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 (ARIPIPRAZOLE) 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

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

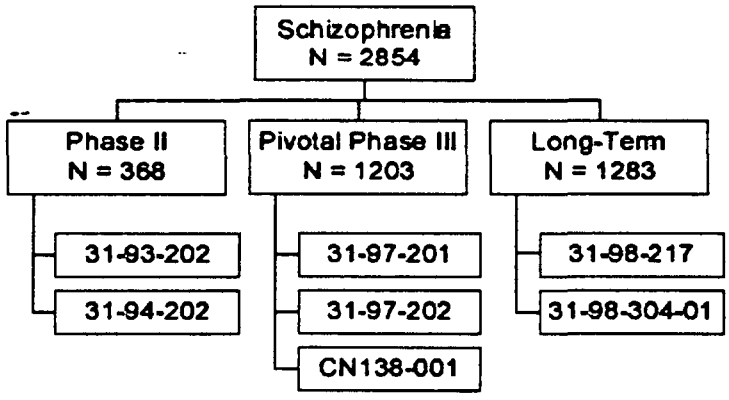
The Otsuka Pharmaceutical Co., Ltd submitted this application to present an overview of all the data in the Abilitat™ (aripiprazole) drug development program demonstrating a positive benefit/risk profile for the treatment of patients with schizophrenia. Collaborative development of aripiprazole between 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 aripriprazole 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, aripiprzole 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 schizophernia. 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 ≥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 method: (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 preciously 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 lat 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 sginificant (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 schizophernia 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 ≥ 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 15 mg N = 102	Aripiprazole 30 mg N=102	Total N = 414
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)
N (%)	Women	32 (30)	36 (35)	26 (25)	32 (31)	126 (30)
Race	White	54 (51)	68 (67)	61 (60)	59 (59)	242 (59)
N (%)	Black	34 (32)	23 (23)	26 (25)	26 (26)	109 (27)
	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	85.3
	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 15 mg N = 74	Aripiprazole 30 mg N=72	Total N = 282
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	Men	56 (75)	43 (70)	59 (80)	51 (71)	209 (74)
N (%)	Women	19 (25)	18 (30)	15 (20)	21 (29)	73 (26)
Race	White	36 (48)	39 (66)	37 (50)	41 (59)	153 (55)
N (%)	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
PANSS Total						
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			
PANSS Positive Sub-Scale Score						
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			
CGI Severity of Illness Score --						
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
PANSS Total					
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)			
PANSS Positive Sub-Scale Score					
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)			
CGI Severity of Illness Score					
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
PANSS Negative Sub-Scale Score						
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
PANSS Negative Sub-Scale Score					
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
Time to Response to Therapy (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
Time to Response to Therapy (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
Time to Discontinuation due to Lack of Efficacy	
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 ≥ 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)
*Percentage of Responders**		
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)
*CGI Improvement Score**		
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
*PANSS-Derived BPRS Core Score**						
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
PANSS Total						
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			
PANSS Positive Sub-Scale Score						
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			
CGI Severity of Illness Score						
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
PANSS Negative Sub-Scale Score						
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)
Percentage of Responders		
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)
CGI Improvement Score		
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
PANSS-Derived BPRS Core Score						
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 20 mg N = 101	Aripiprazole 30 mg N = 101	Total N = 404
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)
N (%)	Women	30 (29)	28 (28)	28 (28)	35 (35)	121 (30)
Race	White	57 (58)	54 (55)	59 (60)	59 (61)	229 (58)
N (%)	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 20 mg N = 101	Aripiprazole 30 mg N = 101	Total N = 404
	Asian/Pacific Islander	3 (3)	2 (2)	2 (2)	2 (2)	9 (2)
	Not recorded	4	1	3	4	12
Weight(kg)	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 20 mg N = 66	Aripiprazole 30 mg N= 71	Total N = 289
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)
N (%)	Women	17 (22)	16 (22)	14 (21)	16 (23)	63 (22)
Race	White	40 (52)	36 (49)	33 (51)	38 (55)	147 (52)
N (%)	Black	32 (42)	33 (45)	28 (43)	27 (39)	120 (42)
	Hispanic	3 (4)	3 (4)	3 (5)	3 (4)	12 (4)
	Asian/Pacific Islander	2 (3)	1 (1)	1 (2)	1 (1)	5 (2)
	Not recorded	1	1	1	2	5
Weight(kg)	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
PANSS Total						
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			
PANSS Positive Sub-Scale Score						
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			
CGI Severity of Illness Score						
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
PANSS Total					
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)			
PANSS Positive Sub-Scale Score					
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)			
CGI Severity of Illness Score					
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-gm 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
PANSS Negative Sub-Scale Score						
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
PANSS Negative Sub-Scale Score					
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
Time to Response to Therapy (Defined as a 30% Decrease from Baseline in PANSS Total Score)	
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
Time to Response to Therapy (Defined as a CGI Global Improvement Score of 1 or 2)	
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
Time to Discontinuation due to Lack of Efficacy	
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 ≥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)
*Percentage of Responders**		
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)
*CGI Improvement Score**		
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
*PANSS-Derived BPRS Core Score**						
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
PANSS Total						
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
PANSS Positive Sub-Scale Score						
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			
CGI Severity of Illness Score						
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
*PANSS Negative Sub-Scale Score**						
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)
*Percentage of Responders**		
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)
*CGI Improvement Score**		
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
*PANSS-Derived BPRS Core Score**						
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 measrues 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)
N (%)	Female	25 (23)	24 (23)	27 (25)	18 (18)	94 (22)
Race	White	49 (45)	53 (50)	57 (54)	52 (52)	211 (50)
N (%)	Black	37 (34)	29 (27)	28 (26)	29 (29)	123 (29)
	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	83.8
(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 statistically 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 N = 107	Aripiprazole 10 mg N = 103	Aripiprazole 15 mg N = 103	Aripiprazole 20 mg N = 97	Ari10 vs Placebo	Ari15 vs Placebo	Ari20 vs Placebo
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

significantly greater improvement compared with the placebo group at Week 6, analysis of this key secondary measure was performed for all three treatment groups versus the placebo group. The results of the analysis of the change in the PANSS-derived BPRS Core Score for the LOCF data set at Week 6 are shown in Table 4.1.3.5. The analysis showed significantly greater improvement for all three aripiprazole treatment groups compared with 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 2 through Week 6. Aripiprazole 15 mg was statistically significantly different from placebo at Week 5 and 6. Aripiprazole 20 mg showed significantly greater improvement compared with placebo from Week 2 through Week 6.

The results of the analysis of the mean change from baseline for the OC data set is shown in Table 4.1.3.6. The aripiprazole 10-mg group showed significantly greater improvement compared with the placebo group at Weeks 2, 3 and 6. Aripiprazole 15 mg did not show significantly greater improvement compared with the placebo group at any week. Aripiprazole 20 mg was statistically significantly different from placebo at Week 3. As expected, Week 4 sample sizes decreased substantially and mean change from baseline PANSS-derived BPRS Core Score improved for all treatment groups when the option to move to open-label aripiprazole could be exercised.

Table 4.1.3.5. Mean Change from Baseline in PANSS-Derived BPRS Core Score, LOCF Data Set, Efficacy Sample for Study CN138-001

	PANSS-Derived BPRS Core Score				Pairwise Comparisons P-values		
	Placebo N = 107	Aripiprazole 10 mg N = 103	Aripiprazole 15 mg N = 103	Aripiprazole 20 mg N = 97	Ari10 vs Placebo	Ari15 vs Placebo	Ari20 vs Placebo
Baseline	16.92	16.99	16.76	16.68	0.857	0.680	0.530
Day 4	-1.09	-1.11	-1.11	-1.46	0.948	0.951	0.258
Week 1	-1.48	-2.22	-1.98	-2.30	0.077	0.236	0.055
Week 2	-1.51	-3.21	-2.39	-2.84	<0.001	0.069	0.007
Week 3	-1.47	-3.45	-2.26	-3.21	<0.001	0.144	0.002
Week 4	-1.57	-3.59	-2.59	-3.16	<0.001	0.073	0.006
Week 5	-1.40	-3.77	-2.67	-3.31	<0.001	0.034	0.002
Week 6	-1.37	-3.91	-2.88	-3.56	<0.001	0.014	<0.001

Table 4.1.3.6 Mean Change from Baseline in PANSS-Derived BPRS Core Score, OC Data Set, Efficacy Sample for Study CN138-001

	PANSS-Derived BPRS Core 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	16.78 (107)	16.87 (103)	16.78 (103)	16.69 (97)	0.825	0.998	0.851
Day 4	-1.03 (100)	-1.09 (97)	-1.10 (100)	-1.40 (94)	0.857	0.847	0.264
Week 1	-1.46 (100)	-2.28 (89)	-1.81 (95)	-1.99 (87)	0.067	0.424	0.242
Week 2	-1.94 (88)	-3.45 (86)	-2.30 (89)	-2.72 (81)	0.005	0.502	0.155
Week 3	-2.23 (82)	-4.04 (78)	-2.49 (82)	-3.62 (68)	0.003	0.662	0.030
Week 4	-4.58 (42)	-5.55 (51)	-5.29 (41)	-4.61 (49)	0.195	0.365	0.967
Week 5	-5.89 (31)	-6.65 (45)	-5.60 (37)	-5.70 (40)	0.342	0.731	0.814
Week 6	-5.78 (30)	-7.53 (42)	-7.18 (34)	-6.40 (39)	0.040	0.113	0.478

The mean change from baseline to Week 6 in the PANSS Negative Sub-Scale Score was the second of two key secondary efficacy measures. Since the first key secondary measure showed significantly greater improvement compared with placebo for all the aripiprazole treatment groups, analysis of this key secondary measure was performed for all treatment groups at Week 6 versus placebo. The results of the analysis of the change in the PANSS Negative Sub-Scale Score for the LOCF data set at Week 6 is shown in Table 4.1.3.7. The analysis showed that all aripiprazole treatments had significantly greater improvement compared with the placebo group. The analysis of the change Scores for the LOCF data set for aripiprazole 10 mg was statistically significantly different from placebo from Week 1 through Week 6. Aripiprazole 15 mg was significantly different from placebo from Week 2 through Week 6. Aripiprazole 20 mg was statistically significantly different from placebo from Day 4 through Week 6.

The results of the analysis of the mean change from baseline for the OC data set is shown in Table 4.1.3.8. The aripiprazole 10-mg treatment group showed significantly greater improvement at Weeks 1 through 3, while aripiprazole 15 mg was not statistically significantly different from placebo at any week. Aripiprazole 20 mg showed significantly greater improvement compared with the placebo group at Day 4 through Week 3.

Table 4.1.3.7 Mean Change from Baseline in PANSS Negative Sub-Scale Total Score, LOCF Data Set, Efficacy Sample for Study CN138-001

	PANSS Negative Scale Total Score				Pairwise Comparisons P-values		
	Placebo N = 107	Aripiprazole 10 mg N = 103	Aripiprazole 15 mg N = 103	Aripiprazole 20 mg N = 97	Ari10 vs Placebo	Ari15 vs Placebo	Ari20 vs Placebo
Baseline	23.16	23.83	23.54	23.59	0.424	0.647	0.611
Day 4	0.10	-0.67	-0.58	-1.26	0.098	0.141	0.004
Week 1	-0.31	-1.65	-1.42	-1.93	0.022	0.059	0.007
Week 2	0.03	-2.41	-1.65	-2.50	0.001	0.022	0.001
Week 3	0.13	-2.80	-1.74	-2.72	<0.001	0.018	<0.001
Week 4	-0.05	-2.94	-2.32	-2.76	0.001	0.008	0.002
Week 5	0.12	-3.31	-2.22	-3.22	<0.001	0.006	<0.001
Week 6	0.08	-3.52	-2.65	-3.33	<0.001	0.002	<0.001

Table 4.1.3.8 Mean Change from Baseline in the PANSS Negative Subscale Total Score, OC Data Set, Efficacy Sample for Study CN138-001

	PANSS Negative Scale 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	22.65 (107)	23.39 (103)	23.37 (103)	23.31 (97)	0.455	0.467	0.511
Day 4	0.12 (100)	-0.68 (97)	-0.64 (100)	-1.27 (94)	0.094	0.110	0.004
Week 1	-0.27 (100)	-1.67 (89)	-1.26 (95)	-1.89 (87)	0.022	0.101	0.009
Week 2	-0.50 (88)	-2.84 (86)	-1.60 (89)	-2.57 (81)	0.004	0.170	0.012
Week 3	-1.13 (82)	-3.58 (78)	-2.06 (82)	-3.53 (68)	0.006	0.287	0.010
Week 4	-3.99 (42)	-5.29 (51)	-4.91 (41)	-5.02 (49)	0.247	0.437	0.362
Week 5	-5.32 (31)	-6.09 (45)	-4.83 (37)	-6.53 (40)	0.486	0.663	0.278
Week 6	-5.21 (30)	-7.37 (42)	-7.28 (34)	-6.89 (39)	0.075	0.102	0.170

Secondary Analyses

Additional secondary outcome measures were the PANSS Positive Sub-Scale Score, CGI Improvement Score Responder rates, CGI Severity Score, MADRS, and discontinuation rates. The results for the additional secondary outcome measures are shown in Table 4.1.3.9.

(Note: In the sponsor's original protocol, the secondary efficacy measures were only specified as the mean change from randomization to Week 6 (not all time points) in CGI Severity score, CGI global improvement score, PANSS-Positive Sub-Scale Total Score, PANSS-Negative Sub-Scale Total Score, and the percentage of responders. So this review only reports the results for the change from randomization to Week 6 for the above mentioned additional secondary outcome measures.)

As we can observe from the table, all the aripiprazole treatment groups showed significantly greater improvement compared to placebo in the change PANSS Positive Sub-Scale Score from randomization to Week 6 for the LOCF data set. The analysis of the mean change from baseline for the OC data indicates aripiprazole 10 mg had significantly greater improvement compared with the placebo group at Week 6, while aripiprazole 15 mg and 20 mg were not statistically significantly different from placebo at Week 6.

The analysis of the mean CGI Improvement Scores for the LOCF data set showed that all aripiprazole groups had significantly greater improvement compared with the placebo group at Week 6. However, none of the aripiprazole groups showed statistical significance in the analysis of the mean score for the OC data.

Response rates were analyzed by evaluating all responders, CGI (Improvement) responders, and PANSS responders. Responders are defined as patients who meet either of the following criteria:

- A rating of very much improved (1) or much improved (2) on the CGI Improvement Score, or
- At least a 30% decrease from baseline in the PANSS Total Score.

For the analysis of percentage of responders in the LOCF data, aripiprazole 10 mg and 20 mg showed significantly greater improvement compared with the placebo group at Weeks 6, while aripiprazole 15 mg was not. None of aripiprazole groups showed statistically significantly different from the placebo at Week 6 for the OC data.

For the analysis of the percentage of CGI (Improvement) responders, results analyzed on the LOCF data set showed that only Aripiprazole 20 mg group had significantly greater improvement compared with the placebo group at Week 6. No treatment groups were statistically significantly different from placebo for the OC data set.

For the analysis of the percentage of PANSS responders in the LOCF data, all aripiprazole groups showed significantly greater improvement compared with the placebo group at Weeks 6. In the analysis of the OC data set, none of treatment groups had statistically significant difference from the placebo.

For the analysis of mean change from baseline in the CGI severity of illness score, all aripiprazole groups showed statistically significantly different from the placebo at Week 6 in the LOCF data. However, none of aripiprazole groups did in the analyses for the OC data.

Since the administration of the MADRS was added to the study several months after study initiation per Amendment 2, a substantial number of patients were not administered the MADRS at either baseline or follow-up or both. The sponsor mentioned that although there was a trend toward significance for the aripiprazole 15-mg group in the LOCF data set at Week 6, no statistical conclusions may be drawn due to the small sample size.

One hundred forty-four patients discontinued the study due to lack of efficacy. This includes patients who discontinued from the trial due to lack of efficacy as well as patients who continued in the study on open-label treatment. Patients not responding at the end of Week 3, as evidenced by a CGI Improvement Score ≥4, were discontinued from blinded therapy and given open-label aripiprazole. The sponsor mentioned that a lower percentage of patients discontinued due to lack of efficacy in all aripiprazole treatment groups compared with placebo. This lower rate of discontinuation was statistically significant for the aripiprazole 10-mg and aripiprazole 20-mg groups.

Table 4.1.3.9 The Summary of Results for the Secondary Analyses for Study CN138-001
For the LOCF Data Set:

Endpoints	N	Baseline	Change from Baseline to Endpoint (i.e., week 6)	Treatment Difference vs. Placebo	95% CI for Difference	P-Value
PANSS Positive Sub-Scale Total Score						
Aripiprazole 10 mg	103	24.53	-4.98	-3.88	(-5.69, -2.08)	<0.001
Aripiprazole 15 mg	103	24.38	-3.81	-2.71	(-4.52, -0.90)	0.003
Aripiprazole 20 mg	97	24.20	-4.51	-3.41	(-5.25, -1.56)	<0.001
Placebo	107	24.47	-1.10			

Endpoints	Mean at Week 6	P-Value (vs. Placebo)
CGI Improvement Score		
Aripiprazole 10 mg (N=103)	3.33	0.004
Aripiprazole 15 mg (N=103)	3.42	0.006
Aripiprazole 20 mg (N=97)	3.31	0.006
Placebo (N=107)	4.00	

Endpoints	Number (Percent) at Week 6	P-Value (vs. Placebo)
Percentage of Responders		
Aripiprazole 10 mg (N=103)	42 (41)	0.038
Aripiprazole 15 mg (N=103)	36 (35)	0.165
Aripiprazole 20 mg (N=97)	44 (45)	0.005
Placebo (N=107)	28 (26)	

Endpoints	Number (Percent) at Week 6	P-Value (vs. Placebo)
Percentage of CGI Responders		
Aripiprazole 10 mg (N=103)	35 (34)	0.134
Aripiprazole 15 mg (N=103)	32 (31)	0.219
Aripiprazole 20 mg (N=97)	41 (42)	0.005
Placebo (N=107)	25 (23)	

Endpoints	Number (Percent) at Week 6	P-Value (vs. Placebo)
Percentage of PANSS Responders		
Aripiprazole 10 mg (N=103)	31 (30)	0.002
Aripiprazole 15 mg (N=103)	26 (25)	0.028
Aripiprazole 20 mg (N=97)	25 (26)	0.025
Placebo (N=107)	14 (13)	

Endpoints	N	Baseline	Change from Baseline to Endpoint (i.e., week 6)	Treatment Difference vs. Placebo	95% CI for Difference	P-Value
CGI Severity of Illness Score						
Aripiprazole 10 mg	103	4.79	-0.65	-0.47	(-0.77, -0.18)	0.002
Aripiprazole 15 mg	103	4.79	-0.51	-0.33	(-0.63, -0.04)	0.028
Aripiprazole 20 mg	96	4.68	-0.64	-0.46	(-0.76, -0.16)	0.003
Placebo	107	4.64	-0.18			

For the OC Data Set:

Endpoints	Baseline & (N)	Change from Baseline to Endpoint & (N)	Treatment Difference vs. Placebo	95% CI for Difference	P-Value
PANSS Positive Sub-Scale Score					
Aripiprazole 10 mg	24.47 (103)	-10.22 (42)	-2.47	(-4.75, -0.19)	0.034
Aripiprazole 15 mg	24.54 (103)	-9.87 (34)	-2.13	(-4.52, 0.26)	0.081
Aripiprazole 20 mg	24.28 (97)	-8.51 (39)	-0.77	(-3.09, 1.55)	0.513
Placebo	24.34 (107)	-7.74 (30)			

Endpoints	Mean at Week 6 (N)	P-Value (vs. Placebo)
CGI Improvement Score		
Aripiprazole 10 mg	1.90 (42)	0.336
Aripiprazole 15 mg	1.85 (34)	0.247
Aripiprazole 20 mg	2.10 (39)	0.909
Placebo	2.13 (30)	

Endpoints	Number Responding/N at Week 6 (%)	P-Value (vs. Placebo)
Percentage of Responders		
Aripiprazole 10 mg	35/42 (83)	0.484
Aripiprazole 15 mg	29/34 (85)	0.381
Aripiprazole 20 mg	33/39 (85)	0.406
Placebo	23/30 (77)	

Endpoints	Number Responding/N at Week 6 (%)	P-Value (vs. Placebo)
Percentage of CGI Responders		
Aripiprazole 10 mg	32/42 (76)	0.560
Aripiprazole 15 mg	28/34 (82)	0.248
Aripiprazole 20 mg	32/39 (82)	0.243
Placebo		

Endpoints	Number Responding/N at Week 6 (%)	P-Value (vs. Placebo)
Percentage of PANSS Responders		
Aripiprazole 10 mg	25/42 (60)	0.178
Aripiprazole 15 mg	20/34 (59)	0.220
Aripiprazole 20 mg	20/39 (51)	0.515
Placebo	13/30 (43)	

Endpoints	Baseline & (N)	Change from Baseline to Endpoint & (N)	Treatment Difference vs. Placebo	95% CI for Difference	P-Value
CGI Severity of Illness Score					
Aripiprazole 10 mg	4.80 (103)	-1.60 (42)	-0.18	(-0.62, 0.27)	0.435
Aripiprazole 15 mg	4.83 (103)	-1.47 (34)	-0.05	(-0.52, 0.41)	0.820
Aripiprazole 20 mg	4.72 (96)	-1.40 (39)	0.02	(-0.43, 0.47)	0.924
Placebo	4.64 (107)	-1.42 (30)			

Endpoints	Number (Percent)	P-Value (vs. Placebo)
Rate of Discontinuation		
Aripiprazole 10 mg (N=103)	28 (27)	0.005
Aripiprazole 15 mg (N=103)	36 (35)	0.140
Aripiprazole 20 mg (N=97)	31 (32)	0.026
Placebo (N=107)	49 (46)	

Subgroup Analysis

Subgroup analyses were performed by gender on the PANSS Total Score and is shown in Table 11 of the Appendix. The sponsor did not make any comment about the results of this analysis.

4.1.3.5 The Sponsor's Overall Efficacy Conclusions

All three fixed doses of aripiprazole were shown to be effective in the treatment of patients with schizophrenia in acute relapse based on the predefined primary and key secondary endpoints of the PANSS Total Score, PANSS-derived BPRS Core Score, and PANSS Negative Sub-Scale Score.

4.2 Phase II Studies

4.2.1 Study 31-93-202

4.2.1.1 Disposition of Patients

A total of 103 patients were randomized into this study: 34 patients in the OPC-14597 group, 34 in the haloperidol group, and 35 patients in the placebo group. A total of 53 patients completed the study: 21 patients in the OPC-14597 group, 20 patients in the haloperidol group, and 12 patients in the placebo group.

4.2.1.2 Demographics and Patient Characteristics

Baseline demographics were determined at the screening visit and included sex, age, weight and race. Of the 103 patients randomized, there were more males (n=91) than females (n=12). There was also a slightly higher number of Caucasians than Blacks, Hispanics or Other, with the majority of the patients being Caucasian (n=54) or Black (n=44). Distribution was generally equivalent across all treatment groups for race. Mean ages and weights by sex were also equivalent across traetment groups with the exception of mean weight for the female haloperidol group. Table 4.2.1.1 presents a summary of patient demographics across all treatment groups.

Table 4.2.1.1 Demographic Characteristics- All Randomized Patients for Study 31-93-202

		OPC-14597		Haloperidol		Placebo	
		Male	Female	Male	Female	Male	Female
Age (years)	N	32	2	30	4	29	6
	Mean	32.4	42.5	38.6	38.8	36.9	42.5
	Min	18	37	21	26	21	31
	Max	57	48	65	46	52	59
Weight (kg)	N	32	2	30	4	29	6
	Mean	84.1	67.4	81	86.6	82.1	65.8
	Min	52.2	57.7	59.5	55.8	50.8	50.4
	Max	158.9	77.2	129.8	116.7	118.5	97.6
Race	Caucasian	18	1	15	2	15	3
	Black	13	1	13	2	12	3
	Hispanic	0	0	1	0	1	0
	Other	1	0	1	0	1	0

4.2.1.3 The Sponsor's Efficacy Results

Primary Efficacy Variables

The primary efficacy variables were 1) change from baseline to last visit in BPRS-total score and 2) a response indicator variable defined by a reduction of at least one point from baseline to last visit in CGI-severity score.

Table 4.2.1.2 shows the sponsor's analysis results for the mean change from baseline in BPRS-total score for each treatment week with p-values for each treatment group and also Table 4.2.1.3 shows the treatment effects (subtracting placebo effect) of OPC-14597 and haloperidol at the last visit. As it was shown in the table, in the OPC-14597 group, improvement in BPRS-total score appeared prominently after Week 2, with a mean decrease of 8.5 points in total score from baseline to Week 3, which continued throughout the remaining treatment period, with a mean decrease of 10.3 points in total score at Week 4. The analyses of last visit results (LOCF), which included data from patients who discontinued the study, also showed an improvement in the BPRS-total score, with a mean decrease of 7.2 points in total score from baseline. In addition, as shown in the table, the superiority of OPC-14597 over placebo with regard to change from baseline to last visit for BPRS-total score was demonstrated with an estimated treatment difference of 6.25 points (p=0.0142). In addition, the superiority of OPC-14597 over placebo with regard to change from baseline to last visit for BPRS-total score was demonstrated with an estimated treatment difference of 6.25 points.

Table 4.2.1.2 BPRS-Total Score- Mean Change from Baseline and p-Values by Week-Observed Cases for Study 31-93-202

Treatment Group	Baseline		Week 1		Week 2		Week 3		Week 4		Last Visit (LOCF)	
	n	Mean	n	Mean	n	Mean	n	Mean	n	Mean	n	Mean
OPC-14597	33	53.0	32	-1.6	28	-2.3	22	-8.5	21	-10.3	33	-7.2
Haloperidol	33	50.3	33	-6.5	30	-6.9	22	-7.4	20	-9.0	33	-8.1
Placebo	35	50.0	35	-3.1	28	-3.6	20	-5.1	14	-9.9	35	-2.1

	2 Sided p-values for Pair-Wise Comparison					
OPC-14597 vs. Placebo	0.1732	0.2370	0.5160	0.1718	0.8863	0.0142
Haloperidol vs. Placebo	0.8939	0.0791	0.1607	0.1891	0.4687	0.0083

Table 4.2.1.3 Treatment Effect Based on the Last Visit Efficacy Analysis BPRS-Total Score for Study 31-93-202

	Estimated Treatment Effect	p-Value	Lower 95% Confidence Limit	Upper 95% Confidence Limit
OPC-14597 vs. Placebo	-6.25	0.0142	-11.21	-1.29
Haloperidol vs. Placebo	-6.41	0.0083	-11.21	-1.70

To evaluate responder rates based on CGI-severity score, as shown in Table 4.2.1.4, both OPC-14597 and haloperidol showed a statistically significant (p=0.035 and p=0.003, respectively) responder rates with 42.4% of the OPC-14597 patients responding to treatment and 54.5% of the haloperidol patients responding to treatment. The placebo patient group had 20% of the response rate.

Table 4.2.1.4 Responder Rates Based on at Least One Point Improvement from Baseline at Last Visit in CGI-Severity Score

	CGI-Severity Score				
Treatment Group	n	No. of Responders	% of Responders	Treatment Comparison	p-Value
OPC-14597	33	14	42.4	OPC-14597 vs. Placebo	0.035
Haloperidol	33	18	54.5	Haloperidol vs. Placebo	0.003
Placebo	35	7	20.0		

Secondary Efficacy Variables

PANSS-total score was based on the severity rating for positive and negative symptoms of schizophrenia and general psychopathology, with a lower score indicating less severe symptoms and a reduction in score over time indicating improvement. As shown in Table 4.2.1.5, improvement in PANSS-total score appeared prominently at Week 3, with a mean decrease of 14.0 points from baseline, and which continued further with a mean decrease of 16.4 points from baseline, and which continued further with a mean decrease of 16.4 points at Week 4. The LOCF analyses, which included data from patients who discontinued the study, also showed an improvement in PANSS-total score, with a mean decrease of 11.1 points from baseline. In addition, as shown in Table 4.2.1.6, the superiority of OPC-14597 over placebo with regard to change from baseline to last visit for PANSS-total score was demonstrated with an estimated treatment difference of 12.01 points (p=0.0080).

Table 4.2.1.5 PANSS-Total Score- Mean Change from Baseline and p-Values By Week - Observed Cases for Study 31-93-202

Treatment	Baseline		Week 1		Week 2		Week 3		Week 4		Last Visit (LOCF)	
Group	n	Mean	n	Mean	n	Mean	n	Mean	n	Mean	n	Mean
OPC-14597	33	91.8	32	-3.1	29	-5.1	22	-14.0	22	-16.4	33	-11.1
Haloperidol	33	89.0	33	-10.8	30	-14.4	22	-14.0	20	-17.1	33	-15.8
Placebo	35	86.5	35	-1.9	27	-4.3	20	-7.0	13	-15.2	35	-1.1

	2 Sided p-values for Pair-Wise Comparison					
OPC-14597 vs. Placebo	0.1742	0.9157	0.9967	0.0879	0.6763	0.0080
Haloperidol vs. Placebo	0.4782	0.0137	0.0252	0.0499	0.2574	0.0004

Table 4.2.1.6 Treatment Effect Based on the Last Visit Efficacy Analysis-PANSS Total Score for Study 31-93-202

	Estimated Treatment Effect	p-Value	Lower 95% Confidence Limit	Upper 95% Confidence Limit
OPC-14597 vs. Placebo	-12.01	0.008	-20.79	-3.24
Haloperidol vs. Placebo	-15.62	0.0004	-24.02	-7.22

PANSS-negative sub-scale score was based on the severity rating for negative symptoms of schizophrenia included in the PANSS-total score, with a lower score indicating less severe symptoms and a reduction in score over time indicating improvement. As shown in Table 4.2.1.7, improvement in PANSS negative sub-scale score appeared prominently at Week 3 with a mean decrease of 4.4 points at the Week 4 visit for the OPC-14597 group. In the weekly analysis, OPC-14597 demonstrated a clear trend of improving the negative symptoms of the disease as measured by the PANSS-negative score. The mean change from baseline under the LOCF analysis was a decrease of 2.8 points. As shown in Table 4.2.1.8, OPC-14597 showed a trend towards superiority over placebo with regard to change from baseline to last visit for PANSS-negative sub-scale score with an estimated treatment effect of 2.71 points (p=0.0642).

Table 4.2.1.7 PANSS-Negative Sub-Scale Score-Mean Change from Baseline and p-Value by Week-Observed Cases for Study 31-93-202

Treatment Group	Baseline		Week 1		Week 2		Week 3		Week 4		Last Visit (LOCF)	
	n	Mean	n	Mean	n	Mean	n	Mean	n	Mean	n	Mean
OPC-14597	33	23.6	32	-1.3	29	-1.7	22	-4.3	22	-4.4	33	-2.8
Haloperidol	33	22.3	33	-2.0	30	-2.9	22	-3.0	20	-3.3	33	-3.4
Placebo	35	22.2	35	0.3	27	-0.9	20	-1.9	13	-5.2	35	-0.9

	2 Sided p-values for Pair-Wise Comparison					
OPC-14597 vs. Placebo	0.3712	0.2675	0.7373	0.0949	0.7438	0.0642
Haloperidol vs. Placebo	0.8519	0.0370	0.1262	0.2408	0.8343	0.0258

Table 4.2.1.8 Treatment Effect Based on the Last Visit Efficacy Analysis-PANSS Negative Sub-Scale Score for Study 31-93-202

	Estimated Treatment Effect	p-Value	Lower 95% Confidence Limit	Upper 95% Confidence Limit
OPC-14597 vs. Placebo	-2.71	0.0642	-5.58	0.16
Haloperidol vs. Placebo	-3.11	0.0258	-5.84	-0.39

4.2.1.4 The Sponsor's Overall Efficacy Conclusions

- OPC-14597 showed statistically significant superiority over placebo in reducing the signs and symptoms of schizophrenia in all illness severity scores as measured by BPRS-total, BPRS-core, CGI-severity, CGI-improvement, and PANSS-total, with efficacy being seen prominently after 2 weeks of treatment and continuing throughout the remainder of the study. This may be attributed to dose escalation in the first two weeks to reach maximum dose.
- OPC-14597 was superior to placebo and comparable to haloperidol with regard to responder rates based on at least one point reduction in the CGI-severity score from baseline to last visit, a 30% reduction in BPRS-total score from baseline to last visit, or

a score of one or two in the CGI-improvement score at last visit in schizophrenic patients.

- Although not-statistically significant, OPC-14597 demonstrated a clear trend of improving negative symptoms of schizophrenia based on PANSS-negative score.
- Patients in the haloperidol group showed improvement in psychosis, confirming that the patient population of this study was responsive to active treatmetn.

4.2.2 Study 31-94-202

4.2.2.1 Disposition of Patients

A total of 176/307 (57.3%) patients completed the study: 37/59 (62.7%) patients in the OPC-2 mg group, 35/60 (58.3%) patients in the OPC-10mg group, 41/61 (67.2%) patients in the OPC-30 mg group, 34/63(54%) patients in the haloperidol group, and 29/64 (45.3%) patients in the placebo group. Patients in the OPC-14597 group, particularly patients in the OPC-30mg group, completed the study at a higher rate (58.3-67.2%) compared to patients in the haloperidol group (54%) and the placebo group (45.3%).

4.2.2.2 Demographics and Patient Characteristics

Table 4.2.2.1 shows the patients' demographic characteristics. Treatment groups were generally comparable for demographic characteristics. Patients were primarily male (247/307, 80.5%) with about one fifth of the patients female (60/307, 19.5%). Mean age ranged from 37.2 to 40.1 years (range: 18-65 years) in males and from 38.8 to 43.2 years (range: 19-63) in females across treatment groups. About half of the patients were Caucasian (159/307, 51.8%) with the rest being black (115/307, 37.5%), Hispanic (24/307, 7.8%), Asian (3/307, 1.0%) and other (6/307, 2.0%). Mean weight ranged from 79.7 to 86.4 kg in males and 68.8-79.1 in females caross treatment groups.

Table 4.2.2.1 Demographic Characteristics- All Randomized Patients for Study 31-94-202

Demographic Characteristics		OPC-14597 mg/day						Haloperidol mg/day		Placebo	
		2 mg		10 mg		30 mg		10 mg			
		M	F	M	F	M	F	M	F	M	F
Age (years)	N	47	12	49	11	46	15	52	11	53	11
	Mean	40.1	38.8	37.2	40.6	38.8	38.9	38.0	43.2	37.5	40.5
	Min	22	19	18	23	18	24	19	25	19	28
	Max	65	51	64	56	61	57	60	63	57	55
Weight (kg)	N	47	12	49	11	46	15	52	11	53	10
	Mean	83.3	77.1	82.9	68.8	79.7	78.6	82.9	79.1	86.4	75.2
	Min	53.1	50.8	54	40.9	54.5	55.4	55.8	62.2	53.6	50.4
	Max	137.1	133.5	129.4	96.7	143	103.5	141.2	101.2	168.0	101.2
Race	Caucasian	22	11	20	6	19	11	28	9	28	5
	Black	19	1	19	5	23	3	19	2	20	4
	Hispanic	5	0	7	0	2	0	4	0	5	1
	Asian	0	0	1	0	1	1	0	0	0	0
	Other	1	0	2	0	1	0	1	0	0	1

4.2.2.3 The Sponsor's Efficacy Results

Primary Efficacy Variable (after excluding Center 003)

The principal investigator at Center 003, Richard L. Borison, M.D., had his employment terminated by the Augusta Veterans Affairs medical Center on June 7, 1996 due to allegations of research misconduct, so the analysis results for this study should be based on the data after excluding the center 003.

Table 4.2.2.2 shows the study results for the primary efficacy variables after excluding Center 003. As it was shown in the table, superiority of the OPC-30 mg group versus placebo ($p<0.05$) was demonstrated at last visit for the primary efficacy variable CGI-improvement after excluding Center 003. This treatment difference was also statistically significant after correction for multiple comparison by Dunnett's method at the two-tailed 0.05 level. The superiority of OPC-30mg over placebo with regard to change from baseline to last visit for BPRS-core score was not demonstrated after excluding Center 003. Significant differences were noted in the comparison of haloperidol versus placebo at last visit for BPRS-Core and trends towards significance for CGI-improvement ($p=0.0811$) after excluding Center 003.

Table 4.2.2.2 Treatment Effects (Last Visit Analysis) of Primary Efficacy Variables Excluding Center 003

Variable	Treatment Comparison	Estimated Treatment Effect	Value of t statistic	P-value	Lower 95% CL	Upper 95% CL
BPRS-core	OPC-14597: 2mg vs. Placebo	-0.31	-0.38	0.7034	-1.94	1.31
	OPC-14597: 10 mg vs. Placebo	-0.11	-0.13	0.8939	-1.75	1.53
	OPC-14597: 30 mg vs. Placebo	-1.29	-1.58	0.1165	-2.89	0.32
	Haloperidol: 10 mg vs. Placebo	-1.61	-1.97	0.0495	-3.22	-0
CGI-Improvement	OPC-14597: 2mg vs. Placebo	-0.15	-0.55	0.5860	-0.69	0.39
	OPC-14597: 10 mg vs. Placebo	-0.33	-1.21	0.2260	-0.87	0.21
	OPC-14597: 30 mg vs. Placebo	-0.75	-2.80	0.0055	-1.29	-0.22
	Haloperidol: 10 mg vs. Placebo	-0.47	-1.75	0.0811	-1.00	0.06

Secondary Efficacy Variables (after excluding Center 003)

Treatment effects for the secondary efficacy variables based on the last visit efficacy analysis excluding Center 003 are summarized in Table 4.2.2.3.

Superiority of the OPC-30mg group versus placebo ($p<0.05$) was demonstrated at last visit for secondary efficacy variables BPRS-total, and PANSS-total. A trend towards superiority of OPC-30 mg versus placebo was noted for PANSS-negative ($p=0.0817$). Trends toward significance were also noted in the comparison of haloperidol versus placebo for PANSS-total ($p=0.0733$) after excluding Center 003.

Table 4.2.2.3 Treatment Effects (Last Visit Analysis) of Secondary Efficacy Variables Excluding Center 003

Variable	Treatment Comparison	Estimated Treatment Effect	P-value	Lower 95% CL	Upper 95% CL
BPRS-total	OPC-14597: 2mg vs. Placebo	-3.09	0.1703	-7.52	1.34
	OPC-14597: 10 mg vs. Placebo	-3.12	0.1675	-7.56	1.32
	OPC-14597: 30 mg vs. Placebo	-6.32	0.0048	-10.69	-1.94
	Haloperidol: 10 mg vs. Placebo	-3.27	0.1415	-7.64	1.10
PANSS-total	OPC-14597: 2mg vs. Placebo	-4.89	0.1849	-12.13	2.35
	OPC-14597: 10 mg vs. Placebo	-5.52	0.1357	-12.78	1.74
	OPC-14597: 30 mg vs. Placebo	-10.66	0.0037	-17.82	-3.50
	Haloperidol: 10 mg vs. Placebo	-6.53	0.0733	-13.67	0.62
PANSS-negative	OPC-14597: 2mg vs. Placebo	-0.70	0.4947	-2.70	1.31
	OPC-14597: 10 mg vs. Placebo	-1.13	0.2680	-3.14	0.88
	OPC-14597: 30 mg vs. Placebo	-1.76	0.0817	-3.74	0.22
	Haloperidol: 10 mg vs. Placebo	-0.43	0.6663	-2.41	1.55

4.2.2.4 The Sponsor's Overall Efficacy Conclusions

- In general, all three dose groups of OPC-14597 (2, 10 or 30 mg/day) were superior to placebo in the treatment of psychosis. Among the three OPC-14597 doses, 30 mg can be distinguished from the other two doses with respect to efficacy
- While no definitive conclusions can be drawn, the results with the 30 mg dose of OPC-14597 are suggestive of an early onset (Week 1) of treatment effect.
- Of all the treatment groups, only the OPC-14597 30 mg dose was found to show significant improvement in the negative symptoms of psychosis.
- OPC-14597 was found to be most effective at a dose of 30 mg/day, in a 4-week duration, for the treatment of schizophrenic patients.
- The patients of the haloperidol group showed improvement in psychosis, which confirmed that the patient population of this study was responsive to an active treatment

(**Note:** The sponsor did not mention but it is clearly that they made the above conclusions based on the whole data. Since the data from Center #3 were invalid, these conclusions were not accurate.)

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

4.3.1 Disposition of Patients

A total of 1452 patients signed the informed consent form; 158 of these patients filed screening and did not enter the placebo washout phase. The remaining 1294 patients underwent placebo washout and were randomized to receive double-blind treatment; 433 to the haloperidol group and 861 to the aripiprazole group. The completion rate was significantly higher for patients on aripiprazole (43%) compared with those on haloperidol (30%). This difference was primarily due to the lower rate of discontinuation for adverse events other than worsening schizophrenia. The disposition of all enrolled patients is

presented in Table 12 of the appendix. The time to discontinuation due to all reasons for the Randomized Sample is presented by treatment group in Figure 4 of the appendix.

4.3.2 Data Sets

The distribution of patients within each of the patient samples is presented by treatment group for all randomized patients in Table 4.3.1.

Table 4.3.1 Number of Patients in Samples for Studies 31-98-217 and 31-98-304-01

Sample	Haloperidol	Aripiprazole	Total
Randomized	433	861	1294
Safety	431	859	1290
Efficacy	430	853	1283

Four of the 1294 randomized patients (two from the haloperidol group and two from the aripiprazole group) were excluded from the Safety Sample because they did not receive study medication according to the dosing record.

4.3.3 Demography and Patient Characteristics

Demographic characteristics are presented by treatment group in Table 4.3.2 for patients in the Randomized Sample. According to the table, the treatment groups were comparable with respect to age, gender, race and weight.

Table 4.3.2 Demographic Characteristics for the Randomized Sample for Studies 31- 98-217 and 31-98-304-01

Variable		Haloperidol N = 433	Aripiprazole N = 861	Total N = 1294
Age (years)	Mean	36.8	37.3	37.1
	Median	36	36	36
	Min-Max	18 - 63	18 - 65	18 - 65
	S.E.	0.5	0.4	0.3
Gender N (%)	Men	247 (57)	511 (59)	758 (59)
	Women	186 (43)	350 (41)	536 (41)
Race N (%)	White	378 (87)	733 (85)	1111 (86)
	Black	41 (10)	99 (11)	140 (11)
	Hispanic	3 (1)	7 (1)	10 (1)
	Asian/Pacific Islander	2 (1)	4 (1)	6 (1)
	Other	9 (2)	18 (2)	27 (2)
Weight (kg)	Mean	73.1	74.5	74.0
	Median	71	72	72
	Min-Max	38 - 153	36 - 143	36 – 153
	S.E.	0.8	0.6	0.5
	Missing	0	3	3

4.3.4 The Sponsor's Efficacy Results

Table 4.3.3 shows the summary results of the primary and supportive efficacy endpoints and Table 4.3.4 shows the summary results of the secondary efficacy endpoints.

Table 4.3.3 Summary of Primary and Supportive Endpoint Efficacy Results for the Randomized Sample for Studies 31-98-217 and 31-98-304-01

Variable	Haloperidol	Aripiprazole	P-value
Number Randomized Patients	433	861	
Umber of Patients in Efficacy Sample	430	853	
Number (%) Responders	298 (69%)	610 (72%)	0.362
Time to Failure to Maintain Response in Responders	Relative Risk (95% CI)		
Treatment (Aripiprazole: Haloperidol)	0.881 (0.645 - 1.204)		0.427
Proportion of Patients Maintaining Response [% (S.D.)]			
Week 8	93% (1.5%)	92% (1.1%)	
Week 26	81% (2.6%)	84% (1.6%)	
Week 52	73% (3.1%)	77% (1.8%)	
Time to Failure in All Patients	Relative Risk (95% CI)		
Treatment (Aripiprazole: Haloperidol)	0.858 (0.721 - 1.021)		0.084
Proportion of Patients not yet Failed [% (S.D.)]			
Week 8	69% (2.3%)	71% (1.6%)	
Week 26	56% (2.6%)	60% (1.7%)	
Week 52	49% (2.7%)	54% (1.8%)	
Proportion of Patients On-treatment and Still in Response [N%]			
Week 8	192 (44%)	449 (52%)	0.005
Week 26	145 (33%)	380 (44%)	<0.001
Week 52	117 (27%)	343 (40%)	<0.001

Table 4.3.4 Summary of Rating Scale Secondary Efficacy Results for the Efficacy Sample by the LOCF for Studies 31-98-217 and 31-98-304-01

Variable	Haloperidol N=430		Aripiprazole N=853
PANSS Total Score			
Mean Baseline	94.7		95.1
Change at Week 8	-20.9		-21.8
95% CI for treatment effect		(-3.41, 1.85)	
P-value		0.560	
Change at Week 26	-20.7		-22.2
95% CI for treatment effect		(-4.27, 1.48)	
P-value		0.341	
PANSS Negative Sub-Scale Score			
Mean Baseline	24.7		24.7
Change at Week 8	-4.2		-4.7
95% CI for treatment effect		(-1.15, 0.14)	
P-value		0.126	
Change at Week 26	-4.4		-5.1
95% CI for treatment effect		(-1.52, -0.08)	
P-value		0.029	
MADRS Total Score			
Mean Baseline	12.8		12.5
Mean at Week 8	-2.6		-3.4
95% CI for treatment effect		-1.74, -0.11	
P-value		0.027	
Change at Week 26	-2.0		-2.9
95% CI for treatment effect		(-1.95, -0.15)	
P-value		0.022	

4.3.4.1 Primary Efficacy Endpoint

Time to failure to maintain response was analyzed only for the responders. Definitions of "response" and "failure to maintain response" can be found in Section 3.3.3. of this review. Worsening schizophrenia was defined by the modified COSTART dictionary terms "psychosis" and "schizophrenic reaction".

Of the 853 patients in the Efficacy Sample that were randomized to aripiprazole, 610 (72%) met the criteria to be classified as responders. Of the 430 patients in the Efficacy Sample that were randomized to haloperidol, 298 (69%) were considered responders.

Out of these responders, the proportion of patients who did not experience failure by Weeks 8, 26, and 52 in summarized in Table 4.3.3. The relative risk for failure for the aripiprazole arm was 88% (95% CI: 65% - 120 %) of that for the haloperidol arm (p=0.4271). It indicated that the risk of failing to maintain response in the aripiprazole group was 12% lower than that of haloperidol.

4.3.4.2 Supportive and Efficacy Endpoints

A numerically greater percentage of randomized patients in the aripiprazole group (54%) had not failed by Week 52 when compared with the haloperidol group (49%). In the analysis of time to failure for all randomized patients, the estimated relative risk (aripiprazole: haloperidol) was 0.858 (95% CI: 0.721, 1.021) indicating that the patients in the aripiprazole group had a 14% lower risk of failure compared to the haloperidol group. This result had a trend towards statistical significance (p=0.084).

A significantly greater percentage of patients randomized to aripiprazole compared to patients randomized to haloperidol who remained on treatment and were in response. This was evaluated at three time points, Week 8 (p=0.005), Week 26 (p<0.001) and Week 52 (p<0.001).

4.3.4.3 Secondary Efficacy Endpoints

Aripiprazole was statistically superior to haloperidol as determined by the time to discontinuatin due to either lack of response to study drug or adverse event (p<0.0010. The risk ratio for this event was 0.692 (95% CI: 0.573 – 0.837) indicating that the risk of discontinuation due to either lack of response to study drug or adverse event was 31% lower for the aripiprazole treated patients relative to the patients treated with haloperidol.

For other secondary time-to-event variables: time to first response (all randomized patients), time to discontinuation due to lack of response to study drug (all randomzied patients), and time from first response to failure to maintain response (responders only), no statistically significant differences were observed between the two treatment groups.

Aripiprazole showed significant improvement over haloperidol in the treatment of negative and depressive symptoms. The improvement in treatment of negative symptoms

was demonstrated by significant differences in the comparison of mean change from baseline in the PANSS Negative Sub-Scale Score at Weeks 26 (p=0.029) and 52 (p=0.011) based on the LOCF data set. The improvemet in treatment of depressive symptoms was demonstrated by statistical differences in the comparison of mean change from baseline in MADRS Total Score at Weeks 8 (p=0.027), 26 (p=0.022), and 52 (p=0.031) (LOCF data set).

No significant differences were observed between treatments in mean change from Baseline in PANSS Total Score, PANSS Positive Sub-Scale Score, CGI Severity of Illness Score, or in mean CGI Improvement Score.

4.3.5 The Sponsor's Efficacy Conclusions

The results from analyses of the primary and supportive efficacy measures demonstrate that aripiprazole was able to provide long-term maintenance therapy to patients who were initially in acute relapse that was similar or superior to the long-term maintenance effects of haloperidol.

- The overall estimated risk ratio (0.881) for failure to maintain response in responders favored aripiprazole, however, this improvement was not statistically significant.
- In the analysis of time to failure in all patients, the estimated relative risk of 0.858 favored aripiprazole and exhibited a trend toward statistical significance (p=0.084).
- Among all randomized patients, a significantly greater percentage of patients treated with aripiprazole demonstrated response at Weeks 8, 26 and 52.

5. Statistical Reviewer's Findings and Comments

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

1. Three primary efficacy endpoints were prospectively specified for Studies 31-97-201 and 31-97-202, but the sponsor did not clearly address either in the protocols or study reports what their decision rules were for these studies. It was indeed mentioned in the protocols and study reports that "The treatment comparisons will be tested by following the step-down procedure, i.e., first aripiprazole 30 mg vs. placebo will be tested at two-tailed 0.05 level; if rejected, aripiprazole 15 mg (or 20 mg for Study 31-97-202) vs. placebo will be tested at two-tailed 0.05 level."

   Now that the sponsor wished to use α=0.05, without any adjustment for testing the results for each primary endpoint of Studies 31-97-201 and 31-97-202, in order to protect the overall type I error rate of 0.05, it was judged by the statistical reviewer that winning on all three primary efficacy endpoints is necessary for claiming a positive study.

2. When the three pivotal phase III studies were evaluated, most of values can be reproduced by this reviewer. There was no inconsistent finding between the reviewer and the sponsor.

3. For Study 31-97-201, an internal audit revealed that data generated at Study Centers 007 and 001 could not be validated, so the sponsor performed the sensitivity analysis of the mean change from baseline for the PANSS Total Score by excluding the 19 patients randomized at these centers. They showed the results in the study report and concluded that they were consistent with those of the overall analysis. This reviewer checked their results and further performed the same kind of sensitivity analyses for the other two primary endpoints. The results did not show much difference to affect the conclusions on the overall analyses for either LOCF and OC data sets.

4. According to Tables 4.1.1.5 and 4.1.2.5, the sponsor had statistically significant results shown on all three primary efficacy endpoints for the LOCF data sets for Studies 31-97-201 and 31-97-202. However, this reviewer noticed that for Study 31-97-202 except the comparison between aripiprazole 20 mg and the placebo on the PANSS Positive Sub-Scale Score (p=0.045), the sponsor had p-values greater than 0.05 for the OC data analyses. So, the dropout cohort analyses were studied to see if the results for the LOCF or OC data analyses were biased. Notice that dropout cohorts were formed by patients that had their last primary efficacy measurement in the same week interval.

   Figures 5.1 to 5.3 showed us the PANSS total score over time for different dropout cohorts from the sponsor. This reviewer confirmed their results. The average changes of PANSS Total Scores from the baseline to each study week in which the patients dropped out the study right after were reported in Tables 5.1 below.

Table 5.1 Average Changes of PANSS Total Score for Dropout Cohort Analyses for Study 31-97-202

Group	Week 1 (n)	Week 2 (n)	Week 3 (n)	Week 4 (n)
Placebo	12.615 (26)	6.287 (21)	-4.1496 (4)	-18.2 (52)
Risperidone 6 mg	-0.873 (11)	-7.302 (16)	8.8141 (7)	-22.7 (61)
Aripiprazole 20 mg	5.265 (20)	-8.308 (9)	2.3 (8)	-23.4 (61)
Aripiprazole 30 mg	5.907 (16)	6.4 (7)	-2.24 (5)	-20.1 (68)

Carefully observing Table 5.1 and Figures 5.1, this reviewer noticed that the average change of PANSS Total Score for the placebo group patients at Week 1 was much bigger than the rest of treatment groups. It tells us that these patients had worse results. Moreover, most of dropout patients in the study were happening in the early two weeks. The placebo group had more dropout patients than the other treatment groups.

With almost 25% of patients dropping out after the first week' evaluation, the bad values carried from the dropout patients in the placebo group at Week 1 could make a difference at the LOCF analyses, especially, in the situation that the placebo patients had improvement as the study continued. On the other hand, with more poorly performed patients dropping out from the placebo group than the other treatment groups, the OC analyses may be biased against the treatment groups.

To investigate the influence of these 26 placebo group patients who dropped out before the second week of the study, this reviewer calculated the unadjusted mean of changes from baseline to Week 1 for the rest of placebo group patients. It was found to

be –7.13. Comparing this value with the OC results (see Table 5.3) after Week 1 (i.e., -9.0 at Week 2, -15.5 at Week 3 and –18.2 at Week 4), the fact that the patients in the placebo group also had improvement as the study continued was confirmed. Moreover, this value of mean change was much closer to the OC values for aripiprazole 20mg and 30mg groups at Week 1. Similarly, this reviewer also calculated the unadjusted mean of changes by Excluding the 47 patients who dropped out before Week 3. The calculated value –14.732 was also much closer to the OC values for aripiprazole 20mg and 30mg groups at Week 2. This tells us that these dropout patients did have worse responses than the average. Therefore, this reviewer suspected that the results from the LOCF analyses and OC analyses for the PANSS total score were both biased.

The other two primary endpoints: changes on the PANSS Positive Sub-Scales and changes on CGI Severity of Illness Score had similar problems. Table 5.4 and 5.5 show the unadjusted means of changes from baseline to each study week for the OC data Sets and Table 5.6 and 5.7 the average changes of scores for dropout cohort analyses.

Table 5.3 Unadjusted Mean Change from Baseline in PANSS Total Score for OC Data Set in Efficacy Sample for Study 31-97-202

		PANSS Total Score							
		Placebo		Risperidone 6 mg		Aripiprazole 20 mg		Aripiprazole 30 mg	
Variable	Week	N	Mean	N	Mean	N	Mean	N	Mean
Mean Baseline		103	95.0	95	93.6	98	94.0	96	92.3
Mean Change	1	102	-2.2	95	-8.0	96	-8.8	95	-8.8
From Baseline	2	77	-9.0	84	-14.1	77	-15.9	79	-13.5
	3	56	-15.5	68	-18.4	68	-18.9	73	-18.6
	4	52	-18.2	61	-22.7	61	-23.4	68	-20.1

Table 5.4 Unadjusted Mean Change from Baseline in PANSS Positive Sub-Scale Score for OC Data Set in Efficacy Sample for Study 31-97-202

		PANSS Positive Sub-Scale Score							
		Placebo		Risperidone 6 mg		Aripiprazole 20 mg		Aripiprazole 30 mg	
Variable	Week	N	Mean	N	Mean	N	Mean	N	Mean
Mean Baseline		103	24.5	95	23.9	98	24.8	96	24.0
Mean Change	1	102	-0.843	95	-3.074	96	-2.656	95	-2.484
From Baseline	2	77	-2.506	84	-4.917	77	-5.299	79	-3.797
	3	56	-4.661	68	-5.765	68	-6.338	73	-4.959
	4	52	-5.346	61	-7.148	61	-7.623	68	-5.662

Table 5.5 Unadjusted Mean Change from Baseline in CGI Severity of Illness Score for OC Data Set in Efficacy Sample for Study 31-97-202

		CGI Severity of Illness Score							
		Placebo		Risperidone 6 mg		Aripiprazole 20 mg		Aripiprazole 30 mg	
Variable	Week	N	Mean	N	Mean	N	Mean	N	Mean
Mean Baseline		103	4.8	95	4.8	98	4.8	96	4.7
Mean Change	1	102	-0.157	95	-0.379	96	-0.281	95	-0.284
From Baseline	2	77	-0.247	84	-0.702	77	-0.649	79	-0.570
	3	56	-0.589	68	-0.838	68	-0.812	73	-0.726
	4	52	-0.712	61	-1.082	61	-0.951	68	-0.853

Figure 5.1 PANSS Total Scores over Time for Different Dropout Cohorts: Placebo vs. Risperidone for Study 31-97-202

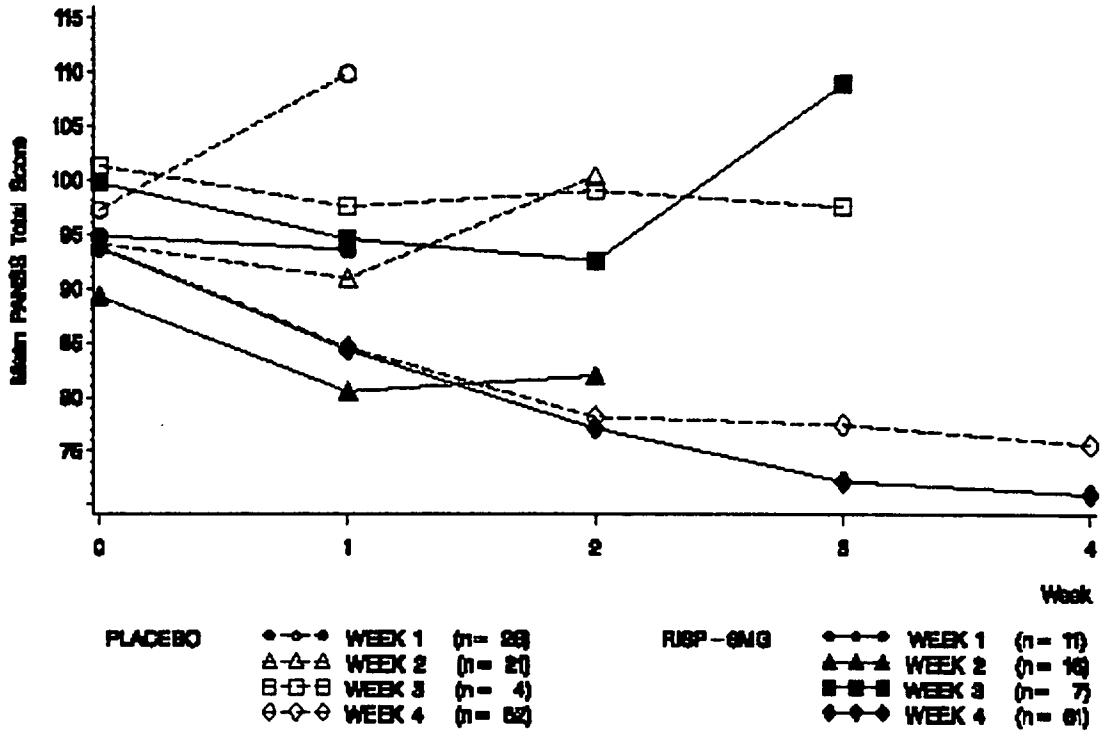


Figure 5.2 PANSS Total Scores over Time for Different Dropout Cohorts: Placebo vs. Aripiprazole 20 mg for Study 31-97-202

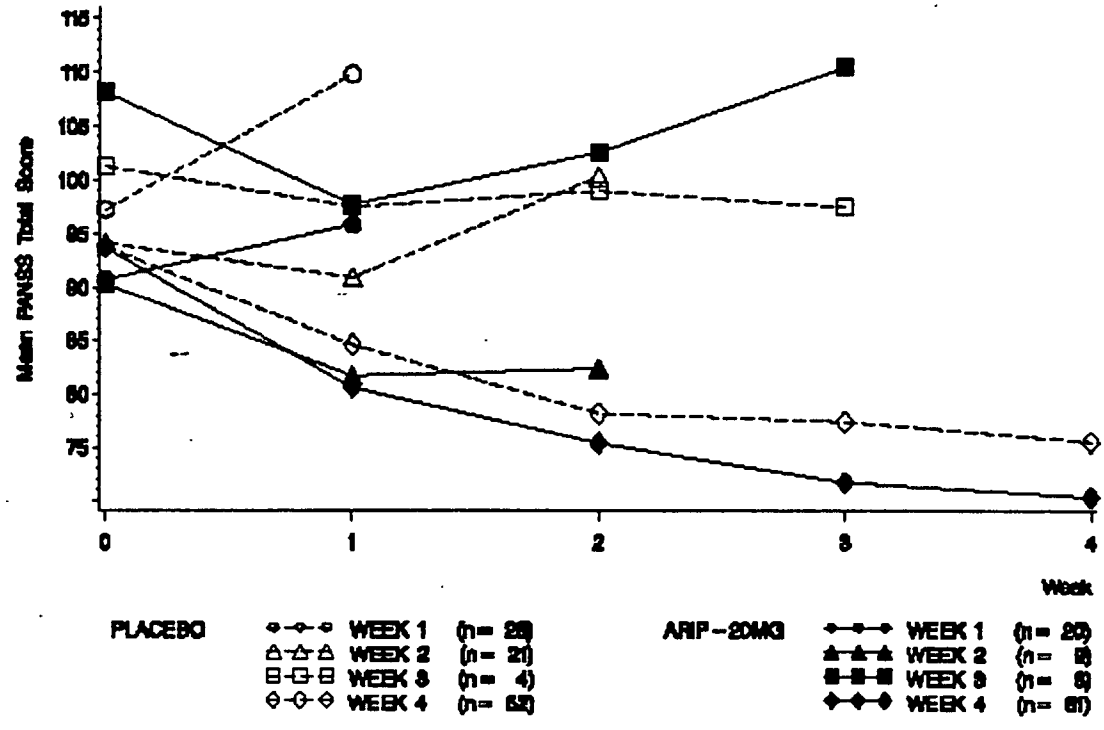


Figure 5.3 PANSS Total Scores over Time for Different Dropout Cohorts: Placebo vs. Aripiprazole 30 mg for Study 31-97-202

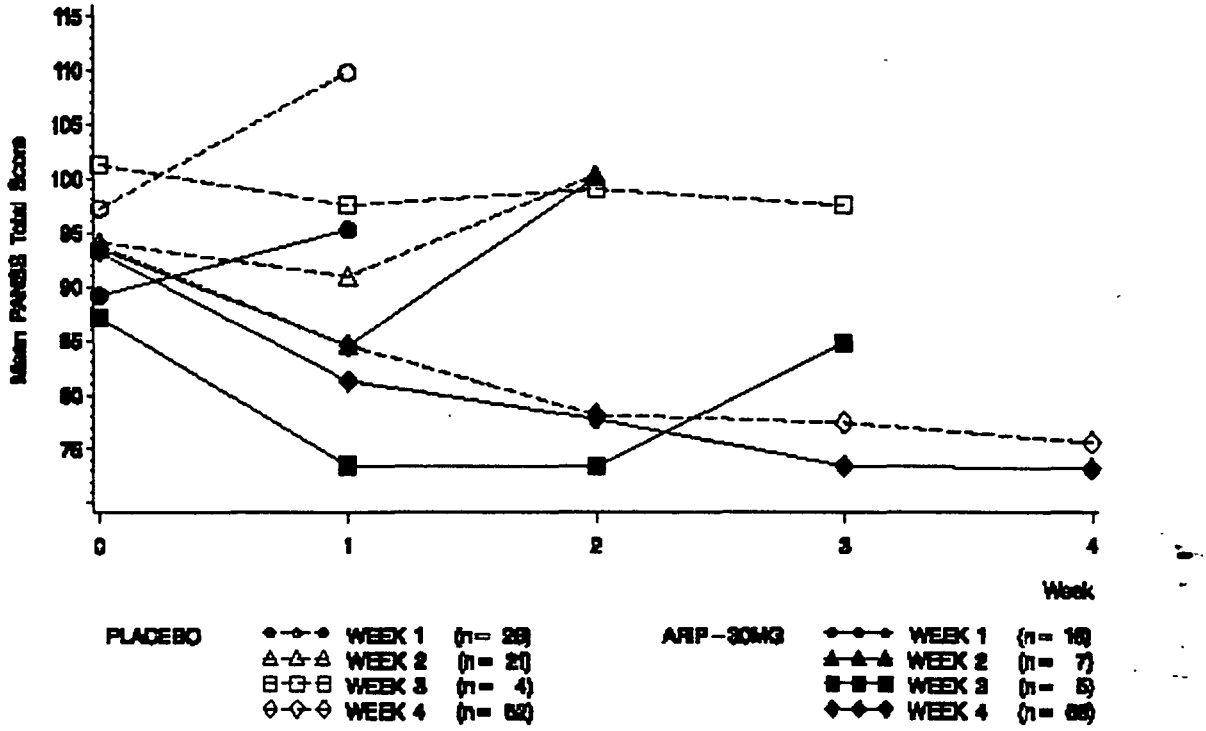


Table 5.6 Average Changes of PANSS Positive Score for Dropout Cohort Analyses for Study 31-97-202

Group	Week 1	Week 2	Week 3	Week 4
Placebo	2.846*	1.334	-0.502	-5.346
Risperidone 6 mg	0.452	-2.81	1.721	-7.148
Aripiprazole 20 mg	2	-5.448	-0.628	-7.623
Aripiprazole 30 mg	1.624	-0.148	1.0102	-5.662

* The values of OC analyses after excluding the dropout patients at week 1 was –2.025 and the unadjusted average changes for the placebo group were –2.506, -4.661 and –5.346 at Weeks 2, 3 and 4.

Table 5.7 Mean Changes of CGI Severity of Illness Score for Dropout Cohort Analyses for Study 31-97-202

Group	Week 1	Week 2	Week 3	Week 4
Placebo	0.347*	0.284	0.249	-0.712
Risperidone 6 mg	-0.182	-0.439	0.0001	-1.082
Aripiprazole 20 mg	0.4	0.006	0.2513	-0.951
Aripiprazole 30 mg	0.315	0.143	-0.199	-0.853

*The values of OC analyses after excluding the dropout patients at week 1 was –0.325 and the unadjusted average changes for the placebo group were –0.247, -0.589 and –0.7115 at Weeks 2, 3 and 4.

5. The first two pivotal phase III studies of aripiprazole, i.e., Studies 31-97-201 and 31-97-202 were designed in the treatment of psychosis. So, the sponsor recruited patients with either schizophrenia or schizoaffective disorder. The sponsor, however, only reported the analyses for all patients (both schizophrenia and schizoaffective diagnoses) and patients with schizophrenia alone in each study's report. They were later requested by us (the clinical reviewer and statistical reviewer) to provide the analyses on the three primary efficacy endpoints for the subgroup of patients with schizoaffetive diagnosis. We are interested to know if the results shown for this subgroup have similar magnitude of the drug/placebo differences as the schizophrenic sample. Since the sample sizes for this subgroup were small in these studies, it is understood that the test results between the treatment groups and placebo may not be significant.

   Table 5.8 and 5.9 show the LOCF data analysis results on the three primary efficacy endpoints for Studies 31-97-201 and 31-97-202, respectively by this reviewer (Note: results were the same as the sponsor's). Comparing the values of treatment effects in Table 5.8 with Table 4.1.1.8 and in Table 5.9 with Table 4.1.2.8. We noticed that for Study 31-97-201, except aripiprazole 30 mg on PANSS Total Scores, other treatment effects in schizoaffective patients were smaller than schizophrenic patients. However, for Study 31-97-202, except aripiprazole 30 mg on PANSS Positive Sub-Scale Score and on CGI Severity of Illness Score, other treatment effects in schizoaffective patients were bigger than patients with schizophrenia. The treatment effects between these two subgroups seemed not much different in both studies.

   Although it is not the purpose, it was noticed that the aripiprazole 15 mg had better improvement results than aripiprazole 30 mg for Study 31-97-201. Similarly, the aripiprazole 20 mg had better improvement results than aripiprazole 30 mg for Study 31-97-202. Moreover, it is interesting to know that for Study 31-97-202 none of the comparisons between the aripiprazole groups and placebo showed a p-value less than 0.05 on any primary endpoint, nevertheless, all three comparisons between risperidone 6 mg and placebo were significant.

   Table 13 in the Appendix shows the Observed Case analysis results on the three primary efficacy endpoints for the subgroup of patients with schizophrenia alone and schizoaffective disorder for both studies. It was noticed that for Study 31-97-202, the OC analysis results did not show separation between aripiprazole and placebo. As a matter of fact, patients in the placebo group even had more average of improvement than those in the aripiprazole groups. The sponsor's explanation was that this may be due to the very small sample sizes (only 9 patients in the placebo group), the very high placebo response in schizoaffective patients in this study and the high discontinuation rate for the placebo group (only 9 patients in the OC analysis compared to 25 in the LOCF).

Table 5.8 Efficacy Analysis Results for the LOCF Data Set for Patients with Schizoaffective for Study 31-97-201

	N	Change from Baseline to Endpoint (i.e., week 4)	Treatment Effect	P-value (vs. placebo)
PANSS Total				
Haloperidol 10 mg	40	-11.8106	-9.4656	0.1452
Aripiprazole 15 mg	27	-14.1604	-11.8154	0.0972
Aripiprazole 30 mg	29	-12.227	-9.882	0.1569
Placebo	28	-2.345		
PANSS Positive Sub-Scale Score				
Haloperidol 10 mg	40	-4.4990	-3.9198	0.0328
Aripiprazole 15 mg	27	-3.7573	-3.1781	0.1125
Aripiprazole 30 mg	29	-3.4067	-2.8275	0.1554
Placebo	29	-0.5792		
CGI Severity of Illness Score				
Haloperidol 10 mg	40	-0.4076	-0.2636	0.3132
Aripiprazole 15 mg	27	-0.6418	-0.4978	0.0837
Aripiprazole 30 mg	29	-0.4548	-0.3108	0.2709
Placebo	29	-0.1440		

Table 5.9 Efficacy Analysis Results for the LOCF Data Set for Patients with Schizoaffective for Study 31-97-202

	N	Change from Baseline to Endpoint (i.e., week 4)	Treatment Effect	P-value (vs. placebo)
PANSS Total				
Risperidone 6 mg	24	-17.34	-15.901	0.0195
Aripiprazole 20 mg	33	-11.27	-9.831	0.1144
Aripiprazole 30 mg	28	-10.74	-9.301	0.1523
Placebo	25	-1.439		
PANSS Positive Sub-Scale Score				
Risperidone 6 mg	24	-5.495	-4.6898	0.0175
Aripiprazole 20 mg	33	-4.303	-3.4978	0.0546
Aripiprazole 30 mg	28	-2.749	-1.9438	0.3001
Placebo	25	-0.8052		
CGI Severity of Illness Score				
Risperidone 6 mg	24	-0.7881	-0.7657	0.0094
Aripiprazole 20 mg	33	-0.3891	-0.3667	0.1762
Aripiprazole 30 mg	28	-0.3459	-0.3235	0.2479
Placebo	25	-0.0224		

6. The primary efficacy variable of Study CN138-001 was the mean change from baseline to Week 6 but it was noticed that in the study design, 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. Due to large amount of patients who chose the open-label aripiprazole treatment during Weeks 4 to 6, the results of OC analysis showed insignificant after Week 4 although the results of LOCF analysis showed significant. Since the results of OC analyses were significant from Week 1 to Week 3 by the Hochberg's procedure, this reviewer thinks that the insignificant results of OC analyses should not be a concern.

7. In conclusion, all three pivotal studies were positive. However, as discussed in Comment #4, the biasness of LOCF and OC analysis results for Study 31-97-202 was a concern to this reviewer.

5.2 Phase II Studies: Studies 31-93-202 and 31-94-202

1. Two primary efficacy variables were defined for Study 31-93-202. They were: (1) change from baseline in BPRS total score at last visit and (2) improvement by at least one point over baseline in CGI severity score at last visit. The analysis method specified in the protocol for variable (1) was the Wilcoxon rank-sum test, and for variable (2) was either Fisher exact test or chi-square test. However, the sponsor's statistical analysis method shown in their study reports for these two variables were ANCOVA and Cochran-Mantel-Haenszel test instead, respectively.

   They were later requested to re-analyze the data by using the protocol-specified methods for the above two primary efficacy variables. The p-value for variable (1) became 0.17 by the Wilcoxon rank-sum test and p-values for variable (2) by Fisher exact test and chi-square test were 0.066 and 0.045, respectively.

   Like pivotal phase III studies 31-97-201 and 31-97-202, they did not pre-specify any method for multiple efficacy endpoints, the significant results shown on both efficacy variables were deemed to be necessary for claiming a positive study. So, it was determined by this reviewer that Study 31-93-202 was a negative study.

2. Since the principal investigator at Center 003, Richard L. Borison, M.D. had his employment terminated by the Augusta Veterans Affairs Medical Center on June 7, 1996 due to allegations of research misconduct, the efficacy analyses for study 31-94-202 should be based on the data without Center 003. According to the data presented in 4.2.2.2 of this review, none of aripiprazole dosage groups showed significant results on the BPRS-Core score, one of two primary efficacy endpoints. Similar to Study 31-93-202, the sponsor did not pre-specify any method for multiple endpoints, so Study 31-94-202 was determined as a negative study.

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

This was a negative study according to the sponsor's test result of p-value, 0.427 on the primary efficacy endpoint: time to failure to maintain response. Although the sponsor had protocol-specified intention to pool data from both studies for efficacy and safety evaluations, we did not usually accept the results by the combined data analyses. Now that the results showed insignificant, there was no need to further discuss this issue.

5.4 Additional Comment (For Subgroup Analysis)

The sponsor reported a table (Table 5.10) 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 the combined studies. For three individual pivotal phase III studies, however, they only performed the subgroup analyses for gender on the PANSS Total Score among those four categories. This reviewer performed the subgroup analyses for gender on the PANSS Positive Sub-Scale Score and CGI Severity of Illness Score for Studies 31-97-201 and 31-97-202, and for age and race for all three primary endpoints for Studies 31-97-201 and 31-97-202 as well as for one primary endpoint for Study CN138-001.

The subgroup analyses for gender are shown in Tables 6, 6A, 8, 8A and 11 of the Appendix. The subgroup analyses for age are shown in the following Table 5.11 and Tables 14 and 15 of the Appendix. The subgroup analyses for race are shown in Tables 16-18 of the Appendix. Note that, the ANOVA model used for obtaining the means of change of scores included the baseline value as a covariate. The sponsor's protocols did not mention any subgroup analysis.

According to Table 5.10, the sponsor summarized in the Integrated Summary of Efficacy that " Efficacy was found to be similar for men and women. For the subset of age, because the number of patients ≥ 65 years was minimal (1%), data was insufficient for useful evaluation of efficacy in that population. In order to evaluate efficacy in older patients, a subset was evaluated at ≥ 50 years. Although aripiprazole patients ≥ 50 years did not show a difference relative to placebo due to a high placebo response, the actual PANSS scores at endpoint for aripiprazole patients who were 50 years or older was similar to those for patients < 50 years old. For the subset of race, efficacy was found to be similar for whites and blacks. In this data set, Hispanic patients (N=42) had a high placebo response compared with other races." and "For baseline psychiatric status, patients who were more severely ill (PANSS Total Score > 91) showed a greater improvement at endpoint compared with patients who were less ill (PANSS Total Score ≤ 91); however this result might be expected because more severely ill patients are able to show a greater change. The PANSS Total Score of 91 was the median value observed in the database."

Since the magnitude of mean change of PANSS Total Score at Endpoint for patients who were less than 50 year old in the placebo group was extremely small comparing to other treatment groups. This reviewer performed the subgroup analyses for age (<50 and ≥50) to observe any difference between these age groups for each pivotal study and showed the results in Table 5.11.

It was interesting to find that for each study the placebo group' magnitude of mean change of PANSS Total Score was greater than one of aripiprazole groups in older patients (age≥50). For Study 31-97-201, the placebo group of older patients had bigger magnitude of mean change of PANSS Total Score than the aripiprazole 30mg group of older patients. For Study 31-97-202, the placebo group of older patients had bigger magnitude of mean change of PANSS Total Score than the aripiprazole 20 mg group of older patients. Also, for Study CN138-001, the placebo group of older patients had bigger magnitude of mean change of PANSS Total Score than the aripiprazole 15 mg group. Although there were not many patients greater or equal to 50 year old in the studies, this consistent finding seems to tell us that aripiprazole may not be an effective drug for the older patients suffering from schizophrenia.

Table 5.10 PANSS Total Score: Model-Based Mean Change from Baseline at Endpoint by Gender, Age, Race and Baseline Score; LOCF Data Set, Efficacy Sample; Short-Term, Placebo-Controlled Efficacy Studies (31-93-202, 31-94-202, 31-97-202 and CN 138-001)

		PANSS Total Score at Endpoint							
Subgroup		**N**	**Placebo**	**N**	**Haloperidol**	**N**	**Risperidone**	**N**	**Aripiprazole**
Gender	Men	301	-2.8	137	-13.4	67	-14.1	661	-12.6
	Women	103	-3.2	49	-14.1	28	-15.3	224	-13.9
Age (years)	< 50	351	-1.8	162	-14.2	87	-14.2	743	-13.3
	≥ 50	53	-9.8	24	-9.4	8	-17.6	142	-10.8
Race	White	204	-2.0	115	-14.4	53	-15.0	492	-12.7
	Black	140	-2.4	51	-11.5	36	-15.8	260	-13.7
	Hispanic	42	-10.9	14	-14.1	4	-1.8	91	-9.0
	Asian	10	14.1	1	-14.0	1	11.4	21	-22.5
Baseline PANSS Total	Above Median (> 91)	196	-5.7	105	-18.1	49	-17.1	433	-17.9
	Below Median (≤ 91)	208	0.1	81	-9.4	46	-11.4	452	-8.0

Table 5.11 Model Based Mean Change of PANSS Total Score for Age Subgroups of Patients for Pivotal Studies

Study 31-97-201	Age ≥50			Age<50		
	n	Mean	SE	n	Mean	SE
Aripiprazole 15mg	12	-18.45	6.36	87	-13.91	2.52
Aripiprazole 30mg	14	-8.69	5.79	86	-10.87	2.53
Haloperidol 10mg	11	-15.65	6.53	88	-12.63	2.50
Placebo	11	-14.44	6.52	91	-0.42	2.46

Study 31-97-202	Age ≥50			Age<50		
	n	Mean	SE	n	Mean	SE
Aripiprazole 20mg	14	-10.11	4.15	84	-14.87	2.44
Aripiprazole 30mg	20	-14	3.42	76	-12.83	2.57
Risperidone 6mg	8	-17.09	5.44	87	-15.16	2.40
Placebo	13	-11.68	4.22	90	-3.54	2.36

CN138-001	Age ≥50			Age<50		
	n	Mean	SE	N	Mean	SE
Aripiprazole 10mg	22	-15.33	5.58	81	-14.14	2.56
Aripiprazole 15mg	17	-3.88	6.33	86	-12.92	2.49
Aripiprazole 20mg	19	-18.51	5.98	78	-13.08	2.61
Placebo	22	-14.30	5.57	85	2.17	2.50

Yeh-Fong Chen, Ph.D.
Mathematical Statistician

Concurrence:

Dr. Jin

Dr. Chi

cc: NDA 21-436
HFD-120/Dr. Katz
HFD-120/Dr. Laughren
HFD-120/Dr. Dubitsky
HFD-120/Mr. Hardeman
HFD-700/Dr. Anello
HFD-710/Dr. Chi
HFD-710/Dr. Jin
HFD-710/Dr. Chen
This review consists of 89 pages. MS Word: C:/yfchen/NDA21436/review.doc

6. Appendices

Table 1 Disposition of Patients, All Patients (Schizophrenia and Schizoaffective Disorder)

	Number of Patients
Enrolled	502
Entered Placebo Washout	460
Discontinued	46
Did not qualify for randomization	18
Adverse event	4
Lost to follow-up	11
Patient withdrew consent	13

	Number (%) of Patients				
		Haloperidol	Aripiprazole		
	Placebo	10 mg	15 mg	30 mg	Total
Randomized Sample	106	104	102	102	414
Completed Study	58 (55)	62 (60)	68 (67)	60 (59)	248 (60)
Discontinued	48 (45)	42 (40)	34 (33)	42 (41)	166 (40)
Adverse event	17 (16)	11 (11)	9 (9)	8 (8)	45 (11)
Lost to follow-up	1 (1)	0	0	1 (1)	2 (< 1)
Patient withdrew consent (personal reasons)	12 (11)	20 (19)	15 (15)	10 (10)	57 (14)
Patient met withdrawal criteria[a]	1 (1)	0	1 (1)	1 (1)	3 (1)
Noncompliance	1 (1)	1 (1)	0	1 (1)	3 (1)
Insufficient clinical response	15 (14)	6 (6)	5 (5)	15 (15)	41 (10)
Patient withdrew consent (lack of effect)	1 (1)	4 (4)	4 (4)	6 (6)	15 (4)

[a] Patient 97201-1-3 in the placebo group was withdrawn for administrative reasons. Patient 97201-1-2 in the aripiprazole 15-mg group and Patient 97201-1-1 in the aripiprazole 30-mg group were withdrawn because the study site was closed.

Figure 1: Time to Discontinuation Due to All Reasons, Randomized Sample for Study 31-97-201

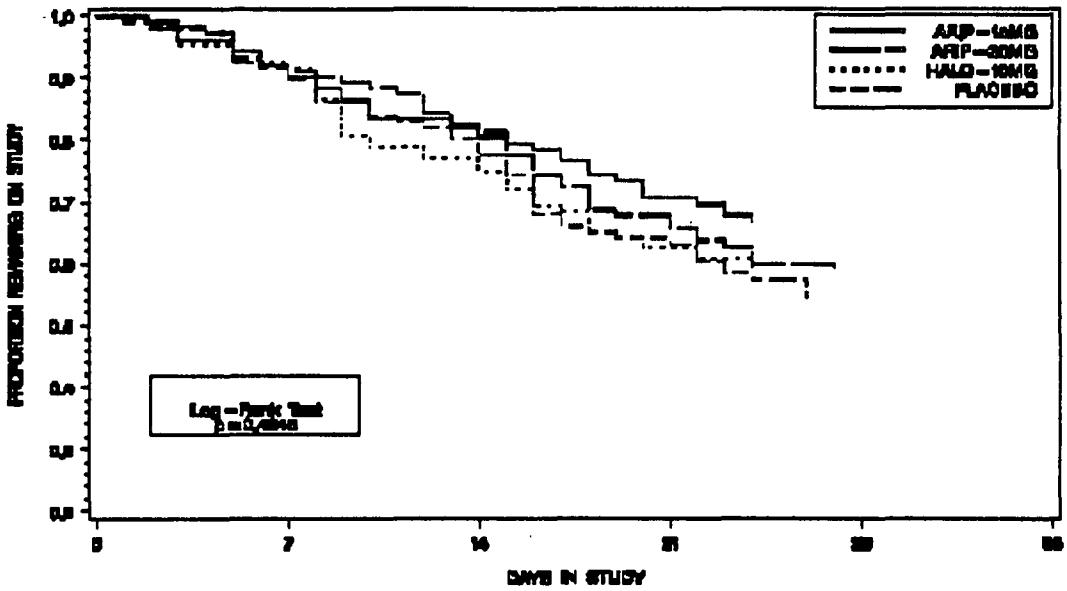


Table 2 Disposition of Patients with Schizophrenia

	Number (%) of Patients				
		Haloperidol	Aripiprazole		
	Placebo	10 mg	15 mg	30 mg	Total
Randomized Sample	75	61	74	72	282
Completed Study	44 (59)	41 (67)	53 (72)	42 (58)	180 (64)
Discontinued	31 (41)	20 (33)	21 (28)	30 (42)	102 (36)
Adverse event	11 (15)	6 (10)	5 (7)	6 (8)	28 (10)
Lost to follow-up	1 (1)	0	0	1 (1)	2 (1)
Patient withdrew consent (personal reasons)	6 (8)	9 (15)	10 (14)	9 (13)	34 (12)
Patient met withdrawal criteria[a]	0	0	1 (1)	0	1 (< 1)
Noncompliance	1 (1)	0	0	0	1 (< 1)
Insufficient clinical response	11 (15)	2 (3)	2 (3)	9 (13)	24 (9)
Patient withdrew consent (lack of effect)	1 (1)	3 (5)	3 (4)	5 (7)	12 (4)

[a] Patient 97201-1-2 in the aripiprazole 15-mg group was withdrawn because the study site was closed.

Table 3 Disposition of Patients: All Patients (Schizophrenia and Schizoaffective Disorder)

	Number of Patients
Enrolled	487
Entered Placebo washout	448
Discontinued	44
Did not qualify for randomization	21
Patient withdrew consent	12
Reasons for withdrawal not noted	11

	Number (%) of Patients				
		Risperidone	Aripiprazole		
	Placebo	6 mg	20 mg	30 mg	Total
Randomized Sample	103	99	101	101	404
Completed Study	52 (50)	62 (63)	61 (60)	67 (66)	242 (60)
Discontinued	51 (50)	37 (37)	40 (40)	34 (34)	162 (40)
Adverse event	17 (17)	8 (8)	11 (11)	8 (8)	44 (11)
Lost to follow-up	0	2 (2)	0	2 (2)	4 (1)
Patient withdrew consent (personal reasons)	11 (11)	12 (12)	18 (18)	9 (9)	50 (12)
Patient met withdrawal criteria[a]	0	0	0	1 (1)	1 (< 1)
Noncompliance	0	1 (1)	1 (1)	2 (2)	4 (1)
Protocol Violation	1 (1)	1 (1)	0	0	2 (< 1)
Insufficient clinical response	17 (17)	8 (8)	9 (9)	8 (8)	42 (10)
Patient withdrew consent (lack of effect)	5 (5)	5 (5)	1 (1)	4 (4)	15 (4)

[a] Patient 97-202-71-22 in the aripiprazole 30-mg group did not receive study medication according to the dosing record.

APPEARS THIS WAY
ON ORIGINAL

Figure 2 Time to Discontinuation Due to All Reasons, Randomized Sample for Study 31-97-202

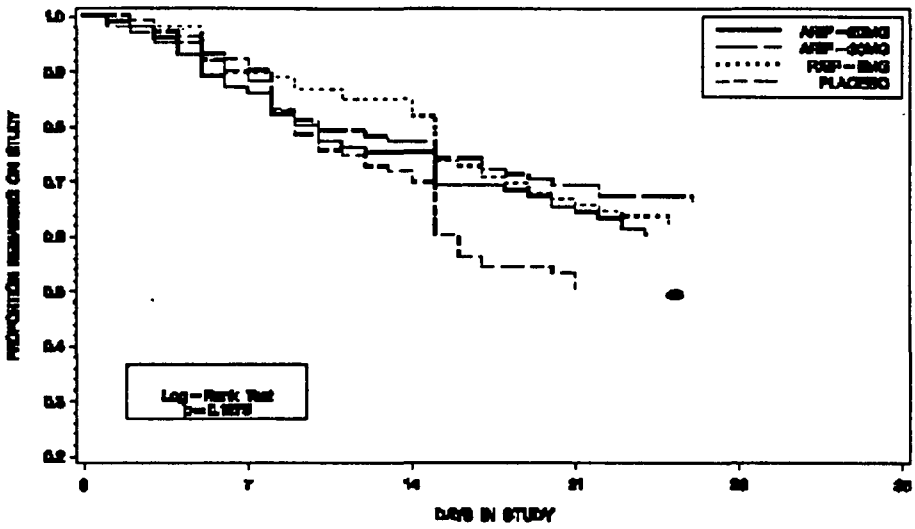


Table 4 Disposition of Patients with Schizophrenia

	Number (%) of Patients				
		Risperidone		Aripiprazole	
	Placebo	6 mg	20 mg	30 mg	Total
Randomized Sample	78	74	66	71	289
Completed Study	43 (55)	47 (64)	42 (64)	51 (72)	183 (63)
Discontinued	35 (45)	27 (36)	24 (36)	20 (28)	106 (37)
Adverse event	10 (13)	5 (7)	6 (9)	5 (7)	26 (9)
Lost to follow-up	0	2 (3)	0	1 (1)	3 (1)
Patient withdrew consent (personal reasons)	9 (12)	9 (12)	13 (20)	5 (7)	36 (12)
Noncompliance	0	1 (1)	0	2 (3)	3 (1)
Protocol Violation	1 (1)	0	0	0	1 (< 1)
Insufficient clinical response	11 (14)	6 (8)	5 (8)	5 (7)	27 (9)
Patient withdrew consent (lack of effect)	4 (5)	4 (5)	0	2 (3)	10 (3)

Table 5 The Summary of Model-Based Mean Change from Baseline in PANSS Total Score by Study Center, LOCF Data Set, Efficacy Sample for Study 31-97-201

Center	Visit	PANSS Total Score							
		Placebo		Haloperidol 10 mg		Aripiprazole 15 mg		Aripiprazole 30mg	
		N	Mean	N	Mean	N	Mean	N	Mean
Overall	Baseline	102	100.9	99	99.9	99	98.8	100	99.6
	Endpoint	102	-2.9	99	-13.8	99	-15.5	100	-11.4
023	Endpoint	7	0.6	7	-13.7	7	-19.1	7	-4.0
039	Endpoint	7	-3.1	7	-11.6	6	-16.5	7	-8.8
043	Endpoint	7	2.5	7	-4.8	7	-7.9	7	-5.7
027	Endpoint	6	-4.1	5	-8.8	6	-4.9	6	-0.7
030	Endpoint	6	0.5	6	-8.0	6	-23.6	6	-14.7
036	Endpoint	6	9.9	6	-9.8	6	3.8	6	-10.1
025	Endpoint	5	-17.8	5	-6.0	5	-32.7	5	-18.8
031	Endpoint	4	5.8	3	-3.4	4	-26.2	3	-18.1
007	Endpoint	4	-6.4	4	-7.8	3	-10.4	4	-2.0
028	Endpoint	4	2.9	4	-3.5	4	-7.5	4	-7.1
020	Endpoint	4	-6.6	4	-16.1	3	-28.5	3	8.5
038	Endpoint	3	22.1	3	-15.8	3	-8.1	3	4.8
026	Endpoint	3	0.1	3	-16.9	3	-29.9	3	-11.0
029	Endpoint	3	40.6	3	-18.5	3	5.7	3	-37.0

Table 6 Mean Change from Baseline in PANSS Total Score by Gender, LOCF Data Set, Efficacy Sample for Study 31-97-201

	PANSS Total Score							
	Placebo		Haloperidol 10 mg		Aripiprazole 15mg		Aripiprazole 30mg	
Variable	Men N=71	Women N=31	Men N=65	Women N=34	Men N=73	Women N=26	Men N=69	Women N=31
Mean Baseline	100.5	99.4	99.1	100.5	96.1	103.0	98.5	98.7
Endpoint (Week4)	-1.2	-3.7	-12.2	-14.1	-13.5	-17.3	-12.9	-5.4

Table 6A Mean Change from Baseline in PANSS Positive Sub-Scale Score and CGI Severity of Illness Scores by Gender, LOCF Data Set, Efficacy Sample for Study 31-97-201

	PANSS Positive Sub-Scale Score							
	Placebo		Haloperidol 10 mg		Aripiprazole 15mg		Aripiprazole 30mg	
Variable	Men N=71	Women N=31	Men N=65	Women N=34	Men N=73	Women N=26	Men N=69	Women N=31
Mean Baseline	25.27	23.97	24.72	25.94	24.40	25.15	24.36	24.42
Endpoint (Week4)	-0.16	-0.60	-4.01	-4.50	-3.62	-5.04	-4.35	-1.77

	CGI Severity of Illness Score							
	Placebo		Haloperidol 10 mg		Aripiprazole 15mg		Aripiprazole 30mg	
Variable	Men N=71	Women N=31	Men N=65	Women N=34	Men N=73	Women N=26	Men N=69	Women N=31
Mean Baseline	4.94	4.94	4.88	4.79	4.90	4.92	4.84	4.74
Endpoint (Week4)	-0.01	-0.17	-0.47	-0.47	-0.64	-0.61	-0.50	-0.12

APPEARS THIS WAY
ON ORIGINAL

Table 7 The Summary of Model-Based Mean Change from Baseline in PANSS Total Score by Study Center, LOCF Data Set, Efficacy Sample for Study 31-97-202

Center	Visit	PANSS Total Score							
		Placebo		Risperidone 6 mg		Aripiprazole 15 mg		Aripiprazole 30mg	
		N	Mean	N	Mean	N	Mean	N	Mean
Overall	Baseline	103	94.1	95	92.6	98	93.5	96	91.6
	Endpoint	103	-5.0	95	-15.7	98	-14.5	96	-13.9
050	Endpoint	7	-1.1	7	-6.6	7	-19.1	7	-6.3
051	Endpoint	3	-6.8	3	-16.3	3	-18.9	3	-15.3
053	Endpoint	3	-3.7	3	-3.2	3	-29.8	3	0.7
059	Endpoint	7	-26.0	5	-1.3	6	-15.9	5	-18.6
067	Endpoint	6	-0.0	5	-14.6	6	-7.3	6	3.0
069	Endpoint	6	-18.4	5	-7.9	6	-16.4	6	-9.9
071	Endpoint	6	11.5	6	-1.9	6	-3.8	5	-8.7
081	Endpoint	4	5.4	5	-25.8	5	-30.6	4	-25.8
084	Endpoint	7	-12.6	7	-20.8	7	-16.1	7	-23.2
093	Endpoint	5	5.0	5	-37.1	5	-15.1	5	-23.9

APPEARS THIS WAY
ON ORIGINAL

Table 8 Mean Change from Baseline in PANSS Total Score by Gender, LOCF Data Set, Efficacy Sample for Study 31-97-202

	PANSS Total Score							
	Placebo		Risperidone 6mg		Aripiprazole 20mg		Aripiprazole 30mg	
Variable	Men N=73	Women N=30	Men N=67	Women N=28	Men N=71	Women N=27	Men N=63	Women N=31
Mean Baseline	95.5	93.9	94.5	91.3	91.2	101.3	91.9	93.1
Endpoint (Week 4)	-4.6	-5.1	-14.4	-18.6	-13.8	-12.8	-9.8	-20.1

Table 8A Mean Change from Baseline in PANSS Positive Sub-Scale Score and CGI Severity of Illness Scores by Gender, LOCF Data Set, Efficacy Sample for Study 31-97-202

	PANSS Positive Sub-Scale Score							
	Placebo		Risperidone 10 mg		Aripiprazole 15mg		Aripiprazole 30mg	
Variable	Men N=73	Women N=30	Men N=67	Women N=28	Men N=71	Women N=27	Men N=63	Women N=33
Mean Baseline	25.07	23.23	24.22	23.25	24.72	25	24.16	23.64
Endpoint (Week4)	-1.40	-2.45	-4.8	-5.79	-5.02	-4.0	-2.54	-6.10

	CGI Severity of Illness Score							
	Placebo		Risperidone 10 mg		Aripiprazole 15mg		Aripiprazole 30mg	
Variable	Men N=73	Women N=30	Men N=67	Women N=28	Men N=71	Women N=27	Men N=63	Women N=33
Mean Baseline	4.86	4.67	4.96	4.57	4.73	4.96	4.78	4.67
Endpoint (Week4)	-0.18	-0.25	-0.73	-0.77	-0.49	-0.49	-0.43	-0.84

APPEARS THIS WAY
ON ORIGINAL

Table 9 Disposition of Patients in Study CN138-001

Patient Status	Number of Patients (%)				
	Placebo	Aripiprazole 10 mg	Aripiprazole 15 mg	Aripiprazole 20 mg	Total
Enrolled Sample	n/a	n/a	n/a	n/a	508
Baseline failures	n/a	n/a	n/a	n/a	88
Randomized	108	106	106	100	420
Discontinued from double-blind treatment[a]	78 (72)	63 (59)	74 (70)	63 (63)	278 (66)
Due to lack of response entered open-label treatment	44 (41)	28 (26)	37 (35)	22 (22)	131 (31)
Adverse event	6 (6)	11 (10)	3 (3)	5 (5)	25 (6)
Lack of efficacy	11 (10)	5 (5)	8 (8)	11 (11)	35 (8)
Patient withdrew consent	13 (12)	18 (17)	24 (23)	18 (18)	73 (17)
Patient unreliability	0	1 (1)	1 (1)	1 (1)	3 (1)
Lost to follow-up	0	0	0	4 (4)	4 (1)
Pregnancy	0	0	0	0	0
Death	0	0	0	0	0
Other known cause	4 (4)	0	1 (1)	2 (2)	7 (2)
Completed double-blind treatment	30 (28)	43 (41)	32 (30)	37 (37)	142 (34)

[a] Patients not responding at the end of Week 3, as indicated by CGI Improvement Score of 4 to 7, were placed on open-label treatment.

APPEARS THIS WAY
ON ORIGINAL

Table 10 Disposition of Patients Who Entered Open-label Treatment for Study CN138001

	Number of Patients (%)				
	Original Randomized Treatment Group				
Patient Status	**Placebo**	**Aripiprazole 10 mg**	**Aripiprazole 15 mg**	**Aripiprazole 20 mg**	**Total**
Entered open-label treatment from Week 3 to Week 5	44	28	37	22	131
Discontinued from open-label treatment	22 (50)	10 (36)	18 (49)	9 (41)	59 (45)
Adverse event	5 (11)	2 (7)	4 (11)	1 (5)	12 (9)
Lack of efficacy	13 (30)	8 (29)	10 (27)	6 (27)	37 (28)
Patient withdrew consent	4 (9)	0	3 (8)	2 (9)	9 (7)
Patient unreliability	0	0	0	0	0
Lost to follow-up	0	0	0	0	0
Pregnancy	0	0	0	0	0
Death	0	0	0	0	0
Other known cause	0	0	1 (3)	0	1 (1)
Completed open-label treatment	22 (50)	18 (64)	19 (51)	13 (59)	72 (55)

Figure 3: Time to Discontinuation for Any Reason, Randomized Sample for Study CN 138-001

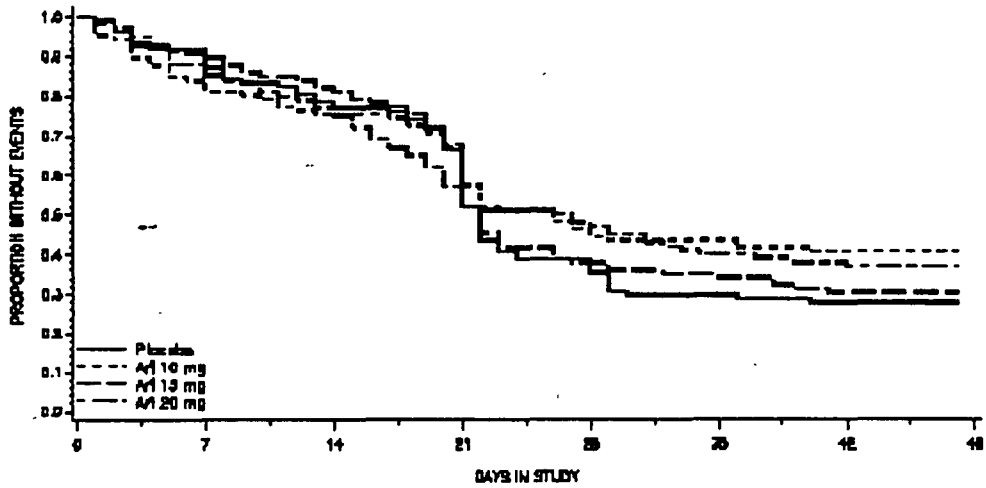


Table 11 Model Based Mean Change from Baseline in PANSS Total Score for Gender in the LOCF Data Set for Study CN138-001

Male	PANSS Total Score			
Phase	Placebo	Aripiprazole 10 mg	Aripiprazole 15 mg	Aripiprazole 20 mg
Variable	N=82	N=80	N=76	N=79
Mean Baseline	92.02	93.44	92.79	97.31
Day 4	-2.28	-4.13	-3.66	-4.74
Week 1	-3.02	-7.51	-7.14	-6.76
Week 2	-1.87	-10.18	-6.66	-9.66
Week 3	-1.40	-12.03	-7.43	-11.89
Week 4	-1.03	-12.82	-9.41	-12.13
Week 5	-0.44	-13.57	-10.26	-13.03
Week 6	-1.05	-14.87	-11.72	-14.31
Female	PANSS Total Score			
Phase	Placebo	Aripiprazole 10 mg	Aripiprazole 15 mg	Aripiprazole 20 mg
Variable	N=25	N=23	N=27	N=18
Mean Baseline	93.64	90.39	94.63	103.89
Day 4	-3.83	-0.99	-7.13	-6.82
Week 1	-0.50	-5.29	-2.76	-11.23
Week 2	-1.03	-9.38	-7.33	-12.26
Week 3	-2.01	-9.38	-9.23	-10.82
Week 4	-1.85	-8.77	-10.51	-10.35
Week 5	-1.55	-10.39	-10.09	-12.89
Week 6	-1.86	-10.99	-10.82	-15.25

Note: The baseline score was used as a covariate in the ANCOVA model.

Figure 4 Time to Discontinuation due to All Reasons in the Randomized Sample for Studies 31-98-217 and 31-98-304-01

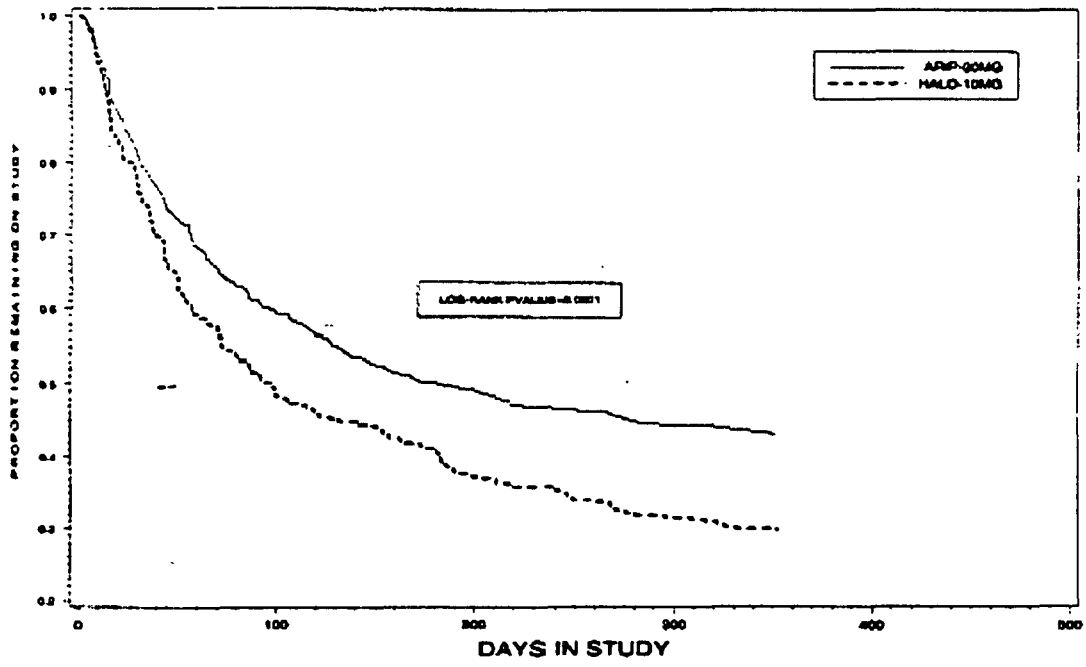


Table 12 Disposition of Patients for Studies 31-98-217 and 31-98-304-01

	Number (%) of Patients		
Patient Status	**Haloperidol**	**Aripiprazole**	**Total**
Enrolled Sample	n/a	n/a	1452
Screening Failures	n/a	n/a	158
Entered Placebo Washout	n/a	n/a	1294
Randomized Sample	433	861	1294
Discontinued From Double-Blind Treatment	305 (70)	494 (57)	799 (62)
Lost to follow-up	10 (2)	24 (3)	34 (3)
Patient withdrew consent (personal reasons)	97 (22)	159 (19)	256 (20)
Insufficient clinical response	38 (9)	63 (7)	101 (8)
Adverse event other than worsening schizophrenia	80 (19)	70 (8)	150 (12)
Adverse event of worsening schizophrenia	58 (13)	143 (17)	201 (16)
Study participation terminated by sponsor	1 (< 1)	2 (< 1)	3 (< 1)
Noncompliance	17 (4)	25 (3)	42 (3)
Protocol violation	3 (1)	8 (1)	11 (1)
Patient met withdrawal criteria	1 (< 1)	0	1 (< 1)
Completed Double-Blind Treatment	128 (30)	367 (43)	495 (38)

Table 13 Subgroup Analysis for the Diagnosis in the OC Data Set for Study 31-97-201 and Study 31-97-202

Study 31-97-201	**Schizophrenia Patients alone**			**Schizoaffective Disorder Patients alone**		
PANSS Total Score	N	Mean	P-Value	N	Mean	P-Value
Haloperidol 10mg	40	-16.5	0.135	21	-16.22	0.9120
Aripiprazole 15mg	53	-21.3	0.005	15	-33.15	0.0227
Aripiprazole 30mg	42	-17.4	0.086	19	-25.15	0.1965
Placebo	46	-10.0		14	-15.44	
PANSS Positive Sub-Scale Score	N	Mean	P-Value	N	Mean	P-Value
Haloperidol 10mg	40	-4.3	0.104	21	-6.35	0.1248
Aripiprazole 15mg	53	-5.8	0.004	15	-8.40	0.0166
Aripiprazole 30mg	42	-5.9	0.004	19	-7.17	0.0643
Placebo	46	-2.2		14	-3.46	
CGI Severity of Illness Score	N	Mean	P-Value	N	Mean	P-Value
Haloperidol 10mg	40	-0.6	0.122	21	-0.63	0.8589
Aripiprazole 15mg	53	-0.8	0.008	15	-1.29	0.0409
Aripiprazole 30mg	41	-0.6	0.095	19	-0.86	0.3942
Placebo	46	-0.3		14	-0.58	

Study 31-97-202	Schizophrenia Patients alone			Schizoaffective Disorder Patients alone		
PANSS Total Score	N	Mean	P-Value	N	Mean	P-Value
Risperidone 6mg	47	-20.6	0.254	14	-30.60	0.6726
Aripiprazole 20mg	42	-23.5	0.069	19	-22.25	0.4146
Aripiprazole 30mg	51	-20.3	0.279	17	-20.11	0.2650
Placebo	43	-16.1		9	-27.65	
PANSS Positive Sub-Scale Score	N	Mean	P-Value	N	Mean	P-Value
Risperidone 6mg	47	-6.6	0.170	14	-9.64	0.3653
Aripiprazole 20mg	42	-7.8	0.023	19	-6.74	0.7037
Aripiprazole 30mg	51	-6.2	0.291	17	-4.46	0.1598
Placebo	43	-4.9		9	-7.57	
CGI Severity of Illness Score	N	Mean	P-Value	N	Mean	P-Value
Risperidone 6mg	47	-1.0	0.072	14	-1.31	0.4519
Aripiprazole 20mg	42	-1.0	0.141	19	-0.92	0.7199
Aripiprazole 30mg	51	-1.0	0.123	17	-0.64	0.2348
Placebo	43	-0.7		9	-1.04	

Table 14 Model Based Mean Change of PANSS Positive Sub-scale Score for Age Subgroups of Patients for Pivotal Studies

Study 31-97-201	**Age ≥50**			**Age<50**		
	N	**Mean**	**SE**	**n**	**Mean**	**SE**
Aripiprazole 15mg	12	-5.69	1.72	87	-3.77	0.74
Aripiprazole 30mg	14	-3.05	1.58	86	-3.63	0.74
Haloperidol 10mg	11	-3.84	1.78	88	-4.23	0.73
Placebo	11	-3.44	1.78	92	0.09	0.72
Study 31-97-202	**Age ≥50**			**Age<50**		
	N	**Mean**	**SE**	**n**	**Mean**	**SE**
Aripiprazole 20mg	14	-3.33	1.49	84	-5.07	0.72
Aripiprazole 30mg	20	-4.44	1.23	76	-3.55	0.76
Risperidone 6mg	8	-5.41	1.99	87	-5.02	0.71
Placebo	13	-5.02	1.52	90	-1.22	0.70

Table 15 Model Based Mean Change of CGI Severity of Illness Score for Age Subgroups of Patients for Pivotal Studies

Study 31-97-201	**Age ≥50**			**Age<50**		
	N	**Mean**	**SE**	**n**	**Mean**	**SE**
Aripiprazole 15mg	12	-1.08	0.29	87	-0.56	0.1
Aripiprazole 30mg	14	-0.29	0.26	86	-0.39	0.1
Haloperidol 10mg	11	-0.50	0.31	88	-0.48	0.1
Placebo	11	-0.33	0.30	92	-0.02	0.1
Study 31-97-202	**Age ≥50**			**Age<50**		
	n	**Mean**	**SE**	**n**	**Mean**	**SE**
Aripiprazole 20mg	14	-0.45	0.23	84	-0.50	0.11
Aripiprazole 30mg	20	-0.67	0.19	76	-0.54	0.12
Risperidone 6mg	8	-0.38	0.30	87	-0.77	0.11
Placebo	13	-0.56	0.23	90	-0.15	0.11

Table16 Model Based Mean Change from Baseline in All Three Primary Endpoints for Race Subgroup Analysis for Study 31-97-201

	PANSS Total Score			
	Placebo	Haloperidol 10 mg	Aripiprazole 15 mg	Aripiprazole 30 mg
White				
Baseline Mean & (N)	98.2 (50)	99.6 (64)	98.7 (58)	97.1 (58)
Endpoint (Week 4)	-0.59	-14.12	-13.42	-9.66
Black				
Baseline Mean & (N)	102.85 (34)	96.14 (22)	98.5 (26)	100.88 (25)
Endpoint (Week 4)	-3.59	-11.52	-19.69	-15.05
Hispanic				
Baseline Mean & (N)	103 (14)	103.11 (9)	92.67 (12)	99.25 (12)
Endpoint (Week 4)	-8.4	-5.57	-9.96	1.03
Asian				
Baseline Mean & (N)	84.67 (3)	106 (1)	97 (3)	109.67 (3)
Endpoint (Week 4)	23.17	-15.10	-14.48	-41.65

	PANSS Positive Sub-Scales			
	Placebo	Haloperidol 10 mg	Aripiprazole 15 mg	Aripiprazole 30 mg
White				
Baseline Mean & (N)	24.27 (51)	25.66 (64)	24.97 (58)	24.4 (58)
Endpoint (Week 4)	0.08	-4.71	-3.39	-3.63
Black				
Baseline Mean & (N)	25.85 (34)	23.91 (22)	23.81 (26)	24.44 (25)
Endpoint (Week 4)	-0.28	-3.61	-4.53	-4.54
Hispanic				
Baseline Mean & (N)	25.21 (14)	24.44 (9)	23.92 (12)	24 (12)
Endpoint (Week 4)	-2.77	-0.99	-4.54	0.51
Asian				
Baseline Mean & (N)	21.33 (3)	19 (1)	27 (3)	25.67 (3)
Endpoint (Week 4)	5.99	-10.71	-7.22	-11.86

	CGI Severity of Illness			
	Placebo	Haloperidol 10 mg	Aripiprazole 15 mg	Aripiprazole 30 mg
White				
Baseline Mean & (N)	4.96 (51)	4.89 (64)	4.97 (58)	4.83 (58)
Endpoint (Week 4)	-0.08	-0.57	-0.63	-0.43
Black				
Baseline Mean & (N)	5 (34)	4.73 (22)	4.89 (26)	4.76 (25)
Endpoint (Week 4)	-0.002	-0.51	-0.64	-0.4
Hispanic				
Baseline Mean & (N)	4.93 (14)	4.78 (9)	4.67 (12)	4.92 (12)
Endpoint (Week 4)	-0.23	0.19	-0.59	0.13
Asian				
Baseline Mean & (N)	4 (3)	5 (1)	5 (3)	5 (3)
Endpoint (Week 4)	0.67	0	-1	-1.33

Table17 Mean Change from Baseline in All Three Primary Endpoints for Race Subgroup Analysis for Study 31-97-202

	PANSS Total Score			
	Placebo	Risperidone 10 mg	Aripiprazole 15 mg	Aripiprazole 30 mg
White				
Baseline Mean & (N)	96.30 (57)	96.06 (53)	94.05 (58)	90.93 (58)
Endpoint (Week 4)	-4.59	-16.18	-13.92	-12.33
Black				
Baseline Mean & (N)	93.83 (35)	88.94 (36)	90.64 (30)	92.83 (29)
Endpoint (Week 4)	-5.83	-17.49	-15.47	-12.11
Hispanic				
Baseline Mean & (N)	94.5 (4)	90.5 (4)	102.17 (6)	106.67 (3)
Endpoint (Week 4)	-2.68	5.36	-1.77	-13.70
Asian				
Baseline Mean & (N)	87.33 (3)	102 (1)	103.5 (2)	109 (2)
Endpoint (Week 4)	8.60	4.51	-23.37	-40.28

	PANSS Positive Sub-Scales			
	Placebo	Risperidone 10 mg	Aripiprazole 15 mg	Aripiprazole 30 mg
White				
Baseline Mean & (N)	24.96 (57)	24.64 (53)	24.47 (58)	22.93 (58)
Endpoint (Week 4)	-1.92	-5.56	-5.07	-3.23
Black				
Baseline Mean & (N)	24.14 (35)	23.06 (36)	25.17 (30)	25.38 (29)
Endpoint (Week 4)	-1.54	-5.47	-4.52	-3.51
Hispanic				
Baseline Mean & (N)	24 (4)	20.5 (4)	29.67 (6)	31 (3)
Endpoint (Week 4)	-1.07	0.74	-1.26	-6.03
Asian				
Baseline Mean & (N)	23 (3)	22 (1)	23 (2)	25 (2)
Endpoint (Week 4)	0.86	0.28	-8.97	-9.97

	CGI Severity of Illness			
	Placebo	Risperidone 10 mg	Aripiprazole 15 mg	Aripiprazole 30 mg
White				
Baseline Mean & (N)	4.95 (57)	4.96 (53)	4.86 (58)	4.64 (58)
Endpoint (Week 4)	-0.20	-0.76	-0.55	-0.56
Black				
Baseline Mean & (N)	4.66 (35)	4.67 (36)	4.63 (30)	4.97 (29)
Endpoint (Week 4)	-0.16	-0.85	-0.46	-0.59
Hispanic				
Baseline Mean & (N)	4.25 (4)	4.75 (4)	5.17 (6)	5 (3)
Endpoint (Week 4)	-0.11	0.24	0.23	-0.63
Asian				
Baseline Mean & (N)	4.67 (3)	5 (1)	5 (2)	4.5 (2)
Endpoint (Week 4)	0.32	0.05	-1.45	-1.05

Table 18 Mean Change from Baseline in All Three Primary Endpoints for Race Subgroup Analysis for Study CN138-001

	PANSS Total Score			
	Ariperidone 10 mg	Ariperidone 15 mg	Ariperidone 20 mg	Placebo
White				
Baseline Mean & (N)	92.25 (53)	92.25 (55)	90.29 (52)	94.69 (49)
Endpoint (Week 4)	-12.63	-7.53	-15.18	2.76
Black				
Baseline Mean & (N)	87.44 (27)	85.67 (27)	89.42 (26)	85.78 (36)
Endpoint (Week 4)	-14.31	-14.23	-12.47	2.27
Hispanic				
Baseline Mean & (N)	99.28 (18)	109.31 (16)	112.83 (12)	98.47 (17)
Endpoint (Week 4)	-20.85	-17.45	-12.18	-23.24
Asian				
Baseline Mean & (N)	80 (1)	92.5 (4)	88.67 (3)	92.5 (4)
Endpoint (Week 4)	-8.37	-20.25	0.13	7.50

APPEARS THIS W
ON ORIGINA

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

/s/

Yeh-Fong Chen
7/15/02 03:47:57 PM
BIOMETRICS

Kun Jin
7/15/02 04:11:46 PM
BIOMETRICS

George Chi
7/16/02 09:07:39 AM
BIOMETRICS

APPEARS THIS WAY
ON ORIGINAL

Statistical Review and Evaluation

Review of Rat Carcinogenicity Studies

NDA#: 21-436

APPLICANT: Otsuka Pharm

NAME OF DRUG: Aripiprazole

INDICATION: Schizophrenia

STUDIES REVIEWED: Rat Studies: Otsuka Study No. 009489 and BMS Study No. 99321 in Volumes 1.75, and 1.80

PHARMACOLOGY REVIEWER: Lois Freed, Ph.D. (HFD-120)

STATISTICAL REVIEWER: Roswitha Kelly, M.S. (HFD-710)

Roswitha Kelly, M.S.
Mathematical Statistician

George Chi, Ph.D.
Director, Division of Biometrics 1

Distribution:

HFD-120/ Arch. NDA 21-436
HFD-120/Barry Rossloff Ph.D.
HFD-120/Lois Freed, Ph.D.
HFD-710/Kun Jin, Ph.D.
HFD-710/George Chi, Ph.D.
HFD-710/Roswitha Kelly, M.S.
HDS-700/Charles Anello, Sc.D.

This review consists of 8 pages of text and 16 pages of Tables and Figures.
File Directory: c:\Data\My Documents\aripiprazole_rats.doc

Table of Contents

Table of Tables

Table of Figures

1.0 Introduction

The sponsor has submitted two rat and two mouse carcinogenicity studies. As there is basically double the information of the usual bioassay, the multiplicity problem inherent in carcinogenicity analyses is increased. This reviewer, however, performed no further adjustment on the usual 0.025 and 0.005 levels of significance in trend for rare and common tumors, as there is no guidance on this issue. High levels of significance and consistency across gender and studies should be considered when interpreting any findings. This reviewer wrote a separate review for each species but presented the main findings from both species in the summary section.

2.0 Otsuka Study No. 009489 in Rats

This was a 104-week carcinogenicity study of OPC-31 in SPF Fischer (F344/DuCrj) rats. The test substance was administered in the diet to 50 animals/sex/group at dose levels of 0, 1, 3, and 10 mg/kg/day. Animals were individually housed and water and feed were available ad lib. Additional 8 satellite animals/sex of the treated groups were maintained for 52 weeks for determination of plasma concentrations. From the main study, surviving animals were terminally sacrificed after 104 weeks of dosing. All animals were subjected to complete necropsy and histopathological examination. Mortality was statistically investigated by life table analysis (two-sided). Overall incidence of neoplastic lesions and number of females with mammary gland tumors were tested one-sided by Cochran-Armitage trend test. Time-related occurrence of mammary gland tumors in females was tested one-sided by Peto's onset rate method.

2.1 Sponsor's Findings for Study 009489

The sponsor observed no significant differences in mortality between the control and treated groups of either gender. The final mortality rates (number of animals killed in extremis or found dead) were as follows:

Dose Group (mg/kg/day)	Males	Females
0	13/50 (26%)	7/50 (14%)
1	7/50 (14%)	8/50 (16%)
3	8/50 (16%)	13/50 (26%)
10	10/50 (20%)	12/50 (24%)

Mean body weight in the male high dose group was slightly, but significantly, lower (3-8%) than controls throughout the study. Mean body weight in the female high dose was slightly, but significantly, higher (4-8%) than controls for weeks 30-68. However, mean body weight was comparable to that of controls at the other times.

Among the neoplastic findings, only females showed any statistically significant increase, namely for fibroadenoma of the mammary gland ($p \leq 0.05$ by Cochran-Armitage trend test) as well as by Fisher's exact test when comparing the high dose with the controls).

Additionally, time to onset of mammary gland tumors in females revealed that these tumors tended to occur earlier in the high-dose group than in the control group (Peto's onset rate method).

2.2 Reviewer's Findings for Study 009489

The intercurrent mortality among the female rats of this study showed a few more animals dying in the mid- and high dose groups than in the controls, but not to a statistically significant degree (p=0.2378, Tables 1-2, Figure 1). There was a statistically significant increase in fibroadenomas of the mammary gland (p=0.002 vs. α=0.005, Table 3).

Among the males, there was a lack of difference between the mortality experience of the various dose groups (p=0.9028, Tables 4-5, Figure 2)). No individual tumor finding reached statistical significance (Table 6).

2.3 Validity of the Male Rats of Study 009489

As there were no statistically significant tumor trends among the male rats in this study, its validity needs to be assessed. Two criteria are set up for this purpose (Haseman[12], Chu et al.[3], and Bart et al.[4]):

i) was a sufficient number of animals exposed long enough to allow for late-developing tumors, and

ii) did the high dose provide a sufficient tumor challenge?

The number of animals and length of exposure can be assessed at weeks 52, 80-90, and at termination, but are generally considered satisfied if 20-30 animals survive through weeks 80-90. Only seven high dose males had died by week 91 and 80% lived to terminal sacrifice, easily satisfying this criterion. The high dose is expected to be close to the MTD to present a sufficient tumor challenge. Suppression in survival when compared to the controls and/or average body weight differences of about 10 percent, especially during the first year of treatment, are indicators that the high dose is close to the MTD. For this study, the mortality pattern of the high dose group did not distinguish itself from the other groups, including the controls. The sponsor reported average body weight data being 3-8 percent lower for the high-dose males than the controls for most of the study, which suggests that the high dose was close to the MTD. Therefore, the long-term administration of aripiprazole to male rats can be considered a valid study.

---

[1] Haseman: Statistical Issues in the Design, Analysis and Interpretation of Animal Carcinogenicity Studies, *Environmental Health Perspectives*, Vol. 58, pp 385-392, 1984.

[2] Haseman: Issues in Carcinogenicity Testing: Dose Selection, *Fundamental and Applied Toxicology*, Vol. 5, pp. 66-78, 1985.

[3] Chu, Cueto, Ward: Factors in the Evaluation of 200 National Cancer Institute Carcinogenicity Bioassays, *Journal of Toxicology and Environmental Health*, Vol. 8, pp 251-280, 1981.

[4] Bart, Chu, Tarone: Statistical Issues in Interpretation of Chronic Bioassay Tests for Carcinogenicity, *Journal of the National Cancer Institute*, pp. 957-974, 1979.

3.0 BMS Study No. 99321 in Rats

Crl:CD®(SD)IGS BR rats were assigned randomly to groups receiving either the vehicle (two groups) or aripiprazole at doses of 10, 20, 40, and 60 mg/kg/day via gavage for two years. Group size was 55 per gender. Implanted microchip identification devices held the permanent identification number. Animals were housed individually and water and food was available ad lib. All tissues were microscopically examined from each animal. Mortality data were evaluated using a two-sided Cox-Tarone test for trend. Differences in non-palpable tumor rates were analyzed using the method of Peto and Pike[5]. The two control groups were pooled for these analyses. Palpable tumors were analyzed by the Cox-Tarone binary regression method using the first palpation time as onset time. Levels of significance were set at p-values of 0.005 and 0.025 for common and rare tumors, respectively.

3.1 Sponsor's Findings for Study 99321

The sponsor observed dose-dependent increases in survival for each gender, though five maximum dose females could not tolerate the dosing during week one and were found dead or euthanatized moribund.

Dose Group (mg/kg/day)	Males	Females
0	34/55 (62%)	33/55 (60%)
0	36/55 (65%)	34/55 (62%)
10	26/55 (47%)	33/55 (60%)
20	26/55 (47%)	23/55 (42%)
40	22/55 (40%)	15/55 (27%)
60	17/55 (31%)	20/55 (36%)

Average bodyweights were decreased in a dose-dependent way for males (9-44% at week 102). Low dose females experienced no effect on bodyweight. Mid-, high-, and maximum-dose females had average bodyweights 17-41% lower than controls at week 102.

Among the females, adrenocortical carcinoma and combined adrenocortical adenomas and carcinomas showed a statistically significant trend. The trend excluding the maximum dose was not statistically significant and the two carcinomas in the 40 mg/kg dose were considered not to be clearly drug related. The maximum dose was considered to markedly exceed the MTD. No other statistically significant positive trends in tumors were observed in either gender.

[5] Peto et al.: Guidelines for simple sensitive significance tests for carcinogenic effects in long-term animal experiments. In: IARC Monographs on the Evaluation of the Carcinogenic Risk of Chemicals to Humans, Supplement 2: Long-term and Short-term Screening Assays for Carcinogens: A Critical Appraisal, Lyon, International Agency for Research on Cancer, 1980: 311-346.

3.2 Reviewer's Findings for Study 99321

The intercurrent mortality for the female rats in this study was extremely statistically significant (p=0.0000, Tables 7-8), but in the direction of better survival with dose. The significant lack of homogeneity indicates that the separation among the survival curves does not strictly follow the increasing doses. The Kaplan-Meier graphs make visually apparent what these p-values convey (Figure 3). Among the tumor findings, malignant carcinoma of the adrenal cortex was highly statistically significant (p=0.0001 vs. α=0.025, Table 9). Combining carcinomas and adenomas of the adrenal cortex resulted in a significant p-value of 0.0002 (based on the asymptotic test due to tumors being both incidental and fatal).

The male rats of this study had similar mortality experience as the female rats, in that the treated animals experienced significantly better survival than the controls (p=0.0000, Tables 10-11). The significant p-value for homogeneity indicates that the survival curves crossed occasionally, but, in general, survival increased with dose. The Kaplan-Meier curves bear out these observations (Figure 4). Among the tumor findings, none reached statistical significance among the male rats (combining benign and malignant pheochromocytoma of the adrenal, medullar, resulted in a p-value of 0.0334, which is not close to statistical significance for common tumors).

3.3 Validity of Male Rat Study 99321

The same criteria as noted above to evaluate the male rat study of Study 009489 are being applied to the male rats of Study 99321, as no statistically significant increase in tumors were observed. Survival was excellent for all groups and the number of animals living long enough is not an issue. Survival was better for the treated than for the control groups. Average bodyweight for the maximum dose animals was 44 % lower than the controls' at week 102. By week 2 of the study, an 18.5% lower average body weight was observed for the maximum dose, and the difference steadily increased. This would indicate as the sponsor had noted, that the maximum dose exceeded the MTD. As there is another valid study in male rats available (Study No. 009489), no further investigation of this study was done (e.g. excluding the top dose and evaluating the remaining data as a potentially valid study).

4.0 Summary -

The following table summarizes the major statistically significant findings of the two rat studies.

	Otsuka Study 009489		BMS Study 99321	
	Females	Males	Females	Males
Survival	NS	NS	Sign. increased	Sign. increased
Mammary Gland, Fibroadenoma	Sign. increased	NS	NS	NS
Adrenal Cortex, Carcinoma	NS	NS	Sign. increased	NS
Validity	N/A	Yes	N/A	MTD exceeded

Otsuka Study 009489 used doses of 0, 1, 3, and 10 mg/kg/day in the diet. Survival was not affected by treatment and the only statistically significant tumor finding were fibroadenomas of the mammary gland among female rats. Among the males, no increase in tumor incidence reached statistical significance, but the study was considered valid based on length of exposure and number of animals available at study end. The high dose was judged to be close to the MTD due to the suppressed average body weights of 3-8 %.

In BMS Study 99321 doses of 10, 20, 40, and 60 mg/kg/day were administered via gavage. Two identical vehicle controls were also available. For either gender, survival was much better for the treated that the control groups. The only statistically significant tumor finding were carcinoma of the adrenal cortex, and carcinoma and adenoma of the adrenal cortex combined. The males showed no statistically significant increase in tumors. The length of exposure and the number of animals available at study end were satisfactory. However, the high average body weight suppression of the top dose compared to the controls suggested that this dose far exceeded the MTD.

The summary results for the two mice studies are given as well:

MICE	Otsuka Study 011487		Otsuka Study 011932	
	Females	Males	Females	Males
Survival	NS	Sign. increased	Sign. decreased	Sign. increased
Anterior Pituitary, Adenoma	Sign. increased	NS	Sign. increased	NS
Mammary Gland, Adenocarcinoma	Sign. increased	NS	Sign. increased	NS
Mammary Gland, Adenoacanthoma	Sign. increased	NS	Sign. increased	NS
Validity	N/A	Yes	N/A	Yes

Overall, it appears that the administration of aripiprazole in the doses given results in increased tumor findings in female rats or mice. The p-values in each case are highly statistically significant. No increase in tumor incidence rates was observed among the males of any of the studies. All but one of these male studies were judged to be valid. The maximum dose in Study No. 99321 was judged to be well beyond the MTD, based on much lower average body weights of these animals compared to the controls.

Table 1: Number of Deaths per Time interval for Female Rats in Study 009489

	Treatment Group				Total
	CTRL	LOW	MED	HIGH	
	N	N	N	N	N
Week					
53-78	.	1	2	1	4
79-91	3	2	4	1	10
92-103	3	4	7	9	23
104-104	44	43	37	39	163
Total	50	50	50	50	200

Table 2: Dose Mortality Trend Test* for Female Rats in Study 009489

Method	Time-Adjusted Trend Test	Statistic	P Value
Cox	Dose-Mortality Trend	1.39	0.2378
	Depart from Trend	2.87	0.2384
	Homogeneity	4.26	0.2346
Kruskal-Wallis	Dose-Mortality Trend	1.29	0.2562
	Depart from Trend	2.88	0.2368
	Homogeneity	4.17	0.2437

* Program used: Trend and Homogeneity Analyses of Proportions and Life Table Data Version 2.1, by Donald G. Thomas, National Cancer Institute.

Figure 1: Kaplan-Meier Survival Curves for Female Rats in Study 009489

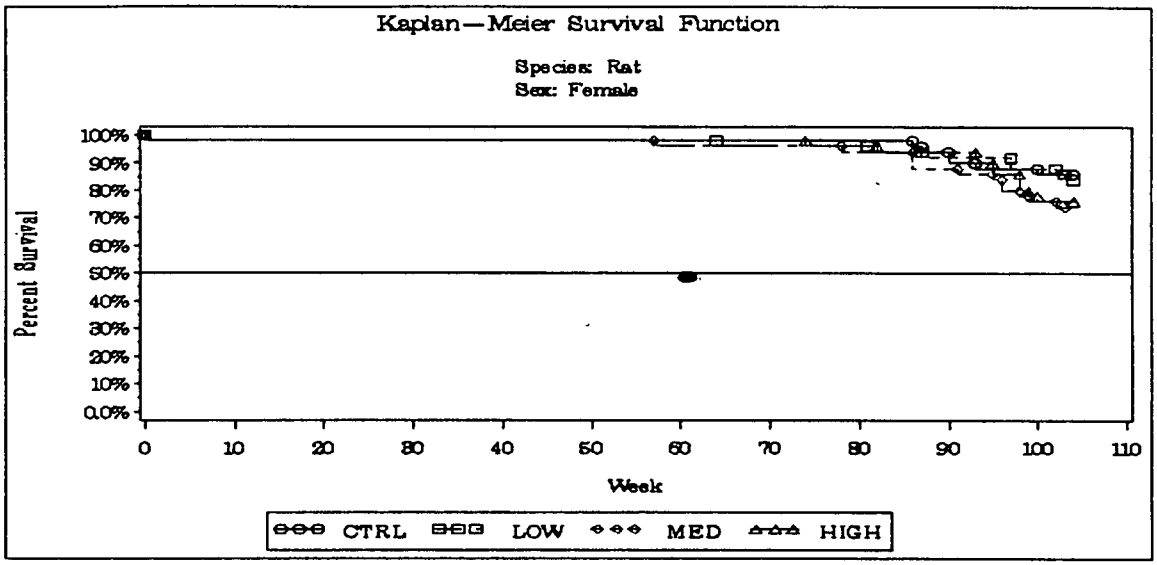


Table 3: Tumor Trend Tests for Female Rats in Study 009489

Organ Name	Organ Code	Tumor Name	Tumor Code	Natural Rate (in ctrl group)	CTRL	LOW	MED	HIGH	Tumor type	pValue (Exact)	pValue (Asymp)
Lung	18	Adenoma (Lung)	1835	2%	1	1	1	1	IN	0.4750	0.4922
Lung	18	Adenocarcinoma (Lung)	1865	2%	1	0	1	1	IN	0.4593	0.4521
Liver	34	Cholangiocarcinoma (Liver	3467	0%	0	0	1	0	IN	0.4663	0.5853
Pancreas	37	Islet cell adenoma (Pancr	3736	0%	0	0	1	2	IN	0.0516	0.0309
Kidney	38	Adenoma (Kidney)	3835	0%	0	0	1	0	IN	0.4663	0.5853
Kidney	38	Liposarcoma (Kidney)	3871	0%	0	0	1	0	FA	0.4888	0.5999
Ovary	52	Granulosa cell tumor (Ova	5232	2%	1	0	1	0	IN	0.7167	0.7770
Uterus	54	Endometrial stromal polyp	5431	12%	6	8	6	1	IN	0.9892	0.9843
Uterus	54	Hemangioma (Uterus)	5442	4%	2	0	0	0	MX	1.0000	0.9085
Clitoral gland	58	Adenoma (Clitoral gland)	5835	4%	2	3	2	5	IN	0.1074	0.1008

Clitoral gland	58	Adenocarcinoma (Clitoral	5865	2%	1	2	0	3	MX	0.1390	0.1195
Pituitary	59	Anterior adenoma (Pituita	5935	50%	25	25	33	32	MX	0.0335	0.0330
Pituitary	59	Anterior adenocarcinoma (	5965	2%	1	0	2	0	FA	0.6548	0.7503
Thyroid	60	C-cell adenoma (Thyroid)	6036	6%	3	2	7	4	IN	0.2438	0.2467
Thyroid	60	C-cell carcinoma (Thyroid	6066	2%	1	3	1	3	IN	0.2105	0.2065
Adrenal	62	Pheochromocytoma (Adrenal	6239	2%	1	4	0	2	IN	0.4892	0.4945
Adrenal	62	Malignant pheochromocytom	6269	.0%	0	2	0	0	IN	0.7850	0.8276
Ear	84	Zymbal's gland carcinoma	8465	2%	1	1	0	0	IN	0.9283	0.8702
Auricle	85	Schwannoma (Auricle)	8550	.0%	0	0	0	1	IN	0.2393	0.0564
Auricle	85	Malignant schwannoma (Aur	8580	.0%	0	0	1	0	FA	0.4889	0.5986
Skin	86	Papilloma (Skin)	8631	2%	1	0	1	1	IN	0.3564	0.3550
Skin	86	Malignant schwannoma (Ski	8680	2%	1	0	0	0	FA	1.0000	0.8471
Mammary gland	95	Adenoma (Mammary gland)	9535	2%	1	3	2	1	IN	0.6872	0.7187
Mammary gland	95	Fibroadenoma (Mammary gla	9539	12%	6	9	8	17	IN	0.0020	0.0015
Mammary gland	95	Fibroma (Mammary gland)	9540	.0%	0	1	0	2	IN	0.0966	0.0584
Mammary gland	95	Adenocarcinoma (Mammary g	9565	2%	1	1	2	1	IN	0.5897	0.6204
Abdominal cavity	98	Paraganglioma (Abdominal	9831	.0%	0	1	0	0	FA	0.7381	0.7664
Abdominal cavity	98	Malignant mesothelioma (A	9886	.0%	0	0	0	1	FA	0.2528	0.0633
General	99	Malignant lymphoma (Gener	9989	.0%	0	0	1	1	MX	0.1764	0.1520
General	99	Mononuclear cell leukemia	9993	12%	6	6	6	7	MX	0.3225	0.3298

Table 4: Number of Deaths per Timer Interval for Male Rats in Study 009489

	Treatment Group				Total
	CTRL	LOW	MED	HIGH	
	N	N	N	N	N
Week 53-78	2	1	1	2	6
79-91	3	2	2	5	12
92-103	7	4	5	3	19
104-104	38	43	42	40	163
Total	50	50	50	50	200

Table 5: Dose Mortality Trend * Test for Male Rats in Study 009489

Method	Time-Adjusted Trend Test	Statistic	P Value
Cox	Dose-Mortality Trend	0.01	0.9028
	Depart from Trend	1.90	0.3874
	Homogeneity	1.91	0.5910
Kruskal-Wallis	Dose-Mortality Trend	0.03	0.8642
	Depart from Trend	1.85	0.3958
	Homogeneity	1.88	0.5971

APPEARS THIS WAY
ON ORIGINAL

Figure 2: Kaplan Meier Survival Curves for Male Rats in Study 009489

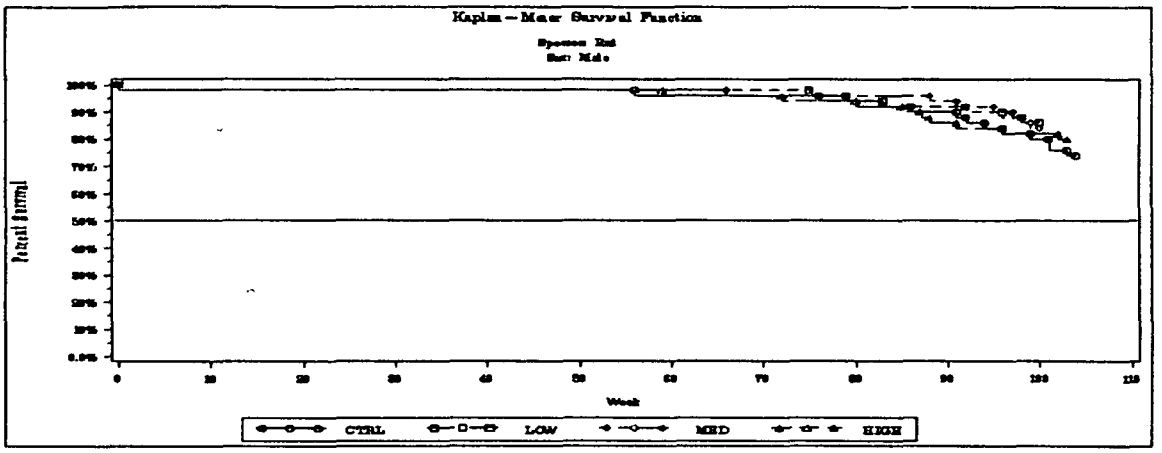


Table 6: Tumor Trend Tests for Male Rats in Study 009489

Organ Code	Organ Name	Tumor Code	Tumor Name	CTRL	LOW	MED	HIGH	pValue (Exact)	pValue (Asymp)	Natural Tumor # in control group	Natural Rate (in ctrl group)	Tumor type
03	Bone marrow(femur)	0342	Hemangioma (Bone marrow)	0	0	1	0	0.5031	0.6012	0	.0%	IN
03	Bone marrow(femur)	0354	Histiocytoma (Bone marrow	1	1	1	0	0.8379	0.8527	1	2%	IN
04	Bone marrow(sternum)	0354	Histiocytoma (Bone marrow	1	1	1	0	0.8379	0.8527	1	2%	IN
05	Bone marrow(vertebra)	0354	Histiocytoma (Bone marrow	1	1	1	0	0.8379	0.8527	1	2%	IN
07	Thymus	0740	Thymoma (Thymus)	0	0	2	0	0.4969	0.6072	0	.0%	IN
13	Spleen	1372	Hemangiosarcoma (Spleen)	0	0	2	0	0.4969	0.6072	0	.0%	IN
18	Lung	1835	Adenoma (Lung)	2	1	4	4	0.1298	0.1280	2	4%	IN
18	Lung	1865	Adenocarcinoma (Lung)	1	1	1	2	0.2580	0.2393	1	2%	IN
31	Small intestine	3165	Adenocarcinoma (Small int	0	1	1	0	0.6337	0.7386	0	.0%	IN

31	Small intestine	3176	Leiomyosarcoma (Small int	0	0	1	0	0.5031	0.6012	0	.0%	IN
34	Liver	3465	Hepatocellular carcinoma	1	0	0	0	1.0000	0.8491	1	2%	IN
37	Pancreas	3735	Acinar cell adenoma (Panc	0	0	1	0	0.5031	0.6012	0	.0%	IN
37	Pancreas	3736	Islet cell adenoma (Pancr	3	2	1	1	0.8189	0.8183	3	6%	IN
38	Kidney	3865	Adenocarcinoma (Kidney)	0	0	1	0	0.5027	0.5969	0	.0%	FA
41	Urinary bladder	4131	Papilloma (Urinary bladde	0	0	1	0	0.4211	0.5092	0	.0%	IN
43	Testis	4337	Interstitial cell tumor (	42	41	46	18	1.0000	1.0000	42	84%	IN
48	Prostate	4835	Adenoma (Prostate)	2	2	1	0	0.9391	0.9244	2	4%	IN
50	Preputial gland	5035	Adenoma (Preputial gland)	1	2	2	1	0.5828	0.6293	1	2%	MX
50	Preputial gland	5065	Adenocarcinoma (Preputial	1	3	2	1	0.6925	0.7229	1	2%	IN
59	Pituitary	5935	Anterior adenoma (Pituita	21	21	14	6	0.9999	0.9997	21	42%	MX
59	Pituitary	5936	Adenoma in intermediate p	0	0	1	0	0.5062	0.6025	0	.0%	IN
60	Thyroid	6036	C-cell adenoma (Thyroid)	10	8	6	8	0.5696	0.5776	10	20%	IN
60	Thyroid	6066	C-cell carcinoma (Thyroid	2	2	3	1	0.7337	0.7563	2	4%	IN
62	Adrenal	6239	Pheochromocytoma (Adrenal	2	3	5	1	0.7839	0.7904	2	4%	IN
62	Adrenal	6268	Malignant ganglioneuroma	0	0	0	1	0.2513	0.0629	0	.0%	FA
62	Adrenal	6269	Malignant pheochromocytom	1	0	2	0	0.6875	0.7726	1	2%	IN
75	Bone(others)	7278	Osteosarcoma (Bone)	0	0	1	0	0.5031	0.6012	0	.0%	IN
79	Skeletal muscle(others)	7977	Rhabdomyosarcoma (Skeleta	0	1	0	0	0.7500	0.8339	0	.0%	IN
85	Auricle	8550	Schwannoma (Auricle)	0	2	0	2	0.1874	0.1621	0	.0%	MX
85	Auricle	8580	Malignant schwannoma (Aur	0	0	1	0	0.5031	0.6012	0	.0%	IN
86	Skin	8631	Papilloma (Skin)	1	3	1	0	0.9216	0.9193	1	2%	IN
86	Skin	8632	Keratoacanthoma (Skin)	1	2	3	2	0.3483	0.3712	1	2%	IN

86	Skin	8633	Trichoepithelioma (Skin)	1	0	0	1	0.4317	0.3220	1	2%	IN
86	Skin	8640	Fibroma (Skin)	8	4	5	3	0.9123	0.9093	8	16%	IN
86	Skin	8641	Lipoma (Skin)	2	0	1	0	0.8970	0.8846	2	4%	IN
86	Skin	8643	Hemangiopericytoma (Skin)	1	0	0	0	1.0000	0.8472	1	2%	FA
86	Skin	8644	Hemangioleiomyoma (Skin)	0	0	0	1	0.2454	0.0600	0	.0%	IN
86	Skin	8650	Schwannoma (Skin)	0	1	0	0	0.7669	0.7805	0	.0%	IN
86	Skin	8660	Squamous cell carcinoma (	0	1	0	0	0.7527	0.7751	0	.0%	FA
86	Skin	8661	Basal cell carcinoma (Ski	0	1	0	0	0.7526	0.7786	0	.0%	FA
86	Skin	8672	Hemangiosarcoma (Skin)	2	0	0	0	1.0000	0.9129	2	4%	FA
95	Mammary gland	9539	Fibroadenoma (Mammary gla	1	3	1	0	0.9066	0.9076	1	2%	IN
95	Mammary gland	9540	Fibroma (Mammary gland)	1	0	2	0	0.6911	0.7754	1	2%	IN
95	Mammary gland	9565	Adenocarcinoma (Mammary g	1	0	0	0	1.0000	0.8503	1	2%	IN
98	Abdominal cavity	9841	Lipoma (Abdominal cavity)	0	0	1	0	0.5031	0.6012	0	.0%	IN
98	Abdominal cavity	9842	Hemangioma (Abdominal cav	1	0	0	0	1.0000	0.8479	1	2%	FA
98	Abdominal cavity	9886	Malignant mesothelioma (A	0	0	1	0	0.5000	0.5985	0	.0%	FA
99	General	9989	Malignant lymphoma (Gener	0	0	0	1	0.2487	0.0617	0	.0%	FA
99	General	9993	Mononuclear cell leukemia	5	3	2	6	0.1971	0.1980	5	10%	MX

Table 7: Number of Deaths per Time Interval for Female Rats in Study 99321

Analysis of Mortality		No. Risk	No. Died	No. Alive	Pct Survival	Pct Mortality
CTR1	0-52	55	1	54	98.2	1.8
	53-78	54	10	44	80.0	20.0
	79-91	44	5	39	70.9	29.1
	92-104	39	17	22	40.0	60.0
	FINALKILL105-106	22	22	0		
CTR2	0-52	55	3	52	94.5	5.5
	53-78	52	12	40	72.7	27.3
	79-91	40	12	28	50.9	49.1
	92-104	28	7	21	38.2	61.8
	FINALKILL105-106	21	21	0		
LOW	0-52	55	1	54	98.2	1.8
	53-78	54	9	45	81.8	18.2
	79-91	45	12	33	60.0	40.0
	92-104	33	11	22	40.0	60.0
	FINALKILL105-106	22	22	0		
MED	53-78	55	5	50	90.9	9.1
	79-91	50	5	45	81.8	18.2
	92-104	45	13	32	58.2	41.8
	FINALKILL105-106	32	32	0		
HIGH	0-52	55	1	54	98.2	1.8
	53-78	54	5	49	89.1	10.9
	79-91	49	5	44	80.0	20.0
	92-104	44	4	40	72.7	27.3
	FINALKILL105-106	40	40	0		
MAX	0-52	55	5	50	90.9	9.1
	53-78	50	3	47	85.5	14.5
	79-91	47	3	44	80.0	20.0
	92-104	44	9	35	63.6	36.4
	FINALKILL105-106	35	35	0		

Table 8: Dose Mortality Trend Test for Female Rats in Study 99321

Dose-Mortality Trend Tests

	Method			
	Cox		Kruskal-Wallis	
	Statistics	P-Value	Statistics	P-Value
Time-Adjusted Trend Test				
Depart from Trend	7.8514	0.0972	8.9348	0.0628
Dose-Mortality Trend	17.8748	0.0000	16.4303	0.0001
Homogeneity	25.7262	0.0001	25.3651	0.0001

Figure 3: Kaplan Meier Curves for Female Rats in Study 00321

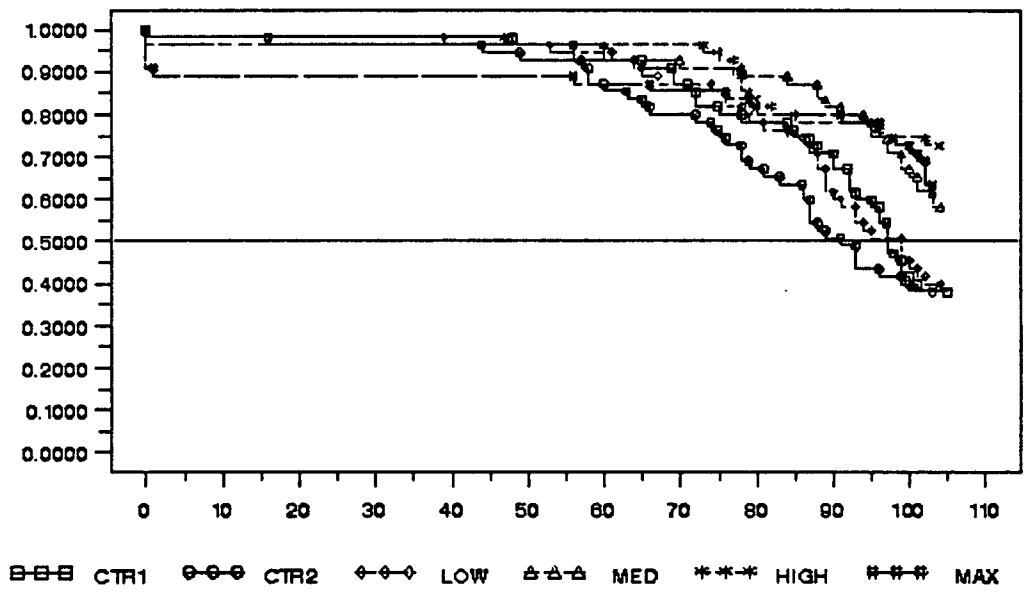


Table 9: Tumor Trend Tests for Female Rats in Study 99321

Organ Code	Organ Name	Tumor Code	Tumor Name	CTRL1	CTRL2	LOW	MED	HIGH	MAX	pValue (Exact)	pValue (Asymp)	Natural Tumor # in control group	Natural Rate (in ctrl group)	Tumor type
AC	ADRENAL, CORTEX	265	B-ADENOMA	2	3	1	3	4	6	0.0495	0.0416	5	5%	IN
AC	ADRENAL, CORTEX	90	M-CARCINOMA	0	0	0	0	2	6	0.0001	0.0001	0	.0%	MX
AS	AUDITORY SEB GL	445	M-CARCINOMA	0	0	0	0	1	0	0.4327	0.2776	0	.0%	IN
BR	BRAIN	121	M-ASTROCYTOMA	0	1	1	1	0	0	0.8689	0.8617	1	.9%	MX
BR	BRAIN	191	B-OLIGODENDROGLIOMA	1	0	1	0	0	0	0.8750	0.8532	1	.9%	IN
BR	BRAIN	192	B-GRANULAR CELL TUMOR	0	1	0	0	0	0	1.0000	0.7721	1	.9%	IN
CV	CERVIX	456	B-POLYP, ENDOMETRIAL STRO	0	0	1	0	0	0	0.7500	0.7791	0	.0%	IN
CV	CERVIX	468	M-LEIOMYOSARCOMA	0	0	0	1	0	0	0.6221	0.6244	0	.0%	IN
CV	CERVIX	54	M-SARCOMA, STROMAL	1	0	0	0	0	0	1.0000	0.8423	1	.9%	FA
HN	HEMATO NEOPLASIA	178	M-LYMPHOMA	0	0	0	1	0	2	0.0753	0.0457	0	.0%	MX
HN	HEMATO NEOPLASIA	264	M-SARCOMA, HISTIOCYTIC	0	2	0	0	1	0	0.8233	0.8024	2	2%	MX
HT	HEART	446	B-RHABDOMYOMA	0	0	0	0	1	0	0.4360	0.2785	0	.0%	IN
LI	LIVER	241	B-ADENOMA, HEPATOCELLULAR	1	3	2	1	0	1	0.9498	0.9362	4	4%	IN
LI	LIVER	451	B-CHOLANGIOMA	0	0	0	0	1	0	0.4360	0.2785	0	.0%	IN
LI	LIVER	466	M-CARCINOMA, HEPATOCELLUL	0	0	0	1	0	0	0.6221	0.6244	0	.0%	IN
LU	LUNG	369	M-CHORDOMA	0	0	0	1	0	0	0.4262	0.4624	0	.0%	IN
MA	ADRENAL, MEDULLA	218	B-PHEOCHROMOCYTOMA	2	1	4	4	2	5	0.2007	0.1867	3	3%	IN
MF	MAMMARY, FEMALE	205	B-ADENOMA	6	6	8	6	4	7	0.6408	0.6317	12	11%	MX
MF	MAMMARY, FEMALE	46	B-FIBROADENOMA	26	22	19	20	19	15	0.9947	0.9938	48	44%	MX
MF	MAMMARY, FEMALE	73	M-CARCINOMA	16	16	15	10	6	2	1.0000	1.0000	32	29%	MX
OV	OVARY	444	M-MALIGNANT GRANULOSA/THE	0	0	0	0	1	0	0.4386	0.2796	0	.0%	IN
OV	OVARY	475	B-ADENOMA	0	1	0	0	0	0	1.0000	0.8888	1	.9%	IN
PA	PANCREAS	244	M-CARCINOMA, ISLET CELL	0	1	1	2	0	0	0.8913	0.8762	1	.9%	MX
PA	PANCREAS	286	B-ADENOMA, ISLET CELL	5	2	3	2	2	1	0.9428	0.9315	7	6%	IN

PI	PITUITARY	292	M-CARCINOMA	2	1	2	0	1	1	0.8580	0.8422	3	3%	MX
PI	PITUITARY	49	B-ADENOMA	50	46	43	49	39	29	1.0000	1.0000	96	87%	MX
PN	PINNA	300	M-FIBROSARCOMA	0	0	0	1	0	0	1.0000	0.8473	0	0%	FA
PT	PARATHYROID	326	B-ADENOMA	0	2	0	1	0	0	0.9402	0.9144	2	2%	IN
SK	SKIN	207	M-FIBROSARCOMA	0	1	0	1	0	0	0.8351	0.8207	1	1%	FA
SK	SKIN	267	B-TRICHOEPITHELIOMA	1	0	0	1	0	0	0.6553	0.6735	1	1%	IN
SK	SKIN	358	M-CARCINOMA, SEBACEOUS GL	0	0	1	0	0	0	0.6066	0.6506	0	0%	IN
SK	SKIN	386	M-CARCINOMA, SQUAMOUS CEL	0	1	0	0	0	0	1.0000	0.8074	1	1%	IN
SM	MUSCLE, SKELETAL	243	M-SARCOMA, UNDIFFERENTIAT	1	0	0	0	0	0	1.0000	0.8546	1	1%	FA
SM	MUSCLE, SKELETAL	277	M-FIBROSARCOMA	0	1	0	0	0	0	1.0000	0.8598	1	1%	FA
SP	SPLEEN	448	B-HEMANGIOMA	0	0	0	0	2	0	0.2649	0.1979	0	0%	IN
SU	STOMACH, NONGL	413	B-PAPILLOMA, SQUAMOUS CEL	0	0	1	1	0	1	0.3188	0.3066	0	0%	IN
TH	THYMUS	318	B-THYMOMA	1	0	0	1	0	0	0.8028	0.7868	1	1%	IN
TH	THYMUS	403	M-THYMIC CARCINOMA	1	0	0	0	0	1	0.5160	0.4248	1	1%	IN
TI	TAIL	342	B-PAPILLOMA, SQUAMOUS CEL	0	0	1	0	0	0	1.0000	0.8461	0	0%	IN
TY	THYROID	128	B-ADENOMA, FOLLICULAR CEL	2	2	0	0	0	0	1.0000	0.9863	4	4%	IN
TY	THYROID	340	M-CARCINOMA, C-CELL	0	0	0	1	1	0	0.3991	0.3060	0	0%	IN
TY	THYROID	81	B-ADENOMA, C-CELL	7	7	14	4	5	1	0.9996	0.9993	14	13%	IN
UT	UTERUS	232	B-POLYP, ENDOMETRIAL, STR	1	1	0	3	5	2	0.0719	0.0593	2	2%	MX
UT	UTERUS	349	M-LEIOMYOSARCOMA	1	0	0	0	0	0	1.0000	0.8806	1	1%	FA
UT	UTERUS	354	M-CARCINOMA	1	0	0	0	0	0	1.0000	0.8074	1	1%	IN
UT	UTERUS	469	M-SARCOMA, ENDOMETRIAL, S	0	0	0	1	0	0	0.6221	0.6244	0	0%	IN
VA	VAGINA	370	B-POLYP, STROMAL	0	0	0	1	0	1	0.2597	0.1951	0	0%	FA

APPEARS THIS WAY
ON ORIGINAL

Table 10: Number of Deaths per Time Interval for Male Rats in Study 99321

Analysis of Mortality		No. Risk	No. Died	No. Alive	Pct Survival	Pct Mortality
CTR1	0-52	55	2	53	96.4	3.6
	53-78	53	9	44	80.0	20.0
	79-91	44	14	30	54.5	45.5
	92-104	30	9	21	38.2	61.8
	FINALKILL10 5-106	21	21	0		
CTR2	0-52	55	3	52	94.5	5.5
	53-78	52	8	44	80.0	20.0
	79-91	44	10	34	61.8	38.2
	92-104	34	15	19	34.5	65.5
	FINALKILL10 5-106	19	19	0		
LOW	0-52	55	4	51	92.7	7.3
	53-78	51	5	46	83.6	16.4
	79-91	46	6	40	72.7	27.3
	92-104	40	11	29	52.7	47.3
	FINALKILL10 5-106	29	29	0		
MED	0-52	55	5	50	90.9	9.1
	53-78	50	3	47	85.5	14.5
	79-91	47	7	40	72.7	27.3
	92-104	40	11	29	52.7	47.3
	FINALKILL10 5-106	29	29	0		
MEDHI	0-52	55	2	53	96.4	3.6
	53-78	53	1	52	94.5	5.5
	79-91	52	7	45	81.8	18.2
	92-104	45	12	33	60.0	40.0
	FINALKILL10 5-106	33	33	0		
HIGH	0-52	55	2	53	96.4	3.6
	53-78	53	3	50	90.9	9.1
	79-91	50	5	45	81.8	18.2
	92-104	45	7	38	69.1	30.9
	FINALKILL10 5-106	38	38	0		

Table 11: Dose Mortality Trend Test for Male Rats in Study 99321

	Method			
	Cox		Kruskal-Wallis	
	Statistics	P-Value	Statistics	P-Value
Time-Adjusted Trend Test				
Depart from Trend	3.1335	0.5357	2.7301	0.6040
Dose-Mortality Trend	16.8969	0.0000	16.5506	0.0000
Homogeneity	20.0304	0.0012	19.2807	0.0017

Figure 4: Kaplan Meier Survival Curves for Male Rats in Study 99321

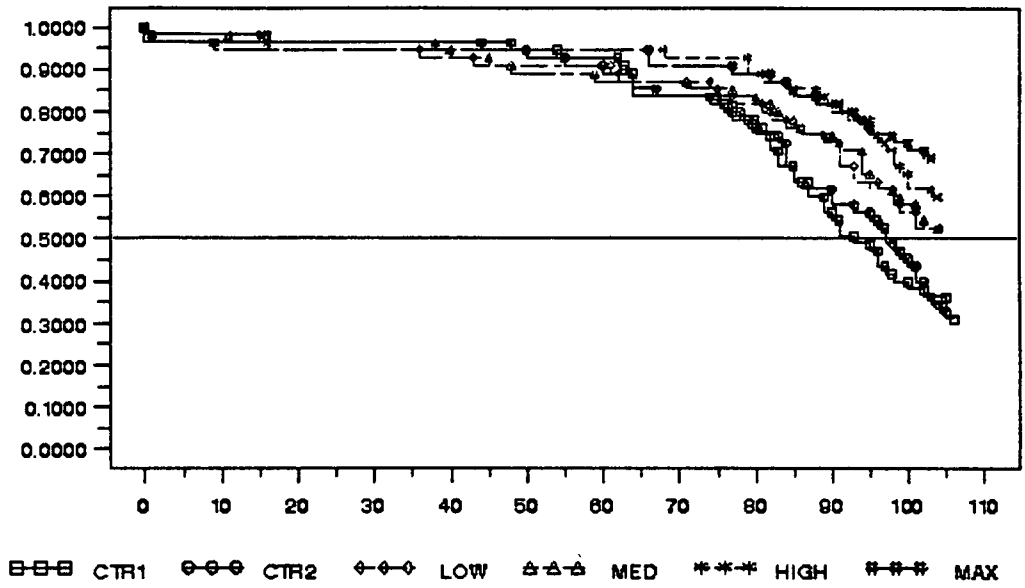


Table 12: Tumor Trend Tests for Male Rats in Study 99321

Organ Name	Organ Code	Tumor Name	Tumor Code	Natural Rate (in ctrl group)	CTRL1	CTRL2	LOW	MED	HIGH	MAX	Tumor type	pValue (Exact)	pValue (Asymp)
ADRENAL, CORTEX	AC	B-ADENOMA	15	5%	3	2	2	1	3	2	IN	0.6410	0.6279
ADRENAL, CORTEX	AC	M-CARCINOMA	424	2%	2	0	0	0	0	2	IN	0.3591	0.3175
AUDITORY SEB GL	AS	M-CARCINOMA, SEBACEOUS-SQ	133	2%	0	2	1	1	0	0	FA	0.9488	0.9326
AUDITORY SEB GL	AS	B-ADENOMA	431	.0%	0	0	0	0	0	1	IN	0.2216	0.0692
BONE, OTHER	BO	B-ODONTOMA	305	.9%	1	0	0	0	0	0	IN	N/A	N/A
BRAIN	BR	M-ASTROCYTOMA	296	4%	1	3	6	1	0	0	MX	0.9985	0.9956
BRAIN	BR	M-GRANULAR CELL TUMOR	332	.9%	0	1	0	0	0	0	FA	1.0000	0.8679
COLON	CO	M-CARCINOMA	447	.0%	0	0	0	0	0	1	IN	0.2249	0.0713
BONE, FEMUR	FE	M-OSTEOSARCOMA	350	.0%	0	0	1	0	0	0	FA	0.7261	0.7462

)OT/FOOTPAD	FP	M-FIBROSARCOMA	395	.0%	0	0	0	0	1	0	FA	0.4062	0.2383
HEMATO NEOPLASIA	HN	M-SARCOMA, HISTIOCYTIC	145	.9%	0	1	2	2	3	1	MX	0.4140	0.3949
HEMATO NEOPLASIA	HN	M-LEUKEMIA, GRANULOCYTIC	262	.0%	0	0	0	0	1	1	FA	0.1090	0.0516
HEMATO NEOPLASIA	HN	M-LYMPHOMA	89	.9%	0	1	2	2	1	0	MX	0.8186	0.8031
HEART	HT	M-ENDOCARDIAL SCHWANNOMA	454	.0%	0	0	0	0	0	1	IN	0.2249	0.0713
ILEUM	IL	M-CARCINOMA	496	.0%	0	0	0	1	0	0	IN	0.5952	0.6239
JEJUNUM	JE	M-CARCINOMA	404	.0%	0	0	0	0	1	0	FA	0.4211	0.2808
KIDNEY	KD	M-RENAL MESENCHYMAL TUMOR	341	.9%	1	0	0	0	1	1	MX	0.3186	0.2667
LIVER	LI	B-ADENOMA, HEPATOCELLULAR	123	5%	2	3	1	2	0	0	IN	0.9939	0.9856
LIVER	LI	M-CARCINOMA, HEPATOCELLUL	342	2%	2	0	1	0	0	0	IN	0.9686	0.9346
IVER	LI	M-HEMANGIOSARCOMA	444	.0%	0	0	0	0	0	1	IN	0.2249	0.0713
LUNG	LU	M-CARCINOMA, BRONCHIOLAR	310	.9%	0	1	0	0	0	0	IN	1.0000	0.7907
ADRENAL, MEDULLA	MA	B-PHEOCHROMOCYTOMA	188	10%	6	5	6	6	10	15	IN	0.0163	0.0141
ADRENAL, MEDULLA	MA	M-MALIGNANT PHEOCHROMOCYT	387	3%	1	2	0	0	0	4	MX	0.1704	0.1467
MAMMARY, MALE	MM	B-FIBROADENOMA	407	3%	1	2	0	0	0	0	IN	1.0000	0.9783
MAMMARY, MALE	MM	B-ADENOMA	433	.0%	0	0	0	0	0	1	IN	0.2333	0.0785
LN, MESENTERIC	MS	M-HEMANGIOSARCOMA	443	.9%	1	0	0	0	0	0	IN	1.0000	0.8832
PANCREAS	PA	M-CARCINOMA, ACINAR CELL	158	2%	2	0	0	0	0	1	MX	0.5759	0.5326
ANCREAS	PA	M-CARCINOMA, ISLET CELL	213	4%	1	3	3	0	1	0	IN	0.9701	0.9552

ANCREAS	PA	B-ADENOMA, ISLET CELL	286	9%	6	4	4	2	2	3	IN	0.9450	0.9366
PANCREAS	PA	B-ADENOMA, ACINAR CELL	52	.0%	0	0	2	0	0	0	IN	0.8340	0.8105
CAVITY, ABDOM	PC	M-MESOTHELIOMA	316	.0%	0	0	0	1	0	0	FA	0.4000	0.3566
CAVITY, ABDOM	PC	M-LIPOSARCOMA	334	.9%	0	1	0	0	0	0	FA	1.0000	0.7861
PITUITARY	PI	B-ADENOMA	61	63%	29	40	26	29	24	15	MX	1.0000	1.0000
PINNA	PN	B-ADENOMA, BASAL CELL	402	.9%	1	0	0	0	0	0	IN	1.0000	0.8643
PINNA	PN	B-PAPILLOMA, SQUAMOUS CEL	403	.9%	1	0	0	0	0	0	IN	1.0000	0.8643
PARATHYROID	PT	B-ADENOMA	102	4%	1	3	1	1	2	0	IN	0.8250	0.8103
SALIV GL, MANDIB	SG	M-CARCINOMA	372	.9%	1	0	0	0	0	0	FA	1.0000	0.8773
SKIN	SK	B-KERATOACANTHOMA	274	3%	2	1	2	1	0	0	IN	0.9748	0.9588
SKIN	SK	M-LIPOSARCOMA	287	.9%	1	0	0	0	0	0	FA	1.0000	0.8585
KIN	SK	M-CARCINOMA, BASAL CELL	412	.0%	0	0	1	1	0	0	IN	0.6767	0.6998
SKIN	SK	M-CARCINOMA, SQUAMOUS CEL	417	.0%	0	0	1	0	0	0	IN	0.6308	0.6792
SKIN	SK	B-TRICHOEPITHELIOMA	418	.0%	0	0	0	1	0	0	IN	0.4615	0.4886
SKIN	SK	B-PAPILLOMA, SQUAMOUS CEL	441	.9%	0	1	1	0	0	3	IN	0.1463	0.1251
SKIN	SK	B-FIBROMA	450	2%	1	1	3	0	1	0	IN	0.9623	0.9495
SKIN	SK	B-LIPOMA	465	.0%	0	0	0	0	2	0	IN	0.2527	0.2013
SKIN	SK	B-ADENOMA, SEBACEOUS/SQUA	470	.9%	0	1	0	0	0	0	IN	1.0000	0.8841
SKIN	SK	M-FIBROSARCOMA	62	4%	2	2	3	0	0	0	MX	0.9972	0.9901
MUSCLE, SKELETAL	SM	M-FIBROSARCOMA	304	.9%	1	0	1	0	0	0	FA	0.9266	0.8959
SPLEEN	SP	M-HEMANGIOSARCOMA	338	2%	1	1	0	0	0	0	IN	1.0000	0.9350

ГОМАСН, GL	ST	M- LEIOMYOSARCOMA -	437	.9%	1	0	0	0	0	0	IN	1.0000	0.8841
TESTIS	TE	B- INTERSTITIAL CELL TUMOR	201	5%	2	3	2	2	0	1	IN	0.9355	0.9209
TAIL	TI	B- KERATOACANTHOMA	493	.0%	0	0	1	0	0	0	IN	0.8333	0.8982
THYROID	TY	B-ADENOMA, FOLLICULAR CEL	108	5%	3	2	2	1	1	2	IN	0.6827	0.6707
THYROID	TY	B-ADENOMA, C-CELL	205	11%	5	7	8	3	3	4	IN	0.9524	0.9455
THYROID	TY	M- CARCINOMA, FOLLICULAR C	480	.0%	0	0	0	0	1	0	IN	0.4201	0.2813
THYROID	TY	M- CARCINOMA, C-CELL	495	.0%	0	0	0	2	0	0	IN	0.6938	0.6633

APPEARS THIS WAY
ON ORIGINAL

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

/s/

Roswitha Kelly
8/13/02 03:45:18 PM
BIOMETRICS

Gang Chen
8/14/02 07:35:38 AM
BIOMETRICS

APPEARS THIS WAY
ON ORIGINAL

- Statistical Review and Evaluation

Review of Mouse Carcinogenicity Studies

NDA#: 21-436

APPLICANT: **Otsuka Pharm**

NAME OF DRUG: **Aripiprazole**

INDICATION: **Schizophrenia**

STUDIES REVIEWED: **Mouse Studies: Otsuka Study No. 011487 and Otsuka Study No. 011932 in Volumes 1.67 and 1.71**

PHARMACOLOGY REVIEWER: **Lois Freed, Ph.D. (HFD-120)**

STATISTICAL REVIEWER: **Roswitha Kelly, M.S. (HFD-710)**

Roswitha Kelly, M.S.
Mathematical Statistician

George Chi, Ph.D.
Director, Division of Biometrics 1

Distribution:

HFD-120/Arch. NDA 21-436
HFD-120/Barry Rossloff Ph.D.
HFD-120/Lois Freed, Ph.D.
HFD-710/Kun Jin, Ph.D.
HFD-710/George Chi, Ph.D.
HFD-710/Roswitha Kelly, M.S.
HDS-700/Charles Anello, Sc.D.

This review consists of 7 pages of text and 14 pages of Tables and Figures.
File Directory: c:\Data\My Documents\aripiprazole_mice.doc

Table of Contents

Table of Tables

Table of Figures

1.0 Introduction

The sponsor has submitted two rat and two mouse carcinogenicity studies. As there is basically double the information of the usual bioassay, the multiplicity problem inherent in carcinogenicity analyses is increased. This reviewer, however, performed no further adjustment on the usual 0.025 and 0.005 levels of significance in trend for rare and common tumors, as there is no guidance on this issue. High levels of significance and consistency across gender and studies should be considered when interpreting any findings. This reviewer wrote a separate review for each species but presented the main findings from both species in the summary section.

2.0 Otsuka Study No. 011487 in Mice

This is a 104 week study of the aripiprazole in the diet of ICR (Crj:CD-1) mice at dose levels of 0, 1, 3, and 10 mg/kg/day. Groups of 60 animals/gender comprised the main study. Additional 8 animals per treatment group were maintained for 52 weeks to study plasma concentrations. The animals were housed individually and food and water were available ad lib. All main-study animals were fully histopathologically examined. The sponsor used a two-tailed log-rank test for mortality and one-tailed Cochran-Armitage trend tests for incidence of neoplastic lesions and number of females with pituitary or mammary gland tumors. Further, one-tailed Peto's onset rate and death rate methods were used for the incidence of mammary gland tumors and pituitary tumors respectively.

2.1 Sponsor's Findings for Study 011487

The sponsor observed no statistically significant differences in mortality between groups for either gender. The final mortality rates (number of animals killed in extremis or found dead) were as follows:

Dose Group (mg/kg/day)	Males	Females
0	40/60 (67%)	36/60 (60%)
1	44/60 (73%)	37/60 (62%)
3	38/60 (63%)	46/60 (78%)
10	31/60 (52%)	43/60 (72%)

Average body weight of the male mice in the high dose was slightly lower (about 5%) than the controls throughout the administration period. Differences reached statistical significance for Week 1 to Week 40. The average body weight of the high dose females was slightly lower than the controls (about 3-4%) in the early weeks, but became similar to the controls for the remainder of the study.

Among the females, incidence of adenocarcinoma and of adenoacanthoma in the mammary glands and of adenoma in the anterior pituitary was significantly higher in the 3- and 10 mg/kg doses when compared to the controls. The corresponding trend tests reached statistical significance as well. No statistically significant increase in neoplastic findings was seen among the males, except for a comparison of hemangiomas in the liver

between male control and mid dose animals. The corresponding trend test or control-high dose comparison was not statistically significant. Similar findings were observed when considering decedents and moribund sacrifices and terminal sacrifices separately.

2.2 Reviewer's Findings for Study 011487

The intercurrent mortality among the female mice of this study showed an increase in mortality with dose, however the trend did not reach statistical significance (p=0.0821, Tables 1-2, Figure 1). Tumor findings in the pituitary and mammary gland were highly statistically significant: adenoma in the anterior pituitary: p=0.0000; adenocarcinoma in the mammary gland: p=0.0000, and adenoacanthoma in the mammary gland: p=0.0011, (Table 3).

Among the males, there was statistically significantly better survival among the treated than among the control group (p=0.0138, Table 4-5, Figure 2). There were no positive increases in tumor findings that reached statistical significance (Table 6).

2.3 Validity of the Male Mouse of Study 011487

As there were no statistically significant tumor trends among the male mice in this study, its validity needs to be assessed. Two criteria are set up for this purpose (Haseman[12], Chu et al.[3], and Bart et al.[4]):

i) was a sufficient number of animals exposed long enough to allow for late-developing tumors, and

ii) did the high dose provide a sufficient tumor challenge?

The number of animals and length of exposure can be assessed at weeks 52, 80-90, and at termination, but are generally considered sufficient if 20-30 animals survive through weeks 80-90. The high dose is expected to be close to the MTD to present a sufficient tumor challenge. Suppression in survival when compared to the controls and/or average body weight differences of about 10 percent, especially during the first year of treatment, are indicators that the high dose is close to the MTD. For this study, 19 animals had died by the end of week 91 and 50% survived till terminal sacrifice. Therefore, there was a sufficient number of animals living long enough to satisfy the first criterion. There was no reduction in survival with dose. The sponsor reported average body weights of the high dose group being below the controls'. Though the difference is only about 5 percent, it was observed very early in the study and was maintained for most of the two years. These findings may sufficiently indicate that the high dose was close to the MTD and that the study can be considered valid.

---

[1] Haseman: Statistical Issues in the Design, Analysis and Interpretation of Animal Carcinogenicity Studies, *Environmental Health Perspectives*, Vol. 58, pp 385-392, 1984.

[2] Haseman: Issues in Carcinogenicity Testing: Dose Selection, *Fundamental and Applied Toxicology*, Vol. 5, pp. 66-78, 1985.

[3] Chu, Cueto, Ward: Factors in the Evaluation of 200 National Cancer Institute Carcinogenicity Bioassays, *Journal of Toxicology and Environmental Health*, Vol. 8, pp 251-280, 1981.

[4] Bart, Chu, Tarone: Statistical Issues in Interpretation of Chronic Bioassay Tests for Carcinogenicity, *Journal of the National Cancer Institute*, pp. 957-974, 1979.

3.0 Otsuka Study No. 011932 in Mice

This study consisted only of two groups of ICR (Crj:CD-1) SPF mice: a control group of 60 animals/gender and a 30 mg/kg/day group of another 60 animals/gender. Animals were randomly allocated on the basis of body weight measured 1 week prior to the assignment. The compound was administered orally in the diet for 104 weeks for the males. Dosing was terminated at week 100 for females due to 75% mortality. The animals were housed individually and water and diet were available ad lib. Additional 8 animals per gender were dosed and maintained for 52 weeks for determination of plasma concentrations. All tissues were histopathologically examined for all animals of the main study. Mortality was assessed by a two-tailed life table analysis. Tumor incidences were analyzed by a one-tailed Peto test.

3.1 Sponsor's Findings of Study 011932

The sponsor found increased survival in the treated males but decreased survival in the treated females compared to their controls. The final mortality rates (number of animals killed in extremis or found dead) at week 104 (males) or week 100 (females) were as follows:

Dose Group (mg/kg/day)	Males	Females
0	45/60 (75%)	35/60 (58%)
30	33/60 (55%)	45/60 (75%)

Mean body weight in the treated males was approximately 10% lower than the controls throughout the treatment period. This difference was statistically significant at almost all weeks. The body weight of the treated females was approximately 5% lower than the controls from Week 1 through Week 16. Thereafter, the average weight became similar to the controls' with no statistically significant differences at any time.

There were no significant differences in neoplastic lesions among the control and treated males. Among the females, there were statistically significant increases in adenoma in the anterior pituitary and in adenocarcinoma and adenoacanthoma in the mammary glands. Also, the number of animals with epithelial mammary gland tumors was significantly greater than the controls'.

3.2 Reviewer's Findings for Study 011932

There was only one control and one treated (30 mg/kg/day) group per gender. Therefore, all statistical tests are pair-wise comparisons, two-sided for mortality and one-sided for tumor findings. The intercurrent mortality for the female mice showed higher mortality in the treated group, which reached statistical significance (p=0.0391, Table 7-8, Figure 3). Among the tumor findings, adenoma in the anterior pituitary (p=0.0023),

adenocarcinoma in the mammary gland (p=0.0000), and adenoacanthoma in the mammary gland (p=0.0000) were highly statistically significant (Table 9). These findings are consistent with the sponsor's.

The male mice of this study had statistically significant better survival in the treated group than in the control (p=0.0234, Table 10-11, Figure 4). Among the tumor findings, none reached statistical significance (Table 12).

3.3 Validity of Male Mouse Study 011932

The same criteria as noted above to evaluate the male mice of Study 011487 are being applied to the male mice of this study, as no statistically significant increase in tumors were observed. Survival was good for both the control and the treated group, and the number of animals living long enough is not an issue. Survival was significantly better for the treated than for the control group, and therefore mortality cannot be used as a criterion for assessing whether 30 mg/kg/day presented a sufficient tumor challenge in these animals. The sponsor's average bodyweight data indicated an early and sustained differential of about 10% for the treated males compared to the controls. This finding implies that the high dose was close to the MTD for these animals.

4.0 Summary

For Study 011487, 60 animals/gender received aripiprazole in the diet at levels of 0, 1, 3, and 10 mg/kg/day. The sponsor observed no statistically significant difference in mortality patterns for either the male or female mice. This reviewer, however, observed that the increased survival among males reached statistical significance and that the decreased survival among females approached statistical significance. This difference in conclusion about survival is minor, since all tumor findings were tested by age-adjusted methods. The conclusions based on the tumor findings are the same as the sponsor's, except that the sponsor did not note the very high levels of statistical significance. Adenoma in the anterior pituitary gland, and adenocarcinoma and adenoacanthoma in the mammary glands were highly statistically significantly increased among the females. Among the male mice, there were no statistically significant increases in tumor findings, however, the length of exposure and number of animals alive at study end were acceptable. Whether the high dose presented a sufficient tumor challenge is assessed by suppressed body weights, since there was no increased mortality for these animals. The sponsor reported an early and sustained reduction of about 5% in average body weights for the high dose males compared to the controls. This reviewer assumes that this differential is sufficient to conclude that this was a valid study.

For Study 011932, where 60 control animals/gender were compared to 60 mice treated with 30 mg/kg/day in the diet, the sponsor's and this reviewer's findings and conclusions agree. Mortality of 75% prompted the sponsor to stop dosing the females at week 100. The increased mortality reached statistical significance. Among the females, highly significant increases in tumor incidence rates were found for adenoma in the anterior

pituitary gland and for adenocarcinoma and adenoacanthoma in the mammary gland. For the males, survival was significantly better among the treated than among the controls. There were no statistically significant increases in neoplastic findings among the males, however, length of exposure, number of animals alive at study end, and suppressed body weights indicated that this was a valid study.

The major findings of the two mouse studies are summarized below:

MICE	Otsuka Study 011487		Otsuka Study 011932	
	Females	Males	Females	Males
Survival	NS	Sign. increased	Sign. decreased	Sign. increased
Anterior Pituitary, Adenoma	Sign. increased	NS	Sign. increased	NS
Mammary Gland, Adenocarcinoma	Sign. increased	NS	Sign. increased	NS
Mammary Gland, Adenoacanthoma	Sign. increased	NS	Sign. increased	NS
Validity	N/A	Yes	N/A	Yes

For completeness, the major findings of the two rat studies are given as well:

RATS	Otsuka Study 009489		BMS Study 99321	
	Females	Males	Females	Males
Survival	NS	NS	Sign. increased	Sign. increased
Mammary Gland, Fibroadenoma	Sign. increased	NS	NS	NS
Adrenal Cortex, Carcinoma	NS	NS	Sign. increased	NS
Validity	N/A	Yes	N/A	MTD exceeded

Overall, it appears that the long-term administration of aripiprazole in the doses given resulted in increased tumor findings in female rats or mice. The p-values in each case are highly statistically significant. No increase in tumor incidence rates was observed among the males of any of the studies. All but one of these male studies were judged to be valid. The maximum dose in Study No. 99321 was judged to be well beyond the MTD, based on much lower average body weights of these animals compared to the controls.

Otsuka Study 011487

Table 1: Study 011487, Number of Animals Dying during Given Time Intervals, Female Mice

	Treatment Group				Total
	CTRL	LOW	MED	HIGH	
	N	N	N	N	N
Week 0-52	4	4	6	7	21
53-78	10	11	12	12	45
79-91	10	9	16	12	47
92-103	11	12	11	12	46
104-104	25	24	15	17	81
Total	60	60	60	60	240

Table 2: Study 011487, Dose-Mortality Trend Tests* for Female Mice

Method	Time-Adjusted Trend Test	Statistic	P Value
Cox	Dose-Mortality Trend	3.02	0.0821
	Depart from Trend	2.83	0.2430
	Homogeneity	5.85	0.1191
Kruskal-Wallis	Dose-Mortality Trend	2.88	0.0898
	Depart from Trend	1.94	0.3785
	Homogeneity	4.82	0.1854

* The results are produced by: Trend and Homogeneity Analyses of Proportions and Life Table Data Version 2:1, by Donald G. Thomas, National Cancer Institute.

APPEARS THIS WAY
ON ORIGINAL

Figure 1: Study 011487, Kaplan Meier Survival Curves in Female Mice

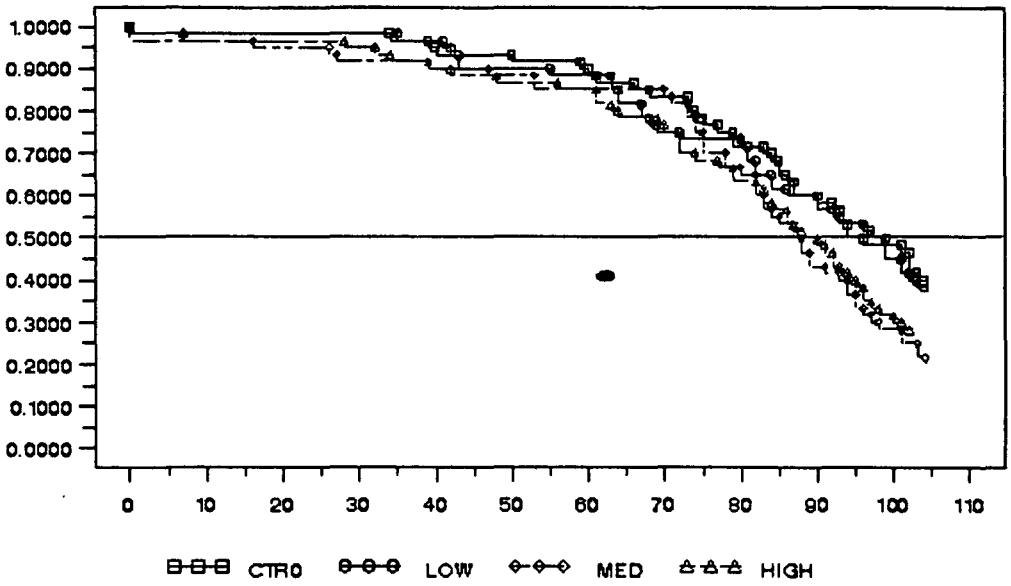


Table 3: Study 011487, Test for Dose-Dependent Linear Trend in Tumors, Female Mice

Organ Code	Organ Name	Tumor Code	Tumor Name	CTRL	LOW	MED	HIGH	pValue (Exact)	pValue (Asymp)	Natural Tumor # in control group	Natural Rate (in ctrl group)	Tumor type
01	Heart	0180	Malignant schwannoma (Hea	0	0	1	0	0.5957	0.6298	0	.0%	IN
03	Bone marrow(femur)	0342	Hemangioma (Bone marrow)	1	0	0	0	1.0000	0.8493	1	2%	IN
07	Thymus	0780	Malignant schwannoma (Thy	0	0	1	0	0.5957	0.6298	0	.0%	IN
07	Thymus	0784	Histiocytic sarcoma (Thym	1	0	0	0	1.0000	0.8678	1	2%	IN
09	Lymph nodes (mesenteric)	0880	Malignant schwannoma (Lym	0	0	0	1	0.2609	0.0698	0	.0%	IN
13	Spleen	1342	Hemangioma (Spleen)	0	2	5	1	0.5631	0.5784	0	.0%	MX
13	Spleen	1372	Hemangiosarcoma (Spleen)	1	0	0	0	1.0000	0.8160	1	2%	FA
13	Spleen	1380	Malignant schwannoma (Spl	0	0	1	1	0.2387	0.1763	0	.0%	IN

18	Lung	1835	Adenoma (Lung)	12	5	12	7	0.6500	0.6572	12	20%	MX
18	Lung	1865	Adenocarcinoma (Lung)	6	7	12	6	0.4467	0.4559	6	10%	MX
18	Lung	1880	Malignant schwannoma (Lun	0	0	1	0	0.5957	0.6298	0	0%	IN
18	Lung	1884	Histiocytic sarcoma (Lung	1	0	0	0	1.0000	0.8678	1	2%	IN
28	Esophagus	2880	Malignant schwannoma (Eso	0	0	0	1	0.2553	0.0668	0	0%	IN
29	Stomach (non-glandular po	2960	Squamous cell carcinoma (	0	1	0	0	0.6914	0.7419	0	0%	IN
30	Stomach (glandular portio	3076	Leiomyosarcoma (Stomach(g	0	0	1	0	0.5000	0.6083	0	0%	IN
31	Small intestine	3176	Leiomyosarcoma (Small int	0	0	1	0	0.5000	0.6083	0	0%	IN
32	Large intestine	3246	Leiomyoma (Large intestin	0	0	0	1	0.2099	0.0413	0	0%	IN
32	Large intestine	3276	Leiomyosarcoma (Large int	0	0	1	0	0.5000	0.6083	0	0%	IN
34	Liver	3435	Hepatocellular adenoma (L	1	1	0	1	0.4846	0.4392	1	2%	IN
34	Liver	3442	Hemangioma (Liver)	2	3	3	2	0.5607	0.5865	2	3%	MX
34	Liver	3467	Cholangiocarcinoma (Liver	0	0	0	1	0.2276	0.0503	0	0%	FA
34	Liver	3472	Hemangiosarcoma (Liver)	1	0	0	0	1.0000	0.8197	1	2%	IN
34	Liver	3476	Leiomyosarcoma (Liver)	0	0	1	0	0.5000	0.6083	0	0%	IN
34	Liver	3484	Histiocytic sarcoma (Live	1	1	0	0	0.9343	0.8764	1	2%	IN
36	Gallbladder	3635	Adenoma (Gallbladder)	0	0	1	1	0.1207	0.1067	0	0%	IN
37	Pancreas	3736	Islet cell adenoma (Pancr	0	0	2	0	0.4603	0.5753	0	0%	IN
37	Pancreas	3750	Schwannoma (Pancreas)	0	0	0	1	0.2099	0.0413	0	0%	IN
37	Pancreas	3776	Leiomyosarcoma (Pancreas)	0	0	1	0	0.5000	0.6083	0	0%	IN
38	Kidney	3835	Adenoma (Kidney)	1	1	0	0	0.9262	0.8653	1	2%	IN
38	Kidney	3880	Malignant schwannoma (Kid	0	0	1	0	0.5957	0.6298	0	0%	IN
41	Urinary bladder	4146	Leiomyoma (Urinary bladde	0	2	1	1	0.3341	0.3989	0	0%	IN
52	Ovary	5235	Luteoma (Ovary)	1	0	0	0	1.0000	0.8197	1	2%	IN
52	Ovary	5236	Adenoma (Ovary)	2	1	2	1	0.5354	0.5852	2	3%	IN
52	Ovary	5242	Hemangioma (Ovary)	1	3	1	0	0.9187	0.9160	1	2%	MX
52	Ovary	5246	Leiomyoma (Ovary)	0	1	0	0	0.7609	0.7825	0	0%	IN
52	Ovary	5264	Malignant granulosa-theca	0	0	1	0	0.4937	0.5987	0	0%	FA
52	Ovary	5276	Leiomyosarcoma (Ovary)	0	0	1	0	0.5000	0.6083	0	0%	IN

54	Uterus	5431	Endometrial stromal polyp	3	5	3	5	0.1945	0.1951	3	5%	IN
54	Uterus	5442	Hemangioma (Uterus)	2	2	0	0	0.9840	0.9510	2	3%	MX
54	Uterus	5446	Leiomyoma (Uterus)	0	0	1	1	0.2387	0.1763	0	.0%	IN
54	Uterus	5465	Adenocarcinoma (Uterus)	0	0	2	1	0.2164	0.2462	0	.0%	MX
54	Uterus	5472	Hemangiosarcoma (Uterus)	0	1	0	0	0.6914	0.7419	0	.0%	IN
54	Uterus	5474	Endometrial stromal sarco	0	0	0	1	0.3333	0.1065	0	.0%	IN
54	Uterus	5476	Leiomyosarcoma (Uterus)	1	0	0	0	1.0000	0.8197	1	2%	IN
54	Uterus	5484	Histiocytic sarcoma (Uter	0	1	0	0	0.6914	0.7419	0	.0%	IN
59	Pituitary	5935	Anterior adenoma (Pituita	2	4	8	14	0.0001	0.0000	2	3%	MX
59	Pituitary	5936	Adenoma in intermediate p	1	0	1	1	0.3182	0.3273	1	2%	MX
60	Thyroid	6035	Follicular adenoma (Thyro	1	0	0	0	1.0000	0.8197	1	2%	IN
60	Thyroid	6036	C-cell adenoma (Thyroid)	0	0	0	1	0.2609	0.0674	0	.0%	IN
62	Adrenal	6236	Subcapsular cell adenoma	0	2	0	0	0.7222	0.7967	0	.0%	IN
62	Adrenal	6239	Pheochromocytoma (Adrenal	0	0	2	0	0.4804	0.5937	0	.0%	IN
64	Cerebrum	6433	Glioma (Cerebrum)	0	0	0	1	0.2435	0.0588	0	.0%	FA
64	Cerebrum	6436	Meningioma (Cerebrum)	0	0	1	0	0.3951	0.5475	0	.0%	IN
68	Spinal cord (thoracic)	6780	Malignant schwannoma (Spi	0	0	0	1	0.2276	0.0503	0	.0%	FA
72	Bone (sternum)	7280	Malignant schwannoma (Bon	0	0	1	0	0.5957	0.6298	0	.0%	IN
74	Bone (vertebra)	7242	Hemangioma (Bone)	0	0	1	0	0.3951	0.5475	0	.0%	IN
81	Harderian gland	8135	Adenoma (Harderian gland)	4	2	1	2	0.6763	0.6887	4	7%	IN
86	Skin	8631	Papilloma (Skin)	1	0	0	0	1.0000	0.8197	1	2%	IN
86	Skin	8635	Sebaceous gland adenoma (	0	0	1	0	0.3951	0.5475	0	.0%	IN
86	Skin	8640	Fibroma (Skin)	1	0	0	0	1.0000	0.8303	1	2%	FA
86	Skin	8642	Hemangioma (Skin)	0	0	0	1	0.2553	0.0668	0	.0%	IN
86	Skin	8650	Schwannoma (Skin)	0	0	1	1	0.1715	0.1505	0	.0%	MX
86	Skin	8654	Histiocytoma (Skin)	0	0	1	0	0.5333	0.6207	0	.0%	IN
86	Skin	8660	Squamous cell carcinoma (	0	0	0	1	0.2412	0.0583	0	.0%	FA
86	Skin	8661	Basal cell carcinoma (Ski	0	1	0	0	0.6907	0.7359	0	.0%	FA
86	Skin	8670	Fibrosarcoma (Skin)	0	1	0	0	0.7872	0.8036	0	.0%	IN
86	Skin	8678	Osteosarcoma (Skin)	0	0	1	0	0.5076	0.5952	0	.0%	FA
86	Skin	8680	Malignant schwannoma (Ski	0	0	1	1	0.2165	0.1746	0	.0%	IN
86	Skin	8684	Histiocytic sarcoma (Skin	0	0	2	0	0.4603	0.5753	0	.0%	IN

95	Mammary gland	9535	Adenoma (Mammary gland)	0	0	0	2	0.0696	0.0147	0	0%	IN
95	Mammary gland	9565	Adenocarcinoma (Mammary g	1	5	13	19	0.0000	0.0000	1	2%	MX
95	Mammary gland	9566	Adenoacanthoma (Mammary g	0	2	15	10	0.0019	0.0011	0	0%	MX
95	Mammary gland	9568	Carcinosarcoma (Mammary g	1	0	1	1	0.3441	0.3376	1	2%	IN
99	General	9988	Myelogenic leukemia (Gene	1	0	0	0	1.0000	0.8437	1	2%	FA
99	General	9989	Malignant lymphoma (Gener	15	17	10	12	0.6388	0.6467	15	25%	MX

Table 4: Study 011487, Number of Animals Dying during Given Time Intervals, Male Mice

	Treatment Group				Total
	CTRL	LOW	MED	HIGH	
	N	N	N	N	N
Week 0-52	8	7	5	6	26
53-78	13	9	12	4	38
79-91	9	12	9	9	39
92-103	9	16	10	11	46
104-104	21	16	24	30	91
Total	60	60	60	60	240

Table 5: Study 011487, Dose-Mortality Trend Tests* for Male Mice

Method	Time-Adjusted Trend Test	Statistic	P Value
Cox	Dose-Mortality Trend	6.06	0.0138
	Depart from Trend	0.72	0.6978
	Homogeneity	6.78	0.0793
Kruskal-Wallis	Dose-Mortality Trend	6.04	0.0140
	Depart from Trend	0.20	0.9026
	Homogeneity	6.24	0.1003

* The results are produced by: Trend and Homogeneity Analyses of Proportions and Life Table Data Version 2:1, by Donald G. Thomas, National Cancer Institute.

Figure 2: Study 011487, Kaplan Meier Survival Curves in Male Mice

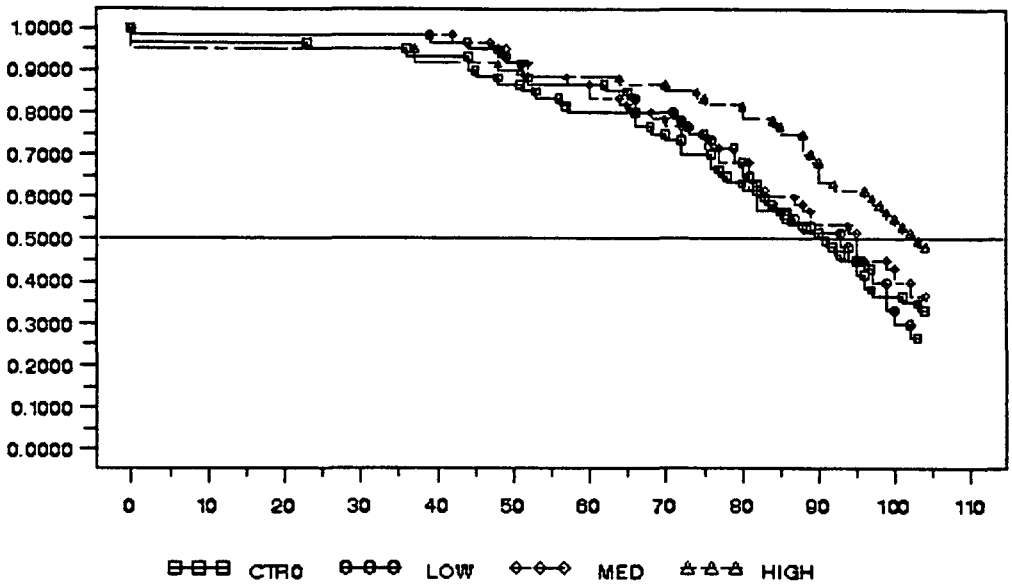


Table 6: Study 011487, Test for Dose-Dependent Linear Trend in Tumors, Male Mice

Organ Code	Organ Name	Tumor Code	Tumor Name	CTRL	LOW	MED	HIGH	pValue (Exact)	pValue (Asymp)	Natural Tumor # in control group	Natural Rate (in ctrl group)	Tumor type
01	Heart	0184	Histiocytic sarcoma (Hear	1	0	0	0	1.0000	0.8246	1	2%	IN
03	Bone marrow(femur)	0342	Hemangioma (Bone marrow)	0	0	0	1	0.3297	0.1043	0	.0%	IN
09	Lymph nodes (mesenteric)	0842	Hemangioma (Lymph nodes)	0	1	0	0	0.7692	0.8170	0	.0%	IN
13	Spleen	1342	Hemangioma (Spleen)	1	1	2	2	0.2600	0.2716	1	2%	IN
13	Spleen	1356	Mesotheliom a (Spleen)	0	0	0	1	0.3082	0.0919	0	.0%	FA
13	Spleen	1372	Hemangiosar coma (Spleen)	1	0	1	0	0.7811	0.8155	1	2%	IN
18	Lung	1835	Adenoma (Lung)	11	10	10	16	0.1381	0.1370	11	18%	MX
18	Lung	1865	Adenocarcin oma (Lung)	16	13	12	15	0.6389	0.6438	16	27%	MX
23	Tongue	2331	Papilloma	0	0	1	0	0.5934	0.6641	0	.0%	IN

			(Tongue)									
31	Small intestine	3165	Adenocarcinoma (Small int	0	0	0	2	0.1062	0.0305	0	.0%	IN
34	Liver	3435	Hepatocellular adenoma (L	9	16	11	11	0.6863	0.6905	9	15%	MX
34	Liver	3442	Hemangioma (Liver)	2	4	8	4	0.5899	0.6003	2	3%	MX
34	Liver	3465	Hepatocellular carcinoma	9	13	10	12	0.3713	0.3765	9	15%	MX
34	Liver	3466	Hepatoblastoma (Liver)	0	1	0	0	0.7758	0.8011	0	.0%	FA
34	Liver	3472	Hemangiosarcoma (Liver)	1	0	4	0	0.8451	0.8420	1	2%	MX
34	Liver	3484	Histiocytic sarcoma (Live	1	0	1	0	0.7402	0.7702	1	2%	MX
36	Gallbladder	3635	Adenoma (Gallbladder)	0	0	1	1	0.2809	0.2513	0	.0%	IN
37	Pancreas	3756	Mesothelioma (Pancreas)	0	0	0	1	0.3082	0.0919	0	.0%	FA
37	Pancreas	3772	Hemangiosarcoma (Pancreas	1	0	0	0	1.0000	0.8737	1	2%	IN
37	Pancreas	3784	Histiocytic sarcoma (Panc	0	0	1	0	0.4565	0.5919	0	.0%	IN
38	Kidney	3835	Adenoma (Kidney)	0	0	1	1	0.2821	0.2488	0	.0%	IN
38	Kidney	3842	Hemangioma (Kidney)	0	0	1	0	0.5934	0.6641	0	.0%	IN
43	Testis	4337	Interstitial cell tumor (	1	0	0	1	0.5531	0.4340	1	2%	IN
43	Testis	4342	Hemangioma (Testis)	1	1	0	0	0.9652	0.8792	1	2%	IN
46	Seminal vesicle	4635	Adenoma (Seminal vesicle)	2	0	0	0	1.0000	0.9384	2	3%	IN
47	Coagulating gland	4735	Adenoma (Coagulating glan	0	0	0	1	0.2391	0.0556	0	.0%	IN
47	Coagulating gland	4756	Mesothelioma (Coagulating	0	0	0	1	0.3082	0.0919	0	.0%	FA
48	Prostate	4856	Mesothelioma (Prostate)	0	0	0	1	0.3082	0.0919	0	.0%	FA
48	Prostate	4872	Hemangiosarcoma (Prostate	1	0	0	0	1.0000	0.8737	1	2%	IN
60	Thyroid	6035	Follicular adenoma (Thyro	1	0	0	1	0.4844	0.3663	1	2%	IN
62	Adrenal	6239	Pheochromocytoma (Adrenal	0	1	0	0	0.8043	0.7745	0	.0%	IN
64	Cerebrum	6433	Glioma	0	1	0	0	0.7692	0.8170	0	.0%	IN

			(Cerebrum)									
64	Cerebrum	6436	Meningioma (Cerebrum)	0	0	1	0	0.5249	0.6227	0	.0%	FA
72	Bone (sternum)	7272	Hemangiosarcoma (Bone)	0	0	1	0	0.5121	0.6112	0	.0%	FA
74	Bone (vertebra)	7278	Osteosarcoma (Bone)	1	0	0	1	0.4399	0.3284	1	2%	FA
81	Harderian gland	8135	Adenoma (Harderian gland)	8	5	7	5	0.7935	0.7972	8	13%	IN
86	Skin	8631	Papilloma (Skin)	1	0	0	0	1.0000	0.8737	1	2%	IN
86	Skin	8640	Fibroma (Skin)	1	0	0	0	1.0000	0.8737	1	2%	IN
86	Skin	8642	Hemangioma (Skin)	0	0	1	0	0.5934	0.6641	0	.0%	IN
86	Skin	8650	Schwannoma (Skin)	0	0	1	2	0.0666	0.0458	0	.0%	IN
86	Skin	8672	Hemangiosarcoma (Skin)	0	0	1	0	0.5739	0.6514	0	.0%	FA
86	Skin	8676	Leiomyosarcoma (Skin)	0	0	0	2	0.0984	0.0264	0	.0%	FA
86	Skin	8680	Malignant schwannoma (Ski	0	1	0	0	0.7784	0.7953	0	.0%	FA
86	Skin	8684	Histiocytic sarcoma (Skin	0	1	1	1	0.4148	0.4604	0	.0%	IN
99	General	9988	Myelogenic leukemia (Gene	0	0	0	2	0.0630	0.0120	0	.0%	FA
99	General	9989	Malignant lymphoma (Gener	7	9	10	6	0.8272	0.8290	7	12%	MX

Table 7: Study 011932, Number of Animals Dying during Given Time Intervals, Female Mice

	Treatment Group		Total
	CTRL1	LOW	
	N	N	N
Week 0-52	1	6	7
53-78	12	12	24
79-91	12	16	28
92-100	10	11	21
101-101	25	15	40
Total	60	60	120

Table 8: Study 011932, Dose-Mortality Trend Tests* for Female Mice

Method	Time-Adjusted Trend Test	Statistic	P Value
Cox	Dose-Mortality Trend	4.25	0.0391
Kruskal-Wallis	Dose-Mortality Trend	4.70	0.0302

* The results are produced by: Trend and Homogeneity Analyses of Proportions and Life Table Data Version 2:1, by Donald G. Thomas, National Cancer Institute.

Figure 3: Study 011932, Kaplan Meier Survival Curves in Female Mice

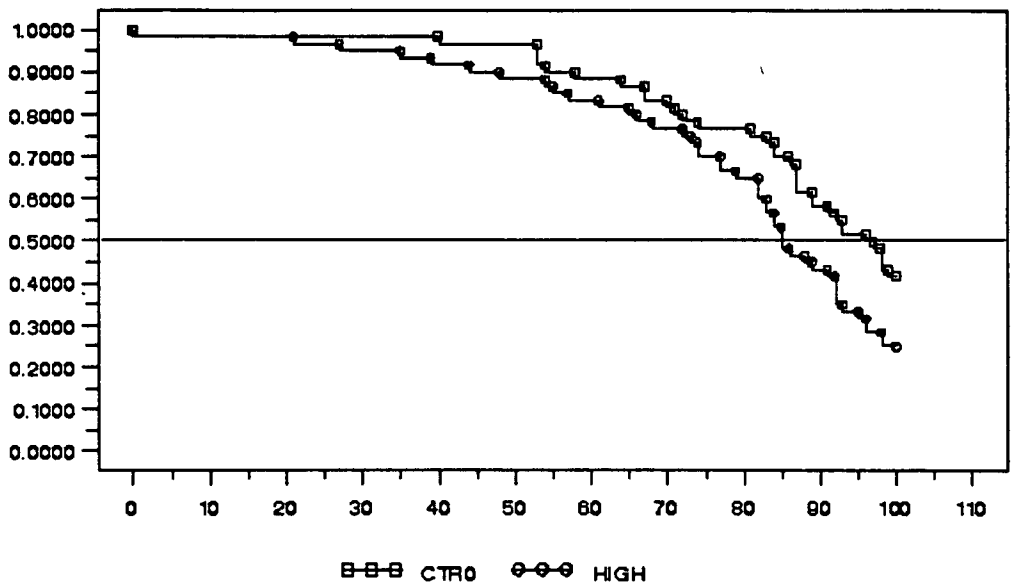


Table 9: Study 011932, Test for Dose-Dependent Linear Trend in Tumors, Female Mice

Organ Code	Organ Name	Tumor Code	Tumor Name	CTRL1	LOW	pValue (Exact)	pValue (Asymp)	Natural Tumor # in control group	Natural Rate (in ctrl group)	Tumor type
13	Spleen	1342	Hemangioma (Spleen)	1	0	1.0000	0.8604	1	2%	IN
13	Spleen	1354	Histiocytoma (Spleen)	0	1	0.5714	0.2026	0	0%	IN
13	Spleen	1372	Hemangiosarcoma (Spleen)	2	0	1.0000	0.9059	2	3%	IN
13	Spleen	1384	Histiocytic sarcoma (Sple	0	2	0.2993	0.1033	0	0%	IN
18	Lung	1835	Adenoma (Lung)	11	12	0.4067	0.3228	11	18%	IN
18	Lung	1865	Adenocarcinoma (Lung)	10	9	0.4906	0.3978	10	17%	MX
18	Lung	1884	Histiocytic sarcoma (Lung	0	1	0.4237	0.1287	0	0%	FA
31	Small intestine	3176	Leiomyosarcoma (Small int	0	1	0.3889	0.1114	0	0%	FA
34	Liver	3435	Hepatocellular adenoma (L	2	0	1.0000	0.8717	2	3%	IN
34	Liver	3442	Hemangioma (Liver)	1	2	0.5612	0.3337	1	2%	IN
34	Liver	3465	Henatocellular	2	0	1.0000	0.8717	2	3%	IN

			carcinoma							
34	Liver	3484	Histiocytic sarcoma (Live	3	1	0.8752	0.7422	3	5%	FA
36	Gallbladder	3635	Adenoma (Gallbladder)	0	1	0.5714	0.2026	0	.0%	IN
37	Pancreas	3735	Acinar cell adenoma (Panc	0	1	0.5238	0.1788	0	.0%	IN
38	Kidney	3835	Adenoma (Kidney)	0	1	0.5238	0.1788	0	.0%	IN
41	Urinary bladder	4146	Leiomyoma (Urinary bladde	0	1	0.3846	0.1092	0	.0%	IN
52	Ovary	5232	Granulosa cell tumor (Ova	2	1	0.9575	0.8395	2	3%	IN
52	Ovary	5236	Adenoma (Ovary)	1	1	0.6154	0.3649	1	2%	IN
54	Uterus	5431	Endometrial stromal polyp	8	2	0.9876	0.9654	8	13%	IN
54	Uterus	5446	Leiomyoma (Uterus)	5	1	0.9670	0.9064	5	8%	IN
54	Uterus	5465	Adenocarcinoma (Uterus)	2	0	1.0000	0.9280	2	3%	MX
54	Uterus	5484	Histiocytic sarcoma (Uter	3	0	1.0000	0.9732	3	5%	IN
59	Pituitary	5935	Anterior adenoma (Pituita	5	14	0.0044	0.0023	5	8%	MX
60	Thyroid	6035	Follicular adenoma (Thyro	2	0	1.0000	0.9543	2	3%	IN
62	Adrenal	6236	Subcapsular cell adenoma	1	0	1.0000	0.7907	1	2%	IN
62	Adrenal	6239	Pheochromocytoma (Adrenal	2	0	1.0000	0.9179	2	3%	IN
81	Harderian gland	8135	Adenoma (Harderian gland)	3	4	0.3930	0.2584	3	5%	IN
86	Skin	8631	Papilloma (Skin)	1	0	1.0000	0.8827	1	2%	IN
86	Skin	8632	Keratoacantoma (Skin)	0	1	0.3750	0.1045	0	.0%	IN
86	Skin	8636	Myxoma (Skin)	1	0	1.0000	0.8604	1	2%	IN
86	Skin	8650	Schwannoma (Skin)	0	1	0.5238	0.1788	0	.0%	IN
86	Skin	8672	Hemangiosarcoma (Skin)	1	0	1.0000	0.7882	1	2%	FA
86	Skin	8676	Leiomyosarcoma (Skin)	1	0	1.0000	0.8148	1	2%	FA
95	Mammary gland	9535	Adenoma (Mammary gland)	0	1	0.5238	0.1788	0	.0%	IN
95	Mammary gland	9542	Hemangioma (Mammary gland	0	1	0.5714	0.2026	0	.0%	IN
95	Mammary	9565	Adenocarcino	1	14	0.0000	0.0000	1	2%	MX

	gland		ma (Mammary g							
95	Mammary gland	9566	Adenoacantho ma (Mammary g	0	11	0.0000	0.0000	0	0%	MX
95	Mammary gland	9568	Carcinosarcom a (Mammary g	1	3	0.2307	0.1102	1	2%	MX
99	General	9984	Histiocytic sarcoma (Gene	0	1	0.4237	0.1287	0	0%	FA
99	General	9989	Malignant lymphoma (Gener	13	9	0.6290	0.5513	13	22%	MX

Table 10: Study 011932, Number of Animals Dying during Given Time Intervals, Male Mice

	Treatment Group		Total
	CTRL1	LOW	
	N	N	N
Week			
0-52	2	4	6
53-78	11	8	19
79-91	14	8	22
92-104	18	12	30
105-105	15	28	43
Total	60	60	120

Table 11: Study 011932, Dose-Mortality Trend Tests* for Male Mice

Method	Time-Adjusted Trend Test	Statistic	P Value
Cox	Dose-Mortality Trend	5.14	0.0234
Kruskal-Wallis	Dose-Mortality Trend	4.32	0.0376

* The results are produced by: Trend and Homogeneity Analyses of Proportions and Life Table Data Version 2:1, by Donald G. Thomas, National Cancer Institute.

Figure 4: Study 011932, Kaplan Meier Survival Curves in Male Mice

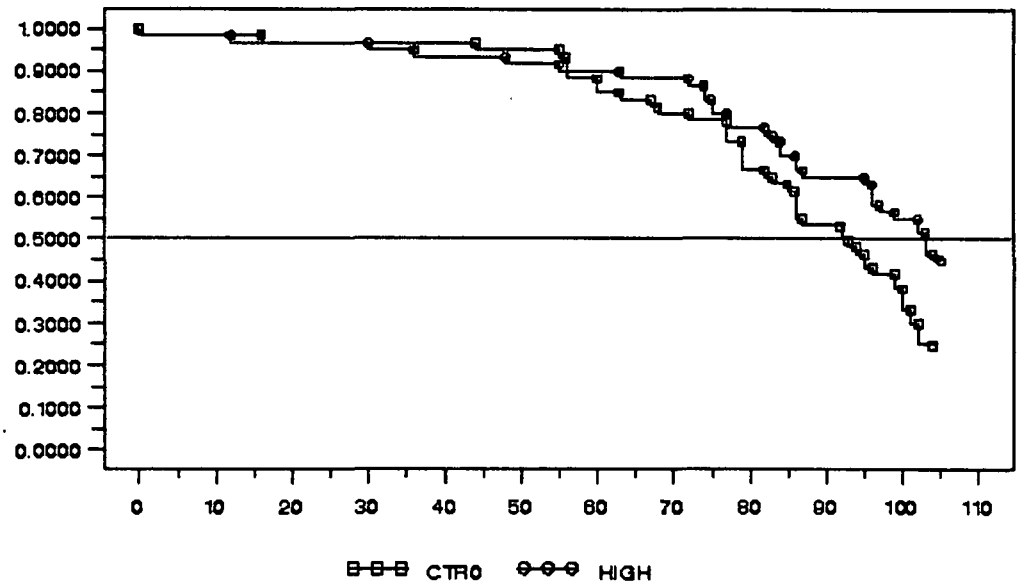


Table 12: Study 011932, Test for Dose-Dependent Linear Trend in Tumors, Male Mice

Organ Name	Organ Code	Tumor Name	Tumor Code	Natural Rate (in ctrl group)	CTRL1	LOW	Tumor type	pValue (Exact)	pValue (Asymp)
Spleen	13	Hemangioma (Spleen)	1342	2%	1	1	IN	0.7907	0.7900
Lung	18	Adenoma (Lung)	1835	15%	9	11	IN	0.3951	0.3955
Lung	18	Adenocarcinoma (Lung)	1865	23%	14	20	MX	0.5342	0.5320
Small intestine	31	Adenoma (Small intestine)	3135	.0%	0	1	IN	0.4000	0.4191
Small intestine	31	Adenocarcinoma (Small int	3165	.0%	0	2	FA	0.2765	0.2627
Large intestine	32	Mucinous carcinoma (Large	3261	.0%	0	1	IN	0.6512	0.6244
Liver	34	Hepatocellular adenoma (L	3435	28%	17	21	IN	0.2079	0.2088
Liver	34	Hemangioma (Liver)	3442	.0%	0	2	IN	0.2742	0.2665
Liver	34	Hepatocellular carcinoma	3465	20%	12	13	MX	0.4242	0.4242
Liver	34	Hemangiosarcoma (Liver)	3472	.0%	0	1	FA	0.5102	0.5081
Gallbladder	36	Adenoma (Gallbladder)	3635	2%	1	3	IN	0.5641	0.5454
Urinary bladder	41	Papilloma (Urinary	4131	2%	1	0	IN	1.0000	0.9669

		bladde							
Testis	43	Interstitial cell tumor (	4337	2%	1	1	IN	0.6483	0.6703
Testis	43	Rete testes adenoma (Test	4339	3%	2	0	IN	1.0000	0.9966
Pituitary	59	Anterior adenoma (Pituita	5935	.0%	0	2	IN	0.2742	0.2665
Thyroid	60	Follicular adenoma (Thyro	6035	.0%	0	1	IN	0.4000	0.4191
Skeletal muscle (m. trice	78	Schwannoma (Skeletal musc	7850	2%	1	0	IN	1.0000	0.9637
Harderian gland	81	Adenoma (Harderian gland)	8135	8%	5	4	IN	0.4950	0.5035
Skin	86	Hemangioma (Skin)	8642	.0%	0	1	IN	0.4000	0.4191
Skin	86	Leiomyoma (Skin)	8646	.0%	0	1	IN	0.3636	0.3884
Skin	86	Schwannoma (Skin)	8650	2%	1	0	IN	1.0000	0.9921
Skin	86	Hemangiosarcoma (Skin)	8672	2%	1	0	IN	1.0000	0.9669
Skin	86	Histiocytic sarcoma (Skin	8684	2%	1	0	IN	1.0000	0.9689
General	99	Histiocytic sarcoma (Gene	9984	3%	2	0	FA	1.0000	0.9918
General	99	Myelogenic leukemia (Gene	9988	2%	1	0	FA	1.0000	0.9873
General	99	Malignant lymphoma (Gener	9989	8%	5	4	MX	0.8114	0.8114

APPEARS THIS WAY
ON ORIGINAL

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

/s/

Roswitha Kelly
8/13/02 03:40:36 PM
BIOMETRICS

Gang Chen
8/14/02 07:32:11 AM
BIOMETRICS

APPEARS THIS WAY
ON ORIGINAL

Executive CAC
Date of Meeting: 7/16/02
Mouse/Rat Carcinogenicity Studies

Committee: Joseph Contrera, Ph.D., HFD-901, Acting Chair
Jim Farrelly Ph.D., HFD-530, Alternate Member
Abby Jacobs, Ph.D., HFD-540, Alternate Member
Barry Rosloff, Ph.D., Supervisory Pharmacologist
Lois Freed, Ph.D.HFD-120, Presenting Reviewer

Author of Draft: Lois M. Freed, Ph.D.

The following information reflects a brief summary of the Committee discussion and its recommendations. Detailed study information can be found in the individual review.

NDA #21-436
Drug Name: aripiprazole
Sponsor: Otsuka Pharmaceuticals

Mouse Carcinogenicity Studies: two 2-yr dietary carcinogenicity studies were conducted in CD-1 mice. Study 1 was conducted at doses of 0, 1, 3, and 10 mg/kg. No dose-limiting toxicities were observed. No drug-related tumors were detected in male mice. In female mice, there were significant increases in anterior pituitary adenomas and mammary gland tumors [adenocarcinoma, adenoacanthoma] at 3 and 10 mg/kg. Study 2 was conducted at doses of 0 and 30 mg/kg. Body weight was reduced in drug-treated males [10%] relative to control males. Mortality rate was significantly increased in drug-treated females. No drug-related tumor findings were detected in male mice. In female mice, there were significant increases in anterior pituitary adenoma and mammary gland tumors [adenocarcinoma, adenoacanthoma] in drug-treated females. The sponsor attributed the neoplastic findings to increases in serum prolactin [not measured in the carcinogenicity studies]; a direct drug-effect on DNA synthesis in the pituitary was also suggested as a possible mechanism underlying the increase in pituitary adenomas.

Rat Carcinogenicity Studies: two 2-yr dietary carcinogenicity studies were conducted in rats. Study 1 was conducted in Fischer 344 rats at doses of 0, 1, 3, and 10 mg/kg. No dose-limiting toxicities were observed. No drug-related tumors were detected in male rats. In female rats, there was a significant increase in mammary gland fibroadenomas at the HD. Study 2 was conducted in Sprague-Dawley rats at doses of 0, 0, 10, 20, 40, and 60 mg/kg. Dose-related decreases in body weight [compared to controls] were observed in both males and females. No drug-related tumor findings were detected in male rats. In female rats, there was a significant increase in adrenocortical tumors [carcinoma, combined adenoma and carcinoma]. No mechanism was proposed by the sponsor for the adrenocortical tumors.

Executive CAC Recommendations and Conclusions: the ExeCAC concluded that the assessment of carcinogenic potential was adequate in both mice and rats based on body wt effects in male mice, male rats, and female rats and on an increase in mortality in female mice at the highest doses tested. Aripiprazole was negative for neoplasms in male mice and rats. In female mice, pituitary adenomas and mammary gland tumors [adenocarcinoma, adenoacanthoma] at 3, 10, and 30 mg/kg were considered drug-related. In female rats, the increase in mammary gland fibroadenomas at 10 mg/kg in Study 1 and the increase in adrenocortical tumors [carcinoma, combined adenoma/carcinoma] at 60 mg/kg in Study 2 were considered drug-related.

The Committee recommended that the sponsor be asked to provide evidence for an association between

mammary gland adenoacanthomas and hyperprolactinemia.

Joseph Contrera, Ph.D.
Acting Chair, Executive CAC

cc:\
/Division File, HFD-120
/BNRosloff, HFD-120
/LMFreed, HFD-120
/SHardeman, HFD-120
/ASeifried, HFD-024

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

/s/

Joe Contrera
7/25/02 11:24:47 AM

APPEARS THIS WAY
ON ORIGINAL