CENTER FOR DRUG EVALUATION AND RESEARCH

APPROVAL PACKAGE FOR:

APPLICATION NUMBER

21-434

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: Neuropharm Drug Products (HFD-120)
Biometrics Division: Division of Biometrics I (HFD-710)

NDA NUMBER:	21-434
DRUG NAME:	Alprazolam XR (XANAX XR)
INDICATION:	Panic Disorder
SPONSOR:	Pharmacia & Upjohn Company
STATISTICAL REVIEWER:	Fanhui Kong, Ph.D. (HFD-710)
DATE OF DOCUMENT:	2/14/2002

DISTRIBUTION:

HFD-120 Anna M. Homonnay Weikel, Project Manager
Robert Levin, M.D., Clinical Reviewer
Thomas Laughren, M.D., Medical Team Leader
Russell Katz, M.D., Division Director

DFD-710 Kun Jin, Ph.D., Statistical Team Leader
George Chi, Ph.D., Division Director

HFD-700 Charles Anello, Sc.D., Deputy Director

TABLE OF CONTENTS

Statistical Review and Evaluation

1. Executive Summary

The current submission NDA 21-434 for XANAX XR (Alprazolam) consists of four phase-III studies to compare the efficacy and safety of alprazolam sustained-release with that of placebo for treating patients with panic disorder with extensive or limited phobic avoidance.

Study M/2000/0369 was a randomized, multicenter, double-blind, placebo-controlled, adjustable-dose six-week study conducted in the United States. A total of 217 patients were randomized and 199 were in the intent-to-treat population.

Study M/2000/0271 was a randomized, multicenter, double-blind, placebo-controlled, adjustable-dose six-week study conducted in the United States and Canada. A total of 212 patients were randomized and 205 were in the intent-to-treat population.

Study M/2002/0003 was a randomized, multicenter, double-blind, placebo-controlled, fixed-dose eight-week study conducted in the United States. A total of 261 patients were randomized and 252 were in the intent-to-treat population.

Study M/2002/0002 was a randomized, multicenter, double-blind, placebo-controlled, fixed-dose eight-week study conducted in the United States. A total of 231 patients were randomized and 225 were in the intent-to-treat population.

In this submission there are seven primary endpoints for Studies M/2000/0369 and M/2000/0271, there are 5 primary endpoints for Studies M/2002/0002 and M/2002/0003. Study M/2000/0369 was positive with p-values below 0.05 in all seven primary outcomes in LOCF analyses. Study M/2000/0271 was positive in 5 of the 7 primary endpoints in LOCF analyses. The other 2 studies (M/2002/0002 and M/2002/0003) were not positive in any of the seven primary outcomes.

2. Introduction

The studies in the current NDA submission were conducted from 1986 to 1991 and were submitted [redacted] when it was not required to specify a single primary endpoint. The previous submission was [redacted] due to the fact that there were only one positive study while two positive studies were required for the approval of the application.

The current submission NDA 21-434 for XANAX XR (Alprazolam) consists of four phase-III studies to compare the efficacy and safety of alprazolam sustained-release with that of placebo for treating patients with panic disorder with extensive or limited phobic avoidance.

Study M/2000/0369 was a randomized, multicenter, double-blind, placebo-controlled, adjustable-dose six-week study conducted in the United States. A total of 217 patients were randomized and 199 were in the intent-to-treat population.

Study M/2000/0271 was a randomized, multicenter, double-blind, placebo-controlled, adjustable-dose six-week study conducted in the United States and Canada. A total of 212 patients were randomized and 205 were in the intent-to-treat population.

Study M/2002/0003 was a randomized, multicenter, double-blind, placebo-controlled, fixed-dose eight-week study conducted in the United States. A total of 261 patients were randomized and 252 were in the intent-to-treat population.

Study M/2002/0002 was a randomized, multicenter, double-blind, placebo-controlled, fixed-dose eight-week study conducted in the United States. A total of 231 patients were randomized and 225 were in the intent-to-treat population.

3. Study M/2000/0369

This study was conducted from June 15, 1988 to January 23, 1990. The final protocol was signed off on April 25, 1988. The only amendment signed off on May 19, 1988 was to add a urine drug toxicology screen to laboratory evaluations and the statistical analysis plan (SAP) was not changed. This study was first submitted [redacted], but the NDA was [redacted] due to the fact that there was only one positive study and per guideline at that time required two positive studies for approval of the submission.

3.1 Study Objectives

The primary objective of this study was to compare the efficacy and safety of alprazolam sustained-release (ASR) with that of placebo for treating patients with panic disorder with extensive or limited phobic avoidance.

3.2 Study Design

This is a randomized, multicenter, double-blind, placebo-controlled, adjustable-dose study that compares ASR with placebo in the treatment of panic disorder. The study was designed as follows: (a) a screening visit; (b) a minimum one-week of drug-free run-in period; (c) after the subjects were randomized to either alprazolam or placebo, they went through a double-blind treatment period of 6-weeks in adjustable, stepped doses not to exceed 10 tablets per day; (d) and a taper period of up to five weeks until no medication was administrated. After that, subjects returned in two weeks for a post-discontinuation evaluation.

It was planned a total sample size of 180 subjects with 90 in each treatment group in the protocol. This was based on the historical data from a previous study 4412 (49% of patients on alprazolam achieved zero panic attacks vs. 29% of patients on placebo). So in order to calculate the sample size, the primary endpoint of achieving zero attacks and the above event rates were used to achieve a power of 75% with a 2-sided significance level of 0.05.

In the real study however, a total sample of 217 with 109 in alprazolam and 108 in placebo were recruited that include an ITT sample of 199.

3.3 Efficacy Measures

The primary efficacy measures include Total Panic Attacks, Clinician's Global Impressions and Overall Phobia State. In these measures Total Panic Attack is the sum of situational panic attacks and unexpected panic attacks. Three sub-measures are used for the Total Panic Attacks: Mean Change from Baseline, ≥ 50% Decrease and Achieved Zero. In the mean time, three sub-measures are used for the Clinician's Global Impression: Status of Mental Illness, Change in Condition and Therapeutic Effect.

The secondary efficacy measures include Panic Scale Variable (which itself consists of Situational Panic Attacks, Spontaneous Panic Attacks and Anticipatory Anxiety), Phobia Scale, Hamilton Anxiety Rating (HAM-A) Scale, Hamilton Depression (HAM-D) Scale, Sheehan Patient-Rated Anxiety Scale-Modified (SPRAS) Scale and Patient's Status.

3.4. Statistical Analysis Plan

Listings and tabular summaries were presented to give a general description of the patients studied. Results of statistical comparisons between treatment groups were presented to identify those differences that were conformed when subjected to probabilistic arguments. Primary and secondary measures were analyzed for the intent-to-treat (ITT) population which was defined as all enrolled subjects who took at least one dose of study medication and who had a baseline and at least one follow-up visit. The Observed Case Analysis (OC) and Last Observation Carried Forward (LOCF) imputation analysis were reported for the primary and secondary endpoint for the first six weeks and the "Last Visit" data point.

Statistical tests were two-sided with a significance level of 0.05, and a marginal significance level of $0.05 < p \leq 0.10$. Chi-square test was used to test for the homogeneity of distributions of discrete endpoint. Type III F-test was used for analysis of variance (ANOVA) for two way fixed effects model for continuous and ordinal endpoints. An additive linear model

$$\text{Delta (Response)} = \text{Mean} + \text{Treatment} + \text{Investigator} + \text{Interaction} + \text{Error}$$

was used for the endpoint of the change from baseline value (Delta), treatment effect (Treatment), the effect of individual investigator (Investigator), and the treatment investigator interaction (Interaction). If an interaction test for a given analysis variable/time point was found ($p \leq 0.1$), then separate analyses were also performed for each investigator or that variable/time point. Continuous secondary endpoints were analyzed with a similar model but with treatment effect only.

The dates were April 21, 1988 for the original protocol and May 19, 1988 for Amendment I.

3.5 Study Population

The study population consisted of male or non-pregnant female outpatients between 18 and 65 years of age. Female patients were to be surgically sterile, postmenopausal, or using adequate contraceptive measures. All patients were required to have a modified DSM-III diagnosis of panic disorder, which was further classified by their level of phobic symptoms as either panic disorder with extensive or limited phobic avoidance. Patients with such a diagnosis must had experienced at least two unexpected four-symptom attacks sometime during the history of the disorder and currently were experiencing at least one four-symptom attack per week for three consecutive weeks.

There were 3 investigative sites that randomized patients into the study. Of the 109 subjects randomized into ASR group, 104 (95%) were included in the ITT population.
Of the 108 subjects randomized into placebo group, 95 (88%) were included in the ITT population.
Seventy-seven (74% of ITT) ASR-treated patients completed 6 weeks of treatment, while 51 (54% of ITT) placebo patients did so. In addition, 10 ASR- and 7 PBO-treated patients completed more than 6 weeks of treatment.

For patient disposition, alprazolam group has protocol completed rate (84%) as compared to placebo (61%). The primary reasons for not completing the treatment were "Lost to Follow-Up" and "Medical events" for ASR group, "Lack of efficacy" and "Patient decision to withdraw" for the placebo group.

Table 3.5.1 Number of Patients Completed 6 weeks

	ASR	Placebo
Screen	109	108
Week 6	77	51
6 Weeks Treatment	87	58
Discontinuation	81	71
Post Discontinuation	67	55

Table 3.5.2 Reasons for Not Complete Treatment – ITT Population

Reason	ASR	Placebo
Lack of efficacy	1	19
Loss to follow-up	5	3
Medical events	4	1
Intercurrent illness	2	1
Protocol violation	2	1
Patient decisions to withdraw	3	9
TOTAL	17	34

For the 199 ITT patients (104 in ASR, 95 in Placebo), the ASR and placebo treatment groups were similar at baseline in age, sex, race, weight and height. The average age of both medication groups was 35 years. Fifty-nine percent of the ASR- and 62% of placebo treated patients were female, and

95% and 97% respectively, were white. Patients in the ASR had a mean weight of 162 lbs, while those in the placebo group had a mean weight of 159 lbs. Patients in both groups had a mean height of 67 inches. Patients in the ASR group smoked a larger number of cigarettes per day than did those in the placebo group (10 vs. 5) and consumed a large amount of beer per day (0.4 vs. 0.15 12-oz cans). There were also significant differences in the number of patients reporting a history of mental illness (9 ASR, 1 placebo) and a history of allergies (24 ASR and 34 placebo).

Table 3.5.3 Demographic and Baseline Characteristics – ITT Population

Variable	ALP XR	Placebo
Age (yrs)	n=104	n=95
Mean	35.0	34.8
Sex, No. (%) pts	n=104	n=95
Male	43 (41)	36 (38)
Female	61 (59)	59 (62)
Race, No. (%) pts	n=104	n=95
White	99 (95)	92 (97)
Black	3 (3)	2 (2)
Hispanic	1 (1)	1 (1)
Oriental	0 (0)	0 (0)
American Indian	0 (0)	0 (0)
Other	1 (1)	0 (0)
Weight (lbs)	n=99	n=94
Mean	162.3	158.8
History of mental illness, No. (%) pts	n=104	n=95
Yes	9 (9)	1 (1)

The ASR dose peaked at Week 5 when the mean number of tablets taken was 4.67 per day, while the highest mean dose taken by placebo-treated patients was 6.94 tablets at Week 6. Seventy-seven (74% of ITT population) ASR-treated patients took study medication through Week 6, while 51 (54% of ITT population) patients in the placebo treatment group did so.

Other protocol violations include: (1) Although the intended sample size of the protocol was 180 patients, 217 were enrolled. (2) Medications that were prohibited in the protocol were used. Some patients were not tapered to zero medication and some patients who had terminated the study were immediately placed on another medication without first being referred to a physician not involved in the study. (3) The sequencing of the discontinuation and post-discontinuation phases was not done as planned in the protocol.

Table 3.5.4 Baseline Efficacy Score - Baseline Severity of Illness -- ITT Population

Primary Efficacy Parameters at Baseline	ASR (N=104)	Placebo (N=95)	P-value
Total Panic Attack Mean (SD)	6.27 (6.14) N=104	5.99 (5.57) N=95	0.82
Overall Phobia State Mean (SD)	6.89 (2.29) N=102	6.95 (1.99) N=93	0.81
Clinician's Global Impression: Status of Mental Illness Mean (SD)	4.54 (0.83) N=103	4.41(0.94) N=95	0.45

3.6 Sponsor's Efficacy Results

3.6.1. Primary Efficacy Results

The results of statistical comparisons of primary endpoints between ASR and placebo groups of the ITT population at the sixth week of treatment are given in Table 3.6.1. In the primary efficacy analysis, given a high percentage of dropouts, the sponsor showed statistically significant differences between treatment groups in favor of ASR with respect to all seven primary efficacy parameters at the sixth week endpoint for LOCF analysis. In OC analysis, significance was found in five of seven primary efficacy measures but not found in the reduction of the Total Panic Attacks and Overall Phobia State. As depicted in Table 3.6.1, significantly more patients in the ASR group achieved 50% reduction in the total number of attacks and achieved zero attacks after six weeks of treatment compared to the placebo group. At the same time, ASR was shown to give significantly more reductions in the mean scores from baseline in the status of mental illness and change in condition of Clinician's Global Impressions compared to placebo group in both OC and LOCF analyses. On the other hand, ASR significantly increased the therapeutic effect compared to the placebo in both OC and LOCF analyses.

Table 3.6.1 Primary Efficacy Significance at Week 6 --- ITT Population

Parameter	ASR		Placebo		P-Value	
	OC	LOCF	OC	LOCF	OC	LOCF
Total Panic Attacks						
Mean Change from Baseline (SD)	-5.13(7.09) N=78	-4.65(6.85) N=104	-3.78(6.63) N=51	-1.95(6.62) N=95	0.416	0.026
≥ 50 % Decrease (% pts)	92.3 N=78	84.3 N=102	78.4 N=51	62.1 N=95	0.023	<0.001
Achieved zero (% pts)	79.5 N=78	70.2 N=104	58.8 N=51	45.3 N=95	0.011	<0.001
Overall Phobia State						
Mean Change from Baseline (SD)	-3.97(2.88) N=72	-3.55(2.91) N=102	-3.13(2.91) N=48	-2.18(2.79) N=93	0.160	0.0028
Clinician's Global Impression						
Status of Mental Illness Mean Change from Baseline (SD)	-2.34(1.21) N=77	-2.03(1.35) N=103	-1.53(1.39) N=51	-1.00(1.38) N=95	0.0045	0.0001
Change in Condition Mean (SD)	1.81(0.97) N=77	2.04(1.24) N=103	2.45(1.36) N=51	3.05(1.59) N=95	0.0083	0.0001
Therapeutic Effect Mean (SD)	4.35(0.84) N=77	4.17(1.04) N=103	3.69(1.24) N=51	3.17(1.41) N=95	0.0022	0.0001

OC = Observed Case; LOCF = Last Observation Carried Forward.

3.6.2. Secondary Efficacy Evaluation:

The results of statistical comparisons of secondary efficacy endpoints between ASR and placebo groups of the ITT population at the sixth week of treatment are summarized in Table 3.6.2. ASR-treated subjects demonstrated significant improvement on the mean percent of time spent worrying (p=0.019). These subjects also demonstrated significantly greater improvement in the fear score and HAM-A score compared with placebo-treated subjects.

Between-group comparisons of outcomes on the mean numbers of situational and spontaneous panic attacks, Hamilton Depression Scale Score and Sheehan Patient-Rated Anxiety Scale did not shown a significant improvement over the placebo-treated subjects. There was no significant difference between groups in the avoidance scores of Phobia and the percentage of personal events which were bothered.

Table 3.6.2 Secondary Efficacy Significance at Week 6 – ITT Population for Observed Case

Parameters	ASR	Placebo	P-value
Panic Attack Scale Variables			
Situational Panic Attack Mean Number per Week (SD)	-2.04(5.45) N=78	-2.27(3.36) N=51	0.78
Spontaneous Panic Attack Mean Number per Week (SD)	-3.09(4.47) N=78	-1.51(5.33) N=51	0.072
Anticipatory Anxiety Mean % Time Spent Worry (SD)	-24.5(29.77) N=78	-12.6(24.37) N=51	0.019
Phobia Scale Variables			
Total Main Phobia Fear Score	-14.0(9.42) N=72	-10.18(10.15) N=50	0.035
Avoidance Score	-3.93(3.60) N=72	-3.35(3.56) N=49	0.38
Hamilton Anxiety Scale Score	-11.77(7.71) N=77	-8.10(8.64) N=51	0.013
Hamilton Depression Scale Score	-5.83(6.41) N=64	-4.85(5.93) N=47	0.41
Sheehan Patient-Rated Anxiety Scale			
Part I (Problem/complaints)	-40.61(28.46) N=77	-35.02(28.96) N=51	0.212
Part II (How Patient Felt)	-9.49(10.40) N=77	-8.06(12.20) N=51	0.477
Patient's Status Variables			
How Much Time Illness Prevent Activity	-1.53(1.26) N=77	-0.98(1.12) N=51	0.013
How Much Personal Events Have Bothered pt	-0.51(1.59) N=77	-0.94(1.38) N=51	0.12

3.7. Reviewer's Analysis

1. The reviewer duplicated the sponsor's analyses according to the protocol and found the sponsor's results to be accurate. The violation of the protocol in collecting data was moderate and should not alter the study conclusion. However, the high percentage and the imbalance of withdrawing between the treatment and placebo groups raise concerns on the reliability of the testing results using LOCF data. Such single imputation method usually reduces the standard deviation estimations. On the other hand, imbalance of dropping out between treatment and placebo groups could lead to the bias in the testing of treatment effect.

   Compared to LOCF analysis, the significance level of OC analysis was dramatically reduced. For primary endpoints Total Panic Attacks and Overall Phobia State, the statistical tests of OC population are not significant. To find out if such an insignificance is caused by the difference between the non-dropouts, who finished six weeks of follow up, and the dropouts, who dropped out before the sixth week, we plotted the weekly means of the two treatment groups

for both dropouts and non-dropouts. The patterns of change are similar for dropouts and non-dropouts. This might suggest that the non-significance of the OC analysis is not caused by the imbalance between the dropout and non-dropouts, instead, rather by the reduction of the number of subjects.

Table 3.7.1 Pattern of Change from Baseline for the Total Panic Attacks for the first Five Weeks - ITT Population

Total Panic Attacks Change from baseline		Week				
		Week 1	Week 2	Week 3	Week 4	Week 5
Completers	ASR	-5.74	-5.48	-5.38	-5.57	-5.07
	Placebo	-3.87	-3.49	-3.87	-3.89	-3.87
Droppers	ASR	-3.0	-5.06	-2.83	-3.40	-4.33
	Placebo	0.53	1.06	-0.42	-1.40	-2.46

2. Of the 217 subjects who were randomized into either treatment or placebo groups, only 199 were included in the ITT population, 18 of them were excluded. The reviewer checked the data set provided by the sponsor and found that 17 of these 18 subjects did not have measurements of the primary endpoint. This seems to be quite high at such an early stage of the study.

3. Normality assumption of continuous and ordinal endpoints failed. So the validity of the test results for treatment effects using ANOVA methods is a concern. Nonparametric methods such as Wilcoxon test were applied by the reviewer to verify the results on the primary endpoints at the sixth week derived from the data provided by the sponsor. The results obtained give the similar conclusion. Results were presented in Table 3.7.3.

4. The Division of Scientific Investigations (DSI) inspected the three clinical sites and found that the source documents for 28 of 37 subjects at Dr. Rosenthal's site were destroyed. They recommended to exclude all data generated at this site and to reanalyze the efficacy data in support of this NDA. So the data for the 37 subjects at the site of Dr. Rosenthal are excluded and the rest of the data are reanalyzed for the primary endpoints at the sixth week by the reviewer. The results are depicted in Table 3.7.2. These results do not change the conclusion derived by the sponsor using the data from all three sites. The non-parametric are depicted in Table 3.7.3.

Table 3.7.2 Primary Efficacy Significance at Week 6 for Two Centers Except that of Rosenthal, — ITT Population.

Parameter	ASR		Placebo		P-Value	
	OC	LOCF	OC	LOCF	OC	LOCF
Total Panic Attacks						
Mean Change from Baseline (SD)	-5.45(6.0) N=62	-4.89(6.1) N=85	-3.87(6.64) N=39	-1.76(6.9) N=76	0.28	0.0026
≥ 50 % Decrease (% pts)	95.2 N=62	85.9 N=85	76.9 N=39	61.8 N=76	0.006	0.0005
Achieved zero (% pts)	82.3 N=62	71.8 N=85	53.9 N=39	42.1 N=76	0.002	0.0001
Overall Phobia State						
Mean Change from Baseline (SD)	-4.0(2.57) N=60	-3.58(2.6) N=83	-3.0(2.78) N=36	-2.14(2.67) N=74	0.066	0.0005
Clinician's Global Impression						
Status of Mental Illness Mean Change from Baseline (SD)	-2.34(1.20) N=62	-2.05(1.33) N=84	-1.28(1.28) N=39	-0.83(1.23) N=78	<0.0001	<0.0001
Change in Condition Mean (SD)	1.85(1.04) N=62	2.07(1.25) N=85	2.67(1.42) N=39	3.21(1.55) N=78	0.0022	<0.0001
Therapeutic Effect Mean (SD)	4.31(0.88) N=62	4.14(1.05) N=85	3.49(1.27) N=39	3.04(1.36) N=78	0.0005	<0.0001

OC = Observed Case; LOCF = Last Observation Carried Forward.

Table 3.7.3 Wilcoxon Test of Primary Endpoints at Week 6 for Two Centers Except That of Rosenthal — ITT Population.

Parameter	P-Value	
	OC	LOCF
Total Panic Attacks		
Mean Change from baseline	0.33	0.016
Overall Phobia State		
Mean Change from baseline	0.12	0.0012
Clinician's Global Impression		
Status of Mental Illness Mean Change from baseline	<0.0001	<0.0001
Change in Condition	0.0028	<0.0001
Therapeutic Effect	0.015	<0.0001

OC = Observed Case; LOCF = Last Observation Carried Forward.

5. Subgroup analyses were performed by the reviewer to see if the treatment effects were concentrated in some particular groups rather than the whole population. The results are briefly described in the following. The population is separated into subgroups according to sex, race, age, and clinical sites. The frequencies for female versus male subjects are around 60% versus 40% who have data for statistical analysis for different primary endpoints. The corresponding frequencies for white versus nonwhite subjects are around 95% versus 5% and the frequencies for subjects below 45 years of age versus those at least 45 years of age are around 90% versus 10% in the population. In addition, the site of Dr. Rickel has 143 subjects and that of Dr.

Patterson has 71 subjects but the corresponding number of subjects available for analysis at the sixth week are around 100 versus 60 for LOCF analyses and 60 versus 50 for OC analyses. Due to the low frequencies in nonwhite group and the group of at least 45 years of age, the subgroup analyses are performed only for sex and sites.

Subgroup analyses by the reviewer for male and female subgroups indicate that the treatment effects of the both subgroups have the same direction yet the female group has a higher significance level, especially for the OC analyses. The results for LOCF analysis were depicted in Table 3.7.4. Similarly, the treatment effects have the same direction for both sites. Although the site of Dr. Patterson has less number of subjects, it has a higher percentage of subjects available for analyses at Week 6 and the treatment effects tend to be higher for the most of the primary endpoints.

Table 3.7.4 Subgroup Analysis for Primary Endpoints for Sex at Week 6 Two Centers Except That of Rosenthal — ITT LOCF Population.

Parameter	ASR		Placebo		P-Value*	
	Male	Female	Male	Female	Male	Female
Total Panic Attacks						
Mean Change from Baseline (SD)	-4.78 N= 36	-4.98 N= 49	-3.04 N= 27	-1.06 N= 49	0.21	0.045
≥ 50 % Decrease (% pts)	88.9 N= 36	83.7 N= 49	74.1 N= 27	55.1 N= 49	0.125	0.002
Achieved zero (% pts)	75.0 N= 36	69.4 N= 49	52 N= 27	36.7 N= 49	0.057	0.0012
Overall Phobia State						
Mean Change from Baseline (SD)	-3.56 N= 34	-3.59 N= 49	-2.46 N= 28	-1.93 N= 46	0.12	0.0035
Clinician's Global Impression						
Status of Mental Illness Mean Change from Baseline (SD)	-2.28 N=36	-1.88 N= 48	-1.03 N= 29	-0.71 N= 49	0.0001	0.0001
Change in Condition Mean (SD)	1.83 N=36	2.24 N= 49	2.97 N= 29	3.35 N= 49	0.0007	0.0007
Therapeutic Effect Mean (SD)	4.3 N=36	4.0 N=49	3.21 N=29	2.94 N=49	0.0005	0.0002

*Wilcoxon test is applied to derive the p-value for continuous outcome variables.

4. Study M/2000/0271

This study is a double-blind, randomized, parallel group design, multicenter Phase III study conducted at three study centers in United States and Canada from July 1986 to March 1989 to compare the efficacy and the safety of ASR with that of the marketed formulation of XANAX Tablets (ACT) and placebo (PBO) in patients with panic disorder with limited or extensive phobic avoidance.

4.1 Basic Study Design and Analysis Procedure

A six-week treatment phase was followed by a tapered discontinuation phase and a four-week post discontinuation phase. Tablets were taken four times daily in a double-dummy design that resulted in a single daily dose of ASR or four daily doses of ACT and PBO; alprazolam doses ranged from 1 to 10 mg per day. Subjects were evaluated at screen, baseline and weekly during the treatment, discontinuation and post-discontinuation phases of the study.

The primary objective of this study was to compare the efficacy and safety of among ASR, ACT and placebo for treating patients with panic disorder with extensive or limited phobic avoidance.

The similar criteria for subjects entering Study-0369 were used for this study. The primary efficacy measures are the same as those in Study-0369. Most of the secondary efficacy measures are the same as those in Study-0369.

A total of 75 patients at each of three sites were to be randomized to one of three treatment groups (25 patients per group per site). A protocol addendum increases the patient number at Site 2 to 90 patients (30 per group).

The same statistical analysis plan as in Study-0369 was adopted for the primary and secondary endpoint analyses.

4.2 Study Population

All 70 of the subjects randomized to ACT group, 68 of the 71 subjects randomized to ASR group, and 67 of the 71 subjects randomized to placebo group were included in the ITT population. Fifty-seven (84% of ITT) ASR-treated patients, 63 (90% of ITT) ACT-treated patients, and 45 (67% of ITT) placebo-treated patients had Week 6 assessments.

The ASR, ACT and PBO groups are comparable at baseline as there were no significant or marginal between-group differences in demographic variables: age, sex, race and body weight. Overall significant differences among groups for mean daily dose were seen throughout the treatment phases. ASR and ACT groups have a significantly lower mean dose than the PBO group. ACT and ASR groups are comparable.

Protocol violation is very moderate. The major statistics related violation is that the protocol specified to enroll 225 patients while the actual enrollment was 212 patients.

4.3 Sponsor's Results

4.3.1 Primary Efficacy Evaluation:

Baseline primary measurements are compatible among the three treatment groups of the ITT population. The statistical comparisons for primary endpoints among the three groups at Week 6 of treatment for LOCF and OC populations are given in Table 4.3.1 and Table 4.3.2. The results are briefly summarized in the following.

No overall significant differences were found among the treatment groups regarding the decreases of the total panic attacks in either LOCF or OC analysis. Both ASR and ACT were significantly superior to placebo in the reduction of at least 50% panic attacks in LOCF analysis, but not in OC analysis. ACT was also found to be significantly superior to placebo in reaching 0 panic attacks in both analyses, but not ASR.

Both ASR and ACT groups had greater mean reductions in overall phobia score than placebo at Week 6 in LOCF analysis but no overall significance was seen in OC analysis. In LOCF analysis, both ASR and ACT groups had greater improvement in all three CGI scores than placebo at Week 6. But in OC analysis, only ASR over performed the placebo on the mean improvement score of CGI, not any other.

4.3.2 Secondary Efficacy Evaluation:

No overall significant or marginal difference was seen among treatment group means at baseline for secondary endpoints.

In Study 0271, ASR and ACT groups had significantly greater improvement over placebo for 12 and 13 efficacy variables, respectively, in the LOCF analysis, but improved placebo only for 4 efficacy variables in the OC analysis. In one case (Other Phobia, fear score), the placebo group had a significantly larger mean decrease than the active treatment groups. One efficacy variable (SPRAS, Part 1 [problems/complaints]) showed that the ACT group had a significantly larger mean decrease than the ASR group (LOCF analysis). Overall, the results of the secondary efficacy variables were positive and supported those of the primary efficacy variables.

Table 4.3.1 Primary Efficacy Significance at Week 6 — ITT Population for LOCF Data

Parameter	ASR	ACT	PBO	P-value
Total Panic Attacks				
Mean change from baseline	-7.28 (20.23) n=67	-5.47 (6.41) n=70	-2.91 (8.93) n=66	Overall: 0.179
% With 50% decrease	81% n=67	90% n=70	63% n=65	Overall: 0.001 ASR vs pbo: 0.025 ACT vs pbo: <0.001 ASR vs ACT: 0.119
% Achieved zero	53% n=68	80% n=70	51% n=67	Overall: <0.001 ASR vs pbo: 0.799 ACT vs pbo: <0.001 ASR vs ACT: 0.001
Overall Phobia State				
Mean change from baseline	-3.37 (2.84) n=67	-3.41 (2.72) n=69	-1.85 (2.57) n=67	Overall: <0.001 ASR vs pbo: 0.002 ACT vs pbo: <0.001 ASR vs ACT: 0.762
Clinician's Global Impressions				
Severity, mean change from Baseline	-1.76 (1.20) n=68	-2.13 (1.52) n=70	-1.12 (1.36) n=65	Overall: <0.001 ASR vs pbo: 0.009 ACT vs pbo: <0.001 ASR vs ACT: 0.076
Improvement, mean score	2.15 (1.08) n=68	1.84 (0.81) n=70	2.92 (1.49) n=65	Overall: <0.001 ASR vs pbo: <0.001 ACT vs pbo: <0.001 ASR vs ACT: 0.127
Therapeutic Effect, mean score	4.13 (1.06) n=68	4.34 (0.80) n=70	3.45 (1.36) n=65	Overall: <0.001 ASR vs pbo: <0.001 ACT vs pbo: <0.001 ASR vs ACT: 0.237

Table 4.3.2 Primary Efficacy Significance at Week 6 — ITT Population, for OC Data

Parameter	ASR	ACT	PBO	P-value
Total Panic Attacks				
Mean change from baseline	-8.18 (21.78) n=57	-5.70 (6.48) n=63	-4.43 (8.62) n=44	Overall: 0.528
% With 50% decrease	86% n=57	92% n=63	77% n=43	Overall: 0.084
% Achieved zero	61% n=57	84% n=63	64% n=45	Overall: 0.013 ASR vs pbo: 0.752 ACT vs pbo: 0.018 ASR vs ACT: 0.005
Overall Phobia State				
Mean change from baseline	-3.71 (2.56) n=56	-3.33 (2.59) n=61	-2.24 (2.43) n=45	Overall: 0.078
Clinician's Global Impressions				
Severity, mean change from Baseline	-1.93 (1.12) n=57	-2.30 (1.44) n=63	-1.55 (1.27) n=44	Overall: 0.061
Improvement, mean score	1.96 (0.87) n=57	1.75 (0.72) n=63	2.44 (1.20) n=45	Overall: 0.037 ASR vs pbo: 0.01 ACT vs pbo: 0.145 ASR vs ACT: 0.219
Therapeutic Effect, mean score	4.28 (0.84) n=57	4.43 (0.73) n=63	3.87 (1.14) n=45	Med. by Inv. Int.

4.4 Reviewer's Evaluation

The reviewer's computations indicate that the results that the sponsor reported are generally reliable.

1. The reviewer duplicated the sponsor's analyses according to the protocol. We found the sponsor's results to be accurate. The violation of the protocol in collecting data was moderate and should not alter the study conclusion. However, the high percentage and the imbalance of withdrawing between two groups raise concerns on the reliability of the test results using LOCF data. Such imbalance of dropping out between treatment and placebo groups could lead to the bias in the testing of treatment effect.

   Compared to LOCF analysis, the significance level of OC analysis was dramatically reduced. For primary endpoints Overall Phobia State and Severity of CGI, the statistical tests of OC population become non-significant. Similar to study 0369, we plotted the weekly means of the two treatment groups for both dropouts and non-dropouts for the mean change from baseline of Overall Phobia State. The patterns of change for non-dropouts indicate that both AST and ACT improve the placebo. But for the dropouts the ASR does not improve the placebo. Since the dropouts are only less than 10% of the total ITT subjects, this suggest that the non-significance

of the OC analysis is not caused by the imbalance between the dropout and non-dropouts, instead, rather by the reduction of the number of subjects.

2. Normality assumption of continuous and ordinal endpoints failed. So nonparametric methods such as Wilcoxon test were applied by the reviewer to verify the results on the primary endpoints at the sixth week derived from the data provided by the sponsor. The results obtained give the similar conclusion. Results were presented in Table 4.4.1.

Table 4.4.1 Primary Efficacy Significance of Wilcoxon Test at Week 6 — ITT Population.

Parameter	P-Value	
	OC	LOCF
Total Panic Attacks		
Mean Change from baseline	0.23	0.044
Overall Phobia State		
Mean Change from baseline	0.02	0.0014
Clinician's Global Impression		
Status of Mental Illness Mean Change from baseline	0.015	0.0002
Change in Condition	0.0039	<0.0001
Therapeutic Effect	0.03	0.0003

5. Study M/2002/0003

This is a fixed-dose, double-blind, randomized, placebo-controlled, parallel-group, multicenter Phase III clinical study. The study was conducted at 16 study centers in United States from May 1990 to October 1991 to evaluate the efficacy and safety of XANAX XR (ASR) tablets in the treatment of panic disorder using once daily dosing.

5.1 Basic Study Design and Analysis Procedure

Patients were randomized into one of the following three treatment groups: 6.0 mg daily fixed-dose XANAX XR (XXR 6 mg); 4.0 mg daily fixed-dose XANAX XR (XXR 4 mg); or placebo (PBO). The duration of the study consists of a two weeks of lead-in; 8 weeks treatment phase followed by 4 weeks taper phase and 2 weeks post-discontinuation phase.

Tablets were taken once daily. Subjects were evaluated at screen, baseline and Weeks 1 through 4, 6 and 8, weekly during taper and at two weeks after drug discontinuation.

The primary objective of this study was to compare the sustained-release formulation of XANAX XR tablets at fixed-doses of 6.0 mg daily, 4.0 mg daily, and placebo in the treatment of panic disorder.

The primary efficacy endpoints consist of major panic attacks (total number of attacks and proportion of patients reaching zero attacks), clinical global impressions and overall phobia state. The secondary efficacy points include panic state variables, phobia scale, disability and quality of life scale. The same criteria for subjects entering Study-0369 and Study-0271 were used for this study.

An original sample size of 210 patients (70 each group) was chosen for the study. The targeted enrollment was later increased to 336 patients across 16 sites. Altogether, 261 patients were enrolled (83 patients into PBO, 89 patients intro 4-mg/day XXR, and 89 patients into 6-mg/day XXR) across 15 sites in the United States.

Similar statistical analysis plan as that in Study-0369 and Study-0271 was adopted in the protocol for the primary and secondary endpoint analyses. However, due to the presence of extreme outliers, additional analysis was performed on the ranks of the total panic attack data on the reduced model in the actual analysis

Response = Mean + Treatment + Error.

5.2 Study Population

Eighty-five of the 89 patients randomized to the 4-mg XXR group, 88 of the 89 patients randomized to the 6-mg XXR group, and 79 of the 83 patients randomized to placebo group were included in the ITT population. Sixty-five (76% of ITT) XXR 4-mg treated patients, 56 (64% of ITT) XXR 6-mg treated patients, and 60 (76% of ITT) placebo-treated patients had Week 8 assessment.

Two protocol violations are related with statistical analyses. The first is that the protocol specified to a total enrollment of 336 patients (112 patients per treatment arm) across 16 sites, while the actual enrollment was 261 patients across 15 sites as mentioned above. The second is the analysis of variance model on the rank of the total panic attacks mentioned above.

5.3 Sponsor's Results

Three treatment groups are comparable at baseline. Overall marginal differences were found for demographic variables age and race but no differences were found for sex make up and body weight.

5.3.1 Primary Efficacy Evaluation

The results of statistical comparisons for primary endpoints among XXR 4-mg, XXR 6-mg and placebo of the ITT population at Week 8 of treatment are given in Table 5.3.1. The primary efficacy analysis results are briefly given as follows:

The active treatment groups did not significantly reduce the mean number of attacks over the placebo group at the eighth week, nor were they significantly superior in reducing the number of attacks to zero. No significantly larger mean decreases in phobia score were observed for both active treatment groups over placebo. The active treatment groups had neither significantly lower mean improvement score (greater improvement) nor significantly smaller means for severity of illness (greater improvement) than placebo group for the eighth week.

Table 5.3.1 Primary Efficacy Significance: Weeks 8 LOCF Data

Variable	ALP XR 4 mg	ALP XR 6mg	Placebo	P-value: Overall
Total Panic Attacks				
Mean change from baseline	-3.44 (8.54) n=86	-3.51 (6.94) n=87	-2.27 (8.15) n=79	0.507
% Achieved zero	50% n=86	40% n=87	46% n=79	0.434
Overall Phobia State				
Mean change from baseline	-14.31 (25.48) n=86	-10.50 (23.76) n=86	-8.28 (24.11) n=79	0.178
Clinical Global Impressions				
Severity, mean change from baseline	-1.30 (1.43) n=86	-1.22 (1.32) n=87	-0.94 (1.11) n=79	0.064
Improvement, mean	2.57 (1.44) n=86	2.56 (1.48) n=87	2.76 (1.19) n=79	0.239

5.3.2 Secondary Efficacy Evaluation

No significant improvements of the active treatment groups over PBO were observed on either unexpected major or limited attacks at the eighth week. No significant group differences in change from baseline for agoraphobia, illness and social phobia were observed at the eighth week. No significant improvements on Sheehan disability scale or quality of life scores were seen for the active treatment groups over placebo through treatment weeks.

5.4 Reviewer's Evaluation

The reviewer's computations indicate that the results that the sponsor reported are accurate.

1. The reviewer duplicated the sponsor's results using the data set provided by the sponsor according to the analytic procedure described in the protocol. The sponsor's results are accurate. The violation of the protocol in the procedure of collecting data was moderate and should not alter the study conclusion.

2. Normality assumption of continuous and ordinal endpoints failed. Nonparametric methods were applied by the reviewer to verify the results on the primary endpoints at the sixth week derived from the data provided by the sponsor. The results obtained give the similar conclusion.

6. Study M/2002/0002

This is a fixed-dose, double-blind, randomized, placebo-controlled, parallel-group, multicenter Phase III clinical study. The study was conducted at 15 study centers in United States from June 1990 to October 1991 to evaluate the efficacy and safety of fixed 4-mg, 6-mg daily XANAX XR (ASR) doses with placebo in the treatment of panic disorder using twice daily dosing.

6.1 Basic Study Design and Analysis Procedure

Patients were randomized into one of the following three treatment groups: 6.0 mg daily fixed-dose XANAX XR (XXR 6 mg); 4.0 mg daily fixed-dose XANAX XR (XXR 4 mg); or placebo (PBO). The duration of the study consists of a one to two weeks of lead-in; 8 weeks treatment phase followed by 4 weeks taper phase and 2 weeks post-discontinuation phase.

Tablets were taken twice daily. Subjects were evaluated at screen, baseline and Weeks 1 through 4, 6 and 8, weekly during taper and at two weeks after drug discontinuation.

The primary objective of this study was to compare the sustained-release formulation of XANAX XR tablets at fixed-doses of 6.0 mg daily, 4.0 mg daily, and placebo in the treatment of panic disorder.

The primary efficacy endpoints consist of major panic attacks (total number of attacks and proportion of patients with zero attacks), clinical global impressions and overall phobia state. The secondary efficacy points include panic state variables, phobia scale, disability and quality of life scale. The same criteria for subjects entering Study-0369 and Study-0271 were used for this study.

A sample size of 231 subjects was enrolled and 225 patients remained in the ITT population (78 patients in 4-mg/day XXR, and 74 patients in 6-mg/day XXR, 73 patients in PBO).

Similar statistical analysis plan as that in Study-0369 and Study-0271 was adopted in the protocol for the primary and secondary endpoint analyses. However, due to the presence of extreme outliers, additional analysis was performed on the ranks of the total panic attack data on the reduced model in the actual analysis (also see Study-0003)

Response = Mean + Treatment + Error.

6.2 Study Population

Three treatment groups were similar at baseline regarding demographic variables.
Seventy-eight of the 79 patients randomized to the 4-mg XXR group, 74 of the 76 patients randomized to the 6-mg XXR group, and 73 of the 76 patients randomized to placebo group were included in the ITT population. Fifty-five (71% of ITT) XXR 4-mg treated patients, 47 (64% of ITT) XXR 6-mg treated patients, and 54 (74% of ITT) placebo-treated patients had Week 8 assessments.

Two protocol violations are related with statistical analyses. The protocol specified a total enrollment of 210 patients (70 per group) across 16 sites, while the actual enrollment was 231

patients across 15 sites as mentioned above. The second is the analysis of variance model on the rank of the total of panic attacks as mentioned above.

6.3 Sponsor's Results

6.3.1 Primary Efficacy Evaluation

The results of statistical comparisons for primary endpoints among XXR 4-mg, XXR 6-mg and placebo of the ITT population at Week 8 of treatment are given in Table 6.3.1. The primary efficacy analysis results are briefly given as follows:

No significantly larger decrease in mean number of attacks was seen for the active treatments than placebo at Week 8 in either OC or LOCF analysis, nor were the active treatments superior to placebo in reducing the panic attacks to zero at the eighth week. No significantly different mean decreases were observed among three treatment groups regarding the mean phobia scores. The active treatment groups had neither significantly lower mean improvement score (greater improvement) nor significantly smaller means for severity of illness (greater improvement) than placebo group for the eighth week.

Table 6.3.1 Primary Efficacy Significance: Weeks 8 LOCF Data

Variable	ALP XR 4 mg	ALP XR 6mg	Placebo	P-value: Overall
Total Panic Attacks				
mean change from baseline	-7.05 (11.47) n=78	-4.08 (6.00) n=74	-4.84 (8.70) n=73	0.107
% Achieved zero	54% n=78	47% n=74	48% n=73	0.671
Overall Phobia State				
mean change from baseline	-14.75 (26.4) n=77	-17.08 (28.73) n=72	-15.79 (22.64) n=73	0.990
Clinical Global Impressions				
Severity, mean change from	-1.64 (1.41) n=78	-1.53 (1.32) n=74	-1.36 (1.31) n=73	0.299
Improvement, mean	2.21 (1.25) n=78	2.16 (1.15) n=74	2.49 (1.34) n=73	0.139

6.3.2 Secondary Efficacy Evaluation

In this study, no differences were found in the LOCF analyses. In the OC analyses at Week 8 however, the ASR 4-mg/day group was found to have significantly larger mean decrease from baseline in work impairment (DISS) (-22.14) than in the ASR 6-mg/day group (-7.28) and placebo group (-8.53).

6.4 Reviewer's Evaluation

The reviewer's computations indicate that the results that the sponsor reported are accurate.

1. The reviewer duplicated the sponsor's results using the data set provided by the sponsor according to the analytic procedure described in the protocol. The sponsor's results are accurate. The violation of the protocol in the procedure of collecting data was moderate and should not alter the study conclusion.

2. Normality assumption of continuous and ordinal endpoints failed. Nonparametric methods were applied by the reviewer to verify the results on the primary endpoints at the sixth week derived from the data provided by the sponsor. The results obtained give the similar conclusion.

7. Conclusion

The sponsor conducted four Phase III, placebo controlled clinical trials for the efficacy study of the alprazolam XR for treating patients with panic disorder with extensive or limited phobic avoidance. In LOCF analysis, Study M/2000/0369 is positive and supports the conclusion that alprazolam XR is more effective than placebo in reducing the number of panic attacks and improving the overall clinical conditions of the patients. This study is positive in 5 of 7 primary endpoints in OC analysis. In LOCF analysis, Study M/2000/0271 is also positive in 5 of 7 primary endpoints and provides supportive evidence for alprazolam XR to be statistically more effective than placebo in the treatment of patients with panic disorder. This study is positive in 2 of 7 primary endpoints in OC analysis. The other 2 studies (M/2002/0002 and M/2002/0003) are not positive and do not provide supportive evidence.

In Study M/2000/0369, of the 217 subjects who were randomized into either treatment or placebo group, only 199 were included in the ITT population, 18 of them were excluded. The reviewer checked the data set provided by the sponsor and found that 17 of these 18 subjects did not have measurements of the primary endpoint, 12 in the placebo group and 5 in the treatment group. The agency has strong concern on the high drop out rate at such an early stage of the study. So we requests the sponsor to clarify this issue by providing detailed information of these subjects.

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

/s/

Fanhui Kong
10/4/02 01:22:54 PM
BIOMETRICS

Kun Jin
10/4/02 02:19:29 PM
BIOMETRICS

George Chi
10/9/02 12:15:58 PM
BIOMETRICS

APPEARS THIS WAY
ON ORIGINAL