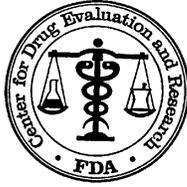


**CENTER FOR DRUG EVALUATION AND
RESEARCH**

APPLICATION NUMBER:

207533Orig1s000

STATISTICAL REVIEW(S)



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/BLA #: NDA 207533/ O-1

Drug Name: ARISTADA (Aripiprazole Lauroxil, extended release suspension for IM injection 441 mg and 882 mg)

Indication(s): Treatment of Schizophrenia

Applicant: Alkermes, Inc.

Date(s): Submission data: August 22, 2014
PDUFA due date: August 22, 2015

Review Priority: Standard

Biometrics Division: Division of Biometrics I

Statistical Reviewer: Jinglin Zhong, Ph.D.

Concurring Reviewers: Peiling Yang, Ph.D.
HM James Hung, Ph.D.

Medical Division: Division of Psychiatry Products

Clinical Team: Lucas Kempf, M.D.

Project Manager: Sharonjit Sagoo, Pharm. D.

Keywords: Clinical Study; NDA review; Schizophrenia

Table of Contents

1	EXECUTIVE SUMMARY	5
2	INTRODUCTION	5
2.1	OVERVIEW.....	5
2.2	DATA SOURCES	5
3	STATISTICAL EVALUATION	5
3.1	DATA AND ANALYSIS QUALITY	5
3.2	EVALUATION OF EFFICACY	6
3.2.1	<i>Study Design and Endpoints</i>	6
3.2.2	<i>Statistical Methodologies</i>	7
3.2.3	<i>Patient Disposition, Demographic and Baseline Characteristics</i>	9
3.2.4	<i>Results and Conclusions</i>	14
4	FINDINGS IN SPECIAL/SUBGROUP POPULATIONS	23
4.1	GENDER, RACE, AGE, AND GEOGRAPHIC REGION	23
4.2	OTHER SPECIAL/SUBGROUP POPULATIONS	27
5	SUMMARY AND CONCLUSIONS	27
5.1	STATISTICAL ISSUES	27
5.2	COLLECTIVE EVIDENCE	28
5.3	CONCLUSIONS AND RECOMMENDATIONS	28
5.4	LABELING RECOMMENDATIONS (AS APPLICABLE).....	28

LIST OF TABLES

Table 1: List of all studies included in analysis	5
Table 2: Missing data handling in responder analysis	8
Table 3: Definition of failure event based on treatment outcome and/or dropout cause in time-to-failure analysis	9
Table 4: Disposition of subjects	10
Table 5: Summary of demographics and baseline characteristics for safety measurements (randomized population).....	11
Table 6: Summary of baseline efficacy parameters (full analysis set).....	12
Table 7: Summary of major protocol violations (randomized population).....	12
Table 8: Summary of patients excluded from FAS.....	13
Table 9: Change from baseline in PANSS total score for each stage and combined stages ANCOVA LOCF (FAS).....	15
Table 10: Change from baseline at Day 85 in PANSS total score MMRM (FAS).....	16
Table 11: Change from baseline at Day 85 in PANSS total score ANCOVA with LOCF (PP).....	17
Table 12: Change from baseline at Day 85 in PANSS total score, non-parametric rank ANCOVA model, LOCF (FAS)	17
Table 13: Non-parametric PANSS responder analysis at Day 85 (FAS).....	17
Table 14: Time to failure analysis (FAS)	18
Table 15: Change from baseline at Day 85 in PANSS total score (missing data is imputed by the mean of change at Day 85 from the placebo arm)	19
Table 16: Change from baseline at Day 85 in PANSS total score (missing data is imputed by the worst observed change at Day 85)	19
Table 17: Change from baseline at Day 85 in PANSS total score MMRM.....	21
Table 18: Change from baseline in PANSS total score using ANCOVA LOCF (Day 22 as Baseline)	22
Table 19: Change from baseline in PANSS total score using ANCOVA LOCF (Day 29 as Baseline)	22
Table 20. CGI-I score at Day 85, LOCF (FAS).....	23

LIST OF FIGURES

Figure 1: Study design schematic	7
Figure 2: Change from baseline in PANSS total score by visit by ANCOVA LOCF (FAS)	14
Figure 3: Change in PANSS total score at endpoint (FAS with Subject 502-002 using LOCF)	20
Figure 4: Change in PANSS total score at endpoint (FAS with Subject 502-002 imputed by the mean of change at Day 85 from the placebo arm)	20
Figure 5: Change in PANSS total score at endpoint (FAS with Subject 502-002 imputed by the worst change at Day 85).....	21
Figure 6. Box plot of change from baseline to endpoint in PANSS total score by gender and treatment (FAS with Subject 502-002 using LOCF).....	24
Figure 7. Box plot of change from baseline to endpoint in PANSS total score by race and treatment (FAS with Subject 502-002 using LOCF).....	25
Figure 8. Observed change from baseline in PANSS total (FAS and North America)	25
Figure 9. Observed change from baseline in PANSS total (FAS and Europe).....	26
Figure 10. Observed change from baseline in PANSS total (FAS and Asia)	26
Figure 11. Box plot of change from baseline to endpoint in PANSS total score by BMI category and treatment (FAS with Subject 502-002)	27

1 EXECUTIVE SUMMARY

This review describes statistical findings about the sponsor’s study report ALK9072-003 supporting the request for approval of aripiprazole lauroxil in subjects with schizophrenia.

This review confirms sponsor’s finding from ALK9072-003 that both aripiprazole lauroxil dose levels (441 mg and 882 mg) were statistically better than placebo as measured by change from baseline to Day 85 in Positive and Negative Syndrome Scale (PANSS) total score in treating adult subjects with schizophrenia.

2 INTRODUCTION

2.1 Overview

Aripiprazole lauroxil has been developed by Alkermes as an intramuscular (IM) injectable extended-release atypical antipsychotic for the treatment of schizophrenia. This NDA submission includes one pivotal safety and efficacy study (ALK9072-003) to support the efficacy of two aripiprazole lauroxil dose levels (441 mg and 882 mg) in subjects with schizophrenia.

Table 1: List of all studies included in analysis

Protocol Number	Phase and Design	Treatment Period	Follow-up Period	# of Subjects per Arm	Study Population
ALK9072-003	Phase 3, double-blind, placebo-controlled, conducted at 107 centers in US, Asia, and Europe	12 weeks double blind treatment	8 weeks	Placebo 208 441 mg 207 882 mg 208	Adults 18 – 70 with acute exacerbation of schizophrenia

2.2 Data Sources

Electronic datasets and study reports are located at:
<\\CDSESUB1\evsprod\NDA207533\0000>

3 STATISTICAL EVALUATION

3.1 Data and Analysis Quality

The data quality is fine. The FDA statistical reviewer can reproduce the primary analysis dataset from the original data source. Final statistical analysis plan (SAP) was submitted prior to unblinding. The blind was maintained until the database was locked on 31 March 2014.

3.2 Evaluation of Efficacy

3.2.1 Study Design and Endpoints

This was a multicenter, randomized, double-blind, placebo-controlled study designed to evaluate aripiprazole lauroxil in subjects with schizophrenia experiencing an acute exacerbation episode.

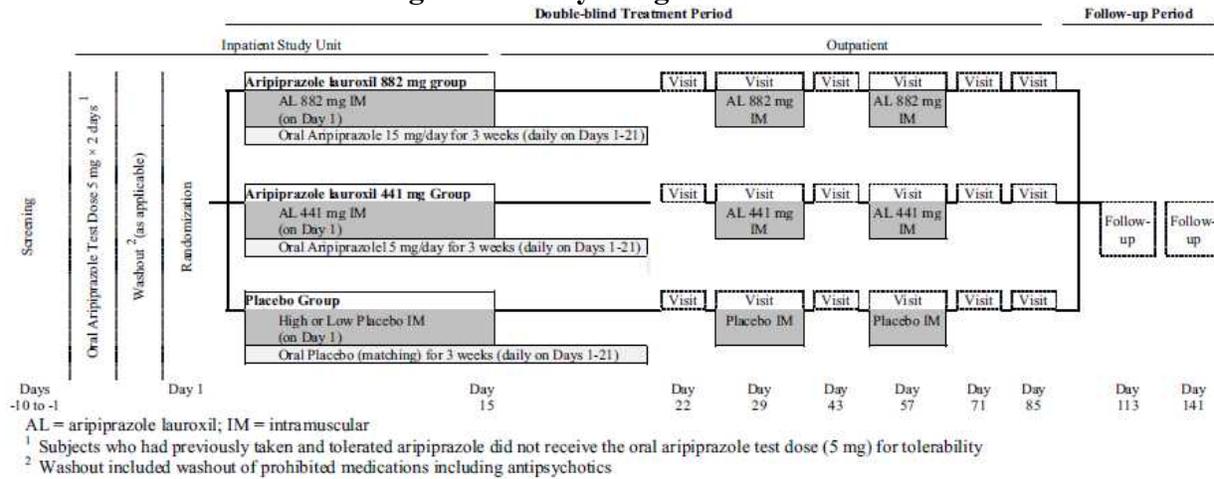
Subjects meeting initial screening eligibility criteria were admitted to an inpatient study unit. Currently prescribed antipsychotics were required to be discontinued after screening, and prior to administration of IM study drug. The allowable washout period was 2-5 days. For subjects who had never taken aripiprazole, a test dose of oral aripiprazole 5 mg was administered by mouth daily for 2 days prior to randomization, in order to assess individual tolerability prior to proceeding to injectable study drug.

Subjects who successfully completed screening and baseline assessments and tolerated the oral aripiprazole test doses, or had a history of safe and well-tolerated exposure to aripiprazole, were randomized on Day 1 in a 1:1:1 ratio to 1 of the following 3 treatment groups: aripiprazole lauroxil 882 mg (equivalent to ALKS 9072 600 mg), aripiprazole lauroxil 441 mg (equivalent to ALKS 9072 300 mg), or placebo (Intralipid®). The first dose of IM study drug was administered on Day 1. The second dose of IM study drug was administered on Day 29. The third and final dose of IM study drug was administered on Day 57.

In addition to IM study drug, subjects received oral study drug daily for the first 3 weeks after randomization. Oral study drug was administered in a double-blind fashion. Subjects randomized to an aripiprazole lauroxil treatment group received oral aripiprazole, and subjects randomized to the placebo group received matching oral placebo.

Subjects remained in the inpatient study unit for at least 2 weeks after administration of the first dose of IM study drug. Subjects were discharged from the inpatient facility when assessed and determined by the study investigator to be clinically stable and appropriate for discharge. Efficacy, safety, tolerability, and PK were measured throughout the treatment period. Two monthly follow-up visits were scheduled after the last IM injection. Subjects were given the option to enroll in an extension study with aripiprazole lauroxil. For subjects who chose to participate in the extension study, follow-up visits were not required. A schematic of the overall study design is shown in Figure 1.

Figure 1: Study design schematic



Source: Sponsor’s Figure 1 in Clinical Study Report (CSR).

The primary efficacy endpoint was change in Positive and Negative Syndrome Scale (PANSS) total score from baseline to Day 85 using an analysis of covariance (ANCOVA) model with last observation carried forward (LOCF) imputation. Clinician Global Impression - Improvement (CGI-I) scores at Day 85 was the secondary efficacy endpoint.

The placebo-controlled, parallel-group design is a typical design for IM injection treatment of schizophrenia. However, in this trial subjects from active arms received oral drug during the first 21 days, while placebo patients did not. The observed treatment effect is likely due to the combination of IM and oral drug. It is hard to separate the treatment effect due to IM from that due to the oral drug. During the review course, a question was raised on whether or not the placebo group should have also received the oral drug instead of the oral placebo during the first 21 days. This reviewer performed several analyses to address this design problem.

3.2.2 Statistical Methodologies

The full analysis set (FAS) consists of all evaluable subjects, defined as all randomized subjects who receive at least 1 dose of IM study drug and have at least 1 primary efficacy assessment after administration of IM study drug. The FAS is the primary efficacy population.

The primary efficacy endpoint was analyzed by the ANCOVA model using the LOCF approach. The ANCOVA model includes change from baseline in the PANSS total score at Day 85 as the dependent variable, with study region and treatment group as fixed effects, and baseline PANSS total score as a covariate.

An unblinded interim analysis on the 271 subjects (50% of planned sample size) was conducted for sample size re-estimation only based on the conditional power. The CHW method together with Hommel procedure was used for controlling the Type 1 error rate due to the interim analysis and multiple comparisons. The Cui, Hung, Wang (CHW) method combined 2 independent statistical results with an equal weight (sqrt(0.5)) from 2 stages. Stage 1 was based on the interim

analysis population (n=271), and Stage 2 was based on the post-interim population (n=325). The subjects in the 2 stages did not overlap.

The sponsor conducted the following sensitivity analyses: 1) mixed model for repeated measures (MMRM) analysis on FAS; 2) an ANCOVA model using the LOCF approach on the per-protocol population; 3) a non-parametric rank ANCOVA model (includes study region and treatment groups as factors and the baseline PANSS total score as a covariate); 4) model-free, non-parametric responder analyses; and 5) time-to-failure analyses. The details of some of the sensitivity analyses are presented below.

MMRM analysis was performed on observed data without imputation of missing data. The MMRM model uses the change from baseline in the PANSS total score at each post-baseline visit as the dependent variable, and includes study region, treatment group, visit, and treatment group-by-visit interaction as factors and baseline PANSS total score as a covariate. An unstructured covariance structure was applied for MMRM.

Model-free, non-parametric responder analyses were used for PANSS response at Day 85. It is difficult to differentiate the missing at random and missing not at random. Therefore, this plan combined these 2 types in the sensitive analyses, by treating either type of missing mechanisms as informative missing. Any informative missing data were treated as non-responders (failures) in the estimation of the treatment response rates. Any non-informative missing data were excluded from calculating the treatment response rates. Furthermore, 2 cutoffs (at least 30% and at least 20% reduction in PANSS total score) were used to define the response. A Chi-square test and interval estimate for the difference of the response rates were computed to compare the treatment response rate between each aripiprazole lauroxil group and the placebo group. Table 2 summarizes the sensitivity responder analysis under various dropout pattern scenarios and definitions of the treatment response to support the primary analysis.

Table 2: Missing data handling in responder analysis

Treatment Responder	Dropout Pattern	Dropouts			
		Lack of Efficacy	Adverse Event	Loss to Follow-up	Consent Withdrawal
≥30%	1	Informative	Non-informative	Non-informative	Non-informative
	2	Informative	Informative	Non-informative	Non-informative
≥20%	1	Informative	Non-informative	Non-informative	Non-informative
	2	Informative	Informative	Non-informative	Non-informative

Source: Sponsor’s Table 2 in SAP.

Time-to-failure analysis used different censoring rule to handle miss data. The detail is provided in Table 3.

Table 3: Definition of failure event based on treatment outcome and/or dropout cause in time-to-failure analysis

Patterns	Completers	Dropouts			
	Reduction from baseline in PANSS total score at week 12	Lack of Efficacy	Adverse Event	Lost to follow up	Consent Withdrawal
1	<20% (Failure)	Failure	Censored	Censored	Censored
2	<20% (Failure)	Failure	Failure	Censored	Censored
3	<30% (Failure)	Failure	Censored	Censored	Censored
4	<30% (Failure)	Failure	Failure	Censored	Censored

Source: Sponsor’s Table 3 in SAP.

When both comparisons (high dose vs. placebo and low dose vs. placebo) were statistically significant for the primary efficacy endpoint, the statistical test was performed on the secondary efficacy endpoint, CGI-I score at Day 85, using Hommel procedure at two-sided alpha of 0.05. The CGI-I score at Day 85 were analyzed using a non-parametric Wilcoxon rank sum test using the LOCF approach, and further confirmed with “no change” imputation for missing data as a sensitivity analysis.

3.2.3 Patient Disposition, Demographic and Baseline Characteristics

The disposition of the subjects is summarized in Table 4. The proportion of subjects who completed the treatment period is higher in the aripiprazole lauroxil groups (62.8% in the 441 mg group and 64.9% in the 882 mg group) than in the placebo group (45.9%). The most common reasons for discontinuation during the Treatment Period include withdrawal by subjects (13.8%), lack of efficacy (10.3%), and adverse event (9.0%). There were more discontinuations due to lack of efficacy in the placebo group (18.4%) than in the aripiprazole lauroxil groups (4.3% in the 441 mg group, 8.2% in the 882 mg group), and more discontinuations due to adverse event in the placebo group (17.4%) than in the aripiprazole lauroxil groups (6.8% in the 441 mg group, 2.9% in the 882 mg group).

The patient demographics are show in Table 5 for the safety population. As shown in Table 6, the mean PANSS score and the mean CGI-S score were similar across the treatment groups at baseline.

A summary of major protocol violations is presented in Table 7. The percentages of subjects with major deviations were similar across the treatment groups. Twelve subjects were discontinued for protocol violations (3 in placebo group, 6 in aripiprazole lauroxil 441 mg group, and 3 in aripiprazole lauroxil 882 mg group). The majority of protocol violations were enrollment criteria deviations or positive urine drug screens.

Table 4: Disposition of subjects

Category	All n (%)	Placebo n (%)	Aripiprazole Lauroxil n (%)	
			441 mg	882 mg
Screened	848			
Randomized	623	208	207	208
Safety Population	622	207	207	208
Treatment Period				
Completed Treatment Period	360 (57.9)	95 (45.9)	130 (62.8)	135 (64.9)
Discontinued Treatment Period	262 (42.1)	112 (54.1)	77 (37.2)	73 (35.1)
Withdrawal by Subject	86 (13.8)	22 (10.6)	35 (16.9)	29 (13.9)
Lack of Efficacy	64 (10.3)	38 (18.4)	9 (4.3)	17 (8.2)
Adverse Event	56 (9.0)	36 (17.4)	14 (6.8)	6 (2.9)
Lost to Follow-up	35 (5.6)	10 (4.8)	10 (4.8)	15 (7.2)
Protocol Violation	12 (1.9)	3 (1.4)	6 (2.9)	3 (1.4)
Other	4 (0.6)	1 (0.5)	0	3 (1.4)
Physician Decision	3 (0.5)	0	3 (1.4)	0
Death	1 (0.2)	1 (0.5)	0	0
Non-Compliance with Study Drug	1 (0.2)	1 (0.5)	0	0
Entered Extension Study	236 (37.9)	55 (26.6)	81 (39.1)	100 (48.1)
Follow-up Period^a				
Entered Follow-up Period	124	40	49	35
Completed Follow-up Period	113 (91.1)	36 (90.0)	46 (93.9)	31 (88.6)
Discontinued Follow-up Period	11 (8.9)	4 (10.0)	3 (6.1)	4 (11.4)
Lost to Follow-up	7 (5.6)	2 (5.0)	3 (6.1)	2 (5.7)
Adverse Event	2 (1.6)	1 (2.5)	0	1 (2.9)
Lack of Efficacy	1 (0.8)	0	0	1 (2.9)
Withdrawal by Subject	1 (0.8)	1 (2.5)	0	0

Note: Percentages are based on the total number of subjects in the safety population.

Safety population includes all subjects who receive at least 1 dose of IM study drug.

^a Percentages for follow-up period are relative to number of subjects in the follow-up period.

Source: Sponsor's Table 3 in CSR.

Table 5: Summary of demographics and baseline characteristics for safety measurements (randomized population)

Variable Category/ Statistics	Placebo (N=208)	Aripiprazole Lauroxil	
		441 mg (N=207)	882 mg (N=208)
Age (years)			
Mean (SD)	39.5 (11.85)	39.9 (10.13)	39.7 (11.06)
Median	38.5	39.0	40.0
Min - Max	18 - 66	18 - 61	20 - 65
Gender (n, %)			
Male	139 (66.8)	141 (68.1)	143 (68.8)
Female	69 (33.2)	66 (31.9)	65 (31.3)
Primary Race (n, %)			
White	94 (45.2)	99 (47.8)	98 (47.1)
Black or African-American	84 (40.4)	83 (40.1)	81 (38.9)
Asian	29 (13.9)	24 (11.6)	28 (13.5)
American Indian or Alaska Native	1 (0.5)	0	1 (0.5)
Native Hawaiian or other Pacific Islander	0	1 (0.5)	0
Region (n, %)			
North America	102 (49.0)	103 (49.8)	102 (49.0)
Europe	80 (38.5)	81 (39.1)	78 (37.5)
Asia	26 (12.5)	23 (11.1)	28 (13.5)
Weight (kg)			
Mean (SD)	78.94 (18.660)	80.78 (17.632)	80.26 (19.306)
Median	77.85	78.00	77.65
Min - Max	46.20 - 148.70	44.50 - 140.00	43.00 - 143.80
Body Mass Index (kg/m²)			
Mean (SD)	26.95 (5.094)	27.68 (5.332)	27.29 (5.673)
Median	26.45	27.10	26.15
Min - Max	18.5 - 39.5	18.5 - 40.3	18.6 - 40.9
Body Mass Index Categories (n, %)			
Normal (<25 kg/m ²)	83 (39.9)	73 (35.3)	81 (38.9)
Overweight (25≤BMI<30 kg/m ²)	70 (33.7)	62 (30.0)	67 (32.2)
Obese (≥30 kg/m ²)	55 (26.4)	72 (34.8)	60 (28.8)
CYP2D6 Predicted Phenotype (n, %)			
Extensive Metabolizer	120 (57.7)	127 (61.4)	117 (56.3)
Intermediate Metabolizer	46 (22.1)	45 (21.7)	52 (25.0)
Poor Metabolizer	6 (2.9)	6 (2.9)	8 (3.8)
Inconclusive	1 (0.5)	0	0
Alcohol Usage (n, %)			
Yes	36 (17.3)	33 (15.9)	30 (14.4)
No	172 (82.7)	174 (84.1)	178 (85.6)
Tobacco Use at Screening (n, %)			
Never	70 (33.7)	65 (31.4)	70 (33.7)
Current	126 (60.6)	135 (65.2)	130 (62.5)
Former	10 (4.8)	5 (2.4)	8 (3.8)

Max=maximum; Min=minimum.

Source: Sponsor's Table 6 in CSR.

Table 6: Summary of baseline efficacy parameters (full analysis set)

Variable Statistics	Placebo (N=196)	Aripiprazole Lauroxil	
		441 mg (N=196)	882 mg (N=204)
PANSS Total Score			
n	196	196	204
Mean (SD)	93.9 (11.28)	92.6 (10.20)	92.0 (10.77)
Median	93.0	92.0	91.5
Min - Max	65 - 143	70 - 119	68 - 119
CGI-S			
n	196	196	204
Mean (SD)	4.9 (0.61)	4.9 (0.59)	4.9 (0.61)
Median	5.0	5.0	5.0
Min - Max	4 - 6	4 - 6	4 - 7

Source: Sponsor's Table 7 in CSR.

Table 7: Summary of major protocol violations (randomized population)

Category	Placebo (N=208)	Aripiprazole Lauroxil	
		441 mg (N=207)	882 mg (N=208)
Have at least one major protocol violation	33 (15.9)	30 (14.5)	34 (16.3)
Positive Urine Drug Screen	14 (6.7)	18 (8.7)	20 (9.6)
Dosing	2 (1.0)	3 (1.4)	5 (2.4)
Visit/ Procedure Requirement	4 (1.9)	0	5 (2.4)
Enrollment Criteria	7 (3.4)	2 (1.0)	4 (1.9)
Other Antipsychotics Use	5 (2.4)	2 (1.0)	3 (1.4)
Concomitant Medication	1 (0.5)	3 (1.4)	1 (0.5)
Non-Compliance	1 (0.5)	1 (0.5)	1 (0.5)
Other	3 (1.4)	1 (0.5)	1 (0.5)
Visit Schedule	0	0	1 (0.5)
Laboratory	1 (0.5)	2 (1.0)	0

Source: Sponsor's Table 4 in CSR.

Among the 623 subjects randomized, one subject (501-015) did not receive IM study drug. The safety population included 622 subjects and 596 subjects were included in sponsor's FAS, which is the primary analysis data set. A summary of the 26 patients excluded from the FAS are listed below in Table 8. All of the 26 patients belong to Stage 2 and all had 1 IM injection. All the subjects except "502-002" had no post baseline PANSS total score. Subject 502-002 had a post baseline PANSS total of 117 after one IM injection. This FDA reviewer asked the sponsor the reason for excluding Subject 502-002 from FAS. The sponsor replied "Subject ALK9072003-502-002 was excluded from the full analysis set (FAS) as Items N5 (Difficulty in abstract thinking) and G12 (Lack of judgment and insight) of the PANSS for Day 85 (Early Termination Visit) was not rated (comments from investigator – unable to assess) and hence total PANSS was not derived for the ADaM dataset (see ALK9072-003 Clinical Study Report, Listing 16.2.6.1, pages 194 and 195)." According to the definition of FAS, this reviewer did not agree with the sponsor's decision of excluding Subject "502-002" from FAS. Therefore, this reviewer repeated all the analyses with and without Subject "502-002". Subject "502-002" was randomized to placebo.

Table 8: Summary of patients excluded from FAS

Subjid	Siteid	Country	Age	Sex	Race	Treatment Arm	Withdraw Reason	Baseline PANSS
103-005	103	USA	42	M	black	ALKS 9072 600 mg	Withdraw by Subject	100
104-002	104	USA	52	F	black	Placebo	Adverse Event	113
108-003	108	USA	42	M	black	ALKS 9072 300 mg	Protocol Violation	101
111-002	111	USA	22	M	black	Placebo	Protocol Violation	95
111-010	111	USA	25	M	black	ALKS 9072 300 mg	Withdraw by Subject	90
111-015	111	USA	36	M	black	ALKS 9072 300 mg	Withdraw by Subject	89
112-001	112	USA	49	F	black	ALKS 9072 300 mg	Withdraw by Subject	91
115-011	115	USA	34	M	black	ALKS 9072 600 mg	Lack of Efficacy	94
116-002	116	USA	29	M	black	ALKS 9072 600 mg	Withdraw by Subject	111
116-003	116	USA	42	M	White	ALKS 9072 300 mg	Withdraw by Subject	82
116-013	116	USA	29	M	White	ALKS 9072 300 mg	Withdraw by Subject	110
119-019	119	USA	48	M	black	ALKS 9072 300 mg	Withdraw by Subject	80
121-037	121	USA	41	F	black	ALKS 9072 600 mg	Withdraw by Subject	96
121-054	121	USA	28	F	black	Placebo	Withdraw by Subject	97
125-016	125	USA	27	M	black	ALKS 9072 300 mg	Withdraw by Subject	101
211-005	211	RUS	47	M	White	Placebo	Withdraw by Subject	88
220-004	220	RUS	24	M	White	Placebo	Adverse Event	88
310-001	310	UKR	38	F	White	ALKS 9072 300 mg	Withdraw by Subject	86
312-004	312	UKR	36	M	White	Placebo	Withdraw by Subject	97
314-003	314	UKR	23	F	White	Placebo	Lack of Efficacy	116
502-002	502	MYS	55	M	Asian	Placebo	Withdraw by Subject	86

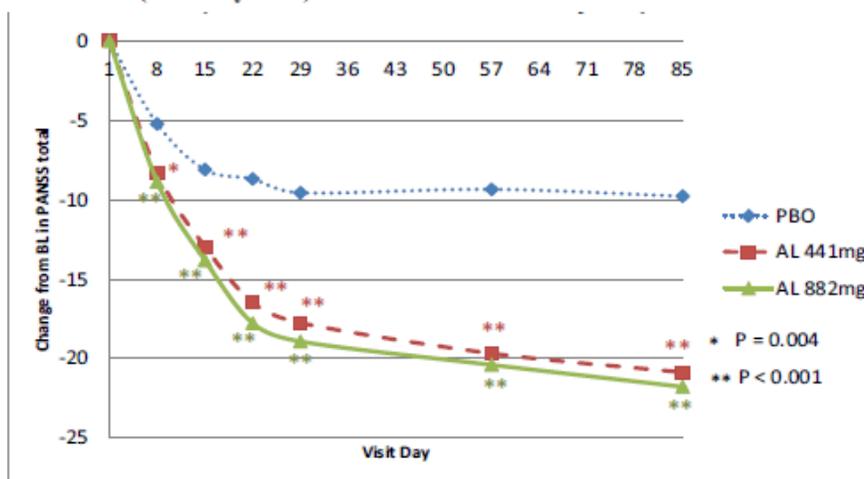
Subjid	Siteid	Country	Age	Sex	Race	Treatment Arm	Withdraw Reason	Baseline PANSS
503-005	503	MYS	31	M	Asian	Placebo	Withdraw by Subject	82
807-003	807	ROU	40	F	White	ALKS 9072 300 mg	LOST TO FOLLOW-UP	100
808-001	808	ROU	38	F	White	ALKS 9072 300 mg	Lack of Efficacy	104
808-003	808	ROU	50	F	White	Placebo	Withdraw by Subject	97
808-004	808	ROU	38	F	White	Placebo	Withdraw by Subject	89

Source: This reviewer.

3.2.4 Results and Conclusions

The change from baseline in PANSS total score by visit using ANOVA LOCF is presented in Figure 2.

Figure 2: Change from baseline in PANSS total score by visit by ANCOVA LOCF (FAS)



Abbreviations: AL=aripiprazole lauroxil; ANCOVA=analysis of covariance; PBO=placebo; PANSS=Positive and Negative Syndrome Scale; LOCF=last observation carried forward.

Source: Sponsor's Figure 2 in CSR.

An unblinded interim analysis on the 271 subjects (50% of planned sample size) was conducted for sample size re-estimation only based on the conditional power. The CHW method together with Hommel procedure was used for controlling the Type 1 error rate due to the interim analysis and multiple comparisons. The Cui, Hung, Wang (CHW) method combined 2 independent statistical results with an equal weight ($\sqrt{0.5}$) from 2 stages. Stage 1 was based on the interim analysis population ($n = 271$), and Stage 2 was based on the post-interim population ($n=325$). The subjects in the 2 stages did not overlap. Table 9 shows the analysis results from ANCOVA model in each stage, in all subjects without applying the CHW method, and in all subjects with applying CHW method.

Using the CHW method, the least square (LS) mean difference (standard error, SE) compared to placebo was -10.9 (1.82) for the aripiprazole lauroxil 441 mg group, and -11.9 (1.81) for the aripiprazole lauroxil 882 mg group. For both aripiprazole lauroxil groups, the difference is statistically significant and corresponds to $p < 0.001$. Similar results and statistical significance were also observed in the analysis without applying the CHW method.

Table 9: Change from baseline in PANSS total score for each stage and combined stages ANCOVA LOCF (FAS)

	Placebo (N=196)	Aripiprazole Lauroxil	
		441 mg (N=196)	882 mg (N=204)
Stage 1 (Interim Analyses subset)			
N	94	89	88
Baseline: Mean (SD)	94.1 (12.04)	93.0 (10.28)	92.1 (10.94)
Change from Baseline at Day 85: LS Mean (SE)	-9.7 (2.27)	-18.2 (2.26)	-19.7 (2.33)
LS Mean Difference against Placebo (SE)	--	-8.5 (2.50)	-10.0 (2.52)
Unadjusted p-value against placebo	-	<0.001	<0.001
Stage 2			
N	102	107	116
Baseline: Mean (SD)	93.7 (10.59)	92.3 (10.18)	92.0 (10.68)
Change from Baseline at Day 85: LS Mean (SE)	-10.0 (2.00)	-23.3 (1.97)	-23.7 (1.89)
LS Mean Difference against Placebo (SE)	--	-13.3 (2.65)	-13.7 (2.60)
Unadjusted p-value against placebo	-	<0.001	<0.001
All Subjects (without penalty)			
N	196	196	204
Baseline: Mean (SD)	93.9 (11.28)	92.6 (10.20)	92.0 (10.77)
Change from Baseline at Day 85: LS Mean (SE)	-9.8 (1.39)	-20.9 (1.39)	-21.8 (1.35)
LS Mean Difference against Placebo (SE)	--	-11.1 (1.84)	-12.0 (1.82)
Unadjusted p-value against placebo	-	<0.001	<0.001
Adjusted p-value against placebo	-	<0.001	<0.001
CHW Method: Combining Stage 1 & 2 (with penalty)			
Weighted LS Means Difference against Placebo (SE)	--	-10.9 (1.82)	-11.9 (1.81)
Test Statistics against Placebo	--	-5.951	-6.538
Test Statistics for Joint Test	-6.329		

Abbreviations: ANCOVA=analysis of covariance; CHW= Cui, Hung, and Wang; PANSS=Positive and Negative Syndrome Scale; LOCF=last observation carried forward; SD=standard deviation; SE=standard error.
 Note: The independent statistical results from two stages are combined using the CHW method with a pre-specified equal weight (sqrt(0.5)). The superiority of a particular aripiprazole lauroxil group is claimed using a closed test strategy if the test statistics against placebo and the test statistics for joint test are smaller than or equal to -1.96 at the 0.05 significance level.

Source: Sponsor's Table 11 in CSR.

Reviewer's note: The primary efficacy analysis uses LOCF to impute the missing data. Given the high dropout rate (42.1%) for this trial, the LOCF method is not very sensible. During the protocol/SAP review stage, FDA raised the concern about the LOCF method and requested the sponsor to use the MMRM method and other methods as sensitivity analyses. FDA also informed the sponsor that FDA "will seriously take the results from the sensitivity analyses into consideration to decide whether or not the primary analysis results are valid and reliable."

MMRM analysis was performed on observed data without imputation of missing data. The change from baseline in PANSS total score using MMRM is presented in Table 10. The least square (LS) mean difference (standard error, SE) compared to placebo was -11.76 (2.14) for the aripiprazole lauroxil 441 mg group, and -12.35 (2.12) for the aripiprazole lauroxil 882 mg group. For both aripiprazole lauroxil groups, the difference is statistically significant and corresponds to $p < 0.001$.

Table 10: Change from baseline at Day 85 in PANSS total score MMRM (FAS)

Visit Statistics	Placebo (N=196)	Aripiprazole Lauroxil	
		441 mg (N=196)	882 mg (N=204)
Baseline: Mean (SD)	93.9 (11.28)	92.6 (10.20)	92.0 (10.77)
Change from Baseline at Day 85			
n	96	133	137
LS Mean (SE)	-10.57 (1.60)	-22.33 (1.48)	-22.92 (1.43)
LS mean Difference against placebo (SE)		-11.76 (2.14)	-12.35 (2.12)
95% CI		-15.97, -7.56	-16.51, -8.19
p-value against placebo		<0.001	<0.001

Abbreviations: CI=confidence interval; LS=least squares; PANSS=Positive and Negative Syndrome Scale; SD=standard deviation; SE=standard error.

Note: Mixed model for repeated measures (MMRM) is based on the observed case without imputation of missing data. The MMRM model uses the change from baseline in the PANSS total score at each post-baseline visit as the dependent variable, and includes study region, treatment group, visit, and treatment group-by-visit interaction as factors and baseline PANSS total score as a covariate. An unstructured covariance structure will be applied for MMRM. The Kenward-Roger approximation is used to adjust the denominator degree of freedom.

Source: Sponsor's Table 13 in CSR.

Other sensitivity analyses were performed by the sponsor on the primary efficacy endpoint. The change from baseline in PANSS total score using ANCOVA model and LOCF approach on the per-protocol population is presented in Table 11. The change from baseline at Day 85 in PANSS total score using non parametric rank ANCOVA model and LOCF approach on FAS is presented in Table 12. The analysis result from the non-parametric PANSS responder analysis at Day 85 on FAS is presented in Table 13. The analysis result from the time to failure analyses is presented in Table 14. All the results show a statistically significant separation between each aripiprazole lauroxil group and placebo in favor of aripiprazole lauroxil groups.

This reviewer repeated the primary analysis and the sensitivity analyses from the raw data and obtained very similar results and the same conclusion. This reviewer included Subject "502-002" in FAS and repeated the analyses. Subject "502-002" received placebo and his last observation was 31 points higher (worse) than his baseline value. The LS mean of change from baseline for the remaining placebo patients is -9.8. Therefore, the analysis results including Subject "502-002" yield slightly more treatment effects for aripiprazole lauroxil groups comparing to placebo and the conclusions are all the same.

Table 11: Change from baseline at Day 85 in PANSS total score ANCOVA with LOCF (PP)

Visit Statistics	Placebo (N=194)	Aripiprazole Lauroxil	
		441 mg (N=193)	882 mg (N=202)
Baseline: Mean (SD)	93.5 (10.72)	92.7 (10.11)	92.1 (10.82)
Change from Baseline at Day 85			
n	192	193	201
LS Mean (SE)	-9.83 (1.414)	-20.48 (1.412)	-21.77 (1.375)
LS Mean Difference against Placebo (SE)	–	-10.65 (1.861)	-11.94 (1.844)
95% CI		-14.30, -6.99	-15.56, -8.32
p-value against placebo		<0.001	<0.001

Abbreviations: CI=confidence interval; LS=least squares; PANSS=positive and negative syndrome scale; SD=standard deviation; SE=standard error.

Note: For the ANCOVA model using the LOCF approach, the dependent variable is change from baseline in the PANSS total score at Day 85, with study region and treatment group as fixed effects and the baseline PANSS total score as a covariate.

Source: Sponsor's Table 14 in CSR.

Table 12: Change from baseline at Day 85 in PANSS total score, non-parametric rank ANCOVA model, LOCF (FAS)

Visit Statistics	Placebo (N=196)	Aripiprazole Lauroxil	
		441 mg (N=196)	882 mg (N=204)
All Subjects			
n	196	196	204
Baseline: Median	93.00	92.00	91.50
Change from Baseline at Day 85			
Median	-7.50	-20.00	-21.00
Median Difference against Placebo	--	-12.00	-12.00
95% CI	--	-16.00, -8.00	-16.00, -8.00
p-value against placebo	--	<0.001	<0.001

Abbreviations: ANCOVA=analysis of covariance; CI=confidence interval; LOCF=last observation carried forward; PANSS=positive and negative syndrome scale.

The median difference against placebo and its 95% confidence interval are based on the Hodges-Lehmann method. For the ANCOVA model using the LOCF approach, the dependent variable is the rank of change from baseline in the PANSS total score at Day 85, with study region and treatment group as fixed effects and the rank of baseline PANSS total score as a covariate.

Source: Sponsor's Table 15 in CSR.

Table 13: Non-parametric PANSS responder analysis at Day 85 (FAS)

PANSS Improvement threshold	Dropout Pattern ^a		Placebo (N=196) n/m (%)	Aripiprazole Lauroxil	
				441 mg (N=196) n/m (%)	882 mg (N=204) n/m (%)
≥30%	1	Response	33/165 (20.0)	63/152 (41.4)	63/157 (40.1)
		Difference against placebo (SE)	–	21.45 (5.066)	20.13 (5.000)
		p-value	–	<0.001	<0.001

≥30%	2	Response	33/131 (25.2)	63/138 (45.7)	63/151 (41.7)
		Difference against placebo (SE)	–	20.46 (5.689)	16.53 (5.522)
		p-value	–	<0.001	0.003
≥20%	1	Response	43/165 (26.1)	92/152 (60.5)	93/157 (59.2)
		Difference against placebo (SE)	–	34.47 (5.234)	33.18 (5.202)
		p-value	–	<0.001	<0.001
≥20%	2	Response	43/131 (32.8)	92/138 (66.7)	93/151 (61.6)
		Difference against placebo (SE)	–	33.84 (5.739)	28.76 (5.701)
		p-value	–	<0.001	<0.001

Abbreviations: n=number of subjects with a response; m=number of subjects with available data; PANSS= positive and negative syndrome scale; SE=standard error.

Note: p-values from Chi-square test.

^a Dropout pattern 1: dropout due to either lack of efficacy or adverse event as informative missing, and dropout due to all other reasons as non-informative missing. Dropout pattern 2 includes dropout due to lack of efficacy as informative missing, and all other reasons as non-informative missing.

Source: Sponsor's Table 16 in CSR.

Table 14: Time to failure analysis (FAS)

PANSS Reduction from Baseline	Dropout Pattern	Statistic	Placebo (N=196)	Aripiprazole Lauroxil	
				441 mg (N=196)	882 mg (N=204)
< 20%	1	Number (%) of subjects with failure	122/165 (73.9%)	60/152 (39.5%)	64/157 (40.8%)
		p-value	–	<.001	<.001
< 20%	2	Number (%) of subjects with failure	88/131 (67.2%)	46/138 (33.3%)	58/151 (38.4%)
		p-value	–	<.001	0.003
< 30%	1	Number (%) of subjects with failure	132/165 (80%)	89/152 (58.6%)	94/157 (59.9%)
		p-value	–	<.001	<.001
< 30%	2	Number (%) of subjects with failure	98/131 (74.8%)	75/138 (54.3%)	88/151 (58.3%)
		p-value	–	<.001	<.001

Note: In dropout pattern 1, dropout due to lack of efficacy or adverse event is considered a failure; in dropout pattern 2, dropouts due to lack of efficacy is considered a failure. p-value is from log rank test.

Source: Sponsor's Table 17 in CSR.

To assess the impact of the drop-outs on the primary efficacy endpoint, this reviewer conducted sensitivity analyses using two different imputing methods on those discontinued in treatment period. 1). Missing data are imputed by the mean of change at Day 85 from the placebo arm (-17.57). 2). Missing data are imputed by the worst observed change at Day 85 (24). The first imputation method keeps the mean of placebo arm unchanged, and moves the means of active arms close to the mean of placebo arm. Hence the treatment effects are reduced. The results are summarized in Table 15. Although the treatment effects are reduced, they are still statistically significant (p-value<0.0001). The second imputation method actually makes the treatment effects slightly larger (See Table 16). By imputing with the worst observed change at Day 85, the smaller magnitude of changes in the placebo group is offset by the larger number of drop outs in placebo group. Therefore, this imputation method also yields statistically significant differences between the active arms and placebo.

This reviewer also plotted the cumulative distribution function (CDF) for the change in PANSS total score from baseline to endpoint by each arm in Figure 3 (with LOCF), in Figure 4 (imputed

with the mean change at Day 85 from the placebo arm), and in Figure 5 (imputed with the worst observed change at Day 85). All three CDF plots show that the probability of having a change in PANSS total score less (better) or equal to any given x are larger for both active treatment arms compared to the placebo group regardless of the three imputation method.

This study planned to enroll 180 patients per arm, 540 patients total. However, the final number of patients in the FAS is 597. The reason of over enrollment is not clear. This reviewer performed the primary analysis using the first 180 patients per arm. Then result is presented in Table 17 and the conclusion is the same as using all the patients.

Table 15: Change from baseline at Day 85 in PANSS total score (missing data is imputed by the mean of change at Day 85 from the placebo arm)

	Placebo	Aripiprazole Lauroxil	
		441 mg	882 mg
Baseline: Mean (SD)	93.7 (11.4)	92.0 (10.73)	92.8 (10.27)
LS Mean (SE)	-19.30 (0.95)	-25.02(0.93)	-24.70(0.95)
LS Mean Diff from Placebo		-5.72(1.25)	-5.40(1.26)
95% CI*		(-8.18, -3.27)	(-7.88, -2.92)
p-value		<.0001	<.0001

*: The confidence intervals and p-values are not adjusted for multiple comparisons.

Source: This reviewer.

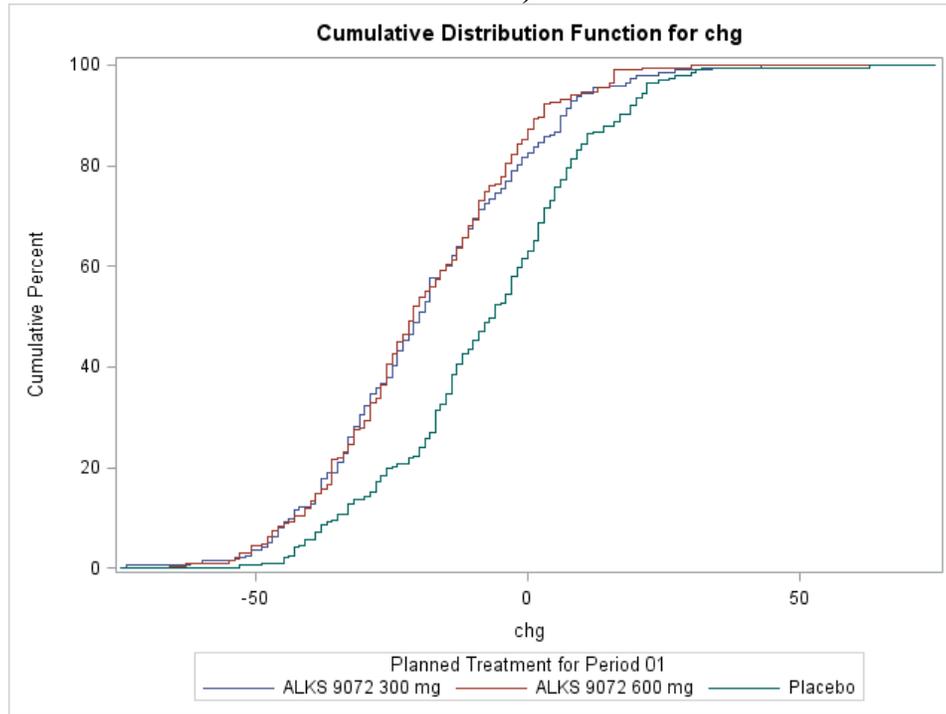
Table 16: Change from baseline at Day 85 in PANSS total score (missing data is imputed by the worst observed change at Day 85)

	Placebo	Aripiprazole Lauroxil	
		441 mg	882 mg
Baseline: Mean (SD)	93.7 (11.4)	92.0 (10.73)	92.8 (10.27)
LS Mean (SE)	0.75 (1.87)	-11.91(1.84)	-11.77(1.89)
LS Mean Diff from Placebo		-12.66(2.48)	-12.51(2.50)
95% CI*		(-17.52, -7.80)	(-17.42, -7.61)
p-value		<.0001	<.0001

*: The confidence intervals and p-values are not adjusted for multiple comparisons.

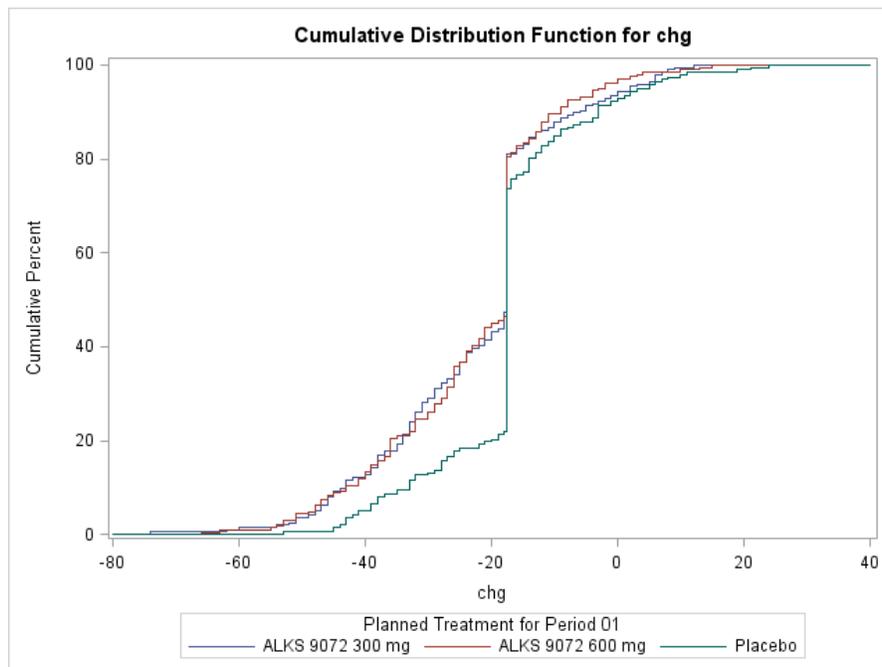
Source: This reviewer.

Figure 3: Change in PANSS total score at endpoint (FAS with Subject 502-002 using LOCF)



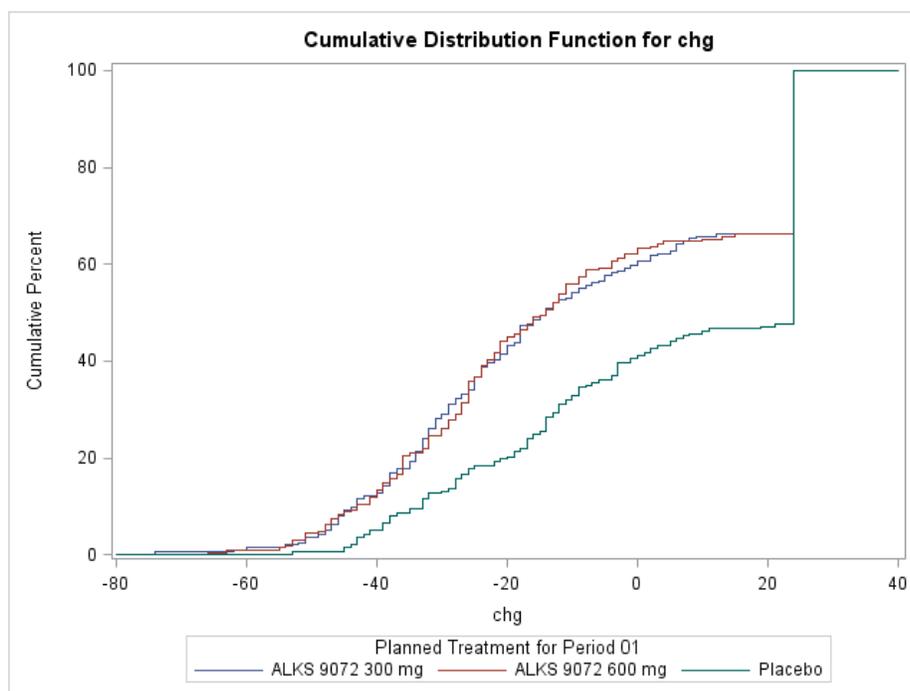
Source: This reviewer.

Figure 4: Change in PANSS total score at endpoint (FAS with Subject 502-002 imputed by the mean of change at Day 85 from the placebo arm)



Source: This reviewer.

Figure 5: Change in PANSS total score at endpoint (FAS with Subject 502-002 imputed by the worst change at Day 85)



Source: This reviewer.

Table 17: Change from baseline at Day 85 in PANSS total score MMRM (First 180 Patients per Arm in FAS)

	Placebo	Aripiprazole Lauroxil	
		441 mg	882 mg
Baseline: Mean (SD)	93.7 (11.4)	92.0 (10.73)	92.8 (10.27)
LS Mean (SE)	-10.7 (1.66)	-21.2(1.54)	-21.6(1.54)
LS Mean Diff from Placebo		-10.5(2.22)	-10.9(2.22)
95% CI*		(-14.9, -6.1)	(-15.3, -6.5)
p-value		<0.0001	<0.0001

*: The confidence intervals and p-values are not adjusted for multiple comparisons.

Source: This reviewer.

This placebo-controlled, parallel-group design is a typical design for IM injection treatment of schizophrenia. However, in this trial subjects from active arms receive oral drug during the first 21 days, while placebo patients did not. The observed treatment effect is likely due to both IM and oral drug. It is hard to separate the treatment effect due to IM from that due to the oral drug. During the review course, a question was raised on whether or not the placebo group should have also received the oral drug instead of the oral placebo during the first 21 days. This reviewer performed several analyses to address this design problem. The primary analysis was repeated 1) using Day 22 as baseline; 2) using Day 29 as baseline. Since the half-life of the oral drug is 3 days, the exposure contribution due to oral is likely to be insignificant by day 29. The FAS population includes 596 patients. When Day 22 or Day 29 is used as baseline, the FAS

population includes 455 patients or 412 patients, respectively. Table 18 and Table 19 summarize the results using Day 22 as baseline and using Day 29 as baseline. The results demonstrate that PANSS score was statistically significantly improved from Day 22 or Day 29 to Day 85 for each active treatment group compared with placebo, which support the claim that both aripiprazole lauroxil dose levels (441 mg and 882 mg) were statistically better than placebo.

Table 18: Change from baseline in PANSS total score using ANCOVA LOCF (Day 22 as Baseline)

	Total (n = 455)		
	Placebo (n=136)	441 mg (n =154)	882 mg (n = 165)
Day 22: Mean (SD)	79.9 (17.3)	74.6 (15.6)	73.3 (15.7)
Day 85: Mean (SD)	79.7 (20.2)	69.7 (18.6)	69.2 (19.0)
Change from Day 22 to Day 85: Mean (SD)	-0.2 (14.7)	-4.8 (11.3)	-4.2 (13.9)
LSmean change from Day 22 to Day 85 Mean(95% CI)	-0.6 (-2.9, 1.6)	-5.9 (-8.1, -3.7)	-5.3 (-7.4, -3.2)
Diff from placebo: LSMean (95% CI) pvalue		-5.3 (-8.3, -2.3) 0.0005	-4.6 (-7.6, -1.7) 0.002

Note: The confidence intervals and p-values are not adjusted by multiple comparisons.
Source: This reviewer.

Table 19: Change from baseline in PANSS total score using ANCOVA LOCF (Day 29 as Baseline)

	Total (n=412)		
	Placebo (n=115)	441 mg (n=144)	882 mg (n=153)
Day 29: Mean (SD)	76.5 (17.2)	72.6 (16.9)	71.5 (16.8)
Day 85: Mean (SD)	76.8 (19.1)	68.8 (18.6)	68.2 (18.9)
Change from Day 29 to Day 85: Mean (SD)	0.2 (13.8)	-3.9 (11.6)	-3.2 (12.1)
LSmean change from Day 29 to Day 85 Mean(95% CI)	-0.01 (-2.3, 2.3)	-4.5 (-6.6, -2.4)	-4.0 (-6.0, -2.0)
Diff from placebo: LSMean (95% CI) pvalue		-4.5 (-7.4, -1.6) 0.0026	-4.0 (-6.9, -1.1) 0.0073

Note: The confidence intervals and p-values are not adjusted by multiple comparisons.
Source: This reviewer.

The secondary efficacy endpoint is the CGI-I score at Day 85. The number (%) of subjects in each CGI-I category at Day 85 using LOCF are shown in Table 20. The CGI-I scores for both

aripiprazole lauroxil groups were statistically significantly lower than that for the placebo group ($p < 0.001$ for each aripiprazole lauroxil group) using a non-parametric Wilcoxon test. In addition to LOCF imputation, using “no change” to impute the missing values was used. The analyses yielded similar results as obtained using the LOCF approach.

This reviewer obtained similar results from raw data with and without subject “502-002”. Subject “502-002” had CGI-I score of 6 on his last visit.

Table 20. CGI-I score at Day 85, LOCF (FAS)

CGI-I Score	Placebo (N=196)	Aripiprazole Lauroxil	
		441 mg (N=196)	882 mg (N=204)
1: Very much improved n(%)	15 (7.7)	27 (13.8)	25 (12.3)
2: Much improved n(%)	33 (16.8)	68 (34.7)	81 (39.7)
3: Minimally improved n(%)	43 (21.9)	45 (23.0)	52 (25.5)
4: No change n(%)	42 (21.4)	32 (16.3)	24 (11.8)
5: Minimally worse n(%)	37 (18.9)	11 (5.6)	16 (7.8)
6: Much worse n(%)	23 (11.7)	12 (6.1)	5 (2.5)
7: Very much worse n(%)	3 (1.5)	1 (0.5)	1 (0.5)
P-value against placebo ¹	-	< 0.001	< 0.001
Adjusted p-value ²		< 0.001	< 0.001

Abbreviations: CGI-I=Clinical Global Impression-Improvement; LOCF=last observation carried forward;

PANSS=positive and negative syndrome scale. Source: [Table 14.2.3.1](#)

¹ p-values are based on non-parametric Wilcoxon rank sum test

² Adjusted p-values using Simes-Hommel approach

4 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

The subgroup analyses presented in this section are all exploratory. The main objective of the exploratory subgroup analysis is to assess consistency across subgroups with respect to the primary analysis results. The sponsor only performed ANCOVA analyses on BMI subgroups and regions, and reported p-values for subgroup factors. Because of the exploratory purpose of the subgroup analyses, those p-values are not presented here.

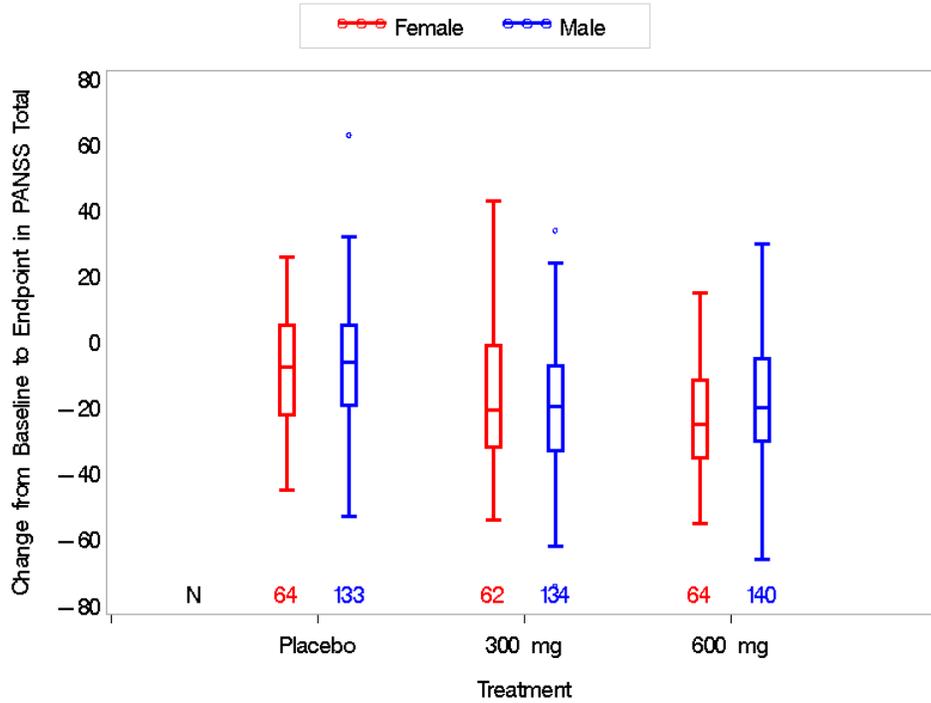
4.1 Gender, Race, Age, and Geographic Region

The age of the subjects ranges from 18 to 66. Therefore, no subgroup analyses by age are relevant.

This reviewer plotted the changes from baseline in PANSS total score for each gender and dose group, and for each race and dose group. There are 2 subjects belong to American Indian or Alaska Native, and 1 subject belongs to Native Hawaiian or other pacific islander. These 3 subjects were combined with Asian subjects. This reviewer also plotted the observed treatment effect over the trial course on the primary endpoint for each region. From the plots, we can see

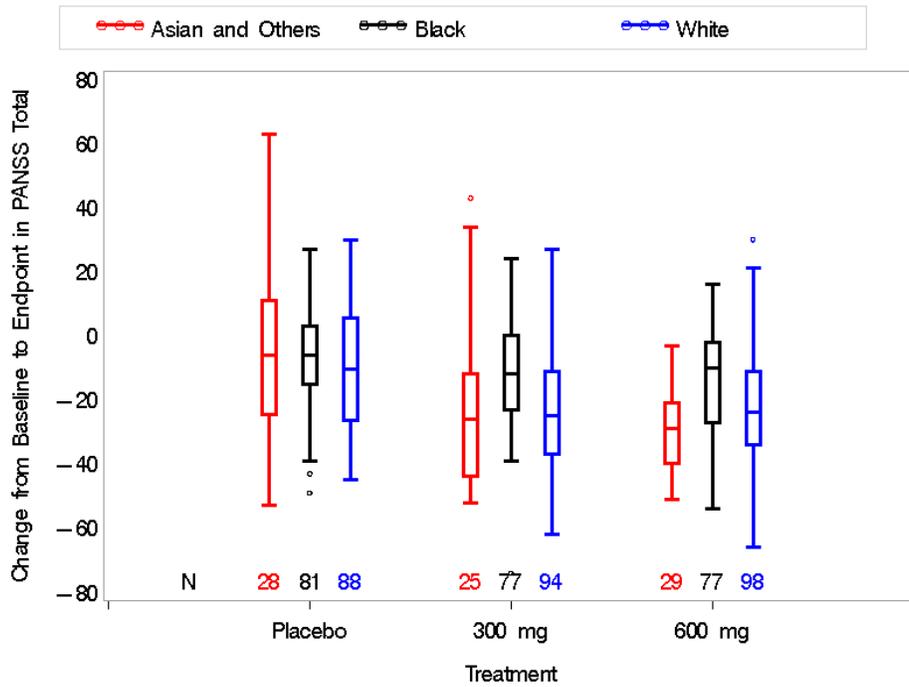
that Black subjects have smaller treatment effect than the other races. North American subjects have smaller treatment effect than the other regions.

Figure 6. Box plot of change from baseline to endpoint in PANSS total score by gender and treatment (FAS with Subject 502-002 using LOCF)



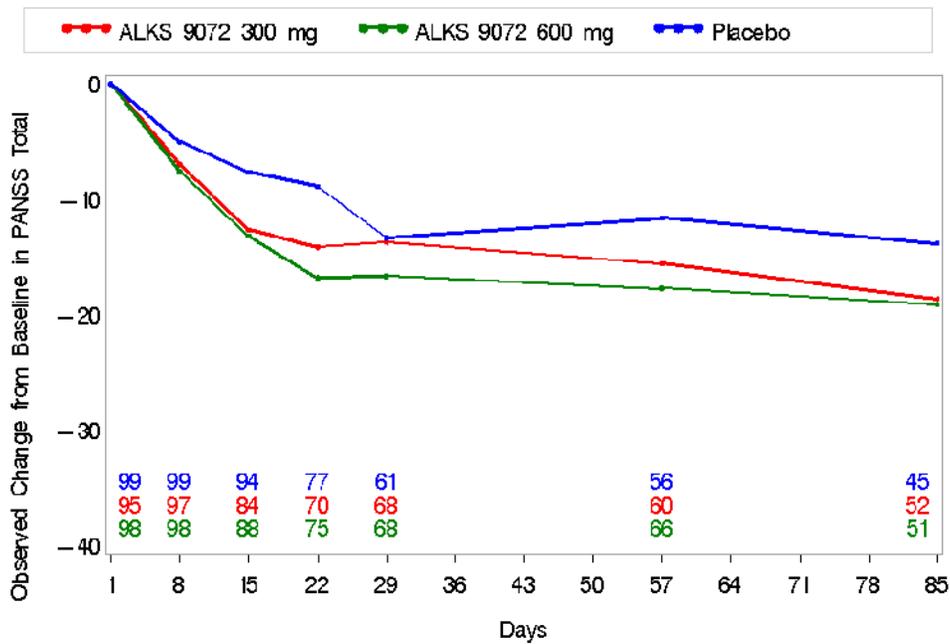
Source: This reviewer.

Figure 7. Box plot of change from baseline to endpoint in PANSS total score by race and treatment (FAS with Subject 502-002 using LOCF)



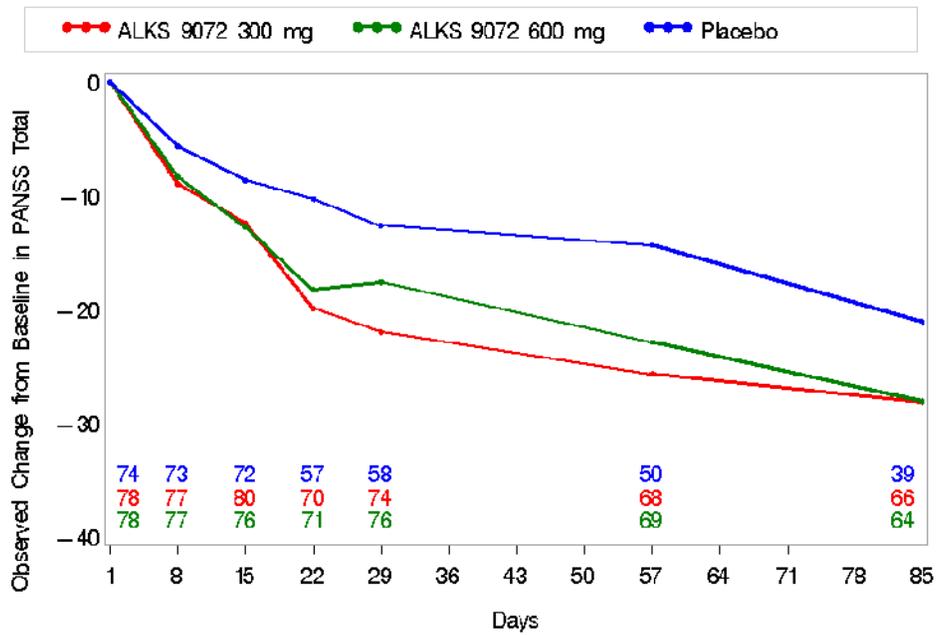
Source: This reviewer.

Figure 8. Observed change from baseline in PANSS total (FAS and North America)
Region: North America



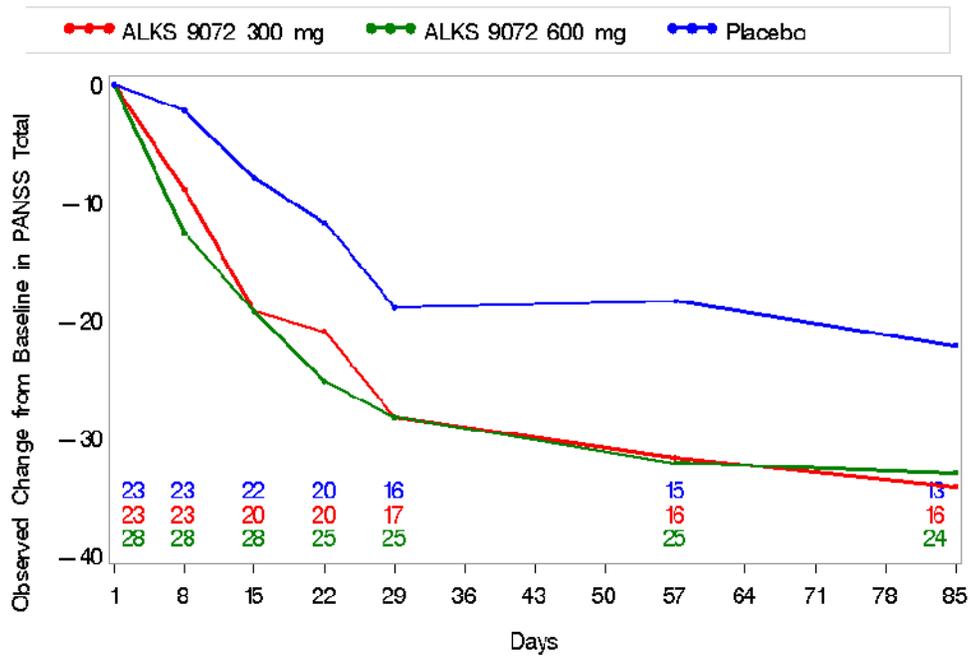
Source: This reviewer.

Figure 9. Observed change from baseline in PANSS total (FAS and Europe)
Region: Europe



Source: This reviewer.

Figure 10. Observed change from baseline in PANSS total (FAS and Asia)
Region: Asia

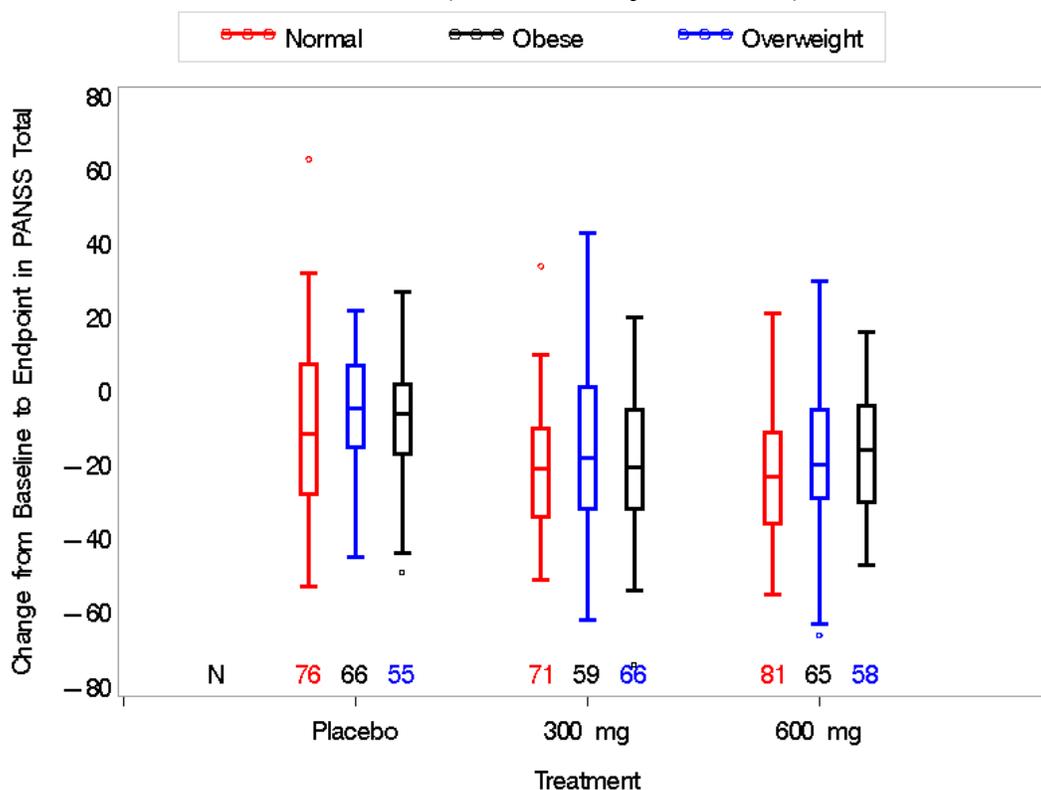


Source: This reviewer.

4.2 Other Special/Subgroup Populations

The change from baseline in PANSS total score was plotted by BMI subgroups with 3 different categories: normal ($<25 \text{ kg/m}^2$), overweight ($25 \text{ kg/m}^2 \leq \text{BMI} < 30 \text{ kg/m}^2$), and obese ($\leq 30 \text{ kg/m}^2$). The treatment effect is consistent across all the BMI subgroups.

Figure 11. Box plot of change from baseline to endpoint in PANSS total score by BMI category and treatment (FAS with Subject 502-002)



Source: This reviewer.

5 SUMMARY AND CONCLUSIONS

5.1 Statistical Issues

This placebo-controlled, parallel-group design is a typical design for IM injection treatment of schizophrenia. However, in this trial subjects from active arms receive oral drug during the first 21 days, while placebo patients did not. The observed treatment effect is likely due to both IM and oral drug. It is hard to separate the treatment effect due to IM from that due to the oral drug in this design. During the review course, a question was raised on whether or not the placebo group should have received the oral drug instead of the oral placebo during the first 21 days. This reviewer performed several analyses to address this design problem. The primary analysis was repeated 1) using Day 22 as baseline; 2) using Day 29 as baseline. Since the half-life of the oral drug is 3 days, the exposure contribution due to oral is likely to be insignificant by day 29. The

results demonstrate that PANSS score was statistically significantly improved from Day 22 or Day 29 to Day 85 for each active treatment group compared with placebo, which support the claim that both aripiprazole lauroxil dose levels (441 mg and 882 mg) were statistically better than placebo.

5.2 Collective Evidence

The primary and the key efficacy endpoint of this study were met. The change from baseline or Day 22 or Day 29 to Day 85 in PANSS total score was statistically significantly better for each active treatment group (aripiprazole lauroxil 441 mg and 882 mg) than for placebo. CGI-I scores at Day 85 was statistically significantly better for each active treatment group (aripiprazole lauroxil 441 mg and 882 mg) than for placebo.

5.3 Conclusions and Recommendations

This reviewer concludes, based on statistical evidence in Trial ALK9072-003, aripiprazole lauroxil 441 mg and 882 mg are effective.

5.4 Labeling Recommendations (as applicable)

NA

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

/s/

JINGLIN ZHONG
06/04/2015

PEILING YANG
06/04/2015
I concur that the efficacy is demonstrated.

HSIEN MING J HUNG
06/08/2015