

CENTER FOR DRUG EVALUATION AND RESEARCH

Application Number 21-436

STATISTICAL REVIEW(S)



DEPARTMENT OF HEALTH AND HUMAN SERVICES
PUBLIC HEALTH SERVICE
FOOD AND DRUG ADMINISTRATION
CENTER FOR DRUG EVALUATION AND RESEARCH

STATISTICAL REVIEW AND EVALUATION

Medical Division: Division of Neuropharmacological Drug Products (HFD-120)

Biometrics Division: Division of Biometrics I (HFD-710)

NDA NUMBER/SERIAL NUMBER:	21-436
DATE RECEIVED BY CENTER:	Oct. 31, 2001
DRUG NAME:	ABILITAT (ARIPIRAZOLE) TABS
INDICATION:	Schizophrenia
SPONSOR:	Otsuka Pharmaceutical Co., Ltd. & Bristol-Myers Squibb
STATISTICAL REVIEWER:	Yeh-Fong Chen, Ph.D.
STATISTICAL TEAM LEADER:	Kun Jin, Ph.D.
BIOMETRICS DIVISION DIRECTOR:	George Chi, Ph.D.
CLINICAL REVIEWER:	Gregory Dubitsky, M.D.
PROJECT MANAGER:	Steve Hardeman

**APPEARS THIS WAY
ON ORIGINAL**

I. TABLE OF CONTENTS

1. Introduction and Background-----	5
2. Summary of the Sponsor's Results and Conclusions-----	6
3. Description of the Sponsor's Studies and Statistical Methodologies-----	7
3.1 Pivotal Phase III Studies-----	7
3.1.1 Study 31-97-201-----	7
3.1.1.1 Study Objectives-----	8
3.1.1.2 Study Design-----	8
3.1.1.3 Efficacy Variables-----	8
3.1.1.4 Statistical Methods-----	9
3.1.2 Study 31-97-202-----	12
3.1.2.1 Study Objectives-----	12
3.1.2.2 Study Design-----	12
3.1.2.3 Efficacy Variables-----	12
3.1.2.4 Statistical Methods-----	12
3.1.3 Study CN 138-001-----	13
3.1.3.1 Study Objectives-----	13
3.1.3.2 Study Design-----	13
3.1.3.3 Efficacy Variables-----	14
3.1.3.4 Statistical Methods-----	14
3.2 Phase II Studies-----	17
3.2.1 Study 31-93-202-----	17
3.2.1.1 Study Objectives-----	17
3.2.1.2 Efficacy Outcome Variables-----	18
3.2.1.3 Study Design-----	18
3.2.1.4 Statistical Methods-----	19
3.2.2 Study 31-94-202-----	20
3.2.2.1 Study Objectives-----	20
3.2.2.2 Efficacy Outcome Variables-----	20
3.2.2.3 Study Design-----	20
3.2.2.4 Statistical Method-----	21
3.3 Long Term Studies: Studies 31-98-217 and 31-98-304-01-----	22
3.3.1 Study Objectives-----	22
3.3.2 Methodology-----	22
3.3.3 Statistical Methods-----	23
4. Detailed Review of the Sponsor's Individual Study Results-----	24
4.1 Pivotal Phase III Studies-----	24
4.1.1 Study 31-97-201-----	24
4.1.1.1 Disposition of Patients-----	24
4.1.1.2 Data Set-----	24
4.1.1.3 Demography and Patient Characteristics-----	25
4.1.1.4 The Sponsor's Efficacy Results-----	26
4.1.1.5 The Sponsor's Overall Efficacy Conclusions-----	34
4.1.2 Study 31-97-202-----	34
4.1.2.1 Disposition of Patients-----	34
4.1.2.2 Data Set-----	35
4.1.2.3 Demography and Patient Characteristics-----	35
4.1.2.4 The Sponsor's Efficacy Results-----	36
4.1.2.5 The Sponsor's Overall Efficacy Conclusions-----	43
4.1.3 Study CN 138-001-----	44
4.1.3.1 Disposition of Patients-----	44
4.1.3.2 Data Set-----	44
4.1.3.3 Demography and Patient Characteristics-----	44

4.1.3.4	The Sponsor's Efficacy Results -----	45
4.1.3.5	The Sponsor's Overall Efficacy Conclusions -----	52
4.2	Phase II Studies -----	53
4.2.1	Study 31-93-202 -----	53
4.2.1.1	Disposition of Patients -----	53
4.2.1.2	Demographics and Patient Characteristics -----	53
4.2.1.3	The Sponsor's Efficacy Results -----	53
4.2.1.4	The Sponsor's Overall Efficacy Conclusions -----	56
4.2.2	Study 31-94-202 -----	57
4.2.2.1	Disposition of Patients -----	57
4.2.2.2	Demographics and Patient Characteristics -----	57
4.2.2.3	The Sponsor's Efficacy Results -----	58
4.2.2.4	The Sponsor's Overall Efficacy Conclusions -----	59
4.3	Long-Term Studies: Studies 31-98-217 and 31-98-304-01 -----	59
4.3.1	Disposition of Patients -----	59
4.3.2	Data Sets -----	60
4.3.3	Demography and Patient Characteristics -----	60
4.3.4	The Sponsor's Efficacy Results -----	60
4.3.5	The Sponsor's Overall Efficacy Conclusions -----	63
5.	Statistical Reviewer's Findings and Comments -----	63
5.1	Pivotal Phase III Studies -----	63
5.2	Phase II Studies -----	70
5.3	The Sponsor's Efficacy Conclusions -----	70
5.4	Additional Comment -----	71
6.	Appendices -----	74

II. EXECUTIVE SUMMARY OF STATISTICAL REVIEWER'S FINDINGS

2.1 For Pivotal Phase III Studies: Studies 31-97-201, 31-97-202 and CN138-001

- The sponsor did not provide decision rules for three prospectively specified primary efficacy endpoints in Studies 31-97-201 and 31-97-202. They did address in the protocol using unadjusted type I error rate $\alpha = 0.05$ to perform the test for each primary endpoint. So, to control for the overall type I error rate, this reviewer made conclusions for both studies based on significant results shown on all three primary endpoints.
- This reviewer generally confirmed the sponsor's statistical results.
- For Study 31-97-201, the sensitivity analyses for three primary endpoints by removing the data from the invalid centers (#007 and #001) led same conclusions as the overall data analyses.
- For Study 31-97-202, except one of three primary efficacy endpoints, the comparisons between aripiprazole group and placebo were significant on the LOCF analyses but insignificant on the OC analyses. So, the dropout cohort analyses were studied to learn the possible bias of LOCF and OC analyses.
- For Studies 31-97-201 and 31-97-202, the patients with schizoaffective diagnosis alone were analyzed to compare with the patients with schizophrenia only. It was found that the treatment effects between these two subgroups did not differ much on all three primary endpoints in both studies.

- Due to a large amount of patients who chose the open-label aripiprazole treatment during Week 4 to 6, the results of OC analysis for the primary endpoint, i.e., PANSS Total Score, showed insignificant after Week 4 although the results of LOCF analyses were significant. Since the results of OC analyses were significant from Week 1 to Week 3, the insignificant results of OC analyses was not a concern.
- In conclusion, all three pivotal studies were positive. However, for Study 31-97-202, this reviewer had a concern about the biasness of the LOCF and OC analysis results.

2.2 For Phase II Studies: Studies 31-93-202 and 31-94-202

- For Study 31-93-202, the sponsor performed different statistical analyses from what was specified in the protocol for two primary efficacy endpoints and showed significant results. After they were requested to perform the protocol specified methods for them, it was found that the study was negative.
- Study 31-94-202 became negative after the invalid Center 003 was removed from the overall data set.

2.3 Long-Term Studies: Studies 31-98-217 and 31-98-304-01

- Despite the sponsor pooled the data from both studies, which was not generally acceptable, the test result for the primary efficacy endpoint was still insignificant. So, there was no question that studies were negative.

2.4 Additional Comment (Subgroup Analyses)

- The sponsor did not perform the complete subgroup analyses for age, gender and race for individual studies. They reported a table for model-based mean change of PANSS Total Score from baseline at endpoint by gender, age, race and baseline score in the LOCF data set of combined studies. According to the table, it was noticed that Hispanic patients and patients who were ≥ 50 year old had high placebo responses. This reviewer performed the detailed subgroup analyses for the above categories and found that in each pivotal study, the placebo group's magnitude of mean change of PANSS Total Score was greater than one of aripiprazole groups in the older patients (age ≥ 50). Since the low magnitude of changes happened across different dosage groups, this reviewer has a concern about the aripiprazole's efficacy for patients who were greater than or equal to 50.

STATISTICAL REVIEW AND EVALUATION OF EVIDENCE

1. Introduction and Background

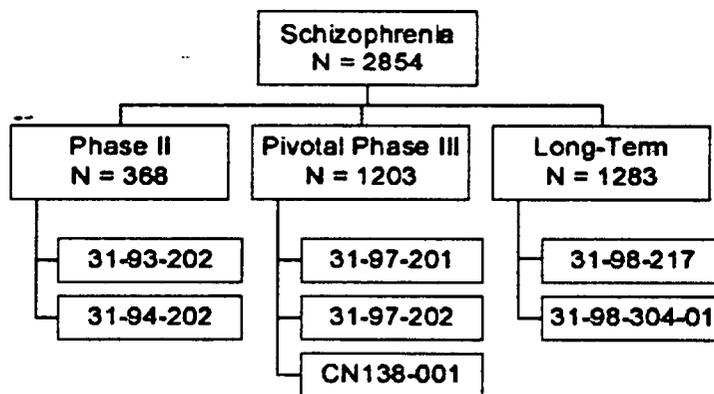
The Otsuka Pharmaceutical Co., Ltd submitted this application to present an overview of all the data in the Abilitat™ (aripiprazole) drug development program demonstrating a positive benefit/risk profile for the treatment of patients with schizophrenia. Collaborative development of aripiprazole between Otsuka Pharmaceutical Company (OPC), Otsuka Maryland Research Institute (OMRI) and Bristol-Myers Squibb (BMS) began in 1999. Collectively, the clinical program comprises 34 clinical pharmacology studies and 13 Phase II/III studies in schizophrenia.

Among those 13 Phase II/III studies, there were five short-term, four long-term and four special studies. According to the sponsor, all of the short-term studies met the FDA-defined criteria for adequate and well-controlled studies. Three of five studies were considered pivotal and two of five were supportive for efficacy analyses. The three pivotal Phase III studies were named Studies 31-97-201 and 31-97-202 (4-week fixed-dose studies, each with an active control) and CN138-001 (a 6-week, fixed-dose study). The two supportive Phase II studies were named Studies 31-93-202 (an ascending-dose study) and 31-94-202 (a fixed-dose study).

At the conclusion of the short-term studies, eligible patients were given the option of continuing on long-term treatment, either in the extension phase of the protocol that the patient had completed (for patients in Study CN138-001) or in an open-label long-term study. The other two double-blind, active-controlled, long-term studies, 31-98-217 and 31-98-304-01, enrolled patients who had not previously participated in an aripiprazole study. These two studies were prospectively designed to be analyzed together. They were 52 weeks in duration and assessed maintenance of efficacy versus haloperidol.

Figure 1 shows the diagram for the sponsor's studies that are pertinent to the efficacy of aripiprazole in the treatment of schizophrenia. This review will mainly focus on the evaluation for these 7 studies.

Figure 1. Studies That Are Pertinent to the Efficacy of Aripiprazole in the Treatment of Schizophrenia; Efficacy Sample



2. Summary of the Sponsor's Efficacy Results and Conclusions

For the Phase III Studies 31-97-201 and 31-97-202, the primary outcome measures were (1) the mean change from baseline to endpoint in the PANSS Total Score, (2) the mean change from baseline to endpoint in the PANSS Positive Sub-Scale Total Score, and (3) the mean change from baseline to endpoint in the CGI Severity of Illness Score. According to the sponsor's protocol, for each primary outcome measure, the treatment comparisons were tested by following the step-down procedure, i.e., first aripiprazole 30 mg vs. placebo was tested at two-tailed 0.05 level; if rejected, aripiprazole 15 mg vs. placebo was tested at two-tailed 0.05 level.

The primary outcome measure for the third Phase III Study, CN138-001, was the mean change from baseline to endpoint in the PANSS Total Score. In order to protect the experiment-wise alpha level at 0.05 level when making three comparisons of aripiprazole fixed doses versus placebo on the primary efficacy analyses, the statistical testing was carried out using Hochberg's sequentially rejective procedure. That is, superiority to placebo was claimed if all three pair-wise comparisons were significant at the 0.05 level, or two out of three were significant at the 0.025 level, or if one out of three was significant at the 0.0167 level.

In the Phase II ascending-dose study, 31-93-202, the primary outcome measures were (1) the mean change from baseline to endpoint in the BPRS Total Score and (2) the percentage of patients having improved by at least one point on the CGI Severity of Illness Score at endpoint. In the Phase II fixed-dose study, 31-94-202, the primary outcome measures were (1) the mean change from baseline to endpoint in the BPRS Core Score, and (2) the mean CGI Improvement Score at endpoint. In either study, no method preplanned for adjusting alpha 0.05 for the two primary endpoints. For Study 31-94-202, the Dunnett's procedure was pre-specified for adjusting the three dosage groups.

Studies 31-98-217 and 31-98-304-01 were designed to demonstrate the efficacy of aripiprazole versus haloperidol in long-term (up to 52 weeks) studies. The sponsor performed the analyses and reported the results based on the combined data. The primary efficacy measure for this combined studies was a time-to-event variable phrased as "time-to-failure to maintain response" in responders (defined in Section 3.3.3).

The summary of p-values for the primary endpoints of the five studies are shown in Tables 2.1 to 2.3. According to the analysis results, the sponsor concluded that aripiprazole is effective in the treatment of patients with schizophrenia. The clinical trial program established that the efficacy of aripiprazole was consistent and reproducible across the three pivotal Phase III studies and the two supportive Phase II studies, as well as the two studies (one analysis) that documented long-term efficacy against an active comparator. Within studies, aripiprazole demonstrated consistent efficacy across outcome measures that assessed positive symptoms (PANSS Positive Sub-Scale, PANSS-derived BPRS Core Score), negative symptoms (PANSS Negative Sub-Scale), and global measures of patient improvement (CGI Severity Score and CGI Improvement Score). However, after these studies were reviewed, it was determined that two phase II studies and the combined

analyses of two long-term studies were negative studies. Three phase III studies were positive but the analysis results shown in Study 31-97-202 seemed to be biased.

Table 2.1 The Sponsor's P-values for the Primary Endpoints of Three Pivotal Phase III Studies: Studies 31-97-201, 31-97-202 and CN138-001

Study	PANSS Total Score	PANSS Positive Sub-Scale Score	CGI Severity of Illness Score
31-97-201			
Aripiprazole 15 mg vs. Placebo	0.0001	0.0001	0.0001
Aripiprazole 30 mg vs. Placebo	0.0089	0.0005	0.0187
Haloperidol 10 mg vs. Placebo	0.0008	0.0001	0.0019
31-97-202			
Aripiprazole 20 mg vs. Placebo	0.0013	0.0006	0.0298
Aripiprazole 30 mg vs. Placebo	0.0029	0.0177	0.0063
Risperidone 6 mg vs. Placebo	0.0004	0.0002	0.0001
CN138-001			
Aripiprazole 10 mg vs. Placebo	0.0036		
Aripiprazole 15 mg vs. Placebo	0.0002		
Aripiprazole 20 mg vs. Placebo	0.0001		

Table 2.2 The Sponsor's P-values for the Primary Endpoints of Two Supportive Phase II Studies: Study 31-93-202 and 31-94-202

Study	BPRS Total Score	Responders (CGI Severity)	PANSS-Derived BPRS-Core Score	CGI-Improvement Score
31-93-202				
Aripiprazole 5-30 mg vs. Placebo	0.0142	0.035		
Haloperidol 5-20 mg vs. Placebo	0.0083	0.003		
31-94-202				
Aripiprazole 2 mg vs. Placebo			0.7034	0.5860
Aripiprazole 10 mg vs. Placebo			0.8939	0.2260
Aripiprazole 30 mg vs. Placebo			0.1165	0.0055
Haloperidol 10 mg vs. Placebo			0.0495	0.0811

Table 2.3 The Sponsor's P-value for the Primary Endpoints of the Long Term Studies: Studies 31-98-217 and 31-98-304-01

Study	P-value of the logrank test for time to failure to maintain response
31-98-217 and 31-98-304-01	0.427

3. Description of the Sponsor's Studies and Statistical Methodologies

3.1 Pivotal Phase III Studies

3.1.1 Study 31-97-201

This study was titled as 'A Phase III, Double-Blind, Placebo-Controlled Study of Aripiprazole in the Treatment of Psychosis' and was conducted at 36 study centers in the United States of America.

3.1.1.1 Study Objectives

The objectives of this study were to compare the safety and efficacy of each of two doses of aripiprazole (15 mg and 30 mg) versus placebo for the treatment of acute psychosis (in schizophrenia or schizoaffective disorder), and to evaluate the efficacy of aripiprazole on the negative symptoms of psychosis and the relationship of aripiprazole doses with time to response.

3.1.1.2 Study Design

This study was a multicenter, 4-week, randomized, double-blind, parallel-group comparison of the safety and efficacy of aripiprazole, haloperidol, and placebo. The active control, haloperidol was included to confirm the validity of the trial. Approximately 400 patients who were in acute relapse with a diagnosis of schizophrenia or schizoaffective disorder, and who had previously responded to neuroleptics were to be enrolled in the study. After a minimum 5-day placebo washout period, each eligible patient was randomized to one of four double-blind treatment groups: aripiprazole 15 mg, aripiprazole 30 mg, haloperidol 10 mg, or placebo. Study medication was administered orally once daily for 4 weeks. Doses of study medication were not modified during the study. Patients who could not tolerate study drug were withdrawn from the study. Every effort was made to keep patients in the study for at least 2 weeks after randomization. Symptoms were assessed before and during double-blind treatment to evaluate clinical response. Blood samples were collected on specified study days for the determination of plasma concentrations of aripiprazole.

3.1.1.3 Efficacy Variables

The Positive and Negative Syndrome Scale (PANSS) consisted of three sub-scales. The severity of each symptom on these sub-scales was rated on a 7-point scale. The symptom constructs for each sub-scale were as follows:

- Positive Sub-Scale (7 positive symptom constructs: delusions, conceptual disorganization, hallucinatory behavior, excitement, grandiosity, suspiciousness/persecution, and hostility);
- Negative Sub-Scale (7 negative symptom constructs: blunted affect, emotional withdrawal, poor rapport, passive pathetic withdrawal, difficulty in abstract thinking, lack of spontaneity and flow of conversation, and stereotyped thinking); and
- General Psychopathology Sub-Scale (16 symptom constructs: somatic concern, anxiety, guilt feelings, tension, mannerism and posturing, depression, motor retardation, uncooperative, unusual thought content, disorientation, poor attention, lack of judgement and insight, disturbance of volition, poor impulse control, preoccupation, and active social avoidance).

The Clinical Global Impression (CGI) consisted of two 7-point sub-scales: Severity of Illness Scale and Global Improvement Scale. The PANSS and CGI scales were to be administered by the same rater for a given patient throughout the study.

Primary measures of efficacy were:

- 1) change from baseline at Week 4 in PANSS Total Score;
- 2) change from baseline at Week 4 in PANSS Positive Sub-Scale Score; and
- 3) change from baseline at Week 4 in CGI Severity of Illness Score.

Secondary measures of efficacy were:

- 1) change from baseline at Week 4 in PANSS Negative Sub-Scale Score
- 2) time to response to therapy. A response was defined as
 - a) a $\geq 30\%$ decrease from baseline in the PANSS Total Score, or
 - b) a score of 1 (very much improved) or 2 (much improved) on the CGI Improvement Scale.
- 3) time to discontinuation due to lack of efficacy.

Non-protocol specified efficacy measures were:

- 1) number and percentage of responders (patients having a response as defined above);
- 2) mean CGI Improvement Score;
- 3) change from baseline in the PANSS-Derived Brief Psychiatric Rating Scale (BPRS) Core Score.

3.1.1.4 Statistical Methods

3.1.1.4.1 Sample Size and Power

The estimation of sample size was based on data obtained from aripiprazole Phase II studies. The planned sample size of 100 patients per treatment group yielded more than 90% power to detect a treatment effect of 12 points in the PANSS Total Score at two-tailed significance level of 0.05 (Last Observation Carried Forward [LOCF] analysis with an estimated standard deviation of 23 points for the change from baseline to last visit). Treatment effect was defined as the mean change from baseline to last visit in an aripiprazole group minus mean change from baseline to endpoint in the placebo group.

3.1.1.4.2 Data Set Descriptions

For purposes of analysis, the following samples were defined. The randomized sample comprised all patients who were randomized to treatment. The safety sample comprised all patients in the randomized sample who took at least one dose of study medication, as indicated on the dosing record. The efficacy sample comprised all patients who had at least one post-randomization efficacy evaluation.

The LOCF data set included data recorded at a given visit or, if no observation was recorded at that visit, data carried forward from the previous visit. To perform an efficacy analysis at Week 4, the primary time point of interest, the last observed value of patients who dropped out of the study before Week 4 was carried forward to Week 4. Baseline data

were not carried forward or averaged with post-treatment data to impute missing values for the LOCF data set. The Observed Cases (OC) data set consisted of the actual observations at each visit.

The randomized sample was used for baseline summaries of demographics, medical history, and psychiatric and previous treatment history. The safety sample was used for the summarization of safety data, concomitant medication, and extent of exposure. All efficacy analyses were performed on the efficacy sample at baseline (except CGI improvement score), at endpoint, and at each specified study week. Efficacy analyses were performed using both the LOCF and OC data sets. The LOCF data set was the primary data set. The analyses of the OC data set were considered secondary and were performed to corroborate those on the LOCF data set.

3.1.1.4.3 Small Centers

For the purpose of efficacy analyses, a small center in this study was defined as a center with no patients in one or more treatment groups. Since LOCF efficacy analyses were adjusted for study center, small centers were pooled to form pseudo-centers so that each treatment group included at least one patient within the center. Pooling was done based on the primary efficacy variable (PANSS Total Score) at Week 4 using the following algorithm:

Based on the number of patients who were eligible for an analysis, small centers were ordered from the largest to the smallest. The pooling process started with the largest of the small centers; i.e., first the largest center was pooled with the smaller centers starting with the smallest until a non-small center was formed. The process was repeated using the centers left out after the first pass. In case of ties in center size, the center with the smallest center code was selected. (For example, between the tied centers 012 and 032, center 012 was selected.) If any centers were left out at the end of this process, they were pooled with the smallest pseudo-center.

Of the 36 centers, 6 centers (numbered 001, 011, 016, 019, 022, 035) were identified as small centers. These centers were pooled to form two pseudo-centers 901 and 902 as follows: 901 = Centers 016 and 019 pooled and 902 = Centers 001, 011, 022 and 035 pooled. These pseudo-centers were used for all the LOCF efficacy analyses when the model was adjusted for baseline values and study center.

3.1.1.4.4 Efficacy Analyses

Primary efficacy measures were the mean change from baseline to Week 4 in the PANSS Total Score, the mean change from baseline to Week 4 in the PANSS Positive Sub-Scale Score, and change from baseline to Week 4 in the CGI Severity of Illness Score. These primary efficacy measures were evaluated by analysis of covariance (ANCOVA) adjusting for baseline values and study center. The treatment-by-center interaction was assessed at endpoint by a secondary analysis of the above model including the treatment-by-center interaction. The check of treatment-by-center interaction was tested at 0.10 level for the

homogeneity of the treatment effect across the centers. The primary endpoint was the Week 4 LOCF analysis.

The primary comparisons of interest were aripiprazole 30 mg versus placebo and aripiprazole 15 mg versus placebo. The treatment comparisons were tested by following the step-down procedure, i.e., first aripiprazole 30 mg versus placebo was tested at two-tailed 0.05 level; if rejected, aripiprazole 15 mg versus placebo was tested at two-tailed 0.05 level.

The unadjusted means of change from baseline in the PANSS Total Score were analyzed by a one-way analysis of variance (ANOVA) and are provided in the supplemental tables of the sponsor's study report. Subgroup analyses were performed by gender and study center. In the report, descriptive statistics are provided for subgroup analyses by gender and study center. Due to inadequate enrollment of adolescent and elderly patients in this study, a by-age analysis was not performed. The ANCOVA model for the gender subgroup analysis included only the baseline value and treatment group.

The dropout cohort analysis was performed to assess effects of dropouts by plotting the change of PANSS Total Score by treatment group using different dropout cohorts. Dropout cohorts were formed by patients that had their last primary efficacy measurement in the same week interval.

Additional longitudinal analyses were performed on the PANSS Total, PANSS Positive Sub-Scale, PANSS Negative Sub-Scale, and CGI Severity of illness Scores. These analyses employed three methods: (1) the method of Wu and Bailey (1989) (2) unweighted least squares, and (3) random effects model (Laird and Ware, 1982). The results from these analyses include estimated treatment effects versus placebo, P-values, and 95% confidence intervals.

Other continuous variables, such as the change from baseline to last observation in the PANSS Negative Sub-Scale Score and PANSS-Derived BPRS Core Score, were analyzed following similar methods as those for the primary efficacy measures except that no adjustment in significance level was made to account for multiple comparisons. Categorical data, such as CGI Improvement and the percentage of responders, were evaluated by the Cochran-Mantel-Haenszel (CMH) method with stratification by center. Analyses were performed at all time points for both LOCF and OC data sets.

The time-to-event variables (i.e., time to response and time to discontinuation due to lack of efficacy) were compared between treatment groups by the log-rank test.

All the OC analyses and subset analyses included only treatment and baseline values in the model. Center effect was not adjusted in the OC and subset analyses due to a large number of small centers, and the pooling algorithm was based on the LOCF data set.

3.1.2 Study 31-97-202

This study was titled as 'A Phase III, Double-Blind, Placebo-Controlled Study of Aripiprazole in the Treatment of Psychosis, with Risperidone as Active Control' and was conducted at 40 study centers in the United States of America.

3.1.2.1 Study Objectives

The objectives of this study were to compare the safety and efficacy of 20-mg and 30-mg aripiprazole versus placebo for the treatment of acute psychosis (in schizophrenia and schizoaffective disorders). In addition, information was gathered on the efficacy of aripiprazole on the negative symptoms of psychosis and the relationship of aripiprazole doses with time to response.

3.1.2.2 Study Design

This study was a multicenter, 4-week, randomized, double-blind, parallel-group comparison of the safety and efficacy of aripiprazole, risperidone, and placebo. Approximately 400 patients who were in acute relapse with a diagnosis of schizophrenia or schizoaffective disorder, and who had previously responded to neuroleptics were to be randomized in the study. After a minimum 5-day placebo washout period, each eligible patient was randomized to one of four double-blind treatment groups: aripiprazole 20mg, aripiprazole 30 mg, risperidone 6 mg, or placebo. Study medication was administered orally twice daily for 4 weeks. Doses of study medication were not modified during the study except that risperidone was titrated upward for the first 3 days of study participation. Patients who could not tolerate study drug were withdrawn from the study. Every effort was made to keep patients in the study for at least 2 weeks after randomization. Symptoms were assessed before and during double-blind treatment to evaluate clinical response. Blood samples were collected on specified study days for the determination of plasma concentrations of aripiprazole.

3.1.2.3 Efficacy Variables

Same as Study 31-97-201 in Section 3.1.1.3.

3.1.2.4 Statistical Methods

3.1.2.4.1 Sample Size and Power

Same as Study 31-97-201 in Section 3.1.1.4.1.

3.1.2.4.2 Data Set Descriptions

Same as Study 31-97-201 in Section 3.1.1.4.2.

3.1.2.4.3 Small Centers

The definition of small center was the same as what was defined in Section 3.1.1.4.3 for Study 31-97-201. The pooling algorithm for small centers was also the same as what was mentioned in the section. However, of the 40 centers, 7 centers (numbered 052, 055, 063, 079, 082, 083, 094) were identified as small centers. These centers were pooled to form three pseudo-centers 901, 902 and 903 as follows: centers 052, 055 and 094 form center 901; centers 063 and 082 form center 902; centers 079 and 083 form center 903. These pseudo-centers were used for all the LOCF efficacy analyses when the model was adjusted for baseline values and study center.

3.1.2.4.4 Efficacy Analyses

Same as Study 31-97-201 in Section 3.1.1.4.4.

3.1.3 Study CN 138-001

This study was titled as 'A Multicenter, Randomized, Double-Blind, Placebo-Controlled Study of Three Fixed Doses of Aripiprazole in the Treatment of Patients with Acute Schizophrenia'. It was conducted at total 57 centers in the United State and Canada (4 centers in Canada).

3.1.3.1 Study Objectives

Primary Objective: This study compared the efficacy of three fixed doses of aripiprazole with placebo in the treatment of acutely relapsed patients with a diagnosis of schizophrenia.

Secondary Objective: This study compared the safety of three fixed doses of aripiprazole with placebo in the treatment of acutely relapsed patients with a diagnosis of schizophrenia.

3.1.3.2 Study Design

This study was a double-blind, placebo-controlled, randomized, multicenter trial with four parallel groups of inpatients (placebo, aripiprazole 10 mg, aripiprazole 15 mg, and aripiprazole 20 mg). The patients in this trial met the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV) criteria for schizophrenia and were in acute relapse. After a minimum 2-day neuroleptic medication washout, patients fulfilling entry criteria were randomized into the 6-week Acute Phase. Patients received blinded, oral fixed doses of 10 mg, 15 mg, or 20 mg aripiprazole or placebo, once daily. Patients who were unable to tolerate the study medication were discontinued from the study. Symptoms were assessed before and during double-blind treatment to evaluate clinical response. Patients remained hospitalized for the duration of the 6-week treatment period.

Patients showing no improvement or a worsening of symptoms (i.e., Clinical Global Impression [CGI] Improvement ≥ 4) at the end of Week 3, were offered the option of open-label aripiprazole treatment during Weeks 4, 5 and 6. Treatment with open-label aripiprazole was initiated at 20 mg with the option of decreasing to 15 mg based on tolerability. Patients still not improving by Week 5 were discontinued from the study.

Patients who completed the 6-week Acute Phase (including patients who received open-label aripiprazole) were eligible to enter the long-term, outpatient Extension Phase in which they were randomized to double-blind aripiprazole at a dose range of either 10 mg to 15 mg or 20 mg to 30 mg per day.

3.1.3.3 Efficacy Variables

The primary measure of efficacy was the mean change from baseline to Week 6 (Last Observation Carried Forward [LOCF] data set) in the PANSS Total Score.

Key secondary efficacy measures were: 1) the mean change from baseline to Week 6 (LOCF data set) in the PANSS Negative Sub-Scale Score (with additional analyses at all time points) and 2) the mean change from baseline to Week 6 (LOCF data set) in the PANSS-derived Brief Psychiatric Rating Scale (BPRS) Core Score calculated from the PANSS.

Additional efficacy endpoints were: 1) the mean change from baseline in the PANSS Positive Sub-Scale Score at all time points, 2) the mean CGI Improvement Score at all time points, 3) the mean change from baseline in the CGI Severity of Illness Score at all time points, 4) the rate of discontinuation due to lack of efficacy or entry into the open-label aripiprazole at/after Week 3 with a CGI Improvement Score of 4 to 7, 5) the mean change from baseline to Week 6 in the MADRS, and 6) response rates at all time points. Responders were patients who met either of the following criteria:

- a rating of very much improved or much improved on the CGI Improvement Score; or
- at least a 30% decrease from baseline in the PANSS Total Score at all time points.

An evaluable patient was one who had taken at least one dose of study medication and received at least one post-randomization efficacy evaluation.

3.1.3.4 Statistical Methods

3.1.3.4.1 Sample Size and Power

The primary efficacy outcome measure was the mean change from baseline to Week 6 (LOCF Data Set) on the PANSS Total Score. The planned sample size of 400 evaluable patients (100 per treatment group) provided 90% power to detect a difference of 12 in the change from Baseline to Week 6 in PANSS Total Score between placebo and each of the three fixed doses of aripiprazole. This assumed a standard deviation of 23 and a two-sided

test at the 0.0167 significance level (0.05 significance level adjusted for three comparisons versus placebo).

3.1.3.4.2 Data Set Descriptions

The definitions of randomized sample, safety sample, efficacy sample, the LOCF data set and OC data set were the same as what were described in Section 3.1.1.4.2 for Study 31-97-201.

All efficacy analyses were performed on the Efficacy sample at Baseline (if evaluated at baseline), at endpoint, and at each specified study week. Efficacy analyses were performed using both the LOCF and OC data sets. The analyses of the LOCF data set were considered primary analyses. The analyses of the OC data set were considered secondary and were performed to corroborate those on the LOCF data set.

For the analyses of the double-blind treatment, data for patients that received open-label aripiprazole after Week 3 were handled in the following manner. LOCF data for the patients on open-label aripiprazole reflected their last double-blind treatment evaluation and OC data were considered missing (i.e., open-label Week 4, 5 and 6 results were not used in the double-blind analysis).

3.1.3.4.3 Small Centers

The definition of small center was the same as what it was defined in Section 3.1.1.4.3 for Study 31-97-201.

Of the 57 centers, 16 centers (numbered 9, 20, 23, 24, 30, 38, 40, 43, 47, 51, 54, 56, 66, 72, 75, 76) were identified as small centers. These centers were pooled to form two pseudo-centers P10 and P11 as follows: P10 = Centers 9, 20, 23, 24, 30, 40, 43, 51, 56, 72, and 76 pooled, and P11 = Centers 38, 47, 54, 66, and 75 pooled. These pseudo centers were used for all LOCF efficacy and safety analyses when the model was adjusted for study center.

3.1.3.4.4 Efficacy Analyses

Primary Efficacy Analysis

The primary efficacy variable in this study was the mean change from baseline to Week 6 in the PANSS-Total Score. The primary efficacy measure was evaluated by Analysis of Covariance (ANCOVA). The model included the baseline (randomization) measure as covariate and the study center and treatment as main effects. The primary presentation of results were the model-based estimates and the 95% confidence intervals (CI) for the treatment differences (aripiprazole-placebo), which were derived from the estimation (ESTIMATE) of the treatment contrast. Change Scores were derived by subtracting the baseline Score from the Score at each follow up visit. Baseline data were evaluated by Analysis of Variance (ANOVA) with treatment and study center as main effects.

In order to protect the experiment-wise alpha level at 0.05 level when making three comparisons of aripiprazole fixed doses versus placebo on the primary efficacy analyses, the statistical testing was carried out using Hochberg's sequentially rejective procedure. Superiority to placebo was claimed if all three pairwise comparisons were significant at the 0.05 level, or two out of three were significant at the 0.025 level, or if one out of three was significant at the 0.0167 level.

In addition to the primary analysis, the following were also performed.

The mean change from baseline in PANSS Total Score was evaluated at all time points on both the LOCF and OC data sets.

To evaluate the dose response effect, a linear trend test using the actual doses (the dose in placebo group was assigned to zero) was performed with and without the placebo group at 0.05 level. The LOCF data set was used and the analysis was performed at all time points.

To corroborate the results of the primary analysis, the primary efficacy measure was also analyzed by a Non-Parametric One-Way test (NPAR1WAY), i.e., Wilcoxon test. The LOCF data set was used and the analysis was performed at all time points.

The unadjusted mean changes were analyzed by a one-way ANOVA for the LOCF data set at all time points.

Subgroup analyses were performed by gender. The ANCOVA model included only the baseline value and treatment group in the model. The LOCF data set was used and the analysis was performed at all time points. Treatment effects only were provided (i.e., no P-values) since this study was not powered to detect treatment differences in this subgroup analysis.

An analysis was performed to assess effects of dropouts by plotting the change of the PANSS Total Score by treatment group using different dropout cohorts. Like Studies 31-97-201 and 31-97-202, dropout cohorts were formed by patients that had their last primary efficacy measurement in the same week interval.

Study centers were not included in any analyses of the OC data set.

Key Secondary Analyses

The key secondary efficacy measures were the mean change from baseline to Week 6 in the PANSS-derived BPRS Core Score and the mean change from baseline to Week 6 in the PANSS Negative Subscale Score in the LOCF data set. A hierarchical testing procedure was used in testing the key confirmatory analyses so that the overall experiment-wise Type I error rate was 0.05, and Hochberg's sequentially rejective procedure was applied. Testing proceeded sequentially. First, the PANSS-derived BPRS Core Score was tested for those treatment groups significantly different versus placebo from the primary analysis. Only those treatment groups for which the PANSS-derived

BPRS Core Score were significantly different versus placebo were tested for the PANSS Negative Sub-Scale Score. The outcome of the tests for the key secondary endpoints did not affect the statistical significance achieved for the primary endpoint. These measures were analyzed by ANCOVA, for Week 6 LOCF data set, only for the appropriate treatment groups.

Other Secondary Analyses

In addition to the key secondary analyses, the following other secondary analyses of the key secondary variables were performed. The mean change from baseline in PANSS Negative Sub-Scale Score and PANSS-derived BPRS Core Score were evaluated at all time points on both the LOCF and OC data sets.

Other secondary efficacy variables, such as the mean change from baseline in the PANSS Positive Sub-Scale Score, mean change from baseline in the MADRS Total Score, and mean change from baseline in the CGI Severity of Illness Score were analyzed following similar methods as those for the primary efficacy variable. The model for the LOCF analysis of the MADRS did not include study center. Categorical data, such as mean CGI Improvement Score, were analyzed within the framework of the generalized CMH procedure, controlling for study center. Analyses on the other secondary efficacy measures were performed at the 5% significance level without adjustment for multiple comparisons.

The time-to-event variable (i.e., time to discontinuation) was evaluated by survival analysis. The survivorship function and estimated survivorship curves were obtained from Kaplan-Meier maximum likelihood estimates. The log rank test was used to compare survival distributions.

3.2 Phase II Studies

3.2.1 Study 31-93-202

The study was titled as “Efficacy and Tolerability of Ascending Doses of OPC-14597 Compared to Placebo and to Haloperidol in Acutely Relapsing Hospitalized Schizophrenic Patients”. There were 10 sites in the United States involved in the study.

3.2.1.1 Study Objectives.

The primary objective of this study was to assess the efficacy of OPC-14597 (aripiprazole) for the treatment of acute schizophrenia and the tolerability of the effective doses.

The secondary objectives of this study were:

- 1) to evaluate the effective dose range of OPC-14597;
- 2) to evaluate whether OPC-14597 was more effective on positive or negative symptoms of the disease;

- 3) to evaluate the pharmacokinetic characteristics of OPC-14597 in schizophrenic patients; and -
- 4) to compare the effects of OPC-14597 to those of haloperidol on serum prolactin concentration in schizophrenic patients.

3.2.1.2 Efficacy Outcome Variables

The primary efficacy variables were (1) change from baseline to last visit in BPRS-total score, and (2) a response indicator variable (with values 'improved' or 'not improved') defined as follows. Patient disease status was categorized as 'improved' if a reduction of at least one point from baseline to last visit in CGI-severity score was recorded; otherwise, the patient condition was categorized as 'not improved'.

The secondary efficacy measurement was based on the score from the Positive and Negative Syndrome Scale (PANSS).

3.2.1.3 Study Design

This was a Phase II, 4-week, multicenter, double-blind, randomized, placebo-controlled, parallel-group, inpatient study of the efficacy and tolerability of ascending doses of OPC-14597 in acutely relapsing schizophrenic patients with a history of responding to antipsychotic drugs.

Following a 3-7 day placebo washout period, patients were randomized to either ascending doses of OPC-14597, ascending doses of haloperidol, or placebo. Patients were evaluated for efficacy and tolerability at the end of each treatment week (± 2 days).

According to the original protocol, the dose of OPC-14597 was to be titrated from 5 mg to 30 mg per day up to Day 13 of the study and, provided tolerability was satisfactory, the 30 mg/day dose was to be maintained for the remaining 15 days of the study. The dose of haloperidol was to be titrated from 5 mg to 20 mg per day up to Day 10 of the study and, provided tolerability was satisfactory, the 20 mg/day dose was to be maintained for the remaining 18 days of the study. Each ascending dose of OPC-14597 or haloperidol was to be given for 3 days. The original protocol was first amended (Amendment 001) to limit the dose of OPC-14597 to a maximum of 20 mg/day. The protocol was amended a second time (Amendment 002) to increase the maximum dose of OPC-14597 back to 30 mg/day, as per the original protocol. To reach therapeutic levels faster, the protocol was further amended (Amendment 003) to decrease the dosing period from 3 days to 2 days for ascending doses of OPC-14597 or haloperidol. Through this titration schedule, a 20 mg/day dose was achieved in the OPC-14597 group on Day 7 and the maximum doses of OPC-14597 (30 mg) and haloperidol (20 mg) were achieved on Day 13 and Day 7 respectively.

3.2.1.4 Statistical Methods

Sample Size

The sample size for this study was calculated based on expected changes in mean BPRS-total score and on an expected 30% dropout rate. Haloperidol and OPC-14597 were expected to induce a 30% decrease in mean BPRS-total score. A 10-15% decrease in mean BPRS-total score was expected in the placebo group. Based on the above assumptions and using crude estimates of variability from the literature, it was determined that 25 patients in each of the two groups would provide greater than 80% power.

Baseline Comparisons

Demographic and baseline psychiatric comparisons were based on information obtained at the screening visit prior to washout for all randomized patients. Mean, minimum and maximum by sex were used to describe continuous variables such as age and weight. Frequency distributions of each treatment group by sex were tabulated for race. Baseline psychiatric characteristics, including age at first hospitalization for psychiatric illness, number of times hospitalized in the past, length of present schizophrenic episode, onset of current condition, and categorization of subchronic/chronic schizophrenia, were tabulated by treatment group for all randomized patients for purposes of comparison.

Population Analyzed

The primary and secondary efficacy analyses were performed on the intent-to-treat (ITT) population at Week 4 by the last observation carried forward (LOCF) method and also for the observed cases at each week.

Small Centers

For the purpose of efficacy analyses, a small center in this study was defined as a center with no patients in one or more treatment groups. Only one center (center 12) was identified as a small center. This center was pooled with center 007, which had the lowest total number of patients.

Analyses for Primary Efficacy Variables

For BPRS-total score, both last visit and by-week observed cases analyses were performed. For the last visit analysis, the change from baseline in BPRS-total score was analyzed by fitting a linear model with terms for treatment, center, center-by-treatment interaction, and baseline value as covariate for the by-week observed cases analyses, the model included only treatment and baseline as covariate except that at week 0 (baseline) only treatment and center were included in the model. The p-value for the primary comparison of OPC-14597 vs. placebo was obtained based on this model (Type III analyses were utilized). For the variable of response indicator, the responder rates of OPC-14597 vs. placebo were compared by the Cochran-Mantel-Haenszel (CMH) test stratified

by center. Results were declared statistically significant relative to a two-tailed nominal significance level of 0.05. Similar methods were also applied for the comparison of haloperidol vs. placebo.

Analyses for Secondary Efficacy Variables

Analyses of the secondary variables were performed in a parallel fashion as in the case of the primary efficacy variable BPRS-total score.

3.2.2 Study 31-94-202

This study was titled as a dose ranging study of the efficacy and tolerability of OPC-14597 in acutely relapsing hospitalized schizophrenic patients. It was a multi-center study with 23 US sites participated.

3.2.2.1 Study Objectives

The primary objective of this study was to determine an optimal dose of OPC-14597 (aripiprazole) for the treatment of acute schizophrenia.

The secondary objectives of this study were: (1) preliminary comparison of the efficacy of OPC-14597 to that of haloperidol on negative symptoms, and (2) comparison of the effects of OPC-14597 to those of haloperidol on serum prolactin levels.

3.2.2.2 Efficacy Outcome Variables

The primary efficacy variables were (1) Change from baseline to last visit in the Psychotic Items Sub-scale of the Brief Psychiatric Rating Scale (BPRS) and (2) Clinical Global Impression (CGI) improvement score at last visit.

The secondary outcome variables will be based on the Positive and Negative Symptom Scale (PANSS) and on the total BPRS scores.

3.2.2.3 Study Design

This was a multi-center, 4-week, double-blind, randomized, placebo-controlled, dose-ranging, parallel-group study of the efficacy and tolerability of three doses of OPC-14597 in chronic schizophrenic patients who have a history of responding to antipsychotics and who present with an acute relapse. OPC-14597 was given in three doses: 2 mg/day (starting with 1 mg on Day 1; followed by 2 mg/day for the rest of the study); 10 mg/day (starting with 5 mg on Day 1; followed by 10 mg/day for the rest of the study); and 30 mg/day (starting with 15 mg on Day 1; followed by 30 mg/day for the rest of the study). Patients were hospitalized throughout the study.

Upon inclusion, patients were submitted to a 3- to 7- day placebo washout period. Every effort was made to washout patients for at least 5 days. Following washout, qualifying

patients were randomized to either one of three fixed doses of OC-14597 (2, 10 or 20 mg/day), a fixed dose of haloperidol (10 mg/day), or placebo. During the double-blind treatment period, patients were evaluated for efficacy and tolerability at the end of each treatment week (± 2 days).

3.2.2.4 Statistical Methods

Sample Size

It was determined initially that 50 patients per treatment group (a total of 250 patients) would provide greater than 80% power to detect a difference of 9 points in mean change scores in BPRS-total between an OPC-14597 group and placebo. The protocol was amended to increase the sample size by 10 patients per group, yielding 60 patients per group (a total of 300 patients) to account for multiple comparisons with placebo.

Baseline Comparisons

Demographic and baseline psychiatric comparisons were based on information obtained at the screening visit prior to washout for all randomized patients. Mean, minimum and maximum by sex were used to describe continuous variables such as age and weight. Frequency distributions of each treatment group by sex were tabulated for race. Baseline psychiatric characteristics, including age at first hospitalization for schizophrenia, number of times hospitalized for schizophrenia in the past, length of present schizophrenic episode, onset of current condition, and categorization of sub-chronic/chronic schizophrenia, were tabulated by treatment group for all randomized patients for purposes of comparison.

Population Analyzed

The primary and secondary efficacy analyses were performed on the intent-to-treat (ITT) population by the last observation carried forward (LOCF) method and for observed cases at each visit.

Small Centers

For the purpose of efficacy analyses, a small center in this study was defined as a center with no patients in one or more treatment groups. Only one center (center 014) was identified as a small center. This center was pooled with center 002, which had the lowest total number of patients.

Primary Efficacy Analyses

For BPRS-core score, the primary analysis at last visit was performed by fitting a linear model to the change score with right hand terms for treatment, center, and the baseline value. For CGI-improvement score at last visit, the model included only terms for treatment and center. Each of the treatment contrasts (i.e., OPC-2 mg vs. placebo, OPC-10 mg vs. placebo, and OPC-30 mg vs. placebo) was estimated by use of Least Squares

Means from a type III analysis, and the p-values were derived from Student's test with appropriate degrees of freedom. Dunnett's method was used for reporting statistically significant (at two-tailed 0.05 level) results corrected for multiple comparisons of the three OPC-14597 groups with placebo.

Secondary Efficacy Analyses

By visit analyses (at Weeks 1, 2, 3 and 4) were performed for those secondary endpoints of changes of scores following parallel methods as described in primary efficacy analyses but no correction to the significance level was made for multiple comparisons. The response to treatment (responder rates) of OPC-14597 dose groups vs. placebo and haloperidol vs. placebo were made by the Cochran-Mantel-Haenszel test stratified by center. The time to discontinuation due to lack of clinical response or marked deterioration in clinical status was plotted by Kaplan-Meier curves and differences in survival between a treatment group and placebo were tested by the log-rank test. Data for the efficacy index in the CGI scale at last visit were summarized by treatment group.

3.3 Long-Term Studies : Studies 31-98-217 and 31-98-304-01

These 52-week, double-blind, haloperidol-controlled long-term studies were nearly identical in design. Study 31-98-217 was conducted in the USA (33 centers) and 31-98-304-01 was a multinational study (137 centers).

3.3.1 Objectives

The primary objective of both studies was to evaluate the long-term maintenance of the acute anti-psychotic effect of aripiprazole, compared with haloperidol, when administered for 52 weeks in patients whose treatment started during an acute relapse of chronic schizophrenia.

The secondary objectives of these studies were to evaluate:

- The efficacy of aripiprazole, compared with haloperidol, in the treatment of patients experiencing an acute relapse of chronic schizophrenia over an 8-week treatment period;
- The efficacy of aripiprazole, compared with haloperidol, for the treatment of positive and negative symptoms of schizophrenia over a 52-week treatment period;
- The safety and tolerability of aripiprazole, compared with haloperidol, in short- and long-term treatment of patients with schizophrenia.

3.3.2 Methodology

For these two studies, the protocol-specified intention was to pool their data for efficacy and safety evaluations. Therefore, these studies are treated throughout this document as though they were a single study except when there were specific differences between them. After a 5-day placebo washout, patients were to be randomized to receive either aripiprazole 30 mg or haloperidol 10 mg, administered orally once daily. Randomization

was done in a 2:1 ratio of aripiprazole to haloperidol. Patients randomized to receive aripiprazole took the 30 mg dose from Day 1 onward while those randomized to haloperidol were to take a 5-mg dose for Days 1 to 3 and the 10-mg dose from Day 4 onward. After the first week of treatment, a one-time dose decrease was allowed if needed for tolerance. Patients who could not tolerate study drug were withdrawn from the study. After randomization all patients were followed for 52 weeks or until early discontinuation.

3.3.3 Statistical Methods

The primary efficacy variable in this study was the “time to failure to maintain response” in responders. Response was defined as a $\geq 20\%$ decrease from baseline in PANSS Total Score and, at the same visit, the patient did not meet any of the following criteria: 1) a CGI Improvement Score of 6 (much worse) or 7 (very much worse); or 2) an adverse event of worsening schizophrenia, or 3) a score of 5 (moderately severe), 6 (severe), or 7 (extreme) in at least one of the four items of the psychotic sub-scale of the PANSS.

Failure to maintain response was defined as (1) a CGI Improvement Score of 6 or 7 in two consecutive evaluations 3 to 5 days apart, or (2) adverse event of worsening schizophrenia, or (3) a score of 5, 6, or 7 in at least one of the four items that constitute the psychotic items sub-scale of PANSS in two consecutive evaluations 3 to 5 days apart. Of the two evaluations, the time-point of the first evaluation was used for determination of failure to maintain response. For patients who had missing data in the second follow-up evaluation to confirm failure to maintain response, the Last Observation Carried Forward (LOCF) imputation method was used and these patients were considered to have failed. The time origin for this time to event measure was the date of randomization. Responders who discontinued from a study without meeting the failure criteria or who completed the study and did not meet the failure criteria at their last visit were treated as censored at the date of discontinuation.

The time to failure to maintain response data was analyzed by fitting the Cox proportional hazard regression model with baseline PANSS Total Score as a covariate and protocol (31-98-217 and 31-98-304-01) as a stratification factor. The null hypothesis of equal hazard rates (i.e., hazard ratio = 1) between the two treatment groups was tested at 0.05 level (two-tailed) and a 95% confidence interval (CI) for the hazard ratio was reported. This analysis was performed only on patients who were considered responders.

Secondary efficacy variables were: 1) change from baseline in PANSS Total Score, 2) change from baseline in PANSS Positive Sub-Scale Score, 3) change from baseline in PANSS Negative Sub-Scale Score, 4) change from baseline in CGI Severity of Illness Score, 5) CGI Improvement Score as recorded, 6) change from baseline in MADRS Total Score, 7) Time to first response, 8) Time from first response to failure to maintain response, 9) Time to discontinuation due to lack of response to study drug, 10) Time to discontinuation due to lack of response to study drug or adverse event.

Of the above, variables (1) through (6) were used to compare efficacy of the acute phase treatment between the two treatment groups at Week 8. Each of the variables (1) through

(4) and variable (6) were summarized by treatment group at each scheduled visit through Week 52. Formal statistical comparisons, by ANCOVA with baseline value as covariate and protocol as classification factor were made at Weeks 8, 26, and 52 for each of these variables. For variable (5), the Cochran-Mantel-Haenszel Mean Score test stratified by protocol was used for treatment group comparison. Variables (7) through (10) were analyzed by plotting Kaplan-Meier curves and by the Cox proportional hazard model with baseline value of PANSS total as covariate and protocol as a stratification factor. Variable (8) was analyzed by similar methods except that the event time was measured from the date of first response.

4. Detailed Review of the Sponsor's Individual Study Results

4.1 Pivotal Phase III Studies

4.1.1 Study 31-97-201

4.1.1.1 Disposition of Patients

A total of 502 patients signed the informed consent form; 42 of these patients failed screening and did not enter the placebo-washout phase. The remaining 460 patients underwent placebo washout; 46 of these patients discontinued from the study prior to randomization.

Four hundred and fourteen patients were randomized to receive double-blind treatment; 106 to the placebo group, 104 to the haloperidol group, 102 to the aripiprazole 15-mg group, and 102 to the aripiprazole 30-mg group. Of these, 248 (60%) patients completed 4 weeks of treatment and 166 (40%) discontinued from the study early. The disposition of all enrolled patients and the time to discontinuation due to all reasons for the Randomization Sample are presented by treatment group in Table 1 and Figure 1 of the Appendix.

Two hundred eighty-two patients who had a DSM-IV diagnosis of schizophrenia were randomized to receive double-blind treatment; 75 to the placebo group, 61 to the haloperidol group, 74 to the aripiprazole 15-mg group, and 72 to the aripiprazole 30-mg group. Of these, 180 (64%) patients completed 4 weeks of treatment and 102 (36%) discontinued from the study early. The disposition of patients with schizophrenia only who were randomized to treatment is presented by treatment group in Table 2 of the Appendix.

4.1.1.2 Data Set

The number of patients within each patient sample was presented by treatment group for all randomized patients in Table 4.1.1.1 as well as for patients with schizophrenia in Table 4.1.1.2.

Table 4.1.1.1 Number of Patients in Different Samples for Study 31-97-201

Sample	Placebo	Haloperidol 10 mg	Aripiprazole		Total
			15 mg	30 mg	
Randomized	106	104	102	102	414
Safety	104	103	102	101	410
Efficacy	103	99	99	100	401

Table 4.1.1.2 Number of Patients with Schizophrenia in Different Samples for Study 31-97-201

Sample	Placebo	Haloperidol 10 mg	Aripiprazole		Total
			15 mg	30 mg	
Randomized	75	61	74	72	282
Safety	74	60	74	71	279
Efficacy	74	59	72	71	276

Four of the 414 randomized patients were excluded from the safety sample because they did not receive study medication according to the dosing record. Moreover, thirteen of the 414 randomized patients were excluded from the efficacy sample because they did not have a post-randomization efficacy evaluation.

4.1.1.3 Demography and Patient Characteristics

Demographic characteristics are presented by treatment group in Table 4.1.1.3 and Table 4.1.1.4 for all randomized patients and for those with schizophrenia in the randomized sample, respectively. The treatment groups were comparable with respect to age, sex, race, and weight.

Table 4.1.1.3 Demographic Characteristics for All Patients in the Randomized Sample for Study 31-97-201

Variable		Placebo N=106	Haloperidol 10 mg N = 104	Aripiprazole		Total N = 414
				15 mg N = 102	30 mg N=102	
Age (yrs)	Mean	38.5	38.9	37.8	39.3	38.6
	Median	38.5	40.0	36.5	40.0	39.0
	Min-Max	19.0-68.0	18.0-59.0	19.0-61.0	19.0-65.0	18.0-68.0
	S.E.	0.9	0.9	1.0	1.0	0.5
Gender	Men	74 (70)	68 (65)	76 (75)	70 (69)	288 (70)
	Women	32 (30)	36 (35)	26 (25)	32 (31)	126 (30)
Race	White	54 (51)	68 (67)	61 (60)	59 (59)	242 (59)
	Black	34 (32)	23 (23)	26 (25)	26 (26)	109 (27)
N (%)	Hispanic	14 (13)	9 (9)	12 (12)	12 (12)	47 (12)
	Asian/Pacific Islander	3 (3)	1 (1)	3 (3)	3 (3)	10 (2)
	Other	1	3	0	2	6
	Weight(kg)	Mean	83.3	84.8	85.3	87.8
Weight(kg)	Median	80.8	81.7	82.4	84.9	82.2
	Min-Max	48.6-204.3	44.4-136.7	48.8-169.8	50.4-202.0	44.4-204.3
	S.E.	2.1	2.0	2.2	2.4	1.1
	Missing	1	3	0	1	5

Table 4.1.1.4 Demographic Characteristics for Patients with Schizophrenia in the Randomized Sample for Study 31-97-201

Variable		Placebo N=75	Haloperidol 10 mg N = 61	Aripiprazole		Total N = 282
				15 mg N = 74	30 mg N=72	
Age (yrs)	Mean	39.2	39.0	37.9	39.8	39.0
	Median	39.0	40.0	37.5	40.0	39.5
	Min-Max	19.0-68.0	23.0-58.0	22.0-61.0	19.0-65.0	19.0-68.0
	S.E.	1.1	1.1	1.1	1.2	0.6
Gender N (%)	Men	56 (75)	43 (70)	59 (80)	51 (71)	209 (74)
	Women	19 (25)	18 (30)	15 (20)	21 (29)	73 (26)
Race N (%)	White	36 (48)	39 (66)	37 (50)	41 (59)	153 (55)
	Black	26 (35)	15 (25)	25 (34)	17 (24)	83 (30)
	Hispanic	12 (16)	5 (8)	9 (12)	10 (14)	36 (13)
	Asian/Pacific Islander	1 (1)	0	3 (4)	2 (3)	6 (2)
	Other	0	2	0	2	4
Weight(kg)	Mean	82.8	83.9	83.4	87.9	84.5
	Median	81.0	80.8	78.5	82.3	81.3
	Min-Max	48.6-204.3	44.4-125.3	52.9-146.6	53.6-202.0	44.4-204.3
	S.E.	2.7	2.6	2.3	3.0	1.3
	Missing	0	1	0	0	1

4.1.1.4 The Sponsor's Efficacy Results

Efficacy analyses were performed using the Efficacy Sample (N=401), which comprised all patients who had baseline and post-randomization efficacy evaluations on at least one of the primary or secondary efficacy variables. In addition, as recommended by European regulatory authorities, efficacy analyses were performed for a subset of patients with schizophrenia (N=276) on the key outcome measures (i.e., PANSS Total Score, PANSS Positive Subscale Score, CGI Severity of Illness Score, PANSS Negative Sub-Scale Score, CGI Improvement Score, percentage of responders, and PANSS-Derived BPRS Core Score) to gather information on the efficacy of aripiprazole in schizophrenia.

4.1.1.4.1 For All Randomized Patients

Primary Efficacy Measures:

Table 4.1.1.5 shows the summary of efficacy analysis results for the three primary endpoints in all schizophrenia and schizoaffective disorder patients. Change in these three primary endpoints were derived by subtracting baseline scores from the score at each study week. This review only reports results of changes from baseline to the endpoint, i.e., week 4. Negative change scores indicate improvement.

According to the sponsor's study report, the analysis of the change in the PANSS Total Score for the LOCF data set showed that patients in the haloperidol group and patients in both aripiprazole groups had significantly greater improvement compared with the placebo group during Weeks 2 through 4. The analysis of the change scores for the OC data set showed that both aripiprazole groups had significantly greater improvement

compared with the placebo group at Weeks 2 and 4, while the haloperidol group improved significantly more than the placebo group at Week 2.

The analysis of the model-based mean change in the PANSS Positive Sub-Scale Score for the LOCF data set showed that both aripiprazole groups had significantly greater improvement compared with the placebo group during Weeks 2 through 4. The haloperidol group showed significantly greater improvement compared with the placebo group during Weeks 1 through 4. Results of the OC analysis showed that both aripiprazole groups had significantly greater improvement compared with the placebo group at Weeks 2 and 4. Significantly greater improvement was seen for the haloperidol groups compared with the placebo group at Weeks 1, 2 and 4.

The analysis of the model-based mean change from baseline in the CGI Severity of Illness Score for the LOCF data set showed that the aripiprazole 15 mg group and the haloperidol group had significantly greater improvement compared with the placebo group during Weeks 1 through 4. The aripiprazole 30mg group showed significantly greater improvement during Weeks 3 and 4. The analysis for the OC data set showed that the aripiprazole 15 mg group had significantly greater improvement compared with the placebo group during Weeks 1 through 4. Significantly greater improvement compared with placebo was seen for the haloperidol group and aripiprazole 30 mg group at Week 2.

Table 4.1.1.5 Efficacy Analysis Results for the Primary Endpoints for Study 31-97-201
For the LOCF Data Set:

Endpoints	N	Baseline	Change from Baseline to Endpoint (i.e., week 4)	Treatment Difference vs. Placebo	95% CI for Difference	P-Value
<i>PANSS Total</i>						
Haloperidol 10 mg	99	99.9	-13.8	-10.8	(-17.2, -4.5)	0.0008
Aripiprazole 15 mg	99	98.8	-15.5	-12.6	(-18.9, -6.3)	0.0001
Aripiprazole 30 mg	100	99.6	-11.4	-8.5	(-14.8, -2.1)	0.0089
Placebo	102	100.9	-2.9			
<i>PANSS Positive Sub-Scale Score</i>						
Haloperidol 10 mg	99	25.0	-4.4	-3.9	(-5.7, -2.0)	0.0001
Aripiprazole 15 mg	99	24.5	-4.2	-3.7	(-5.5, -1.8)	0.0001
Aripiprazole 30 mg	100	24.4	-3.8	-3.3	(-5.1, -1.4)	0.0005
Placebo	103	24.8	-0.6			
<i>CGI Severity of Illness Score --</i>						
Haloperidol 10 mg	99	4.9	-0.5	-0.4	(-0.7, -0.2)	0.0019
Aripiprazole 15 mg	99	4.9	-0.6	-0.6	(-0.8, -0.3)	0.0001
Aripiprazole 30 mg	100	4.8	-0.4	-0.3	(-0.6, -0.1)	0.0187
Placebo	103	5.0	-0.1			

For the OC Data Set:

Endpoints	Baseline & (N)	Change from Baseline to Endpoint & (N)	Treatment Difference vs. Placebo	95% CI for Difference	P-Value
<i>PANSS Total</i>					
Haloperidol 10 mg	99.6 (N=99)	-16.6 (N=61)	-5.2	(-12.4, 2.1)	0.163
Aripiprazole 15 mg	97.9 (N=99)	-24.3 (N=68)	-12.8	(-19.9, -5.8)	<0.001
Aripiprazole 30 mg	98.5 (N=100)	-19.1 (N=61)	-7.7	(-15.0, -0.4)	0.040
Placebo	100.2 (N=102)	-11.4 (n=60)			
<i>PANSS Positive Sub-Scale Score</i>					
Haloperidol 10 mg	25.1 (N=99)	-5.0 (N=61)	-2.5	(-4.6, -0.4)	0.023
Aripiprazole 15 mg	24.6 (N=99)	-6.4 (N=68)	-3.8	(-5.9, -1.8)	<0.001
Aripiprazole 30 mg	24.4 (N=100)	-6.2 (N=61)	-3.7	(-5.8, -1.6)	0.001
Placebo	24.9 (N=103)	-2.6 (N=60)			
<i>CGI Severity of Illness Score</i>					
Haloperidol 10 mg	4.9 (N=99)	-0.6 (N=61)	-0.2	(-0.5, 0.1)	0.147
Aripiprazole 15 mg	4.9 (N=99)	-0.9 (N=68)	-0.5	(-0.8, -0.2)	0.001
Aripiprazole 30 mg	4.8 (N=100)	-0.7 (N=60)	-0.3	(-0.6, 0.0)	0.053
Placebo	4.9 (N=103)	-0.4 (N=60)			

An internal audit revealed that data generated at Study Centers 007 and 011 could not be validated. Therefore, a sensitivity analysis of the mean change from baseline in the PANSS Total Score was performed by excluding the 16 patients randomized at Center 007 and the three patients randomized at Center 011. Results of the sensitivity analysis were consistent with those of the overall analysis.

The trial was not designed to compare treatment effects between the aripiprazole 15mg and 30mg groups. However, the change from baseline in the PANSS and CGI scores for the aripiprazole 15 mg group relative to placebo was quantitatively greater than that for the aripiprazole 30 mg group. Exploratory evaluations showed that the difference seen from baseline to Week 4 between the aripiprazole 15 and 30 mg groups was largely driven by the negative and general PANSS items, especially for those patients that completed the study.

Secondary Efficacy Measures:

The summaries of efficacy analysis results for the protocol specified secondary endpoints for all patients are shown in Table 4.1.1.6. According to the sponsor's study report, the analysis of the model-based mean change in the PANSS Negative Sub-Scale Score for the LOCF data set showed significantly greater improvement compared with the placebo group for the aripiprazole 15-mg group at Weeks 2 and 4. Although treatment differences between the aripiprazole 30-mg and placebo groups did not reach statistical significance, the magnitude of change in the PANSS Negative Sub-Scale Score for the aripiprazole 30-mg group was substantial. The haloperidol group showed significantly greater improvement compared with the placebo group during Weeks 2 and 4. The analysis of the

OC data set showed significantly greater improvement compared with the placebo group at Week 2 for both aripiprazole groups and the haloperidol group. The aripiprazole 15-mg group also showed significantly greater improvement at Week 4 compared with the placebo group.

The time-to-response analysis was performed separately for the CGI Improvement Score (CGI response) and the PANSS Total Score (PANSS response) using survival analysis. Only the aripiprazole 15-mg group showed a significant difference from the placebo group ($p=0.0122$) in the time to CGI response. There were no other significant differences between treatment groups in the time to PANSS response.

Lack of efficacy is defined as insufficient clinical response or withdrawal of consent by the patient due to lack of effect. Results of the analysis of time to discontinuation due to lack of efficacy showed a between-treatment difference overall. This difference was contributed by two pairwise comparisons: aripiprazole 15 mg versus placebo ($p=0.003$) and haloperidol versus placebo ($p=0.009$).

Table 4.1.1.6 Efficacy Analysis Results for the Secondary Efficacy Endpoints for Study 31-97-201

For the LOCF Data Set:

Endpoints	N	Baseline	Change from Baseline to Endpoint (i.e., week 4)	Treatment Difference vs. Placebo	95% CI for Difference	P-Value
<i>PANSS Negative Sub-Scale Score</i>						
Haloperidol 10 mg	99	26.2	-2.9	-1.8	(-3.5, -0.1)	0.043
Aripiprazole 15 mg	99	25.8	-3.6	-2.4	(-4.1, -0.7)	0.006
Aripiprazole 30 mg	100	26.2	-2.3	-1.1	(-2.8, 0.6)	0.213
Placebo	102	26.5	-1.2			

For the OC Data Set:

Endpoints	Baseline & (N)	Change from Baseline to Endpoint & (N)	Treatment Difference vs. Placebo	95% CI for Difference	P-Value
<i>PANSS Negative Sub-Scale Score</i>					
Haloperidol 10 mg	25.7 (N=99)	-3.7 (N=61)	-0.2	(-2.3, 1.9)	0.829
Aripiprazole 15 mg	25.1 (N=99)	-6.3 (N=68)	-2.8	(-4.8, -0.7)	0.008
Aripiprazole 30 mg	25.5 (N=100)	-3.9 (N=61)	-0.4	(-2.5, 1.8)	0.735
Placebo	25.9 (N=102)	-3.5 (N=60)			

Endpoints	P-Value by the Log-Rank Test
<i>Time to Response to Therapy</i> (Defined as a 30% Decrease from Baseline in PANSS Total Score)	
Haloperidol 10 mg vs. Placebo	0.5569
Aripiprazole 15 mg vs. Placebo	0.0925
Aripiprazole 30 mg vs. Placebo	0.2413

Endpoints	P-Value by the Log-Rank Test
<i>Time to Response to Therapy</i> (Defined as a CGI Global Improvement Score of 1 or 2)	
Haloperidol 10 mg vs. Placebo	0.7396
Aripiprazole 15 mg vs. Placebo	0.0122
Aripiprazole 30 mg vs. Placebo	0.1878

Endpoints	P-Value by the Log-Rank Test
<i>Time to Discontinuation due to Lack of Efficacy</i>	
Haloperidol 10 mg vs. Placebo	0.0092
Aripiprazole 15 mg vs. Placebo	0.0031
Aripiprazole 30 mg vs. Placebo	0.2290

Other Efficacy Measures:

Other efficacy measure results are shown in Table 4.1.1.7. A patient who had a CGI Improvement Score of 1 (very much improved) or 2 (much improved), or a $\geq 30\%$ decrease from baseline in the PANSS Total Score was considered a responder. Both the LOCF and OC analyses showed that the aripiprazole 15-mg group had a significantly greater percentage of responders at Week 4. A greater percentage of responders was seen for the aripiprazole 30-mg group compared with the placebo group at Week 4 for the LOCF data set; however, the OC analysis showed no statistically significant differences between these treatment groups. There were no significant differences in the percentage of responders between the haloperidol and placebo groups at any time during the study for either the LOCF or OC data set.

The analysis of the mean CGI Improvement Score for the LOCF data set showed that the aripiprazole 15-mg group had significantly greater improvement compared with the placebo group throughout the 4-week study. Significantly greater improvement compared with placebo was seen at Weeks 3 and 4 for the aripiprazole 30-mg group and at Weeks 2 through 4 for the haloperidol group. In the OC analysis, significantly greater improvement compared with the placebo group was seen at Week 4 for the aripiprazole 30-mg group and at Weeks 1 through 3 for the haloperidol group. Results of the OC analysis for the aripiprazole 15-mg group were consistent with those of the LOCF analysis; significantly greater improvement compared with placebo was evident throughout the study.

The analysis of the model-based mean change in the PANSS-Derived BPRS Core Score for the LOCF data set showed significantly greater improvement compared with the placebo group for the aripiprazole 15-mg group at Weeks 2 through 4 and for the aripiprazole 30-mg group at Weeks 3 and 4. The haloperidol group showed significantly greater improvement compared with the placebo group during Weeks 1 through 4. The analysis of the OC data set showed significantly greater improvement compared with the placebo group at Weeks 1, 2 and 4 for the aripiprazole 15-mg group and haloperidol group, and at Weeks 2 through 4 for the aripiprazole 30-mg group.

Table 4.1.1.7 Efficacy Analysis Results for the Other Efficacy Measures for Study 31-97-201

Endpoints	Number (%) of Responders at Week 4	P-Value (vs. Placebo)
<i>Percentage of Responders*</i>		
Haloperidol 10 mg (N=99)	26 (26)	0.089
Aripiprazole 15 mg (N=99)	35 (35)	0.002
Aripiprazole 30 mg (N=100)	28 (28)	0.050
Placebo (N=103)	17 (17)	

Endpoints	Mean at Week 4	P-Value (vs. Placebo)
<i>CGI Improvement Score*</i>		
Haloperidol 10 mg (N=99)	3.7	0.002
Aripiprazole 15 mg (N=99)	3.5	<0.001
Aripiprazole 30 mg (N=100)	3.8	0.016
Placebo (N=103)	4.3	

Endpoints	N	Baseline	Change from Baseline to Endpoint (i.e., week 4)	Treatment Difference vs. Placebo	95% CI for Difference	P-Value
<i>PANSS-Derived BPRS Core Score*</i>						
Haloperidol 10 mg	99	17.1	-3.5	-2.4	(-3.5, -1.2)	<0.001
Aripiprazole 15 mg	99	16.8	-3.1	-2.0	(-3.1, -0.8)	0.001
Aripiprazole 30 mg	100	16.9	-3.0	-1.9	(-3.1, -0.8)	0.001
Placebo	103	17.0	-1.1			

* The results shown above are for the LOCF data set.

Subgroup Analysis:

The subgroup analysis results by gender and study center for the PANSS Total Score in the LOCF analysis are shown in Table 5 and 6 of the Appendix. In addition to showing some descriptive statistics, the sponsor commented that there was no significant treatment-by-center interaction.

4.1.1.4.2 For Patients with Schizophrenia

Primary Efficacy Measures:

The LOCF analysis results of the model-based mean change in the PANSS Total Score, the PANSS Positive Sub-Scale Score and the CGI Severity of Illness Score for patients with schizophrenia are shown in Table 4.1.1.8. As they are shown in the table, all three groups of patients show significantly greater improvement compared with placebo on all PANSS Total Score, the PANSS Positive Sub-Scale Score and the CGI Severity of Illness Score in the LOCF analysis at the Endpoint, i.e., Week 4. For the OC analyses, the patients with schizophrenia showed a significantly greater improvement PANSS Total Score and CGI Severity of Illness Score compared with placebo at Week 4 in the aripiprazole 15-mg group but not in the aripiprazole 30-mg group and the haloperidol group. For the PANSS Positive Subscale at Week 4, the OC analyses (Table 13 of the Appendix) showed significance on both aripiprazole groups but not in the haloperidol group.

Table 4.1.1.8 Efficacy Analysis Results of the Primary Endpoints for the LOCF Data Set for Patients with Schizophrenia for Study 31-97-201

Endpoints	N	Baseline	Change from Baseline to Endpoint (i.e., week 4)	Treatment Difference vs. Placebo	95% CI for Difference	P-Value
<i>PANSS Total</i>						
Haloperidol 10 mg	59	101.7	-13.8	-12.1	(-19.7, -4.5)	0.002
Aripiprazole 15 mg	72	96.7	-14.6	-12.9	(-20.1, -5.7)	0.001
Aripiprazole 30 mg	71	99.2	-9.9	-8.2	(-15.4, -0.9)	0.027
Placebo	74	100.8	-1.7			
<i>PANSS Positive Sub-Scale Score</i>						
Haloperidol 10 mg	59	25.6	-4.0	-3.8	(-6.0, -1.6)	0.001
Aripiprazole 15 mg	72	24.3	-4.1	-3.9	(-6.1, -1.8)	<0.001
Aripiprazole 30 mg	71	24.6	-3.6	-3.5	(-5.6, -1.3)	0.002
Placebo	74	24.8	-0.2			
<i>CGI Severity of Illness Score</i>						
Haloperidol 10 mg	59	4.9	-0.51	-0.48	(-0.8, -0.2)	0.003
Aripiprazole 15 mg	72	4.9	-0.62	-0.6	(-0.9, -0.3)	<0.001
Aripiprazole 30 mg	71	4.9	-0.35	-0.32	(-0.6, -0.03)	0.032
Placebo	74	5.0	-0.03			

Other Efficacy Measures:

Table 4.1.1.9 shows the other efficacy analysis results for patients with schizophrenia only in the LOCF data set. As we can observe in the table, for the change of the PANSS Negative Sub-Scale Score from the baseline to Week 4, the difference between the aripiprazole-30 mg group and placebo did not reach statistical significance. The other two comparisons between the aripiprazole 15 mg group and the haloperidol group versus

placebo did show significantly greater improvement. The OC analyses performed similarly to the LOCF analyses.

For the percentage of responders, both the LOCF and OC analyses showed a significant results at Week 4 for the aripiprazole 15-mg group. There were no significant differences in the results between the haloperidol and placebo groups as well as aripiprazole 30-mg and placebo group at Week 4 for either the LOCF or OC data set.

For the endpoint of mean CGI Improvement Score and the change of PANSS-Derived BPRS Core Score, all three treatment groups showed significant test results in the LOCF data analyses. For the OC analyses, the difference between the aripiprazole 15-mg group and placebo showed statistical significance on the CGI Improvement Score, and the differences between each aripiprazole group and placebo showed statistical significance on the PANSS-Derived BPRS Core Score. Other comparisons between treatment and placebo on these two scores did not reach statistical significance.

Table 4.1.1.9 Other Efficacy Analysis Results for Patients with Schizophrenia only in the LOCF Data set for Study 31-97-201

Endpoints	N	Baseline	Change from Baseline to Endpoint (i.e., week 4)	Treatment Difference vs. Placebo	95% CI for Difference	P-Value
<i>PANSS Negative Sub-Scale Score</i>						
Haloperidol 10 mg	59	26.8	-3.3	-2.2	(-4.3, -0.1)	0.043
Aripiprazole 15 mg	72	25.1	-3.8	-2.6	(-4.6, -0.6)	0.011
Aripiprazole 30 mg	71	25.9	-2.1	-1.0	(-3.0, 1.1)	0.354
Placebo	74	26.4	-1.2			
Endpoints	Number (%) of Responders at Week 4			P-Value (vs. Placebo)		
<i>Percentage of Responders</i>						
Haloperidol 10 mg (N=59)	15 (25)			0.129		
Aripiprazole 15 mg (N=72)	25 (35)			0.006		
Aripiprazole 30 mg (N=71)	19 (27)			0.078		
Placebo (N=74)	11 (15)					
Endpoints	Mean at Week 4			P-Value (vs. Placebo)		
<i>CGI Improvement Score</i>						
Haloperidol 10 mg (N=59)	3.7			0.003		
Aripiprazole 15 mg (N=72)	3.4			<0.001		
Aripiprazole 30 mg (N=71)	3.8			0.023		
Placebo (N=74)	4.4					

Endpoints	N	Baseline	Change from Baseline to Endpoint (i.e., week 4)	Treatment Difference vs. Placebo	95% CI for Difference	P-Value
<i>PANSS-Derived</i>						
<i>BPRS Core Score</i>						
Haloperidol 10 mg	59	17.9	-3.3	-2.6	(-4.0, -1.1)	<0.001
Aripiprazole 15 mg	72	16.7	-3.0	-2.3	(-3.7, -0.9)	0.001
Aripiprazole 30 mg	71	16.9	-2.8	-2.1	(-3.5, -0.8)	0.002
Placebo	74	17.0	-0.7			

4.1.1.5 The Sponsor's Overall Efficacy Conclusions

- Both doses of aripiprazole were shown to be effective in the treatment of patients with schizophrenia and schizoaffective disorder in acute relapse based on the predefined primary efficacy measures of PANSS Total Score, PANSS Positive-Sub-scale Score, and CGI Severity of illness Score.
- Early onset of efficacy was demonstrated by Week 2 for the aripiprazole treatment groups as demonstrated by the PANSS-Positive Sub-Scale Score.
- Aripiprazole 15 mg improved negative symptoms of schizophrenia and schizoaffective disorder as measured by the PANSS Negative Sub-Scale Score.

4.1.2 Study 31-97-202

4.1.2.1 Disposition of Patients

A total of 487 patients signed the informed consent form; 39 of these patients failed screening and did not enter the placebo-washout phase. The remaining 448 patients underwent placebo washout; 44 of these patients discontinued from the study prior to randomization.

Four hundred and four patients were randomized to receive double-blind treatment; 103 to the placebo group, 99 to the risperidone 6-mg group, 101 to the aripiprazole 20-mg group and 101 to the aripiprazole 30-mg group. Of these, 242 (60%) patients completed 4 weeks of treatment and 162 (40%) discontinued from the study early. The disposition of all enrolled patients is presented in Table 3 of the Appendix. The time to discontinuation due to all reasons for the Randomized Sample is presented by treatment group in Figure 2 of Appendix.

Two hundred eighty-nine patients who had a DSM-IV diagnosis of schizophrenia were randomized to receive double-blind treatment: 78 to the placebo group, 74 to the risperidone group, 66 to the aripiprazole 20-mg group, and 71 to the aripiprazole 30-mg group. Of those, 183 (63%) completed 4 weeks of treatment and 106 (37%) discontinued from the study early. The disposition of schizophrenic patients randomized to treatment is presented by treatment group in Table 4 of the Appendix.

4.1.2.2 Data Set

The number of patients in each sample is presented by treatment group for all randomized patients in Table 4.1.2.1 and for patients with schizophrenia in Table 4.1.2.2

One of the 404 patients, i.e., Patient 97-202-71-22 in the aripiprazole 30-mg group, was excluded from the Safety Sample because he (she) did not receive study medication according to the dosing record.

Eleven (3%) of the 403 patients in the Safety Sample were excluded from the Efficacy Sample because they did not have a post-randomization efficacy evaluation.

Table 4.1.2.1 Number of Patients in Different Samples for Study 31-97-202

Sample	Placebo	Risperidone 6 mg	Aripiprazole		Total
			20 mg	30 mg	
Randomized	103	99	101	101	404
Safety	103	99	101	100	403
Efficacy	103	95	98	96	392

Table 4.1.2.2 Number of Patients with Schizophrenia in Different Samples for Study 31-97-202

Sample	Placebo	Risperidone 6 mg	Aripiprazole		Total
			20 mg	30 mg	
Randomized	78	74	66	71	289
Safety	78	74	66	71	289
Efficacy	78	71	65	68	282

4.1.2.3 Demography and Patient Characteristics

Demographic characteristics are presented by treatment group in Table 4.1.2.3 for all patient in the Randomized Sample and in Table 4.1.2.4 for patients with schizophrenia in the Randomized Sample.

Table 4.1.2.3 Demographic Characteristics for All Patients in Randomized Sample for Study 31-97-202

Variable		Placebo N = 103	Risperidone 6 mg N = 99	Aripiprazole		Total N = 404
				20 mg N = 101	30 mg N = 101	
Age (yrs)	Mean	38.8	38.6	38.1	40.2	38.9
	Median	39.0	39.0	39.0	41.0	39.0
	Min-Max	18.0-62.0	18.0-64.0	18.0-57.0	20.0-65.0	18.0-65.0
	S.E.	1.0	0.9	0.9	1.1	0.5
Gender	Men	73 (71)	71 (72)	73 (72)	66 (65)	283 (70)
	Women	30 (29)	28 (28)	28 (28)	35 (35)	121 (30)
Race	White	57 (58)	54 (55)	59 (60)	59 (61)	229 (58)
	Black	35 (35)	38 (39)	31 (32)	33 (34)	137 (35)
	Hispanic	4 (4)	4 (4)	6 (6)	3 (3)	17 (4)

Variable		Placebo N = 103	Risperidone 6 mg N = 99	Aripiprazole		Total N = 404
				20 mg N = 101	30 mg N = 101	
Weight(kg)	Asian/Pacific					
	Islander	3 (3)	2 (2)	2 (2)	2 (2)	9 (2)
	Not recorded	4	1	3	4	12
	Mean	85.2	82.4	87.2	84.0	84.7
	Median	81.7	79.5	84.0	82.2	81.9
	Min-Max	48.8-132.5	54.0-145.3	49.7-194.8	44.5-158.0	44.5-194.8
	S.E.	1.9	1.7	2.2	2.0	1.0
	Missing	1	0	2	0	3

Table 4.1.2.4 Demographic Characteristics for Patients with Schizophrenia in Randomized Sample for Study 31-97-202

Variable		Placebo N=78	Risperidone 6 mg N = 74	Aripiprazole		Total N = 289
				20 mg N = 66	30 mg N= 71	
Age (yrs)	Mean	39.7	39.2	38.0	40.9	39.5
	Median	39.5	39.5	38.5	42.0	40.0
	Min-Max	18.0-62.0	19.0-64.0	18.0-57.0	20.0-63.0	18.0-64.0
	S.E.	1.2	1.0	1.1	1.2	0.6
Gender	Men	61 (78)	58 (78)	52 (79)	55 (77)	226 (78)
	Women	17 (22)	16 (22)	14 (21)	16 (23)	63 (22)
Race	White	40 (52)	36 (49)	33 (51)	38 (55)	147 (52)
	Black	32 (42)	33 (45)	28 (43)	27 (39)	120 (42)
N (%)	Hispanic	3 (4)	3 (4)	3 (5)	3 (4)	12 (4)
	Asian/Pacific					
	Islander	2 (3)	1 (1)	1 (2)	1 (1)	5 (2)
Weight(kg)	Not recorded	1	1	1	2	5
	Mean	82.5	83.4	85.1	82.0	83.2
	Median	78.4	82.2	83.8	81.7	81.7
	Min-Max	48.8-132.5	54.0-145.3	49.7-134.4	44.5-158.0	44.5-158.0
	S.E.	2.0	2.0	2.2	2.5	1.1

4.1.2.4 The Sponsor's Efficacy Results

Efficacy analyses were performed using the Efficacy Sample (N=392), which comprised all patients who had a baseline and a post-randomization efficacy evaluation on at least one of the primary or secondary efficacy variables. In addition, post hoc efficacy analyses were performed for a subset of patients with schizophrenia (N=282) on the key outcome measures to gather information on the efficacy of aripiprazole in schizophrenia as recommended by European regulatory authorities.

4.1.2.4.1 For All Randomized Patients

Primary Efficacy Measures:

The summaries of the analysis results for the three primary endpoints for all patients are shown in Table 4.1.2.5.

Table 4.1.2.5 Efficacy Analysis Results for the Primary Endpoints for Study 31-97-202
For the LOCF Data Set:

Endpoints	N	Baseline	Change from Baseline to Endpoint (i.e., week 4)	Treatment Difference vs. Placebo	95% CI for Difference	P-Value
<i>PANSS Total</i>						
Risperidone 6 mg	95	92.6	-15.7	-10.7	(-16.6, -4.9)	0.0004
Aripiprazole 20 mg	98	93.5	-14.5	-9.6	(-15.4, -3.8)	0.0013
Aripiprazole 30 mg	96	91.6	-13.9	-9.0	(-14.8, -3.1)	0.0029
Placebo	103	94.1	-5.0			
<i>PANSS Positive Sub-Scale Score</i>						
Risperidone 6 mg	95	23.7	-5.2	-3.4	(-5.2, -1.6)	0.0002
Aripiprazole 20 mg	98	24.6	-4.9	-3.1	(-4.9, -1.4)	0.0006
Aripiprazole 30 mg	96	23.7	-3.9	-2.2	(-3.9, -0.4)	0.0177
Placebo	103	24.2	-1.8			
<i>CGI Severity of Illness Score</i>						
Risperidone 6 mg	95	4.8	-0.7	-0.6	(-0.8, -0.3)	0.0001
Aripiprazole 20 mg	98	4.8	-0.5	-0.3	(-0.6, -0.0)	0.0298
Aripiprazole 30 mg	96	4.7	-0.6	-0.4	(-0.7, -0.1)	0.0063
Placebo	103	4.8	-0.2			

For the OC Data Set:

Endpoints	Baseline & (N)	Change from Baseline to Endpoint & (N)	Treatment Difference vs. Placebo	95% CI for Difference	P-Value
<i>PANSS Total</i>					
Risperidone 6 mg	93.6 (N=95)	-22.7 (N=61)	-4.5	(-11.3, 2.3)	0.191
Aripiprazole 20 mg	94.0 (N=98)	-23.4 (N=61)	-5.2	(-12.0, 1.6)	0.132
Aripiprazole 30 mg	92.3 (N=96)	-20.2 (N=68)	-2.0	(-8.6, 4.6)	0.552
Placebo	95.0 (N=103)	-18.2 (N=52)			
<i>PANSS Positive Sub-Scale Score</i>					
Risperidone 6 mg	23.9 (N=95)	-7.3 (N=61)	-1.9	(-4.0, 0.2)	0.073
Aripiprazole 20 mg	24.8 (N=98)	-7.5 (N=61)	-2.2	(-4.3, -0.1)	0.045
Aripiprazole 30 mg	24.0 (N=96)	-5.7 (N=68)	-0.4	(-2.5, 1.7)	0.700
Placebo	24.5 (N=103)	-5.3 (N=52)			
<i>CGI Severity of Illness Score</i>					
Risperidone 6 mg	4.8 (N=95)	-1.1 (N=61)	-0.4	(-0.7, -0.0)	0.043
Aripiprazole 20 mg	4.8 (N=98)	-1.0 (N=61)	-0.2	(-0.6, 0.1)	0.165
Aripiprazole 30 mg	4.7 (N=96)	-0.9 (N=68)	-0.2	(-0.5, 0.2)	0.335
Placebo	4.8 (N=103)	-0.7 (N=52)			

According to the sponsor's study report, the analysis of the mean change in the PANSS Total Score for the LOCF data set showed that the risperidone group and both aripiprazole groups had significantly greater improvement compared with the placebo group during

Weeks 1 through 4. The analysis of the mean change score for the OC data set showed that both aripiprazole groups and the risperidone group had significantly greater improvement compared with the placebo group at Week 1, while only the aripiprazole 20-mg group improved significantly more than the placebo group at Week 2. There were no significant differences among any of the treatment groups and placebo at Weeks 3 and 4.

The analysis of the model-based mean change in the PANSS Positive Sub-Scale Score for the LOCF data set showed that both aripiprazole groups and the risperidone group had significantly greater improvement compared with the placebo group during Weeks 1 through 4. Results of the OC analysis showed that the aripiprazole 20-mg group had significantly greater improvement compared with the placebo group at Weeks 1, 2 and 4. Patients treated with aripiprazole 30-mg had significantly greater improvement compared with placebo at Week 1 only. Significantly greater improvement was seen for the risperidone group compared with the placebo group at Weeks 1 and 2.

The analysis of the model-based mean change from baseline in the CGI Severity of Illness Score for the LOCF data set showed that the risperidone group had significantly greater improvement compared with the placebo group during Weeks 1 through 4. The aripiprazole 20-mg and 30-mg groups showed significantly greater improvement compared with the placebo group during Weeks 2 through 4. The analysis for the OC data set showed that the aripiprazole 20-mg and 30-mg groups only had significantly greater improvement compared with the placebo group during Week 2. Significantly greater improvement compared with placebo was seen for the risperidone group at Weeks 1, 2 and 4.

Secondary Efficacy Measures:

The summaries of efficacy analysis results for the secondary efficacy measure in LOCF data set for all schizophrenia and schizoaffective disorder patients are shown in Table 4.1.2.6.

According to the sponsor's study reports, the analysis of the model-based mean change in the PANSS Negative Sub-Scale Score for the LOCF data set showed significantly greater improvement compared with the placebo group for both aripiprazole groups at Weeks 1 through 4. The risperidone group showed significantly greater improvement compared with the placebo group during Weeks 2 through 4. The analysis of the OC data set showed significantly greater improvement compared with the placebo group at Weeks 1 and 2 for both aripiprazole groups. The risperidone group was comparable to placebo at all time points.

The time-to-response analysis was performed separately for the PANSS Total Score (PANSS response) and the CGI Improvement Score (CGI response) using survival analysis. Only the aripiprazole 30-mg group showed a significant difference from the placebo group ($p=0.0278$) in the time to PANSS response (30% decrease from baseline) and in the time to an improvement of 1 or 2 in CGI global score ($p=0.0430$).

Lack of efficacy is defined as insufficient clinical response or withdrawal of consent by the patient due to lack of effect. Results of the analysis of time to discontinuation due to lack of efficacy (log-rank test) showed a between-treatment difference in the comparison between aripiprazole 20 mg and the placebo (p=0.0256).

Table 4.1.2.6 Efficacy Analysis Results for the Secondary Efficacy Endpoints for Study 31-97-202

For the LOCF Data Set:

Endpoints	N	Baseline	Change from Baseline to Endpoint (i.e., week 4)	Treatment Difference vs. Placebo	95% CI for Difference	P-Value
<i>PANSS Negative Sub-Scale Score</i>						
Risperidone 6 mg	95	24.3	-3.1	-2.3	(-3.9, -0.7)	0.005
Aripiprazole 20 mg	98	23.6	-3.4	-2.6	(-4.1, -1.0)	0.002
Aripiprazole 30 mg	96	23.0	-3.4	-2.5	(-4.1, -1.0)	0.002
Placebo	103	23.5	-0.8			

For the OC Data Set:

Endpoints	Baseline & (N)	Change from Baseline to Endpoint & (N)	Treatment Difference vs. Placebo	95% CI for Difference	P-Value
<i>PANSS Negative Sub-Scale Score</i>					
Risperidone 6 mg	24.3 (N=95)	-4.7 (N=61)	-0.9	(-2.9, 1.0)	0.352
Aripiprazole 20 mg	23.5 (N=98)	-5.6 (N=61)	-1.8	(-3.8, 0.1)	0.064
Aripiprazole 30 mg	23.0 (N=96)	-5.0 (N=68)	-1.2	(-3.1, 0.7)	0.203
Placebo	23.5 (N=103)	-3.7 (N=52)			

Endpoints	P-Value by the Log-Rank Test
<i>Time to Response to Therapy (Defined as a 30% Decrease from Baseline in PANSS Total Score)</i>	
Risperidone 6 mg vs. Placebo	0.0574
Aripiprazole 20 mg vs. Placebo	0.1111
Aripiprazole 30 mg vs. Placebo	0.0278

Endpoints	P-Value by the Log-Rank Test
<i>Time to Response to Therapy (Defined as a CGI Global Improvement Score of 1 or 2)</i>	
Risperidone 6 mg vs. Placebo	0.6164
Aripiprazole 20 mg vs. Placebo	0.0611
Aripiprazole 30 mg vs. Placebo	0.0430

Endpoints	P-Value by the Log-Rank Test
<i>Time to Discontinuation due to Lack of Efficacy</i>	
Risperidone 6 mg vs. Placebo	0.0662
Aripiprazole 20 mg vs. Placebo	0.0256
Aripiprazole 30 mg vs. Placebo	0.0501

Other Efficacy Measures:

Table 4.1.2.7 shows the summaries of the sponsor's results for other efficacy measures. A patient who had a CGI Improvement Score of 1 (very much improved) or 2 (much improved), or a $\geq 30\%$ decrease from baseline in the PANSS Total Score was considered a responder. Within the LOCF data set, the aripiprazole 30-mg group had a significantly greater percentage of responders compared with the placebo group at all time points. The aripiprazole 20-mg group and the risperidone group had a significantly greater percentage of responders compared with the placebo group at Weeks 2 through 4. Within the OC data set, only the aripiprazole 30-mg group at Week 2 had a significantly greater percentage of responders compared with placebo.

The analysis of the unadjusted mean CGI Improvement Score for the LOCF data set showed that all treatment groups had significantly greater improvement compared with the placebo group throughout the 4-week study. In the OC analysis, significantly greater improvement compared with the placebo group was seen at Weeks 1 and 2 for both aripiprazole groups and at Week 2 for the risperidone group.

The analysis of the model-based mean change in the PANSS-Derived BPRS Core Score for the LOCF data set showed significantly greater improvement compared with the placebo group for both aripiprazole groups as well as the risperidone group at Weeks 1 through 4. The analysis of the OC data set showed significantly greater improvement compared with the placebo group at Weeks 1 and 2 for the aripiprazole 20-mg group and the risperidone group, and at Week 1 for the aripiprazole 30-mg group.

Table 4.1.2.7 Efficacy Analysis Results for the Other Efficacy Measures for Study 31-97-202

Endpoints	Number (%) of Responders at Week 4	P-Value (vs. Placebo)
<i>Percentage of Responders*</i>		
Risperidone 6 mg (N=95)	38 (40)	0.008
Aripiprazole 20 mg (N=98)	35 (36)	0.043
Aripiprazole 30 mg (N=96)	39 (41)	0.005
Placebo (N=103)	24 (23)	
Endpoints	Mean at Week 4	P-Value (vs. Placebo)
<i>CGI Improvement Score*</i>		
Risperidone 6 mg (N=95)	3.3	<0.001
Aripiprazole 20 mg (N=98)	3.4	0.005
Aripiprazole 30 mg (N=96)	3.3	0.001
Placebo (N=103)	4.0	

Endpoints	N	Baseline	Change from Baseline to Endpoint (i.e., week 4)	Treatment Difference vs. Placebo	95% CI for Difference	P-Value
<i>PANSS-Derived BPRS Core Score*</i>						
Risperidone 6 mg	95	16.4	-3.9	-2.2	(-3.4, -1.0)	<0.001
Aripiprazole 20 mg	98	16.7	-3.5	-1.8	(-3.0, -0.6)	0.004
Aripiprazole 30 mg	96	16.5	-3.3	-1.5	(-2.7, -0.3)	0.013
Placebo	103	16.6	-1.7			

* The results shown above are for the LOCF data set.

Subgroup Analysis:

The subgroup analysis results by gender and study center for the PANSS Total Score in the LOCF analysis are shown in Table 7 and 8 of the Appendix. In addition to showing some descriptive statistics, the sponsor commented that there was no significant treatment-by-center interaction.

4.1.2.4.2 For Patients with Schizophrenia

Primary Efficacy Measures:

The LOCF analysis results of the model-based mean change in the PANSS Total Score, the PANSS Positive Sub-Scale Score and the CGI Severity of Illness Score for patients with schizophrenia are shown in Table 4.1.2.8. As we can see in the table, the LOCF analysis of the model-based mean change in the PANSS Total Score and the PANSS Positive Sub-Scale Score for patients with schizophrenia showed that all three treatment groups had significantly greater improvement compared with the placebo group at Week 4. For the change of CGI Severity of Illness Score, the aripiprazole 30-mg group and the risperidone group had significantly greater improvement compared with the placebo group at Week 4 but not the aripiprazole 20-mg group. For the OC analyses (Table 13 of the Appendix), only the comparison between aripiprazole 20-mg group and the placebo group showed statistical significance in the change of the PANSS Positive Sub-Scale Score.

Table 4.1.2.8 Efficacy Analysis Results of the Primary Endpoints for the LOCF Data Set for Patients with Schizophrenia for Study 31-97-202

Endpoints	N	Baseline	Change from Baseline to Endpoint (i.e., week 4)	Treatment Difference vs. Placebo	95% CI for Difference	P-Value
<i>PANSS Total</i>						
Risperidone 6 mg	71	94.4	-15.0	-9.5	(-16.3, -2.8)	0.006
Aripiprazole 20 mg	65	92.2	-15.0	-9.5	(-16.4, -2.6)	0.007
Aripiprazole 30 mg	68	92.7	-14.5	-9.0	(-15.8, -2.2)	0.009
Placebo	78	94.4	-5.5			

Endpoints	N	Baseline	Change from Baseline to Endpoint (i.e., week 4)	Treatment Difference vs. Placebo	95% CI for Difference	P-Value
<i>PANSS Positive Sub-Scale Score</i>						
Risperidone 6 mg	71	24.3	-4.9	-3.0	(-5.0, -0.9)	0.005
Aripiprazole 20 mg	65	24.6	-5.1	-3.1	(-5.2, -1.0)	0.004
Aripiprazole 30 mg	68	24.3	-4.2	-2.2	(-4.3, -0.1)	0.038
Placebo	78	24.7	-2.0			
<i>CGI Severity of Illness Score</i>						
Risperidone 6 mg	71	4.9	-0.7	-0.5	(-0.8, -0.1)	0.005
Aripiprazole 20 mg	65	4.8	-0.6	-0.3	(-0.6, 0.0)	0.083
Aripiprazole 30 mg	68	4.8	-0.7	-0.4	(-0.7, -0.1)	0.016
Placebo	78	4.8	-0.3			

Other Efficacy Measures:

Table 4.1.2.9 shows the other efficacy analysis results for patients with schizophrenia only in the LOCF data set.

Table 4.1.2.9 Other Efficacy Analysis Results for Patients with Schizophrenia only in the LOCF Data set for Study 97-202

Endpoints	N	Baseline	Change from Baseline to Endpoint (i.e., week 4)	Treatment Difference vs. Placebo	95% CI for Difference	P-Value
<i>PANSS Negative Sub-Scale Score*</i>						
Risperidone 6 mg	71	24.8	-3.3	-2.5	(-4.3, -0.6)	0.008
Aripiprazole 20 mg	65	23.1	-3.7	-2.8	(-4.7, -1.0)	0.003
Aripiprazole 30 mg	68	23.5	-3.4	-2.5	(-4.3, -0.7)	0.008
Placebo	78	23.3	-0.9			

Endpoints	Number (%) of Responders at Week 4	P-Value (vs. Placebo)
<i>Percentage of Responders*</i>		
Risperidone 6 mg (N=71)	23 (32)	0.278
Aripiprazole 20 mg (N=65)	26 (40)	0.046
Aripiprazole 30 mg (N=68)	29 (43)	0.019
Placebo (N=78)	19 (24)	

Endpoints	Mean at Week 4	P-Value (vs. Placebo)
<i>CGI Improvement Score*</i>		
Risperidone 6 mg (N=71)	3.3	0.012
Aripiprazole 20 mg (N=65)	3.4	0.034
Aripiprazole 30 mg (N=68)	3.3	0.010
Placebo (N=78)	3.9	

Endpoints	N	Baseline	Change from Baseline to Endpoint (i.e., week 4)	Treatment Difference vs. Placebo	95% CI for Difference	P-Value
<i>PANSS-Derived</i>						
<i>BPRS Core Score*</i>						
Risperidone 6 mg	71	16.6	-3.8	-2.1	(-3.5, -0.7)	0.004
Aripiprazole 20 mg	65	16.7	-3.7	-2.0	(-3.4, -0.5)	0.008
Aripiprazole 30 mg	68	16.8	-3.3	-1.6	(-3.0, -0.2)	0.028
Placebo	78	16.9	-1.7			

*The results shown above are for the LOCF data set.

As we can observe in the table, all three treatment groups had significantly greater improvement compared with the placebo group at Week 4 in the LOCF analyses of PANSS Negative Sub-Scale. The OC analysis for the PANSS Negative Sub-Scale, however, only the comparison between Aripiprazole 20-mg and Placebo showed statistical significance.

For the percent of responders, the LOCF data analysis with patients with schizophrenia only showed a significantly greater at Week 4 for both aripiprazole groups and the risperidone group. In the OC analysis, none of the comparisons between treatment groups and the placebo group showed significant results at Week 4.

For the CGI Improvement Score, all treatment groups showed significantly greater improvement compared with the placebo group in the LOCF analysis with schizophrenia patients only. However, none of the comparisons between the treatment groups and placebo showed significantly improvement in the OC analysis.

For the change in the PANSS-Derived BPRS Core Score, all treatment groups showed significantly greater improvement compared with the placebo group at Week 4 in the LOCF analysis with schizophrenia patients only. The OC analysis, however, there is only one significant result which was shown in the aripiprazole 20-mg group.

4.1.2.5 The Sponsor's Overall Efficacy Conclusions

- Both doses of aripiprazole were shown to be effective in the treatment of patients with schizophrenia and schizoaffective disorder in acute relapse based on the predefined primary efficacy measures of PANSS Total Score, PANSS Positive Sub-Scale Score, and CGI Severity of Illness Score.
- Early onset of efficacy was seen by Week 1 for the aripiprazole treatment groups as demonstrated by the PANSS Positive Sub-Scale Score.
- Both doses of aripiprazole improved negative symptoms of schizophrenia and schizoaffective disorder as measured by the PANSS Negative Sub-Scale Score.

4.1.3 Study CN 138-001

4.1.3.1 Disposition of Patients

Five hundred eight patients were enrolled in the study. Of these, 420 were randomized to receive double-blind treatment; 108 to the placebo group, 106 to the aripiprazole 10-mg group, 106 to the aripiprazole 15-mg group, and 100 to the aripiprazole 20-mg group. Of the 420 randomized patients, 214 (51%) completed 6 weeks of treatment and 206 (49%) discontinued from the study early.

Of the 420 randomized patients, 142 (34%) completed double-blind treatment and 278 (66%) discontinued double-blind treatment. Of the 420 randomized patients, 131 (31%) switched from double-blind treatment to open-label treatment at end of Weeks 3 or 4. Seventy-two (55%) of these patients completed the Acute Phase (and are included in the number of completers above); 59 (45%) of these patients discontinued open-label treatment before Week 6. The disposition of all patients enrolled in the study as specified on the end-of-study CRF is presented by treatment group in Table 9 and in Table 10 of Appendix for patients who entered open-label treatment.

The time to discontinuation for any reason is presented in Figure 3 of Appendix.

4.1.3.2 Data Set

The distribution of all randomized patients within each of the patient samples is presented by treatment group in Table 4.1.3.1.

Table 4.1.3.1 Number of Patients in Samples for Study CN138-001

Sample	Placebo	Aripiprazole 10 mg	Aripiprazole 15 mg	Aripiprazole 20 mg	Total
Randomized	108	106	106	100	420
Safety	107	105	105	98	415
Efficacy	107	103	103	97	410

Five of the 420 randomized patients were excluded from the Safety Sample because they did not receive study medication according to the dosing record. Five of the 415 patients in the Safety Sample were excluded from the Efficacy Sample because they did not have a post-randomization efficacy rating.

4.1.3.3 Demography and Patient Characteristics

Demographic characteristics for the Randomized Sample are presented by treatment group in Table 4.1.3.2 for the Randomized Sample. According to the table, the treatment groups were comparable with respect to age, gender, race, and weight.

Table 4.1.3.2 Demographic Characteristics, Randomized Sample for Study CN138-001

Variable		Placebo N=108	Aripiprazole 10 mg N=106	Aripiprazole 15 mg N=106	Aripiprazole 20 mg N=100	Total N=420
Age (yrs)	Mean	41.2	40.0	40.0	40.4	40.4
	Median	41.0	39.5	41.0	40.0	41.0
	Min-Max	19.0-76.0	18.0-73.0	19.0-68.0	19.0-69.0	18.0-76.0
	S.E.	1.1	1.1	1.1	1.1	0.5
Gender	Male	83 (77)	82 (77)	79 (75)	82 (82)	326 (78)
	Female	25 (23)	24 (23)	27 (25)	18 (18)	94 (22)
Race	White	49 (45)	53 (50)	57 (54)	52 (52)	211 (50)
	Black	37 (34)	29 (27)	28 (26)	29 (29)	123 (29)
N (%)	Asian/Pacific Islander	4 (4)	1 (1)	4 (4)	3 (3)	12 (3)
	Hispanic/ Latino	17 (16)	19 (18)	16 (15)	12 (12)	64 (15)
	American/ Alaskan Native	0	0	0	1 (1)	1(15)
	Other	1 (1)	4 (4)	1(1)	3 (3)	9 (2)
	Weight	Mean	84.1	82.9	81.5	86.7
(kg)	Median	81.0	80.7	79.3	83.3	81.0
	Min-Max	45.0-143.3	44.8-142.2	49.5-147.2	36.5-164.7	36.5-164.7
	S.E.	1.9	2.0	1.9	2.4	1.0
	Missing	1	2	0	2	5

4.1.3.4 The Sponsor's Efficacy Results

Primary Efficacy Measure: Mean Change from Baseline in PANSS Total Score

Change in PANSS Total Scores were derived by subtracting baseline PANSS Total Scores from the PANSS Total Score at each study week. Negative change Scores indicate Improvement. The mean change from baseline to Week 6 in the PANSS Total Score was the primary efficacy measure. Results of the analysis of the mean change in the PANSS Total Score are shown by treatment group and study week in Table 4.1.3.3 for the LOCF data set and 4.1.3.4 for the OC data set.

As we can observe from the tables, the analysis of the change in the PANSS Total Score for the LOCF data set at Week 6 showed that patients in all three aripiprazole treatment groups had significantly greater improvement compared with patients in the placebo group. The analysis of the change Scores for the LOCF data set for aripiprazole 10 mg showed significantly greater improvement compared with the placebo group from Week 1 through Week 6. Aripiprazole 15 mg was statistically significantly different from placebo from Week 3 through Week 6. Aripiprazole 20 mg showed significantly greater improvement compared with placebo from Week 1 through Week 6.

The analysis of the mean change from baseline for the OC data showed that aripiprazole 10 mg showed significantly greater improvement compared with the placebo group at Weeks 1, 2 and 3. Aripiprazole 15 mg was not statistically significantly different than placebo at any week. Aripiprazole 20 mg showed significantly greater improvement compared with the placebo group at Weeks 1 and 3. As expected, at Week 4 sample sizes

decreased substantially and mean change from baseline PANSS Total Score improved for all treatment groups when the option to move to open-label aripiprazole could be exercised.

Additionally, results for the analysis of the unadjusted mean change from baseline in the PANSS Total Score was consistent with those of the adjusted mean change Scores. Results from the NPAR1WAY analysis of PANSS Total Score in the LOCF Data Set generally support the primary efficacy analysis. However, aripiprazole 15 mg did not achieve statistical significance until Week 4 and aripiprazole 20 mg did not achieve statistical significance until Week 2. Moreover, results of the linear trend test showed that when placebo was included there was a linear trend across the four treatment groups starting at Week 1 but when placebo was not included there was no linear trend across the three aripiprazole treatment groups at any study week.

Table 4.1.3.3 Mean Change from Baseline in PANSS Total Score, LOCF Data Set, Efficacy Sample for Study CN138-001

	PANSS Total Score				Pairwise Comparisons P-values		
	Placebo	Aripiprazole	Aripiprazole	Aripiprazole	Ari10 vs Placebo	Ari15 vs Placebo	Ari20 vs Placebo
	N = 107	10 mg N = 103	15 mg N = 103	20 mg N = 97			
Baseline	92.63	92.90	92.42	91.91	0.902	0.925	0.746
Day 4	-2.78	-3.47	-4.35	-5.22	0.650	0.304	0.116
Week 1	-3.32	-7.89	-6.47	-8.32	0.023	0.116	0.015
Week 2	-3.27	-11.63	-7.76	-10.80	0.001	0.068	0.003
Week 3	-2.73	-12.66	-8.50	-11.79	<0.001	0.038	0.001
Week 4	-2.72	-13.30	-10.40	-12.15	<0.001	0.010	0.002
Week 5	-2.02	-14.14	-10.71	-13.30	<0.001	0.005	<0.001
Week 6	-2.33	-15.04	-11.73	-14.44	<0.001	0.004	<0.001

Table 4.1.3.4 Mean Change from Baseline in PANSS Total Score, OC Data Set, Efficacy Sample for Study CN138-001

	PANSS Total Score				Pairwise Comparisons P-values		
	Placebo (N)	Aripiprazole 10 mg (N)	Aripiprazole 15 mg (N)	Aripiprazole 20 mg (N)	Ari10 vs Placebo	Ari15 vs Placebo	Ari20 vs Placebo
Baseline	92.40 (107)	92.76 (103)	93.27 (103)	92.29 (97)	0.902	0.763	0.969
Day 4	-2.61 (100)	-3.40 (97)	-4.56 (100)	-5.12 (94)	0.603	0.198	0.103
Week 1	-3.21 (100)	-8.46 (89)	-6.08 (95)	-7.44 (87)	0.013	0.169	0.047
Week 2	-5.30 (88)	-12.87 (86)	-7.31 (89)	-10.55 (81)	0.005	0.451	0.055
Week 3	-7.45 (82)	-15.69 (78)	-10.59 (82)	-14.99 (68)	0.008	0.300	0.018
Week 4	-18.96 (42)	-23.69 (51)	-22.51 (41)	-20.86 (49)	0.212	0.373	0.619
Week 5	-26.41 (31)	-27.78 (45)	-23.89 (37)	-25.91 (40)	0.704	0.500	0.891
Week 6	-26.86 (30)	-33.42 (42)	-31.92 (34)	-28.91 (39)	0.113	0.242	0.624

Key Secondary Analyses: the PANSS-Derived BPRS Core Score and the PANSS Negative Sub-Scale Score

The mean change from baseline to Week 6 in the PANSS-derived BPRS Core Score was the first of two key secondary measures. Since the primary efficacy measure showed