

**CENTER FOR DRUG EVALUATION AND  
RESEARCH**

*APPLICATION NUMBER:*  
**21-196**

**STATISTICAL REVIEW(S)**

## Statistical Review and Evaluation

MAR 23 2001

<b><u>NDA#:</u></b>	21-196
<b><u>DRUG COMPANY:</u></b>	Orphan Medical, Inc.
<b><u>NAME OF DRUG:</u></b>	Xyrem (Sodium Oxybate)
<b><u>INDICATION:</u></b>	Narcolepsy
<b><u>STUDIES REVIEWED:</u></b>	OMC -GHB-2, Scrima, Lammers, and SXB-21
<b><u>DOCUMENTS REVIEWED:</u></b>	Sponsor's original NDA submission
<b><u>MEDICAL REVIEWER</u></b>	Ranjit Mani, M.D.

### 1. Introduction

Narcolepsy is a chronic neurological disorder characterized by excessive daytime sleepiness, disturbed nocturnal sleep, cataplexy, sleep paralysis, and hypnagogic hallucinations (Aldrich, 1990). The usual treatment for narcolepsy includes symptomatic treatment of daytime sleepiness with stimulants. The symptoms of cataplexy, hypnagogic hallucinations, and sleep paralysis are typically treated with tricyclic antidepressants (TCAs) or selective serotonin reuptake inhibitors (SSRIs). The effectiveness of these treatments is limited by frequent undesirable adverse events and tolerance.

As part of the sodium oxybate development project, Orphan Medical, Inc. acquired the rights to the Scrima and Lammers data for this NDA. Results from these two trials and along with Orphan Medical's own trial OMC-GHB-2 were submitted in this NDA to form the data sets to establish the efficacy of sodium oxybate in the treatment of narcolepsy symptoms. After the submission of the NDA, another long-term efficacy trial SXB-021 was completed. Results of the study SXB-021 were later submitted to be included in this NDA.

### 2. Specifications and Findings of Study OMC-GHB-2

#### 2.1 Objectives

The primary objective of this study was to evaluate and compare the efficacy of 3 doses (3g, 6g, and 9g) of GHB and placebo in the treatment of the symptoms of narcolepsy.

The secondary objective of this study was to evaluate and compare the safety of GHB and placebo when used in a narcoleptic patient population.

#### 2.2 Study Design

This was a randomized, double blind, placebo controlled, parallel-group, multi-center study of 3 doses of GHB and placebo in the treatment of patients with narcolepsy. The study was conducted in 16 centers and a total of 136 patients were enrolled. The study was divided into 5 periods as follows:

Table 1. Study flow

Screening	Washout	Baseline	Double-blind Treatment	Follow-up
1 day - 4 wks	5 - 28 days	2 - 3 wks	4 wks	3 - 5 days
Visit 1	Visit 2	Visit 3	Visits 4, 5, 6	Visit 7
Withdrawal of TCAs	No treatment for cataplexy		Placebo or GHB 3g, 6g, or 9g	No Treatment for cataplexy

1. Screening period - lasted 1 day to 4 weeks. Tricyclic antidepressants (TCAs) or other drugs used to treat cataplexy were gradually withdrawn from patients on these drugs. Patients not on TCAs proceeded directly to the next study period if they met entry criteria. Patients were permitted to continue taking stable doses of stimulant medication throughout the study.
2. Washout period - lasted 5 - 28 days. This period allowed time to eliminate any clinical effects of TCAs, for rebound cataplexy (cataplexy that is spontaneous, unprovoked, and with greater frequency and severity than usual) to abate, and to train patients on the use of diary.
3. Baseline period - lasted 2 to 3 weeks. This period was an opportunity to assess the patients' cataplexy attacks and to establish a stable number of attacks. Eligibility for admission into the double-blind treatment period required patients to report an average of 3 or more complete and/or partial cataplexy attacks per week, during the last two weeks of the baseline period.
4. Double-blind treatment period - lasted 4 weeks. Eligible patients were randomly assigned to receive each night 3g, 6g, or 9g GHB or placebo in 2 divided doses. Patients returned approximately every 2 weeks during this period to assess safety and efficacy.
5. Follow-up period - a visit for final assessment 3 - 5 days after study medication was discontinued.

Approximately 104 patients (26 in each of the 4 treatment groups) were planned to be enrolled and 136 patients were actually enrolled and assigned to receive 4 weeks of treatment with study medication.

The study was initiated on February 7, 1997 and completed on February 9, 1998.

### 2.3 Main Inclusion Criteria

Patients were included in the study if they met the following criteria:

- Were 18 years of age or older;
- Had not received investigational therapy within 30 days prior to study entry;
- Had a history of excessive daytime sleepiness;
- Had a history of sudden loss of voluntary muscle control or muscle weakness (cataplexy) localizable to a specific muscle group(s) or part(s) of the body during which the patient was lucid and not experiencing an inadvertent nap or micro sleep;

- Had a valid polysomnography (PSG) within the last 5 years and a current diagnosis of narcolepsy for at least 6 months according to the following 2 items of Criteria A as established by the American Sleep Disorder Association (ASDA):
  - a) Recurrent daytime naps or lapses into sleep that occur almost daily for at least 3 months;
  - b) Sudden bilateral loss of postural muscle tone in association with intense emotion (Cataplexy).

In addition, a patient must have recorded an average of 3 or more complete and/or partial cataplexy attacks per week during the last 2 weeks of the baseline period to be eligible for entry into the randomized portion of the study.

## **2.4 Efficacy Variables**

### **2.4.1 Primary Efficacy Variables**

The primary efficacy variable for this study as defined in the protocol was the change from baseline in the total number of cataplexy attacks, which is the sum of complete and partial cataplexy attacks that occurred.

The endpoint was defined as the last 2-week period of the treatment (Visit 6) and the baseline was defined as the last 2-week period before the treatment (Visit 4).

A cataplexy attack, episode, or event was defined as a sudden loss of voluntary muscle tone usually triggered by emotions such as those associated with laughter, anger, elation, fear or surprise. These events could range from a brief experience of partial muscle weakness to an almost complete loss of muscle control lasting for several minutes and resulting in total physical collapse during which time the patient was unable to move or speak. To be classified as cataplexy for this study the patient must have been aware of the time and place during the event. The event must have been of sudden onset and localizable to a specific muscle group(s) or part of the body.

### **2.4.2 Secondary Efficacy Variables**

Secondary measures of efficacy include following variables:

- Change in the number of complete and number of partial cataplexy attacks;
- Daytime sleepiness as measured by the Epworth Sleepiness Scale and number and duration of inadvertent naps;
- Quality of nighttime sleep as measured by number of awakenings during the nights and total amount of sleep;
- Incident of hypnagogic hallucination;
- Incident of sleep paralysis;
- Change in severity of the patients' narcolepsy symptoms as measured by the Clinical Global Impression of Change

## **2.5 Statistical Method**

### **2.5.1 Primary Efficacy Analysis**

The efficacy analyses were to be done on the intent-to-treat (ITT) population. The planned analyses called for an analysis of variance (ANOVA) on the change from baseline to endpoint including in the model the factors of treatment, site, and their interaction. The interaction term was then to be removed if found to be not statistically significant. In addition, an analysis of covariance (ANCOVA) was planned for the primary efficacy variable using the baseline value as a covariate. The factor of site was to be removed from this ANCOVA model, as indicated in the statistical analysis plan. An additional analysis would examine for a possible dose response relationship.

It was stated in the Statistical Analysis Plan, documented July 1997, that the method of statistical analysis implied normal distribution assumption. If the data were not normal then the most suitable data transformation method (log, square root, etc.) was to be applied. If the assumption of normality and the data transformation did not appear to be satisfied, then appropriate non-parametric methods were to be used in the analyses.

### **2.5.2 Analysis of Secondary Efficacy Variables**

All secondary measures of efficacy except CGI-c were to be analyzed in a similar way as to the primary efficacy variable.

CGI-c was to be analyzed using Fisher's exact test and Cochran-Mantel-Haenszel (CMH) test for nonzero correlation.

### **2.5.3 Center Grouping**

The study was to be conducted in 16 centers. In the event of any site(s) fail to reach a minimum of 8 patients, these sites were to be pooled and treated as a single site for purpose of statistical analysis.

### **2.5.4 Interim Analysis**

No interim analyses were planned or performed for this study.

## **2.6 Results: Sponsor's Analysis**

### **2.6.1 Patient Disposition**

The sponsor reported that a total of 136 patients were enrolled from 16 centers. The number of patients entered by each center ranged from 1 to 21. Of the 136 patients enrolled, 16 withdrew from the study before completion. Disposition of patients is summarized in Table 2. The primary reason of withdrawal from the study was adverse

experience (10 patients). Patient withdrawal for adverse events was more frequent in the 9g GHB dose group than in the other 3 treatment groups.

Table 2. Patient disposition

Disposition	Placebo	GHB Dose			All Patients
		3g	6g	9g	
Receive study medication	34	34	33	35	136
Withdrawn from the study					
Adverse event	1	1	2	6	10
Protocol deviation		1			1
Patient request		1		1	2
Lost of follow-up			1		1
Lack of efficacy		1	1		2
Total Withdrawals	1	4	4	7	16
Completed study	33	30	29	28	120

### 2.6.2 Patient Baseline and Demographic Characteristics

Demographic characteristics of the 136 patients who received study medication are summarized in Table 3. Significant differences between groups in sex and height were noted by the sponsor. The 6g GHB group was predominantly male, while the placebo and 3g GHB groups were predominantly female. Consistent with the larger proportion of females in 3g GHB group, the mean height of this group was less than the other treatment groups. The sponsor reported that there were no difference between treatment groups in baseline blood pressures, pulse rates, respiration, or ECGs.

APPEARS THIS WAY  
ON ORIGINAL

APPEARS THIS WAY  
ON ORIGINAL

Table 3. Demographic characteristics by treatment group

Characteristic	Total	Placebo	GHB Dose			p-value <sup>1</sup>
			3g	6g	9g	
Age						0.2737
N	136	34	34	33	35	
Mean (year)	43.06	40.82	47.06	43.52	40.91	
SD	15.03	14.33	16.89	14.98	13.53	
Sex (n)						0.0027
Male	57	12	7	21	17	
Female	79	22	27	12	18	
Race (n)						0.1379
Caucasian	124	29	33	31	31	
African American	9	4	0	1	4	
Asian	1	0	0	1	0	
Other	2	1	1	0	0	
Height						0.0283
N	131	31	33	33	34	
Mean (cm)	170.91	171.97	166.7	173.1	171.9	
SD	9.53	8.18	8.78	10.39	9.64	
Weight						0.4847
N	134	34	33	33	34	
Mean (kg)	82.87	83.98	78.86	85.04	83.56	
SD	17.36	18.89	15.65	15.54	19.08	

1. p-value based on ANOVA (GLM)

The severity of narcolepsy in the patient population was assessed by documenting the frequency of symptoms that were reported in the 3 months prior to screening. The sponsor reported that the distribution of patient's symptoms was similar across the 4 treatment groups (see Table 4). Cataplexy, excessive daytime sleepiness, and inadvertent naps/sleep attacks were among the most common symptoms. When focusing on the key symptom of interest in this study, cataplexy, the sponsor reported that the 4 treatment groups were similar in the average number of cataplexy attacks per week over the last 3 months reported by the patients at the screening visit. The mean numbers of the reported cataplexy attacks for the 4 treatment groups were 15.5 for the placebo group and 12.3, 16.6, and 16.7 for the 3g, 6g, and 9g GHB group, respectively.

Table 4. Number of patients reporting symptoms in the 3 months prior to screening

Symptoms	Number (%) of patient with symptoms			
	Placebo n=34	3g GHB n=34	6g GHB n=33	9g GHB n=35
Cataplexy	34 (100)	34 (100)	32 (97)	35 (100)
Excessive daytime Sleepiness	34 (100)	34 (100)	32 (97)	34(97)
Awakening at night	27 (79)	31 (91)	27 (82)	30 (86)
Inadvertent naps/sleep attacks	32 (94)	33 (97)	31 (94)	33 (94)
Sleep paralysis	26 (76)	25 (74)	25 (76)	24 (69)
Hypnagogic hallucinations	27 (79)	29 (85)	26 (79)	26 (74)

A summary of baseline narcolepsy symptoms by treatment group is presented in Table 5. The sponsor noted that the increase in cataplexy attacks at baseline (Visit 4) was due to the discontinuation of anti-cataplexy medication that was present at screening (Visit 1). The sponsor reported that the values of total, complete, and partial cataplexy attacks did not vary in a statistically significant manner across treatment groups, though it was noted that, in general, the median number of partial cataplexy attacks was almost 5 times greater than the median number of complete cataplexy attacks.

**Table 5. Summary of baseline (Visit 4) narcolepsy symptoms by treatment group**

Type of event	Placebo (n=34)	3g (n=33)	GHB Dose 6g (n=33)      9g (n=35)		p-value Kruskal-Wallis
Total cataplexy/wk					0.7749
Mean (SD)	34.27 (46.63)	28.57 (30.53)	38.85 (55.04)	34.60 (33.92)	
Median	20.21	20.00	23.00	23.50	
Complete cataplexy/wk					0.5151
Mean (SD)	6.86 (12.37)	7.08 (8.50)	15.26 (27.53)	8.61 (14.01)	
Median	1.12	4.50	4.85	2.00	
Partial cataplexy/wk					0.7289
Mean (SD)	27.44 (42.08)	21.49 (28.30)	23.59 (29.01)	26.12 (26.14)	
Median	15.03	15.00	16.15	18.79	
Hypnagogic hallucination/day					0.9766
Mean (SD)	0.57 (0.74)	0.56 (0.68)	1.14 (3.72)	0.53 (0.70)	
Median	0.23	0.43	0.33	0.29	
Sleep paralysis/day					0.9597
Mean (SD)	0.51 (0.74)	0.42 (0.55)	0.73 (1.84)	0.41 (0.60)	
Median	0.26	0.14	0.08	0.10	
Inadvertent naps/day					0.7008
Mean (SD)	1.71 (0.96)	1.91 (1.43)	1.70 (1.12)	1.72 (1.56)	
Median	1.57	1.93	1.45	1.27	
Excessive day-sleep <sup>1</sup>					Not reported
Mean (SD)	18.47 (3.13)	17.06 (3.71)	17.26 (3.49)	16.66 (4.07)	

1. By Epworth Sleepiness Scale

### 2.6.3 Efficacy Evaluation

The sponsor reported that at the time of analysis, each of the primary and secondary efficacy variables was assessed for normality and whether a log transformation would improve the distribution. This assessment was based on using Wilk-Shapiro test for normality on the residuals from the fitted model and a plot of the residuals against the predicted response, also from the fitted model. The following measures were analyzed using the log transformation:

- Total number of cataplexy attacks
- Partial cataplexy attacks
- Complete cataplexy attacks
- Duration of inadvertent naps/sleep attacks/day
- Sleep paralysis (episodes/day)
- Hypnagogic hallucinations

- Number of awakenings

For each of these measures the value of 1 was added prior to the log transformation because 0 was possible. Therefore, the variable analyzed was  $\log(\text{endpoint}+1) - \log(\text{baseline}+1)$ . The ANCOVA model used to assess overall treatment group comparisons include treatment, site, and  $\log(\text{baseline}+1)$ . The interaction of treatment and site and treatment with  $\log(\text{baseline}+1)$  were included in the model and then removed when found to be not statistically significant. Comparisons of GHB dose to placebo were performed using least-square means with Dunnett's adjustment. The significance of the median change from baseline for each treatment group was assessed using a paired t-test.

Several measures did show a normal distribution without a log transformation. They included

- Epworth Sleepiness Scale;
- Total amount of sleep/night;
- Number of inadvertent naps/day.

For these variables, the analysis procedures were consistent with those previously described, but were based on the untransformed values.

The Clinical Global Impression of Change (CGI-c) was assessed by correlation analysis using Cochran-Mantel-Haenszel (CMH) test for Nonzero Correlation between the CGI-c score and treatment.

### **2.6.3.1 Primary Efficacy Variable: Total Number of Cataplexy Attacks**

The primary efficacy variable was the change from baseline to the endpoint in the total number of cataplexy attacks per week. The sponsor reported that there was a significant difference ( $p=0.0021$ ) among treatment groups in change from baseline to endpoint in total number of cataplexy attacks. The analysis was based on an ANCOVA model on log-transformed data including factors of treatment, site, and  $\log(\text{baseline}+1)$ .

The sponsor reported that a dose response relationship in improvement was observed across all dosage groups. The change in the total number of cataplexy attacks exceeded placebo and was in a clinically meaningful range with all GHB treatment groups, was of marginal statistical significance in the 6g GHB group ( $p=0.0529$ ) and was of unambiguous statistical significance in the 9g GHB group ( $p=0.0008$ ). In a *post hoc* analysis, there was no evidence of rebound cataplexy on withdrawal of GHB at the end of the study.

Table 6. Total number of cataplexy attacks

Dose group	Statistics	Observed		Change from baseline	p-value
		Baseline	Endpoint		
Placebo	N	33	33	33	0.5235
	Mean (SD)	35.1 (47.1)	24.0 (28.4)	-11.1 (27.7)	
	Median	20.5	16.3	-4.3	
	p-value			0.028	
3g	N	33	33	33	0.0529
	Mean (SD)	28.6 (30.5)	19.5 (27.5)	-9.1 (22.4)	
	Median	20.0	9.5	-7.0	
	p-value			0.026	
6g	N	31	31	31	0.0008
	Mean (SD)	33.8 (45.6)	24.6 (62.9)	-9.2 (27.3)	
	Median	23.0	8.0	-9.9	
	p-value			0.070	
9g	N	33	33	33	<0.001
	Mean (SD)	35.7 (34.5)	14.4 (19.3)	-21.3 (29.8)	
	Median	2.5	8.7	-16.1	
	p-value			<0.001	

The sponsor reported that placebo, 3g, 6g, and 9g GHB groups showed percent reductions in the median number of cataplexy attacks of -28%, -40%, -50%, and -55% after 2 weeks of treatment. There were no further reductions in the median number of cataplexy attacks from 2 to 4 weeks in the placebo and 6g GHB groups. The 3g and 9g GHB groups had further reductions in the median number of cataplexy attacks to -49% and -69% respectively. The overall pattern is that the greater part of the response in all groups was seen by 2 weeks with smaller changes by 4 weeks.

### 2.6.3.2 Secondary Efficacy Analysis

#### Complete cataplexy attacks

The sponsor reported that complete cataplexy attacks were much less frequent than partial cataplexy attacks. At baseline the median number of complete cataplexy attacks was 1.2, 4.5, 4.7, and 2.0 in the placebo, 3g, 6g, and 9g GHB treatment groups, respectively. At the endpoint that number changed by 0, -1.00, -1.62, and -1.62 in the placebo, 3g, 6g, and 9g GHB treatment groups, respectively. None of the decreases in the median number of complete cataplexy attacks reached statistical significance when compared to placebo.

#### Partial cataplexy attacks

The sponsor reported that from baseline to endpoint the median number of partial cataplexy attacks changed by -2.72, -3.69, -6.35, and -10.00 in the placebo, 3g, 6g, and

9g GHB treatment groups, respectively. The difference between placebo and 9g GHB treatment groups carried a p-value of 0.0009 from the ANCOVA model.

### Excessive daytime sleepiness

Excessive daytime sleepiness as assessed by Epworth Sleepiness Scale improved in all GHB treated groups, as reported by the sponsor. The overall treatment group comparison on the change from baseline to endpoint on excessive daytime sleepiness carried a p-value of 0.0006 while the comparison between 9g GHB group and placebo carried a p-value of 0.0001.

### Clinical Global Impression of Change (CGI-c)

The CGI-c was a 7-point scaled measure of improvement/deterioration based on the clinical investigator's overall impression of the change in the patient's condition. The sponsor reported that a highly significant treatment effect was noted on the CGI-c scale (p=0.0010 from overall treatment group comparison based in Cochran-Mantel-Haenzel test). The following table presents the summary of Clinical Global Impression of Change at endpoint by treatment group.

Table 7. Summary of Clinical Global Impression of Change at endpoint by treatment group

Impression	Placebo	GHB Dose		
		3g	6g	9g
Very much improved	3 (9%)	3 (10%)	5 (16%)	11 (37%)
Much improved	8 (24%)	11 (37%)	11 (35%)	13 (43%)
Minimally improved	8 (24%)	9 (30%)	9 (29%)	3 (10%)
No change	12 (35%)	6 (20%)	5 (16%)	1 (3%)
Minimally worse	2 (6%)	1 (3%)	0	2 (7%)
Much worse	0	0	0	0
Very much worse	1 (3%)	0	1 (3%)	0

The CGI-c data was also analyzed by categorizing a patient as a responder or not. A responder was defined as a patient falling into the much improved or very much improved category. A p-value of 0.0014 was reported for the overall treatment group comparison based on Fisher's exact test. A p-value of 0.0002 was reported by comparing 9g treatment with placebo group on Fisher's exact test.

### Other Secondary Measures of Efficacy

The sponsor reported that compared with placebo, a significant decrease in the number of inadvertent naps/sleep attacks was seen in both 6g and 9g GHB groups ( $p=0.0497$  and  $p=0.0122$  respectively). A significant decrease in the number of awakenings was seen in the 9g GHB group ( $p=0.0035$ ). No significant differences between treatments were seen in the change from baseline of hypnagogic hallucinations, sleep paralysis episodes, total amount of sleep, and duration of inadvertent naps/sleep attacks.

## 2.7 Reviewer's Analysis

### 2.7.1 Primary Efficacy Analysis

The primary efficacy measure is the change from baseline to endpoint in the total number of cataplexy attacks. An ANOVA model including factors of treatment, site, and treatment-by-site interaction was applied. It was found that the assumption of normality was violated ( $p=0.0017$ ), indicating that the ANOVA might not be appropriate for the analysis.

As specified in the analysis plan, a log transformation was employed to endpoint as well as to baseline. The dependent variable of  $\log(\text{endpoint}+1) - \log(\text{baseline}+1)$  was used in the analysis. After the log transformation, the normal assumption was no longer violated ( $p=0.7365$ ; Shapiro-Wilk). The treatment-by-site interaction was not statistically significant, and was removed from the model. The final model contained factors of treatment and site showed a significant treatment effect with a p-value of 0.0034. The effect of site was also significant with a p-value of 0.0321. The descriptive statistics of the change from baseline in cataplexy attacks were the same as obtained by the sponsor (Table 6) except for the 9g GHB group, where the patient #0824 was included in the reviewer's analysis. The mean, SD, and median number for the 9g GHB group obtained by this reviewer were -20.5, 29.7, and -15.6, respectively.

Note that the p-value of 0.0034 for the treatment effect obtained by this reviewer is different from the one reported by the sponsor ( $p=0.0021$ ), although both p-values reached significance level of 0.05. The difference is due to the following reasons:

- 1) The data submitted in this NDA contained patient #0824, which was not contained in the data set in the original treatment IND. The reviewer's analysis was based on this new data set of 131 patients, while the sponsor's analysis was based on the data set of 130 patients that did not include patient #0824.
- 2) The protocol and analysis plan called for an ANOVA with factors of treatment and site. An additional analysis of ANCOVA would include baseline value as a covariate, but would exclude site as a factor. However, the sponsor used an ANCOVA model that included both site and baseline value in the primary analysis.

### Dose Response Relationship

The dosage of GHB studied in this trial included 3 dose levels: 3g, 6g, and 9g. The significance level of the effect of each dose as compared to placebo was to be adjusted by Dunnett multiplicity adjustment. Among the 3 dose levels of GHB, 9g reached

significance level by Dunnett adjustment ( $p=0.0021$ ). Dose groups of 3g and 6g did not reach the significance level ( $p=0.6358$  for GHB 3g and  $p=0.0772$  for GHB 6g). The 95% confidence interval for the mean difference of the dependent variable as compared to placebo was (-0.6658, 0.2755) for GHB 3g, (-0.9196, 0.0368) for GHB 6g, and (-1.1478, -0.2134) for GHB 9g, given by the Dunnett adjustment.

### Summary Statistics by Demographic and Baseline Characteristics

Means and medians of the change in the number of cataplexy attacks were calculated for the subgroups of gender, age, and baseline total cataplexy attacks (Table 8). Median age of 42 years and median baseline cataplexy attacks of 21 were used as cutoff points. There were no substantial discrepancies between the subgroups of gender and age in the change of cataplexy numbers. Patients who had fewer cataplexy attacks at baseline appeared to have less change in the number of cataplexy attacks during the treatment as compared to patients who had more than 21 cataplexy attacks per week at baseline.

Table 8. Means and medians of the change in the number of cataplexy attacks by subgroups of sex, age, and baseline number of cataplexy attacks.

Characteristics	Change in the Number of Cataplexy Attacks			
	Placebo	3g GHB	6g GHB	9g GHB
Sex				
Male	-10.8 (0.2)	-9.5 (-5.0)	-7.0 (-10.6)	-27.7 (-16.8)
Female	-11.3 (-5.5)	-8.9 (-7.1)	-13.3 (-9.7)	-13.4 (-8.8)
Age <sup>1</sup> (years)				
≤42	-13.0 (-3.7)	-8.7 (-7.0)	-7.9 (-6.4)	-18.2 (-11.3)
>42	-8.3 (-4.5)	-9.2 (-5.1)	-10.6 (-12.3)	-23.2 (-16.8)
Baseline # of attacks <sup>2</sup>				
≤21	-1.4 (-3.2)	-4.1 (-3.7)	-7.5 (-8.3)	-5.6 (-4.3)
>21	-24.3 (-19.4)	-15.1 (-17.5)	-10.9 (-17.9)	-33.8 (-31.0)

1. Median age of 42.5 was used as a cutoff point;
2. Median number of 21 was used as a cutoff point.

### **2.7.2 Analysis of Secondary Efficacy Variables**

A total of 10 secondary efficacy measures were specified for this study. Among those 10 efficacy variables, number of complete cataplexy attacks, number of partial cataplexy attacks, daytime sleepiness as measure by Epworth sleepiness scale, and CGI-c were analyzed by this reviewer. The results of these analyses are discussed in this section. A table of summary is presented at the end of this section.

#### Number of Complete and Partial Cataplexy Attacks

The number of complete and partial cataplexy attacks were analyzed separately using the same method as to the total number of cataplexy attacks. In both analyses of complete and partial cataplexy attacks, the normal assumption for the parametric ANOVA model

did not hold and log transformations were applied. Using log (endpoint+1) - log (baseline+1) as the dependent variable, the normal assumption was no longer violated in the ANOVA model for analyses of both complete and partial cataplexy attacks.

From the ANOVA model based on the log transformed data the treatment-by-site interaction was not significant, and was removed from the analyses of complete and partial cataplexy attacks.

For the number of complete cataplexy attacks, treatment effect carried a p-value of 0.2362. A p-value of 0.0204 was obtained from the analysis of partial cataplexy attacks. In both analyses, the effect of center was significant.

Daytime Sleepiness Measured by Epworth Sleepiness Scale

Daytime sleepiness, as measured by Epworth Sleepiness Scale, was to be analyzed similarly to the primary endpoint. Original scale was used first. The interaction of treatment-by-site was found not significant. After removing the interaction term from the model, the normal assumption was no longer held by the residuals. The log transformation made no improvement in the normality of the distribution and non-parametric method of Kruskal-Wallis test was employed. A p-value of 0.0109 was obtained from the Kruskal-Wallis test for the treatment effect across 4 treatment groups.

Clinical Global Impression of Change (CGI-c)

CGI-c was to be analyzed by Fisher's exact test and Cochran-Mantel-Haenszel (CMH) test for nonzero correlation, as specified in the protocol. Due to the large size of the categorical table, the Fisher's exact test was not applicable, and CMH test for nonzero correlation was employed. A p-value of 0.322 was obtained, indicating that the hypothesis of no association between treatment groups and CGI-c values was not rejected.

Table 9. Summary of secondary efficacy results (not including CGI-c)

Variable	Change from Baseline				p-value
	Mean (Median)				
	Placebo	3g GHB	6g GHB	9g GHB	
Complete attacks	-3.0 (0)	-2.3 (-1)	-3.9 (-1.6)	-6.5 (-1.6)	0.2362
Partial attacks	-8.1 (-2.7)	-6.7 (-3.7)	-5.3 (-6.3)	-14.0 (-9.9)	0.0204
Epworth Sleepiness Scale	-1.2 (-1)	-2.5 (-1)	-2.4 (-2)	-4.5 (-3)	0.0109

Note that the p-values obtained by this reviewer in the analyses of above secondary efficacy measures are different from those reported by the sponsor due to the differences in the data set and analysis method as described in Section 2.7.1. For results of inadvertent naps and number of awakenings, the p-values reported by the sponsor were

not from comparisons across 4 treatment groups. Analyses for these two variables were not performed by this reviewer.

### **2.7.3 Center Effect**

The study was conducted in 16 centers. The number of patients enrolled in each center ranged from 1 to 21. For the purpose of statistical analysis, eight centers that had less than 8 patients were pooled together to form a larger center with 30 patients altogether, which was planned in the protocol. In the analysis of primary efficacy variable and two secondary efficacy variables (number of complete and partial cataplexy attacks), center effect was significant. Summary statistics of change in the number of cataplexy attacks within each center were explored. No evidence was found that efficacy was driven by any individual center(s).

### **2.8 Reviewer's Conclusion for Study OMC-GHB-2**

In this study 136 patients were enrolled and randomized into 3g, 6g, and 9g GHB and placebo treatment groups. The primary efficacy variable was the change from baseline in the total number of cataplexy attacks. Based on the protocol defined primary analysis, a p-value of 0.0034 was obtained by this reviewer for the overall treatment effect. Therefore, this reviewer would conclude that the study has provided sufficient evidence that GHB, given in the dose of 3g, 6g, and 9g, was efficacious in treating patients with symptoms of narcolepsy.

Although the study has shown the efficacy of GHB successfully, it was not clear which dose level would provide most benefit with limited risk. The 9g GHB was considered as the high end of tolerable dose, while the 6g GHB failed to show a significant treatment effect. The safe and effective dose has yet to be determined.

A total of 10 secondary efficacy measures were specified in the protocol. None of the differences between treatment groups reached significance level when adjusted by Bonferroni procedure. Therefore, this reviewer would conclude that the evidence of the treatment effect with regard to any of the secondary efficacy measures was not sufficient.

### **3. Specifications and Findings of Lammers' Study**

The study was carried out by Dr. G.J. Lammers, neurologist in training, together with other investigators. Dr. Lammers has prepared, in collaboration with the Dutch Narcolepsy Association, this double-blind study to investigate gammahydroxybutyrate. Polysomnography was analyzed at the \_\_\_\_\_ Clinical monitoring was done in \_\_\_\_\_

Orphan Medical analyzed and reported the Lammers trial results to the FDA in October 1998 as part of the package for approval of a Treatment IND. The Lammers data were analyzed in two different ways. The analysis using investigator-designated methodology

failed to reach the significance level for cataplexy. However, the primary analysis method proposed in the protocol was not appropriate for the crossover design. The sponsor also analyzed data employing the same methodology used in study OMC-GHB-2. Although it showed a significant treatment effect, the sponsor acknowledged that the analysis was not prospectively defined. This reviewer's analysis using a similar model showed a contrary result. The Lammers study was submitted as part of this NDA.

### **3.1 Trial Objectives**

The purpose of the trial was to investigate whether:

- Gammahydroxybutyrate, in a double-blind study, has an effect on the REM dissociation phenomena and possibly on the elevated inclination to sleep during the day;
- Gammahydroxybutyrate affects alertness during the day; and
- Gammahydroxybutyrate has a mood-improving effect.

It was noted that the second objective listed above was not investigated following the discussions among the investigators concerning the difficulties in defining and measuring the alertness during the day.

### **3.2 Study Design**

The study was a double blind, crossover study that consisted of a 1-week baseline observation period followed by 2 medication periods of 4 weeks each with a washout period of 4 weeks. During the week prior to each medication period and during each medication period, a diary was completed by each patient.

Following recruitment, the patients were entered into a baseline observation period of 1 week. Immediately afterward, patients were randomly assigned to two groups and treatment was started for a period of 4 weeks. A washout period of 4 weeks followed. Crossover took place at the end of second baseline week and the study completed after 4 weeks of treatment of the second period.

At the end of each treatment period a global therapeutic impression was assessed by patients. Co-medication was continued and unchanged throughout the study.

The dose of gammahydroxybutyrate was 30 mg/kg (0.3 ml/kg) to be taken once just before going to sleep and once 4 hours later at night. The actual dose received by patients ranged from 3.78 to 5.52 grams per day, with a mean of 4.75 grams per day.

### **3.3 Main Inclusion Criteria**

Patients were selected by Dr. Lammers from a group of patients of the Dutch Narcolepsy Association. The main inclusion criteria were as follows:

- Patients must have a combination of sleep attacks during day and at least one of the “REM dissociation phenomena” (cataplexy, hypnagogic hallucination and sleep paralysis) or;
- In case of doubt, a positive multiple sleep latency test as recorded with a 24-hour EEG was required;
- Patients must have provided written informed consent.

### **3.4 Efficacy Variables**

The protocol-defined criteria for the response of the patients were the following:

- The opinion of the patients on the benefit of the medication (global therapeutic impression);
- The opinion of the physician (global clinical impression);
- The number of cataplexy attacks;
- The number of sleeping attacks during the day;
- The feeling of sleepiness during the day (decrease of 20% on the visual analogue scale);
- MSLT improvement of the two shortest latencies with a minimum of 4 minutes in total;
- The stability of alertness during the day;
- The duration of the slow wave sleep;
- A decrease of the number of phase shifts at night.

A positive result of the primary efficacy was defined as an improvement in all of the first three criteria. The other variables were of secondary importance and were analyzed only if effect was found in the primary variables.

### **3.5 Statistical Method**

The protocol stated that the differences between placebo and gammahydroxybutyrate were to be tested by means of the Wilcoxon signed rank test.

### **3.6 Results from Sponsor’s Analysis**

#### **3.6.1 Disposition of Patients**

A total of 25 patients were included in the study. All of them completed the study. One patient (patient #13) failed to keep his diary and could not be included in the analysis. There was clinical doubt of whether patient #16 suffered from narcolepsy. This doubt was confirmed by MSLT findings. This patient was included in the ITT patient analysis.

#### **3.6.2 Patients Demographic and Baseline Characteristics**

The mean age of the patients was 40 years, ranging from 16 to 65 years. Thirteen patients were male and 12 were female.

### **3.6.3 Efficacy Results (Sponsor's Analysis)**

A total of 24 patients were included in the ITT analysis of parameters measured in the diary. Patient #13 was excluded because he failed to keep the diary of the first period during which he received gammahydroxybutyrate. Patient #25 had no baseline data for all efficacy parameters in the gammahydroxybutyrate period. The baseline data of the placebo period was used.

For each of the efficacy variables, an analysis of covariance (ANCOVA) was used employing a model for a crossover design. The model included factors of treatment order (GHB followed by placebo or the reverse), patient within treatment order, treatment group, period, and baseline value for the efficacy variable. The significance of the covariate was also examined. The usual assumption associated with the analysis of covariance of normal distribution and constant variance were examined, and if not satisfied, a log transformation of the data was considered.

#### **3.6.3.1 Analysis of Primary Efficacy Variables**

Primary efficacy variables include the first 3 parameters listed in Section 3.4. A positive result of the primary efficacy was defined as an improvement in all of the first three criteria. For the physician clinical global impression parameter, the investigator only reproduced the opinion of the patient and could not translate this information into his own opinion. Therefore, only the global therapeutic impression (GTI) of the patient has been recorded and analyzed.

##### Primary Efficacy Variable: Number of Cataplexy Attacks

The sponsor reported that during the gammahydroxybutyrate treatment period, patients had significantly fewer cataplexy attacks per day ( $p=0.002$ ). The p-value from the Shapiro-Wilk test was 0.88 and the period effect carried a p-value of 0.39. The baseline effect was found to be significant with a p-value of 0.0001.

##### Primary Efficacy Variable: Global Therapeutic Impression (GTI)

The sponsor reported that the test for the difference between GHB and placebo with regard to the GTI gave a p-value of 0.001 for the scores rated by patients on the hand written sheet at the end of the study, and a p-value of 0.021 for the scores rated by patients on the diary form directly after the treatment. Both analyses used McNemar' test. The results were presented in the following tables.

Table 10. GTI scored on the hand written sheet at the end of the study

Placebo period	GHB period		Total
	No beneficial effect	Beneficial effect	
No beneficial effect	8	15	23
Beneficial effect	1	1	2
Total	9	16	25

Table 11. GTI scored on the diary form directly after the treatment period

Placebo period	GHB period		Total
	No beneficial effect	Beneficial effect	
No beneficial effect	11	10	21
Beneficial effect	2	2	4
Total	13	12	25

### 3.6.3.2 Analysis of Secondary Efficacy Variables

The results of the analyses of covariance for secondary efficacy variables were presented in Table 12. Included in the table are results of the Shapiro-Wilk test for normality, a comparison of GHB with placebo (treatment), an assessment of period effect comparing the results of the first period with the second period (period), and the effect attributable to the covariate of baseline.

Table 12. Analysis of covariance of the change from baseline (ITT patient population)

Change at week 4	Shapiro-Wilk test	p-values of the factors in the model		
		Treatment	Period	Baseline
Severity of daytime sleepiness	Not normal P=0.0005	0.029	0.61	0.0002
Daytime sleep attacks	Normal P=0.93	0.0008	0.12	0.050

The distribution of the residuals of the severity of daytime sleepiness was not normally distributed with and without log transformation, so non-parametric method of Wilcoxon paired sample test was used.

### 3.7 Reviewer's Analysis

In this study neither the primary efficacy variables nor the primary efficacy analysis were defined specifically and appropriately. For example, it was not stated in the protocol whether the number of cataplexy attacks was to be measured as 4-week average or last week's average, whether the data during the treatment period or the change from baseline

should be used as the endpoint. For the GTI data, there were scores on the hand written sheet at the end of the study and scores on the diary form directly after the treatment period.

The primary analysis for the number of cataplexy attacks specified in the protocol was the Wilcoxon signed rank test. However, this study used a crossover design, in which the possible period effect needed to be taken into account in the analysis. The period effect could not be examined or incorporated in the Wilcoxon test. Therefore, the protocol specified Wilcoxon test was not considered as an appropriate analysis.

The commonly used statistical method for data from a 2 by 2 crossover design is the analysis of variance that take into account of the effects of period, sequence, and variations of subjects within sequence. However, such model is subject to specifications of dependent variable as well as factors and covariates.

The sponsor used change from baseline to week 4 in the number of cataplexy attacks as the dependent variable, and baseline value of the cataplexy attacks as a covariate. A p-value of 0.002 was reported to represent the treatment effect.

This reviewer used two models of the same type in analyzing the data of cataplexy attacks. One model used difference from the baseline in the number of cataplexy attacks as the dependent variable without any covariate, and a p-value of 0.1789 was obtained for the treatment effect. The other model used log transformation of the data, and  $\log(\text{endpoint}+1) - \log(\text{baseline}+1)$  was used as the dependent variable without covariates. The p-value from the latter model was 0.1233. For both models, the normal assumption was not violated, although log transformation made improvement in the distribution of the data to be more close to normal. Both models are considered to be appropriate, as well as the sponsor's model, for the design. The difference between the models used by this reviewer and the one used by the sponsor is that baseline was used as a covariate in the sponsor's model, but not in the reviewer's models. However, baseline value was taken into account in the reviewer's models because the difference from the baseline to endpoint was used as dependent variable.

In addition to the two models used by this reviewer and the one used by the sponsor, there are many other ways to analyze the data. For example, one may use Week 4 average or average of the 4 weeks instead of change from baseline as the endpoint. One may also use repeated measure analysis with or without baseline as a covariate. The model could also include demographic characteristics, such as gender and age, as factors. The models used by this reviewer are just examples to show that without an appropriately prospectively defined method, it is just not possible to give a definitive conclusion for this study.

The same complication occurred to GTI, which is also designated as one of the primary efficacy parameters. It was not mentioned that GTI would be scored in two different ways. The analysis of McNemar's test was not prospectively defined. In addition, the two categories of beneficial or not were not defined in protocol.

In summary, no conclusions on efficacy can be reached for this study based on scientific grounds. A table of summary for the mean values of cataplexy attacks during baseline and treatment period is presented in Table 13.

Table 13. Mean values of number of cataplexy attacks per day at baseline and endpoint by period and sequence order.

	Sequence 1: GHB/Placebo			Sequence 2: Placebo/GHB		
	Baseline	Endpoint	Difference	Baseline	Endpoint	Difference
Period 1	1.35	0.56	-0.79	1.17	0.75	-.42
Period 2	1.17	1.19	0.02	0.78	0.45	-.33

#### 4. Specifications and Findings of Scrima Study

This study was conducted under IND \_\_\_\_\_, FDA Office of Orphan Product Development Grant #FD-R-000115-01, at the Sleep Disorders and Research Center of the University of Arkansas for Medical Sciences (UAMS) under the principal investigator Dr. Lawrence Scrima.

In 1988, \_\_\_\_\_ assumed sponsorship for IND patient studies using GHB for the treatment of cataplexy/narcolepsy, incorporating data from this trial.

On November 1, 1991, \_\_\_\_\_ was contracted by \_\_\_\_\_ for preparation of case report forms, data analysis, and report writing. \_\_\_\_\_ performed data entry and data management services.

In 1996, Orphan Medical, Inc., began IND patients trial related to GHB in the treatment of narcolepsy, and the trial that is subject of this report is being incorporated in the NDA submission.

#### 4.1 Objectives

The objectives of the study were as follows:

- To evaluate as primary variables average number of cataplexy attacks and objective daytime sleepiness using the sleepiness index determined by the Multiple Sleep Latency Test (MSLT) in narcolepsy patients during treatment with GHB as compared to placebo and baseline;
- To evaluate as secondary variables average number of sleep attacks per day, average number of awakenings per night, need for methylphenidate, feeling on awakening, mood in the morning and evening, sleep patterns identified on the polysomnogram (PSG), and average number of REM onsets determined by the MSLT during treatment with GHB as compared to placebo and baseline.

## 4.2 Study Design

This single-center study utilized a randomized, double blind, two-way, crossover design, balanced for sequence group and gender. The study consisted of a 14-day baseline period, an initial 29-day treatment period, a 6-day washout period, a subsequent 29-day treatment period, and a washout/follow-up period of at least 5 days. On each day of the two treatment periods, patients were to take 25 mg/kg of the randomly assigned study drug, GHB or placebo, orally at bedtime and approximately 3 hours later for a total GHB dose of 50 mg/kg.

Twenty patients were to be enrolled in the study, 10 males and 10 females. All patients were to be selected by the principal investigator from patients diagnosed at the UAMS Sleep Disorder's Center, where the study was to be conducted.

### Method of Assigning Patients to Treatment Groups

On the last night of the baseline period (Day 14), patients were to be randomly assigned by the UAMS pharmacy to one of the two treatment groups: Treatment Group A (GHB:PLC) or Treatment Group B (PLC:GHB). Each group was to be composed of 5 males and 5 females. In Treatment Group A, patients were to receive GHB in the first treatment period and placebo in the second; in Treatment Group B, patients were to receive placebo first and GHB second.

All research staff and patients participating in the study were to be blinded to study medication assignment, which was known only to the UAMS Pharmacy.

## 4.3 Main Inclusion Criteria

Patients meeting all of the following criteria were to be considered for admission to the study:

- Age of 18 – 65 years male or female;
- A history of narcolepsy and cataplexy diagnosed by an accredited clinical polysomnographer;
- Rapid eye movement (REM) onsets  $\geq 2$  on the diagnostic MSLT;
- A sleepiness index  $\geq 75$  on the diagnostic MSLT;
- A minimum of 10 cataplexy attacks subjectively reported on a daily log during a 14-day baseline period

## 4.4 Concomitant Medication

Patients were to be free of anticataplexy medication, such as tricyclic antidepressants, 15 days prior to baseline PSG and MSLT and to refrain from using diuretics, alcohol, sleeping pills, or other CNS depressants throughout the duration of the study. Methylphenidate (up to 30 mg daily) was permitted.

## **4.5 Efficacy Variables**

Case report form to capture efficacy and safety data for this study was retrospectively developed from patient data transcribed from source documents by \_\_\_\_\_

Efficacy data during the study were to be collected from patients' descriptive comments on a daytime questionnaire and nightly sleep log. The PSG was to be taken on the last night of baseline and on the first and last night of each treatment period. The MSLT was to be taken on the day following each PSG.

### **4.5.1 Primary Efficacy Parameter**

The primary measures of efficacy were (1) number of daily cataplexy attacks and (2) objective daytime sleepiness as measured by the sleepiness index determined from MSLT.

### **4.5.2 Secondary Efficacy Variables**

Secondary measures of efficacy include (1) total sleep time; (2) number of arousal from sleep; (3) REM sleep; (4) delta sleep; (5) hypokalemia; (6) EKG arrhythmias; (7) obstructive sleep apnea; (8) arousals due to nocturnal myoclonus; (9) mood; (10) cognitive and motor skill.

## **4.6 Statistical Method**

### **4.6.1 Analysis of Primary Efficacy Variables**

The original statistical analysis plan was submitted for FDA's review on December 4, 1986. The analysis called for a repeated measure ANOVA where the frequency of cataplexy attacks were to be measured as means of 1) two-week baseline period, 2) last 2-week GHB treatment, and 3) last 2-week placebo treatment. The final statistical analysis plan was submitted to the FDA on March 6, 1987 in response to FDA statistical reviewer's requesting for clarification of the statistical model. The statistical method proposed were as follows:

GHB efficacy for narcolepsy was to be evaluated with a 1\*2 mixed design, 1 between patient factor, treatment order, and 2 within patient factors, substance (GHB versus placebo) and time (1<sup>st</sup> week versus 4<sup>th</sup> week for patient measures and 1<sup>st</sup> day versus 28<sup>th</sup> day for sleep study measures). Effect of withdrawal from GHB was to be evaluated with a separate paired t-test contrasting the washout following GHB versus the washout week following placebo, thereby reflecting potential longer-term carry-over effects in the order \* substance interaction. All dependent measures were to be in the form of change scores from their values during the baseline period.

### **4.6.2 Analysis of Secondary Efficacy Variables**

Multi-factorial repeated-measure ANOVA was to be run on all of the dependent measures obtained in the sleep lab except EKG arrhythmias, which was to be examined separately. Subject measures from patients daily log were to be averaged into weekly means and repeated-measure ANOVA was to be performed on these dependent measures.

**4.6.3 Interim Analysis**

No interim analysis was performed during the trial.

**4.7 Results: Sponsor’s Analysis**

**4.7.1 Patient Disposition**

A total of twenty patients, 10 males and 10 females, were enrolled in the study and all of them completed study. No records were available regarding how many patients were screened and were not selected.

**4.7.2 Protocol Violation**

Patient #18 was 16 years of age when she began the study, which was a violation of the inclusion criteria. Patient #1, #8 and #13 had a sleepiness index of 64, 53, and 46, respectively. It was required that the index be  $\geq 75$ . Patient #17 continued taking propranolol 40 mg daily for hypertension throughout the study. The sponsor stated that this medication has been used for treating narcoleptics at doses of 80-480 mg/day.

**4.7.3 Patient Baseline and Demographic Characteristic**

All patients enrolled in the study were included in the efficacy analyses except as noted for individual variables. The following table summarizes patient demographics at baseline.

Table 14. Patient Demographic and Baseline Characteristics by Treatment

Characteristics	GHB-PLC			PLC-GHB			Combined		
	N	Mean	SD	N	Mean	SD	N	Mean	SD
Age (yrs)	10	53.2	9.0	10	42.8	14.9	20	48.0	13.1
Weight (lbs)	10	199.8	22.9	10	165.9	27.2	20	182.9	30.0
Age at Diagnose (yrs)	10	28.1	7.8	9	25.2	14.7	19	26.7	11.3
Cata attacks (per day)	9	3.5	2.9	10	3.0	2.6	19	3.2	2.7

All patients were white except one, a black male. The average age was 48 years (ranged 16 to 64) for females and 50 (ranged 21 to 64) for males. The average age at diagnosis was 26.7 years.

The sponsor reported that there was no significant sequence group effect for age (p=0.089). However, GHB:PLC group (Group A) was somewhat older than the PLC:GHB group (Group B) (mean age 53.2 years vs. 42.8 years). Patients in treatment Group A were also significantly heavier (p=0.008) than those in treatment Group B (mean 199.9 lb. Vs. 165.9 lb.). No other statistically significant pre-study differences were identified.

**4.7.4 Efficacy Results (Sponsor’s Analysis)**

**4.7.4.1 Primary Efficacy Analysis**

The primary efficacy variables of interest were the number of cataplexy attacks per day and the sleepiness index as determined by the MSLT during the GHB treatment as compared to placebo and baseline.

Number of Cataplexy Attacks

Eighteen patients were included in the analysis. Patient #6 was excluded because there were no proximal baseline data. Patient #13 was excluded because there were no diary data collected during the study. Patient #11 (Treatment Group B) had no pre-treatment entries for cataplexy recorded in her diary. However, the technician noted on the diary that the patient was having frequent cataplexy attacks and recorded 10 attacks per day throughout the baseline period.

During each treatment week, the 19 patients evaluated recorded their daily number of attacks for at least 5 days. Thus the computation of average number of cataplexy attacks per day for each week was always based on five or more entries.

The following table summarizes the mean number of cataplexy attacks per day during treatment periods and baseline.

Table 15. Mean number of cataplexy attacks per day

Treatment Group	Pre-treatment	Mean Number of Cataplexy Attacks Per Day				Overall (SE)	Baseline to Endpoint
		Week 1 (SE)	Week 2 (SE)	Week 3 (SE)	Week 4 (SE)		
GHB	2.9 (0.5)	1.4 (0.2)	1.4 (0.2)	0.9 (0.2)	0.9 (0.2)	1.2 (0.2)	2.9 to 1.2 (p=0.007)
Placebo		1.5 (0.2)	2.0 (0.3)	2.1 (0.4)	1.9 (0.3)	1.9 (0.3)	2.9 to 1.9 (p=0.117)
p-value between treatments	---	n.s.	n.s.	0.005	0.004	0.013	---

n.s. – not significant

The sponsor reported that during active treatment periods over 4 weeks, a mean of 1.2 cataplexy attacks per day was reported by patients during the GHB treatment compared

to 1.9 during the placebo treatment, representing a mean decrease from baseline of 1.6 during GHB treatment ( $p=0.007$ ) and of 1.0 during placebo treatment ( $p=0.117$ ).

It was also reported that by week 4, treatment with GHB was superior to placebo for 84% (16/19) of patients, with a mean of 0.9 cataplexy events per day after treatment with GHB compared to 1.9 per day after treatment with placebo.

There were also significantly fewer ( $p=0.013$ ) cataplexy attacks per day during GHB treatment overall compared to placebo. The interaction of treatment with week was not statistically significant ( $p=0.071$ ). However, the data do suggest an interaction, i.e., there was very little difference between treatments at week 1 ( $p=0.735$ ; GHB 1.4, PLC 1.5) and a greater difference at week 2 ( $p=0.073$ ; GHB 1.4, PLC 2.0). At week 3 and 4, significant differences were detected ( $p=0.005$ ; GHB 0.9, PLC 2.1; and  $p=0.004$ ; GHB 0.9, PLC 1.9, respectively). No other significant main effects or interactions were identified, in particular sequence group ( $p=0.775$ ) or treatment x sequence group interaction ( $p=0.713$ ). No evidence of carryover effect was detected.

The sponsor reported that during washout, patients previously taking GHB had an average of 1.0 cataplexy attacks per day compared to 1.7 for patients taking placebo. The mean change from baseline was -1.6 attacks per day during washout from GHB, and -1.4 attacks per day during washout from placebo. There was a significant increase in attacks for both sequence groups from day 1 to day 5 ( $p=0.040$ ). By day 5, 25% of GHB patients and 50% of placebo patients reported at least as many cataplexy attacks as at baseline. The sponsor stated that both sequence groups appeared to approach their baseline levels by day 5, and there was no evidence that the washout was too short to eliminate the effects of previous treatment.

The sponsor reported that there was no statistically significant differences between the first 5 days of follow-up and washout ( $p=0.355$ ) nor for the interaction with sequence group ( $p=0.147$ ).

#### Objective Daytime Sleepiness – MSLT Sleepiness Index

Prior to treatment period 1, the mean MSLT sleepiness index at pre-study was 88.5. No statistically significant differences were detected for baseline comparability between sequence groups, gender or their interaction.

No significant treatment effects ( $p=0.085$ ), day effects ( $p=0.057$ ), or interaction of sequence group x gender ( $p=0.081$ ) were detected overall, with mean sleepiness index during GHB treatment (87.2) less than placebo (90.3). There were no statistically significant effects by day.

#### **4.7.4.2 Analysis of Secondary Efficacy Variables**

The sponsor reported that not all of the secondary measures of efficacy outlined in the protocol were analyzed for this report.

The sponsor reported that the mean number of sleep attacks per day during the 4 weeks of treatment decreased significantly from baseline for both GHB and placebo groups, but differences between treatments were not significant. No patient was free of sleep attacks under either regimen.

There was no significant difference compared to baseline in the mean number of subjective awakenings at night for either GHB or placebo, but significantly fewer awakenings during GHB treatment versus placebo (p=0.042).

Mood in the evening was significantly improved for GHB compared to baseline, but not for placebo versus baseline or for GHB versus placebo.

There were no significant differences between GHB and placebo in amount of methylphenidate taken, how patients felt upon awakening, or average morning mood. Results from analyses of secondary efficacy parameters are summarized in Table 16.

There were no significant differences between washout values versus baseline for these secondary variables except for sleep attacks per day. For sleep attacks per day, the washout data versus baseline data was significant for both GHB and placebo.

Table 16. Summary of Secondary Efficacy Variables from Diaries

Variable	Baseline (mean)	Treatment	Overall 4 weeks of Trt	p-value versus baseline	p-value between Trt
Sleep attacks/day	2.8	PLC	2.1	0.007	0.530
		GHB	1.9	0.002	
Awakenings at night	3.0	PLC	3.7	0.095	0.042
		GHB	2.4	0.091	
Methylphenidate (mg/day)	12.5	PLC	14.8	0.094	0.792
		GHB	14.4	0.690	
Patient felt upon awakening	2.8	PLC	2.7	0.472	0.219
		GHB	2.5	0.075	
Mood in the morning	2.0	PLC	2.2	0.867	0.142
		GHB	2.8	0.348	
Mood in the evening	0.8	PLC	1.5	0.253	0.210
		GHB	2.2	0.007	

## 4.8 Reviewer's Analysis

### 4.8.1 Analysis of Primary Efficacy Variables

As described in Section 4.6, the final statistical plan was submitted to FDA on March 6, 1987, in which the analysis model for the primary efficacy variable was detailed. It was

stated that all dependent measures were to be in the form of change scores from their values during the baseline period. A repeated measure ANOVA using 1st week vs. 4<sup>th</sup> week measures (1<sup>st</sup> day vs. 28<sup>th</sup> day for sleeping study measures) was designated. Sex was NOT to be included as a factor.

However, none of the p-values displayed in Table 15 represented results from the protocol specified analysis method. A SAS program was submitted to FDA containing the primary analysis, in which numbers of cataplexy attacks at week 1 **through** week 4 were used as dependent variables and sex was included as a factor. Measures of change from baseline were not used.

This reviewer evaluated primary efficacy parameter using the model specified by the final version of the statistical analysis plan.

Eighteen patients were included in the analysis of cataplexy attacks. Patient #6 did not have baseline value and patient #13 did not have measurement during treatment period. These two patients were excluded from the analysis. During the first period of treatment, a mean decrease in the number of cataplexy attacks of 1.8 and 1.1 occurred at week 4 for patients receiving GHB and placebo, respectively. During the second period, a mean decrease of 2.1 and 0.9 was observed at week 4 for patients receiving GHB and placebo, respectively. A p-value of 0.0372 was obtained for the treatment effect from the repeated measure ANOVA. No period effect was detected from the analysis. The residual of the model has been tested by Wilk-Shapiro method, and it seemed that normal assumption held pretty well as indicated by a p-value of 0.8894.

MSLT was taken at baseline and the 1<sup>st</sup> and 28<sup>th</sup> day of each treatment period. No patients were excluded from the MSLT sleepiness index data, which was analyzed similarly as to the number of cataplexy attacks. During the first period, a mean decrease of 0.5 and 4.8 was observed for patients receiving GHB and placebo, respectively. During the second period, a mean decrease of 0.2 in the index was observed in the GHB group, and a mean increase of 2.5 was observed in the placebo group (see Table 17). The difference between the treatment groups failed to reach the significance level, as the p-value obtained from the repeated ANOVA was 0.0688.

Table 17. Primary efficacy results (Mean (SD)) by period and treatment

	Sequence 1: GHB-Placebo			Sequence 2: Placebo-GHB		
	Baseline	Change at week 1	Change at week 4	Baseline	Change at week 1	Change at week 4
Number of cataplexy attacks						
Period 1	2.6 (1.6)	-0.8 (1.7)	-1.8 (1.6)	3.0 (2.6)	-1.4 (2.3)	-1.1 (2.6)
Period 2	2.6 (1.6)	-1.3 (1.8)	-0.9 (1.9)	3.0 (2.6)	-1.9 (2.8)	-2.0 (2.4)
MSLT sleepiness index		Change at Day 1	Change on Day 28	Baseline	Change on Day 1	Change on Day 28
Period 1	89.5 (15.7)	1.0 (9.0)	-0.5 (11.7)	87.4 (15.9)	-0.7 (7.5)	-4.8 (14.6)
Period 2	89.5 (15.7)	2.1 (14.0)	2.5 (11.1)	87.4 (15.9)	2.7 (7.8)	-0.2 (10.3)

The sponsor reported that patient #17 continued taking propranolol 40 mg daily for hypertension throughout the study and propranolol has been used for treating narcoleptics at doses of 80-480 mg/day. This patient received GHB first followed by placebo. The observed values of the primary efficacy variables for this patient were presented in the following table.

Table 18. Observed values of patient # 17

	Baseline	Period 1: GHB		Period 2: Placebo	
		Change at week 1	Change at week 4	Change at week 1	Change at week 4
Cataplexy	5.4	-4.26	-5.28	-4.90	-3.11
MSLT	96	-13	-17	-18	-6

Due to the fact that the degree of decrease for both cataplexy attacks and MSLT sleepiness index for this patient was far larger than the average of the others, this reviewer re-analyzed the data by excluding patient #17. The obtained p-value was 0.0705 for the number of cataplexy attacks and was 0.0897 for the MSLT sleep index. The following table shows the corresponding results as to Table 17 by excluding patient #17.

Table 19. Primary efficacy results (Mean (SD)) by period and treatment

	GHB			Placebo		
	Baseline	Change at week 1	Change at week 4	Baseline	Change at week 1	Change at week 4
Number of cataplexy attacks						
Period 1	2.2 (1.2)	-0.3 (0.9)	-1.2 (0.8)	3.0 (2.6)	-1.4 (2.3)	-1.1 (2.6)
Period 2	3.0 (2.6)	-1.9 (2.8)	-2.0 (2.4)	2.2 (1.2)	-0.8 (1.2)	-0.5 (1.8)
MSLT sleepiness index		Change at Day 1	Change on Day 28	Baseline	Change on Day 1	Change on Day 28
Period 1	88.8 (16.4)	2.6 (8.0)	1.3 (10.8)	87.4 (15.9)	-0.7 (7.5)	-4.8 (14.6)
Period 2	87.4 (15.9)	2.7 (7.8)	-0.2 (10.3)	89.5 (15.7)	4.3 (12.8)	3.4 (11.3)

#### 4.8.2 Analysis of Secondary Efficacy Variables

There were at least 10 secondary efficacy measures listed in the protocol. None of the efficacy results as reported by the sponsor reached the significance level when adjusted by Bonferroni procedure.

No analyses were performed by this reviewer for secondary efficacy variables.

#### 4.9 Reviewer's Efficacy Conclusion for Scrima Study

In this study of crossover design, two primary efficacy variables were designated. The difference between the treatment groups in the number of cataplexy attacks reached

significance level of 0.05 without multiple testing adjustment, while the trial failed to show significant difference between the treatment groups in MSLT sleepiness index.

The protocol did not address the issue of multiple testing for the two primary efficacy parameters. It is difficult to determine which adjustment method is appropriate after the break of blind. Any conclusion of the successfulness of the study based on a retrospectively defined adjustment is not statistically sound, unless the most conservative method is used. The study parameters failed to reach the significance level, should the conservative method of Bonferroni adjustment be applied. Therefore, this reviewer would not conclude that the evidence showed in the study is sufficient for claiming the effectiveness of the treatment on either of the primary efficacy variables.

## **5. Specifications and Findings of Study SXB-021**

### **5.1 Study Objectives**

The primary objective of this trial was to provide evidence for the long-term efficacy of Xyrem based on the return of cataplexy symptoms upon cessation of a minimum of 6 months of open-label treatment with Xyrem. The secondary objective of the trial was to evaluate the safety of Xyrem.

### **5.2 Study Design**

SXB-021 was designed as a randomized, double blind, placebo-controlled, multi-center trial to assess the long-term efficacy of orally administered Xyrem, compared to placebo, for the treatment of narcolepsy. Patients were drawn from a pool of current patients in study OMC-SXB-7 (the open-label extension to OMC-SXB-3 and OMC-SXB-6).

The trial consisted of 3 phases (4 visits): a 3 - 5 day screening phase, a 2-week lead-in phase, and a 2-week double-blind phase.

Phase I (screening) started at visit 1 and it lasted 3 to 5 days. During this phase, patients were to continue Xyrem at the same dosage they were taking in OMC-SXB-7. Patients were evaluated for inclusion and, if eligible, randomized to one of the two treatment groups of GHB or placebo. Phase I ended at visit 2.

In Phase II (lead-in), patients received single blind Xyrem at previously established dose for 2 weeks. Phase II ended at visit 3.

During Phase III (double blind), half of the patients, who were randomized to GHB, were to continue on Xyrem at the established dose. The other half of the patients, who were randomized to placebo, were to receive placebo for an equivalent volume of their established dose of Xyrem. Phase III lasted 2 weeks and ended at visit 4.

Patient dosage of Xyrem was determined by the stable dosage established in the OMC-SXB-7 trial (3.0, 4.5, 6.0, 7.5, or 9.0 g/d). The total dosage was divided into 2 equal

doses and added to 2 oz (4 tablespoons) of water administered orally. The first dose was taken at bedtime, and the second dose was taken 2.5 to 4 hours later.

The randomization was performed in a centralized manner. The randomization code was developed to account for the current Xyrem dosage for patients in the 4.5, 6.0, 7.5, and 9.0 g/d groups. Only two patients were taking Xyrem 3.0 g/d at screening, and they were included in the 4.5 g/d randomization scheme, although they continued to take 3.0 g/d.

Enrollment of up to 80 patients was planned for this trial. Fifty-six patients were enrolled and 55 patients were actually treated; all completed the trial. The trial was conducted in 14 sites.

### **5.3 Main Inclusion Criteria**

Patients were included in the trial if they met the following criteria:

- Were 16 years of age or older;
- Had at least 5 cataplexy attacks per week prior to receiving any treatment;
- Had been treated continuously for the symptoms of narcolepsy with GHB for a period of 6 months to 3.5 years. The patients must have been previously enrolled in Orphan Medical clinical trials OMC-GHB-3 or OMC-GHB-6.

### **5.4 Concomitant Medication**

Patients were not allowed to be on TCAs, SSRIs, or any other anti-cataplexy medications other than Xyrem for any reason in the 30 days prior to Visit 1. Patients were allowed to continue stable dosages of stimulants for the treatment of daytime sleepiness associated with narcolepsy. Patients were advised not to consume any alcoholic beverages during the entire course of the trial.

### **5.5 Efficacy Variable**

#### Primary Efficacy Variable

The primary efficacy variable for this trial was the change in the number of cataplexy attacks from baseline (2-week lead-in, single-blind treatment phase) to endpoint (2-week double-blind treatment phase).

The number of cataplexy attacks was taken from each patient's diary, which was completed by the patient each night before bedtime during phase II and phase III of the trial.

A cataplexy attack was defined as a sudden bilateral loss of voluntary muscle tone. To be classified as a cataplexy attack for this trial, the event must have been of sudden onset and localized to a specific group(s) of part of the body, and the patient must have been aware of the time and the place during the event (i.e., not a sleep attack or microsleep).

## Secondary Efficacy Variables

No secondary efficacy variables were proposed for this trial.

## **5.6 Statistical Methods**

### Primary Efficacy Analysis

The primary efficacy variable, the change from baseline in the number of cataplexy attacks, was to be analyzed using a non-parametric analysis of covariance (ANCOVA). The corresponding ranks of the change in the number of cataplexy attacks and baseline number of attacks were to be used to replace the original numbers. The model was to include ranks of baseline cataplexy attacks, treatment group, and baseline-by-treatment interaction. The interaction term was to be removed if it was not statistically significant at 0.10 level.

The primary efficacy analysis was to be performed using the intent-to-treat patient population, which include patients who received 1 or more doses of double-blind trial medication, and had baseline and post-baseline cataplexy measurements.

The number of cataplexy attacks was based on a 14-day diary in phase II and a 14-day diary in phase III. If fewer or greater than 14 days of data were available for phase II or phase III, the average number of cataplexy attacks per day was determined based on the number of days the cataplexy data were recorded, and that daily average was multiplied by 14 to determine the number of cataplexy attacks for that phase.

### Sample Size Determination

The sample size for this trial was determined using the data from the trial OMC-GHB-2 (see Section 2 for details for study OMC-GHB-2). A standard deviation of 0.30 based on a log transformation of data from OMC-GHB-2 and a desired power of 80% were used in calculating the sample size for this trial. It should be noted that the primary analysis for this trial was to use rank transformation instead of log transformation. Therefore, the sample size estimated for this trial may not be accurate as desired.

## **5.7 Results: Sponsor's Analysis**

### **5.7.1 Disposition of Patients**

The trial was conducted in 14 sites. The number of patients enrolled in each site ranged from 1 to 7 patients. A total of fifty-six patients were screened and randomized. One patient failed screening due to concomitant use of an SSRI and was never treated. All remaining 55 patients were treated and completed the trial. The number of days the study medication actually administered ranged from 11 days to 18 days during phase III. The blind was broken on one patient after completion of the trial due to SAE (the patient complete trial on 7/28/00 and the SAE occurred on 8/1/00).

## 5.7.2 Demographic and Other Baseline Characteristics

The sponsor reported that although there were some differences between the treatment groups in sex, race, and number of cataplexy attacks during the 2-week baseline period, none of the differences were statistically significant.

Overall, patients had a mean age of 47.7 +/- 16.66 years, a mean body weight of 80.5 +/- 20.09 kg, and a mean height of 170.1 +/- 10.25 cm. Forty-two percent (42%) of patients were men and 58% were women. The majority of the patients were Caucasian (95%), with 2 African-American patients (2%) and 1 Hispanic patient (2%). Overall, patients had been on Xyrem 7 to 44 months, and had a historical mean number of cataplexy attacks of 12.6 during the 2-week baseline.

A list of patient demographic and current dosage at screening by treatment group is presented in the following table.

Table 20. Demographic and Baseline Characteristics by Treatment Group

Characteristics	Total (n=55)	Treatment Group		P-value
		Xyrem (n=26)	Placebo (n=29)	
Age (years)				
Mean (SD)	47.7 (16.66)	47.9 (17.06)	47.6 (16.60)	0.955
Range	16.3 - 82.6	19.1 - 82.6	16.3 - 70.0	
Sex (n, %)				
Male	23 (42%)	8 (31%)	15 (52%)	0.172
Female	32 (58%)	18 (69%)	14 (48%)	
Time on Xyrem (months)				
Mean (SD)	21.22 (12.28)	23.27 (12.36)	19.38 (12.13)	
Range	7 - 44	8 - 38	7 - 44	ND
Cataplexy attacks (2-wk baseline)				
Mean (SD)	12.6 (31.75)	9.0 (19.25)	15.7 (39.88)	0.436
Median	3.0	1.9	4.0	
Dosage of Xyrem at screening (n, %)				
3.0 g/d	2 (4%)	1 (4%)	1 (3%)	ND
4.5 g/d	9 (16%)	4 (15%)	5 (17%)	
6.0 g/d	15 (27%)	7 (27%)	8 (28%)	
7.5 g/d	15 (27%)	7 (27%)	8 (28%)	
9.0 g/d	14 (25%)	7 (27%)	7 (24%)	

ND = Not determined

## 5.7.3 Efficacy Results (Sponsor's Analysis)

The sponsor reported that there was no change in the number of cataplexy attacks from baseline to endpoint in the Xyrem group, while cataplexy attacks increased by a median

of 21.0 in the placebo group during withdrawal. This difference was statistically significant ( $p < 0.001$ ) when analyzed by an ANCOVA model containing rank baseline, treatment group, and baseline-by-treatment interaction. The following table presents the change in the number of cataplexy attacks.

Table 21. Number of Cataplexy Attacks (Per 2 Weeks) by Treatment Group

	Xyrem (n=26)			Placebo (n=29)		
	Baseline	Double blind	Change	Baseline	Double blind	Change
Mean (SD)	9.0 (19.3)	12.6 (30.3)	3.6 (20.7)	15.7 (39.9)	50.4 (81.1)	34.6 (55.7)
Median	1.9	1.1	0.0	4.0	21.0	21.0
Minimum	_____					
Maximum	_____					

### 5.8 Reviewer's Analysis

This study was to investigate withdrawal effect after long-term treatment of GHB. Patients enrolled in this study had been in stable dosage of GHB at 3.0, 4.5, 6.0, 7.5, or 9.0 g/d during the trial with dosage established from the studies OMC-GHB-3 or OMC-GHB-6. During the protocol stage, the agency communicated with the sponsor regarding the effect of different dosage on the primary efficacy endpoint, and a randomization scheme that accounting for dosage was adopted by the sponsor based on FDA's recommendations. Therefore, no dosage unbalance was expected.

There were substantial discussions between the sponsor and the agency concerning the design of the trial. The agency has raised several issues and provided many suggestions to the sponsor. One of the issues raised by the agency is the issue of normality assumption for the proposed ANCOVA model, as the data from the Lammers study is quite skewed. The issue was resolved when a non-parametric model using rank transformed data was specified to replace the original proposed model. A SAS program containing the code for the final model was submitted to the agency dated April 13, 2000 (serial # 085).

The reviewer's analysis for the only efficacy variable, change for the number of cataplexy attacks, was based on the specific model provided by the sponsor in the form of SAS source code.

Change in the number of cataplexy attacks from baseline to endpoint was the only efficacy variable proposed for this trial. The baseline was defined as the 2-week period of phase II and the endpoint was defined as the 2-week period of phase III. For each of phase II and phase III, an average number of cataplexy attacks per day was calculated for each patient based on the number of days the diary data recorded. This average number was then multiplied by 14 to represent the number of cataplexy attacks for the phase.

The average numbers of cataplexy attacks for baseline and endpoint were transformed to their corresponding ranks. An ANCOVA model specified by the sponsor included treatment, baseline, and treatment-by-baseline interaction. The interaction term was to be removed if not significant at 0.10 level.

The analysis showed that the interaction of baseline by treatment was not significant ( $p=0.8292$ ). From the final analysis model after removing the interaction term, a p-value of 0.0001 representing the treatment effect was obtained. The effect of baseline was not significant ( $p=0.8264$ ), which could be interpreted that the effect of baseline had already been taken care of by using the change from baseline as the dependent variable.

Summary statistics for the number of cataplexy attacks at endpoint and in change from baseline were calculated by breaking down the patient population into subgroups by gender and age. The median age of the whole patient population was 51 years old, and it was used as the cut-off point for the age groups. No substantial discrepancies between the groups in gender or in age were found in the frequency of cataplexy attacks at the endpoint and in change from baseline. Table 22 presents the summary statistics for the primary efficacy variable.

Table 22. Summary Statistics of Cataplexy Attack Frequencies

	Placebo (n=29)			GHB (n=26)		
	Mean	SD	Median	Mean	SD	Median
All Patients						
Baseline	15.75	39.88	4.00	8.99	19.25	1.88
Endpoint	50.36	81.09	21.00	12.57	30.34	1.08
Change from Baseline	34.61	55.72	21.00	3.57	20.73	0.00
Male	N=15			N=8		
Endpoint	33.28	60.88	13.07	2.78	4.86	0.54
Change from Baseline	22.77	37.73	8.91	0.46	1.79	0.00
Female	n=14			n=18		
Endpoint	68.65	97.33	22.30	16.92	35.76	1.62
Change from Baseline	47.30	69.42	21.42	4.96	24.97	0.00
Age <sup>1</sup> <=51	N=16			N=12		
Endpoint	47.38	76.95	20.14	8.67	17.76	1.58
Change from Baseline	30.32	49.11	19.27	-3.93	8.65	0.00
Age > 51	n=13			n=14		
Endpoint	54.02	88.96	21.00	15.91	38.43	1.08
Change from Baseline	39.89	64.62	21.00	10.01	25.81	0.00

1. The cut-off point uses median age of 51 years old from all patient population.

Recall that patients were on different dosage of GHB during the lead-in phase as well as during the double-blind phase, if they were randomized to GHB group. An efficacy analysis using a similar model to the primary model of ANCOVA with an additional

factor of dose group was performed. The treatment effect remained with the same significance level ( $p \leq 0.0001$ ).

Descriptive statistics of the change from baseline in the number of cataplexy attacks by dose was calculated to examine the efficacy on different dose level. The results are summarized in Table 23.

**Table 23. Descriptive statistics of the number of cataplexy attacks by patient's GHB dose**

Treatment		GHB dose (gram)			
		3.0 and 4.5	6.0	7.5	9.0
All patients		N=11	N=15	15	14
Baseline	Mean (SD)	10.9 (14.6)	3.2 (3.9)	21.6 (54.4)	14.2 (24.8)
	Median	3.3	1.8	0.0	5.2
Week 4	Mean (SD)	44.1 (94.0)	15.8 (20.0)	41.9 (88.4)	31.1 (35.3)
	Median	1.1	7.5	7.5	21.3
Change	Mean (SD)	33.3 (84.3)	12.6 (18.1)	20.3 (40.7)	16.9 (26.5)
	Median	-0.8	3.5	3.0	14.3
GHB		N=5	N=7	N=7	N=7
Baseline	Mean (SD)	5.3 (9.1)	3.8 (4.3)	0.7 (1.3)	25.2 (32.3)
	Median	1.1	1.8	0.0	10.5
Week 4	Mean (SD)	0.8 (0.9)	10.2 (21.7)	1.6 (2.1)	34.2 (50.5)
	Median	1.0	0.0	1.1	14.0
Change	Mean (SD)	-4.4 (8.3)	6.5 (18.9)	0.9 (1.1)	9.0 (35.7)
	Median	-0.9	0	1.1	0
Placebo		N=6	N=8	N=8	N=7
Baseline	Mean (SD)	15.5 (17.3)	2.7 (3.7)	39.9 (71.5)	3.2 (3.3)
	Median	12.5	1.4	5.7	3.0
Week 4	Mean (SD)	80.2 (119.3)	20.8 (18.2)	77.1 (112.3)	28.0 (11.4)
	Median	6.9	15.6	20.4	26.9
Change	Mean (SD)	64.7 (107.4)	18.0 (51.1)	37.2 (51.1)	24.8 (10.2)
	Median	0.5	15.6	17.4	23.9

Patients who withdrew from GHB had a larger number of rebound cataplexy attacks across all dose level compared to patients stayed on GHB. In terms of median numbers, patients who were on larger dose of GHB seemed to have a larger number of rebound cataplexy attacks after withdrew from GHB.

### 5.9 Reviewer's Efficacy Conclusion for Study SXB-21

This study was to evaluate the long-term efficacy of GHB. Because of the difficulty in placing patients in the treatment of placebo in a long term, a withdrawal design was used.

After the withdrawal from GHB, a median increase of 21 cataplexy attacks per 2 weeks was observed from patients who had been taking GHB at a stable dose. For patients who had continuously taking GHB at stable dose, no change was observed in the median number of the cataplexy attacks. The difference between the treatment groups reached a

statistically significant level from the protocol-specified analysis. This reviewer would conclude that the study has provided sufficient evidence that the treatment of GHB was efficacious with regard to the number of cataplexy attacks in patients who had been on GHB treatment.

## **6. Reviewer's Overall Summary and Conclusion**

In this NDA submission, 4 efficacy studies were reviewed. Study OMC-GHB-2 was a dose finding study, which used a parallel design of 3g, 6g, and 9g GHB treatment groups plus placebo. The study was the largest one among the 4 efficacy studies and 136 patients were enrolled and received 4 weeks of double-blind treatment of GHB or placebo. This study has demonstrated that GHB was efficacious with regard to the total number of cataplexy attacks in treating patients with narcolepsy symptoms, although the most favorable safe and effective dose could not be determined by this reviewer.

When break down the total number of cataplexy attacks into complete and partial cataplexy attacks, no significant treatment difference was found in the number of complete cataplexy attacks. The treatment difference in partial cataplexy attacks carried a p-value of 0.0204, which is subjected to multiplicity adjustment.

Daytime sleepiness was measured by Epworth sleepiness index. The p-value of the treatment effect for this parameter was 0.0109, which is also subjected to multiplicity adjustment. Because 10 secondary efficacy parameters were proposed, none of the secondary efficacy had met the significance criteria after Bonferroni adjustment, although numerically the data showed benefit for the treatment of 9g GHB in partial cataplexy attacks and daytime sleepiness. The comparison between treatment of 9g GHB and placebo on secondary efficacy parameters was not proceeded due to the fact that the overall comparisons on those parameters could not be concluded as positive.

Lammers study used a crossover design. Twenty-five patients received both GHB and placebo treatment of 4 weeks each. The study had two primary efficacy parameters: number of cataplexy attacks and global therapeutic impression. Neither the primary efficacy parameters nor the primary analysis were appropriately defined. No conclusion could be reached, as whether or not the GHB was efficacious as to the cataplexy attacks or global therapeutic impression.

Scrima study was a crossover study in which 25 patients received both GHB and placebo treatment of 4 weeks each. The study designated two primary efficacy parameters: number of cataplexy attacks and MSLT sleepiness index. The treatment effect showed a p-value of 0.0372 for the number of cataplexy attacks and a p-value of 0.0688 as for the MSLT sleepiness index. It was not specified in the protocol how the alpha should be adjusted for the two parameters. The usual practice is 1) either parameter win at significance level of 0.025 or; 2) both parameters win at significance level of 0.05. None of the criteria were met by the efficacy results. Therefore, the study could not be concluded as showing significant treatment effect.

Study SXB-21 included a 2-week period of double blind treatment of GHB or placebo to examine the long-term effect of GHB after withdrawal from GHB for patients who had been on stable dose of GHB. The primary efficacy parameter was the change in the number of cataplexy attacks from baseline, which was the only parameter designated in the study. A total of 55 patients received double-blind treatment. The study demonstrated that patients who had been staying on the stable dose of GHB had less cataplexy attacks compared to patients who withdrew from GHB treatment. The p-value obtained from the primary analysis of cataplexy was 0.0001.

Reviewer's Overall Conclusion

This reviewer concludes that among the 4 studies reviewed, Study OMC-GHB-2 and Study SXB-21 have demonstrated that the treatment of GHB was effective in patients with symptoms of narcolepsy. Sufficient evidence was found from those two studies that patients receiving GHB treatment had fewer cataplexy attacks on average, as compared to patients receiving placebo. Lammers study could not be conclude as a positive study for reasons of 1) the primary analysis prospectively defined was negative and not appropriate; 2) all analysis results reported by the sponsor were not prospectively defined; and 3) the results from post-hoc analyses using different methodologies were not consistent. For Scrima study, this reviewer could not reach a conclusion that the study had showed significant treatment effect.

*/S/*

---

Sharon Yan, Ph.D.  
Mathematical Statistician

Concur:

*/S/*

---

Kun Jin, Ph.D.  
Team Leader

*/S/*

---

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

Cc: NDA# 21-196

HFD-120/Dr. Katz

HFD-120/Dr. Feeney

HFD-120/Dr. Mani

HFD-120/Ms. Homonnay (CSO)

HFD-710/Dr. Anello

HFD-710/Dr. Chi

HFD-710/Dr. Jin

HFD-710/Dr. Yan

This review consists of 38 pages. 3/22/2001