outcomes in treatment groups differentially, especially in the maintenance phase. In the acute phase, there were no numerical or statistically significant differences between groups stratified by rapid-cycling. In the maintenance phase, however, manic symptoms were elevated in rapid-cycling but not in non–rapid-cycling subjects (Supplementary Appendix 2). In the rapid-cycling subgroup, estimated mean (95% CI) change in MRS-SADS score from baseline to final visit was 1.9 (−0.3 to 4.0) for citalopram versus 0.1 (−1.9 to 2.0) for placebo; in the non–rapid-cycling subgroup, respective values were −1.8 (−3.3 to −0.2) for citalopram versus −3.6 (−5.2 to −2.1) for placebo. Hence, in the rapid-cycling subgroup, MRS-SADS scores worsened in the citalopram-treated group versus placebo by a mean of almost 2 points, whereas in the non–rapid-cycling group, MRS-SADS scores improved in both groups (albeit more with placebo than with citalopram).

Other Scales

Supplementary Appendix 2 provides summary statistics for outcomes on other secondary efficacy scales, specifically the Clinical Global Impression scale for depression (CGI-Depression),[21] CGI-Mania,[21] CGI-Overall,[21] Global Assessment of Functioning,[22] Sheehan Disability Scale,[23] and 36-Item Short Form Health Survey.[24] No notable numerical differences were present in the acute or maintenance phase, consistent with findings on the MADRS.

Concomitant Medications

As noted in Table 1, in the citalopram group, there was greater carbamazepine and lamotrigine prescription as a baseline mood stabilizer and greater benzodiazepine use as a concomitant medication, while in the placebo group, there was greater lithium prescription as a baseline mood stabilizer. Some patients were treated with 2 mood stabilizer agents, such as lithium plus divalproex; this occurred in 7 subjects in the citalopram group (11.7%) versus 4 subjects in the placebo group (6.8%; OR = 1.81; 95% CI, 0.43 to 8.92).

Side Effects and Termination

Overall, side effects were observed in 36·1% of subjects in the entire sample, and similarly in both groups (35% with citalopram vs 37.3% with placebo). In the acute phase, 28.3% of citalopram subjects (17/60) and 32.2% of placebo subjects (19/59) terminated the study before 6 weeks. Of subjects who entered the maintenance phase, 67.4% of citalopram subjects (29/43) and 70.0% of placebo subjects (28/40) terminated the study before 1 year. See Supplementary Appendix 1 for further details.

DISCUSSION

Citalopram, added to standard mood stabilizers, did not have clinically meaningful benefit versus placebo for both acute and maintenance treatment of bipolar depression. Citalopram was not associated with notable risk of acute mania/hypomania, nor with greater benefit in type II versus type I bipolar illness. In the maintenance phase, there was some manic symptomatic worsening overall, especially in rapid-cycling subjects.

In the acute phase, it is notable that both citalopram and placebo groups improved similarly, which suggests non-pharmacologic reasons for improvement, such as natural remission of bipolar depressive episodes. The natural history of bipolar depression indicates resolution of an episode typically within 6 months or less,[25] which is consistent with improvement in a notable subgroup of placebo-treated patients in acute 6- to 8-week trials. This RCT is consistent with the most recent meta-analysis of bipolar depression,[4] which found benefit with placebo, very likely reflecting natural history, that is not exceeded by antidepressants. In clinical practice, given the absence of placebo comparison, this benefit will be attributed to antidepressant use, producing the clinical impression of drug efficacy.

Regarding possible type II false-negative error, this study used standard definitions of clinically meaningful improvement of depression symptoms. In the acute phase, the overall improvement seen with citalopram was 1.7 points more than with placebo, a small effect size, below the 3-point or larger clinically meaningful standard.[20] Even with a much larger sample, that effect size of less than 2 points' benefit would not be clinically meaningful. These results are consistent with those of other meta-analyses[4,6,26] and provide further evidence of lack of clinically meaningful efficacy of antidepressants in bipolar depression. Thus, this study was not underpowered for a clinically meaningful outcome.

It could be argued that the small effect size of benefit shown here is similar to what is seen in meta-analyses[27,28] of modern antidepressants in "major depressive disorder," which is similar to but not exactly the same concept as what had been called "unipolar" depression prior to *DSM-III* in 1980.[29] This fact could be interpreted not as more reason to use antidepressants in bipolar depression, however, but rather perhaps as less rationale for using them in unipolar depression as well, given the low level of clinical benefit.

Regarding secondary outcomes, this RCT found that type II bipolar depression subjects did not respond better to antidepressants compared to those with type I bipolar depression. This result contradicts common belief, but it should be noted that other studies which find benefit with antidepressants in type II bipolar depression[30–32] do not have direct comparisons in the same sample with subjects with type I bipolar depression. When such direct comparison has been made, no differences have been found.[33]

The immediate risk of full episodes of acute mania/hypomania with citalopram, in the setting of concomitant mood stabilizer treatment, was low in this study, in agreement with findings of other RCTs of serotonin reuptake inhibitors.[9,10] However, in the maintenance phase, manic rating scale scores were higher with citalopram compared to with placebo overall, and especially so in rapid-cycling subjects. This worsening is consistent with all prior randomized data on antidepressants in rapid-cycling subjects,[11,12] which indicate more mood episodes with maintenance continuation of antidepressants compared to placebo or antidepressant discontinuation. Oft-cited data to the contrary[34] are observational and non-randomized, and thus less valid than randomized results.

Limitations

Potential limitations of a type II error of false-negative findings are addressed in the preceding paragraphs, and limitations also include concerns regarding effects of concomitant medications as well as potential false-positive findings in subgroup analysis. Another potential limitation involves dropouts in long-term follow-up studies. Maintenance studies of bipolar illness often have high termination rates; a rate of 68.7% was found with the present protocol. It is important to realize that this termination was mostly protocol-defined, meaning subjects were taken out of the study if any intervention other than study medication was deemed clinically necessary. This termination rate does not primarily reflect "dropout" defined as a patient decision to stop treatment. Further, this observed termination rate is similar to the lowest rates seen in maintenance trials of bipolar illness, such as the BALANCE (Bipolar Affective disorder: Lithium/ANticonvulsant Evaluation) study[35] (about 60% termination), and is notably lower than rates found in most maintenance RCTs of bipolar illness, including major highly cited studies used for US Food and Drug Administration indications of other agents (eg, lamotrigine, quetiapine[36–38]) (94%–100% dropout at 1 year).

CONCLUSION

Citalopram, added to standard mood stabilizers, did not have clinically meaningful benefit versus placebo for either acute or maintenance treatment of bipolar depression. Worsening of manic symptoms with maintenance treatment was seen, especially in rapid-cycling subjects. The results of this RCT would suggest caution in the use of antidepressants in bipolar illness, mainly because of lack of clinically meaningful efficacy for acute and maintenance treatment and secondarily due to long-term manic worsening in rapid-cycling subjects.

Submitted: October 26, 2019; accepted July 20, 2020.

Published online: January 12, 2021.

Author contributions: Drs Ghaemi and Patkar were the two co–Principal Investigators (PIs) of this study. They recruited and treated all research subjects and jointly collected data, analyzed data, and wrote and revised the manuscript. Dr Whitham was a research coordinator during the study, collecting and entering data. Later, she assisted with data analysis, manuscript preparation, and manuscript revision. Drs Amerio and Barroilhet were research fellows during the project and assisted with patient recruitment, data cleanup, and data analysis and assisted with writing and revising the manuscript. Dr Sverdlov was involved in final statistical analysis and manuscript preparation in 2019 and was not involved in the study previously.

Potential conflicts of interest: **Dr Ghaemi** is an employee of Novartis Institutes for Biomedical Research (NIBR), Cambridge, MA, which began in October 2017. The current study was funded by the National Institute of Mental Health (NIMH) in 2007, with Dr Ghaemi as the PI, and was conducted and completed in 2007–2016, before Dr Ghaemi's employment at NIBR. **Dr Patkar** has received grant support from Pfizer, Sunovion, Forest Research Institute, Daiichi Sankyo, and Forum Pharma and honoraria for being a consultant/member of advisory boards or speakers bureaus for Otsuka, Biodelivery Sciences, and Acadia Pharma. **Dr Sverdlov** is employed by NIBR. **Drs Whitham**, **Amerio**, and **Barroilhet** have no potential conflicts of interest to disclose.

Funding/support: Funded by NIMH grant 5R01MH078060-05.

Role of the sponsor: The NIMH had no role in the conduct or publication of the study.

Supplementary material: Available at PSYCHIATRIST.COM.

REFERENCES

1. El-Mallakh RS, Ghaemi SN, eds. *Bipolar Depression: A Comprehensive Guide*. Washington, DC: American Psychiatric Publishing; 2006.
2. Baldessarini R, Henk H, Sklar A, et al. Psychotropic medications for patients with bipolar disorder in the United States: polytherapy and adherence. *Psychiatr Serv*. 2008;59(10):1175–1183.
3. Pacchiarotti I, Bond DJ, Baldessarini RJ, et al. The International Society for Bipolar Disorders (ISBD) task force report on antidepressant use in bipolar disorders. *Am J Psychiatry*. 2013;170(11):1249–1262.
4. McGirr A, Vöhringer PA, Ghaemi SN, et al. Safety and efficacy of adjunctive second-generation antidepressant therapy with a mood stabiliser or an atypical antipsychotic in acute bipolar depression: a systematic review and meta-analysis of randomised placebo-controlled trials. *Lancet Psychiatry*. 2016;3(12):1138–1146.
5. Ghaemi SN, Hsu DJ, Soldani F, et al. Antidepressants in bipolar disorder: the case for caution. *Bipolar Disord*. 2003;5(6):421–433.
6. Ghaemi SN, Wingo AP, Filkowski MA, et al. Long-term antidepressant treatment in bipolar disorder: meta-analyses of benefits and risks. *Acta Psychiatr Scand*. 2008;118(5):347–356.
7. Goldberg JF, Truman CJ. Antidepressant-induced mania: an overview of current controversies. *Bipolar Disord*. 2003;5(6):407–420.
8. Tondo L, Vázquez G, Baldessarini RJ. Mania associated with antidepressant treatment: comprehensive meta-analytic review. *Acta Psychiatr Scand*. 2010;121(6):404–414.
9. Frye MA, Helleman G, McElroy SL, et al. Correlates of treatment-emergent mania associated with antidepressant treatment in bipolar depression. *Am J Psychiatry*. 2009;166(2):164–172.
10. Leverich GS, Altshuler LL, Frye MA, et al. Risk of switch in mood polarity to hypomania or mania in patients with bipolar depression during acute and continuation trials of venlafaxine, sertraline, and bupropion as adjuncts to mood stabilizers. *Am J Psychiatry*. 2006;163(2):232–239.
11. El-Mallakh RS, Vöhringer PA, Ostacher MM, et al. Antidepressants worsen rapid-cycling course in bipolar depression: a STEP-BD randomized clinical trial. *J Affect Disord*. 2015;184:318–321.
12. Wehr TA, Goodwin FK. Can antidepressants cause mania and worsen the course of affective illness? *Am J Psychiatry*. 1987;144(11):1403–1411.
13. First MB, Gibbon M. The Structured Clinical Interview for *DSM-IV* Axis I Disorders (SCID-I) and the Structured Clinical Interview for

DSM-IV Axis II Disorders (SCID-II). In: Hilsenroth M, Segal D, eds. *Comprehensive Handbook of Psychological Assessment*. New York, NY: John Wiley and Sons Inc; 2004:134–143.
14. Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. *Br J Psychiatry*. 1979;134(4):382–389.
15. Endicott J, Spitzer RL. A diagnostic interview: the schedule for affective disorders and schizophrenia. *Arch Gen Psychiatry*. 1978;35(7):837–844.
16. Ghaemi SN, Selker HP. Maintenance efficacy designs in psychiatry: Randomized discontinuation trials - enriched but not better. *J Clin Transl Sci*. 2017;1(3):198–204.
17. Tohen M, Vieta E, Calabrese J, et al. Efficacy of olanzapine and olanzapine-fluoxetine combination in the treatment of bipolar I depression. *Arch Gen Psychiatry*. 2003;60(11):1079–1088.
18. Loebel A, Cucchiaro J, Silva R, et al. Lurasidone monotherapy in the treatment of bipolar I depression: a randomized, double-blind, placebo-controlled study. *Am J Psychiatry*. 2014;171(2):160–168.
19. Keefe RS, Kraemer HC, Epstein RS, et al. Defining a clinically meaningful effect for the design and interpretation of randomized controlled trials. *Innov Clin Neurosci*. 2013;10(5–6 suppl A):4S–19S.
20. National Collaborating Centre for Mental Heath. *National Institute for Health and Clinical Excellence: Guidance.* Leicester, UK: British Psychological Society; 2010.
21. Guy W, ed. *Clinical Global Impressions*. *ECDEU Assessment Manual for Psychopharmacology: Publication ADM 76-338*. Rockville, MD: National Institute of Mental Health, Psychopharmacology Research Branch; 1976:217–222.
22. Spitzer RL, Gibbon M, Williams JBW, et al. Global Assessment of Functioning (GAF) Scale. In: Sederer LJ, Dickey B, eds. *Outcome Assessment in Clinical Practice*. 1st ed. Baltimore, MD: Williams and Wilkins; 1996.
23. Sheehan DV, Harnett-Sheehan K, Raj BA. The measurement of disability. *Int Clin Psychopharmacol*. 1996;11(suppl 3):89–95.
24. Ware JE, Snow KK, Kosinski M, et al. *SF-36 Health Survey Manual and Interpretation Guide*. Boston, MA: The Health Institute; 1993.
25. Goodwin FK, Jamison KR, Ghaemi SN. *Manic-Depressive Illness: Bipolar Disorders and Recurrent Depression*. 2nd ed. New York, NY: Oxford University Press; 2007.
26. Sidor MM, Macqueen GM. Antidepressants for the acute treatment of bipolar depression: a systematic review and meta-analysis. *J Clin Psychiatry*. 2011;72(2):156–167.
27. Jakobsen JC, Katakam KK, Schou A, et al. Selective serotonin reuptake inhibitors versus placebo in patients with major depressive disorder: a systematic review with meta-analysis and trial sequential analysis. *BMC Psychiatry*. 2017;17(1):58.
28. Cipriani A, Furukawa TA, Salanti G, et al. Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis. *Lancet*. 2018;391(10128):1357–1366.
29. Ghaemi SN, Dalley S. The bipolar spectrum: conceptions and misconceptions. *Aust N Z J Psychiatry*. 2014;48(4):314–324.
30. Amsterdam JD, Brunswick DJ. Antidepressant monotherapy for bipolar type II major depression. *Bipolar Disord*. 2003;5(6):388–395.
31. Amsterdam JD, Shults J. Comparison of fluoxetine, olanzapine, and combined fluoxetine plus olanzapine initial therapy of bipolar type I and type II major depression—lack of manic induction. *J Affect Disord*. 2005;87(1):121–130.
32. Amsterdam JD, Shults J. Comparison of short-term venlafaxine versus lithium monotherapy for bipolar II major depressive episode: a randomized open-label study. *J Clin Psychopharmacol*. 2008;28(2):171–181.
33. Vöhringer PA, Ostacher MJ, El-Mallakh RS, et al. Antidepressants in type ii versus type i bipolar depression: a randomized discontinuation trial. *J Clin Psychopharmacol*. 2015;35(5):605–608.
34. Coryell W, Solomon D, Turvey C, et al. The long-term course of rapid-cycling bipolar disorder. *Arch Gen Psychiatry*. 2003;60(9):914–920.
35. Geddes JR, Goodwin GM, Rendell J, et al; BALANCE investigators and collaborators. Lithium plus valproate combination therapy versus monotherapy for relapse prevention in bipolar I disorder (BALANCE): a randomised open-label trial. *Lancet*. 2010;375(9712):385–395.
36. Calabrese JR, Suppes T, Bowden CL, et al. A double-blind, placebo-controlled, prophylaxis study of lamotrigine in rapid-cycling bipolar disorder: Lamictal 614 Study Group. *J Clin Psychiatry*. 2000;61(11):841–850.
37. Bowden CL, Calabrese JR, Sachs G, et al; Lamictal 606 Study Group. A placebo-controlled 18-month trial of lamotrigine and lithium maintenance treatment in recently manic or hypomanic patients with bipolar I disorder. *Arch Gen Psychiatry*. 2003;60(4):392–400.
38. Suppes T, Vieta E, Liu S, et al; Trial 127 Investigators. Maintenance treatment for patients with bipolar I disorder: results from a north american study of quetiapine in combination with lithium or divalproex (trial 127). *Am J Psychiatry*. 2009;166(4):476–488.

See supplementary material for this article at PSYCHIATRIST.COM.

THE JOURNAL OF
CLINICAL PSYCHIATRY
THE OFFICIAL JOURNAL OF THE AMERICAN SOCIETY OF CLINICAL PSYCHOPHARMACOLOGY

Supplementary Material

Article Title: Citalopram for Acute and Preventive Efficacy in Bipolar Depression (CAPE-BD): A Randomized, Double-Blind, Placebo-Controlled Trial

Author(s): S. Nassir Ghaemi MD MPH; Elizabeth A. Whitham MD; Paul A. Vohringer MD MPH; Sergio A. Barroilhet MD; Andrea Amerio MD; Oleksandr Sverdlov PhD; and Ashwin A. Patkar MD

DOI Number: https://doi.org/10.4088/JCP.19m13136

List of Supplementary Material for the article

1. **Appendix 1** Supplementary methods and results
2. **Appendix 2** Summary statistics of primary and secondary rating scales
3. **Appendix 3** Analysis of covariance of MRS: Outcome = change from baseline to final visit; treatment group as a classification factor, Rapid cycling and baseline MRS as covariates

Disclaimer

This Supplementary Material has been provided by the author(s) as an enhancement to the published article. It has been approved by peer review; however, it has undergone neither editing nor formatting by in-house editorial staff. The material is presented in the manner supplied by the author.

Appendix 1. Supplementary methods and results

Study design and setting

This study was funded by the National Institute of Mental Health (5R01MH078060-05) as a double-blind, randomized, placebo-controlled, parallel-group trial conducted in the United States (patient recruitment 2007-2014; completion of follow up until 2015). Two research groups, located at three sites, participated: one research group initially at Emory University (Atlanta, Georgia), and then Tufts Medical Center (Boston, Massachusetts), and a second research group at Duke University (Durham, North Carolina).

Intervention

119 patients were block-randomized in a 1:1 ratio to receive R, S-citalopram or placebo, stratified by rapid cycling status and type I versus type I subtypes. Random number sequences were generated by web-based random allocation algorithm.

The Tufts Medical Center research pharmacy generated the random allocation sequence, and then notified trial investigators about allocation status, after which subjects were assigned to treatment. Participants and investigators were masked to treatment allocation during the entire trial.

Identical placebo pills were prepared by the Tufts Medical Center research pharmacy, and then distributed to trial investigators. All pills were distributed in opaque containers, and investigators did not visualize the pills utilized during the study. Patients were not exposed to pill type of any particular patient, did not participate in group cross-over, and had no interaction or knowledge of other enrolled study patients.

Citalopram and placebo dose was 10 mg/day for the first week, and then increased by 10 mg per week to a maximum of 60 mg/day, based on clinical judgment.

Dosing was analyzed as follows. In the acute phase: For each subject in the citalopram arm, their dose at Week 6 (or last available observation during acute phase) was retrieved. Then summary statistics were calculated across all available subject data. In the maintenance phase: For each subject in the citalopram arm, their dose recorded at the final/termination visit was retrieved. Then summary statistics were calculated across all available subject data. The highest or final dose, whichever was higher, was used in the analysis.

Ethical considerations

This study was approved by the Institutional Review Board of each participating site. All subjects provided informed consent.

Statistical analysis (further details)

To account for the correlated nature of the data, data were analyzed as repeated measures (using unstructured covariance structure) and included site of data collection as a random effect. Two separate models were developed for acute (6 weeks) and maintenance (up to 1 year of follow up) phases. The primary model for each stage was built using MADRS scores over time as the response variable, and citalopram (drug) and weeks/months (time) as the main explanatory variables. Time was also included in the model as an interaction term with the drug arm. Race, diagnosis (bipolar type I vs bipolar type II), gender, rapid-cycling status were in turn included as dependent variables in a backward selection model.

A further sensitivity analysis for both phases of treatment (acute and maintenance) was performed with baseline MADRS measures, race, diagnosis, and interaction between diagnosis, drug and rapid-cycling status. Significant deviations from model assumptions were not found. Additional sensitivity analysis assessing robustness of results were done. Dropout rates were not significantly different overall. For dropouts by randomization arm and CONSORT data, see figure 1. New onset of manic/hypomanic outcomes at any time of the trial for each group were reported as safety assessments.

Side effects and termination (further details)

With citalopram, the most common side effects were sedation (n=8, 13.3%), dry mouth (n=6, 10%), sexual dysfunction (n=5, 8.3%), and headache (n=4, 6.7%). Other side effects were insomnia, nausea, itchiness, cognitive impairment (n=3 each); weight gain/increased appetite, shakiness/jitteriness (n=2 each); and diarrhea, stomach cramps, polyuria, myalgia, stiffness, and akathisia (n=1 each). With placebo, the most common side effects were nausea and headache (n=6 each, 10.2%), diarrhea (n=5, 8.5%), and tremor (n=4, 6.8%). Other side effects were sedation, ataxia, polyuria (n=3 each); cognitive impairment, dizziness, manic symptoms, weight gain (n=2 each); and tinnitus, shakiness, stiffness, sweating, abdominal pain, flatulence, acne, myalgia, and akathisia (n=1 each). Comparing the two groups, the main differences in side effects involved more sexual dysfunction, insomnia, and sedation with citalopram, and more nausea, tremor, ataxia, and polyuria with placebo. Some of these differences may be related to differences in concomitant medications, as noted above. Thus, more lithium use in the placebo group may explain some of the above side effects, while more benzodiazepine use in the citalopram group may explain a greater amount of sedation.

The main side effects differentiated between the two groups that were unlikely to be related to concomitant medications were sexual dysfunction (8.3% vs 0% for citalopram versus placebo and insomnia (5% vs 0% for citalopram versus placebo), in both cases present in some patients with citalopram and absent with placebo.

Figure 1 describes similar reasons for termination including manic episodes, continued or worsening depression, loss to follow-up with appointment non-adherence, or adverse events such as nausea. As noted previously, eight subjects completed six weeks of treatment but did not continue into the maintenance phase, six with placebo and two with citalopram, mostly for continued depression.

Figure 1 describes some numerical differences for early maintenance termination, with somewhat more loss to follow up in the citalopram group (27.9%, 12/43), and somewhat more discontinuation due to patient decision and non-adherence in the placebo group. Both groups had similar rates of discontinuation due to side effects, mood episodes, and suicidality (see eAppendix).

Specific frequencies for reasons for discontinuation are provided in the figure, except for the "other" category. For citalopram, other reasons for discontinuation included unknown (non-adherence, sexual dysfunction, patient decision, and protocol violation; n = 1 each; and inadequate documentation, n=6); for placebo, other reasons for discontinuation included study relocation (n = 2), and rash and lack of efficacy (n = 1 each), and inadequate documentation (n=3). Overall 21.0% (25/119) of subjects completed the entire trial for 12 months (21.6%, 13/60, with citalopram, and 20.3%, 12/59, with placebo).

One subject, who was randomized to citalopram, committed suicide. He was recently started on lamotrigine, and was judged to have a severe melancholic depressive episode that predated entry into the study and did not change after four weeks of acute treatment. The subject reported mild suicidal ideation but consistently denied plan or intent at all times, to clinicians as well as family members. There was no evidence of any manic symptoms or a mixed state, and the suicide was judged not to be related to the study intervention.

Appendix 2. Summary statistics of primary and secondary rating scales

For all tables, the final termination visit is the final visit seen in the study, at whatever point in the study, which would include early termination.

Summary of total MADRS score over time by treatment group

Acute phase	Citalopram (n=60)		Placebo (n=59)	
	N	Mean (SD)	N	Mean (SD)
Week 0	58	27.4 (9.1)	58	27.4 (7.3)
Week 1	43	22.1 (9.8)	36	22.7 (9.6)
Week 2	38	16.8 (8.3)	47	21.5 (9.9)
Week 3	42	17.6 (11.1)	44	19.7 (10.7)
Week 4	38	18.3 (10.0)	39	16.2 (8.3)
Week 5	32	14.4 (7.8)	40	15.8 (8.3)
Week 6	35	13.1 (8.4)	50	15.2 (9.9)
Maintenance phase				
Visit 1 (Week 8)	34	12.8 (9.3)	33	17.1 (8.8)
Visit 2 (Week 10)	33	11.6 (8.7)	28	15.4 (6.7)
Visit 3 (Week 12)	17	10.6 (7.5)	24	13.1 (7.5)
Visit 4 (Week 14)	20	8.5 (6.9)	20	13.8 (5.3)
Visit 5 (Week 18)	21	11.3 (8.4)	22	13.5 (8.4)
Visit 6 (Week 22)	14	13.4 (7.1)	21	12.8 (7.8)
Visit 7 (Week 26)	17	10.8 (7.7)	18	13.2 (5.4)
Visit 8 (Week 30)	16	10.9 (7.8)	12	11.3 (6.3)
Visit 9 (Week 34)	17	9.6 (7.8)	12	11.7 (6.0)
Visit 10 (Week 38)	13	11.9 (7.3)	6	10.3 (5.5)
Visit 11 (Week 42)	11	9.2 (5.1)	8	11.0 (7.5)
Final/Termination Visit	59	15.8 (11.5)	59	18.0 (13.0)

Summary statistics of MRS over time, by treatment group

Acute phase	Citalopram (n=60)		Placebo (n=59)	
	N	Mean (SD)	N	Mean (SD)
Week 0	58	7.2 (5.3)	58	8.2 (6.5)
Week 1	43	5.6 (4.4)	36	7.5 (5.6)
Week 2	38	4.9 (4.0)	47	5.8 (5.4)
Week 3	42	6.1 (6.2)	44	5.9 (4.6)
Week 4	38	5.8 (5.8)	39	5.9 (4.9)
Week 5	32	6.4 (6.5)	40	5.1 (4.3)
Week 6	35	5.2 (6.7)	50	6.1 (4.4)
Maintenance phase				
Visit 1 (Week 8)	34	4.2 (3.9)	33	7.2 (5.6)
Visit 2 (Week 10)	32	5.1 (4.1)	28	7.8 (5.5)
Visit 3 (Week 12)	17	4.1 (3.7)	23	6.0 (4.6)
Visit 4 (Week 14)	20	3.8 (4.1)	20	5.8 (4.6)
Visit 5 (Week 18)	24	6.0 (5.8)	20	5.2 (3.7)
Visit 6 (Week 22)	21	4.4 (3.6)	22	6.7 (4.5)
Visit 7 (Week 26)	14	6.6 (4.5)	21	6.9 (5.2)
Visit 8 (Week 30)	17	7.9 (7.9)	18	6.6 (4.4)
Visit 9 (Week 34)	16	6.5 (4.7)	12	6.3 (3.7)
Visit 10 (Week 38)	17	5.9 (2.7)	12	7.6 (3.7)
Visit 11 (Week 42)	12	6.1 (4.1)	7	7.1 (4.5)
Visit 12 (Week 46)	11	6.3 (3.1)	8	7.5 (3.6)
Final/Termination Visit	59	7.3 (7.6)	59	5.4 (4.4)

Summary statistics of MRS, by treatment group and Rapid Cycling feature

	Overall		Rapid Cycling		Non-rapid Cycling	
	Citalopram (n=60)	**Placebo (n=59)**	**Citalopram (n=14)**	**Placebo (n=19)**	**Citalopram (n=46)**	**Placebo (n=40)**
Baseline MRS						
N	58	58	14	19	44	39
Mean (SD)	7.2 (5.3)	8.2 (6.5)	9.1 (5.5)	6.4 (4.9)	6.5 (5.2)	9.1 (7.1)
Median	6	6.5	8	6	5.5	7.0
IQR	3–9.8	4–10.8	5.3–13.3	3–8	2–9	4.5–14.5
Range	0–20	0–26	2–20	0–21	0–20	0–26
Final visit MRS						
N	59	59	14	19	45	40
Mean (SD)	7.3 (7.6)	5.4 (4.4)	11.6 (9.1)	6.0 (5.0)	6.0 (6.6)	5.2 (4.1)
Median	5	5	11.5	4	4	5
IQR	1.5–10.5	2–7.5	5–19	2–10.5	1–8	2–6
Range	0–30	0–17	0–27	0–15	0–30	0–17
Change (Final visit – Baseline)						
N	58	58	14	19	44	39
Mean (SD)	0.1 (7.8)	-2.8 (6.4)	2.5 (9.9)	-0.4 (5.2)	-0.6 (6.9)	-4 (6.7)
Median	-0.5	-3	0	0	-1	-3
IQR	-4–2	-6.8–0	-4.5–5.3	-4.5–1	-4–2	-7.5–0
Range	-15–21	-20–15	-8–21	-8–11	-15–21	-20–15

Summary statistics of GAF over time, by treatment group

Acute phase	Citalopram (n=60)		Placebo (n=59)	
	N	**Mean (SD)**	**N**	**Mean (SD)**
Week 0	57	52.7 (7.4)	56	53.3 (5.9)
Week 1	43	57.3 (8.4)	35	56.3 (11.1)
Week 2	38	63.8 (9.4)	47	59.5 (9.8)
Week 3	42	63.8 (12.9)	43	61.7 (11.9)
Week 4	38	64.4 (9.5)	39	64.9 (9.3)
Week 5	32	67.8 (9.7)	39	64.7 (10.0)
Week 6	33	69.3 (10.5)	50	65.5 (10.0)
Maintenance phase				
Visit 1 (Week 8)	34	70.3 (12.2)	33	66.5 (9.3)
Visit 2 (Week 10)	31	70.1 (13.4)	27	65.5 (7.2)
Visit 3 (Week 12)	16	73.4 (10.7)	23	68.2 (9.1)
Visit 4 (Week 14)	20	72.7 (9.8)	20	67.4 (8.2)
Visit 5 (Week 18)	23	73.6 (11.0)	20	67.7 (11.0)
Visit 6 (Week 22)	21	72.6 (10.0)	22	68.1 (10.9)
Visit 7 (Week 26)	14	70.6 (9.9)	21	69.1 (12.8)
Visit 8 (Week 30)	17	69.3 (10.4)	18	67.2 (8.0)
Visit 9 (Week 34)	16	69.3 (10.9)	10	71.0 (7.7)
Visit 10 (Week 38)	17	70.4 (10.6)	12	68.0 (6.2)
Visit 11 (Week 42)	13	68.2 (7.9)	7	70.1 (5.7)
Visit 12 (Week 46)	11	67.9 (5.7)	8	70.0 (7.1)
Final/Termination Visit	58	62.4 (16.2)	56	62.5 (14.7)

Summary statistics of CGI-Mania over time, by treatment group

Acute phase	Citalopram (n=60)		Placebo (n=59)	
	N	Mean (SD)	N	Mean (SD)
Week 0	56	2.3 (1.0)	56	2.2 (0.9)
Week 1	42	2.0 (0.9)	35	2.3 (0.9)
Week 2	38	1.9 (0.9)	47	2.0 (0.8)
Week 3	41	2.0 (1.0)	44	2.2 (1.1)
Week 4	34	2.0 (0.9)	39	2.0 (0.9)
Week 5	31	1.9 (0.9)	38	1.8 (0.8)
Week 6	33	1.7 (0.9)	49	2.1 (0.7)
Maintenance phase				
Visit 1 (Week 8)	32	1.6 (0.9)	33	2.1 (0.7)
Visit 2 (Week 10)	30	1.7 (0.9)	28	2.1 (1.0)
Visit 3 (Week 12)	17	1.8 (0.7)	24	2.1 (0.8)
Visit 4 (Week 14)	19	1.5 (0.7)	19	2.0 (0.8)
Visit 5 (Week 18)	23	2.2 (1.0)	20	2.1 (0.8)
Visit 6 (Week 22)	21	1.8 (0.8)	21	2.2 (0.9)
Visit 7 (Week 26)	13	2.1 (1.0)	20	2.2 (0.7)
Visit 8 (Week 30)	15	2.3 (0.9)	18	2.4 (0.7)
Visit 9 (Week 34)	15	2.3 (0.9)	12	2.3 (0.5)
Visit 10 (Week 38)	16	2.1 (0.8)	10	2.4 (0.5)
Visit 11 (Week 42)	13	2.2 (0.8)	7	2.6 (0.5)
Visit 12 (Week 46)	10	2.3 (0.8)	7	2.7 (0.5)
Final/Termination Visit	55	1.9 (1.3)	58	2.1 (0.9)

Summary statistics of CGI-Depression over time, by treatment group

Acute phase	Citalopram (n=60)		Placebo (n=59)	
	N	**Mean (SD)**	**N**	**Mean (SD)**
Week 0	56	4.5 (0.7)	56	4.5 (0.7)
Week 1	43	3.8 (1.0)	35	4.1 (1.0)
Week 2	38	3.3 (1.0)	47	3.8 (1.0)
Week 3	41	3.2 (1.2)	44	3.6 (1.1)
Week 4	36	3.2 (0.9)	39	3.2 (0.8)
Week 5	31	2.9 (1.0)	38	3.2 (0.9)
Week 6	33	2.8 (1.1)	49	3.1 (1.1)
Maintenance phase				
Visit 1 (Week 8)	34	2.6 (1.1)	33	3.2 (1.0)
Visit 2 (Week 10)	32	2.5 (1.1)	28	3.1 (0.7)
Visit 3 (Week 12)	17	2.5 (1.2)	24	3.0 (0.8)
Visit 4 (Week 14)	20	2.5 (0.9)	19	3.1 (0.8)
Visit 5 (Week 18)	24	2.5 (1.1)	20	3.0 (0.9)
Visit 6 (Week 22)	21	2.5 (1.0)	21	3.0 (1.0)
Visit 7 (Week 26)	13	2.9 (1.0)	20	3.0 (0.9)
Visit 8 (Week 30)	17	2.5 (1.1)	18	3.1 (0.6)
Visit 9 (Week 34)	15	2.5 (1.4)	12	2.8 (0.6)
Visit 10 (Week 38)	17	2.5 (1.1)	10	2.7 (0.8)
Visit 11 (Week 42)	13	2.5 (1.1)	7	2.6 (0.5)
Visit 12 (Week 46)	10	2.5 (0.7)	7	2.9 (0.7)
Final/Termination Visit	55	3.1 (1.3)	58	3.4 (1.4)

Summary statistics of CGI-Overall over time, by treatment group

Acute phase	Citalopram (n=60)		Placebo (n=59)	
	N	Mean (SD)	N	Mean (SD)
Week 0	55	4.5 (0.8)	56	4.5 (0.7)
Week 1	42	3.9 (1.0)	34	4.1 (1.0)
Week 2	38	3.3 (1.0)	47	3.8 (1.0)
Week 3	41	3.4 (1.2)	42	3.6 (1.1)
Week 4	35	3.2 (1.0)	38	3.2 (0.8)
Week 5	32	2.9 (1.0)	37	3.2 (0.9)
Week 6	33	2.8 (1.1)	47	3.1 (1.0)
Maintenance phase				
Visit 1 (Week 8)	34	2.7 (1.1)	32	3.2 (0.9)
Visit 2 (Week 10)	32	2.6 (1.1)	27	3.1 (0.7)
Visit 3 (Week 12)	16	2.6 (1.1)	23	3.0 (0.8)
Visit 4 (Week 14)	20	2.5 (0.9)	19	3.0 (0.8)
Visit 5 (Week 18)	23	2.7 (1.1)	19	3.0 (1.0)
Visit 6 (Week 22)	21	2.6 (1.0)	20	3.0 (1.0)
Visit 7 (Week 26)	14	2.8 (1.2)	19	2.9 (0.9)
Visit 8 (Week 30)	17	2.5 (1.0)	18	3.1 (0.7)
Visit 9 (Week 34)	16	2.8 (1.1)	12	2.8 (0.6)
Visit 10 (Week 38)	17	2.5 (1.1)	10	2.8 (0.8)
Visit 11 (Week 42)	12	2.7 (1.0)	7	2.7 (0.5)
Visit 12 (Week 46)	10	2.5 (0.7)	7	3.0 (0.6)
Final/Termination Visit	55	3.3 (1.4)	59	3.5 (1.4)

Summary statistics of Sheehan disability scale, by treatment group

Variable	Citalopram (n=60)	Placebo (n=59)
Baseline Sheehan		
N	43	44
Mean (SD)	20.7 (5.6)	19.4 (7.0)
Median	21	20.5
IQR	16–24	16.5–24
Range	9–30	1–30
Final visit Sheehan		
N	33	40
Mean (SD)	12.9 (7.5)	12.5 (9.7)
Median	12	14
IQR	7–20	3.5–20
Range	0–27	0–30
Change (Final visit – Baseline)		
N	26	33
Mean (SD)	-6.6 (9.0)	-5.8 (9.9)
Median	-6	-4
IQR	-12.8–0	-11–1
Range	-30–9	-30–13

95% CI for the mean treatment difference (based on 2-sample t-test): (-5.8, 4.2)

Summary statistics of SF-36 survey, by treatment group

Variable	Citalopram (n=60)	Placebo (n=59)
Baseline SF-36		
N	33	38
Mean (SD)	1768 (512)	1899 (597)
Median	1925	1780
IQR	1390–2095	1438–2329
Range	580–2695	935–3205
Final visit SF-36		
N	25	31
Mean (SD)	2248 (648)	2153 (699)
Median	2215	2320
IQR	1760–2675	1695–2530
Range	1020–3285	490–3465
Change (Final visit – Baseline)		
N	18	25
Mean (SD)	427 (655)	277 (592)
Median	123	295
IQR	13–891	-140–815
Range	-330–2110	-730–1460

95% CI for the mean treatment difference (based on 2-sample t-test): (-237, 536)

Appendix 3. Analysis of covariance of MRS: Outcome = change from baseline to final visit; treatment group as a classification factor, Rapid cycling and baseline MRS as covariates

```
Residuals:
   Min     1Q Median     3Q    Max
-9.583 -3.131 -1.337  2.438 17.417

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    3.47003    0.97325   3.565  0.00055 ***
AD vs placebo -1.84167    1.01818  -1.809  0.07337 .
rapidcycling1  3.68357    1.09386   3.367  0.00106 **
MRS.Wk0       -0.69631    0.08603  -8.093 1.15e-12 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘
’ 1

Residual standard error: 5.221 on 104 degrees of freedom
  (11 observations deleted due to missingness)
Multiple R-squared:  0.4486,    Adjusted R-squared:  0.4327
F-statistic:  28.2 on 3 and 104 DF,  p-value: 2.002e-13
```

Estimated overall treatment effects: mean (95%CI):

```
> emmeans(ancova.mrs, ~Group)
 Group  emmean    SE  df lower.CL upper.CL
 1      0.0508 0.770 104    -1.48    1.579
 2     -1.7909 0.723 104    -3.22   -0.358

Results are averaged over the levels of: rapidcycling
Confidence level used: 0.95
```

Estimated treatment effects for rapid-cycling and non-rapid-cycling groups: mean (95%CI):

```
> emmeans(ancova.mrs, ~Group*rapidcycling)
 Group rapidcycling  emmean    SE  df lower.CL upper.CL
 1     0            -1.7910 0.785 104   -3.347   -0.235
 2     0            -3.6326 0.793 104   -5.206   -2.059
 1     1             1.8926 1.081 104   -0.252    4.037
 2     1             0.0509 1.006 104   -1.945    2.047

Confidence level used: 0.95
```