Antidepressant Discontinuation in Bipolar Depression: A Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD) Randomized Clinical Trial of Long-Term Effectiveness and Safety

S. Nassir Ghaemi, MD, MPH; Michael M. Ostacher, MD, MPH; Rif S. El-Mallakh, MD; David Borrelli, MD†; Claudia F. Baldassano, MD; Mary E. Kelley, PhD; Megan M. Filkowski, BA; John Hennen, PhD†; Gary S. Sachs, MD; Frederick K. Goodwin, MD; and Ross J. Baldessarini, MD

Objective: To assess long-term effectiveness and safety of randomized antidepressant discontinuation after acute recovery from bipolar depression.

Method: In the Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD) study, conducted between 2000 and 2007, 70 patients with *DSM-IV*–diagnosed bipolar disorder (72.5% non–rapid cycling, 70% type I) with acute major depression, initially responding to treatment with antidepressants plus mood stabilizers, and euthymic for 2 months, were openly randomly assigned to antidepressant continuation versus discontinuation for 1–3 years. Mood stabilizers were continued in both groups.

Results: The primary outcome was mean change on the depressive subscale of the STEP-BD Clinical Monitoring Form. Antidepressant continuation trended toward less severe depressive symptoms (mean difference in *DSM-IV* depression criteria = −1.84 [95% CI, −0.08 to 3.77]) and mildly delayed depressive episode relapse (HR = 2.13 [1.00–4.56]), without increased manic symptoms (mean difference in *DSM-IV* mania criteria = +0.23 [−0.73 to 1.20]). No benefits in prevalence or severity of new depressive or manic episodes, or overall time in remission, occurred. Type II bipolar disorder did not predict enhanced antidepressant response, but rapid-cycling course predicted 3 times more depressive episodes with antidepressant continuation (rapid cycling = 1.29 vs non–rapid cycling = 0.42 episodes/year, $P = .04$).

Conclusions: This first randomized discontinuation study with modern antidepressants showed no statistically significant symptomatic benefit with those agents in the long-term treatment of bipolar disorder, along with neither robust depressive episode prevention benefit nor enhanced remission rates. Trends toward mild benefits, however, were found in subjects who continued antidepressants. This study also found, similar to studies of tricyclic antidepressants, that rapid-cycling patients had worsened outcomes with modern antidepressant continuation.

Trial Registration: clinicaltrials.gov Identifier: NCT00012558

J Clin Psychiatry 2010;71(4):372–380

Submitted: *November 19, 2008; accepted May 12, 2009 (doi:10.4088/JCP.08m04909gre).*

†Deceased.

Corresponding author: *S. Nassir Ghaemi, MD, MPH, Mood Disorders Program, Department of Psychiatry, Tufts Medical Center, 800 Washington St, Box 1007, Boston, MA 02111 (nghaemi@tuftsmedicalcenter.org).*

In the United States, most persons with bipolar disorder receive antidepressants as initial treatment, often without mood stabilizers, and frequently long-term.[1,2] This practice may be an understandable response to depressive episodes, or chronic subsyndromal depression, a common outcome in treated bipolar illness.[2–4] Despite being particularly difficult to treat,[3,5] associated with comorbidities,[6] disability,[7] cognitive dysfunction,[8] and suicide,[9] bipolar depression remains poorly studied,[10] with effectiveness and safety of antidepressants, particularly long-term, uncertain.[10–12] Resolving these questions is a public health challenge of high priority.

Some,[12] but not all,[13] randomized clinical trials (RCTs), including modern antidepressants (like serotonin reuptake inhibitors, bupropion, and venlafaxine), indicate probable short-term efficacy in acute bipolar depression, as well as at least moderate risk of inducing manic or mixed states with some agents.[14]

Once antidepressants are initiated for acute treatment, the question of how long they should be continued arises. Prior RCTs of antidepressant discontinuation are limited to tricyclic antidepressants, all of which found no benefit from continuing antidepressants compared to lithium.[11] With modern antidepressants, some nonrandomized observational studies report benefit from continuing antidepressants after recovery from the acute major depressive episode.[15,16] Other observational data fail to find such benefit.[17]

This is the first RCT to assess discontinuation of modern antidepressants after acute treatment for bipolar depression. Our main hypothesis was that antidepressant continuation would have mild to moderate benefits in depressive symptom reduction in bipolar disorder.

For Clinical Use

- The first randomized discontinuation study with modern antidepressants showed no statistically significant symptomatic benefit with those agents in the long-term treatment of bipolar disorder.
- Trends toward modest symptomatic benefits were found in subjects who continued antidepressants.
- Patients with rapid-cycling had worsened outcomes with continuation of modern antidepressants.

METHOD

Study Design

This report provides final results of an unblinded, randomized trial within the Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD) study cohort.[18] Subjects were patients with a *Diagnostic and Statistical Manual of Mental Disorders*, Fourth Edition (*DSM-IV*)[19] diagnosis of bipolar disorder (N = 70) who achieved clinical recovery (at least 2 months of euthymia) from an index episode of acute bipolar major depression while treated with an antidepressant and a mood stabilizer. They were then randomly assigned to antidepressant continuation (n = 32) or antidepressant discontinuation (n = 38), while their mood stabilizer was continued, for up to 3 years.

Study Subjects

Between 2000 and 2007, patients were recruited within the STEP-BD study from 4 collaborating sites: Cambridge Health Alliance (CHA; Cambridge, Massachusetts), the University of Pennsylvania Hospital (Philadelphia, Pennsylvania), University of Louisville Medical Center (Louisville, Kentucky), and Massachusetts General Hospital (Boston, Massachusetts). In the final year of the study, further subjects were also recruited at Emory University School of Medicine (Atlanta, Georgia). There were no detectable site differences in effects. Diagnoses included *DSM-IV* bipolar disorder types I, II, or not otherwise specified (NOS). Mood stabilizers allowed were lithium carbonate, divalproex, carbamazepine, and lamotrigine. Other putative mood-stabilizing agents were allowed in type II/NOS subjects if past inefficacy or intolerance had occurred with all of the preceding 4 agents. In 53% of study patients, an antidepressant was added after they became depressed while taking a mood stabilizer; another 13% were taking antidepressants without benefit for depression and later given a mood stabilizer; 7% had received no recent medication (within 6 months) and were given a mood stabilizer and antidepressant simultaneously for depression; 6% received mood stabilizer and antidepressant combinations with continued depression that responded to alterations in type or dose of 1 or both agents; details regarding immediate prior use of an antidepressant or mood stabilizer were not available in 21% of the sample.

Study Procedures

Simple randomization with a computer-generated list was conducted. Specific antidepressants were chosen by agreement between each patient and treating physician and prescribed in accord with local clinical practice to treat an index episode of major depression to the point of remission as determined by total scores ≤ 8 on the 21-item Hamilton Depression Rating Scale,[20] sustained for ≥ 8 weeks. Discontinuation of antidepressant treatment was by gradual dose reduction to 0 mg/d over 1–4 weeks (average 2 weeks). Other currently prescribed psychotropic agents (excluding any nonstudy antidepressants) could continue and be used or changed at the discretion of each patient's prescribing physician, as in standard clinical practice (specific agents used are described in results).

Clinical Assessments

The primary clinical assessment instrument was the STEP-BD Clinical Monitoring Form (CMF), an outcome assessment instrument with extensive testing for reliability and validity, described in more detail elsewhere,[21] in which depressive and manic symptoms are rated on a severity scale from −2 to +2, with 0 meaning no symptoms, and +1 or −1 meaning *DSM* threshold criteria. Clinical Monitoring Form depressive and manic scores correlated strongly with Montgomery-Åsberg Depression Rating Scale[22] (mean $r = 0.87$) and Young Mania Rating Scale[23] (mean $r = 0.84$) scores, respectively. A total CMF score was calculated by adding the item scores for the 9 depressive symptoms (with the absolute value of each item used) for the CMF depression score and the 7 mania symptoms for the CMF mania score.

To assess potential open-label bias, patients completed a 4-question visual analog scale at randomization, with [−] scores indicating negative, 0 meaning neutral, and [+] indicating positive attitudes toward antidepressants (range, −3 to +3). The measures demonstrated a generally positive attitude toward antidepressant treatment throughout the sample (Table 1).

Primary and Secondary Outcomes

The primary outcome was mean change on the depressive subscale of the CMF. The study evaluated subsyndromal as

Table 1. Baseline Characteristics of Subjects: Antidepressants Continued Versus Discontinued

Measure	All Subjects	Continued	Discontinued	Difference or Ratio[a] (95% CI)
Subjects enrolled, N	70	32	38	…
Men, %	47.1	53.1	42.1	0.64 (0.22 to 1.83)
Age, mean ± SD, y	43.0 ± 13.4	45.6 ± 14.1	40.8 ± 12.8	4.78 (–1.42 to 11.0)
Age-at-onset, mean ± SD, y	18.7 ± 9.7	19.2 ± 10.2	18.3 ± 9.2	0.91 (–3.92 to 5.74)
Baseline CMF scores, mean ± SD				
Total	1.94 ± 1.86	2.25 ± 1.84	1.68 ± 1.86	0.57 (–0.32 to 1.45)
Depression	1.45 ± 1.40	1.77 ± 1.44	1.18 ± 1.33	0.59 (–0.07 to 1.25)
Mania	0.51 ± 0.84	0.52 ± 0.79	0.50 ± 0.89	0.02 (–0.38 to 0.43)
Patient AD attitude score, mean ± SD	3.14 ± 2.43	3.06 ± 0.00	3.21 ± 0.00	–0.15 (–1.36 to 1.06)
Type I bipolar disorder, %	70.0	68.8	71.1	1.12 (0.35 to 3.51)
Prior psychosis, %	22.9	21.9	23.7	0.90 (0.25 to 3.18)
Prior substance use, %	42.9	37.5	47.4	0.67 (0.23 to 1.93)
Prior AD-associated mania, %	38.5	44.8	33.3	1.63 (0.53 to 5.02)
Prior rapid cycling, %	24.6	22.6	26.3	0.82 (0.23 to 2.82)
Prior AD-associated rapid cycling, %	16.4	18.2	15.2	1.24 (0.22 to 6.64)
Prior loss of AD benefit, %	30.8	33.3	29.0	1.22 (0.31 to 4.70)
AD reassigned in trial, %				
All reassignments	47.1	46.9	47.4	0.98 (0.34 to 2.79)
Weeks to reassignment, mean ± SD	18.11 ± 12.04	21.83 ± 14.62	15.02 ± 8.62	6.81 (–1.54 to 15.17)

[a]Mean differences (continuous variables) and risk ratios (categorical variables) consider the antidepressant discontinued arm as the reference group.
Abbreviations: AD = antidepressant, CMF = Clinical Monitoring Form.

well as syndromal depression. We addressed the less-studied subsyndromal component as the primary outcome, due to its clinical importance[24] and also to increase statistical power (as a continuous measure). We also focused a priori on outcomes in the first 12 months as dropouts were expected to be higher at longer follow-up.

Secondary outcome measures included depressive and manic subscores of morbidity ratings, the frequency and severity of new episodes, weeks to new episodes, and weeks in remission. Two a priori subgroup analyses were planned to avoid inflating positive findings (type II error): rapid cycling (≥ 4 recurrences within the previous year) and bipolar disorder diagnostic subtype, based on reports that rapid cycling worsens, and type II bipolar disorder improves, antidepressant responses.[17]

Ethical Considerations

The study procedures and consent forms were approved by the institutional review boards of the collaborating sites. The study was not blinded to limit ethical risks that might arise from either continuing or stopping antidepressant treatment over the prolonged follow-up, as well as to enhance generalizability by allowing commonly employed treatments that each patient and treating clinician was free to select. Moreover, the protocol allowed for discontinuing or restarting antidepressant treatment on ethical grounds, based on clinical judgment. Such patients continued to be analyzed in the original randomized group at 1-year outcome, using intent-to-treat (ITT) methods (see below).

Enrollment and Generalizability

The study sample consisted primarily of patients treated in academic specialty clinics. At the CHA and Emory sites, where 51% (n = 36) of patients were recruited, another 55 patients were excluded as follows: patient not interested in participating in the study or refused protocol treatment conditions (34.6%), actively abused substances currently or within 1 month (21.8%), was considered unlikely to be compliant with appointments or lived far away (18.2%), did not meet *DSM-IV* criteria for bipolar disorder (14.5%), remained depressed (12.7%) or became manic (1.8%) with antidepressant treatment, or was lost to follow-up (3.6%); several patients met more than 1 exclusion criterion. Thus, overall, 39.6% (36/91) of patients initially treated for acute bipolar depression in these 2 sites ultimately entered the randomized discontinuation protocol. Excluded subjects entered alternative STEP-BD research protocols or continued standard clinical care.

Statistical Considerations

The primary outcome of CMF change and the secondary outcomes of time to relapse, time in remission, and number of mood episodes were all planned a priori. The subgroup analyses of rapid cycling and type II patients were also planned a priori. Other subgroup analyses were conducted post hoc. In the context of a pilot study, the planned a priori secondary and subgroup analyses do not warrant correction for multiple comparisons, since this study is not definitively testing those hypotheses but, rather, examining their effect sizes.[25] Thus, confidence intervals are reported in all those results, and sole focus on *P* values would be unwarranted. Since CMF ratings had a high proportion of zero values (eg, 60% of mania ratings at baseline), which tend to limit the value of mean scores, we also dichotomized (present/absent) time-specific CMF measures of depression and mania, following commonly employed precedents[26,27] including both means

and proportions in longitudinal assessments of morbidity. In addition, cyclic or random variation in the CMF measures over time precluded use of linear growth-curve analysis. Accordingly, to test time-related group differences, we considered the data in 6-month intervals, using available-case, random-effects, mixed models, with time as a categorical covariate rather than a continuous function. Secondary outcomes were assessed with standard methods, using Poisson tests for episode counts, Wilcoxon tests for continuous measures, and log rank (χ^2) tests for time-to-event measures.

Intent-to-treat analyses with a primary endpoint of 12 months were employed, irrespective of how long patients remained on their original antidepressant continuation or discontinuation randomized assignment. Treatment assignment was changed in 33/70 subjects (47.1%), about equally in both arms (Table 1; 18/38, 47.4%, clinically were prescribed antidepressants after initial randomization to antidepressant discontinuation; 15/32, 46.9%, clinically stopped antidepressants after initial randomization to antidepressant continuation), though somewhat earlier in those who clinically were prescribed antidepressants after initial randomization to antidepressant discontinuation (mean ± SD = 15.0 ± 8.62 vs 21.8 ± 14.6 weeks in those who clinically stopped antidepressants after initial randomization to antidepressant continuation). In the antidepressant discontinuation arm, treatment reassignments were associated almost exclusively with the clinical impression of newly emerging depression (17/18, 94%; 1 case due to patient choice). Among patients randomly assigned to continue antidepressant treatment, the most common reason to stop antidepressants was for emergence of hypomanic/manic/mixed states (7/15, 47%), followed by patient choice (5/15, 33%) and new depression (3/15, 20%).

The statistical literature[27–29] indicates that ITT analysis is less biased and generally more conservative than completer or other analyses that do not preserve initial randomization. Alternatives to ITT analysis in this study would have been a censoring of subjects after change in randomization, which would have markedly reduced sample size and power (only 43% of the original sample would have remained), or conducting post hoc "as-treated" (or "per-protocol") analyses of non–randomly assigned patients, using all available data and coding for change in treatment assignment when it happened. The former post hoc censoring analysis would be prone to false-negative results due to insufficient power, and the "as-treated" approach would tend to yield false-positive results, with inflated effect sizes due to violation of randomization. Nonetheless, to see if they were similar to ITT results, post hoc non-ITT analyses were conducted and are reported.

RESULTS

Characteristics of the Sample

At intake, 70 patients were randomly assigned to continue (n = 32, antidepressant continuation group) or to discontinue (n = 38, antidepressant discontinuation group) treatment with antidepressants after attaining sustained recovery from an index episode of acute major depression. Demographic and clinical features of the sample are in Table 1. The dropout rate was 61.4% (43/70) by 12 months and 88.6% (62/70) by 3 years.

The most frequently employed antidepressant class was serotonin reuptake inhibitors (52%). Common specific agents were bupropion and paroxetine (22% each) and citalopram and venlafaxine (19% each). No tricyclic antidepressants were used. Choices of mood stabilizers ranked: lithium carbonate (44%) > lamotrigine (41%) > divalproex (23%), with a total > 100%, since some patients received > 1 mood stabilizer. Among other psychotropics, 39% of patients also received atypical neuroleptics, most commonly quetiapine (17%), followed by risperidone (10%) and aripiprazole (9%). Only 1 patient received a traditional neuroleptic (haloperidol). No or minor differences existed between the 2 randomized arms in distribution of mood stabilizers or neuroleptics (eg, lamotrigine was used in 47% of antidepressant continuation group vs 37% in antidepressant discontinuation group; quetiapine was used in 16% of antidepressant continuation group vs 18% in antidepressant discontinuation group.) Post hoc analyses did not find any notable changes in main outcomes after adjustment for specific mood stabilizers or neuroleptics used.

Primary Outcome

The primary outcome was mean change on the depressive subscale of the CMF. Intent-to-treat analysis of CMF depressive scores over time showed no difference between groups, but with a trend toward moderate benefit with antidepressant continuation in the first 12 months (Table 2).

Secondary Outcomes

Secondary ITT analysis of the prevalence, as opposed to severity (the primary outcome), of mood symptoms (comparing any depressive or manic symptoms vs none over the first year) again found no differences between groups, with minimal benefit with antidepressant continuation (relative risk of CMF depression between groups at 12 months: OR = 4.51 [95% CI, 0.40–51.0]; relative risk of CMF mania between groups at 12 months: OR = 1.23 [95% CI, 0.08–19.6]). As shown in Table 3, other secondary outcomes, except for survival analysis (see below), also found no or little benefit with antidepressant continuation: specifically, there was no benefit for episode incidence or time in remission. Descriptively, patients spent most of the follow-up year symptomatic (Table 3), and new bipolar disorder episodes occurred in 54.3% of study patients: 45.7% experienced at least 1 depressive episode, 15.7% a manic episode, and 8.6% a mixed episode within the first year.

As seen in Table 3 and in contrast to the above secondary outcomes, Kaplan-Meier survival analysis found benefit with antidepressant continuation for delay in occurrence

Table 2. Severity of Symptoms With Versus Without Long-Term Antidepressant Treatment

	Continued[a]			Discontinued[a]			Depressive[b] Difference	Manic[b] Difference
Month	n[c]	Depressive	Manic	n[c]	Depressive	Manic	(95% CI)	(95% CI)
0	32	1.78 ± 1.44	0.52 ± 0.79	38	1.33 ± 1.47	0.54 ± 0.90	−0.45 (−1.54 to 0.64)	0.02 (−0.54 to 0.58)
6	19	2.35 ± 2.67	1.10 ± 1.72	17	3.50 ± 3.32	0.66 ± 1.59	1.15 (−0.34 to 2.63)	−0.44 (−1.16 to 0.36)
12	10	1.93 ± 2.26	0.48 ± 0.57	16	3.25 ± 3.56	0.73 ± 1.44	1.32 (−0.38 to 3.16)	0.25 (−0.66 to 1.15)
Change by 6 mo[b]	19	0.57 ± 0.59	0.61 ± 0.30	17	2.17 ± 0.61	0.19 ± 0.30	1.60 (−0.07 to 3.26)	−0.41 (−1.25 to 0.42)
Change by 12 mo[b]	10	0.15 ± 0.76	0.04 ± 0.38	16	1.99 ± 0.62	0.27 ± 0.31	1.84 (−0.08 to 3.77)[d]	0.23 (−0.73 to 1.20)[d]

[a]Data are mean ± SD of Clinical Monitoring Form–based ratings of the severity of depressive or manic symptoms at defined times, up to 1 year of follow-up.
[b]Model estimates (mean ± SD) for differences in symptomatic ratings are adjusted (random-effects) for individual-level variation in baseline morbidity.
[c]Number of patients remaining at each time.
[d]Significance of 12-month interaction effects: depression, $P = .06$; mania, $P = .64$.

Table 3. One-Year Outcomes: Antidepressant Treatment Continued Versus Discontinued

Outcome	Continued[a]	Discontinued[a]	Mean Difference (95% CI)
No. of episodes			
All episodes	1.00 ± 1.50	0.97 ± 1.20	0.03 (−0.62 to 0.67)
Manic/hypomanic episodes	0.31 ± 0.78	0.11 ± 0.31	0.20 (−0.07 to 0.48)
Depressive episodes	0.59 ± 0.91	0.79 ± 1.14	−0.20 (−0.70 to 0.30)
Mixed episodes	0.09 ± 0.30	0.11 ± 0.39	−0.02 (−0.18 to 0.16)
% Weeks depressed[b]	76.2 ± 25.3	80.5 ± 25.3	−4.30 (−16.4 to 7.80)
% Weeks manic/hypomanic[b]	50.5 ± 29.5	41.7 ± 33.5	8.80 (−6.40 to 23.9)
			Hazard Ratio (95% CI)
Weeks to first new episode	34.7 ± 3.4	28.5 ± 3.3	1.53 (0.80 to 2.94)
Weeks to first depressive episode	41.4 ± 3.0	31.5 ± 3.3	2.13 (1.00 to 4.56)*
Weeks to first manic episode	45.2 ± 2.8	47.0 ± 2.3	0.84 (0.24 to 2.89)

[a]Data are mean ± SD except for time to event outcomes estimated as mean ± SE. From baseline to 12 months, Ns ranged from 32 to 11 in the continuation group and 38 to 16 in the discontinuation group.
[b]Percent weeks manic/hypomanic + depressed exceeds 100% due to subsyndromal and syndromal mixed states.
*$P < .05$.

of a depressive episode (mean ± SE = 41.4 ± 3.0 in antidepressant continuation vs 31.5 ± 3.3 weeks in antidepressant discontinuation; $\chi^2 = 4.01$, $P = .045$), though less for delay of overall mood episodes, including manic episodes (mean ± SE latency to a first recurrence of any polarity, 34.7 ± 3.4 vs 28.5 ± 3.3 weeks; $\chi^2 = 1.69$, $P = .19$). Most new episodes in the first year were depressive (32/43 = 74.4%), compared to only 9 cases of mania or hypomania (20.9%) and only 2 mixed episodes (4.65%). Some of this apparent delay of illnesses of the same polarity as the index episode probably represented relapses into the recent depressive episode, since illness latency was small and, when events in the first 2 months were removed, the beneficial effect of continued antidepressant treatment was more limited (with vs without antidepressant: 42.6 ± 2.9 vs 36.0 ± 3.1 weeks; $\chi^2 = 2.20$, $P = .138$).

Moderators of Treatment Effects

Two a priori subgroup analyses were planned: rapid cycling and bipolar diagnostic type (I vs II). As shown in Figure 1, a significant interaction between randomized treatment group and rapid cycling was found for the number of depressive episodes, with 3-fold more depressive recurrences/year in the antidepressant continuation group (rapid cycling = 1.29 vs non–rapid cycling = 0.42 episodes/year), but not among the antidepressant discontinuation group (rapid cycling = 0.82 vs non–rapid cycling = 0.70 episodes/year; statistical difference is significant for an association between rapid-cycling status and antidepressant use and of major depressive episodes based on the interaction effect: $z = -2.04$, $P = .04$). Rapid cycling was itself also an independent predictor of poor prognosis (compared to non–rapid cycling: shorter median latency to episodes, 23.7 vs 33.9 weeks, adjusted HR = 3.1, $P = .03$; more depressive episodes within a year, 0.94 vs 0.63, $z = 2.45$, $P = .01$; and fewer weeks in remission, 66.9 vs 79.2, $F = 3.82$, $P = .06$).

Interactions of randomized treatment group and diagnostic type were not found with any secondary outcome measure, including latency to a new depressive episode (adjusted hazard ratio [HR] = 0.66, $P = .57$), number of depressive episodes ($z = -1.05$, $P = .30$), or percent time in depressive illness ($F = 0.43$, $P = .51$).

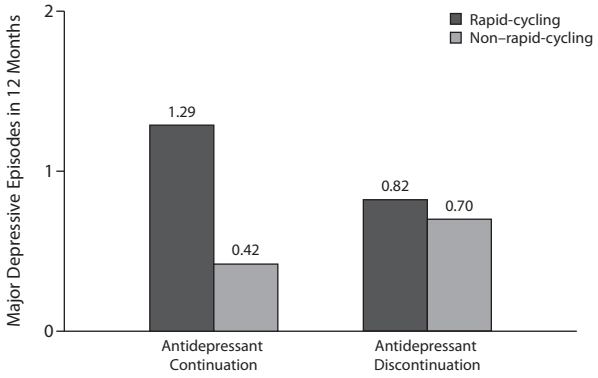
Post Hoc Non–Intent-to-Treat Analyses

To address the question of whether patient- or clinician-driven changes in randomized treatment strategy may have affected the ITT outcomes, as-treated (or per-protocol) analyses were conducted and no longer found the modest benefits for depressive symptoms (adjusted CMF change at 12 months = −0.72; 95% CI, −2.71 to 1.26) and time to depressive relapse (HR = 0.71; 95% CI, 0.34–1.46) seen with antidepressant continuation in the ITT analyses.

DISCUSSION

This is the first long-term RCT of modern antidepressant discontinuation in bipolar disorder, just after recovery from a major depressive episode. In the context of an enriched sample (including only those who tolerated antidepressants, without major side effects or induction of

Figure 1. Increase in *DSM-IV* Major Depressive Episodes Over 1 Year With Antidepressant Continuation in Rapid-Cycling Bipolar Disorder (N = 17)[a]



[a]Significance of 12-month interaction effect between randomized treatment groups: $z = -2.04$, $P = .04$.

mania/hypomania, and subsequently achieved a durable recovery, remaining euthymic for at least 2 months), antidepressant continuation may mildly delay new depressive episodes in bipolar disorder, without increasing manic morbidity, with a trend toward limiting depressive morbidity. However, there was no decrease in prevalence or severity of new depressive episodes and no increased time in remission. Planned secondary analyses found that prior rapid cycling was associated with more depressive illness overall, as expected,[30,31] but much more in association with continued antidepressant treatment, suggesting an interaction of risk factors. No specific benefit or risk was encountered with antidepressant use in type II bipolar disorder.

These results extend results from the only previous, double-blind, antidepressant-discontinuation RCT, which involved tricyclic antidepressants in type I bipolar disorder.[11] It found little benefit in depressive prevention, but greater manic risk, when imipramine was added to or compared to lithium alone. With modern antidepressants, the main prior long-term antidepressant discontinuation study,[15] reported from the Stanley Network, was not randomized; it found that early relapses into bipolar depression were more likely after stopping antidepressant treatment, especially within 6 months of recovery from the acute depressive episode. The present results agree with the general direction of the Stanley findings, but with a smaller effect size of benefit for delayed relapse and with little or no benefit for overall prevalence or severity of depressive episodes or time in remission. Our findings also indicate worse antidepressant outcomes in rapid-cycling bipolar disorder, a group excluded from the Stanley study.

Further, these results should be interpreted in the context of a recent STEP-BD study, the largest RCT of antidepressants in acute bipolar depression treated with standard mood stabilizers.[13] In that report, modern antidepressants (bupropion or paroxetine) were not more effective than placebo acutely, with about 25% of patients improving to remission overall. Similar low efficacy rates were seen in the only maintenance RCT with modern antidepressants (bupropion, sertraline, or venlafaxine added to standard mood stabilizers) prior to this study, in which only 15% of patients remained euthymic for up to 1 year, with little difference among antidepressants.[32] Our results agree with both reports, since only about 40% of patients initially treated for acute bipolar depression entered our study (see methods), and only a portion of that group experienced modest antidepressant continuation benefits. The effect size was modest because it only involved benefit in about 2 depressive criteria, with 5 or more criteria reflecting a full depressive syndrome, and did not reach statistical significance. Since subsyndromal depression is a major problem in the long-term course of bipolar disorder, nevertheless, this mild benefit may be useful. On the other hand, it is not robust enough to support larger claims about the benefits of antidepressants, such as the belief that they may produce complete remission or that they are protective in prevention of full depressive episodes.

Another relevant feature is that these results represent the average results for the entire sample. If a small subgroup had notable benefit, but most patients had little or none, then this apparent modest effect overall would be diluted. In a larger sample, multivariable predictive models might be able to pick out the features of such a potential responsive subgroup. It is possible that the modest antidepressant benefits seen in this study might be generalizable to a minority of the bipolar population, perhaps best estimated at about 20% of patients, far below the 50%–80% antidepressant usage rate routinely seen in practice-pattern studies across many nations.[4,15,33]

The present observations in patients with rapid-cycling bipolar disorder may be particularly important clinically, since previous, observational studies have yielded inconsistent findings concerning antidepressant effects in rapid-cycling patients.[34–38] The only previous RCT, using a double-blind on-off-on-off design, found more recurrences with tricyclic antidepressants than placebo.[39] Our study is a randomized replication of that study with modern antidepressants, and further shows that a mood-destabilizing effect of antidepressants increases risk of recurrent depression as well as mania, even despite concomitant mood stabilizer treatment. Since the rapid-cycling subgroup in our study was small, these positive secondary outcomes should be replicated again, if possible, with a larger study, specifically in a rapid-cycling population.

In contrast, our failure to confirm improved antidepressant responses in type II versus type I bipolar disorder contradicts some other randomized studies, which are either much smaller than the present study[40] or do not use mood stabilizer cotherapy.[41]

Methodological Considerations

All studies have limitations, but their relevance depends on the context of the clinical literature. This study is an improvement over others in the literature because it is randomized, unlike all reports except one (which used tricyclic antidepressants and, thus, is not generalizable to new antidepressants, as in this study).[39] Thus, the methodological limitations of this study need to be weighed against the reality of absence of better data. Although a more homogeneous sample (perhaps only bipolar type I, perhaps only a single antidepressant or a single mood stabilizer) might have allowed for more internal validity, such homogeneity is not what occurs in clinical practice and, thus, would have severely limited the generalizability of the results, a common critique of RCTs.[42] We acknowledge that although the sample size is large enough to detect moderate effects on the primary outcome, it is small for subgroup analysis. However, this would likely affect only negative results and not positive ones, such as the rapid-cycling interaction shown here.[43] Lack of blinding allowed greater generalizability in this study; and confounding bias, corrected by randomization, is generally viewed as a greater bias than measurement bias, corrected by blinding.[25] Thus, open randomization is notably more valid than nonrandomized data, and fully blinding for a single antidepressant may be a useful step in the future, after showing results generalizable to most antidepressants, as in this study.

In other words, this was a randomized trial with an effectiveness design, that is, a randomized trial conducted in a real-world population openly and naturalistically, not a standard, double-blind, randomized efficacy study conducted in a highly selective research cohort. The STEP-BD was, in fact, designed to be a platform for just this kind of effectiveness trial, which has the advantage of moving randomized data closer to the real world, making it more generalizable to actual clinical practice. Rather than being limited to the rarefied RCT patient population so common in pharmaceutical industry–sponsored trials, the purpose of this trial was to inform actual clinical practice.

The dropout rate, although high, is better than most randomized maintenance studies of bipolar disorder (61% within 12 months here).[44] As with all randomized clinical trials, one cannot ethically force patients to remain on randomized treatments. Change in randomized treatment, after the study begins, is common with many types of research, most notably surgical trials.[45] In this study, such change in randomization appeared roughly equal in both subgroups, indicating at least a limited bias in treatment-related change. Intent-to-treat analyses, as mentioned above, are standard practice in clinical trials,[45] preserve randomization, and allow us to say something about the real-world results based on how clinicians intend to treat their patients. The alternative, an "as treated" analysis, is known to be biased in favor of the experimental treatment.[45] Therefore, despite the design concerns, the most adequate analysis is the ITT analysis, and in this particular case, any bias would have been against the experimental intervention: antidepressant discontinuation.[27] Since, in this study, antidepressant discontinuation was mildly less beneficial for subsyndromal depressive symptoms than antidepressant continuation, any ITT-related bias would be in underreporting rather than overreporting benefits with antidepressant discontinuation.[27] As always, replication is the best solution; and, thus, further trials including these newer agents should be performed.

CONCLUSIONS

This first randomized discontinuation study with modern antidepressants found no significant symptomatic benefit with those agents in the long-term treatment of bipolar disorder, along with neither robust depressive episode prevention benefits nor enhanced remission rates. Trends toward subsyndromal benefits, however, were found in subjects who continued antidepressants. Given the other STEP-BD data suggesting no benefit for the use of adjunctive antidepressants for acute bipolar depression, this study does not lend robust support for the use of standard antidepressants in the maintenance treatment of bipolar disorder. It also found, similar to tricyclic antidepressants, that rapid-cycling patients had worsened outcomes with serotonin reuptake inhibitors and other modern antidepressant continuation.

Drug names: aripiprazole (Abilify), bupropion (Aplenzin, Wellbutrin), carbamazepine (Carbatrol, Equetro, and others), citalopram (Celexa and others), divalproex (Depakote and others), haloperidol (Haldol and others), imipramine (Tofranil and others), lamotrigine (Lamictal and others), lithium (Eskalith, Lithobid, and others), paroxetine (Paxil, Pexeva, and others), quetiapine (Seroquel), risperidone (Risperdal and others), sertraline (Zoloft and others), venlafaxine (Effexor and others).
Disclosure of off-label usage: The authors have determined that, to the best of their knowledge, aripiprazole, bupropion, carbamazepine, citalopram, divalproex, haloperidol, imipramine, lamotrigine, lithium, paroxetine, quetiapine, risperidone, sertraline, and venlafaxine are not approved by the US Food and Drug Administration for the treatment of bipolar depression.
Author affiliations: Mood Disorders Program, Department of Psychiatry, Tufts Medical Center (Dr Ghaemi); Department of Psychiatry, Harvard Medical School (Drs Ostacher, Borrelli, Hennen, Sachs, and Baldessarini); Bipolar Clinic and Research Program, Massachusetts General Hospital (Drs Ostacher, Borrelli, and Sachs), Boston, Massachusetts; University of Louisville School of Medicine, Kentucky (Dr El-Mallakh); Hospital of the University of Pennsylvania, Philadelphia (Dr Baldassano); Department of Biostatistics, Rollins School of Public Health (Dr Kelley), and Department of Psychiatry (Ms Filkowski), Emory University, Atlanta, Georgia; International Consortium for Bipolar Disorder Research, McLean Division of Massachusetts General Hospital, Belmont (Drs Baldessarini and Hennen); and George Washington University School of Medicine, Washington, DC (Dr Goodwin).
Financial disclosure: **Dr Ghaemi** has received honoraria in the past from Pfizer, GlaxoSmithKline, AstraZeneca, Bristol-Myers Squibb, Janssen, and Abbot; has received grants from Pfizer; and has grants pending with both Pfizer and Janssen. He is not a member of speakers bureaus nor a paid consultant to any pharmaceutical company, and he holds no equity positions in pharmaceutical or biomedical companies. **Dr Ostacher** is a consultant for Pfizer; has received grant/research

support from the National Institute on Alcohol Abuse and Alcoholism; and is a member of the speakers/advisory boards for AstraZeneca, Bristol-Myers Squibb, Eli Lilly, and Pfizer. **Dr El-Mallakh** is a member of the speakers/advisory boards for AstraZeneca, Abbott, Bristol-Myers Squibb, GlaxoSmithKline, Pfizer, and Eli Lilly. **Dr Baldassano** is a member of the speakers/advisory boards for Pfizer, GlaxoSmith-Kline, and AstraZeneca. **Dr Sachs** has received research support from Abbott, AstraZeneca, Bristol-Myers Squibb, Eli Lilly, GlaxoSmithKline, Janssen, Memory Pharmaceuticals, National Institute of Mental Health, Novartis, Pfizer, Repligen, Shire, and Wyeth; is a member of the speakers bureau for Abbott, AstraZeneca, Bristol-Myers Squibb, Eli Lilly, GlaxoSmithKline, Janssen, Memory Pharmaceuticals, Novartis, Pfizer, sanofi-aventis, and Wyeth; is a member of the advisory board/is a consultant for Abbott, AstraZeneca, Bristol-Myers Squibb, Cephalon, CNS Response, Elan, Eli Lilly, GlaxoSmithKline, Janssen, Memory Pharmaceuticals, Merck, Novartis, Organon, Otsuka, Pfizer, Schering-Plough, Sepracor, Repligen, sanofi-aventis, Shire, Sigma-Tau, Solvay, and Wyeth; and his spouse/partner is an equity/stock shareholder of Concordant Rater Systems. **Dr Goodwin** is a consultant for Pfizer and Bristol-Myers Squibb and has received honoraria from and is a member of the speakers/advisory board for GlaxoSmithKline. **Dr Baldessarini** has been a consultant for or collaborated in research with AstraZeneca, Auritec, Biotrofix, Janssen, JDS-Noven, Eli Lilly, Luitpold, Merck, NeuroHealing, Novartis, and SK-BioPharmaceutical Corporation and has taught or prepared educational materials for the New England Educational Institute and Pri-Med CME organizations, but is not a member of speakers' bureaus, and neither he nor family members hold equity positions in pharmaceutical or biomedical corporations. **Dr Kelley** and **Ms Filkowski** have no personal affiliations or financial relationships with any commercial interest to disclose relative to the article.

Funding/support: This work was supported by the National Institutes of Health grant MH-64189 (Dr Ghaemi) and a grant from the Bruce J. Anderson Foundation and by the McLean Private Donors Psychopharmacology Research Fund (Dr Baldessarini).

Acknowledgment: Our esteemed collaborators Drs David Borrelli, MD, from the Department of Psychiatry, Harvard Medical School, and the Bipolar Clinic and Research Program, Massachusetts General Hospital, Boston, Massachusetts, and John Hennen, PhD, from the Department of Psychiatry, Harvard Medical School, Boston, Massachusetts, and the International Consortium for Bipolar Disorder Research, McLean Division of Massachusetts General Hospital, Belmont, are now deceased. They deserve credit for their efforts in designing and conducting this complex study.

REFERENCES

1. Baldessarini RJ, Leahy L, Arcona S, et al. Patterns of psychotropic drug prescription for U.S. patients with diagnoses of bipolar disorders. *Psychiatr Serv*. 2007;58(1):85–91.
2. Baldessarini RJ, Henk HJ, Sklar AR, et al. Psychotropic medications for patients with bipolar disorder in the United States: polytherapy and adherence. *Psychiatr Serv*. 2008;59(10):1175–1183.
3. Judd LL, Schettler PJ, Akiskal HS, et al. Long-term symptomatic status of bipolar I vs bipolar II disorders. *Int J Neuropsychopharmacol*. 2003;6(2):127–137.
4. Baldessarini RJ, Salvatore P, Khalsa H-MK, et al. Morbidity in 303 first-episode bipolar I disorder patients. *Bipolar Disord*. 2010; In press.
5. Goodwin F, Jamison K. *Manic Depressive Illness*. 2nd ed. New York, New York: Oxford University Press; 2007.
6. Rihmer Z, Szádóczky E, Füredi J, et al. Anxiety disorders comorbidity in bipolar I, bipolar II and unipolar major depression: results from a population-based study in Hungary. *J Affect Disord*. 2001;67(1-3): 175–179.
7. Altshuler LL, Post RM, Black DO, et al. Subsyndromal depressive symptoms are associated with functional impairment in patients with bipolar disorder: results of a large, multisite study. *J Clin Psychiatry*. 2006;67(10):1551–1560.
8. Goldberg J, Burdick K. *Cognitive Dysfunction in Bipolar Disorder: A Guide for Clinicians*. Washington, DC: American Psychiatric Press; 2008.
9. Tondo L, Lepri B, Baldessarini RJ. Suicidal risks among 2826 Sardinian major affective disorder patients. *Acta Psychiatr Scand*. 2007;116(6):419–428.
10. El-Mallakh RS, Ghaemi SN. *Bipolar Depression*. Washington, DC: American Psychiatric Press; 2006.
11. Ghaemi SN, Lenox MS, Baldessarini RJ. Effectiveness and safety of long-term antidepressant treatment in bipolar disorder. *J Clin Psychiatry*. 2001;62(7):565–569.
12. Baldessarini RJ, Vieta E, Calabrese JR, et al. Bipolar depression: overview and commentary. *Harv Rev Psychiatry*. 2010; In press.
13. Sachs GS, Nierenberg AA, Calabrese JR, et al. Effectiveness of adjunctive antidepressant treatment for bipolar depression. *N Engl J Med*. 2007;356(17):1711–1722.
14. Goldberg JF, Truman CJ. Antidepressant-induced mania: an overview of current controversies. *Bipolar Disord*. 2003;5(6):407–420.
15. Altshuler L, Suppes T, Black D, et al. Impact of antidepressant discontinuation after acute bipolar depression remission on rates of depressive relapse at 1-year follow-up. *Am J Psychiatry*. 2003;160(7):1252–1262.
16. Joffe RT, MacQueen GM, Marriott M, et al. One-year outcome with antidepressant—treatment of bipolar depression. *Acta Psychiatr Scand*. 2005;112(2):105–109.
17. Ghaemi SN, Hsu DJ, Soldani F, et al. Antidepressants in bipolar disorder: the case for caution. *Bipolar Disord*. 2003;5(6):421–433.
18. Sachs GS, Thase ME, Otto MW, et al. Rationale, design, and methods of the Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD). *Biol Psychiatry*. 2003;53(11):1028–1042.
19. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*, Fourth Edition. Washington, DC: American Psychiatric Association; 1994.
20. Maier W, Buller R, Philipp M, et al. The Hamilton Anxiety Scale: reliability, validity and sensitivity to change in anxiety and depressive disorders. *J Affect Disord*. 1988;14(1):61–68.
22. Montgomery SA, Åsberg M. A new depression scale designed to be sensitive to change. *Br J Psychiatry*. 1979;134(4):382–389.
23. Young RC, Biggs JT, Ziegler VE, et al. A rating scale for mania: reliability, validity and sensitivity. *Br J Psychiatry*. 1978;133(5):429–435.
24. Judd LL, Akiskal HS, Schettler PJ, et al. The long-term natural history of the weekly symptomatic status of bipolar I disorder. *Arch Gen Psychiatry*. 2002;59(6):530–537.
25. Rothman KJ. *Epidemiology: An Introduction*. Oxford, United Kingdom: Oxford University Press; 2002.
26. Lachenbruch PA. Analysis of data with excess zeros. *Stat Methods Med Res*. 2002;11(4):297–302.
27. Bang H, Davis CE. On estimating treatment effects under non-compliance in randomized clinical trials: are intent-to-treat or instrumental variables analyses perfect solutions? *Stat Med*. 2007;26(5):954–964.
28. Tanaka Y, Matsuyama Y, Ohashi Y; MEGA Study Group. Estimation of treatment effect adjusting for treatment changes using the intensity score method: application to a large primary prevention study for coronary events (MEGA Study). *Stat Med*. 2008;27(10):1718–1733.
29. Peduzzi P, Detre K, Wittes J, et al. Intent-to-treat analysis and the problem of crossovers: an example from the Veterans Administration coronary bypass surgery study. *J Thorac Cardiovasc Surg*. 1991;101(3): 481–487.
30. Baldessarini RJ, Tondo L, Floris G, et al. Effects of rapid cycling on response to lithium maintenance treatment in 360 bipolar I and II disorder patients. *J Affect Disord*. 2000;61(1-2):13–22.
31. Tondo L, Hennen J, Baldessarini RJ. Rapid-cycling bipolar disorder: effects of long-term treatments. *Acta Psychiatr Scand*. 2003;108(1): 4–14.
32. Post RM, Altshuler LL, Leverich GS, et al. Mood switch in bipolar depression: comparison of adjunctive venlafaxine, bupropion and sertraline. *Br J Psychiatry*. 2006;189(2):124–131.
33. Ghaemi SN, Goodwin FK. Long-term naturalistic treatment of depressive symptoms in bipolar illness with divalproex vs lithium in the setting of minimal antidepressant use. *J Affect Disord*. 2001; 65(3):281–287.
34. Kukopulos A, Caliari B, Tundo A, et al. Rapid cyclers, temperament, and antidepressants. *Compr Psychiatry*. 1983;24(3):249–258.
35. Altshuler LL, Post RM, Leverich GS, et al. Antidepressant-induced mania and cycle acceleration: a controversy revisited. *Am J Psychiatry*. 1995;152(8):1130–1138.
36. 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.

37. Ghaemi SN, Boiman EE, Goodwin FK. Diagnosing bipolar disorder and the effect of antidepressants: a naturalistic study. *J Clin Psychiatry*. 2000;61(10):804–808, quiz 809.
38. Schneck CD, Miklowitz DJ, Miyahara S, et al. The prospective course of rapid-cycling bipolar disorder: findings from the STEP-BD. *Am J Psychiatry*. 2008;165(3):370–377, quiz 410.
39. Wehr TA, Sack DA, Rosenthal NE, et al. Rapid cycling affective disorder: contributing factors and treatment responses in 51 patients. *Am J Psychiatry*. 1988;145(2):179–184.
40. Parker G, Tully L, Olley A, et al. SSRIs as mood stabilizers for bipolar II disorder? a proof of concept study. *J Affect Disord*. 2006;92(2–3): 205–214.
41. Amsterdam JD, Brunswick DJ. Antidepressant monotherapy for bipolar type II major depression. *Bipolar Disord*. 2003;5(6):388–395.
42. Essock SM. Enhancing generalizability: stepping up to the plate. *Psychiatr Serv*. 2006;57(1):141, author reply 141–142.
43. Feinstein A. *Clinical Biostatistics*. St. Louis, MO: Mosby; 1977.
44. Calabrese JR, Rapport DJ. Mood stabilizers and the evolution of maintenance study designs in bipolar I disorder. *J Clin Psychiatry*. 1999;60(suppl 5):5–13, discussion 14–15.
45. Peduzzi P, Wittes J, Detre K, et al. Analysis as-randomized and the problem of non-adherence: an example from the Veterans Affairs Randomized Trial of Coronary Artery Bypass Surgery. *Stat Med*. 1993;12(11):1185–1195.

For the CME Posttest for this article, see pages 511–512.