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:	20-151/S-024
DRUG NAME:	Effexor (venlafaxine HCI) Tablets
INDICATION:	Pediatric GAD
SPONSOR:	Wyeth-Ayerst Laboratories
STATISTICAL REVIEWER:	Fanhui Kong, Ph.D. (HFD-710)
DATE OF DOCUMENT:	September 25, 2002

DISTRIBUTION:

HFD-120	Paul David, Project Manager
	Glenn B. Mannheim, M.D., Clinical Reviewer
	Paul Andreason, 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

This submission of efficacy study consists of two 8-week, Phase III, randomized, double-blind, parallel group multi-center, placebo-controlled, flexible dose studies that evaluates the anxiolytic efficacy and safety of Effexor ER (venlafaxine HCI) (37.5 mg to 225 mg/day) versus placebo in the treatment of children (6 to 11 years of age) and adolescents (12 to 17 years of age) with Generalized Anxiety Disorder (GAD). Study 0600B2-396-US was conducted at 36 centers in the United States while Study 0600B2-397-US was conducted at 29 centers in the United States. In Study 0600B2-396-US, a total of 164 patients were randomized and 160 were in the intent-to-treat population. In Study 0600B2-397-US, a total of 158 patients were randomized and 153 were in the intent-to-treat population.

In this submission the primary endpoint was based on the reduction of the C-KIDDIE-SADS GAD total score for the 9 delineated items at Week 8 on therapy evaluation. Study 0600B2-396-US did not give a statistically significant result in the reduction of primary endpoint. It gave a p-value of 0.06 in ITT-LOCF analysis. On the contrary, Study 0600B2-397-US gave a positive result in the reduction of primary endpoint with p-values below 0.001 in ITT-LOCF analysis.

2. Introduction

The current submission of NDA 20-151 for venlafaxine ER consists of two phase-III studies to compare the efficacy and safety of venlafaxine ER with placebo in children and adolescents with generalized anxiety disorder (GAD).

> **Study 0600B2-396-US** is a Phase III, multicenter, randomized, double-blind, parallel-group, placebo controlled, flexible-dose study to evaluate anxiolytic efficacy and safety of venlafaxine ER (37.5 mg to 225 mg/day) versus placebo in the treatment of children and adolescents with GAD.
>
> **Study 0600B2-397-US** is a Phase III, multicenter, randomized, double-blind, parallel-group, placebo controlled, flexible-dose study to evaluate anxiolytic efficacy and safety of venlafaxine ER (37.5 mg to 225 mg/day) versus placebo in the treatment of children and adolescents with GAD.

In the LOCF analysis, Study 0600B2-397-US significantly reduced the C-KIDDIE-SADS GAD total score for the 9 delineated items at Week 8 and therefore supports the conclusion that venlafaxine ER is more effective than placebo in improving clinical conditions of the pediatric patients with GAD. Study 0600B2-396-US does not produce statistically significant results in the same primary endpoint, therefore does not provide enough evidence supporting the conclusion that venlafaxine ER is more effective than placebo in improving clinical conditions of the pediatric patients with GAD.

3 Study 0600B2-396-US

The study period was between August 2000 and September 2001. The final protocol was signed off on August 23, 2001. There are 2 amendments to the original protocol. Significant changes were made in these amendments including changing the sample size and revising the primary endpoint.

3.1 Study Objectives

This study is to compare the anxiolytic efficacy and safety of venlafaxine extended release (ER) with placebo in the treatment of children and adolescents with Generalized Anxiety Disorder (GAD).

3.2 Study Design

This was a multicenter, parallel-group, randomized, double-blind, placebo-controlled, flexible-dose study in 164 pediatric outpatients with GAD. Following a 7±3 day single-blind placebo lead-in period, eligible patients were randomized to a double-blind phase receiving venlafaxine ER or placebo for up to 8 weeks, followed by a taper period of up to 14 days in duration. Patients were to return for a post study evaluation 4 to 10 days after taking the last dose of study medication, regardless of the length of time the study medication was taken.

Patients were evaluated using rating instruments, including the C-KIDDIE-SADS GAD, PARS, HAM-A, SCARED, CGI-S and CGI Global Improvement (CGI-I). The primary outcome measure was a subset of 9 delineated items of the C-KIDDIE-SADS GAD. It was chosen because it is the most specific instrument for GAD and because it correlates with the DSM-IV diagnostic criteria for GAD in children. Five (5) of the items comprise the severity component, and 4 items comprise the impairment component. By administering these items on a weekly basis, the clinician could note changes in both the severity of anxiety symptoms and the improvement associated with these symptoms.

This is a flexible dose treatment in which the dose schedules are different for the patients with different body weight. A detailed schedule is given in Table 6.4.1A in the Final Report of Study 0600B2-396-US.

It is assumed that 60% would respond for venlafaxine ER and 30% would respond for placebo on the primary endpoint. A sample size of 63 patients per group was considered to be sufficient for declaring a significant difference between the venlafaxine ER group and the placebo group at the 5% level with a power of 90%. In order to compensate for patients who might fail to qualify for the ITT analysis (5% of all patients), 70 patients were to be randomly assigned to each group. In the study, 272 patients were screened and 165 completed the single-blind placebo lead-in period and were randomly assigned to double-blind treatment. These patients were recruited from 36 centers in the United States.

There were two amendments to the original protocol on May 22, 2001 and August 23, 2001, respectively. Significant changes were made in these amendments. In the first amendment, the number of patients changed from 132 to approximately 140 and the number of centers changed from 20 to approximately 40. In the second amendment, the primary endpoint and primary time point were revised and clarified. Due to the large number of sites, the treatment by investigator interaction was removed from the ANCOVA model.

3.3 Efficacy Measures

The primary efficacy measure was the C-KIDDIE-SADS GAD total score for the 9 delineated items at Week 8 on therapy evaluation.

Secondary efficacy variables were the C-KIDDIE-SADS GAD Complete score, Severity component (5 delineated items) and Impairment component (4 delineated items) scores, the Pediatric Anxiety Rating Scale (PARS) total score, the Hamilton Rating Scale for Anxiety (HAM-A) total score, the Self Report for Childhood Anxiety Related Disorder (SCARED) Parent and Patient total scores, and the CGI-S and CGI-I scores.

3.4 Statistical Analysis Plan

The primary efficacy analysis population was the intent-to-treat (ITT) population, which included all patients who had entered the double-blind period, had taken at least 1 dose of their assigned medication, had at least 1 baseline evaluation for the primary efficacy variable, and had at least 1 evaluation for the primary efficacy variable either during therapy or within 3 days of the last day of treatment.

The intent-to-treat 2 (ITT2) population consists of all patients who were enrolled in double-blind therapy, had at least 1 baseline evaluation on the primary efficacy variable, had taken at least 1 dose of their assigned study medication, and had at least 1 post-baseline evaluation on the primary efficacy variable. This population is the same as the primary ITT, except that the efficacy evaluation during the double-blind period need not be “on-therapy” to be considered a valid post-baseline observation. Thus, a patient who had only a single post-baseline observation that was more than 3 days after their final dose would be included in the ITT2 population, but not in the ITT population.
A total of 161 patients were included in the ITT2 population.

The primary time point was the Week 8 last-observation-carried-forward (LOCF) on-therapy evaluation. An observation was considered on therapy if it occurred within 3 days of the last full dose. In addition to an LOCF analysis, an analysis of the observed data at each time point was employed. The observed data and LOCF data analyses were applied to the primary and secondary variables. Statistical tests for efficacy measures were two-sided and performed at the 0.05 level of significance. Tests of interaction were performed at the 0.1 significance level.

Changes from baseline for the primary and secondary efficacy variables were analyzed using a parametric 2-way analysis of covariance (ANCOVA) with treatment and investigator as factors and the associated baseline as the covariate. The CGI-I was analyzed by using the same model as the CGI-S, except that there was no baseline CGI-I to enter into the model. The assumptions of the primary ANCOVA model (ie, normality, homogeneity of variance, and parallelism of slopes) were tested. Observed case analyses were also performed using the same methods.

The primary efficacy variable from the population of all randomly assigned patients was also analyzed by using the Entsuah RANKing procedure (ETRANK). ETRANK uses a nonparametric technique that analyzes incomplete repeated measures data when the pattern of withdrawal is treatment or outcome related. The method uses either the observed full data (without inputting or estimating the missing data) or the endpoint data and creates efficient scoring systems.

In this analysis, different weights are assigned to patients' efficacy data at each time point. Those patients who withdrew before completion of a clinical trial were weighted more heavily at early time points than patients who remained in the trial, although the actual weights assigned to the efficacy data were contingent on the reasons underlying withdrawal (eg, intolerance, lack of therapeutic effect, etc.). However, patients who completed the clinical trial were weighted more heavily at the final time point than patients who withdrew before the final time point. The scoring systems generated were either categorical, time-related ranks, or observed levels, and were used to obtain a p-value or empirical significance level with time point descriptive statistics.

For all scales that required a total score, the total score was the mean of all items multiplied by the total number of items in the scale. If more than 50% of the items were missing from a scale, then the total score for that time point was not used in the analysis and was treated as missing. An additional approach to deal with missing items was used for the C-KIDDIE-SADS GAD. Specifically, with regard to the C-KIDDIE-SADS GAD (complete), if 3 or more of the individual Severity items, 2 or more of the Impairment items, or 50% or more of the total 29 items were missing, then the total score for this time point was not used in the analysis.

3.5 Study Population

The target population for this study consisted of patients who were from 6 to 17 years of age with at least 6 months of severe anxiety and worry. Patients were 1-1 randomized to venlafaxine ER and placebo. Patients were recruited from 36 centers in the United States.

Two hundred and seventy-two patients were screened, and 165 patients completed the single-blind placebo lead-in period and were randomly assigned to double-blind treatment. One patient did not return and was lost to follow-up and four patients had no primary evaluation on-therapy or within 3 days of study discontinuation. The ITT population consists of 160 patients for primary evaluation. One hundred twenty-nine patients completed 8 weeks of double-blind treatment.

For patient discontinuation (Table 3.5.1), venlafaxine group has protocol completed rate (81%) as compared to placebo (77%) which are not significant (p=0.70). The primary reason for early discontinuation was "Failed to return" for both groups.

Protocol deviation is one of major reasons that led to withdraw. Patients with major protocol violations are listed in Table 8.1.2A by the sponsor in the New Drug Application Final Report CSR-44723. Major protocol violations include receiving excluded concomitant medications, violating inclusion criteria and having errors in randomization, etc.

Table 3.5.1 Reasons for Discontinuations from Study - All Randomized Subjects

	Placebo		**Venlafaxine ER**	
Primary Reason for Discontinuation	(n=84) n	(%)	(n=80) n	(%)
Lack of efficacy	2	2	3	4
Adverse reaction	2	2	2	3
Protocol violation	3	4	2	3
Fail to return	5	6	5	6
Patient request unrelated to study	2	2	3	4
Other	2	2	3	4
TOTAL	16	19	18	23

Baseline patient characteristics including age, sex, ethnic origin, height, weight and duration of illness appeared to be comparable across treatment groups. The only notable difference was in the distribution of sex. This difference was statistically significant (p=0.006).

Baseline severity of illness based on both primary efficacy measure (C-KIDDIE-SADS GAD total score for 9 delineated items) and secondary efficacy measures (C-KIDDIE-SADS GAD Complete score, Severity component (5 delineated items) and Impairment component (4 delineated items) scores, PARS, HAM-A, SCARED total scores, CGI-S and CGI-I) was compared for ITT population. It appeared to be comparable across treatment groups.

Table 3.5.2 Baseline Demographic Characteristics – ITT Population

VARIABLE	**Placebo N=82**	**Venlafaxine ER N=78**
Age (yr)		
Mean (SD)	11.1 (2.6)	11.4 (3.2)
Min, Max	6.0 – 17.0	6.0 - 17.0
Sex, n (%)		
Female	46 (56)	27 (35)
Male	36 (44)	51 (65)

Ethnic origin, n (%)		
White	59 (72)	56 (72)
Black	7 (9)	7 (9)
Hispanic	14 (17)	13 (17)
Asian	0 (0)	1 (1)
Other	2 (2)	1 (1)
Weight (kg)		
Mean (SD)	51.5 (20.6)	48.7 (17)
Min, Max	23.0 – 105.0	25.0 - 94.0
Height(cm)	(n=81)	(n=78)
Mean (SD)	148.1 (15.5)	148.2 (15.1)
Min, Max	112.0 – 172.0	118.0 – 178.0
Duration of episode, weeks		
Mean (SD)	169 (128)	179 (150)
Min, Max	30.0 – 535.0	27.0 - 697.0

Table 3.5.3 Baseline Efficacy Score - Baseline Severity of Illness in ITT Population

Efficacy Parameters at Baseline	**Placebo** N = 82	**Venlafaxine ER** N = 78	**P-value**
C-KIDDIE-SADS (9 delineated items)			0.533[t]
Mean (SD)	39.7 (4.8)	39.3 (3.7)	
Range	28.0-52.0	32.0-48.0	
C-KIDDIE-SADS (complete)	n = 81	n = 78	0.401[t]
Mean (SD)	74.3 (6.0)	73.5 (4.8)	
Range	57.0-91.0	61.0-85.0	
C-KIDDIE-SADS Impairment (4 items)			
Mean (SD)	16.2 (2.3)	15.9 (2.2)	0.391[t]
Range	9.0-21.0	9.0-20.0	
C-KIDDIE-SADS Severity (5 items)			0.777[t]
Mean (SD)	23.5 (3.1)	23.4 (2.6)	
Range	17.0-31.0	18.0-29.0	
PARS			0.842[t]
Mean (SD)	23.7 (3.2)	23.8 (3.0)	
Range	9.0-31.0	14.0-30.0	
HAM-A			0.459[t]
Mean (SD)	18.8 (7.2)	19.6 (6.2)	
Range	6.0-48.0	5.0-35.0	
SCARED parent			0.788[t]
Mean (SD)	35.9 (11)	36.4 (12)	
Range	8.0-63.0	7.0-71.0	
SCARED patient	(n = 82)	(n = 77)	0.938 [t]
Mean (SD)	33.1 (15.1)	33.3 (15.4)	
Range	2.0-68.0	3.0-71.0	

CGI-Severity			0.817[t]
Mean (SD)	4.4 (0.6)	4.5 (0.6)	
Range	4.0-6.0	4.0-6.0	
CGI-Severity, n (%)			
0			0.614[c]
4	52 (63)	46 (59)	
5	24 (29)	28 (36)	
6	6 (7)	4 (5)	

[c]and [t] indicate p-value for chi-square or paired t-test, respectively.

3.6 Sponsor's Efficacy Results

3.6.1 Primary Efficacy Results

Change from baseline of the C-KIDDIE-SADS GAD (9 delineated items) was analyzed using ANCOVA with treatment and investigator as main effects and the baseline as covariates. The normality, homogeneity of variances and parallelism of regression lines are verified for the primary endpoint at Week 8. None of these assumptions were seriously violated by the significance level of the testing statistics. So the statistical model used for testing the treatment effect of venlafaxine ER is acceptable. Table 3.6.1 gives the results of LOCF analysis for the significance of the changes from baseline of the C-KIDDIE-SADS GAD at Week 8. Patients in the venlafaxine ER group had marginally lower scores (p = 0.06) on the C-KIDDIE-SADS GAD than those in the placebo group.

Table 3.6.1 Reduction of C-KIDDIE-SADS GAD 9 Delineated Items from Baseline to Week 8 (ITT Population) – LOCF Analysis

Primary Efficacy Parameters	**Placebo** (N=82)	**Venlafaxine ER** (N=78)	**P-value**[a]
Baseline			0.533
Mean (SD)	39.7 (4.8)	39.3 (3.7)	
Min, Max	28.0-52.0	32.0-48.0	
Week 8			
Mean (SE)	26.7 (1.17)	23.5 (1.12)	0.06
95% Confidence Interval[c]	(-0.15, 5.9)		

[a] Comparison of treatment groups using ANCOVA (with treatment, baseline). [c] Computed for difference between changes of primary endpoint in venlafaxine and placebo.

3.6.2 Secondary Efficacy Results

Changes from baseline for the secondary efficacy variables (i.e., C-KIDDIE-SADS GAD

(complete, 5 delineated severity items, and 4 delineated impairment items), PARS, HAM-A, SCARED, and CGI-Severity score) were analyzed separately at Week 8. They were assessed using analysis of covariance (ANCOVA) with treatment and investigator as main effects and their respective baseline score as the covariate. The CGI-Improvement subscale was analyzed using an analysis of variance (ANOVA) with treatment and investigator as main effects. A summary of the LOCF results of these analyses appears in Tables 3.6.2. As depicted in the table, the effect of treatment was significant at Week 8 for the C-KIDDIE-SADS Severity (5 severity items), CGI-S and CGI-I.

Table 3.6.2 Secondary Efficacy Measure at Endpoint for ITT Population—LOCF Analysis

Secondary Efficacy Parameters At Endpoint	**Placebo** N = 82 N (%)	**Venlafaxine ER** N = 78 N (%)	**P-value** [b]
C-KIDDIE-SADS (complete)			
Mean change from baseline (SE)[a]	-17.6 (1.64)	-21.4 (1.58)	0.077
95% confidence interval	(-0.4, 7.9)		
N	76	74	
C-KIDDIE-SADS Impairment (4 items)			
Mean change at baseline (SE) [a]	-5.1 (0.56)	-6.0 (0.49)	0.215
95% confidence interval	(-0.5, 2.3)		
N	82	78	
C-KIDDIE-SADS Severity (5 items)			
Mean change at Week 8 (SE) [a]	-7.6 (0.65)	-9.7 (0.66)	0.017
95% confidence interval	(0.4, 4.0)		
N	82	78	
PARS			
Mean change at baseline (SE) [a]	-8.2 (0.82)	-10.2 (0.80)	0.071
95% confidence interval	(-0.1, 4.1)		
N	82	78	
HAM-A			
Mean change at baseline (SE) [a]	-9.1 (0.90)	-10.6 (0.79)	0.141
95% confidence interval	(-0.5, 3.7)		
N	77	74	
SCARED parent			
Mean change at Week 8 (SE) [a]	-10.7 (1.72)	-14.9 (1.71)	0.065
95% confidence interval	(-0.2, 8.7)		
N	77	73	
SCARED patient			
Mean change at Week 8 (SE) [a]	-11.5 (1.94)	-15.2 (1.63)	0.102
95% confidence interval	(-0.7, 8.1)		
N	77	73	
CGI-Severity			
Mean change at Week 8 (SE) [a]	-1.3 (0.14)	-1.7 (0.15)	0.038
95% confidence interval	(0.0, 0.8)		
N	82	78	
CGI-Improvement			
Mean score at Week 8 (SE) [a]	2.5 (0.12)	2.1 (0.11)	0.018
95% confidence interval	(0.1, 0.8)		
N	82	78	

(a) These are the least square adjusted means and standard errors. (b) The p-values are derived based on the least square adjusted means and standard errors.

3.7 Reviewer's Analysis

Using the ITT-LOCF data set provided by the sponsor, the reviewer duplicated the sponsor's analysis according to the protocol and obtained the same results. The results are depicted in Table 3.7.1.

Table 3.7.1 Reduction of C-KIDDIE-SADS GAD (9 delineated items) from Baseline to Week 8 (ITT Population) – LOCF Analysis

Primary Efficacy Parameters	**Placebo** (N=82)	**Venlafaxine ER** (N=78)	**P-value**
Baseline			0.533
Mean (SD)	39.7 (4.8)	39.3 (3.7)	
Min, Max	28.0-52.0	32.0-48.0	
Week 8			
Mean change from baseline (SE)	-12.6 (1.28)	-15.5 (1.24)	0.06
95% Confidence Interval[c]	(-0.12, 6.0)		

[a] Comparison of treatment groups using ANCOVA (with treatment, investigator, baseline). [c] Computed for difference between changes of primary endpoint in venlafaxine and placebo.

Normality assumption is tested for the primary endpoint of reduction in C-KIDDIE-SADS GAD (9 delineated items) from baseline. Although the Kolmogorov-Smirnov D test gives p-values of about 0.15 for both treatment and placebo groups, the Shapiro-Wilks test gives p-values of 0.05 and 0.03 for treatment and placebo group, respectively. It indicates that the normality assumption of the primary endpoint is problematic. The reviewer performed nonparametric tests on the reduction from baseline of the primary endpoint of C-KIDDIE-SADS GAD (9 delineated items). Both the Wilcoxon and Kruskal-Wilks tests give p-values around 0.069. The Wilcoxon test confirms the testing results in Table 3.7.1. Parallelism of the regression line for the placebo and venlafaxine ER treatment groups were tested by the reviewer by testing the interaction between the baseline covariate and the treatment indicator. This test yields a nonsignificant result with a p-value of 0.61. This indicates an acceptable assumption of parallelism between the regression lines of two treatment groups.

There are 36 investigators in total therefore it is hard to do a subgroup analysis for each site.

The following table gives t-test results for the treatment differences by sex. DIFF is the mean change from baseline to Week 8 on the primary endpoint of C-KIDDIE-SADS GAD (9 delineated items). VENLADIFF is the difference between DIFF of venlafaxine and Placebo.

Table 3.7.2 Treatment Effect by Sex for the Reduction of C-KIDDIE-SADS GAD (9 delineated items)

Sex	Therapy	Patient	DIFF	VENLADIFF	t-Value
Male	Venlafaxine ER	51	-15.9	-2.03	-0.98
	Placebo	36	-13.9		(p=0.33)
Female	Venlafaxine ER	27	-15.63	-3.24	-1.43
	Placebo	46	-12.39		(p=0.16)

The above table shows that venlafaxine ER has some treatment effect in both male and female groups but it does not have statistically significant result in either of these groups.

The following table gives t-test results for the treatment differences by age groups. The population is separated into two age groups: young children from 6 to11 years of age and adolescents from 12 to 17 years of age. DIFF is the mean change from baseline to Week 8 on the primary endpoint of C-KIDDIE-SADS GAD (9 delineated items). VENLADIFF is the difference between DIFF of venlafaxine ER and placebo.

Table 3.7.3 Treatment Effect by Age Group for the Reduction of C-KIDDIE-SADS GAD (9 delineated items)

Age	Therapy	Patient	DIFF	VENLADIFF	t-Value
6-11 years	Venlafaxine ER	44	-16.61	-2.5	-1.28
	Placebo	45	-14.11		(p=0.20)
12-17 years	Venlafaxine ER	34	-14.77	-3.03	-1.32
	Placebo	37	-11.74		(p=0.19)

So there was some treatment effect of venlafacxine ER in both age groups but no effect in these groups is statistically significant.

There are 72% White and 28% nonwhite patients in both the venlafaxine ER and placebo groups. We performed a group analysis for White and nonwhite groups. The results are depicted in Table 3.7.4.

Table 3.7.4 Treatment Effect by Ethnic Groups for the Reduction of C-KIDDIE-SADS GAD (9 delineated items)

Race	Therapy	Patient	DIFF	VENLADIFF	t-Value
White	Venlafaxine ER	56	-17.34	-4.81	-2.64
	Placebo	59	-12.53		(p=0.01)
nonwhite	Venlafaxine ER	22	-11.9	2.44	1.02
	Placebo	23	-14.35		(p=0.31)

The results indicate that there is a statistically significant treatment effect of venlafaxine ER in the White group but not in the nonwhite group. In fact, the treatment in the nonwhite patients has an opposite effect in which patients in the placebo group has a more significant reduction compared to the venlafaxine group in the C-KIDDIE-SADS GAD (9 delineated items) even though this effect is not statistically significant.

4 Study 0600B2-397-US

The study period was between April 2000 and August 2001. The final protocol was signed off on August 23, 2001. There are 3 amendments to the original protocol. Significant changes were made in these amendments including changing the sample size and revising primary endpoint.

4.1 Study Objectives

This study is to compare the anxiolytic efficacy and safety of venlafaxine extended release (ER) with placebo in the treatment of children and adolescents with Generalized Anxiety Disorder (GAD).

4.2 Study Design

Same as Study 0600B2-396, this was a multicenter, parallel-group, randomized, double-blind, placebo-controlled, flexible-dose study in 158 pediatric outpatients with GAD. Following a 7±3 day single-blind placebo lead-in period, eligible patients were randomized to a double-blind phase receiving venlafaxine ER or placebo for up to 8 weeks, followed by a taper period of up to 14 days in duration. Patients were to return for a post study evaluation 4 to 10 days after taking the last dose of study medication, regardless of the length of time the study medication was taken. Again, this study is a flexible dose treatment in which the dose schedules are different for the patients with different body weight. A detailed schedule is given in Table 6.4.1A in the Final Report of Study 0600B2-397-US.

It is assumed that 60% would respond for venlafaxine ER and 30% would respond for placebo on the primary endpoint. A sample size of 63 patients per group was considered to be sufficient for declaring a significant difference between the venlafaxine ER group and the placebo group at the 5% level with a power of 90%. In order to compensate for patients who might fail to qualify for the ITT analysis (5% of all patients), 70 patients were to be randomly assigned to each group. In the study, 245 patients were screened and 158 completed the single-blind placebo lead-in period and were randomly assigned to double-blind treatment. One hundred and fifty-six patients received randomly assigned study medication under double-blind conditions and are included in safety analyses. Three of them did not meet ITT criteria and were excluded from primary efficacy evaluation. A total of 112 patients completed 8 weeks of double-blind treatment.

There were three amendments to the original protocol on April 11, 2001 and June 6, 2001 and August 23, 2001, respectively. Significant changes were made in these amendments. In the first amendment, the number of patients was changed from132 to approximately 140 and the number of centers changed from 20 to approximately 40. In the third amendment, the primary endpoint and primary time point were revised and clarified. Due to the large number of sites, the treatment by investigator interaction was removed from the ANCOVA model.

4.3 Efficacy Measures

The primary efficacy measure was the C-KIDDIE-SADS GAD total score for the 9 delineated items at Week 8 on therapy evaluation.

Secondary efficacy variables were the C-KIDDIE-SADS GAD Complete score, Severity component (5 delineated items) and Impairment component (4 delineated items) scores, the Pediatric Anxiety Rating Scale (PARS) total score, the Hamilton Rating Scale for Anxiety (HAM-A) total score, the Self Report for Childhood Anxiety Related Disorder (SCARED) Parent and Patient total scores, and the CGI-S and CGI-I scores.

4.4 Statistical Analysis Plan

The primary efficacy analysis population was the intent-to-treat (ITT) population, which included all patients who had entered the double-blind period, had taken at least 1 dose of their assigned medication, had at least 1 baseline evaluation for the primary efficacy variable, and had at least 1 evaluation for the primary efficacy variable either during therapy or within 3 days of the last day of treatment.

The intent-to-treat 2 (ITT2) population consists of all patients who were enrolled in double-blind therapy, had at least 1 baseline evaluation on the primary efficacy variable, had taken at least 1 dose of their assigned study medication, and had at least 1 post-baseline evaluation on the primary efficacy variable. A total of 155 patients were included in the ITT2 population.

The primary time point was the Week 8 last-observation-carried-forward (LOCF) on-therapy evaluation. In addition to an LOCF analysis, an analysis of the observed data at each time point was employed. The observed data and LOCF data analyses were applied to the primary and secondary variables. Statistical tests for efficacy measures were two-sided and performed at the 0.05 level of significance. Tests of interaction were performed at the 0.1 significance level.

Changes from baseline for the primary and secondary efficacy variables were analyzed using a parametric 2-way analysis of covariance (ANCOVA) with treatment and investigator as factors and the associated baseline as the covariate. The CGI-I was analyzed by using the same model as the CGI-S, except that there was no baseline CGI-I to enter into the model. The assumptions of the primary ANCOVA model (ie, normality, homogeneity of variance, and parallelism of slopes) were tested. Observed case analyses were also performed using the same methods.

The primary efficacy variable from the population of all randomly assigned patients was also analyzed by using the Entsuah Ranking procedure (ETRANK). ETRANK uses a nonparametric technique that analyzes incomplete repeated measures data when the pattern of withdrawal is treatment or outcome related. The method uses either the observed full data (without inputting or estimating the missing data) or the endpoint data and creates efficient scoring systems.

In this analysis, different weights are assigned to patients' efficacy data at each time point. Those patients who withdrew before completion of a clinical trial were weighted more heavily at early time points than patients who remained in the trial, although the actual weights assigned to the efficacy data were contingent on the reasons underlying withdrawal (eg, intolerance, lack of therapeutic effect, etc.). However, patients who completed the clinical trial were weighted more heavily at the final time point than patients who withdrew before the final time point. The scoring systems generated were either categorical, time-related ranks, or observed levels, and were used to obtain a p-value or empirical significance level with time point descriptive statistics.

For all scales that required a total score, the total score was the mean of all items multiplied by the total number of items in the scale. If more than 50% of the items were missing from a scale, then the total score for that time point was not used in the analysis and was treated as missing. An additional approach to deal with missing items was used for the C-KIDDIE-SADS GAD. Specifically, with regard to the C-KIDDIE-SADS GAD (complete), if 3 or more of the individual Severity items, 2 or more of the Impairment items, or 50% or more of the total 29 items were missing, then the total score for this time point was not used in the analysis.

4.5 Study Population

The target population for this study consisted of patients from 6 to 17 years of age with at least 6 months of severe anxiety and worry. Patients were 1-1 randomized to venlafaxine ER and placebo. Patients were recruited from 29 centers in the United States.

Two hundred and forty-five patients were screened, and 158 of them completed the single-blind placebo lead-in period and were randomly assigned to double-blind treatment. Two patients did not return and were lost to follow-up. The ITT population consists of 153 patients for primary evaluation. One hundred and twelve patients completed 8 weeks of double-blind treatment.

For patient discontinuation (Table 4.5.1), venlafaxine group has protocol completion rate (68%) as compared to placebo (75%) which are not statistically significant (p=0.376). The primary reasons for early discontinuation were "Failed to return", "Adverse reaction" and "Unsatisfactory response" for treatment group and "Failed to return" and "Unsatisfactory response" for placebo group.

Protocol deviation was one of major reasons that led to withdraw. Patients with major protocol violations are listed in Table 8.1.2A by the sponsor in the New Drug Application Final Report CSR-44734. Major protocol violations include receiving excluded concomitant medications, violating inclusion criteria etc.

Table 4.5.1 Reasons for Discontinuations from Study - All Randomized Subjects

	Placebo		Venlafaxine ER	
Primary Reason for Discontinuation	(n=79) n	(%)	(n=77) n	(%)
Unsatisfactory response	5	(6)	5	(6)
Adverse reaction	7	(9)	1	(1)
Protocol violation	1	(1)	4	(5)
Fail to return	9	(11)	7	(9)
Patient request	2	(3)	1	(1)
Other	1	(1)	1	(1)
TOTAL	25	(32)	19	(25)

Baseline patient characteristics including age, sex, ethnic origin, height, weight and duration of illness appeared to be comparable across treatment groups. In fact, none of the differences was statistically significant.

Baseline severity of illness based on both primary efficacy measure (C-KIDDIE-SADS GAD total score for 9 delineated items) and secondary efficacy measures (C-KIDDIE-SADS GAD Complete score, Severity component (5 delineated items) and Impairment component (4 delineated items) scores, PARS, HAM-A, SCARED total scores, CGI-S and CGI-I) was compared for ITT population. It appeared to be comparable across treatment groups.

Table 4.5.2 Baseline Demographic Characteristics – ITT Population

VARIABLE	Placebo N=77	Venlafaxine ER N=76
Age (yr)		
Mean (SD)	11.3 (2.8)	11.7 (3.0)
Min, Max	7.0 - 17.0	6.0 – 17.0
Sex, n (%)		
Female	31 (40)	29 (38)
Male	45 (60)	47 (62)
Ethnic origin, n (%)		
White	63 (82)	59 (78)
Black	9 (12)	10(13)
Hispanic	3 (4)	5 (7)
Asian	1 (1)	1 (1)
Other	1 (1)	1 (1)
Weight (kg)		
Mean (SD)	48.8 (18.7)	52.4 (21.1)
Min, Max	25 – 108	25 – 113
Height(cm)		
Mean (SD)	148.3 (14.9)	151.9 (17.0)
Min, Max	122 – 184	104 – 185
Duration of episode, weeks		
Mean (SD)	163 (143)	172 (161)
Min, Max	27 - 642	0.0 – 654

Table 4.5.3 Baseline Efficacy Score - Baseline Severity of Illness in ITT Population

Efficacy Parameters at Baseline	Placebo N = 77	Venlafaxine ER N = 76	P-value
C-KIDDIE-SADS (9 delineated items)			0.959[t]
Mean (SD)	40.3 (3.7)	40.4 (3.7)	
Range	32 - 48	28 - 50	
C-KIDDIE-SADS (complete)			0.954[t]
Mean (SD)	75.3 (4.9)	75.2 (5.2)	
Range	63 - 85	54 – 88	
C-KIDDIE-SADS Impairment (4 items)			
Mean (SD)	16.6 (1.7)	16.6 (2.3)	0.982[t]
Range	13 - 21	10 – 23	
C-KIDDIE-SADS Severity (5 items)			0.953[t]
Mean (SD)	23.8 (2.6)	23.8 (2.3)	
Range	16 - 29	18 – 28	
PARS			0.436[t]
Mean (SD)	23.9 (3.0)	24.3 (3.6)	
Range	14 -19	11 - 34	
HAM-A			0.382[t]
Mean (SD)	18.8 (5.8)	19.7 (5.9)	
Range	7 - 36	6 - 32	
SCARED parent	(N=77)	(N=74)	0.281[t]
Mean (SD)	38.3 (12.5)	36.2 (12)	
Range	13 – 65	3 – 64	
SCARED patient	(N=77)	(N=75)	0.229[t]
Mean (SD)	33.4 (14.1)	36.1 (12.8)	
Range	3 - 66	7 – 63	
CGI-Severity			0.755[t]
Mean (SD)	4.5 (0.6)	4.4 (0.7)	
Range	4 – 6	2 - 7	
CGI-Severity, n (%)			
0			0.430[c]
2	0 (0)	1 (1)	
3	0 (0)	0 (0)	
4	44 (57)	47 (62)	
5	30 (39)	22 (29)	
6	3 (4)	5 (7)	
7	0 (0)	1 (1)	

[c] and [t] indicate p-value for chi-square or paired t-test, respectively.

4.6 Sponsor's Efficacy Results

4.6.1 Primary Efficacy Results

Change from baseline of the C-KIDDIE-SADS GAD (9 delineated items) was analyzed using ANCOVA with treatment and investigator as main effects and the baseline as covariates. The normality, homogeneity of variances and parallelism of regression lines were verified for the primary endpoint at Week 8. None of these assumptions were seriously violated at the significance level of the testing statistics except the test of normality using Kolmogorov-Smirnov D test was close to be significant (p=0.11 for placebo group and p=0.06 for treatment group). So the statistical model used for testing the treatment effect of venlafaxine ER is acceptable. Table 4.6.1 gives the results of LOCF analysis for the significance of the changes from baseline of the C-KIDDIE-SADS GAD at Week 8. Patients in the venlafaxine ER group had significantly lower scores ($p < 0.001$) on the C-KIDDIE-SADS GAD than those in the placebo group.

Because the p-value was close to the 0.05 level for testing the normality of the venlafaxine ER group (indicating a marginally significant skewness of the data), the ranks of the C-KIDDIE-SADS GAD (9 delineated items) scores were submitted to an ANCOVA with treatment and investigator as main effects and the baseline score as the covariate. As may be seen in Table 4.6.2, a statistically significant advantage of venlafaxine ER over placebo is still observed.

The observed case analysis for the primary endpoint was also performed by the sponsor at Week 8. The result was significant with a p-value of 0.005 that supported the significance of venlafaxine ER compared to placebo in reducing the C-KIDDIE-SADS GAD at Week 8. The result is depicted in Table 4.6.3.

Table 4.6.1 Reduction of C-KIDDIE-SADS GAD (9 delineated items) from Baseline to Week 8 (ITT Population) – LOCF Analysis

Primary Efficacy Parameters	**Placebo** (N=77)	**Venlafaxine ER** (N=76)	**P-value**[a]
Baseline			0.959
Mean (SD)	40.3 (3.7)	40.4 (3.7)	
Min, Max	32 - 48	28 - 50	
Week 8			
Mean change from baseline (SE[b])	-12.4 (1.18)	-18.6 (1.16)	<0.001
95% Confidence Interval[c]	(3.0, 9.5)		

[a] Comparison of treatment groups using ANOVA (with treatment, baseline). [b] Standard error is not adjusted by covariates. [c] Computed for difference between changes of primary endpoint in venlafaxine and placebo.

Table 4.6.2: C-KIDDIE-SADS GAD (9 delineated items) Week 8 ANCOVA on ranked data

Primary Efficacy Parameters	Placebo (N=77)	Venlafaxine ER (N=76)	p-Value
C-KIDDIE-SADS GAD (9 delineated items) Adjusted Means	78.14	51.74	0.0002

Table 4.6.3 Reduction of C-KIDDIE-SADS GAD (9 delineated items) from Baseline to Week 8 (ITT Population) – Observed Case Analysis

Primary Efficacy Parameters	Placebo (N=49)	Venlafaxine ER (N=53)	P-value[a]
Mean change from baseline (SE[b])	-15.2 (1.54)	-21.1 (1.23)	0.005
95% Confidence Interval[c]	(1.9, 9.8)		

[a] Comparison of treatment groups using ANOVA (with treatment, baseline). [b] Standard error is not adjusted by covariates. [c] Computed for difference between changes of primary endpoint in venlafaxine and placebo.

4.6.2 Secondary Efficacy Results

Changes from baseline for the secondary efficacy variables (i.e., C-KIDDIE-SADS GAD (complete, 5 delineated severity items, and 4 delineated impairment items), PARS, HAM-A, SCARED, and CGI-Severity score) were analyzed at Week 8. They were assessed using analysis of covariance (ANCOVA) with treatment and investigator as main effects and their respective baseline score as the covariate. The CGI-Improvement subscale was analyzed using an analysis of variance (ANOVA) with treatment and investigator as main effects. A summary of the LOCF results of these analyses appears in Tables 4.6.3. As may be seen, the effect of treatment was significant (all p's < 0.01) for all secondary variables at the primary endpoint (Week 8). These results provide supporting evidence to that obtained on the primary variable indicating that patients receiving venlafaxine ER showed greater improvements than patients receiving placebo.

Table 4.6.3 Secondary Efficacy Measure at Endpoint for ITT Population—LOCF Analysis

Secondary Efficacy Parameters At Endpoint	Placebo N = 77 N (%)	Venlafaxine ER N = 76 N (%)	P-value[b]
C-KIDDIE-SADS (complete)			
Mean change from baseline (SE)[a]	-20.2 (1.82)	-28.3 (1.74)	0.001
95% confidence interval	(3.4, 12.9)		
N	67	65	
C-KIDDIE-SADS Impairment (4 items)			
Mean change at baseline (SE)[a]	-5.3 (0.53)	-7.5 (0.5)	0.002
95% confidence interval	(0.8, 3.7)		

N	77	76	
C-KIDDIE-SADS Severity (5 items) Mean change at Week 8 (SE) [a] 95% confidence interval N	 -7.2 (0.69) (2.0, 5.9) 77	 -11.2 (0.69) 76	 <0.001
PARS Mean change at baseline (SE) [a] 95% confidence interval N	 -7.3 (0.86) (1.8, 6.3) 77	 -11.3 (0.81) 76	 <0.001
HAM-A Mean change at baseline (SE) [a] 95% confidence interval N	 -9.5 (0.97) (1.2, 5.8) 67	 -13.0 (0.78) 65	 0.003
SCARED parent Mean change at Week 8 (SE) [a] 95% confidence interval N	 -12.4 (1.75) (1.8, 10.4) 67	 -18.4 (1.58) 63	 0.007
SCARED patient Mean change at Week 8 (SE) [a] 95% confidence interval N	 -10.1 (2.04) (2.7, 11.5) 67	 -17.2 (1.62) 64	 0.002
CGI-Severity Mean score at Week 8 (SE) [a] 95% confidence interval N	 -1.2 (0.14) (0.3, 1.0) 77	 -1.9 (0.13) 76	 <0.001
CGI-Improvement Mean score at Week 8 (SE) [a] 95% confidence interval N	 2.7 (0.16) (0.3, 1.1) 77	 2.0 (0.12) 76	 <0.001

(a) These are the least square adjusted means and standard errors. (b) The p-values are derived based on the least square adjusted means and standard errors.

4.7 Reviewer's Analysis

Using the ITT data set provided by the sponsor, the reviewer duplicated the sponsor's analysis results according to the protocol. The results are reported in Table 4.7.1.

Table 4.7.1 Reduction of C-KIDDIE-SADS GAD (9 delineated items) from Baseline to Week 8 (ITT Population) – LOCF Analysis

Primary Efficacy Parameters	**Placebo** (N=77)	**Venlafaxine ER** (N=76)	**P-value**[a]
Baseline Mean (SD) Min, Max	 40.3 (3.7) 32 - 48	 40.4 (3.7) 28 – 50	0.959
Week 8 Mean Change from Baseline (SE[b]) 95% Confidence Interval[c]	 -12.4 (1.18) (3.0, 9.5)	 -18.6 (1.16) 	 <0.001

[a] Comparison of treatment groups using ANOVA (with treatment, baseline). [b] Standard error is not adjusted by covariates. [c] Computed for difference between changes in venlafaxine and placebo values.

Normality assumption was tested for the primary endpoint of the reduction in C-KIDDIE-SADS GAD (9 delineated items) from baseline. The Kolmogorov-Smirnov D test gave p-values of 0.06 and 0.11 for treatment and placebo group, respectively. The Shapiro-Wilks test gave p-values of 0.04 and 0.01 for treatment and placebo group, respectively. So it is problematic to assume the normality of population.

The reviewer performed nonparametric tests on significance of the reduction from baseline of the primary endpoint of C-KIDDIE-SADS GAD (9 delineated items). Both the Wilcoxon and Kruskal-Wilks tests gave p-values around 0.001. These tests confirmed the significance results in Table 4.7.1. Parallelism of the regression lines for the placebo and venlafaxine ER treatment groups was tested by testing the significance of an interaction between baseline covariate and treatment indicator. This test yields a nonsignificant result with a p-value of 0.93. This indicates an acceptable assumption of parallelism between the regression lines of two treatment groups.

There are 29 investigators in total. Therefore it is hard to do a subgroup analysis for each site.

The following table gives t-test results for the treatment differences by sex. DIFF is the mean change from baseline to Week 8 on the primary endpoint of C-KIDDIE-SADS GAD (9 delineated items). VENLADIFF is the difference between DIFF of venlafaxine ER and Placebo.

Table 4.7.2 Treatment Effect by Sex for the Reduction of C-KIDDIE-SADS GAD (9 delineated items)

Sex	Therapy	Patient	DIFF	VENLADIFF	t-Value
Male	Venlafaxine ER	47	-15.51	-4.69	-2.23
	Placebo	46	-10.83		(p=0.028)
Female	Venlafaxine ER	29	-15.55	-6.36	-2.50
	Placebo	31	-9.19		(p=0.015)

The above table shows that venlafaxine ER has statistically significant treatment effect in both male and female groups.

The following table gives t-test results for the treatment differences by age groups. The population is separated into two age groups: young children from 6 to11 years of age and adolescents from 12 to 17 years of age. DIFF is the mean change from baseline to Week 8 on the primary endpoint of C-KIDDIE-SADS GAD (9 delineated items). VENLADIFF is the difference between DIFF of venlafaxine ER and Placebo.

Table 4.7.3 Treatment Effect by Age Group for the Reduction of C-KIDDIE-SADS GAD (9 delineated items)

Age	Therapy	Patient	DIFF	VENLADIFF	t-Value
6-11 years	Venlafaxine ER	39	-14.21	-4.90	-2.14 (p=0.036)
	Placebo	42	-9.31		
12-17 years	Venlafaxine ER	37	-16.92	-5.72	-2.52 (p=0.014)
	Placebo	35	-11.2		

So there are statistically significant treatment effects of venlafacxine ER in both age groups.

There are 78% White and 22% nonwhite patients in the venlafaxine ER group, respectively. There are 82% White and 18% in the placebo group, respectively. Although the percentage of nonwhites in both treatment and placebo groups are rather low, we performed a group analysis for these two racial groups. The results are depicted in Table 4.7.4.

Table 4.7.4 Treatment Effect by Ethnic Groups for the Reduction of C-KIDDIE-SADS GAD (9 delineated items)

Race	Therapy	Patient	DIFF	VENLADIFF	t-Value
White	Venlafaxine ER	59	-15.25	-5.49	-3.00 (p=0.003)
	Placebo	63	-9.76		
Nonwhite	Venlafaxine ER	17	-16.47	-4.47	-1.29 (p=0.21)
	Placebo	14	-12		

The results indicate that the treatment effect of venlafaxine ER in White patients is highly significant. Although such an effect is not statistically significant in the nonwhite group, it seems to be caused by the low sample size. In fact, the treatment effect in nonwhite group is in the same direction and has the similar magnitude as in the White group.

5 Conclusion

In this submission, the sponsor conducted two Phase III, placebo controlled clinical trial studies that evaluate the anxiolytic efficacy and safety of Effexor ER (venlafaxine HCI) (37.5 mg to 225 mg/day) versus placebo in the treatment of pediatric patients (6 through 17 years of age) with Generalized Anxiety Disorder (GAD).

In the LOCF analysis, Study 0600B2-397-US significantly reduced the C-KIDDIE-SADS GAD total score for the 9 delineated items at Week 8 and therefore supports the conclusion that venlafaxine ER is more effective than placebo in improving clinical conditions of the pediatric patients with GAD. The model assumptions made by the sponsor on the primary endpoint were checked by the reviewer and were found to be acceptable. However, the normality assumption was found to be problematic so the Wilcoxon nonparametric test was used to test the significance of the treatment effect. The results supported the sponsor's conclusion. However, in Study 0600B2-396-US, the treatment was not found to be statistically significant so it did not support the conclusion that venlafaxine ER is more

effective than placebo in improving patient's condition by reducing the C-KIDDIE-SADS GAD total score for the 9 delineated items at Week 8 among pediatric patients.

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
2/13/03 05:07:04 PM
BIOMETRICS

Kun Jin
2/14/03 09:56:42 AM
BIOMETRICS

George Chi
2/14/03 03:47:26 PM
BIOMETRICS