U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Science
Office of Biostatistics

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

NDA/Suppl. Number:	21-436/S-0029, 21-713/S-0021, 21-729/S-0014, 21-866/S-0016
Drug Name(s):	Aripiprazole (Abilify®, OPC-14597/BMS-337039) Tablets
Indication(s):	Treatment of Schizophrenia and Bipolar Mania Patients
Applicant:	Otsuka Pharmaceutical Company, Ltd.
Submission Date:	April 16, 2010
Review Priority:	Standard
Biometrics Division:	**Division of Biometrics I (HFD-710)**
Statistical Reviewer:	Yang Yang, Ph.D.
Concurring Reviewers:	Peiling Yang, Ph.D., Team Leader
	Kooros Mahjoob, Ph.D., Deputy Division Director
Medical Division:	**Division of Psychiatry Products (HFD-130)**
Clinical Team:	Maju Mathews, M.D., Medical Reviewer
	Jing Zhang, M.D., Medical Term Leader
Project Manager:	Sonny Saini, Pharm.D.

Keywords: Informative missing data, log-rank test, competing risk, ANCOVA

Table of Contents

Reference ID: 2869701

LIST OF TABLES

Reference ID: 2869701

LIST OF FIGURES

Reference ID: 2869701

1. EXECUTIVE SUMMARY

1.1 Conclusions and Recommendations

Results from Study CN138189 demonstrated that aripiprazole was superior to placebo as adjunctive therapy with lithium or valproate in delaying the time to relapse of any mood episode, and in relieving symptoms of bipolar mania as assessed by the mean change from baseline to Week 52 in CGI-BP Severity of Illness (mania) score.

1.2 Brief Overview of Clinical Studies

The sponsor submitted a Phase IIIb, 52-week, randomized, double-blind, multi-center, parallel-group, placebo-controlled, outpatient study CN138189 to evaluate the efficacy of aripiprazole in combination with lithium or valproate, compared with placebo in combination with lithium or valproate, as long-term maintenance therapy for patients with Bipolar I Disorder, manic or mixed, with or without psychotic features. Study CN138189 consisted of 3 phases: 2- to 8-week screening, washout, and confirmation of partial non-response (Phase 1), 13- to 24-week stability and maintenance of stability (Phase 2), and 52-week assessment of relapse (Phase 3).

A total of 1270 patients were enrolled and 686 patients who completed Phase 1 entered Phase 2. Of the 346 patients who completed Phase 2, 337 patients (placebo: 169; aripiprazole: 168) were randomized and entered Phase 3. Finally, 192 (57% of randomized patients) patients completed the study.

The primary efficacy variable was time to relapse to any mood episode and it was evaluated by a stratified log-rank test on the Randomized Sample. Patients in the Randomized Sample had a mean age of 39 years at baseline; 54.9% were female and 68.2% were white.

1.3 Statistical Issues and Findings

This reviewer confirmed the sponsor's primary finding that the time to relapse of any mood episode observed in Phase 3 was statistically significantly delayed in the aripiprazole group, as compared with the placebo group ($p = 0.014$; Randomized Sample).

However, there are several complications that need consideration:

Key Secondary Endpoint: The sponsor designated "change from baseline to Week 52 in CGI-BP Severity of Illness (mania) score" as a key secondary endpoint. Although the sponsor's primary results for this "key" secondary efficacy endpoint were statistically significant ($p = 0.013$; Phase 3 Efficacy Sample), they should not be described in the Clinical Studies Section. This is because that it is questionable to use CGI-BP severity of illness

Reference ID: 2869701

(mania) score as a key secondary endpoint in such a maintenance study due to potentially informative dropouts, particularly when the dropout rate is expected to be large. This reviewer explored the response profiles of a total of 134 (41.4%) dropouts out of the 336 evaluable patients (see Section 3.1.5.4.2) and found that, for those dropouts, the change-from-baseline measures of CGI-BP severity of illness (mania) seemed to be correlated with relapse status at discontinuation. Therefore, the missing data in this secondary variable were likely to be informative; meanwhile, such a large dropout rate (41.4%) also renders the validity of LOCF approach questionable.

In addition, the normality assumption of the primary ANCOVA appeared to be problematic, and the sponsor's non-parametric method missed nominal significance ($p = 0.054$).

Large Proportion of Randomized Patients with Protocol Deviations: There were a total of 133 (39.5%) randomized patients with protocol deviations relating to the inclusion/exclusion criteria, study conduct, patient management, or patient assessments. The primary stratified log-rank test on the Phase 3 Per-protocol Sample failed to reach nominal significance ($p = 0.060$), but this may be due to insufficient sample size.

Relapse of Mood Episode of a Specific Type: The primary efficacy endpoint (time to relapse of any mood episode) was a composite endpoint, defined as time to a manic, mixed, or depressive mood episode observed in Phase 3. A smaller proportion of relapses were observed in the aripiprazole group (25/168; 14.9%) than in the placebo group (43/169; 25.4%) in Phase 3, favoring aripiprazole. However, it was inappropriate to use Kaplan-Maier (KM) method to estimate the time to relapse by type of mood episode (manic or depressive) for this study. Since the 68 randomized patients who discontinued the study after relapse of a mood episode were not followed up till the end of Phase 3, their relapses of the first mood episode of other types that occurred within 52 weeks after randomization were not counted in the KM estimation. Therefore, it was uncertain whether aripiprazole was effective in delaying the time to relapse of manic (or depressive) mood episode in this study. Please refer to Section 3.1.5.2.3 for more details.

Treatment Effects among Subgroups: In this study, the treatment difference (aripiprazole vs. placebo) appeared to be greater in males than in females, in lithium subgroup than in valproate subgroup, and in patients with manic index mood episode than in patients with mixed index mood episode, as assessed by both the primary and the key secondary endpoints. The treatment difference in the primary efficacy endpoint seemed to greater in non-US patients than US patients; while on the contrary, the treatment difference in the key secondary endpoint seemed to be greater in US patients than in non-US patients.

2. INTRODUCTION

2.1 Overview

Bipolar I Disorder is a lifelong episodic illness characterized by manic or depressive episodes followed by symptom-free periods. Psychotic symptoms (delusions, hallucinations, thought disorders) often accompany the manic phase of bipolar disorder. The lifetime prevalence of bipolar disorder is estimated to be 0.4% to 1.6%.

Lithium carbonate was approved in the early 1970's for the treatment of mania. Approximately 20% to 40% of patients with acute mania fail to respond to lithium and adverse effects that may lead to patient noncompliance are quite common. A slow onset of action and narrow therapeutic window are undesirable characteristics of lithium use. The anticonvulsant, *valproate*, is also approved in some countries for the acute treatment of mania. Liver toxicity is a rare but recognized adverse effect associated with valproate, and monitoring of liver enzymes is recommended. When monotherapy fails, the guidelines recommend combination therapies. *Aripiprazole* is currently approved as monotherapy for the short-term and maintenance treatment of patients with Bipolar I Disorder (manic or mixed). In addition, aripiprazole is approved for short-term use (6 weeks) as adjunctive treatment with lithium or valproate in this population.

Reference is made to Otsuka Pharmaceutical Company Ltd.'s (OPC) NDA 21-436 for aripiprazole tablets, which was originally submitted to the FDA in October 2001. Further reference is made to IND 42,776/SN223 (dated November 16, 1999) in which OPC informed the FDA of the collaborative agreement between OPC and Bristol-Myers Squibb Company (BMS) such that BMS would be delegated to act on behalf of OPC in correspondence with Division of Psychiatry Products. Additional reference is made to IND 73,863 for aripiprazole for the treatment of Bipolar I Disorder.

2.2 Data Sources

The submitted data and study report are located in the following directory (EDR location): \\Cdsesub1\evsprod\NDA021436\0004.

Clinical study report (CSR): \\Cdsesub1\evsprod\NDA021436\0004\m5\53-clin-stud-rep\535-rep-effic-safety-stud\adjunctive-bipolar-maintenance\5351-stud-rep-contr\cn138189\cn138189-csr.pdf. The statistical analysis plan (SAP) is in Appendix 1.11 (Pages 3901-3976) of the CSR. Erratum to final CSR: \\Cdsesub1\evsprod\NDA021436\0004\m5\53-clin-stud-rep\535-rep-effic-safety-stud\adjunctive-bipolar-maintenance\5351-stud-rep-contr\cn138189\cn138189-csr-erratum.pdf.

The derived efficacy data and the sponsor's SAS codes are contained in \\Cdsesub1\evsprod\NDA021436\0004\m5\datasets\cn138189\analysis.

3. STATISTICAL EVALUATION

There is one Phase III study (CN138189) only in this submission. The study was initiated on 29 September 2005 and ended on 02 June 2009.

3.1 Evaluation of Efficacy

3.1.1 Objectives

Primary Objective: to evaluate the efficacy of aripiprazole in combination with lithium or valproate as compared with placebo in combination with lithium or valproate, as long-term maintenance therapy for patients with Bipolar I Disorder, manic or mixed, with or without psychotic features.

Secondary Objective: to evaluate the safety and tolerability of aripiprazole in combination with lithium or valproate as long-term maintenance therapy in this patient population.

3.1.2 Study Design

This was a Phase IIIb, randomized, double-blind, multi-center, parallel-group, placebo-controlled, outpatient study involving a total of 1270 patients with bipolar I disorder. The study consisted of 3 phases:

- Phase 1: 2- to 8-week screening, washout, and confirmation of partial non-response
  Patients with acute manic or mixed symptomatology were assigned to a mood stabilizer of the investigator's choice (either lithium or valproate). Patients who were currently receiving lithium or valproate were also eligible for the study if their symptoms met entry criteria.

- Phase 2: 13- to 24-week stability and maintenance of stability
  Two weeks after confirmation of a therapeutic level (lithium: 0.6 – 1.0 mmol/L or valproate 50 – 150 µg/ml), patients with a partial non-response (YMRS ≥ 16) to treatment with mood stabilizer monotherapy were eligible to enter Phase 2, in which aripiprazole (10 to 30 mg/day; starting dose: 15 mg/day) was assigned in a single-blind fashion. Patients responding to the combination of aripiprazole and mood stabilizer, and maintaining response for 12 weeks (with 1 excursion [defined as a YMRS and/or MADRS score at a given visit of > 12] allowed, except at the last visit), were eligible for Phase 3.

- Phase 3: 52-week assessment of relapse (13 post-baseline visits)
  A fixed-block randomization schedule stratified by study center was used to randomize 337 eligible patients from Phase 2 to receive aripiprazole (10 to 30 mg/day

Reference ID: 2869701

[starting dose: that prescribed as of the end of Phase 2]) plus mood stabilizer or placebo plus stabilizer in a 1:1 ratio. Patients were then followed for a maximum of 52 weeks, or until a relapse occurred.

Criteria for **relapse** included one or more of the following:

- Hospitalization for a manic, mixed or depressive episode
- Serious adverse event of worsening disease under study accompanied by a YMRS > 16 and/or a MADRS > 16
- Discontinuation due to lack of efficacy as determined by investigator accompanied by a YMRS > 16 and/or a MADRS > 16.

Hospitalization for psychosocial support of or any other medical-related reasons were not considered a relapse.

3.1.3 Sample Size Calculation and Power Consideration

The primary efficacy measure was time from randomization to relapse to any mood episode during Phase 3. The sponsor assumed a relapse rate of 16% for aripiprazole plus mood stabilizer and a relapse rate of 31% for mood stabilizer alone (i.e., a hazard ratio for relapse of 0.47). They expected that a total of 62 relapse (events) would provide 80% power to show such a difference in the primary efficacy measure between the two treatment groups using the log-rank test (2-sided, $\alpha = 0.05$). Assuming a 40% dropout rate for reasons other than relapse, the sponsor planned a total of 336 randomized patients (168 per arm). They expected that approximately 1100 to 1300 patients would need to be enrolled.

3.1.4 Efficacy Variables and Sponsor's Analyses

3.1.4.1 Primary Efficacy Variable and Sponsor's Analyses

Primary efficacy variable: time from randomization to relapse of any mood episode

- **Primary Analysis**: log-rank test on the Randomized Sample, stratified by type of mood stabilizer (lithium or valproate) and type of index mood episode the patient was displaying upon study entry (manic or mixed). A Cox regression analysis stratified by type of mood stabilizer and type of mood episode was performed to yield hazard ratio estimates.

- **Supportive Analyses**
  - Interaction: treatment-by-mood stabilizer (or treatment-by-index mood episode) interaction was assessed by a Cox regression analysis with the interaction term in the model.

Reference ID: 2869701

- Proportional Hazards assumption: was assessed by the plots of log(-log(estimated survival function) versus log(time) for the two treatment groups.
- Impact of Censoring: was assessed by the Kaplan-Meier survival curves where censoring was considered as an event and relapse as censoring.

- **Sensitivity Analyses:**
  - The impact of weighting of events was assessed by a Wilcoxon test stratified by type of mood stabilizer and type of index mood episode.
  - The effect of the adjustments was assessed by an unstratified log-rank test.
  - The effect of including patients who did not receive any double-blind medication was assessed by repeating the same survival analysis using the Phase 3 Safety Sample instead of the Randomized Sample, and using the "as treated" treatment group instead of the "as randomized" treatment group.

- **Subgroup analyses:** the primary efficacy endpoint was analyzed by gender, age (≤ 50 years or > 50 years), race (white or non-white), region (US or Non-US), type of mood stabilizer, and type of index mood episode.

- **Censoring:** patients who had not relapsed, including those patients who discontinued early for reasons other than relapse, were censored on the date of last efficacy evaluation or their last dose of study medication, whichever was later. Patients whose relapses occurred more than 7 days after the last dosing date of double-blind medication would be censored. Any randomized patients who did not receive any double-blind medication and did not experience a relapse before discontinuation were censored on their randomization date.

- **Two time-to-event efficacy variables:** time from randomization to relapse of *manic* episode and time from randomization to relapse of *depressive* episode. The analyses were similar to those of the primary efficacy measure.

3.1.4.2 Key Secondary Efficacy Variable and Sponsor's Analyses

KEY secondary efficacy variable: change from baseline to Week 52 in the CGI-BP Severity of Illness (mania) score. Its assessment schedule in Phase 3 is shown in Table 1.

- **Primary Analysis**: LOCF ANCOVA on the Phase 3 Efficacy Sample
  - **Model:** factors: treatment, type of mood stabilizer, type of index mood episode; covariate: baseline score.

Reviewer's Note: Patients who had a relapse of a mood episode would be discontinued from this maintenance study. Hence, the CGI-BP Severity of Illness Scores (mania) after discontinuation would be missing for those patients with relapse. Since this outcome for a dropout patient is likely to be correlated with the relapse status at discontinuation, the missing data in this secondary variable is likely to be informative and may not be

Reference ID: 2869701

accounted for by a typical statistical analysis. This concern would be aggravated if the dropout rate is large. As a result, it is questionable to consider this as a key secondary endpoint (see also: FDA advice letter dated 09 July 2009 for IND 73,863/SN036).

Table 1: Efficacy Assessment Schedule in Phase 3

	Phase 3: Assessment of Relapse (up to 52 weeks)												
	STUDY VISITS												
	Week 4	Week 8	Week 12	Week 16	Week 20	Week 24	Week 28	Week 32	Week 36	Week 40	Week 44	Week 48	Week 52/ End of Study
EFFICACY													
Y-MRS	X	X	X	X	X	X	X	X	X	X	X	X	X
CGI-BP	X	X	X	X	X	X	X	X	X	X	X	X	X
MADRS	X	X	X	X	X	X	X	X	X	X	X	X	X

(Source: Sponsor's Final Clinical Study Report; Study CN138189 Module 5; Tables 7.1.3, page 2097)

- **Supportive Analyses:**
  - Interaction: treatment-by-mood stabilizer (or treatment-by-index mood episode) interaction was assessed by an ANCOVA with the interaction term in the model.
  - Assumptions for ANCOVA model: were assessed by distribution of residuals, test of homoscedasticity, and plot or residuals versus predicted values.
  - Equality of Baseline Slopes: was assessed by an ANCOVA with the baseline-by-treatment term in the model.

- **Sensitivity Analyses:**
  - MMRM analysis with unstructured covariance matrix on observed cases (OC) (model: fixed effects: treatment, mood stabilizer, index mood episode, time, baseline-by-time, treatment-by-time; covariate: baseline score).
  - Nonparametric Wilcoxon rank-sum analysis stratified by mood stabilizer and index mood episode (i.e., the van Elteren test) using the LOCF data

- **Multiplicity Adjustment:** a hierarchical testing procedure (primary variable → key secondary variable)

3.1.5 Efficacy Results

3.1.5.1 Study Populations

- Randomized Sample: comprised all patients who were randomized in Phase 3
- Phase 3 Per-Protocol Sample: comprised all patients who were in the Randomized Sample and did not have any protocol deviations
- Phase 3 Safety Sample: comprised all patients who were in the Randomized Sample and took at least one dose of double-blind medication in Phase 3
- Phase 3 Efficacy Sample: comprised all patients who were in the Phase 3 Safety Sample and had at least one efficacy evaluation after taking the Phase 3 medication

Reference ID: 2869701

Table 2 summarizes the patient disposition in Phase 3. The treatment allocation was well-balanced between the two treatment groups (169/168 per group). The dropout rate in the placebo group was 47.3%, numerically larger than that in the aripiprazole group (38.7%). The proportions of dropouts due to lack of efficacy in the placebo group (18.3%) was numerically larger than that in the aripiprazole group (8.3%). On the contrary, the proportion of dropouts due to adverse event(s) in the placebo group was 8.9%, numerically smaller than that in the aripiprazole group (11.3%). Note that those who relapsed during Phase 3 were considered as dropouts, not as completers.

Table 2: Summary of Patient Disposition in Phase 3

	Placebo N = 169 n (%)	**Aripiprazole N = 168 n (%)**	**Total N = 337 n (%)**
Randomized Sample	169 (100.0)	168 (100.0)	337 (100.0)
Safety Sample	166 (98.2)	167 (99.4)	333 (98.8)
Efficacy Sample	166 (98.2)	166 (98.8)	332 (98.5)
Completers[a]	89 (52.7)	103 (61.3)	192 (57.0)
Early termination	80 (47.3)	65 (38.7)	145 (43.0)
Discontinuation Reason[b]**:**			
Lack of efficacy	31 (18.3)	14 (8.3)	45 (13.4)
Adverse event	15 (8.9)	19 (11.3)	34 (10.1)
Subject withdrew consent	14 (8.3)	15 (8.9)	29 (8.6)
Death	0	1 (0.6)	1 (0.3)
Lost to follow-up	7 (4.1)	6 (3.6)	13 (3.9)
Poor/Non-compliance	5 (3.0)	3 (1.8)	8 (2.4)
Pregnancy	2 (1.2)	1 (0.6)	3 (0.9)
Subject no longer meets study criteria	2 (1.2)	1 (0.6)	3 (0.9)
Administrative reason by sponsor	0	2 (1.2)	2 (0.6)
Other	4 (2.4)	3 (1.8)	7 (2.1)

[a] does not include the relapsed patients
[b] The dropout reason for any relapsed patient was either 'Lack of Efficacy' or 'Adverse Event'.
(Source: Sponsor's Final Clinical Study Report; Study CN138189 Module 5; Tables 5.1, page 0090)

Table 3 summarizes the demographic, baseline characteristics, and psychiatric evaluation at baseline for the Randomized Sample. These 337 randomized patients had a mean age of 39.0 years and a mean weight of 81.0 kg; 54.9% were female and 68.2% were white. It appears that the two treatment groups were balanced in demographic, baseline characteristics, and psychiatric evaluation at baseline. The majority of patients had the manic bipolar disorder (placebo: 70.4%; aripiprazole: 66.1%) and all others had the mixed bipolar disorder.

Table 3: Demographic, Baseline Characteristics, and Psychiatric Evaluation (Randomized Sample)

		Placebo N = 169	Aripiprazole N = 168	Total N = 337
Age (years)	Mean (SD)	38.8 (12.29)	39.2 (12.43)	39.0 (12.34)
Gender				
Male	N (%)	71 (42.0)	81 (48.2)	152 (45.1)
Female	N (%)	98 (58.0)	87 (51.8)	185 (54.9)
Race				
White	N (%)	112 (66.3)	118 (48.2)	230 (68.2)
Black/African American	N (%)	19 (11.2)	12 (7.1)	31 (9.2)
Asian	N (%)	33 (19.5)	34 (20.2)	67 (19.9)
Other	N (%)	5 (3.0)	4 (2.4)	9 (2.7)
Weight (kg)	Mean (SD)	81.3 (25.11)	80.6 (18.89)	81.0 (22.20)
BMI (kg/m^2)	Mean (SD)	28.7 (7.72)	28.5 (6.00)	28.6 (6.90)
CGI-BP Severity of Illness Score				
Overall	Mean (SD)	1.6 (0.76)	1.7 (0.83)	1.6 (0.79)
Mania	Mean (SD)	1.5 (0.72)	1.5 (0.72)	1.5 (0.72)
Depression	Mean (SD)	1.3 (0.57)	1.4 (0.70)	1.4 (0.64)
YMRS Total Score	Mean (SD)	4.1 (3.31)	4.1 (3.56)	4.1 (3.43)
MADRS Total Score	Mean (SD)	3.7 (3.45)	4.1 (3.82)	3.9 (3.64)
DSM-IV-TR Classification				
Bipolar Manic	N (%)	119 (70.4)	111 (66.1)	230 (68.2)
Bipolar Mixed	N (%)	50 (29.6)	57 (33.9)	107 (31.8)
Age at Onset				
Manic or Mixed Symptoms*	Mean (SD)	26.9 (11.34)	26.3 (11.64)	26.6 (11.48)
Depressive Symptoms**	Mean (SD)	25.9 (12.00)	25.5 (11.63)	25.7 (11.80)
Number of Mood Episodes				
In the Past 12 Months	Mean (SD)	1.8 (0.98)	1.7 (0.88)	1.7 (0.93)
In the Past 10 Years	Mean (SD)	7.8 (8.70)	8.8 (8.47)	8.3 (8.58)
Rapid Cycle in the Past 12 Months				
Yes	N (%)	10 (5.9)	3 (1.8)	13 (3.9)
No	N (%)	159 (94.1)	165 (98.2)	324 (96.1)

*1 missing record in the aripiprazole group
**16 missing records in the placebo group and 17 missing records in the aripiprazole group

(Source: Sponsor's Final Clinical Study Report; Study CN138189 Module 5; Tables 5.3.1 [pages 0094-0095], 5.3.2 [pages 0098-0099], and 5.3.3.1 [pages 0101-0102])

Reference ID: 2869701

Table 4 summarizes protocol deviations relating to the inclusion or exclusion criteria, study conduct, patient management, or patient assessments that could have potentially affected the interpretability of study results. There were a total of 133 (39.5%) randomized patients with reported protocol deviations. The most two frequent reasons for protocol deviations in Phase 3 were 1) mood stabilizer noncompliance (placebo: 21/169 [12.4%], aripiprazole: 20/168 [11.9%]) and 2) use of prohibited medications on or after the randomization date (placebo: 21/169 [12.4%], aripiprazole: 20/168 [11.9%]). The proportions of these deviations appeared to be balanced between groups. No patient was excluded from the sponsor's analyses because of protocol deviation(s).

Table 4: Summary of Protocol Deviations in Phase 3 by Treatment Group (Randomized Sample)

	Placebo N = 169 n (%)	Aripiprazole N = 168 n (%)	Total N = 337 n (%)
Randomized Patients with Protocol Deviations	70 (41.4)	63 (37.5)	133 (39.5)
Eligibility (inclusion or exclusion criteria)	25 (14.8)	14 (8.3)	39 (11.6)
Errors in treatment assignment	5 (3.0)	4 (2.4)	9 (2.7)
Poor study drug compliance	32 (18.9)	35 (20.8)	67 (19.9)
Use of prohibited concomitant medications	33 (19.5)	34 (20.2)	67 (19.9)
Other relevant deviations	4 (2.4)	5 (3.0)	9 (2.7)

Note: some randomized patients had more than one protocol deviation.
(Source: Sponsor's Final Clinical Study Report; Study CN138189 Module 5; Tables 4.3, pages 0081-0083; Reviewer's Result using SAS 9.2, 2010)

3.1.5.2 Sponsor's Efficacy Results for Primary Efficacy Variable

3.1.5.2.1 Primary Efficacy Analysis

Table 5 shows that the treatment difference between placebo and aripiprazole in the primary efficacy measure was statistically significant, favoring aripiprazole ($p = 0.014$). The proportion of relapses observed during the double-blind treatment period was 14.9% (25/168) in the aripiprazole group and 25.4% (43/169) in the placebo group. The Kaplan-Meier curves of time to relapse are presented in Figure 1.

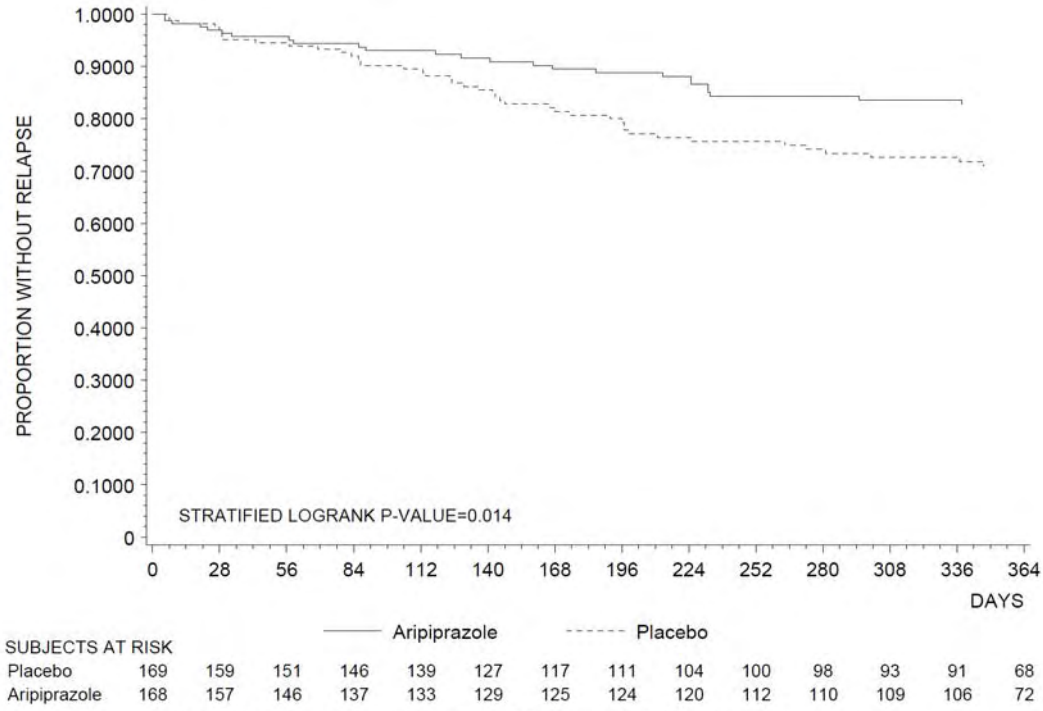
Table 5: Sponsor's Results for Primary Efficacy Variable (Randomized Sample)

Variable	Placebo	Aripiprazole
Primary Efficacy Endpoint		
Time from randomization to relapse to any mood episode (overall)[a]		
No. Events/No. Patients (%)	43/169 (25.4)	25/168 (14.9)
Hazard Ratio		0.54
(95% CI)		(0.33, 0.89)
p-value		0.014

(Source: Sponsor's Final Clinical Study Report; Study CN138189 Module 5; Table 3, page 0008)

Reference ID: 2869701

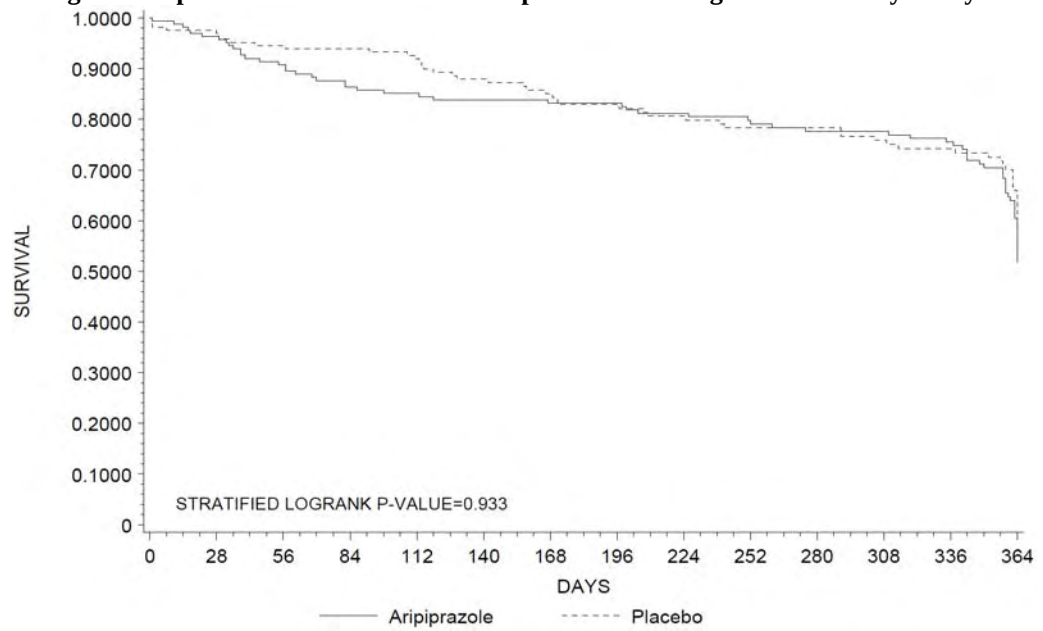
Figure 1: Sponsor's Kaplan-Meier Curves for Relapse to Any Mood Episode (Randomized Sample)



(Source: Sponsor's Final Clinical Study Report; Study CN138189 Module 5; Figure 7.2, page 0123)

The sponsor assessed the impact of censoring on the primary analysis by comparing the censoring distributions between groups. As shown in Figure 2, the pattern of censoring was similar in both treatment groups in Phase 3, which was consistent with the assumption of independent censoring.

Figure 2: Sponsor's Assessment of the Impact of Censoring on the Primary Analysis

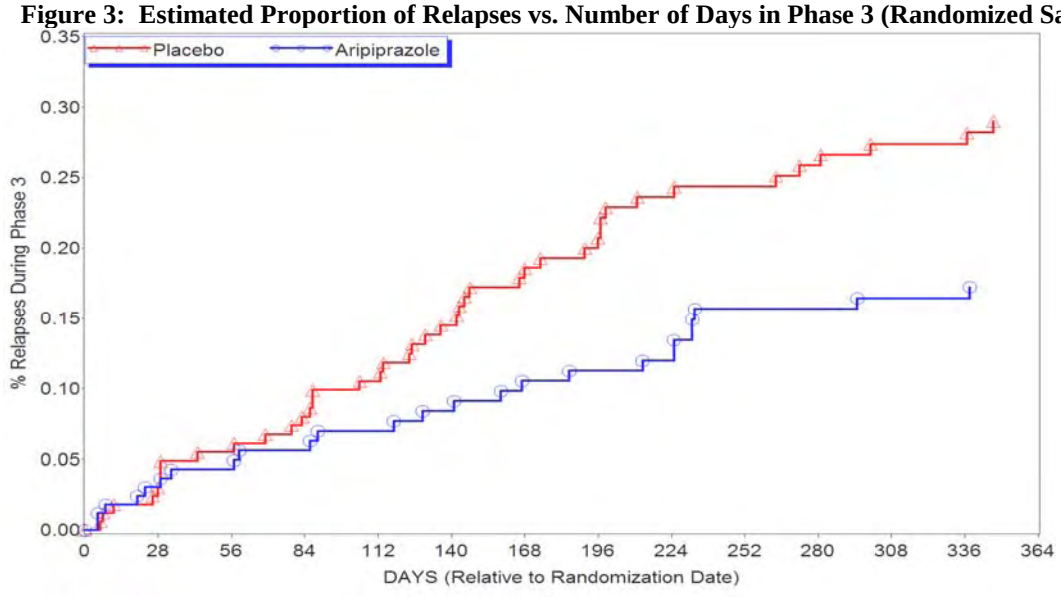


(Source: Sponsor's Final Clinical Study Report; Study CN138189 Module 5; Figure S.5.5, page 2012)

FDA's Comments: This reviewer confirmed the sponsor's primary finding that the time to relapse of any mood episode was statistically significantly delayed in the aripiprazole group as compared with the placebo group during the 52-week double-blind treatment period.

To make the comparison of survival curves (Figure 1) easier to understand, this reviewer compared the estimated proportions of relapses (calculated as 1 – estimated survival probability) between the two treatment groups. Figure 3 shows that the estimated proportion of relapses in the placebo group was consistently higher than that in the aripiprazole group during the course of the 52-week double-blind treatment, favoring aripiprazole. The proportion of relapses observed by endpoint was 29% in the placebo group and 17% in the aripiprazole group.

Figure 3: Estimated Proportion of Relapses vs. Number of Days in Phase 3 (Randomized Sample)



(Source: Reviewer's Result using SAS 9.2, 2010)

It is interesting to found out that, of all the 68 relapsed patients, those in the aripiprazole group appeared to relapse sooner after randomization than those in the placebo group. However, those quantiles were uncertain due to the small number of observed relapses.

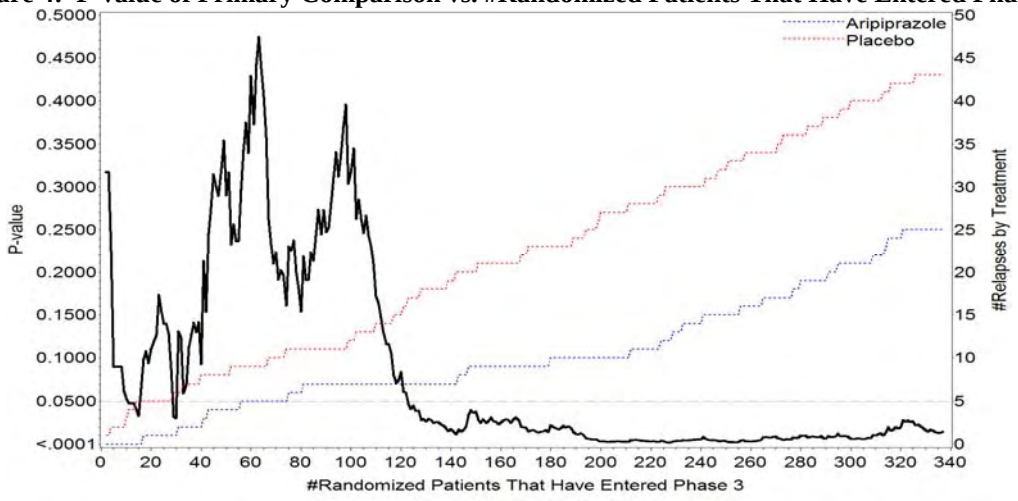
Table 6: Quantiles of Time to Relapse in Days by Treatment in Relapsed Patients

Quantile	Placebo (n = 43)	Aripiprazole (n = 25)
100% Max	347	338
99%	347	338
95%	300	295
90%	273	233
75% Q3	197	225
50% Median	130	129
25% Q1	69	33
10%	28	8
5%	11	5
1%	6	5
0% Min	6	5

(Source: Reviewer's Result using SAS 9.2, 2010)

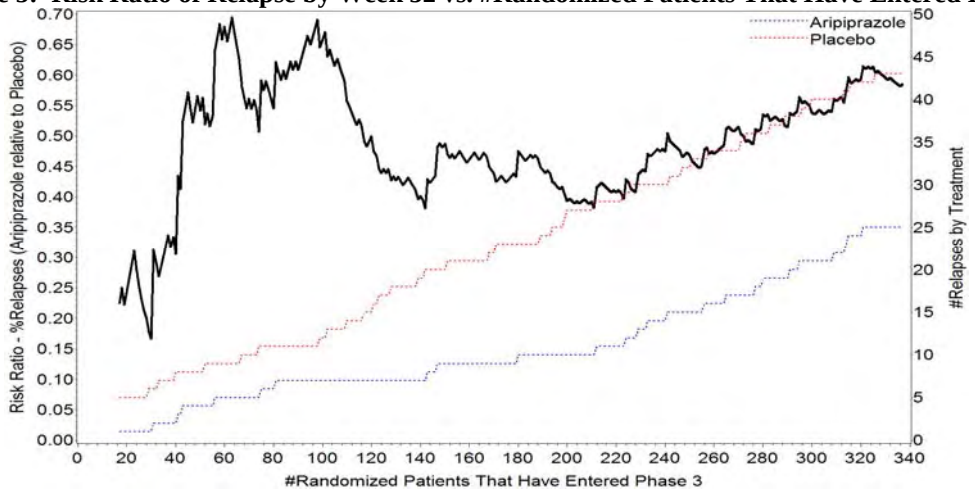
To assess the robustness of the *p*-value from the sponsor's primary log-rank test, this reviewer sorted the 337 patients in the Randomized Sample by the randomization date (from earliest [08 June 2006] to latest [23 July 2008]) and repeated the primary log-rank test on the first *n* randomized patients (*n* ranged from 2 to 337 with increment by 1). Figure 4 shows the *p*-value for the primary endpoint as a function of *n*. It appears that the *p*-value varied dramatically when *n* was small but became stable and consistently below 0.05 when *n* was equal or larger than 123. In addition, the number of observed relapses in the placebo-treated patients (red dot plot) was consistently numerically larger than that in the aripiprazole-treated patients (blue dot plot) for those 336 analysis sets. Figure 5 shows the risk ratio of observed relapse (aripiprazole versus placebo) by Week 52 as a function of *n*. The risk ratio in each analysis set was consistently numerically smaller than 1, indicating a smaller proportion of observed relapses in the aripiprazole-treated patients. All these findings supported the efficacy of aripiprazole as adjunctive treatment with lithium or valproate.

Figure 4: P-value of Primary Comparison vs. #Randomized Patients That Have Entered Phase 3



(Source: Reviewer's Result using SAS 9.2, 2010)

Figure 5: Risk Ratio of Relapse by Week 52 vs. #Randomized Patients That Have Entered Phase 3



(Source: Reviewer's Result using SAS 9.2, 2010)

Reference ID: 2869701

3.1.5.2.2 Sensitivity Analyses for Primary Efficacy Variable

As shown in Tables S.5.6 (page 344), S.5.7 (page 345), and S.5.8 (page 346) in the sponsor's clinical study report for Study CN138189, all the sensitivity analyses (stratified Wilcoxon test, unstratified log-rank test, and stratified log-rank test on Phase 3 Safety Sample) supported the sponsor's primary efficacy findings with *p*-values ≤ 0.020.

FDA's Comments: This reviewer duplicated the sponsor's results of the three sensitivity analyses. Considering that the proportion of randomized patients with protocol deviations (39.5%) was large, this reviewer repeated the primary log-rank test on the Phase 3 Per-protocol Sample (n = 204). Results for the Phase 3 Per-protocol Sample as shown in Table 7 were consistent with the sponsor's primary findings except that this test failed to reach nominal significance (*p* = .060), but it may be due to insufficient sample size.

Table 7: Reviewer's Results for Primary Efficacy Variable (Phase 3 Per-Protocol Sample)

	Placebo	Aripiprazole
Primary Efficacy Endpoint		
Events/Total N (%)	24 / 99 (24.2)	14 / 105 (13.3)
Hazard Ratio*		0.54
(95% CI)		(0.28, 1.04)
P-value of Stratified Log-Rank		0.060

*the proportional hazards assumption appeared to be acceptable.

(Source: Reviewer's Result using SAS 9.2, 2010)

3.1.5.2.3 Time to Relapse by Type of Mood Episode

Each relapse episode observed in Phase 3 was classified as a manic, depressive, or mixed mood episode. As shown in Table 8, the majority (85.3%) of the first observed episodes in Phase 3 were either manic (38.2%) or depressive (47.1%). It appears that more placebo-treated (n = 19) than aripiprazole-treated patients (n = 7) experienced the relapse of a manic episode before discontinuing the double-blind treatment, while the numbers of patients who experienced the relapse of a depressive episode before dropout were similar between the placebo group (n = 18) and the aripiprazole group (n = 14).

Table 8: Type of Observed Relapses in Phase 3 (Randomized Sample)

	Number Relapsed (%)					
	Placebo			Aripiprazole		
Type of Mood Episode	Lithium N = 66	Valproate N = 103	Total N = 169	Lithium N = 70	Valproate N = 98	Total N = 168
Manic	13 (19.7)	6 (5.8)	19 (11.2)	4[b] (5.7)	3[d] (3.1)	7 (4.2)
Depressive	9[a] (13.6)	9 (8.7)	18 (10.7)	5 (7.1)	9[c] (9.2)	14 (8.3)
Mixed	4 (6.1)	2 (1.9)	6 (3.6)	1 (1.4)	3 (3.1)	4 (2.4)
Total	26 (39.4)	17 (16.5)	43 (25.4)	10 (14.3)	15 (15.3)	25 (14.9)

* A total of 5 randomized patient who relapsed more than 7 days after the last dosing date of double-blind study medication were not counted in this table: a. n = 1; b. n = 1; c. n = 1; and d. n = 2.

(Source: Sponsor's Final Clinical Study Report; Study CN138189 Module 5; Table S.5.23, page 0361)

Reviewer's Note: A total of 5 relapses (placebo: 1; aripiprazole: 4) occurred more than 7 days after the last dosing date of double-blind study medication, and they were censored in the sponsor's efficacy analyses. If those 5 relapses were counted as events in the primary log-rank test, the *p*-value would jump from 0.014 to 0.043.

Table 9 and Figure 6 show the sponsor's results for time to relapse by type of mood episode (manic episode: $p = 0.013$; depressive episode: $p = 0.384$).

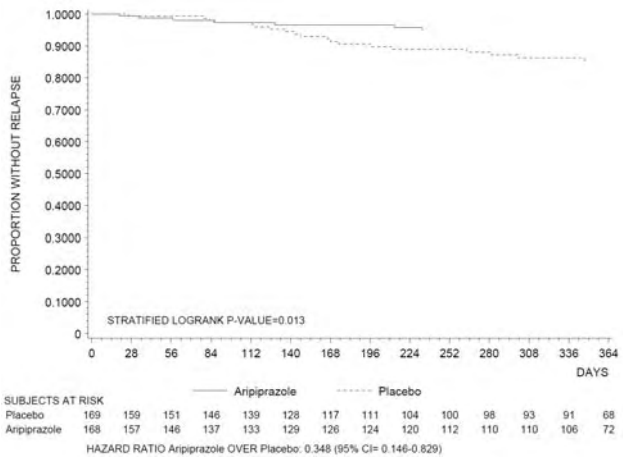
Table 9: Sponsor's Results for Time to Relapse by Type of Mood Episode (Randomized Sample)

Measurement Mood Stabilizer	Placebo Number of Events / Number of Patients (%)	Aripiprazole Number of Events / Number of Patients (%)	Treatment Comparison (a) Aripiprazole/Placebo Hazard Ratio	(95% CI)	p-value
Time from Randomization to Relapse of Manic Episode					
Overall	19/ 169 (11.2)	7/ 168 (4.2)	0.35	(0.15,0.83)	0.013
Lithium	13/ 66 (19.7)	4/ 70 (5.7)	0.27	(0.09,0.84)	0.015
Valproate	6/ 103 (5.8)	3/ 98 (3.1)	0.53	(0.13,2.13)	0.365
Time from Randomization to Relapse of Depressive Episode					
Overall	18/ 169 (10.7)	14/ 168 (8.3)	0.73	(0.36,1.48)	0.384
Lithium	9/ 66 (13.6)	5/ 70 (7.1)	0.46	(0.15,1.37)	0.152
Valproate	9/ 103 (8.7)	9/ 98 (9.2)	1.05	(0.42,2.64)	0.922

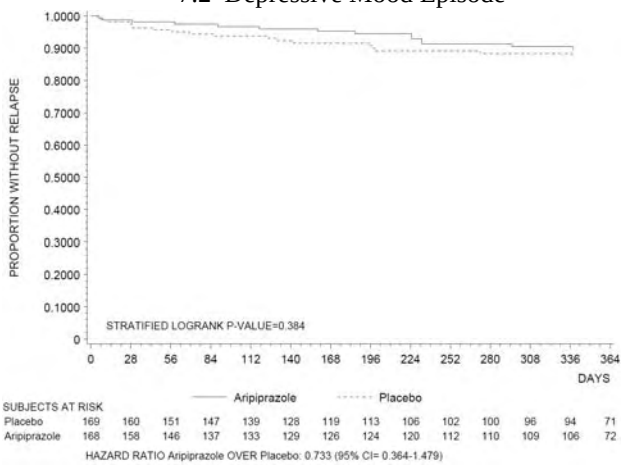
(Source: Sponsor's Final Clinical Study Report; Study CN138189 Module 5; Table 7.1A, page 0114)

Figure 6: Kaplan-Meier Curves of Time to Relapse of the First Manic/Depressive Mood Episode Observed in Phase 3 (Randomized Sample)

7.1 Manic Mood Episode



7.2 Depressive Mood Episode



(Source: Sponsor's Final Clinical Study Report; Study CN138189 Module 5; Figure 7.4.11, page 0146; Figure S.5.12, page 2019)

Reviewer's Note: The sponsor's analysis of time to relapse of manic (or depressive) episode was not about time to the first manic (or depressive) mood episode during the 52-week double-blind treatment period, since patients who had a relapse of manic (or depressive) of mood episode were discontinued from their double-blind medication and were not followed up until their first depressive (or manic) episode by Week 52. Thus, this analysis did not count those potential manic (or depressive) mood episodes, and the results are difficult to interpret.

BEST POSSIBLE COPY

FDA's Comments: It is worth exploring the proportions of observed relapses by type of mood episode. However, the potential competing risk issue rendered the Kaplan-Meier (KM) method inappropriate for estimating the time to relapse by type of mood episode. Since the 68 patients who discontinued their double-blind medication after relapse were not followed up till the endpoint (Week 52), their potential relapses of other type of mood episode were not counted. In addition, with only 26 manic (or 32 depressive) episodes observed in Phase 3, the information conveyed by the KM curves was unreliable. Therefore, the KM curves in Figure 6 should not be used for comparing the time to relapse of manic (or depressive) episode between the treatment groups.

3.1.5.3 Sponsor's Efficacy Results for Key Secondary Efficacy Variable

Table 10 shows the sponsor's LOCF ANCOVA results for the key secondary endpoint on the Phase 3 Efficacy Sample.

Table 10: Sponsor's ANCOVA Results for the Key Secondary Endpoint (Phase 3 Efficacy Sample)

Variable	Placebo	Aripiprazole
Key Secondary Endpoint		
CGI-BP Severity of Illness (Mania)[b]		
N	164	162
Mean baseline (SE)	1.54 (0.059)	1.54 (0.058)
Mean change Week 52 (SE)	0.32 (0.083)	0.04 (0.082)
Difference from Placebo		-0.28
(95% CI)		(-0.50, -0.06)
p-value		0.013

(Source: Sponsor's Final Clinical Study Report; Study CN138189 Module 5; Table 3, page 0008)

FDA's Comments: Although the sponsor's primary results for the "key" secondary efficacy endpoint were statistically significant, they should not be described in the Clinical Studies Section. This is because that it is questionable to use CGI-BP severity of illness (mania) score as a key secondary endpoint in such a maintenance study due to potentially informative dropouts, particularly when the dropout rate is expected to be large. This reviewer explored the response profiles of a total of 134 (41.4%) dropouts out of the 336 evaluable patients (see Section 3.1.5.4.2) and found that, for those dropouts, the change-from-baseline measures of CGI-BP severity of illness (mania) seemed to be correlated with relapse status at discontinuation. Therefore, the missing data in this secondary variable were likely to be informative; meanwhile, such a large dropout rate (41.4%) also renders the validity of LOCF approach questionable.

In addition, the normality assumption of the primary ANCOVA appeared to be problematic, and the sponsor's non-parametric method missed nominal significance (p = 0.054).

3.1.5.4 Statistical Reviewer's Additional Findings and Comments

3.1.5.4.1 *Consistency of Treatment Effect among Centers*

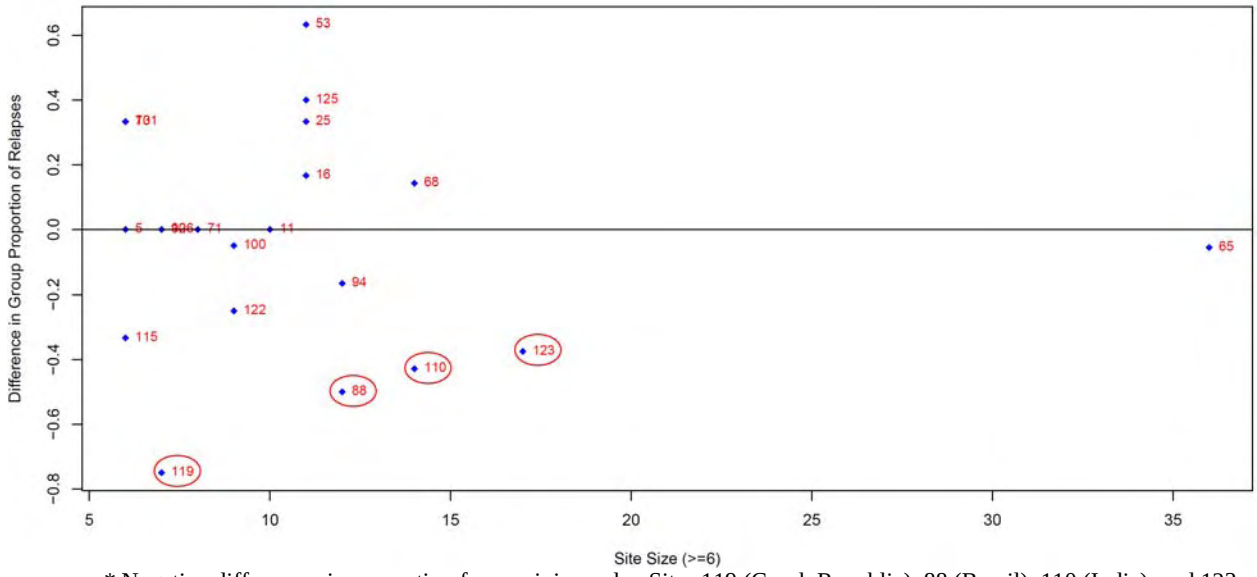
The 337 randomized patients were from 64 centers. The number of randomized patients in each center ranged from 1 to 36. The median site size was 4. Table 11 lists the center IDs, center sizes, and the total number of randomized patients in each country. Four centers (119/Czech Republic, 88/Brazil, 110/India, and 123/United States) most favored aripiprazole (Figure 7), as assessed by the difference in proportion of observed relapses between the aripiprazole and the placebo groups. However, excluding each of those four centers from the primary efficacy analysis did not change the conclusion ($p < 0.040$).

Table 11: Centers by Country

Country	Center ID (n)	#Randomized
Brazil	42(2), 64(1), 65(36), 87(4), 88(12)	55
Croatia	66(2), 67(1), 80(4)	7
Czech Republic	51(3), 52(5), 53(11), 77(1), 119(7), 127(5)	32
France	33(1), 34(1), 79(2), 115(6), 116(2), 117(5)	17
India	89(5), 90(7), 93(1), 94(12), 95(3), 97(2), 100(9), 101(6), 102(4), 103(2), 104(1), 108(1), 110(14)	67
Russia	68(14), 70(2), 71(8), 73(6), 74(2)	32
South Africa	40(4), 55(2), 58(2)	8
United States	1(1), 5(6), 6(3), 10(5), 11(10), 15(1), 16(11), 17(1), 18(4), 19(2), 20(1), 21(1), 23(2), 24(2), 25(11), 28(4), 62(3), 122(9), 123(17), 124(5), 125(11), 126(7), 128(2)	119
Total	*64 centers*	*337*

(Source: Reviewer's Result using SAS 9.2, 2010)

Figure 7: Site Size (≥ 6) vs. Difference in Group Proportion of Observed Relapses (Aripiprazole/Placebo)



* Negative differences in proportion favor aripiprazole. Sites 119 (Czech Republic), 88 (Brazil), 110 (India), and 123 (United States) most favored aripiprazole with a large absolute difference in proportion of relapses between the treatment groups.

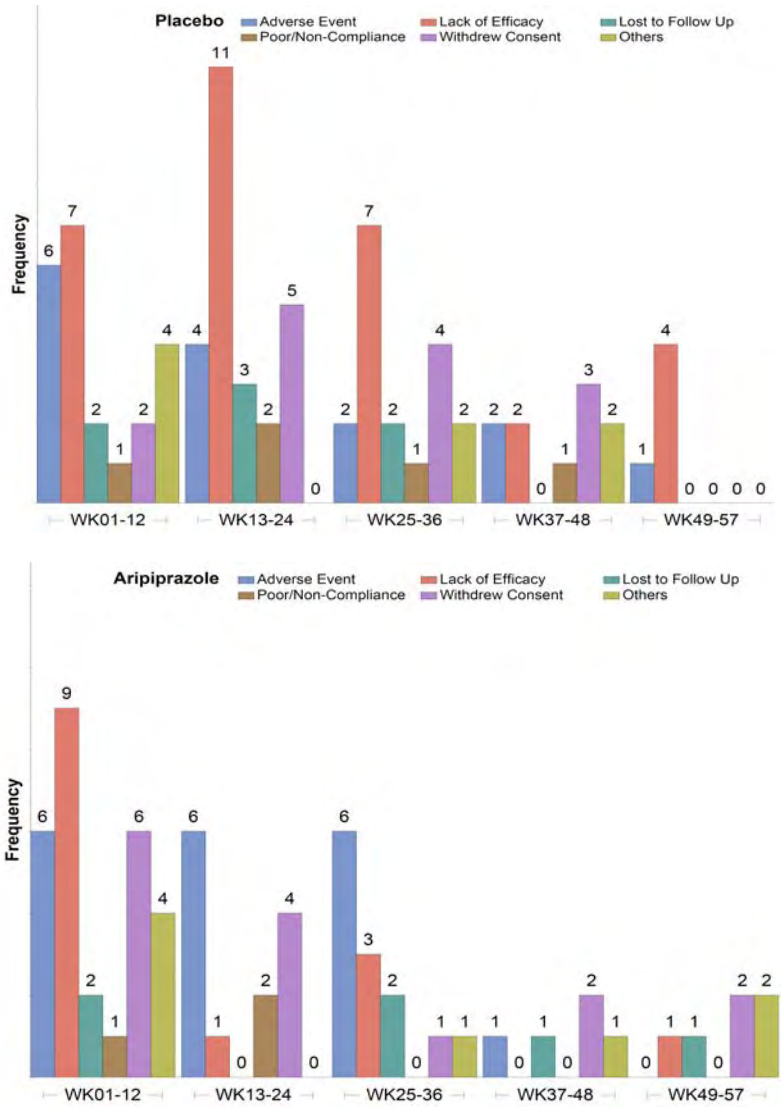
(Source: Reviewer's Result using SAS 9.2, 2010)

Reference ID: 2869701

3.1.5.4.2 Response Profiles of Dropouts

As shown in Table 2, there were a total of 145 (43.0%) dropouts (placebo: 80; aripiprazole: 65) in the Randomized Sample. The dropout rate in the aripiprazole group (38.7%) was numerically smaller than that in the placebo group (47.3%). Figure 8 summarized the number of dropouts by dropout reason within each time interval (Weeks 01-12, Weeks 13-14, Weeks 25-36, Weeks 37-48, or Weeks 49-57). It appears that majority of dropouts occurred within the first 36 weeks after randomization.

Figure 8: Dropouts by Dropout Reason during Each Time Interval (Phase 3 Efficacy Sample)

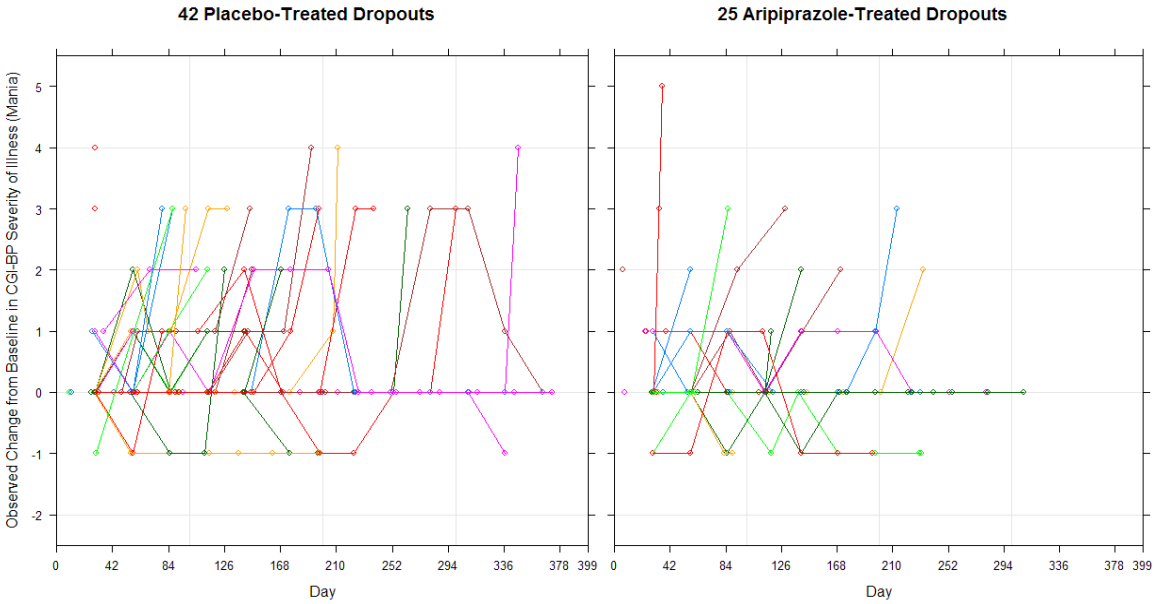


(Source: Reviewer's Result using SAS 9.2, 2010)

Of those 145 dropouts in the Randomized Sample, 11 did not have baseline or any post-baseline measure of CGI-BP Severity of Illness (mania) score. This reviewer then explored the response profiles of the remaining 134 dropouts (placebo: 75; aripiprazole:

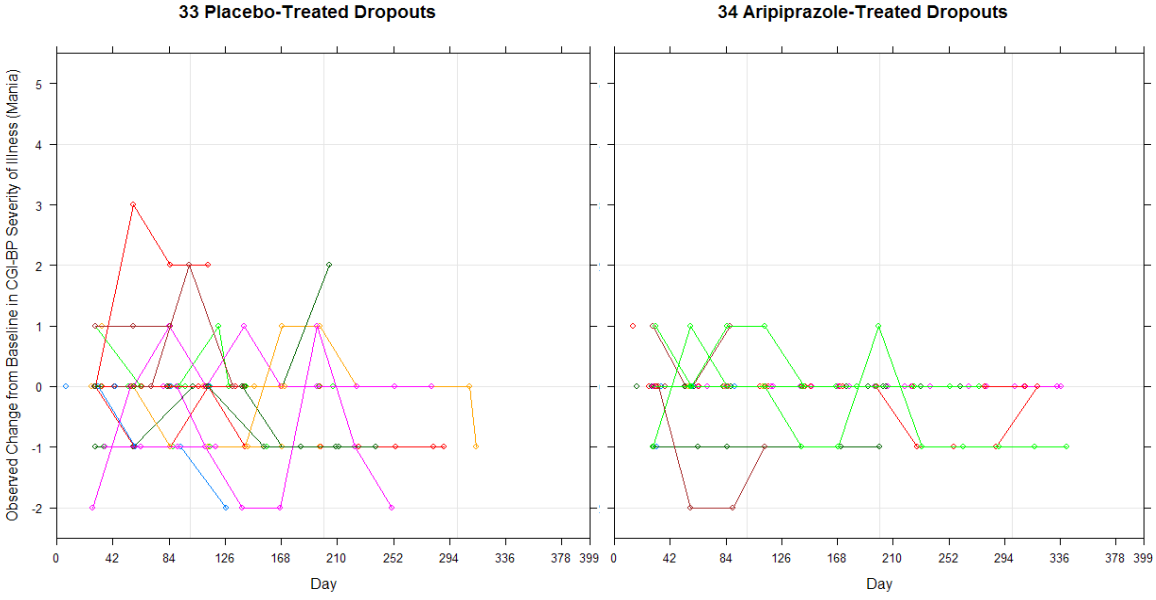
59) in the Phase 3 Efficacy Sample by relapse status at discontinuation (Figure 9 and Figure 10). Patients whose relapses occurred more than 7 days after the last dosing date of double-blind treatment were considered to be relapse-free at discontinuation.

Figure 9: Observed Changes from Baseline in CGI-BP Severity of Illness (Mania) Score for Dropouts with a Relapse at Discontinuation (Phase 3 Efficacy Sample)



Note: the larger the observed change in CGI-BP Severity of Illness (mania), the worse the illness severity.
(Source: Reviewer's Result using SAS 9.2 and S-Plus, 2010)

Figure 10: Observed Changes from Baseline in CGI-BP Severity of Illness (Mania) Score for Dropouts without a Relapse at Discontinuation (Phase 3 Efficacy Sample)



Note: the larger the observed change in CGI-BP Severity of Illness (mania), the worse the illness severity.
(Source: Reviewer's Result using SAS 9.2 and S-Plus, 2010)

Although the dropout rate in the placebo group (75/164 = 45.7%) was numerically larger than that in the aripiprazole group (59/162 = 36.4%), the numbers of dropouts without a relapse at discontinuation were well-balanced between the placebo (n = 33) and aripiprazole (n = 34) groups. In each treatment group, patients who were discontinued after relapse appeared to have worse illness severity on average, as compared with those who were relapse-free at discontinuation. This indicates that, for dropouts, change from baseline in CGI-BP severity of illness (mania) score seemed to be correlated with the relapse status at discontinuation, and the missing mechanisms may be different between the two dropout subgroups.

3.2 Evaluation of Safety

Safety was not reviewed here. Please see Clinical Review.

4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

4.1 Gender, Race, Age, Type of Mood Stabilizer, and Type of Index Mood Episode

4.1.1 Subgroup Analysis of Primary Efficacy Variable

This reviewer replicated the sponsor' subgroup analysis results (Table 12) and confirmed that the proportions of observed relapses in Phase 3 were consistently numerically smaller in the aripiprazole-treated patients than in the placebo-treated patients, across various subgroup categories (gender, age, race, geographic region, type of mood stabilizer, and type of index mood episode).

Table 12: Proportions of Observed Relapses in Phase 3 by Subgroup (Randomized Sample)

Double-Blind Treatment Subgroup (N)	Placebo Events/Total N (%)	Aripiprazole Events/Total N (%)	Log-rank Test P-value*
Gender			
Male (152)	19 / 71 (26.8)	8 / 81 (9.9)	0.005
Female (185)	24 / 98 (24.5)	17 / 87 (19.5)	0.330
Race			
White (230)	31 / 112 (27.7)	20 / 118 (16.9)	0.074
Non-White (107)	12 / 57 (21.1)	5 / 50 (10.0)	0.044
Age at Date of Informed Consent			
≤ 50 Years (267)	36 / 135 (26.7)	20 / 132 (15.2)	0.010
> 50 Years (70)	7 / 34 (20.6)	5 / 36 (13.9)	0.555
Region			
US (119)	14 / 58 (24.1)	10 / 61 (16.4)	0.550
Non-US (218)	29 / 111 (26.1)	15 / 107 (14.0)	0.012
Type of Mood Stabilizer			
Lithium (136)	26 / 66 (39.4)	10 / 70 (14.3)	0.002
Valproate (201)	17 / 103 (16.5)	15 / 98 (15.3)	0.824
Type of Index Mood Episode			
Manic (230)	33 / 119 (27.9)	14 / 111 (12.6)	0.004
Mixed (107)	10 / 50 (20.0)	11 / 57 (19.3)	0.951

*only for exploratory purposes; the primary log-rank test was used for subgroup analyses by gender, race, and age. Log-rank test stratified by type of index mood episode was used for subgroup analysis by type of mood stabilizer. Log-rank test stratified by type of mood stabilizer was used for subgroup analysis by type of index mood episode.

(Source: Sponsor's Final Clinical Study Report; Study CN138189 Module 5; Table S.5.1, page 0335; Tables S.5.9-S.5.13, pages 0347-0351; Reviewer's Results using SAS 9.2, 2010)

Gender: A total of 152 (45.1%) randomized patients were males. The proportions of observed relapses were similar between males (26.8%) and females (24.5%) in the placebo group, while males in the aripiprazole group (9.9%) appeared to have a lower proportion of observed relapses than females (19.5%). Moreover, the treatment difference in the primary efficacy endpoint seemed to be greater in males than in females.

Race: A total of 230 (68.2%) randomized patients were white. A numerically smaller proportion of relapses was observed in non-whites (placebo: 21.1%; aripiprazole: 10.0%) than in whites (placebo: 27.7%; aripiprazole: 16.9%) in each treatment group. Due to insufficient sample size, it is uncertain how well these proportions reflected the true disparities in proportion of relapses among the four race-by-treatment sub-populations.

Reference ID: 2869701

Age: Patient's age at date of informed consent ranged from 18 to 74 in the Randomized Sample. The average and median age was 39. Since there were only 4 randomized patients older than 65, this reviewer did not perform the subgroup analysis by age category.

Type of Mood Stabilizer: There were 136 (40.4%) randomized patients who had lithium as their mood stabilizer. The log-rank test stratified by type of index mood episode reached nominal significance for the lithium subgroup ($p = 0.002$), but not for the valproate subgroup ($p = 0.824$). The sponsor remarked that those results were inconsistent with the results in a previous aripiprazole study (CN138134), where nominal significance was seen in the valproate subgroup, but not in the lithium subgroup.

Type of Index Mood Episode: A total of 230 (68.2%) randomized patients displayed a manic mood episode upon study entry. The log-rank test stratified by type of mood stabilizer reached nominal significance for the manic episode subgroup ($p = 0.004$), but not for the mixed episode subgroup ($p = 0.951$).

4.1.2 Subgroup Analysis of Key Secondary Efficacy Variable

This reviewer performed subgroup analyses of the key secondary endpoint across various subgroup categories. Table 13 shows that the raw mean changes from baseline to Week 52 (LOCF) in CGI-BP Severity of Illness (mania) score were consistently numerically smaller in the aripiprazole-treated patients than in the placebo-treated patients, for each subgroup of interest.

Table 13: Raw Mean Change (SD) from Baseline to Week 52 (LOCF) in CGI-BP Severity of Illness (Mania) Score by Subgroup (Phase 3 Efficacy Sample)

Double-Blind Treatment	Placebo		Aripiprazole		ANCOVA
Subgroup (N)	N	Mean (SD)	N	Mean (SD)	P-value*
Gender					
Male (146)	68	0.40 (1.26)	78	-0.09 (0.97)	0.005
Female (180)	96	0.20 (1.17)	84	0.11 (0.86)	0.438
Race					
White (222)	109	0.31 (1.24)	113	-0.02 (0.90)	0.022
Non-White (104)	55	0.22 (1.13)	49	0.08 (0.98)	0.219
Region					
US (113)	56	0.30 (1.29)	57	-0.09 (0.83)	0.020
Non-US (213)	108	0.27 (1.16)	105	0.07 (0.96)	0.137
Type of Mood Stabilizer					
Lithium (133)	65	0.66 (1.40)	68	0.12 (0.99)	0.005
Valproate (193)	99	0.03 (0.99)	94	-0.06 (0.87)	0.605
Type of Index Mood Episode					
Manic (225)	117	0.32 (1.21)	108	0.03 (0.99)	0.048
Mixed (101)	47	0.17 (1.20)	54	-0.02 (0.76)	0.142

*only for exploratory purposes

(Source: Reviewer's Results using SAS 9.2, 2010)

Reference ID: 2869701

4.2 Other Special/Subgroup Populations

The Randomized Sample included patients from 8 countries. The number of patients from a country ranged from 7 to 119 (median size = 32). To check the consistency in treatment effect among countries, this reviewer compared the treatment difference in proportion of observed relapses and the key secondary endpoint among countries. The estimated hazard ratios were not compared due to the possible violation of proportional hazards assumption. The percentages of observed relapses in Phase 3 were consistently numerically smaller in the aripiprazole-treated patients in all the countries except Russia (Table 14). The raw mean changes in the key secondary endpoint in the aripiprazole-treated patients were numerically smaller in all the countries except Croatia and Russia (Table 15 and Figure 11). However, the numbers of patients and/or events in the Russia and the Croatia subgroups were too small to detect the treatment difference.

Table 14: Percentages of Observed Relapses in Phase 3 by Country (Randomized Sample)

Double-Blind Treatment Subgroup	Placebo Events/Total N (%)	Aripiprazole Events/Total N (%)
Country (N)		
USA (119)	14 / 58 (24.1)	10 / 61 (16.4)
Brazil (55)	7 / 28 (25.0)	4 / 27 (14.8)
Croatia (7)	2 / 3 (66.7)	1 / 4 (25.0)
Czech Republic (32)	6 / 18 (33.3)	4 / 14 (28.6)
France (17)	5 / 8 (62.5)	2 / 9 (22.2)
India (67)	8 / 33 (24.2)	2 / 34 (5.9)
Russia (32)	0 / 17 (0.0)	2 / 15 (13.3)
South Africa (8)	1 / 4 (25.0)	0 / 4 (0.0)

(Source: Reviewer's Result using SAS 9.2, 2010)

Table 15: Raw Mean Change (SD) from Baseline to Week 52 (LOCF) in CGI-BP Severity of Illness (Mania) Score by Country (Phase 3 Efficacy Sample)

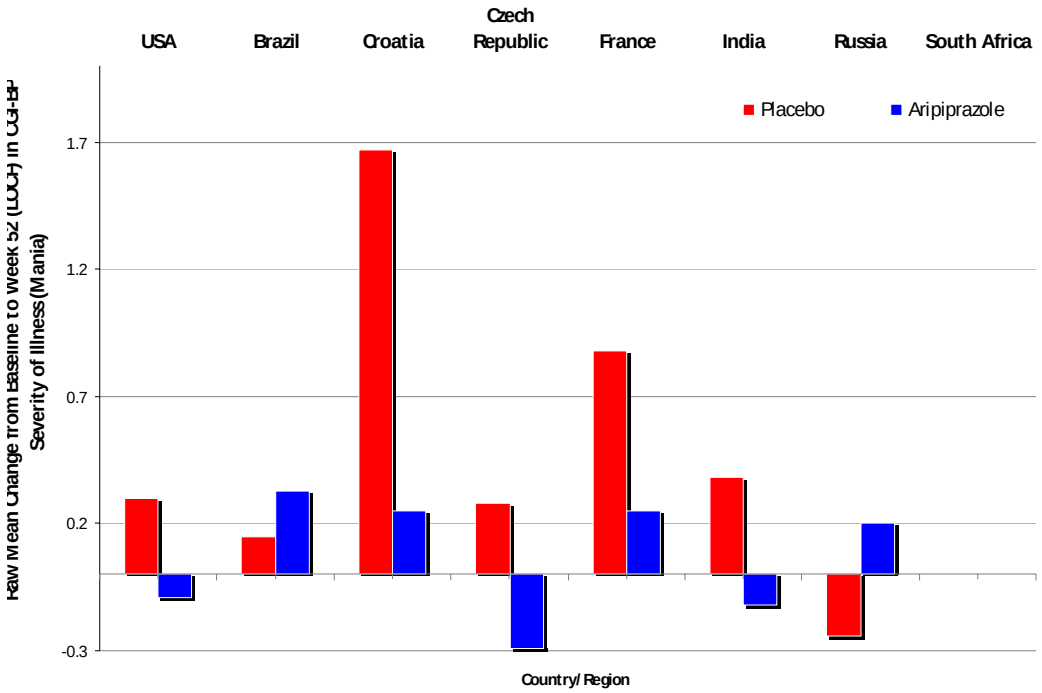
Double-Blind Treatment	Placebo		Aripiprazole	
Subgroup	N	Mean (SD)	N	Mean (SD)
Country (N)				
USA (113) 0.020	56	0.30 (1.29)	57	-0.09 (0.83)
Brazil (53)	26	0.15 (1.01)	27	0.33 (1.18)
Croatia (7)	3	1.67 (1.53)	4	0.25 (0.50)
Czech Republic (32)	18	0.28 (1.07)	14	-0.29 (0.99)
France (16)	8	0.88 (1.73)	8	0.25 (0.89)
India (65) 0.0445	32	0.38 (1.31)	33	-0.12 (0.60)
Russia (32)	17	-0.24 (0.66)	15	0.20 (1.32)
South Africa* (8)	4	0.0 (0.00)	4	0.0 (0.00)

*All the 8 randomized patients had the same baseline and endpoint (LOCF) measures (2 or 1) in CGI-BP Severity of Illness.

(Source: Reviewer's Result using SAS 9.2, 2010)

Reference ID: 2869701

Figure 11: Raw Mean Change from Baseline to Week 52 (LOCF) in CGI-BP Severity of Illness (Mania) Score by Country (Phase 3 Efficacy Sample)



(Source: Reviewer's Result using Excel 2007)

5. SUMMARY AND CONCLUSIONS

5.1 Statistical Issues and Collective Evidence

This reviewer confirmed the sponsor's primary finding that the time to relapse of any mood episode observed in Phase 3 was statistically significantly delayed in the aripiprazole group, as compared with the placebo group ($p = 0.014$; Randomized Sample).

However, there are several complications that need consideration:

Key Secondary Endpoint: The sponsor designated "change from baseline to Week 52 in CGI-BP Severity of Illness (mania) score" as a key secondary endpoint. Although the sponsor's primary results for this "key" secondary efficacy endpoint were statistically significant ($p = 0.013$; Phase 3 Efficacy Sample), they should not be described in the Clinical Studies Section. This is because that it is questionable to use CGI-BP severity of illness (mania) score as a key secondary endpoint in such a maintenance study due to potentially informative dropouts, particularly when the dropout rate is expected to be large. This reviewer explored the response profiles of a total of 134 (41.4%) dropouts out of the 336 evaluable patients (see Section 3.1.5.4.2) and found that, for those dropouts, the change-from-baseline measures of CGI-BP severity of illness (mania) seemed to be correlated with relapse status at discontinuation. Therefore, the missing data in this secondary variable were likely to be informative; meanwhile, such a large dropout rate (41.4%) also renders the validity of LOCF approach questionable.

In addition, the normality assumption of the primary ANCOVA appeared to be problematic, and the sponsor's non-parametric method missed nominal significance ($p = 0.054$).

Large Proportion of Randomized Patients with Protocol Deviations: There were a total of 133 (39.5%) randomized patients with protocol deviations relating to the inclusion/exclusion criteria, study conduct, patient management, or patient assessments. The primary stratified log-rank test on the Phase 3 Per-protocol Sample failed to reach nominal significance ($p = 0.060$), but this may be due to insufficient sample size.

Relapse of Mood Episode of a Specific Type: The primary efficacy endpoint (time to relapse of any mood episode) was a composite endpoint, defined as time to a manic, mixed, or depressive mood episode observed in Phase 3. A smaller proportion of relapses were observed in the aripiprazole group (25/168; 14.9%) than in the placebo group (43/169; 25.4%) in Phase 3, favoring aripiprazole. However, it was inappropriate to use Kaplan-Maier (KM) method to estimate the time to relapse by type of mood episode (manic or depressive) for this study. The information conveyed by the sponsor's KM curves was unreliable due to the small number of mood episodes of a specific type (manic: 26; depressive: 32) observed in Phase 3. In addition, since the 68 randomized patients who discontinued the study after relapse of a mood episode were not followed up till the end of Phase 3, their relapses of the first mood episode of other types that occurred within 52 weeks after randomization were not counted in the KM estimation. Therefore, it was uncertain

Reference ID: 2869701

whether aripiprazole was effective in delaying the time to relapse of manic (or depressive) mood episode in this study. Please refer to Section 3.1.5.2.3 for more details.

Treatment Effects among Subgroups: In this study, treatment difference (aripiprazole vs. placebo) appeared to be greater in males than in females, in lithium subgroup than in valproate subgroup, and in patients with manic index mood episode than in patients with mixed index mood episode, as assessed by both the primary and the key secondary endpoints. The treatment different in the primary efficacy endpoint seemed to greater in non-US patients than US patients; while on the contrary, the treatment difference in the key secondary endpoint seemed to be greater in US patients than in non-US patients.

5.2 Conclusions and Recommendations

Results from Study CN138189 demonstrated that aripiprazole was superior to placebo as adjunctive therapy with lithium or valproate in delaying the time to relapse of any mood episode, and in relieving symptoms of bipolar mania as assessed by the mean change from baseline to Week 52 in CGI-BP Severity of Illness (mania) score.

This is a representation of an electronic record that was signed electronically and this page is the manifestation of the electronic signature.

/s/

YANG YANG
11/29/2010

PEILING YANG
11/29/2010
I concur.

KOOROS MAHJOOB
11/29/2010
We have discussed the review and my views are incorporated in this version. I concur with thisversion.