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

STATISTICAL REVIEW AND EVALUATION

Clinical Studies

NDA#:	21,520 (S-091)
Supplement #:	39
Drug Name:	SYMBYAX (Olanzapine and Fluoxetine HCL)
Indication(s):	(b) (4) Depressive Episodes Associated with Bipolar I Disorder
Applicant:	Eli Lilly and Co.
Date of Submission:	9/28/12
Review Priority:	Standard
Biometrics Division:	Division of Biometrics I
Statistical Reviewer:	Andrejus Parfionovas, Ph.D.
Concurring Reviewers:	Peiling Yang, Ph.D., H.M. James Hung, Ph.D.
Medical Division:	Division of Psychiatry Products
Clinical Team:	Lucas Kempf M.D., Robert Levin M.D.
Project Manager:	Keith Kiedrow, Pharm. D.

Keywords:

NDA review, mixed models, multi-center, pediatric exclusivity, intent-to-treat

Table of Contents

LIST OF TABLES

LIST OF FIGURES

1 EXECUTIVE SUMMARY

Based on the MMRM analysis results, the reviewer confirms sponsor's findings that Symbyax (Olanzapine and fluoxetine combination) administered orally, once a day 6/25, 12/25, 6/50, or 12/50 mg was statistically significantly superior to placebo in the acute treatment of (b) (4) depressive episodes associated with bipolar I disorder in patients 10 to 17 years of age. From the statistical perspective, the study fulfills the Post-Marketing Commitment requested under the Pediatric Research Equity Act (PREA).

2 INTRODUCTION

2.1 Overview

Symbyax (Olanzapine and fluoxetine combination, OFC) has already been approved for use in adult patients to treat depressive episodes associated with bipolar I disorder or treatment-resistant depression. OFC has not been studied in younger patients; however, the individual components have been evaluated:

- Olanzapine is approved for treatment of schizophrenia and manic or mixed episodes of bipolar I disorder in adolescents.
- Fluoxetine is approved for treatment of MDD and obsessive compulsive disorder (OCD) in children.

On April 09, 2007 the Agency sent the Approval Letter, indicating that the sponsor needs to complete a post-marking commitment study under Pediatric Research Equity Act (PREA). In the current submission the sponsor submits the study H6P-MC-HDAX (HDAX) to fulfill this commitment.

This is a Phase 4 multicenter, randomized, placebo-controlled, parallel-group study designed to demonstrate the efficacy and safety of OFC in the pediatric patient population (ages 10 to 17 years) with bipolar I disorder with a current major depressive episode compared to placebo during 8 weeks of double-blind treatment. Patients were randomized in 2:1 ratio to OFC or placebo, respectively. A total of 291 patients were enrolled with 190 patients completing this study at 41 sites in the United States, Puerto Rico, Mexico and Russia. The summary of the study is presented in Table 1.

Table 1. List of all studies included in analysis.

	Phase and Design	**Treatment Period**	**Follow-up Period**	**# of Subjects per Arm (ITT)**	**Study Population**
H6P-MC-HDAX	Phase 4	8 Weeks	none	194 OFC 97 placebo	Bipolar I Disorder in pediatric patients

Source: Computed by the reviewer.

During the course of the trial the sponsor proposed to reverse the prioritization so that the ITT population (randomized patients who received at least one drug dose) would be used for sensitivity analyses, while mITT (modified ITT population excluding the patients from two sites with GCP violation) would be used as the primary dataset for analyses. Further details clarifying this issue are provided in Section 3.2.2.

2.2 Data Sources

The sponsor's electronic submission was stored in the FDA network with the following Electronic Document Room (EDR) location: \\CDSESUB1\EVSPROD\NDA021520\0091

3 STATISTICAL EVALUATION

3.1 Data and Analysis Quality

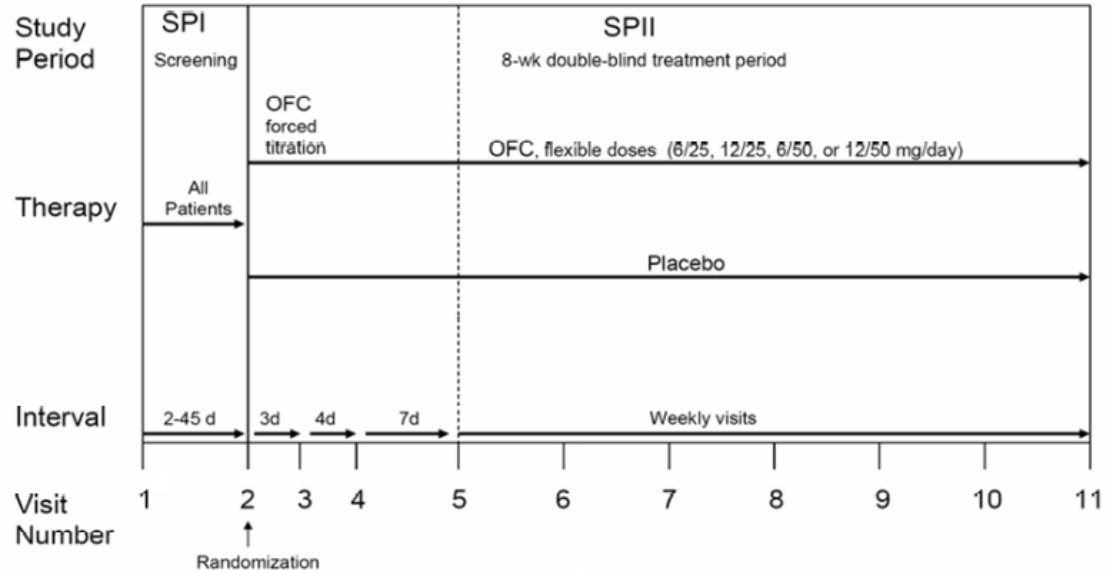
The reviewer found the quality and integrity of the submitted data satisfying and acceptable for the analysis. The reviewer was able to reproduce the primary analysis dataset from the raw data and trace how the primary endpoint was derived.

The sponsor noted that "minor raw data errors" were discovered in the reporting database after it was locked (13/04/12). These errors remain in the reporting database that was used for all analyses in the submitted clinical study report. The sponsor believes these errors are not clinically significant and do not affect any conclusions in the submitted clinical study report. The most notable error was that one patient who was intentionally listed as a screen failure on the eCRF was unintentionally randomized in IVRS. This patient (Patient ID 2308) did not undergo any study procedures for Study Period II and never received study drug.

3.2 Evaluation of Efficacy

This was a Phase IV, multicenter, randomized, double-blind, placebo-controlled, parallel-group study comparing the efficacy and safety of OFC versus placebo in patients aged 10 to 17 years meeting the diagnostic criteria for a current major depressive episode of bipolar I disorder according to the DSM-IV-TR. Patients were randomized 2:1 to OFC or placebo arms. Following a 2- to 45-day screening period (Study Period I), qualified patients were randomized and entered a 2-week forced-titration period followed by a 6-week flexible dosing period (total 8 weeks, Study Period II). Patients were evaluated for safety and efficacy at each visit. The study design is visually represented on Figure 1.

Figure 1. Study design for Protocol H6P-MC-HDAX(a).



Source: Figure HDAX.1 (Protocol H6P-MC-HDAX(a), pg. 22).

3.2.1 Study Design and Endpoints

The primary objective of this study was to assess the superiority of olanzapine and fluoxetine combination (OFC) therapy compared with placebo in the treatment of patients with a current major depressive episode of bipolar I disorder in improving overall symptomatology based on mean change from baseline to Week 8 on the Children's Depression Rating Scale-Revised (CDRS-R) total score, using mixed-model repeated measures (MMRM) analysis.

The primary efficacy measure was the change from baseline to Week 8 in Children's Depression Rating Scale-Revised (CDRS-R) total score. The CDRS-R is modeled after the Hamilton Depression Rating Scale for adults, but also includes questions about school. The scale consists of 17 items scored on a 1 to 5 or 1 to 7 scale. The null hypothesis was that no difference exists between patients receiving OFC and patients receiving only placebo in change from Baseline to Week 8 in CDRS-R total score.

No key secondary endpoints were prespecified.

3.2.2 Statistical Methodologies

The primary efficacy analysis was restricted maximum likelihood (REML)-based, mixed-effects model approach (MMRM) with unstructured covariance. The model includes treatment, visit, treatment-by-visit, country as fixed categorical effects; and baseline CDRS-R as a covariate. Initially, the primary analysis set was on the ITT population (defined as all randomized patients who received at least one drug dose), and analysis performed on the mITT population was considered as sensitivity analysis. The mITT population is the ITT population with the exclusions of patients from the sites prematurely closed due to the GCP violations. During the course of this trial (first patient enrolled on 03/17/2009, last patient completed on 02/02/2012), the sponsor conducted audits for a few sites and found serious CGP violations in two of them (sites 210 and 215, see Table 2 for details).

Table 2. Summary of the sites with GCP violations.

Site	GCP issues revealed	Actions taken	Number of subjects		
			Symbyax (OFC)	Placebo	Total
215	02/02/2010	Enrollment halted on 02/17/2010	7	4	11
210	06/30/2010	Enrollment restricted	14	7	21

Source: Summarized by the reviewer.

As a result, sponsor requested to reverse the prioritization such that the mITT, not ITT, population will be the primary analysis set (refer to meeting minutes dated 08/01/2012 under IND 103074). We commented that in principle the primary analysis set should be performed on the ITT analysis regardless of protocol violations, in addition that FDA had not yet confirmed the violations from these two sites. We noted that it was an unusual circumstance to request revision of the primary analysis set before the sNDA submission and recommended that the sponsor keep ITT as the primary analysis set and what analysis results will be included in the labeling description will eventually be a matter of review. However, the sponsor justified the seriousness

of violations that data from these two sites appear to be problematic. We accepted the sponsor's proposal to use the mITT as the primary set for the efficacy analysis, but noted that we would also review the primary analysis performed with the ITT population.

3.2.3 Patient Disposition, Demographic and Baseline Characteristics

Patients' disposition between the treatment arms during the trial is summarized in Table 3 .

Table 3. Patients' disposition during the trial.

Patients disposition	Symbyax (OFC)	Placebo	Total
Planned	200	100	300
Randomized (in 2:1 ratio)	194	97	291
ITT (received at least 1 dose)	191	96	287
mITT (excluding sites 210 and 215)	170	85	255
Completed the study	116	60	176

Source: summarized by the reviewer from the Clinical Study Report data.

There was also one patient (PID 2308) who was listed as a screen failure on the eCRF, but was unintentionally randomized. This patient did not undergo any study procedures for Study Period II and never received study drug. The data that was provided by the sponsor contains no records of this patient, which makes the total numbers of ITT and mITT patients 286 and 254 respectively. The demographic and baseline characteristics for the mITT patients are summarized in Table 4.

Table 4. Baseline demographic characteristics of the patients in the mITT population.

	Symbyax (OFC) N = 170	Placebo N = 85	Total N = 255
Age *years*			
Mean (SD)	14.59 (2.30)	15.03 (2.13)	14.74 (2.25)
Min – Max	10.04 – 17.90	10.03 – 17.98	10.03 – 17.98
Gender *n (%)*			
Male	84 (49.4)	46 (54.1)	130 (51.0)
Female	86 (50.6)	39 (45.9)	125 (49.0)
Ethnicity *n (%)*			
Hispanic/Latino	38 (22.4)	23 (27.1)	61 (23.9)
Not Hispanic/Latino	132 (77.6)	61 (71.8)	193 (75.5)
Race *n(%)*			
American Indian/Alaska Native	9 (5.3)	5 (5.9)	14 (5.5)
Asian	0 (0.0)	0 (0.0)	0 (0.0)
Black/African American	23 (13.5)	11 (12.9)	34 (13.3)
Native Hawaiian/Pacific Islander	1 (0.6)	0 (0.0)	1 (0.4)
White	119 (70.0)	61 (71.8)	180 (70.6)
Multiple	15 (8.8)	7 (8.2)	22 (8.6)
Weight *kg*			
Mean (SD)	62.74 (19.01)	65.27 (19.33)	63.59 (19.12)
Min – Max	24.00 – 119.00	31.30 – 133.60	24.00 – 133.60

Height *cm*			
Mean (SD)	162.02 (12.76)	164.11 (12.65)	162.72 (12.74)
Min – Max	123 – 192.00	127 – 190.00	123.00 – 192.00
BMI *kg/m^2*			
Mean (SD)	23.53 (5.54)	24.00 (5.81)	23.69 (5.62)
Min – Max	14.34 – 48.45	16.13 – 43.69	14.34 – 48.45

Source: Table HDAX.11.2 (Clinical Study Report, pg.80-81)

The disposition of the mITT patients with respect to the reasons for discontinuation is summarized in Table 5.

Table 5. Discontinuation reasons for patients in the mITT population.

Patients Disposition *n (%)*	**Symbyax (OFC)** N = 170	**Placebo** N = 85	**Total** N = 255
Number of Completers	116 (68.2)	60 (70.6)	176 (69.0)
Number of Dropouts	54 (31.8)	25 (29.4)	79 (31.0)
Dropout Reasons:			
Adverse Event	24 (14.1)	5 (5.9)	29 (11.4)
Death	0 (0.0)	0 (0.0)	0 (0.0)
Sponsor's Decision	1 (0.6)	1 (1.2)	2 (0.8)
Physician Decision	2 (1.2)	0 (0.0)	2 (0.8)
Parent/Caregiver Decision	6 (3.5)	2 (2.4)	8 (3.1)
Subject Decision	8 (4.7)	7 (8.2)	15 (5.9)
Protocol Violation	6 (3.5)	0 (0.0)	6 (2.4)
Lack of Efficacy	5 (2.9)	5 (5.9)	10 (3.9)
Clinical Relapse	0 (0.0)	1 (1.2)	1 (0.4)
Entry Criteria Not Met	0 (0.0)	1 (1.2)	1 (0.4)
Lost to Follow Up	2 (1.2)	3 (3.5)	5 (2.0)

Source: Table HDAX.10.1 (Clinical Study Report, pg. 68)

3.2.4 Sponsor's Results and Conclusions

The sponsor found statistically significantly difference (p-value = 0.003) between OFC and placebo in the acute treatment of major depressive episodes associated with bipolar I disorder in patients 10 to 17 years of age. The results of the primary efficacy analysis (MMRM analysis using contrast between OFC and placebo on mean change from baseline in the CDRS-R total score at the Week 8) are presented in Table 6.

Table 6. CDRS-R total score change from baseline to Week 8 MMRM analysis results (mITT set).

Variable Analyzed: CDRS-R Total Score

Treatment	Baseline						Endpoint (Week 8)					Change to Endpoint				
	N	Mean	SD	Min	Median	Max	Mean	SD	Min	Median	Max	Mean	SD	Min	Median	Max
1) OFC	170	54.58	9.95	28.0	53.00	85.0	25.92	9.51	17.0	23.00	69.0	-29.49	12.12	-61.0	-30.00	4.0
2) Placebo	84	53.71	8.16	40.0	53.00	77.0	29.67	11.36	17.0	25.50	59.0	-24.60	11.89	-57.0	-25.50	-2.0

Raw Data

All Effects (Type III SS) *a

Baseline	F = 39.27	df = 1, 234	p = <.001
Treatment	F = 18.02	df = 1, 198	p = <.001
Visit	F = 48.04	df = 8, 194	p = <.001
Treatment * Visit	F = 1.77	df = 8, 194	p = .085
Country	F = 0.59	df = 3, 229	p = .624

Least Squares Means for Change *b

1) OFC	-28.43	(SE = 1.13)	p = <.001
2) Placebo	-23.40	(SE = 1.49)	p = <.001

Comparison of LS Means between Treatments *c

OFC - Placebo	diff = -5.04	Two-sided 95% CI: (-8.29, -1.79)	t = -3.07	p = .003

Source: Table HDAX.11.14 (Clinical Study Report, pg. 104)

3.2.5 Results Verification by the Reviewer

The reviewer confirm the sponsor's analysis results for the primary efficacy endpoint using both ITT and mITT datasets. The reviewer's results of the MMRM analysis are summarized in Table 7.

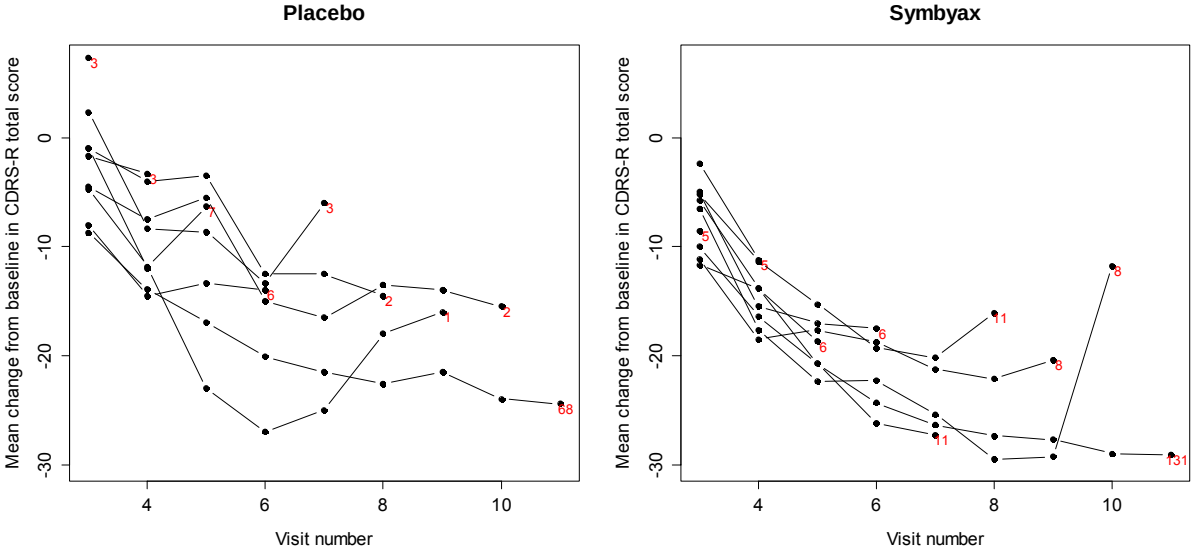
Table 7. LS Mean Differences for the Primary Efficacy Endpoint (ITT and mITTsets).

Sample	Effect	Visit	Difference	(SE)	DF	t-Value	p-value	95% CI
mITT	OFC vs. Placebo	11	-5.0388	(1.6439)	158	-3.07	0.0026	(-8.2856, -1.7919)
ITT	OFC vs. Placebo	11	-4.8018	(1.5954)	187	-3.01	0.0030	(-7.9491, -1.6545)

Source: Computed by the reviewer

The reviewer explored the potential impact of the dropouts on the efficacy results by comparing the average change from the baseline in CDRS-R total score in treatment arm versus placebo for each drop-out date (see Figure 2). The two graphs (one for each treatment arm) show the average change from baseline in the primary efficacy measure (CDRS-R total score) computed for the patients, after they were grouped according to the date of their drop-out. Each curve is labeled with the number of patients in the group. The visual analysis of the data did not appear to indicate an obvious deviation from missing at random (MAR) assumption.

Figure 2. Change from baseline in CDRS-R total score for patients grouped by drop-out date (ITT set).



Source: Computed by the reviewer.

4 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

4.1 Gender, Race, Age, and Geographic Region

This section contains the reviewer's results of the exploratory analysis for Study H6P-MC-HDAX. The data were grouped by gender, age group, weight group, race, ethnicity, and country. The analyses were performed on both ITT and mITT datasets (see Table 8 and Table 9 respectively).

Table 8. Primary efficacy endpoint analysis by subgroup (ITT set).

	N	LS difference (SE)	95% CI	p-value
Full ITT dataset	286	-4.80 (1.60)	(-7.95, -1.65)	0.0030
Gender				
Male	144	-6.49 (2.28)	(-11.02, -1.96)	0.0055
Female	142	-3.27 (2.25)	(-7.73, 1.20)	0.1500
Age category				
7-11 years	42	-4.88 (6.54)	(-18.96, 9.20)	0.4684
12-17 years	244	-4.56 (1.71)	(-7.94, -1.17)	0.0086
Weight category				
≤ 50 kg	74	-8.33 (3.77)	(-16.00, -0.66)	0.0343
> 50 kg	212	-3.77 (1.82)	(-7.38, -0.17)	0.0404
Race/Ethinicity				
Black	37	-1.81 (4.65)	(-11.38, 7.76)	0.7002
White	205	-5.73 (1.91)	(-9.51, -1.94)	0.0033
Multiple	24	-1.49 (5.38)	(-12.87, 9.88)	0.7848
Hispanic/Latino	65	-1.86 (2.53)	(-7.00, 3.29)	0.4689
Not Hispanic/Latino	220	-5.89 (1.93)	(-9.70, -2.08)	0.0027
Country				
Mexico	23	-0.75 (6.26)	(-14.65, 13.14)	0.9066
Russia	44	-10.73 (2.42)	(-15.76, -5.71)	0.0002
USA	209	-4.43 (1.92)	(-8.22, -0.63)	0.0226

Source: computed by the reviewer.

Table 9. Primary efficacy endpoint analysis by subgroup (mITT set)

	N	LS difference (SE)	95% CI	p-value
Full mITT dataset	254	-5.04 (1.64)	(-8.29, -1.79)	0.0026
Gender				
Male	130	-7.52 (2.37)	(-12.22, -2.81)	0.0021
Female	124	-2.60 (2.28)	(-7.14, 1.94)	0.2574
Age category				
7-11 years	39	-3.12(7.75)	(-20.51, 14.27)	0.6964
12-17 years	215	-4.98 (1.77)	(-8.48, -1.49)	0.0056
Weight category				
≤ 50 kg	65	-8.28 (4.54)	(-17.59, 1.02)	0.0789
> 50 kg	189	-3.89 (1.87)	(-7.59, -0.19)	0.0396
Race/Ethinicity				
Black	34	0.73 (4.26)	(-8.07, 9.53)	0.8655
White	180	-5.83 (1.98)	(-9.76, -1.90)	0.0040
Multiple	21	-8.17 (6.28)	(-21.68, 5.35)	0.2153
Hispanic/Latino	61	-2.06 (2.65)	(-7.45, 3.33)	0.4430
Not Hispanic/Latino	192	-6.31 (2.01)	(-10.30, -2.32)	0.0022
Country				
Mexico	23	-0.75 (6.26)	(-14.65, 13.14)	0.9066
Russia	44	-10.73 (2.42)	(-15.76, -5.71)	0.0002
USA	177	-4.87 (2.04)	(-8.90, -0.84)	0.0183

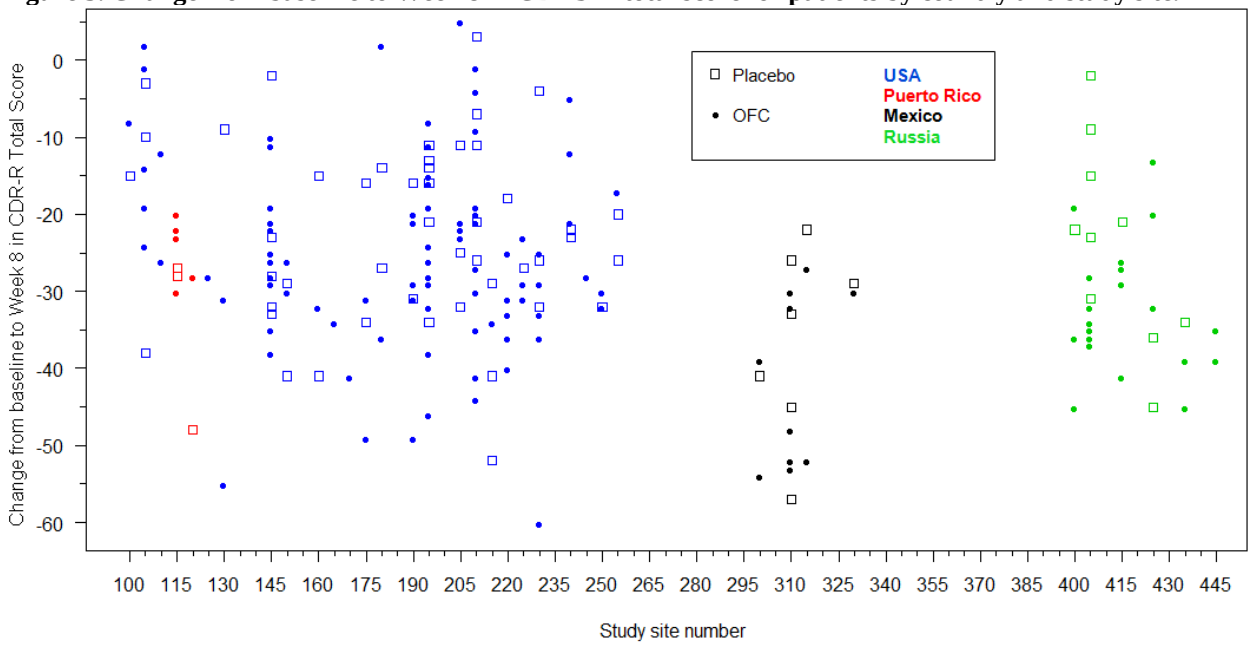
Source: computed by the reviewer.

Based on this reviewer's analysis, there does not appear to be substantial heterogeneity in treatment efficacy among the subgroups. The only subgroup that had opposite sign for numeric estimate for the efficacy effect was the Black subgroup in the mITT population. This was inconsistent with the rest of the subgroups, including the Black subgroup effect in the ITT population. The reason for that could possibly be attributed to the relatively large variance in quite a small subgroup.

4.2 Other Special/Subgroup Populations

The scatterplot of the primary endpoint data (change from baseline to Week 8 in CDRS-R total score) for each patient grouped by country and study site is presented on Figure 3. The visual analysis of the data for each site did not appear to indicate any suspicious patterns that could suggest data manipulation or unusual distribution. The primary endpoint data is also summarized in Table 10 for every study site.

Figure 3. Change from baseline to Week 8 in CDRS-R total score for patients by country and study site.



Source: computed by the reviewer.

5 SUMMARY AND CONCLUSIONS

5.1 Statistical Issues

In the original SAP, the sponsor planned to perform all analyses on the ITT population and only provide modified ITT (mITT) as a sensitivity analysis for the primary efficacy endpoint and three adverse event reports. The mITT population is the ITT population with the exclusion of sites with GCP violations. The sponsor proposed to reverse this prioritization so that the main analyses will be the mITT analyses; the ITT would be used for sensitivity analyses. The issue, however, seems to have no impact on the efficacy analysis results.

5.2 Collective Evidence

There is only one study reviewed in this submission. The primary analysis results of this study are statistically significant whether based on ITT or mITT populations. The exploratory subgroup analyses did not reveal noticeable heterogeneity with respect to the primary efficacy measure (change from baseline to Week 8 in CDRS-R total score).

5.3 Conclusions and Recommendations

The reviewer confirmed the sponsor's analysis results that Symbyax (OFC) was statistically significantly superior to placebo (p-value = 0.003) in the acute treatment of (b) (4) depressive episodes associated with bipolar I disorder in patients 10 to 17 years of age. From the statistical perspective, there is no evidence against fulfilling the postmarketing commitment under the PREA.

APPENDIX

Table 10. Change from the baseline to Week 8 in CDRS-R total score for each treatment arm by study site.

Site	Country	N	Symbyax (OFC) Mean (SD)	Placebo Mean (SD)	OFC–Placebo Difference
145	USA	28	-25.85 (8.95)	-23.60 (12.70)	-2.25
210	USA	21	-24.42 (13.79)	-12.40 (11.48)	-12.02
195	USA	20	-25.46 (10.74)	-18.17 (8.47)	-7.29
205	USA	16	-18.67 (11.13)	-22.67 (10.69)	4.00
405	Russia	14	-34.89 (2.89)	-16.00 (11.40)	-18.89
105	USA	13	-12.20 (11.30)	-17.00 (18.52)	4.80
310	Mexico	12	-42.86 (10.65)	-40.25 (13.65)	-2.61
190	USA	11	-31.00 (11.66)	-23.50 (10.60)	-7.50
215	USA	11	-35.00 (NA)	-40.67 (11.50)	5.67
220	USA	9	-34.00 (5.61)	-18.00 (0)	-16.00
230	USA	9	-37.60 (13.72)	-20.67 (14.74)	-16.93
160	USA	8	-33.00 (NA)	-28.00 (18.38)	-5.00
115	Puerto Rico	7	-24.74 (4.35)	-27.50 (0.71)	2.76
415	Russia	7	-31.75 (6.95)	-21.00 (NA)	-10.75
175	USA	6	-41.00 (12.73)	-25.00 (12.73)	-16.00
180	USA	6	-18.00 (26.87)	-20.50 (9.19)	2.50
225	USA	6	-28.67 (4.16)	-27.00 (NA)	-1.67
425	Russia	6	-22.67 (9.61)	-40.50 (6.36)	17.83
100	USA	5	-9.00 (NA)	-15.00 (NA)	6.00
150	USA	5	-29.00 (2.82)	-35.00 (8.49)	6.00
165	USA	5	-35.00 (NA)	NA (NA)	NA
240	USA	5	-13.67 (8.02)	-22.50 (0.71)	8.83
255	USA	5	-18.00 (NA)	-23.00 (4.24)	5.00
400	Russia	5	-34.33 (13.20)	-22.00 (NA)	-12.33
445	Russia	5	-38.00 (2.83)	NA (NA)	NA
130	USA	4	-44.00 (16.97)	-9.00 (NA)	-35.00
315	Mexico	4	-36.33 (14.43)	-22.00 (NA)	-93.33
330	Mexico	4	-31.00 (NA)	-29.00 (NA)	-2.00
110	USA	3	-20.00 (9.90)	NA (NA)	NA
120	Puerto Rico	3	-29.00 (NA)	-48.00 (NA)	19.00
245	USA	3	-29.00 (NA)	NA (NA)	NA
250	USA	3	-32.00 (1.41)	-32.00 (NA)	0.00
300	Mexico	3	-47.50 (10.61)	-41.00 (NA)	-6.50
435	Russia	3	-43.00 (4.24)	-34.00 (NA)	-9.00
125	USA	2	-29.00 (NA)	NA (NA)	NA
200	USA	2	NA (NA)	NA (NA)	NA
420	Russia	2	NA (NA)	NA (NA)	NA
440	Russia	2	NA (NA)	NA (NA)	NA
155	USA	1	NA (NA)	NA (NA)	NA
170	USA	1	-42.00 (NA)	NA (NA)	NA

Source: computed by the reviewer. The missing data indicate that no patients were available by Week 8.

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

/s/

ANDREJUS PARFIONOVAS
05/09/2013

PEILING YANG
05/09/2013
I concur with the review.

HSIEN MING J HUNG
05/09/2013