

PRELIMINARY REVIEW

EFFICACY REVIEW OF NEW DRUG APPLICATION

NDA	21196
Sponsor:	Orphan Medical, Inc.
Drug:	Xyrem
Proposed Indication:	Narcolepsy
Material Submitted:	Original New Drug Application
Correspondence Date:	9/30/00
Date Received / Agency:	10/3/00
Date Preliminary Review Completed	2/14/01
Reviewer:	Ranjit B. Mani, M.D.

1. BACKGROUND.....	4
2. REVIEW SOURCES.....	5
2.1 MATERIALS FROM NDA	5
2.2 RELATED REVIEWS, CONSULTS	5
3. TABULAR SUMMARY OF EFFICACY STUDIES IN ORIGINAL NDA SUBMISSION	5
3.1 STUDY OMC-GHB-2	5
3.2 SCRIMA STUDY.....	5
3.3 LAMMERS STUDY	6
4. RATING SCALES USED.....	6
4.1 EPWORTH SLEEPINESS SCALE.....	6
4.2 CLINICAL GLOBAL IMPRESSION OF SEVERITY (CGI-S)	6
4.3 CLINICAL GLOBAL IMPRESSION OF CHANGE (CGI-C)	7
5. HUMAN PHARMACOKINETICS	7
6. STUDY # OMC-GHB-2	8
6.1 OBJECTIVES.....	8
6.2 DESIGN.....	8
6.3 INCLUSION CRITERIA.....	8
6.4 EXCLUSION CRITERIA	9
6.5 CONCOMITANT MEDICATION.....	9
6.6 DOSAGE.....	10
6.7 SCHEDULE	10
6.8 EFFICACY OUTCOME MEASURES.....	11
6.9 ANALYSIS PLAN.....	11
6.10 PROTOCOL CHANGES.....	13
6.11 EFFICACY RESULTS.....	13
6.12 SAFETY RESULTS	24
6.13 SPONSOR’S CONCLUSION.....	24
6.14 REVIEWER’S COMMENTS	24
7. SCRIMA STUDY	25
7.1 REVIEWER’S COMMENTS REGARDING STUDY PROTOCOL	26
7.2 OBJECTIVE	26
7.3 DESIGN.....	27
7.4 DURATION	28
7.5 DOSAGE.....	28

7.6	SAMPLE SIZE	28
7.7	MAIN INCLUSION CRITERIA	28
7.8	MAIN EXCLUSION CRITERIA	28
7.9	CONCOMITANT MEDICATIONS	28
7.10	SCHEDULE	29
7.11	EFFICACY OUTCOME MEASURES	30
7.12	ANALYSIS PLAN.....	31
7.13	PROTOCOL AMENDMENTS	35
7.14	EFFICACY RESULTS.....	35
7.15	SAFETY RESULTS	41
7.16	SPONSOR’S CONCLUSIONS	41
7.17	REVIEWER’S COMMENTS	41
8.	LAMMERS STUDY.....	42
8.1	OBJECTIVES.....	42
8.2	DESIGN.....	43
8.3	INCLUSION CRITERIA	43
8.4	EXCLUSION CRITERIA	43
8.5	CONCOMITANT MEDICATION.....	43
8.6	DOSAGE.....	43
8.7	SCHEDULE	44
8.8	EFFICACY MEASURES	44
8.9	ANALYSIS PLAN.....	45
8.10	PROTOCOL CHANGES.....	46
8.11	EFFICACY RESULTS.....	47
8.12	SAFETY RESULTS	51
8.13	INVESTIGATOR’S CONCLUSIONS	51
8.14	EARLIER DIVISIONAL COMMENTS AND CONCLUSIONS ABOUT LAMMERS STUDY.....	52
8.15	REVIEWER’S COMMENTS ABOUT SECONDARY EFFICACY MEASURES	52
9.	STUDY OMC-SXB-20	53
10.	STUDY OMC-SXB-21	53
10.1	TABULAR SUMMARY	53
10.2	TITLE	53
10.3	OBJECTIVE.....	54
10.4	DESIGN	54
10.5	DURATION	54
10.6	SAMPLE SIZE	54
10.7	SELECTION.....	54
10.8	DOSAGE.....	55
10.9	SCHEDULE	56
10.10	OUTCOME MEASURES	56
10.11	SAFETY MONITORING	56
10.12	ANALYSIS PLAN.....	57
10.13	PROTOCOL AMENDMENTS	58
10.14	ACTUAL ANALYSES PERFORMED.....	58
10.15	EFFICACY RESULTS.....	58
10.16	SAFETY RESULTS	62
10.17	SPONSOR’S CONCLUSIONS REGARDING EFFICACY	62
10.18	REVIEWER’S COMMENTS	62
11.	OVERALL COMMENTS REGARDING EFFICACY	63
12.	LABELING REVIEW	64

13. OVERALL COMMENTS REGARDING SAFETY OF XYREM®	64
13.1 CLINICAL SAFETY	64
13.2 WITHDRAWAL PHENOMENA AND ABUSE POTENTIAL	64
13.3 RISK MANAGEMENT PROGRAM	65
14. RISK-BENEFIT EQUATION	65
15. STUDY SITE INSPECTIONS	65
16. FINANCIAL DISCLOSURE CERTIFICATION	65
16.1 COMPONENTS OF CERTIFICATION	65
16.2 REVIEWER’S COMMENT.....	66
17. CONCLUSIONS	66
18. RECOMMENDATIONS	66

1. Background

This submission contains an original New Drug Application for Xyrem® (sodium oxybate) oral solution.

This drug has been developed by Orphan Medical, Inc. for the treatment of narcolepsy under IND # 49641 and Treatment IND # 57271. Data obtained from individual sponsor-investigator INDs #s 21654 (M. Scharf) and 19911 (L. Scrima) have also been used in support of this application.

Narcolepsy is a chronic neurological disorder characterized by excessive daytime sleepiness, disturbed nocturnal sleep, cataplexy, sleep paralysis and hypnagogic hallucinations. The prevalence of this condition in the United States, as per a publication cited by the sponsor is 0.02% and 0.07%. According to the sponsor, current treatments for this condition are limited in effectiveness and have frequent undesirable adverse events.

Through the current application, the sponsor wishes to pursue the following claim:

“Xyrem® (sodium oxybate) oral solution is indicated to reduce the incidence of cataplexy and to improve the symptom of daytime sleepiness in patients with narcolepsy.”

Please see also see my reviews of submissions under IND # 49641 and my review of Treatment IND # 57271 for details about this drug.

In this review the words/phrases “ γ -hydroxybutyrate (GHB)”, “sodium oxybate” and “Xyrem®” have been used interchangeably.

The original NDA submission of 9/30/00 has been followed by an additional submission of 12/18/00 that contains a further efficacy study, OMC-SXB-21

2. Review Sources

2.1 Materials from NDA

In reviewing this application I have read the following volumes of the NDA submission of 9/30/00. These volumes have been read almost entirely in electronic format. They include

- Volume 1, containing the application summary
- Volumes 36-63, containing the full reports for the following studies: OMC-GHB-2, Scrima and Lammers
- Volume 100 containing the Integrated Summary of Efficacy

I have also reviewed the following:

- A separate submission dated 12/16/00 containing the final report for the following long-term efficacy trial: OMC-SXB-21. This application also contained updated draft labeling
- The sponsor's responses to a number of requests for information from this reviewer

2.2 Related Reviews, Consults

I have utilized the many reviews that I have done, since 1997, of submissions under IND # 49641 and Treatment IND # 57271 for details about this drug.

3. Tabular Summary Of Efficacy Studies In Original NDA Submission

3 studies have been used in the main submission to support the efficacy of Xyrem® in the treatment of narcolepsy.

3.1 Study OMC-GHB-2

Study #	OMC-GHB-02 Orphan Medical			
Design	Randomized, double-blind, placebo-controlled, parallel-arm			
Duration	4 weeks			
Dosage	9 g	6 g	3 g	Placebo
Number randomized	35	33	34	34
Number completed	28	29	30	33
Main inclusion criteria	Narcolepsy for at least 6 months with both excessive daytime sleepiness and cataplexy			
Primary outcome measures	Total number of cataplexy attacks			
Main efficacy analysis (statistically significant results)	9 g dose superior to placebo, based on ANCOVA (p = 0.0008)			

3.2 Scrima Study

Study #	Scrima	
Design	Randomized, double-blind, placebo-controlled, cross-over	
Duration	4 weeks	
Dosage	50 mg/kg/day	Placebo
Number randomized	20	20
Number completed	20	20
Main inclusion criteria	Excessive daytime sleepiness, a history of cataplexy with ≥ 10 cataplexy attacks over the 2 week baseline period and ≥ 2 REM onsets and a sleepiness index of ≥ 75 on the a multiple sleep latency test	
Primary outcome measures	Total number of cataplexy attacks per day	

Study #	Scrima
	Sleepiness Index on Multiple Sleep Latency Test
Main efficacy analysis (statistically significant results)	GHB superior to placebo on total number of cataplexy attacks (p = 0.013)

3.3 Lammers Study

Study #	N -1 (R 55 667 082) Lammers et al	
Design	Randomized, double-blind, placebo-controlled, cross-over	
Duration	4 weeks	
Dosage	4.75 g *	Placebo
Number randomized	25	25
Number completed	25 **	25 **
Main inclusion criteria	Excessive daytime sleepiness and at least one of the following: cataplexy, hypnagogic hallucinations, and sleep paralysis	
Primary outcome measures	Total number of cataplexy attacks Global therapeutic impression (patient) Global clinical impression (clinician)	
Main efficacy analysis (statistically significant results)	GHB superior to placebo on first two of above measures, numbered as above p = 0.002 (ANCOVA)*** p = 0.001 (McNemar's test) Not measured	

*This dose is the mean of the protocol-specified dose of 60 mg/kg/day (range 3.78 to 5.52 g/day)

** The number included in the efficacy analysis was 24 for reasons which are described below in a more detailed review of the study

***This was not the protocol specified analysis. The ANCOVA was performed by the current sponsor several years after the study blind was broken and after the initial report of this study was published. The protocol-specified analysis (which was cited in the publication) was the Wilcoxon Signed Rank Test which yielded a p-value of 0.42, but which may have been an inappropriate analysis.

4. Rating Scales Used

4.1 Epworth Sleepiness Scale

The Epworth Sleepiness Scale is a patient-rated measure of daytime sleepiness. Patients are asked to rate their chances of dozing during each of the following 8 activities on a scale from 0-3 (0=never; 1=slight; 2=moderate; 3=high): sitting and reading; watching TV, sitting inactive in a public place; as a passenger in a car for an hour without a break; lying down to rest in the afternoon when circumstances permit; sitting talking to someone; sitting quietly after a lunch without alcohol; in a car while stopped for a few minutes in traffic

4.2 Clinical Global Impression Of Severity (CGI-S)

The following description applies to this measure as used in the OMC-GHB-2 study

This measure involved rating the severity of the patient's narcolepsy at baseline. The rating was made in relation to the investigator's total experience with the narcoleptic population and graded with one of the following:

- Not assessed
- Normal-no signs of illness
- Borderline ill
- Slightly ill
- Moderately ill
- Markedly ill

Among the most extremely ill.

4.3 Clinical Global Impression Of Change (CGI-C)

The following description applies to this measure as used in the OMC-GHB-2 study

This measure assessed the overall change in the patient's severity of narcolepsy using the following rating scale:

Very much improved
Much improved
Minimally improved
No change
Minimally worse
Much worse
Very much worse

5. Human Pharmacokinetics

The following pharmacokinetic summary is based on a summary supplied by the sponsor in this submission.

Orally administered GHB is rapidly absorbed with a t_{max} of 30 - 75 minutes and to a similar degree in narcoleptic and other patient populations; absorption characteristics are similar in males and females and are not altered by chronic dosing; t_{max} is delayed, at higher doses (suggesting a limited absorption capacity) and by the administration of food. C_{max} and $AUC_{0-\infty}$ are reduced by the administration of the drug with food. The absolute bioavailability of the drug is < 30%.

The apparent volume of distribution divided by absolute bioavailability (V_d/F) ranges between 190 and 384 mL/kg. Inter-subject variability in the volume of distribution is high as indicated by the coefficient of variation which ranges between 16% and 84%. The drug readily crosses the placental and blood-brain barriers. Protein binding has been estimated at about 1%.

Less than 5% of an oral dose of GHB is excreted unchanged in the urine. Based on a review of the scientific literature the sponsor states that the end-product of metabolism, regardless of biotransformation pathway, is carbon dioxide. 2 main biotransformation pathways have been identified:

- A β -oxidation pathway
- A pathway involving the entry of succinic acid into the tricarboxylic acid cycle, through the initial formation of succinic semialdehyde

First-pass metabolism occurs with orally administered GHB, probably through the β -oxidation pathway, resulting in an oral bioavailability of < 30%. Intermediate compounds in the metabolic pathways for GHB do not appear to be pharmacologically active

The pharmacokinetics of GHB are non-linear. Plasma clearance is dose-dependent across the therapeutic range: following a total dose of 9 g (2 doses of 4.5 g each administered 4 hours apart) the apparent elimination half-life of GHB was 0.83 hours, which was approximately 40% longer than the mean elimination half-life following a total dose of 4.5 g (2 doses of 2.25 g each administered 4 hours apart). Chronic dosing with GHB did not alter its pharmacokinetics in a clinically significant manner: treatment with this drug for 8 weeks resulted in 13% and 16% increases in $AUC_{0-\infty}$ and C_{max} , respectively; these increases were not considered clinically significant.

There are no significant gender differences in the pharmacokinetics of GHB. Neither are there significant differences in pharmacokinetics between healthy subjects and narcoleptic patients, and between healthy patients and those who are alcohol-dependent. Oral clearance of GHB is altered

in the presence of cirrhosis with or without ascites. Renal disease is not expected to alter the pharmacokinetics of GHB; studies in that setting have therefore not been carried out.

Formal studies indicated that GHB had no interactions with protryptiline, zolpidem and modafinil. In-vitro pooled human liver microsomal studies showed that GHB did not significantly inhibit or enhance the activities of human CYP450 isoenzymes.

6. Study # OMC-GHB-2

6.1 Objectives

6.1.1 Primary

To evaluate and compare the efficacy of 3 doses (3 g, 6 g and 9 g) of GHB and placebo in the treatment of the symptoms of narcolepsy

6.1.2 Secondary

To evaluate and compare the safety of 3 doses (3 g, 6 g and 9 g) of GHB and placebo in the treatment of the symptoms of narcolepsy

6.2 Design

Randomized, double-blind, placebo-controlled, parallel-group, 4-arm, multicenter study

The study comprised 5 phases

Screening Period: This was intended to last 1 day to 4 weeks and permitted withdrawal of tricyclic antidepressants and other drugs used to treat cataplexy

Washout Period: This was intended to last 5 - 28 days and allowed for the clinical effects of tricyclic antidepressants and other medications for cataplexy to be eliminated, for rebound cataplexy to abate and to train patients in the use of the diary; this lasted a minimal period of 5 days in those in whom no medications were being withdrawn so that they could be trained in the use of the diary

Baseline Period: This period was intended to last 2 - 3 weeks and was intended to assess the patient's periods of cataplexy and establish a stable number of attacks; if the frequency of attacks was not stable at the end of the 3-week period the patient was discharged from the study; the judgement of the investigator was to determine whether the frequency of attacks of cataplexy was stable

Double-blind Treatment Period: This was intended to last 4 weeks

Follow-up Period: This was proposed as a period of 3 -5 days after study medication was stopped

6.3 Inclusion Criteria

- Informed consent
- Age \geq 18 years
- Willing and able to complete the entire trial
- History of excessive daytime sleepiness
- History of cataplexy. In addition, patients must have recorded a period of 3 or more complete and/or partial cataplexy attacks per weeks during the last 2 weeks of the baseline period
- Valid polysomnography (PSG) scores
- Current diagnosis of narcolepsy for at least 6 months, according to the criteria below

- Recurrent naps or lapses into sleep that occur almost daily for at least 6 months
- Sudden loss of bilateral postural muscle tone in association with intense emotion
- **If female must be**
 - Surgically sterile OR
 - 2 years post-menopausal OR
 - If of child-bearing potential must be using effective contraception and must continue this treatment during the study

6.4 Exclusion Criteria

- Unstable diseases in any body system, other than narcolepsy, which would place the patient at risk or compromise the trial objectives
- Psychiatric disorders that would preclude participation in, or completion of, the trial
- History of seizure disorder or of anticonvulsant therapy
- History of substance abuse, as defined by DSM-IV, currently or within the past year
- Serum creatinine > 2 mg/dl; AST or ALT > 2 x upper limit of normal; pre-trial electrocardiogram results demonstrating a clinically significant arrhythmia or 2nd or 3rd degree A-V block; history of myocardial infarction within the past 6 months
- Any untreated disorder other than narcolepsy that could be considered a primary cause of excessive daytime sleepiness, including sleep apnea syndrome (criteria specified)
- Occupation requiring variable shift or routine night shift work
- Participation and use of GHB in a previous study
- No clinically significant history of head trauma or of a seizure disorder
- Inadequate support for the duration of the study
- Willing to not operate a car or heavy machinery if the clinical investigator feels such a restriction is warranted
- Use of medication for narcolepsy during baseline period, other than a stable dose of stimulant medication ("stable dose" defined as one without any significant change in dose for the 5 - day period just prior to the baseline period)
- Use of hypnotics, tranquilizers, antihistamines (except for the non-sedating variety of such drugs) and clonidine at the start of the baseline period.

6.5 Concomitant Medication

- The following medications were prohibited: hypnotics, tranquilizers, antihistamines (except for the non-sedating variety of such drugs), clonidine, tricyclic antidepressants, selective serotonin reuptake inhibitors, MAO inhibitors, tetracyclic antidepressants, anticonvulsants and alcohol.
- Patients were cautioned regarding the use of opioid analgesics and alcohol
- Oral contraceptives were permitted in women of child-bearing potential
- Over-the-counter medications needed careful review by the clinical investigator prior to use

6.6 Dosage

- No study medication was to be used during the screening, washout, baseline and follow-up periods of the study
- Those entering the randomized, double-blind phase of the study were assigned to one of the following arms:
 - Placebo
 - GHB 3 g daily
 - GHB 6 g daily
 - GHB 9 g daily
- The total nightly dose was divided into 2 equal portions. These were dissolved in water and administered at bedtime and again 2.5 to 4 hours later

6.7 Schedule

- 7 study visits were scheduled in all as follows; patients were due to be contacted at least 3 times per week during all study periods

Visit Number	Timing
1	Screening period
2	Washout period
3	Baseline period
4	Beginning of double-blind phase
5	Day 14 of double-blind phase
6	Day 28 of double-blind phase
7	During follow-up period and within 3-5 days of Visit 6

- At the screening visit the following were to be obtained or checked: informed consent, medical history, inclusion and exclusion criteria, concomitant medication, determination of duration of screening and washout phases, plan for withdrawal of tricyclic or other prohibited medication and other procedures listed below
- Physical examinations (including neurological examinations), and electrocardiograms were to be performed at screening and at Visit 6
- Vital signs were to be checked at every visit
- Clinical laboratory tests (hematology, clinical chemistry, urinalysis and urine pregnancy test) were to be checked at screening and at all subsequent visits, other than Visit 2
- Adverse events and concomitant medication use were to be checked at every visit
- Fresh daily diaries were to be provided at each visit; at each visit the study diary for the period just completed was to be collected; instructions regarding the use of the diaries was provided at every visit
- The state of the patient's cataplexy was to be assessed at Visit 2 and every subsequent visit
- The Epworth Sleepiness Scale was to be assessed at Visits 4, 6 and 7
- A Clinical Global Impression of Severity of the patient's narcolepsy was to be assessed at Visit 4
- A Clinical Global Impression of Change in the severity of the patient's narcolepsy was to be assessed at Visits 6 and 7

6.8 Efficacy Outcome Measures

6.8.1 Primary

Total number of cataplexy attacks (sum of complete and partial cataplexy attacks) per week

6.8.2 Secondary

- Complete cataplexy attacks
- Partial cataplexy attacks
- Daytime sleepiness as assessed with the Epworth Sleepiness Scale
- Clinical Global Impression of Severity of the patient's narcolepsy (not, strictly speaking, an outcome measure). This was made in relation to the investigator's total experience with the narcoleptic population and graded with one of the following: not assessed, normal-no signs of illness, borderline ill, slightly ill, moderately ill, markedly ill, and among the most extremely ill.
- Clinical Global Impression of Change in the severity of the patient's narcolepsy using the following rating scale: very much improved, much improved, minimally improved, no change, minimally worse, much worse, and very much worse
- Number and duration of awakenings each night
- Total amount of sleep each night
- Number and duration of inadvertent naps and sleep attacks
- Number and occurrences of hypnagogic hallucinations or sleep paralysis
- Quality of sleep, level of alertness and overall ability to concentrate using the following scale: 1-excellent, 2-good, 3-fair and 4-poor (this analysis was intended to be purely "exploratory")

6.8.3 Safety Outcome Measures

Adverse events, vital signs, electrocardiograms, safety laboratory tests and concomitant medication use

6.9 Analysis Plan

All statistical tests were to be declared significant if the two-sided p-value was < 0.05

6.9.1 Demographic And Baseline Characteristics

Quantitative variables were to be analyzed using ANOVA or the Kruskal-Wallis test as appropriate; qualitative variables were to be analyzed using Fisher's exact test. If any site failed to reach a minimum of 8 patients, these sites were to be pooled and treated as a single site for purposes of statistical analysis.

6.9.2 Primary Efficacy Analysis

- The primary efficacy analysis was to be on the intention-to-treat population defined as all patients who received study drug on whom a post-treatment evaluation visit was completed.
- As noted above the primary efficacy variable was the total number of cataplexy attacks per week.

- The primary efficacy analysis was directed to comparing the treatment groups for the change from baseline to endpoint in the total number of cataplexy attacks. Baseline was defined as the 2-week period immediately prior to receiving the study drug. Endpoint was defined as the final 2-week period on study drug.
- **The primary efficacy analysis was to be based on ANOVA with the model including treatment group, trial site and treatment-by-site interaction; if the interaction was found not to be statistically significant, the analysis would be rerun excluding that term from the model.**
- **If the ANOVA demonstrated statistically significant differences among the products, each of the active treatment groups were to be compared with the placebo group using least significant differences**
- **A within-group analysis was also to be performed comparing assessing the significance of the median change from baseline using the paired t-test**
- **If the assumptions for the above proposed between-group and within-group analyses were not met, the between-group analysis would use a blocked Wilcoxon test, and the within-group analysis would be based on the Wilcoxon signed-rank test**
- For the primary efficacy variable (between-group comparison) an ANCOVA was **also** to be performed using the baseline observation as a covariate; no further specifications about the ANCOVA are in the original protocol (in its revised version dated 12/5/96), or in the amendment submitted as serial # 007 on 2/7/97; however in the study report the sponsor states that “prior to the completion of the study and database lock, an analysis plan was written and approved” (although seemingly not in a submitted protocol amendment) that detailed performing a log transformation of the data if the assumptions for ANCOVA were not satisfied. The sponsor further states that “it was anticipated that for many, if not all, of the efficacy variables, the log transformation would result in a more normal distribution conforming to the requirements of the ANCOVA”. Such a log transformation was eventually needed for all the primary and secondary efficacy variables except for the Epworth Sleepiness Scale, total amount of sleep each night and number of inadvertent naps per day; Fisher’s Exact Test was used for the Clinical Global Impression of Change. Eventually an ANCOVA on log-transformed data was used as the primary efficacy analysis, as presented in the study report
- An additional analysis was also to be performed to look for a possible dose-response relationship

6.9.3 Secondary Efficacy Measures

The secondary efficacy measures were to be analyzed using measures similar to the above.

6.9.4 Sample Size Calculation

The sample size calculation was based on the primary efficacy variable

In the calculation:

- Based upon previous trials with GHB, a mean reduction of at least 2 cataplexy events per week was to be expected over a one-month treatment period, with a standard deviation of 2.5.
- Using a power of 80 % and a 2-sided significance level of 0.05, it was calculated that 25 patients would be needed per treatment group to detect a difference of 2 with respect to change in cataplexy events

6.9.5 Safety Analysis

Safety data are discussed as part of the NDA Safety Review

6.10 Protocol Changes

The above represents the **revised** protocol. The original protocol was submitted 1/10/96. Major changes to the original protocol, which were all made in a single submission (serial # 007 under IND # 49641, amendment # 1; dated 2/7/97) are as follows:

- In the original protocol, 3 primary efficacy variables were listed: number of cataplexy attacks, number of sleep attacks and duration of daytime naps. The revised protocol uses only a single primary outcome measure: the total number of cataplexy attacks
- In the original protocol the interval between the two nightly doses of medication was 4 hours. The revised protocol changed the dosing interval to 2.5 to 4 hours
- In the original protocol valid Multiple Sleep Latency Test scores were considered an inclusion criterion; in the revised protocol this requirement was dropped as the other inclusion criteria satisfied the need for an appropriate diagnosis.

The above changes were deemed by this Division to be acceptable provided the sponsor had not already examined the data for this trial.

Note also that primary efficacy analysis, both in the original and in the amended protocols, was not to be based on ANCOVA, which was intended to be an additional analysis. However in the primary efficacy analysis in the study report ANCOVA, using log-transformed data was used for the primary efficacy analysis and for the analysis of most secondary efficacy measures.

6.11 Efficacy Results

6.11.1 Number of Patients and Disposition

136 patients from 16 centers were enrolled in the study. 120 patients completed the study. Their disposition is outlined in the table below which has been copied from the submission

Disposition	All patients	Placebo	GHB Dose (g/day)		
			3	6	9
RECEIVED STUDY MEDICATION	136	34	34	33	35
WITHDREW FROM STUDY					

Disposition	All patients	Placebo	GHB Dose (g/day)		
			3	6	9
			Adverse Event	10	1
Protocol Deviation	1	0	1	0	0
Patient Request	2	0	1	0	1
Lost to Follow-up	1	0	0	1	0
Lack of Efficacy	2	0	1	1	0
Total Withdrawals	16	1	1	4	7
COMPLETED STUDY	120	33	30	29	28

As the table indicates the primary reason for withdrawal was the development of adverse events; these were more frequent in the 9 g/day group than in any other group. The highest proportion (20 %) of withdrawals due to all causes occurred in that group

6.11.2 Duration of Treatment

The mean and standard deviation for duration of treatment in each treatment group is indicated in the following table

Treatment Group	Number	Mean (days)	Standard deviation (days)
All subjects	136	27.39	7.22
Placebo	34	29.00	1.95
3 g	34	28.03	7.72
6 g	33	26.82	7.48
9 g	35	25.43	9.18

As the table indicates the groups are comparable as regards their means, although the standard deviation for the placebo group is substantially less than that for the placebo group

6.11.3 Protocol Deviations

Major protocol deviations are summarized in the following table

Deviation	All patients	Placebo	GHB Dose (g/day)		
			3	6	9
Concomitant Medications	9	1	5	1	2
Dosing Error	6	2	2	2	0
Laboratory Procedures	7	1	1	1	4

Errors in concomitant medication were thus the commonest form of protocol deviation. In the above each protocol deviation was counted separately; thus if a single patient had more than 1 protocol deviation of the same kind, each of these deviations was considered a separate event; the total number of events in each category is listed in the above table

6.11.4 Dataset Analyzed

Only an intention-to-treat analysis was performed. This dataset was defined in the protocol as consisting of all those who received a dose of study drug and had a post-treatment follow-up evaluation carried out. No patients were excluded from this dataset whose overall size and distribution by treatment group varied slightly depending upon the outcome measure analyzed. Their distribution according to treatment group is indicated under “Number of Patients and Disposition” above

6.11.5 Demographic Characteristics

These are summarized in the following table

Disposition	All patients	Placebo	GHB Dose (g/day)			p-value (ANOVA)
			3	6	9	
			Randomized	136	34	34
Mean Age (years)	43.06	40.82	47.06	43.52	40.91	0.2737
Male (%)	41.9	35.2	20.5	63.6	48.6	0.0027
Caucasian (%)	91.2	85.3	97.1	93.9	88.6	0.1379
Mean Height (cm)	170.91	171.97	166.7	173.1	171.9	0.0283
Mean Weight (kg)	82.87	83.98	78.86	85.04	83.56	0.4847

As the above table indicates there were statistically significant differences in gender and height between treatment groups

6.11.6 Baseline Severity of Narcolepsy

The baseline severity of patient symptoms is indicated in part by the following table which depicts the number of patients reporting symptoms in each category during the 3 months prior to screening. The table has been copied from the submission

Symptoms	Number (%) of patients 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 day-time sleepiness	34 (100%)	34 (100%)	32 (97%)	34 (97%)
Awakenings 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%)

As the above table indicates the severity of narcolepsy appears to have been roughly similar across treatment groups at baseline based on the above measures; however a statistical analysis comparing treatment groups does not appear to have been performed.

Since the total number of cataplexy attacks per week was the designated primary outcome measure, the mean number of cataplexy attacks per week over the 3-month period prior to the screening visit has also been used to compare the treatment groups as shown in the following table, copied from the sponsor's submission.

Statistic	GHB Dose (g)			
	Placebo	3	6	9
N	34	33	33	35
Mean	15.5	12.3	16.6	16.7

NOTE: The 4 treatment groups appear roughly comparable in regard to this measure, but a statistical analysis comparing this measure among the treatment groups does not appear to have been performed. This measure was recorded based on patient recall, as part of the narcolepsy history, at the screening visit; its validity and reliability are questionable.

A more reliable and valid estimate of the baseline severity of cataplexy is from data obtained at the baseline visit for the double-blind phase (Visit 4) which is derived from diary records and after stability of the patient's condition was established. These data are provided in the next table which indicates that the treatment groups were comparable, based on the p-value derived from the Kruskal-Wallis test. The table is copied from the submission.

Type of event	Placebo	GHB Dose (g)			p-value Kruskal-Wallis
		3	6	9	
Total cataplexy attacks/week					0.7749
N	34	33	33	35	
Mean	34.27	28.57	38.85	34.60	
Median	20.21	20.00	23.00	23.50	
SD	46.63	30.53	55.04	33.92	
Complete cataplexy attacks/week					0.5151
N	34	33	33	35	
Mean	6.86	7.08	15.26	8.61	
Median	1.12	4.50	4.85	2.00	
SD	12.37	8.50	27.53	14.01	
Partial cataplexy attacks/week					0.7289
N	34	33	33	35	
Mean	27.44	21.49	23.59	26.12	
Median	15.03	15.00	16.15	18.79	
SD	42.08	28.30	29.01	26.14	

Other baseline narcolepsy symptoms based on diary recordings assessed at Visit 4 are depicted in the following table; again the p-value based on the Kruskal-Wallis test indicates that the groups are comparable

Event (Mean Daily Frequency)	Placebo	GHB Dose (g/day)			p-value (Kruskal-Wallis)
		3	6	9	
Hypnagogic hallucinations	0.57	0.58	1.14	0.53	0.9766
Sleep paralysis episodes	0.51	0.42	0.73	0.47	0.9597
Inadvertent naps daily	1.71	1.91	1.70	1.72	0.7008

Baseline Epworth Sleepiness Scale results are depicted in the next table copied from the submission; here again the groups appear to be comparable

Statistic	Placebo	GHB Dose (g)		
		3	6	9
N	34	34	32	35
Mean	18.47	17.06	17.28	16.66
SD	3.13	3.71	3.49	4.07

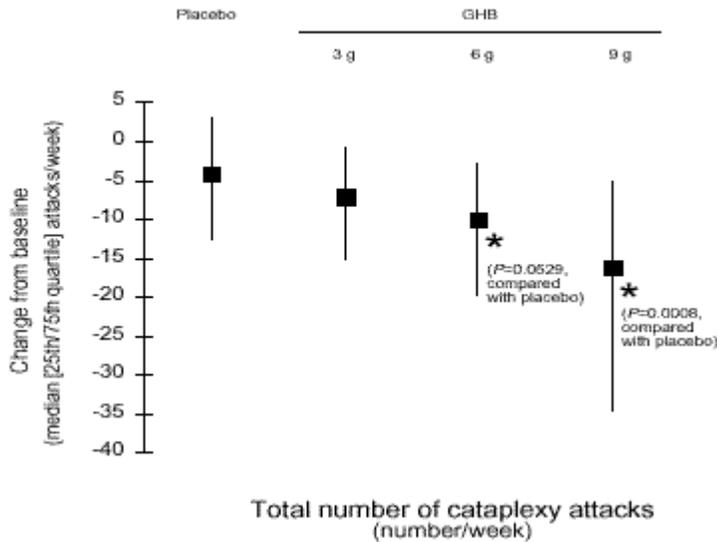
Baseline Clinical Global Impression of Severity results are in the next table, copied from the submission. A statistical analysis comparing treatment groups has not been performed but the placebo group had a distinctly higher proportion of patients in the “extremely ill” category as compared with each of the groups receiving GHB.

Treatment	Normal	Borderline	Slightly ill	Moderately ill	Markedly ill	Extremely ill
Placebo	0	2	2	8	12	10
3g GHB	0	1	1	11	17	4
6g GHB	1	1	0	14	11	6
9g GHB	0	1	2	13	15	4
Total	1	5	5	46	55	24

6.11.7 Primary Efficacy Analysis

The results of the primary efficacy analysis, comparing the change in the total number of cataplexy attacks per week from baseline to endpoint between treatment groups, are presented in the next table and figure, both of which have been copied from the submission. **Note that the results of the analysis are based on ANCOVA using log-transformed data, and not on what was planned as the primary efficacy analysis in the original and amended protocols.**

Dose group	Statistic	Observed		Change from baseline to endpoint	Comparison with placebo (p-value)
		Baseline	Endpoint		
Placebo	N	33	33	33	
	Mean	35.1	24.0	-11.1	
	Median	20.5	16.3	-4.3	
	SD	47.1	28.4	27.7	
	p-value			0.028	
3g	N	33	33	33	
	Mean	28.6	19.5	-9.1	
	Median	20.0	9.5	-7.0	0.5235
	SD	30.5	27.5	22.4	
	p-value			0.026	
6g	N	31	31	31	
	Mean	33.8	24.6	-9.2	
	Median	23.0	8.0	-9.9	0.0529
	SD	45.6	62.9	27.3	
	p-value			0.070	
9g	N	33	33	33	
	Mean	35.7	14.4	-21.3	
	Median	23.5	8.7	-16.1	0.0008
	SD	34.5	19.3	29.8	
	p-value			< 0.001	

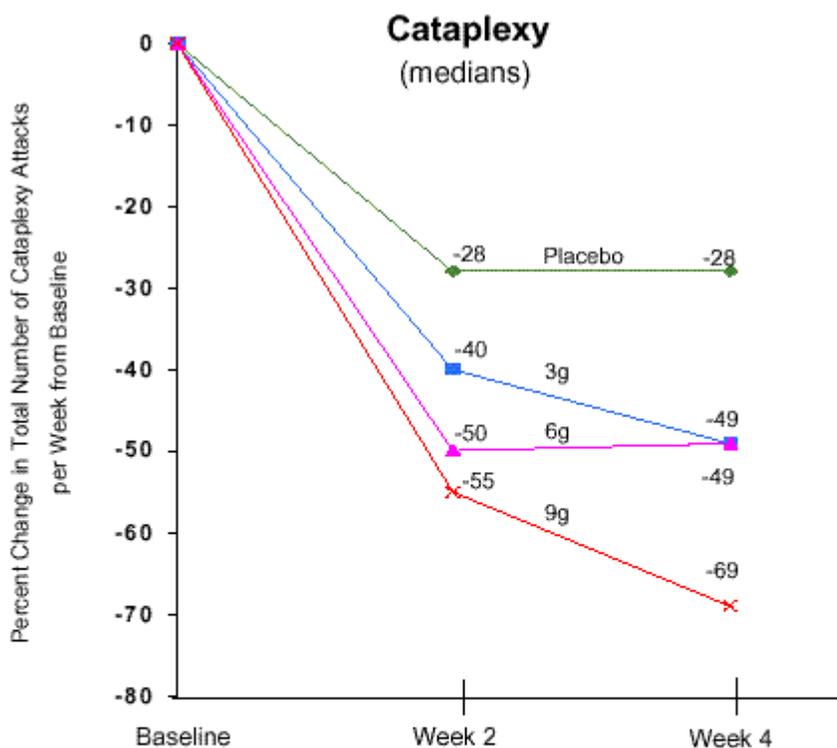


The overall treatment group comparison by ANCOVA was statistically significant at $p = 0.021$

As the table and figure above indicate:

- a dose-response relationship was observed across all treatment groups, based on the median change in total number of cataplexy attacks
- when each GHB dose-group was compared with placebo, the 9 g group showed a definite statistically significant superiority ($p = 0.0008$), and the superiority of the 6 g group approached statistical significance ($p = 0.0529$); however the difference between the 3 g and placebo groups was not statistically significant (the sponsor considers the difference between each of the GHB groups and placebo “clinically meaningful”; this may have a basis if the median change, but not the mean change is considered)

The time-course for the changes in the total number of cataplexy attacks was analyzed by the sponsor as follows. The percentage change in the median total number of cataplexy attacks was calculated based on the distribution of change values for each individual patient. The median of the individual differences is different from the median of the group differences. The mean change in the total number of cataplexy attacks (calculated by this method) at Weeks 2 and 4 for each treatment group is shown in the next figure and accompanying table, which I have copied from the submission.



Treatment Group	Median Baseline Value	Median Week 2 Value	Median % Change from Baseline to Week 2	Median Endpoint Value	Median % Change from Baseline to Week 4
Placebo	20.2	15.0	28.0 %	15.2	28.1 %
3g	20.0	8.75	39.7 %	9.9	49.3 %
6g	23.0	10.0	50.0 %	8.0	49.2 %
9g	23.5	16.69	54.6 %	8.0	68.6 %

The table and figure above indicate that the greater part of the change in all groups was evident by 2 weeks with smaller changes occurring between 2 and 4 weeks.

The within-group change in median from baseline to endpoint, in the total number of cataplexy attacks, was analyzed using the paired-t test. The change in each group was statistically significant as indicated by the following table

Treatment Group	P-value for within-group change from baseline to endpoint
Placebo	0.028
GHB 3 g	0.026
GHB 6 g	0.070
GHB 9 g	<0.001

6.11.8 Secondary Efficacy Analysis

The results of the secondary efficacy analysis are presented, in summary only, below. Most parameters are outlined in the next table copied from the submission.

Parameters	Treatment	Change in Medians from baseline to endpoint	P-value for overall comparison *	P-value GHB group vs placebo
Excessive Daytime Sleepiness (Epworth Scale)	Placebo	-2.0	0.0006	
	3 g	-1.0		0.1137
	6 g	-3.5		0.1860
	9 g	-5.0		0.0001
Frequency of Daytime Sleep Attacks	Placebo	-0.26	0.0101	
	3 g	-0.20		0.1022
	6 g	-0.48		0.0497
	9 g	-0.48		0.0122
Duration of Daytime Sleep Attacks	Placebo	-3.10	0.0282	
	3 g	-5.00		0.9995
	6 g	-9.75		0.4413
	9 g	-7.95		0.0689
Number of Awakenings at Night	Placebo	+0.20	0.0217	
	3 g	-0.25		0.7628
	6 g	-0.21		0.5516
	9 g	-0.91		0.0035
Number of Hypnagogic Hallucinations	Placebo	-0.02	0.3092	
	3 g	-0.10		
	6 g	-0.15		
	9 g	-0.10		
Number of Sleep Paralysis Episodes	Placebo	0.00	0.3326	
	3 g	-0.07		
	6 g	0.00		
	9 g	-0.06		
Total Amount of Sleep Each Night	Placebo	8.57	0.6921	
	3 g	13.66		
	6 g	9.12		
	9 g	9.25		
Quality of Sleep	Placebo	-0.04	0.0009	
	3 g	-0.18		0.2446
	6 g	-0.42		0.0028
	9 g	-0.54		0.0010
Level of Alertness in the Morning	Placebo	0.00	0.0001	
	3 g	-0.13		0.6043
	6 g	-0.49		0.0006
	9 g	-0.42		0.0004
Ability to Concentrate	Placebo	-0.05	0.0012	
	3 g	-0.08		0.5440
	6 g	-0.29		0.0229
	9 g	-0.50		0.0007
Complete Cataplexy Attacks	Placebo	0	0.2131	
	3 g	-1.00		
	6 g	-1.62		
	9 g	-1.62		
Partial Cataplexy Attacks	Placebo	-2.72	0.0032	
	3 g	-3.69		0.5017
	6 g	-6.35		0.1494
	9 g	-10.00		0.0009

* based on ANCOVA

The above analysis indicates that at least one GHB dose group was superior to placebo at a nominally statistically significant ($p < 0.05$) level for the following parameters: excessive daytime sleepiness, frequency of daytime sleep attacks, duration of daytime sleep attacks, number of awakenings at night, partial cataplexy attacks, quality of sleep, level of alertness in the morning, and ability to concentrate. However, given that there were 12 secondary efficacy measures, when adjustment was made for multiple comparisons only the following GHB-placebo comparisons continued to show statistical significance at the same level: Epworth Sleepiness Scale, quality of sleep, level of alertness, and ability to concentrate; for the Epworth Sleepiness Scale, quality of sleep and ability to concentrate it did appear that there was a dose-response with the 9 g/day dose being the most effective; for the Epworth Sleepiness Scale only the 9 g/day dose showed a statistically significant superiority to placebo.

The results of the Clinical Global Impression of Change in the severity of narcolepsy between baseline and endpoint are summarized in the following table which used the original categorical scale, and which I have copied from the submission. The p-value for the overall treatment group comparison was 0.0010 based on the Cochran-Mantel-Haenszel test by Non-zero correlation

Impression	Placebo	GHB Dose (g)		
		3	6	9
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

A dichotomized analysis was then carried out. Responders were those in the “much improved” or “very much improved” original categories; those in all other categories were considered non-responders for purposes of the analysis presented below. The results of this analysis are presented below in a table that was copied from the submission. The p-value is based on Fisher’s exact test

Category	GHB Dose (g)				p-value* (overall comparison)
	Placebo	3	6	9	
Responders	11 (32%)	14 (47%)	16 (52%)	24 (80%)	0.0014
Nonresponders	23 (68%)	16 (53%)	15 (48%)	6 (20%)	
p-value (group vs. Placebo)		0.3075	0.1368	0.0002	

As will be seen from the table above there was a statistically significant difference between the proportions of responders (who are much more numerous) and non-responders in the 9 g dose group

6.11.9 Statistical/Analytical Issues

- The data for a number of outcome measures, including the total number of cataplexy attacks, were not normally distributed; a log transformation was therefore performed
- Missing data were excluded from calculations; 16 patients received study medication but did not complete the study; these were included in the analysis at points prior to their discontinuation from the study
- No inter-site variability was seen
- Since the number of patients treated was small; no formal analyses were conducted between efficacy response and concomitant therapy or between response and past/concurrent illnesses

6.12 Safety Results

These were incorporated into the NDA Safety Review

6.13 Sponsor’s Conclusion

- The 6 g nightly dose of GHB was marginally statistically significantly superior to placebo in regard to the reduction in total number of cataplexy attacks and the number of inadvertent naps during the day
- The 9 g nightly dose of GHB was superior to placebo at a statistically significant level in reducing the total number of cataplexy attacks, number of awakenings during the night, number of inadvertent daytime naps and excessive daytime sleepiness as measured by the Epworth Scale; and in reducing the patient’s overall severity of illness as assessed by the CGI-C

6.14 Reviewer’s Comments

6.14.1 Primary Efficacy Measures

- As noted above the primary efficacy analysis was not performed as specified in the original and amended protocols. Dr Sharon Yan, the statistical reviewer of this submission has however reproduced the per-protocol primary efficacy

analysis using the sponsor's datasets. As the assumptions for ANOVA were not met, she has performed a Wilcoxon Rank Sum test. The p-value for the overall comparison of the 4 treatment groups, using the latter non-parametric test was 0.0101. She then compared each GHB group with placebo and the p-values for each of these comparisons was as follows

Comparison	p-value
GHB 3 g vs placebo	0.4684
GHB 6 g vs placebo	0.1450
GHB 9 g vs placebo	0.0033

- Thus, according to the protocol-specified primary efficacy analysis, only the 9 g/day dose showed a statistically significant superiority to placebo in reducing the total number of cataplexy attacks.

6.14.2 Secondary Efficacy Measures

- In this application the sponsor has sought a claim for Xyrem® in treating daytime sleepiness accompanying narcolepsy.
- The secondary efficacy measures used to assess daytime sleepiness included the Epworth Sleepiness Scale, the frequency of sleep attacks (inadvertent naps) during the day and the duration of daytime sleep attacks (inadvertent naps)
- A nominally statistically significant superiority ($p < 0.05$) of GHB over placebo was seen on the Epworth Sleepiness Scale, the frequency of daytime sleep attacks and the duration of daytime sleep attacks, as measures of excessive daytime sleepiness. However given that there were 12 secondary efficacy measures, only the analysis of the Epworth Sleepiness Scale was still statistically significant after adjustment for multiple comparisons.
- The pairwise comparisons for the Epworth Sleepiness Scale indicated that only the 9 g/day dose of GHB was superior to placebo
- Concomitant stimulant medication was permitted during the study, provided the dose was stable. Although the study was a randomized one, it is unclear to what extent the treatment groups were matched for concomitant stimulant medication and to what extent such medication use confounded the results of the analysis of the Epworth Scale.
- The results of this study, nevertheless, do provide at least some support for the efficacy of GHB in a dose of 9 g/day in treating excessive daytime sleepiness.
- Dr Sharon Yan, Agency Statistical Reviewer finds the analysis of secondary efficacy measures for this study problematical for the following reasons
 - There are many secondary efficacy measures
 - The methods of analysis were not stated in detail a priori

7. Scrima Study

This study was conducted by an individual investigator, Lawrence Scrima, PhD, under IND # 19911 which resides in this Division. The sponsor of this NDA subsequently bought the data for this study.

7.1 Reviewer's Comments Regarding Study Protocol

- A formal study protocol and protocol amendments have not been provided
- In lieu of the study protocol and protocol amendments, the following have been submitted
 - A copy of a grant application from Dr Scrima (#FD-R-000115-01) to this Agency dated February 2, 1985 describing the initial version of the protocol
 - A subsequent FDA letter of approval for this grant application.
 - Multiple subsequent items of correspondence between this Agency and Dr Scrima
 - Multiple subsequent letters between Dr Scrima and the University of Arkansas Medical Sciences Human Research Advisory Committee
- The above items included the detailed statistical analysis plan submitted on December 4, 1986; and further changes to the analysis plan, made in response to FDA comments, in a letter from Dr Scrima dated March 6, 1987. The final prospectively-designated elements of the protocol appear to be in the communications of 2/13/87 and 3/6/87.
- The study itself is reported to have been conducted between 1986 and 1987.
- The Scrima study has been the subject of 2 publications
 - Scrima L, Hartman PG, Johnson FH, Thomas EE, Hiller FC. The effects of gamma-hydroxybutyrate on the sleep of narcolepsy patients: a double-blind study. *Sleep* 1990;13:479-90
 - Scrima L, Hartman PG, Johnson FH, Hiller FC. Efficacy of gamma-hydroxybutyrate versus placebo in treating narcolepsy-cataplexy: double-blind subjective measures. *Biol Psychiatry* 1989;26:331-43
- Case Report Forms were developed retrospectively (i.e., after the study was completed) by Berger-Boyer and Associates as part of a contract with Biocraft, a company that had earlier been developing GHB. The Case Report Forms were developed based on source documents provided by Dr Scrima. Berger-Boyer also performed a statistical analysis that incorporated procedures additional to those reported in the final analysis plan submitted in March 1987. Other elements of the study reviewed by Berger-Boyer and Associates included: study enrollment status, Institutional Review Board, informed consent, drug accountability, clinical findings, results and safety considerations. The Berger-Boyer and Associates analysis was reported in December 1992 (the entire review by this company appears to have been initiated in 1992). The additional elements in this analysis are described below.
- The protocol described below is that outlined in the study report. Any major differences between the analysis plan described in the study report and that designated prospectively are outlined. Major differences between the protocol and study report are also outlined.

7.2 Objective

7.2.1 Primary

To evaluate the following as primary variables in narcolepsy patients during treatment with GHB as compared with treatment with placebo

- Average number of cataplexy attacks as compared with baseline
- Objective sleepiness index as determined by the Multiple Sleep Latency Test as compared with baseline

7.2.2 Secondary

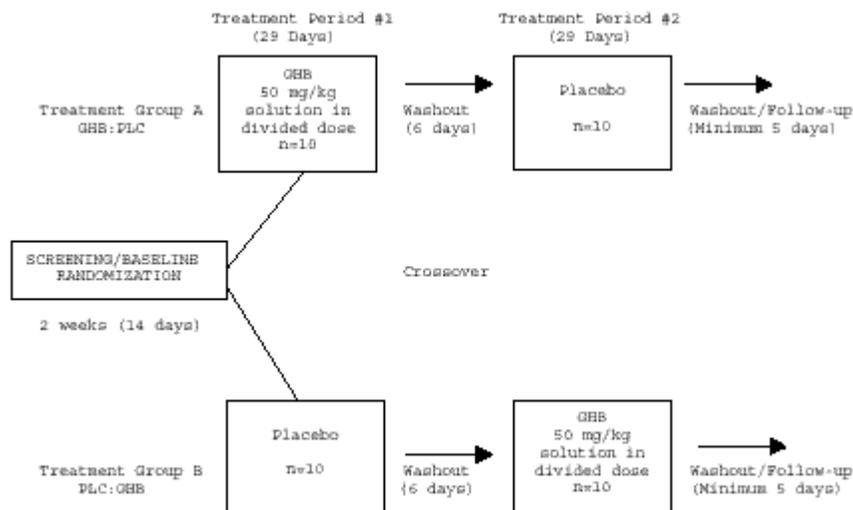
To evaluate the following as secondary variables in narcolepsy patients during treatment with GHB as compared with treatment with placebo (the change from baseline is to be compared in the case of each treatment)

- Average number of sleep attacks per day
- Average number of awakenings per night
- Need for methylphenidate
- Feelings on awakening
- Mood in the morning and evening
- Sleep patterns identified on the polysomnogram
- Average number of REM onsets by Multiple Sleep Latency Test

7.3 Design

Randomized, double-blind, placebo-controlled, single-center, cross-over study comparing the effect of GHB 50 mg/kg total daily dose in with placebo in 20 patients with narcolepsy.

A schematic outline of the study design is presented in the figure below which I have copied from this submission.



Randomization was to be such that half of the men and half of the women participating in the study would receive GHB during the first 29-day double-blind treatment period, and placebo during the second. The remaining patients were to receive GHB first and placebo later.

7.4 Duration

4 weeks of double-blind treatment during each cross-over period

7.5 Dosage

During each period of double-blind treatment, each participating patient was to take

GHB 25 mg/kg at bedtime, and about 3 hours later (total dose: 50 mg/kg/day)

OR

Matching placebo

Both GHB and placebo were to be administered as an oral solution. The study medication was prepared by the pharmacy at the study site (University of Arkansas) from bulk drug

7.6 Sample Size

20 patients (10 male and 10 female)

7.7 Main Inclusion Criteria

- Age: 18-70 years
- Male or female
- Birth control or sterility in females of child-bearing potential
- History of narcolepsy and cataplexy diagnosed by an accredited clinical polysomnographer
- History of cataplexy with ≥ 10 cataplexy attacks over the 2 week baseline period
- ≥ 2 REM onsets on the diagnostic Multiple Sleep Latency Test
- A sleepiness index of ≥ 75 on the Multiple Sleep Latency Test
- Informed consent

7.8 Main Exclusion Criteria

- Lactating females
- Clinically significant abnormalities in medical history or on physical examination or ongoing conditions that interfere with study participation
- Other sleep disorder, except those commonly associated with narcolepsy, including
 - Sleep paralysis
 - Mild-to-moderate obstructive sleep apnea (arterial oxygen saturation $\leq 80\%$)
 - Nocturnal myoclonus
- Use of anti-cataplexy medications for 15 days prior to the baseline Multiple Sleep Latency Test and polysomnogram
- Current use of diuretics

7.9 Concomitant Medications

The following is specifically stated

7.9.1 Prohibited Medications

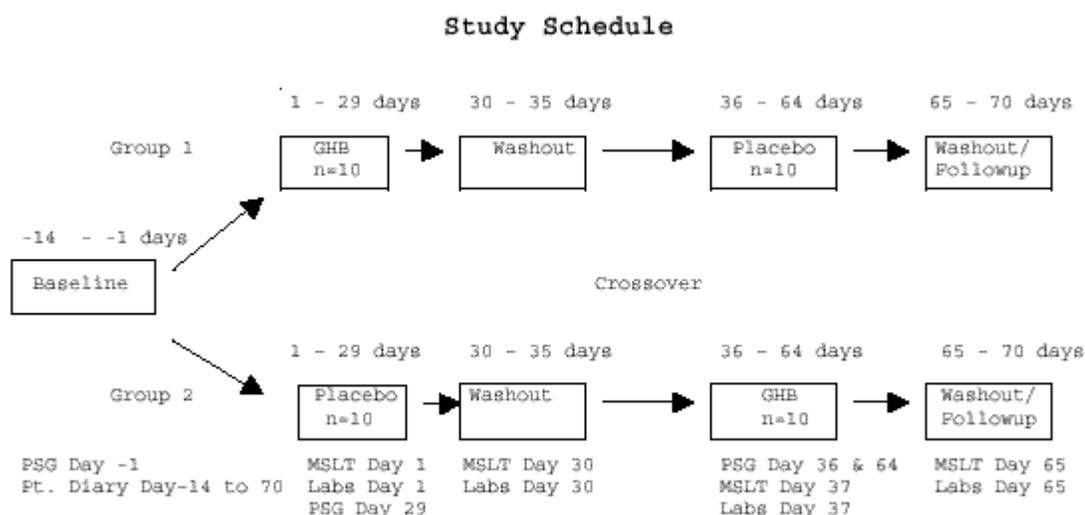
- Other medications for cataplexy (e.g., imipramine or protryptiline) were to be avoided for 15 days prior to the baseline assessment
- Diuretics, alcohol and other sedative hypnotics were to be avoided

7.9.2 Permitted Medications

Methylphenidate in a dose of up to 30 mg/day was to be permitted, but was not to be taken after 6 PM daily

7.10 Schedule

A schematic outlining the schedule for this study is presented below. The schematic is copied from the submission.



In the above schematic

MSLT stands for Multiple Sleep Latency Test
PSG stands for polysomnogram

- Screening assessments consisted of a medical history, sleep disorders interview, physical examination, an overnight polysomnogram and Multiple Sleep Latency Test (the last 2 tests were for diagnostic purposes only)
- Polysomnograms (overnight) were performed on the last day of the baseline period, the first night (Day 1) of each treatment period and the last night (Day 29) of each treatment period (note that this differs somewhat from what is stated in the above schematic)
- Multiple Sleep Latency Tests were performed at some time during the baseline period, on the morning-afternoon after the Day 1 overnight polysomnogram and on Day 29 of each treatment period. On each day the test was performed 5 times, each time over a 20 minute period. The timing of these epochs was as follows: 0800, 1000, 1200, 1400 and 1600 hours.

- A nighttime sleep log and daytime questionnaire were maintained by each patient throughout each study period (baseline, treatment period 1, washout, treatment period 2 and washout/follow-up). These assessed the following: Nighttime sleep onset latency, nighttime arousals, total sleep time, feelings on awakening, number of sleep attacks, number of cataplexy attacks, number of naps, methylphenidate use, sleep/awake patterns, and mood
- Pre-sleep questionnaires were administered before each overnight polysomnogram. These assessed the following: General physical and mental health, food intake, activities on test days, sleep quality, and feelings
- Post-sleep questionnaires were administered on the morning after each overnight polysomnogram, after completion of the Multiple Sleep Latency Test. They assessed the following: Sleep quality, and feelings
- Adverse events were to be recorded continually based on patient interviews, diaries and telephone contacts. The period of observation for these began at the time of obtaining informed consent and extended through at least the first 5 days of the washout/follow-up period at the conclusion of Treatment Period 2
- Standard safety laboratory tests (hematology, chemistry and thyroid functions) were checked at the beginning and end of each treatment period
- During each polysomnogram the following additional items were monitored:
 - Cardiac rhythm by electrocardiogram
 - Respiration using either intercostal EMG or diaphragm bellows
 - Ventilation by nasal O₂ and/or thermocouples
 - Blood arterial O₂ saturation by ear oximetry
 - Cataplexy by EEG, chin-EMG, EOG and electrocardiogram
- Women of child-bearing potential were to have a urine pregnancy test performed 30-45 days and 7 days prior to the baseline polysomnogram.

7.11 Efficacy Outcome Measures

The efficacy outcome measures below are those listed in the study report. The primary outcome measures match those in the final analysis plan submitted by Dr Scrima in December 1986, clarified further in March 1987. The secondary efficacy measures only partly match those listed in the final analysis plan.

7.11.1 Primary Efficacy Measures

- Number of cataplexy attacks per day
- Objective daytime sleepiness as measured by the Multiple Sleep Latency Test* Sleepiness Index

7.11.2 Secondary Efficacy Measures Based On Patient Diaries

- The number of awakenings from sleep at night
- The number of sleep attacks during the day
- Methylphenidate use (mg/day)

- Patient sense of alertness on awakening on a scale from 1 (feeling active, alert, vital and wide awake) to 8 (asleep): this is also referred to as the **Stanford Sleepiness Scale**
Feeling active, vital, alert, or wide awake: 1
Functioning at high levels, but not at peak; able to concentrate: 2
Awake, but relaxed; responsive but not fully alert: 3
Somewhat foggy, let down: 4
Foggy; losing interest in remaining awake; slowed down: 5
Sleepy, woozy, fighting sleep; prefer to lie down: 6
No longer fighting sleep, sleep onset soon; having dream-like thoughts: 7
Asleep: 8
- Mood in the morning recorded on a scale from –10 (extremely negative) to +10 (extremely positive) with 0 representing a neutral mood
- Mood in the evening recorded on a scale from –10 (extremely negative) to +10 (extremely positive) with 0 representing a neutral mood

7.11.3 Secondary Efficacy Measures Based On Polysomnograms

- Sleep efficiency (%)
- Sleep latency (minutes)
- Stage I (%)
- Stage 2 (%)
- Stage 3 (%)
- Stage 4 (%)
- Stage shifts (the number of times during sleep that sleep transitions from one sleep stage to another sleep stage)
- REM sleep (%)
- REM latency (minutes)
- Number of awakenings

*7.11.4 Secondary Efficacy Measures Based On Multiple Sleep Latency Test**

7.11.5 Number of REM onsets

7.11.6 Safety Outcome Measures

Adverse events

Changes in clinical laboratory parameters (adverse events and safety laboratory parameters)

*The **Multiple Sleep Latency Test** consisted of 5 separate sessions on a single day, each session lasting 20 minutes. Measurements in minutes derived from the test included the following:

- Latency to sleep onset (time from lights out to the first epoch of sleep)
- Latency to REM sleep (time from the beginning of sleep onset to the first epoch of REM sleep)
- Average latency to sleep onset (total latencies to sleep/number of naps)
- Number of naps with REM across five naps
- Sleepiness Index = $100 - (5 \times \text{average latency to sleep})$

7.12 Analysis Plan

Defining the analysis plan actually used for the study took considerable effort

7.12.1 General Considerations

7.12.1.1 Efficacy Analyses

- All randomized patients were to be included in the primary and secondary efficacy analyses, except that patients missing an entire treatment period were not included in the post-treatment analysis.
- Patients included in the analysis of the effects of GHB/placebo withdrawal during the washout and washout/follow-up periods must have had both baseline data and at least 3 of the 5 days following treatment periods for each period
- Missing data were substituted for 1 or 2 weeks during the treatment periods or 1 or 2 days during the washout or washout/follow-up periods. Missing data substitutions were made only for the diary data
- All efficacy analyses were to be based on change of data from baseline to post-treatment for each study period

7.12.1.2 Safety Analysis

Any patient who received study medication was to be included in the safety analysis

7.12.1.3 Study Periods

These were defined as follows

- The baseline period was the 2-week pre-treatment period over which patient diaries were recorded. The 5-days preceding the first dose of study medication were considered an adequate baseline for data collected on patients' daily subjective assessments
- GHB/placebo were administered in Treatment Periods 1 and 2 (see schematic in Section 7.3)
- The washout period was the 5-day period after Treatment Period 1
- The follow-up period (also referred to as the washout/follow-up period) was the 5-day period after Treatment Period 2

7.12.1.4 Others

- Sleep stages were converted to percentages of sleep time for analysis purposes
- Diary dates were converted to treatment day as follows
 - The date in the diary for the last day of the baseline period was the first treatment day
 - The date in the diary for the last day of treatment for each treatment period was the day after treatment was stopped
- The criterion for a statistically significant difference was $p < 0.05$ (2-sided?)

7.12.2 Demographic And Baseline Characteristics

- Age, weight, age at diagnosis, and number of sleep and cataplexy attacks were analyzed using a 2-factor ANOVA. The effects in this model were sequence group, gender and the interaction of gender and sequence group

- The distributions of patients with or without histories of hypnagogic hallucinations or sleep paralysis was tested for independence from gender and sequence group using contingency table methods
- Only patients with baseline data who were included in the post-treatment analysis were analyzed for baseline comparability (of narcolepsy-related parameters).
- Note that the above methods for analyzing baseline and demographic characteristics were not specified prospectively.

7.12.3 Primary Efficacy Parameters

- The primary efficacy parameters consisted of the change from baseline scores for the following: the number of daily cataplexy occurrences and the sleepiness index measured from the multiple sleep latency test.
- Change from baseline for the frequency of cataplexy attacks was to be measured as per the final prospectively-designated Scrima analysis plan submitted 12/4/86 as follows
 - The mean daily number of cataplexy attacks was to be calculated from patient diaries for the following periods: the last 2 weeks of the baseline period, the last 2 weeks of the GHB treatment period and the last 2 weeks of the placebo treatment period
 - The change in mean daily cataplexy frequency from the last 2 weeks of the baseline period to the last 2 weeks of the GHB treatment period and to the last 2 weeks of the placebo treatment period was then calculated
- The sponsor states that in Dr Scrima's final analysis plan, the "data called for analyzing only Weeks 1 and 4." What is stated in Dr Scrima's submission of March 6, 1987, is that 2 within patient factors would be included in the ANOVA model: substance and time (1st week versus 4th week for the patient measures, and 1st day versus 28th day for the sleep study measures). If the intention was to use the mean change from baseline to Weeks 1 and 4 in cataplexy frequency as a primary efficacy measure, it is not clearly stated which of these 2 mean changes (i.e., baseline to Week 1 or baseline to Week 4) would be considered the primary efficacy parameter. In the first of 2 publications describing this study (Scrima et al 1989, see Section 7.1) the sponsor states that the change in mean daily cataplexy frequency from baseline to Week 1 and baseline to Week 4 were both used.
- At any rate, the post-hoc Berger-Boyer analysis (see Section 7.12.7) is what was described in the study report. In the study report the following appear to have been done in analyzing the change in frequency of cataplexy attacks
 - The baseline mean daily cataplexy frequency was calculated based on the last 5 days of the baseline period (this is not specified in any prospective version of the analysis plan; all such versions called for the baseline frequency to be calculated using the last 14 days of the baseline period)
 - The mean daily cataplexy frequency was calculated for each of the 4 weeks of treatment with GHB or placebo.
 - The overall mean daily frequency of cataplexy attacks was calculated for each treatment period in their entirety.

- The change from baseline for each treatment modality was then calculated based on the difference between the mean daily number of cataplexy attacks during the baseline period and the mean daily number of cataplexy attacks overall during the GHB and placebo treatment periods. The 2 treatments were then compared
- The change from baseline for the Multiple Sleep Latency Test Sleepiness Index was to be calculated, as per the final prospectively-designated Scrima analysis plan submitted 12/4/86, based on the difference in scores from the Multiple Sleep Latency Test performed on the last day of the baseline period to the last day of the GHB treatment period, and from the last day of the baseline period to the last day of the placebo treatment period. However the analysis presented in the study report appears to imply that the primary analysis of the Sleepiness Index was based upon the difference in mean scores between baseline and the entire period of treatment with either GHB or placebo.
- A multifactorial repeated measures ANOVA was used for this analysis; the “original” model submitted in the final prospective Scrima analysis plan included one between-subject factor (treatment order) and 2 within-subject factors [treatment and time (week)]. In addition to the above factors the following interactions were to be tested: treatment x order; treatment x time; and treatment x substance x time (null hypotheses are stated). The Berger-Boyer analysis (see Section 7.12.7) added a between-patient factor of gender, and all possible interactions with gender; this analysis also pooled various interaction terms related to the within-patient error term.
- Since the term week was frequently significant either as part of a main effect or interaction, further repeated-measures analyses for the individual weeks were performed to support the overall analyses
- In the final prospective Scrima analysis plan, the effect of GHB withdrawal was to be evaluated using a paired t-test contrasting the washout week following withdrawal of GHB with the washout week following withdrawal of placebo. In the actual analysis the effect of GHB withdrawal was assessed as follows
 - The analysis was confined to diary-derived data since only diary-based data were recorded during the washout and washout/follow-up periods
 - One analysis compared the 2 treatment sequences over the initial washout period, in regard to change from baseline. The analysis was based on a repeated-measures ANOVA. The model had a single within-patient factor, days, and two between-patient factors, sequence group and gender. Separate univariate supportive analyses were performed for washout days 1 to 5 with sequence group, gender and their interaction as factors. The intercept was tested in each model to identify departure from baseline
 - A second analysis compared the washout from treatment in Period 1 and the washout/follow-up period after Treatment Period 2. A repeated-measures ANOVA was performed on the change from baseline data. There were two between-patient factors, treatment group and gender, and two within-patient factors, Day 1 to 5 of washout or follow-up

7.12.4 Secondary Efficacy Parameters

- These were analyzed using methods largely similar to the primary efficacy analysis
- Conventional levels of statistical significance (i.e., $p < 0.05$) were retained for the comparisons (Dr Scrima, had in the final prospective analysis plan, stated that the a Bonferroni correction would not be used but the results would have to be interpreted with caution unless replicated)

7.12.5 Safety Parameters

- The GHB and placebo treatment periods were compared in regard to the incidence of adverse events and changes in clinical laboratory results
- Treatment-related adverse events were grouped based on body system, severity and relationship to study drug.
- The clinical laboratory parameters were examined with respect to reported values that were outside the normal range.

7.12.6 Sample Size Rationale

A formal a priori sample size calculation was not performed “as this was the first clinical trial examining GHB in the treatment of narcolepsy”

7.12.7 Berger-Boyer Analysis

This analysis, performed, as noted earlier, in 1992, differed from the original analysis plan as follows

- An analysis of data for Weeks 1, 2, 3, and 4 was performed for the frequency of cataplexy attacks, instead of what was specified prospectively.
- The original analysis plan addressed only the analysis of male patients (it stated that since the male narcoleptic group was reaching completion at that time, the first reports were to be based on a single gender group and gender was not to be included as a factor in the analysis). At the time of the Berger-Boyer analysis, data from both the male and female patients were available and the model used at that time to analyze data from this study included a between-patient factor of gender, as well as all possible interactions with gender
- The model used in the original analysis plan included various interaction terms related to the within-patient error term. The Berger-Boyer analysis pooled these interaction terms so as to simplify the model.

7.13 Protocol Amendments

There were no formal protocol amendments and no comprehensive final version of the protocol. Changes to the protocol have been inferred from correspondence between Dr Scrima and the Agency.

7.14 Efficacy Results

7.14.1 Patient Disposition

- 20 patients (10 men and 10 women) were enrolled in and completed the study

- Of these 10 patients (5 men and 5 women) were randomized to the GHB:placebo sequence and 10 patients (5 men and 5 women) were randomized to the placebo:GHB sequence
- All 20 patients completed the study

7.14.2 Protocol Deviations

- In the case of one 16-year old patient randomized to the placebo:GHB sequence, neither the patient nor her parents signed the original informed consent document; her parents did sign the informed consent attachment for minors and/or fertile females
- 3 patients (1 in the GHB:placebo sequence and the remaining 2 in the placebo:GHB sequence) had a baseline sleepiness index that fell below the minimum required for study entry
- 1 patient randomized to the GHB:placebo sequence took a diuretic throughout the study
- 1 patient randomized to the GHB:placebo sequence continued taking propranolol 40 mg daily for hypertension throughout the study. Propranolol has apparently been used off-label for treating narcolepsy in doses of 80-480 mg/day
- 2 patients (#s 6 and 13; one in each treatment sequence) had no diary record for cataplexy attacks during the pre-treatment period; in one of these instances (a patient, # 13, in the placebo:GHB sequence) the technician noted that the patient had had frequent cataplexy attacks and recorded that she had had 10 cataplexy attacks during the 2-week pre-treatment period, the minimum required to qualify for the study. Another patient in the GHB:placebo sequence (# 11) had attacks that were too frequent to count during the baseline period, and was therefore recorded by the technician to have 10 attacks so as to enable her to enter the study
- 1 patient in the placebo:GHB sequence got drunk during the second treatment period
- Another patient in the placebo:GHB sequence had an extended baseline period because he was initially taking > 100 mg of methylphenidate daily and was slowly tapered to a dose of 30 mg daily. He also had a 4-week washout period between treatment periods. In Weeks 1 and 2 of Treatment Period 1 his methylphenidate dose exceeded protocol guidelines
- Another patient in the placebo:GHB sequence had a 4-week washout period between treatments
- 5 patients in the GHB:placebo sequence and 4 patients in the placebo:GHB sequence took methylphenidate after 6 PM
- 2 patients in the placebo:GHB sequence and 1 patient in the GHB:placebo sequence received their numbers out of sequence

7.14.3 Baseline And Other Demographic Characteristics

These are summarized in the following 2 tables which compares the 2 treatment sequences. There were 5 men and 5 women in each treatment sequence

Variable	GHB:placebo sequence Mean (SD)	Placebo:GHB sequence Mean (SD)
----------	-----------------------------------	-----------------------------------

Variable	GHB:placebo sequence Mean (SD)	Placebo:GHB sequence Mean (SD)
Age (years)	53.2 (9.0)	42.8 (14.9)
Weight (lbs)	199.8 (22.9)	165.9 (27.2)
Age at diagnosis (years)	28.1 (7.8)	25.2 (14.7)
Number of cataplexy attacks per day	3.5 (2.9)	3.0 (2.6)
Sleep attacks per day	2.6 (0.8)	3.1 (1.5)

Variable	GHB:placebo sequence Number of patients	Placebo:GHB sequence Number of patients
History of hypnagogic hallucinations	6	4
History of sleep paralysis	7	6

As the tables above indicate

- The GHB:placebo sequence had an older mean age than the placebo:GHB sequence, although the sponsor points out that the difference was not statistically significant ($p = 0.089$)
- The placebo:GHB sequence had a mean lower weight than the GHB:placebo sequence, and the sponsor acknowledges that the difference was statistically significant ($p = 0.008$)

Analyses of concomitant medications and illnesses are not provided

7.14.4 Medication Compliance

A single patient (#6 participating in the GHB:placebo sequence) failed to bring his supply of study drug to the site on the last day of Treatment Period 1. The pharmacy at the study site prepared the last dose for that study period.

No data are available for the exact dose in grams that each patient received.

7.14.5 Primary Efficacy Analysis

7.14.5.1 Cataplexy Attacks

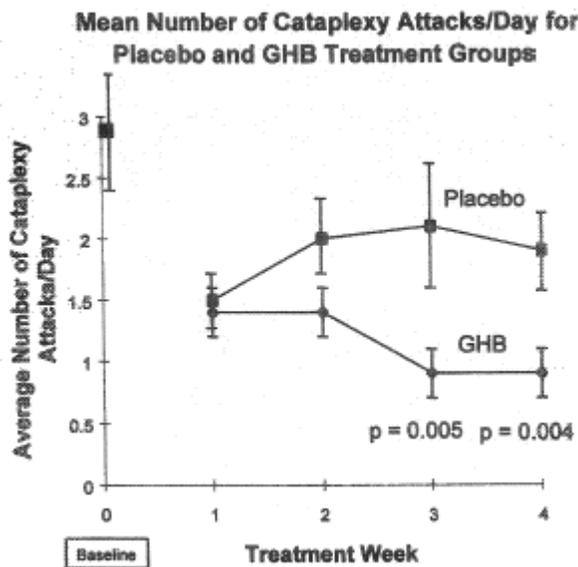
7.14.5.1.1 Baseline Comparability

An analysis of baseline comparability was initially performed, excluding patients 6 and 13 who had no diary records for the baseline period. An analysis of baseline comparability on the remaining 18 patients showed no significant differences between sequence groups, genders nor their interaction for the mean number of daily cataplexy attacks

7.14.5.1.2 Active Treatment Period

Treatment Group	Mean Number of Cataplexy Attacks Per Day						
	Pre-Treatment	Treatment Phase				Overall (SE)	Baseline to Endpoint
	Baseline (SE)	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 analysis of the frequency of cataplexy attacks (the primary efficacy analysis) is summarized in the above table and figure.

The p-value for the overall GHB-placebo comparison was 0.013. The same comparisons for individual weeks were statistically significant at Weeks 3 and 4, although the overall interaction of treatment with week was not statistically significant (p = 0.071). Among other comparisons made by the sponsor: 21% of GHB-treated patients and 5% of placebo-treated patients were cataplexy-free by Week 4.

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); there was thus no evidence of a carry-over effect

As the above table and figure indicate the mean decrease from baseline in the daily frequency of cataplexy attacks in each treatment group was

- 1.7 during GHB treatment (p = 0.007 for change from baseline)
- 1.0 during placebo treatment (p = 0.117 for change from baseline)

7.14.5.1.3 Washout And Washout/Follow-Up Period

Patient #s 6 and 13 were excluded from the analysis of the washout and washout/ follow-up periods, on account of a lack of baseline data and a lack of diary entries, altogether.

For the initial washout period

- Patients taking GHB had an average of 1.0 cataplexy attacks per day versus 1.7 days for those taking placebo
- During the washout period there was a mean decrease from baseline for the GHB group of 1.6 attacks per day (p = 0.054) and for placebo of 1.4 attacks per day (p = 0.065) .
- The difference between treatment groups for mean change from baseline was not statistically significant
- There was a significant increase in attack frequency for both treatment groups from Day 1 to Day 5 of the washout period. By Day 5, 25% of GHB patients and 50% of placebo patients reported as many cataplexy attacks as at baseline.

No statistically significant differences were detected between the 5 days of washout and the 5 days of washout/follow-up overall nor for the interaction with sequence group

7.14.5.2 Sleepiness Index

7.14.5.2.1 Baseline Comparability

No significant differences were detected for baseline comparability between sequence groups, genders or their interaction.

7.14.5.2.2 Active Treatment Period

The results of the analysis are summarized in the following table

Treatment Group	GHB N = 20	Placebo N = 20
Mean Baseline Sleepiness Index	88.5	
Mean Day 1 Sleepiness Index	88.6	90.9
Mean Change From Baseline At Day 1	0.1	2.4
Mean Day 29 Sleepiness Index	85.8	89.6
Mean Change From Baseline At Day 29	-2.7	1.1
Mean Overall Sleepiness Index During Treatment	87.2	90.3
Mean Overall Change From Baseline During Treatment	-1.3	1.8
GHB-Placebo Difference For Overall Treatment Effect	-3.1	
P-value for overall GHB-placebo difference	0.085	

As the table above indicates the difference between the GHB and placebo groups was not statistically significant overall. The decrease in sleepiness index with GHB did appear to be greater at Day 29 than at Day 1. There were however no statistically significant day effects (p = 0.057) or interaction for treatment group x gender (p = 0.081).

7.14.6 Analysis Of Secondary Efficacy Measures

Note that there were a total of 16 secondary efficacy measures recorded in the study report. These were in 2 categories: those based on patient diaries and those based on polysomnogram data

7.14.6.1 Secondary Efficacy Measures Based On Patient Diaries

These analyses are summarized in the next table copied from the submission. The treatment effect (GHB or placebo) listed in the table is the overall effect for 4 weeks of treatment.

The table below indicates that GHB was superior to placebo at a nominal level of statistical significance ((p = 0.042) only in regard to subjective awakenings at night. When adjusted for multiple comparisons this difference was no longer statistically significant.

Summary of Secondary Efficacy Variables from Patient Diaries					
Variable	Baseline (mean)	Treatment Group	Overall (4 Weeks of TX)	P Value Vs. Baseline	P Value Btwn TXs
Sleep Attacks/Day	2.8	PLC	2.1	p=0.007*	0.530
		GHB	1.9	p=0.002*	
Subjective Awakenings at Night	3.0	PLC	3.7	p=0.095	0.042**
		GHB	2.4	p=0.091	
Methylphenidate (mg/day)	12.5	PLC	14.8	p=0.094	0.792
		GHB	14.4	p=0.690	
How Patients Felt Upon Awakening	2.8	PLC	2.7	p=0.472	0.219
		GHB	2.5	p=0.075	
Mood in the Morning	2.0	PLC	2.2	p=0.867	0.142
		GHB	2.8	p=0.348	
Mood in the Evening	0.8	PLC	1.5	p=0.253	0.210
		GHB	2.2	p=0.007*	
TX - Treatment * Significant compared to baseline ** Significant difference between treatments					

7.14.6.2 Secondary Efficacy Measures Based On Polysomnograms

These analyses are summarized in the next table copied from the submission. While the GHB-placebo difference for the overall treatment effect was nominally statistically significant (p < 0.05) for a number of polysomnographic variables, only the effect on Stage 3 sleep (increased with GHB relative to placebo) remained so after adjustment for multiple comparisons

Summary of Secondary Efficacy Variables - PSG Results								
Variable	Baseline	TX Group	Overall Days 1 & 29	Overall Days 1 and 29 Btwn TXs	Day 1	Day 1 Btwn TXs	Day 29	Day 29 Btwn TXs
Sleep Efficiency (%)	83.5	PLC	88.1	p=0.023**	87.8	p=0.019**	88.3	p=0.120
		GHB	84.4		83.5		85.3	
Sleep Latency (min)	4.4	PLC	2.5	p=0.028**	2.5	p=0.310	2.5	p=0.164
		GHB	3.5		3.6		3.4	
Stage 1 Sleep (%)	28.9	PLC	27.0	p=0.042**	25.4	p=0.328	28.4	p=0.026**
		GHB	23.0		21.7		24.3	
Stage 2 Sleep (%)	40.7	PLC	45.4	p=0.732	46.3	p=0.863	44.4	p=0.442
		GHB	44.6		46.7		42.5	
Stage 3 Sleep (%)	3.5	PLC	2.8	p=0.003**	3.2	p=0.360	2.3	p=0.001**
		GHB	5.5		4.3		6.7	
Stage 4 Sleep (%)	4.4	PLC	4.2	p=0.321	3.7	p=0.200	4.7	p=0.623
		GHB	5.4		5.6		5.2	
Stage Shifts	122.2	PLC	129.2	p=0.006**	125.1	p=0.005**	133.2	p=0.051
		GHB	109.9		102.9		116.9	
REM Sleep (%)	23.0	PLC	20.9	p=0.625	21.9	p=0.845	19.9	p=0.577
		GHB	21.5		22.3		20.7	
REM Latency (min)	51.2	PLC	40.6	p=0.133	36.4	p=0.267	44.8	p=0.117
		GHB	24.4		23.6		25.3	
Number of Objective Awakenings	27.2	PLC	26.9	p=0.012**	25.1	p=0.049**	28.6	p=0.042**
		GHB	21.6		20.7		22.8	
		GHB	87.2		4.0		3.4	

TX - Treatment
 ** Significant difference between treatments

7.15 Safety Results

These are described in the NDA Safety Review

7.16 Sponsor's Conclusions

The sponsor's key conclusions may be summarized as follows:

- After 4 weeks of treatment GHB showed a statistically significant superiority to placebo in reducing the frequency of all cataplexy attacks. This effect appeared to be greatest during Weeks 3 and 4 of treatment.
- GHB did not demonstrate a statistically significant superiority to placebo on objective daytime sleepiness as measured by the Multiple Sleep Latency Test Sleepiness Index
- The sponsor has made a number of conclusions regarding efficacy based on
 - Nominally statistically significant results for the GHB-placebo comparisons for secondary efficacy measures
 - Within group changes from baseline for both primary and secondary efficacy measures

7.17 Reviewer's Comments

- The Scrima study has many deficiencies. The main ones are as follows
 - The protocol and its subsequent changes are not presented in an entirely coherent or easily understood form
 - The last prospectively-described modification of the protocol does not clearly define the endpoints for the primary efficacy analysis
 - The primary efficacy analysis described in the study report is post-hoc, and not identical to the prospectively-defined analysis plan

- Some carry-over effects might have occurred during the initial 5-day washout period between treatment sequences (see Section 0)
- Despite the above deficiencies it does appear that, regardless of how the study endpoint was measured, GHB showed an at least nominally statistically significant superiority to placebo in reducing the total frequency of cataplexy attacks
- Dr Sharon Yan, Agency Statistical Reviewer, has conveyed the following to me
 - This study had 2 primary efficacy measures, as already noted
 - It was not specified a priori that a “win” on both measures at a $p < 0.05$ level of significance would be required to declare the study positive
 - Therefore after adjusting for multiple comparisons the level of significance for declaring a “win” on either primary efficacy measure would have to be < 0.025
 - Whereas the sponsor’s analysis of cataplexy attacks showed a drug-placebo difference at a p-value of 0.013, her own analysis has yielded a p-value for the same comparison of 0.0372
 - She is therefore of the view that the study does not have adequate evidence of the efficacy of GHB in treating cataplexy
- The sponsor’s analysis has revealed that GHB did not have a statistically significant superiority to placebo on the Sleepiness Index, calculated from the Multiple Sleep Latency Test, as a measure of excessive daytime sleepiness. Dr Sharon Yan agrees with this conclusion
- Individual absolute doses (i.e., g/day) of GHB are not stated in the protocol. The total nightly dose of GHB was 50 mg/kg as per the protocol. However based on individual weight data listings in the study report
 - Weights of patients participating in the study ranged at entry from 119 to 249 lbs (54.1 to 113.2 kg)
 - Doses of GHB can therefore be estimated to have ranged from 2.7 to 5.7 g/day
 - The mean weight for patients participating in the study was 199.8 lbs (90.8 kg)
 - The mean dose of GHB might therefore be estimated as 4.5 g/day
- Note that in the 2 published reports of this study
 - There was no distinction between the primary and secondary efficacy variables
 - Each report deals with a different set of outcome variables
 - Neither report mentions the Multiple Sleep Latency Test Sleepiness Index as an outcome variable

8. Lammers Study

This study was done outside IND purview in the Netherlands. The results of this study are available both as a study report, as presented in submission # 37 under IND # 49641, and as an article in the medical literature (Lammers GJ et al. Gammahydroxybutyrate and Narcolepsy: A Double-Blind, Placebo-Controlled Study. *Sleep* 1993; 16:216-220). The analysis and conclusions presented in the study report differ somewhat from that in the publication

8.1 Objectives

To demonstrate that GHB in narcoleptics

- Has an effect on REM dissociation phenomena, and on excessive daytime sleepiness
- Has an effect on alertness during the day
- Has a mood-improving effect

8.2 Design

Randomized, double-blind, placebo-controlled, cross-over study

The study consisted of the following phases:

A **baseline** observation period lasting 1 week followed by
2 **treatment** periods each lasting 4 weeks, separated by a
A **washout** period lasting 4 weeks

Note that the last week of the washout period was considered a baseline observation period for the second treatment phase

During the initial treatment period a patient was to be randomly assigned to either GHB or placebo; during the second treatment period that patient would cross over to whatever treatment had not been administered during the first phase. The order in which the patient received either GHB or placebo was therefore randomly decided.

8.3 Inclusion Criteria

- Any race or gender
- Written informed consent
- Combination of sleep attacks during the day, and at least one of the REM dissociation phenomena (cataplexy, hypnagogic hallucinations and sleep paralysis); in case of doubt regarding the diagnosis of narcolepsy, a positive Multiple Sleep Latency Test, as recorded with a 24 hour electroencephalogram, was required

8.4 Exclusion Criteria

- Atypical narcolepsy; narcolepsy was to be considered atypical if
 - The above combination of 2 clinical criteria was absent
 - The Multiple Sleep Latency Test was negative
- Serious liver, kidney or cardiac disorders
- Pregnant patients, or those liable to become pregnant during the study
- One or more serious suicide attempts in the past

8.5 Concomitant Medication

All concomitant medications were to be continued but no alteration of dose was permitted during the study.

8.6 Dosage

GHB 30 mg/kg taken twice each night, with the initial dose at bedtime and the second dose 4 hours later

or matching placebo

The mean daily dose of GHB actually used in the study was 4.75 g/day

8.7 Schedule

- Physician assessments appear to have been planned at recruitment, at the end of the baseline period, at the end of the first treatment period, at the end of the washout period and at the end of the second treatment period (the assessment at the end of the washout period was to serve as the baseline for the second treatment period)
- Telephone contact was to be made between the study physician and patient every week
- A daily diary was to be maintained by the patient during the baseline period, during the last week of the washout period and throughout each treatment period recording the number of cataplexy attacks, the number of daytime sleep attacks, number of awakenings at night, whether refreshed on awaking in the morning, severity of daytime sleepiness, and the number of episodes of hypnagogic hallucinations, sleep paralysis and automatic behavior
- The severity of narcolepsy was assessed at the baseline measurement using a narcolepsy questionnaire, a sample of which was supplied with this submission. A sleep score (Rechtschaffen and Kales) and a vigilance score (Simon and Schulz) were also completed at that time
- A mood rating scale was completed at baseline and during every other week of each treatment period
- A 24 hour cassette electroencephalogram recording and a Multiple Sleep Latency Test were performed at baseline, during the one-week period before the second treatment phase and at the end of each treatment phase
- The opinions of the patient (Global Therapeutic Impression) and clinician (Global Clinical Impression) as to whether improvement had occurred or not were to be obtained at the end of each treatment period.
- Safety monitoring procedures are not specified

8.8 Efficacy Measures

Note that the distinction between primary and secondary efficacy variables is not completely clear in the original protocol. The efficacy measures are listed in the original protocol in the same sequence as they are listed below, but not under separate "Primary" and "Secondary" headings. The protocol however states: "If patients show an improvement on *all* (the emphasis is mine) of the first 3 criteria, this is regarded as a positive result" and "the other variables are of secondary importance and will only be analyzed if effect is found in the primary variables". Based on these statements in the protocol, I have listed certain outcome measures under the "Primary" heading and the remainder under the "Secondary" heading

8.8.1 Primary

- Number of cataplexy attacks per week
- Global Therapeutic Impression (patient): this was scored at the end of the entire study on a hand-written sheet and at the end of each treatment period in the daily diary.

- Global Clinical Impression (clinician)

8.8.2 Secondary

- Number of sleep attacks during the day
- Feeling of sleepiness during the day (using a visual analogue scale)
- Multiple Sleep Latency Test improvement of the two shortest latencies with a minimum of 4 minutes in total
- Stability of alertness during the day (based on electroencephalogram)
- Duration of slow wave sleep
- Decrease in number of phase shifts at night
- Change in mood (?) - this is not consistently specified to be an efficacy parameter in the protocol.

The sponsor states that the secondary efficacy variables were to be analyzed only if an “effect” was found in the primary efficacy variables

8.8.3 Safety Measures

Not specified, except that adverse events were to be recorded in the daily diary supplied to participating subjects. There is no evidence from either the study report or publication that vital signs, safety laboratory tests, electrocardiograms, or physical examinations were checked and recorded.

8.9 Analysis Plan

- **The analysis plan is not presented in a well-organized manner in the version of the protocol that is in the study report, which appears to be largely, but not entirely, a description of the analysis as actually performed; the original protocol (presented in an appendix) contains only the following statements under the heading “Statistical Analysis”:**
“Differences between placebo and gammahydroxybutyrate will be tested by means of the Wilcoxon signed rank test ($\alpha = 0.05$; two-sided). The diary data will be subjected to a trend analysis (including a rank order test for changes within the group and between groups)”
- The median/mean score for the variables was calculated for the baseline observational period that preceded each treatment period, and for each of the weeks during the treatment periods. Next, the change from baseline to study endpoint (Week 4) during the corresponding treatment period was calculated
- The analysis of most of the primary and secondary outcome measures, including the total number of cataplexy attacks, consisted of comparing the median change from baseline to endpoint for the GHB and placebo treatment periods using the Wilcoxon Signed Rank Test with a 2-sided type 1 error of 0.05. For the total number of cataplexy attacks and daytime sleep attacks, data were analyzed for treatment and period effect using the method of Pocock
- The change in mean total and individual item scores on the mood rating scale from baseline was compared for the GHB and placebo periods using the paired t-test

- For each of the efficacy variables, an ANCOVA was also performed using a model appropriate for a cross-over design which included the following factors: treatment order, patient position within treatment order, treatment group, period (?) and baseline value for the efficacy variable; the significance of the covariate was also examined. If the usual assumptions required for ANCOVA were not satisfied, a log transformation of the data was to be considered and the residuals were to be analyzed using the Shapiro-Wilk test. If the data were not normally distributed, a non-parametric test was performed. The ANCOVA was considered an “additional” analysis and was not mentioned at all in the original protocol
- The rating scales for, and methods of analyzing, the Global Clinical Impression and Global Therapeutic Impression are not specified in the original protocol
- “Intent-to-treat” was the only dataset analyzed as per the study report: however the criteria for including patients in this dataset are not defined. In the study report it is stated that an intention-to-treat to treat analysis was also used in the publication. No specific dataset for analysis is specified in the original protocol. The number of patients in the intention-to-treat datasets used for the study report is different from that in the publication, and the further clarified by me under “Dataset Analyzed”, a subheading under “Efficacy Results” below

8.10 Protocol Changes

- Daytime alertness was dropped as a secondary outcome measure, reportedly because the investigators had difficulty defining and measuring alertness during the day
- The Global Clinical Impression was dropped as a primary outcome measure as it was not apparently recorded (“the investigator only reproduced the opinion of the patient and could not translate this information into his own opinion”)
- “Polysomnogram” (presumably the author is referring to 24 hour cassette electroencephalogram recording, as true polysomnographic data was not obtained, as per the study report) data were “no longer available” and were therefore not analyzed.
- **From this reviewer’s perspective all elements of the analysis plan, such as ANCOVA, that were outlined in the study report, but not included in the original protocol as presented in an appendix, may be considered changes in the analysis plan**
- **This reviewer requested the sponsor to clarify whether any formal protocol amendments exist for this study. In submission # 39 under IND # 49641, dated November 6, 1998, the sponsor responded to this request and clarified that to the best of their knowledge and that of the Dutch clinical contractor (Verum Mirai) that assisted in the preparation of the study report for this submission, there were no protocol amendments. The sponsor also clarified that Dr Lammers decided to**

drop several procedures that were specified in the original protocol; I have already outlined these procedures above.

- **In submission # 39 under IND # 49641 (dated 11/6/98) the sponsor further indicated that an ANCOVA was used for the analysis of the data from this study “to maintain consistency with the statistical analysis of the OMC-GHB-2 clinical trial data”.** In this connection I would point out that whereas the results of the Lammers study were first published in 1993 (the actual study period was from November 1987 to December 1988), the protocol for OMC-GHB-2 was finalized (with its last amendment) only in February 1997. Thus the ANCOVA-based analysis of the Lammers data was carried out several years after the data were first unblinded.

8.11 Efficacy Results

8.11.1 Number of Patients and Disposition

- 25 patients were enrolled in and completed both treatment periods. Their randomized distribution, according to treatment sequence, was as follows

First Treatment Period	Second Treatment Period	Number of Patients
GHB	Placebo	13
Placebo	GHB	12

- 1 patient failed to maintain a diary for the first treatment period and was therefore excluded from the intent-to-treat analysis (as well as from a separate analysis for publication). Thus 24 patients were included in the intention-to-treat analysis in the study report
- Another patient was later determined not to have narcolepsy but was included in the intention-to-treat analysis in the study report (a separate intention-to-treat analysis for publication purposes excluded this patient as well; thus the intention-to-treat analysis in the publication had 23 patients)

8.11.2 Protocol Deviations

These include only the two patients referred to above, under “Number of Patients and Disposition”

8.11.3 Dataset Analyzed

- Only an intention-to-treat analysis was carried out, although a definition of this dataset was not provided in the study report or publication
- As indicated above, 24 patients were included in the intention-to-treat analysis that was in the study report and 23 patients in the intention-to-treat analysis in the publication

8.11.4 Demographic Baseline Variables

These are indicated for each treatment order in the next table, copied from the submission. “GHB - Placebo” refers to the sequence where GHB was administered first and placebo second. “Placebo - GHB” refers to the sequence where placebo was administered first and GHB second.

	GHB - Placebo (n=13)	Placebo - GHB (n=12)	Total (n=25)
Age (yrs)	41 {14} ³	39 {15}	40 {14}
Male	8 (32%)	5 (20%)	13 (52%)
Female	5 (20%)	7 (28%)	12 (48%)
Weight (Kg) (n=24) ¹	79 {11}	79 {9}	79 {10}
Height (Cm) (n=24) ²	175 {8}	175 {7}	175 {7}

¹ Weight was not recorded in one patient

² Height was not recorded in one patient

³ Standard deviations are indicated in parentheses

As the table indicates an imbalance of gender was present between sequences with a greater proportion of men in the “GHB - Placebo” sequence and a greater proportion of women in the “Placebo - GHB” sequence. The other demographic variables appear to have been balanced between treatment sequences.

Since this was a cross-over study, the placebo and GHB groups were identical

8.11.5 Primary Efficacy Analysis

This section will describe the intention-to-treat analysis for the following outcome measures: total number of cataplexy attacks per week, and Global Therapeutic Impression.

8.11.5.1 Total Number of Cataplexy Attacks per Week

The following table presents the overall results of this analysis; the results of the study report analysis, and publication analysis are presented in separate rows.

Source	Treatment Group	Median/Mean of Total Number of Cataplexy Attacks per Week *			p-value for Change from Baseline to Endpoint (GHB vs placebo)
		Baseline	Endpoint	Baseline-Endpoint Change	
Study Report (n = 24)	Placebo	5.53	3.01	-2.52	0.002 (ANCOVA)
	GHB	3.99	1.47	-2.52	
Publication (n = 21)	Placebo	1.56	1.24	-0.32	0.42 (Wilcoxon)
	GHB	1.26	0.56	-0.70	

*For the study report, medians are indicated; for the publication, means are indicated

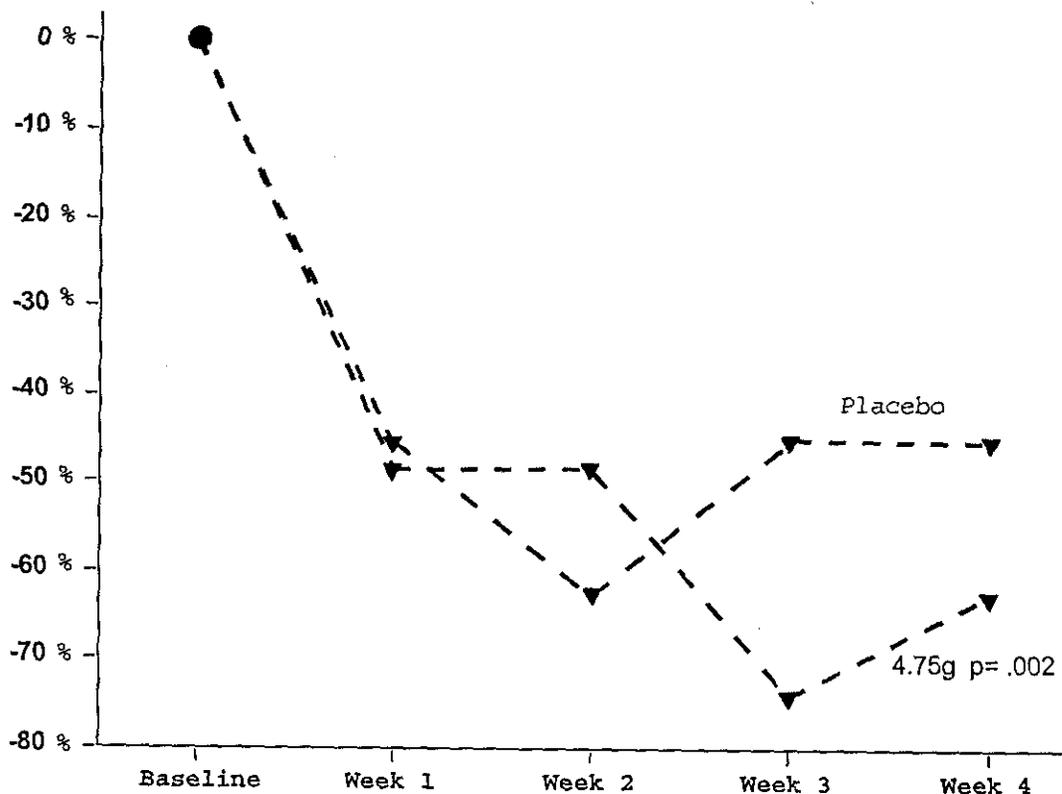
As noted earlier, the primary efficacy analysis as specified in the original protocol was to be based on the Wilcoxon signed rank test and not on ANCOVA; the latter analysis was performed after the study blind was broken.

The next table presents the full details of the ANCOVA used for this outcome measure

Distribution of residual Shapiro-Wilk test	p-values of the factors in the model		
	Treatment	Period	Covariate (Baseline value)

Normal (p = 0.88)	0.002	0.39	0.0001
-------------------	-------	------	--------

The percentage change in the median total number of cataplexy attacks for each week of the study is presented graphically below, for the placebo and GHB groups. The graph is copied from the submission.



The term “4.75 g” in the graph refers to the mean daily dose of GHB.

8.11.5.2 Global Therapeutic Impression

NOTE:

- In the publication, no distinction was made between the Global Therapeutic Impression recorded on a hand-written sheet at the end of the entire study, and that recorded in the daily diary at the end of each treatment period. However the results in the publication appear to correspond to those recorded on a handwritten sheet at the end of the study.
- As noted earlier, the original protocol does not mention how this outcome measure was to be rated or analyzed. In the publication, the Global Therapeutic Impression is described as having been rated on a 4 point scale as in the table below, although for analysis purposes the scale appears to have been dichotomized into “beneficial” or “not beneficial”; in a communication dated November 29, 1998 the sponsor has indicated that the original 4-point rating scale was collapsed so that patients reporting “not beneficial” and “possibly (presumably) beneficial” responses were considered non-responders, whereas those reporting “beneficial” and “strongly beneficial”

responses were considered responders. A similar dichotomous scale has been described in the analysis detailed in the study report

No effect at all	0
Possibly beneficial	1
Beneficial	2
Strongly beneficial	3

The results of the Global Therapeutic Impression, as recorded on a handwritten sheet at the end of the entire study are presented in the table below, copied from the submission

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

As indicated above the results are favorable for GHB in relation to placebo; this difference is statistically significant ($p = 0.001$; McNemar's test)

The results of the Global Therapeutic Impression, as recorded in the daily diary at the end of each treatment period are presented in the table below, again copied from the submission

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

As indicated above the results are again favorable for GHB in relation to placebo; this difference is statistically significant ($p = 0.021$; McNemar's test).

In a communication dated 11/29/98 the sponsor has provided, at the request of the Agency's statisticians, evidence of why McNemar's test is an appropriate analysis in these circumstances.

8.11.6 Secondary Efficacy Analysis

The results for most secondary outcome measures are listed in the following table. Note that the outcome measures listed differ, in part, from those listed in the protocol. With the exception of the "Feeling Refreshed in the Morning" category a negative change represented an improvement. The "Feeling

Refreshed in the Morning” measure was rated on a scale from 0 to 4 with 0 = not and 4 = very good.

Measure	Treatment Group	Median/Mean of Daily Score *			p-value for Change from Baseline to Endpoint (GHB vs placebo)
		Baseline	Endpoint	Baseline-Endpoint Change	
Severity of Daytime Sleepiness (n =24)	Placebo	1.60	1.59	-0.01	0.034 (Wilcoxon)
	GHB	1.60	1.28	-0.32	
Daytime Sleep Attacks (Frequency) (n =24)	Placebo	1.83	2.14	0.31	0.0008 (ANCOVA)
	GHB	2.17	1.36	-0.81	
Nocturnal Awakenings (n =24)	Placebo	2.71	3.31	0.60	0.011 (ANCOVA)
	GHB	3.39	2.00	-1.39	
Refreshed in the Morning (n =24)	Placebo	1.50	1.71	0.21	0.13 (ANCOVA)
	GHB	1.86	2.29	0.43	
Hypnagogic Hallucinations (n =24)	Placebo	0.14	0.00	-0.14	0.056 (ANCOVA)
	GHB	0.21	0.00	-0.21	
Sleep Paralysis (n =23)	Placebo	0.03	0.01	-0.02	0.56 (Wilcoxon)
	GHB	0.07	0.02	-0.05	
Automatic Behavior (n =24)	Placebo	0.27	0.30	0.03	0.23 (Wilcoxon)
	GHB	0.35	0.21	-0.14	

*Median values are indicated for all items other than the severity of daytime sleepiness, sleep paralysis and automatic behavior for which means are indicated

As the above table indicates nominally statistically significant improvement with GHB relative to placebo was seen for the following measures: severity of daytime sleepiness, daytime sleep attacks and nocturnal awakenings; that for hypnagogic hallucinations approached nominal statistical significance.

According to the sponsor, “there were no differences in changes from baseline to Week 4 between the GHB period and placebo period with regard to the total score of the mood rating scale and the separate items” on that scale. The details of this analysis are not presented. All 25 patients were apparently included in the analysis of the mood rating scale.

8.12 Safety Results

These are described further in the NDA Safety Review

8.13 Investigator’s Conclusions

These are summarized below

- Primary efficacy measures
 - GHB showed a statistically significant benefit relative to placebo on the Global Therapeutic Impression
 - GHB showed a statistically significant benefit relative to placebo in reducing the total number of attacks of cataplexy
- Secondary efficacy measures
 - GHB showed a statistically significant benefit relative to placebo on the following measures: severity of daytime sleepiness, number of daytime sleep attacks, and number of nighttime awakenings
 - GHB showed a “marked” benefit relative to placebo on hypnagogic hallucinations

- A statistically significant benefit for GHB relative to placebo was not seen for the following parameters: feeling refreshed in the morning, sleep paralysis, automatic behavior, total score on mood rating scale, scores for individual items on mood rating scale

8.14 Earlier Divisional Comments And Conclusions About Lammers Study

These comments were made at the time Treatment IND # 57271 was reviewed in November-December 1998

- The protocol-specified analysis of one primary efficacy variable, the total number of cataplexy attacks per week, did not reveal any superiority of GHB, in the dose used, to placebo. However, this analysis, the Wilcoxon signed rank test, did not appear appropriate for a cross-over study since it did not examine cross-over and period effects
- Although not listed as a method of analysis in the protocol, an ANCOVA, may have been the appropriate means of analysis for the total number of cataplexy attacks since it could adjust for the baseline differences for the GHB and placebo treatment periods. While this analysis clearly showed a statistically significant superiority of GHB over placebo, it required the assumption that the data were normally distributed. However, the data did not appear normally distributed on visual inspection, even after log transformation; although a formal test did not reject the hypothesis that the data were normally distributed, this test may not have had sufficient power to detect potentially important deviations from normality. In addition the ANCOVA assumed the absence of important carry-over effects; the data suggested that such carry-over effects might be present and that formal testing for normality might not be sufficiently powered to reject such effects. For these reasons it was unclear that the ANCOVA was an appropriate analysis.
- A standard ANOVA used for cross-over study designs yielded a marginal mean difference in the number of cataplexy attacks between treatment groups of -0.359 in favor of GHB ($p = 0.18$)
- Any analysis other than the Wilcoxon signed rank test must be considered post-hoc, although given that the protocol-specified efficacy analysis might not be appropriate, alternative analyses might .
- The other primary efficacy variable that was actually used in the analysis was the Global Therapeutic Impression which showed a statistically significant superiority for GHB over placebo. However it was unclear as to precisely what manifestations of narcolepsy this assessment was measuring
- Based on the above, the Division had expressed the opinion that it was not possible to conclude that the Lammers study was an adequate and well-controlled study contributing reliable evidence of efficacy

8.15 Reviewer's Comments About Secondary Efficacy Measures

- There were at least 9 secondary efficacy measures described in the study report. These were only partly consistent with those described in the protocol

- A nominally statistically significant ($p < 0.05$) improvement with GHB relative to placebo was seen for the following measures: severity of daytime sleepiness, frequency of daytime sleep attacks and nocturnal awakenings
- However, when adjustment was made for multiple comparisons only the effect of GHB on the frequency of daytime sleep attacks remained statistically significant ($p = 0.0008$); this was a protocol-specified outcome measure. This result was however based on ANCOVA which was not a protocol specified analysis. The severity of daytime sleepiness, measured on a visual analogue scale (which was not further described) improved more with GHB than with placebo ($p = 0.034$; Wilcoxon) but this difference did not remain statistically significant (at a $p < 0.05$ level) when adjusted for multiple comparisons
- Note that the sponsor is currently seeking a claim for the use of GHB in treating daytime sleepiness accompanying narcolepsy. The results of this study would, at best, only partly support such a claim.
- Also note that the study protocol called for continuing all concomitant medication during the study, including stimulant drugs, provided the dose was not altered. Since the study had a cross-over design, the effect of GHB on measures of daytime sleepiness is not likely to have been confounded by the concurrent use of stimulant medication unless medication changes were made during the study.

9. Study OMC-SXB-20

This was an open-label study that was intended to evaluate the effects of 4 doses of Xyrem® on sleep architecture. The study report was submitted on 12/16/00, i.e., after the original NDA submission. The sponsor desires that the results of this study be included in labeling.

10. Study OMC-SXB-21

The final report of this efficacy study was submitted on 12/16/00, i.e., after the original NDA submission.

10.1 Tabular Summary

Study #	OMC-SXB-21 Orphan Medical	
Design	Randomized, double-blind, placebo-controlled, parallel-arm, randomized withdrawal study after long-term open label treatment	
Duration	2 weeks (withdrawal phase)	
Study Arms	GHB	Placebo
Number receiving study drug	26	29
Number completed	26	29
Main inclusion criteria	Continuous treatment with GHB for narcolepsy for 6 to 35 months	
Primary outcome measures	Total number of cataplexy attacks	
Main efficacy analysis (statistically significant results)	GHB superior to placebo, based on ANCOVA ($p < 0.001$)	

10.2 Title

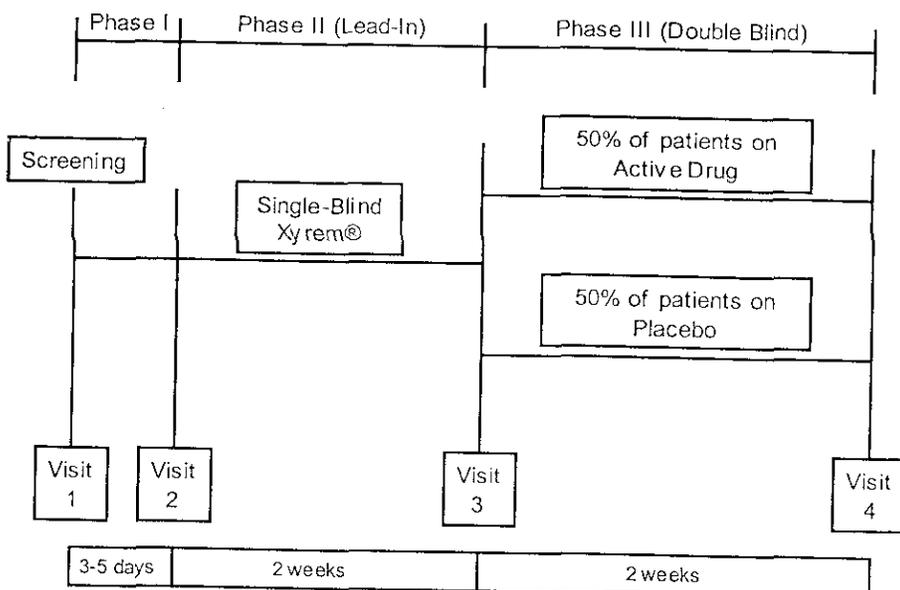
A Randomized, Double-Blind, Placebo-Controlled, Multi-Center Trial To Assess The Long-Term Efficacy Of Orally Administered Xyrem® (Sodium Oxybate) When Compared To Placebo

10.3 Objective

To provide evidence for the long-term efficacy of Xyrem® based upon the return of cataplexy symptoms upon cessation of a minimum of 6 months of open-label treatment with active drug

10.4 Design

The design of the study is schematically summarized below. The schematic is copied from the submission



10.5 Duration

4 weeks (2 weeks of a double-blind withdrawal phase)

10.6 Sample Size

60 patients, with 30 in each treatment group in Phase 3, of the study will be included in the trial

10.7 Selection

10.7.1 Key Inclusion Criteria

- Informed consent
- Age \geq 16 years
- Willing and able to complete the entire trial
- At least 5 cataplexy attacks per week prior to receiving any treatment (tricyclic antidepressants, selective serotonin uptake inhibitors, or Xyrem®) for cataplexy
- If female must be

- Surgically sterile OR
- 2 years post-menopausal OR
- If of child bearing potential must be using a medically accepted means of birth control and must agree to continue such treatment for the duration of the study
- Treated continuously for the symptoms of narcolepsy with Xyrem® for at least 6 months, and not more than 3.5 years
- Willing to not operate a car or heavy machinery if the clinical investigator feels such a restriction is warranted
- Adequate support for the duration of the trial

10.7.2 Key Exclusion Criteria

- Unstable diseases in any body system, other than narcolepsy, which would place the patient at risk or compromise the trial objectives
- Psychiatric disorders that would preclude participation in, or completion of, the trial
- History of substance abuse, as defined by DSM-IV, currently or within the past year
- Serum creatinine > 2 mg/dl; AST or ALT > 2 x upper limit of normal; serum bilirubin > 1.5 x upper limit of normal; pre-trial electrocardiogram results demonstrating a clinically significant arrhythmia or 2nd or 3rd degree A-V block; history of myocardial infarction within the past 6 months
- Any untreated disorder other than narcolepsy that could be considered a primary cause of excessive daytime sleepiness, including sleep apnea syndrome (criteria specified)
- Occupation requiring variable shift or routine night shift work
- Use of tricyclic antidepressants, selective serotonin uptake inhibitors or medications for cataplexy other than Xyrem® in the 30 days prior to Visit 1 of the study
- Clinically significant history of head trauma; previous invasive cranial surgery; seizure disorder; use of anticonvulsant medication

10.7.3 Concomitant Medications

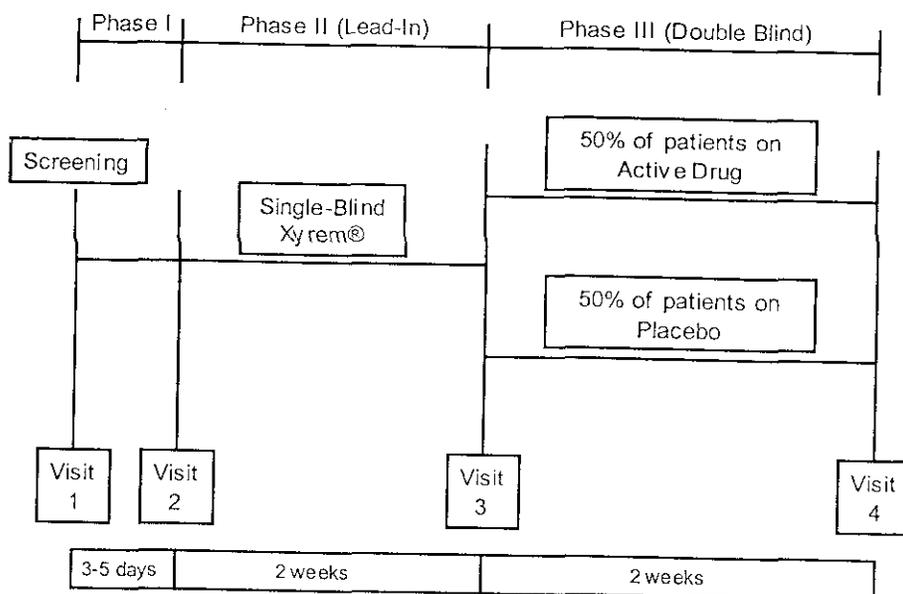
- The following medications are prohibited during the trial: selective serotonin uptake inhibitors and tricyclic antidepressants.
- Patients will be cautioned regarding the use of opioid analgesics and skeletal muscle relaxants
- Alcohol is prohibited during the trial
- Over-the-counter medications need careful review by the clinical investigator prior to use; non-sedating alternatives may be used wherever possible
- Stable doses of stimulant medication may be used to treat excessive daytime sleepiness as clinically indicated

10.8 Dosage

Previously established dose of Xyrem® ranging from 3 to 9 grams daily

10.9 Schedule

The study schematic is reproduced again here for convenience



- The visit schedule was as in the schematic above.
- The following were to be checked at Visit 1 alone: informed consent; selection criteria, medical history, cataplexy history prior to use of any medications, and “support systems”.
- Physical examinations, including neurological examinations were to be performed at Visits 1 and 4
- Daily diaries were to be provided and/or checked at visits 2, 3, and 4. Diaries were to record cataplexy and adverse events.
- Concurrent medications, vital signs and adverse events were to be checked at every visit
- A pregnancy test were to be checked if applicable at Visit 1
- Routine hematology and chemistry were to be checked at Visits 1 and 4

10.10 Outcome Measures

10.10.1 Primary Efficacy

Frequency of cataplexy attacks

10.10.2 Secondary Efficacy

None stated

10.10.3 Safety

Adverse events, laboratory data

10.11 Safety Monitoring

Vital signs, safety laboratory tests, and physical examinations.

10.12 Analysis Plan

10.12.1 Demographic And Baseline Variables

- The 2 double-blind period treatment groups were to be compared in regard to demographic and baseline variables
- Quantitative variables were to be analyzed using either a t-test or a Wilcoxon rank sum test as appropriate
- Qualitative variables were to be analyzed using Fisher's exact test

10.12.2 Primary Efficacy Parameter

- The primary efficacy parameter was the change in the number of cataplexy attacks per week in the 2-week period following Visit 3 (endpoint), compared with the 2-week period prior to Visit 3 (baseline). If a subject withdrew prior to Visit 4 the weekly average would be calculated based upon the data that were available
- The efficacy population was to consist of all those randomized at Visit 2 who had some post-baseline efficacy data
- The above change in the weekly number of cataplexy attacks was to be analyzed using a non-parametric ANCOVA as follows
 - The baseline number of cataplexy attacks and the change in the weekly number of cataplexy attacks were to be replaced by their corresponding ranks (mean ranks will be used when ties occur).
 - The ANCOVA would be constructed from the residuals derived from the ordinary least squares prediction of the change in the weekly number of cataplexy attacks based on a simple linear model
 - The treatment groups would then be compared with respect to these residuals using the Wilcoxon rank sum test.
 - Prior to completion of the analysis a test would be performed to compare the slopes for the 2 treatment groups.
- The significance of the mean change from baseline for each treatment group would be determined using the Wilcoxon signed rank test

10.12.3 Safety Parameters

- The safety population would consist of all those randomized to receive drug at Visit 3 who had some post-baseline safety data
- Adverse events would be summarized by treatment group and organized by preferred term and body system. Treatment groups would be compared to the incidence of each adverse event using Fisher's exact test
- Laboratory data would be summarized in tabular form as well as with the use of shift tables. Treatment groups would be compared in regard to the mean change from baseline using ANOVA. Within each treatment group the significance of the mean change from baseline will be analyzed using a paired t-test

10.12.4 Sample Size Rationale

- The sample size calculation was based on the change in weekly cataplexy attacks comparing the 2 weeks prior to randomization and the 2 weeks after randomization

- The assumptions for the sample size calculation were as follows
 - Power of 80 %
 - 2-sided α of 0.05
 - A 50 % increase in the total number of cataplexy attacks in the placebo group, and a 10 % increase in a Xyrem® group
 - A standard deviation, based on a log transformation, of about 0.30 for the change in total number of cataplexy attacks (based on a previous study)
- Based on the above, a sample size of 22 patients would be required per treatment group to detect a treatment difference.
- To allow for a minor departure from the above assumptions a total of 30 patients would be randomized to each treatment group

10.13 Protocol Amendments

These have been incorporated into the above

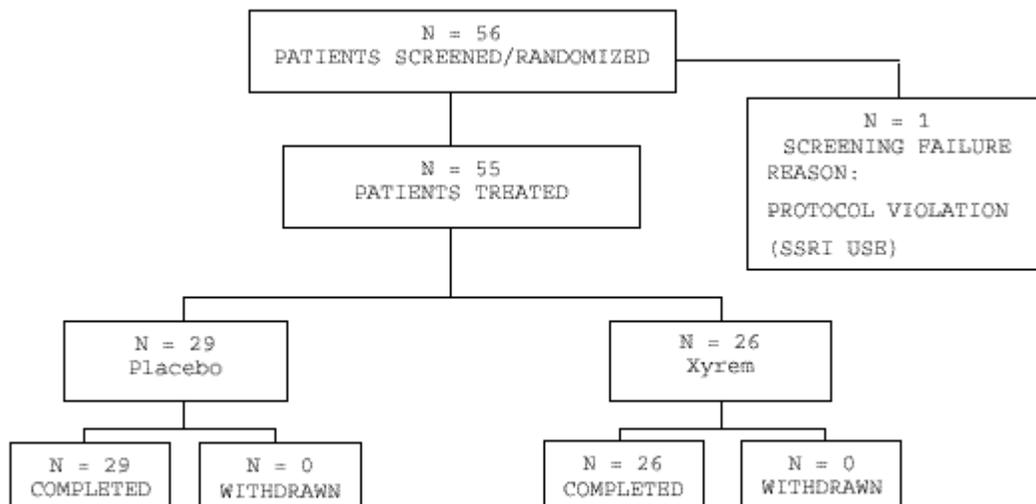
10.14 Actual Analyses Performed

10.15 Efficacy Results

The study was conducted at 14 centers. Each center enrolled between 1 and 7 patients

10.15.1 Patient Disposition

Patient disposition is summarized in the following schematic copied from the submission



Note that 1 randomized patient failed screening because of concomitant use of a selective serotonin re-uptake inhibitor (paroxetine). The blind was broken on 1 patient shortly after completion of the trial on account of a serious adverse event.

10.15.2 Protocol Deviations

- One patient was allowed into the trial despite having been treated with GHB for 3.7 years (the inclusion criteria specified that the duration of treatment should be from 0.5 to 3.5 years)
- One patient was allowed to continue in the trial despite receiving bupropion as a medication for cataplexy
- 3 patients overmedicated
- For “efficiency” 2 patients who were taking 3 g/day at study entry and continued to take that dose during the study were listed as taking 4.5 g/day
- For a number of patients Visits 1 and 2 were combined.

10.15.3 Medication Compliance

As the following table, copied from the submission, indicates medication compliance was comparable for the 2 Phase III treatment groups

Trial Medication Administration	Xyrem (N=26)			Placebo (N=29)		
	Phase II	Phase III	Total	Phase II*	Phase III	Total
Days Treated						
11	0	2		0	3	
12	1	1		0	0	
13	1	5		4	5	
14	14	13		20	13	
15	4	3		0	6	
16	1	0		0	0	
17	4	1		4	1	
18	1	1		1	1	
Duration of Treatment (Nights)						
Mean	14.7 ± 1.43	13.9 ± 1.48	28.6 ± 2.50	14.4 ± 1.35	14.0 ± 1.50	28.4 ± 1.95
Range	12-18	11-18	24-36	13-18	11-18	24-36
Compliance (%)						
Mean ± SD	105.9 ± 17.24	106.1 ± 18.80	106.0 ± 17.44	99.7 ± 6.07	102.4 ± 15.12	101.1 ± 9.28
Range	95-171	85-185	90-178	85-119	72-167	82-139

* Placebo group patients received Xyrem during Phase II.
 SD = Standard deviation.

10.15.4 Baseline And Other Demographic Characteristics

These characteristics are summarized in the next 2 tables copied from this submission. Although gender, and baseline frequency of cataplexy attacks were not entirely balanced between the treatment groups the sponsor describes the differences as not being statistically significant. Note that the daily dose of Xyrem® did appear balanced between the Phase III treatment groups.

Characteristics	Total	Treatment Group		p-Value
	(N=55)	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%)	
Weight (kg)				
Mean ± SD	80.5 ± 20.09	83.8 ± 24.31	77.6 ± 15.22	0.250
Range	54.0 - 142.0	54.0 - 142.0	55.0 - 127.0	
Height (cm)				
Mean ± SD	170.1 ± 10.25	169.6 ± 10.42	170.6 ± 10.24	0.710
Range	152.0 - 188.0	152.0 - 188.0	155.0 - 188.0	
Race (n, %)				
Caucasian	52 (95%)	23 (88%)	29 (100%)	0.099
African-American	2 (4%)	2 (8%)	0	
Asian	0	0	0	
Hispanic	1 (2%)	1 (4%)	0	
Other	0	0	0	
Time on Xyrem (months)				
Mean ± SD	21.22 ± 12.28	23.27 ± 12.36	19.38 ± 12.13	ND
Range	7 - 44	8 - 38	7 - 44	

(continued)

Characteristics	Total	Treatment Group		p-Value
	(N=55)	Xyrem (N=26)	Placebo (N=29)	
Cataplexy attacks (2-week baseline)				
N	55	26	29	0.436
Mean	12.6	9.0	15.7	
SD	31.75	19.25	39.88	
Median	3.0	1.9	4.0	
Minimum	0.0	0.0	0.0	
Maximum	197.0	86.8	197.0	
Daily 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. SD = Standard deviation.

10.15.5 Primary Efficacy Analysis

An intent-to-treat analysis was performed as specified in the protocol comprising all patients who received one or more doses of trial medication during the double blind withdrawal period and had recorded baseline and post-baseline efficacy measures

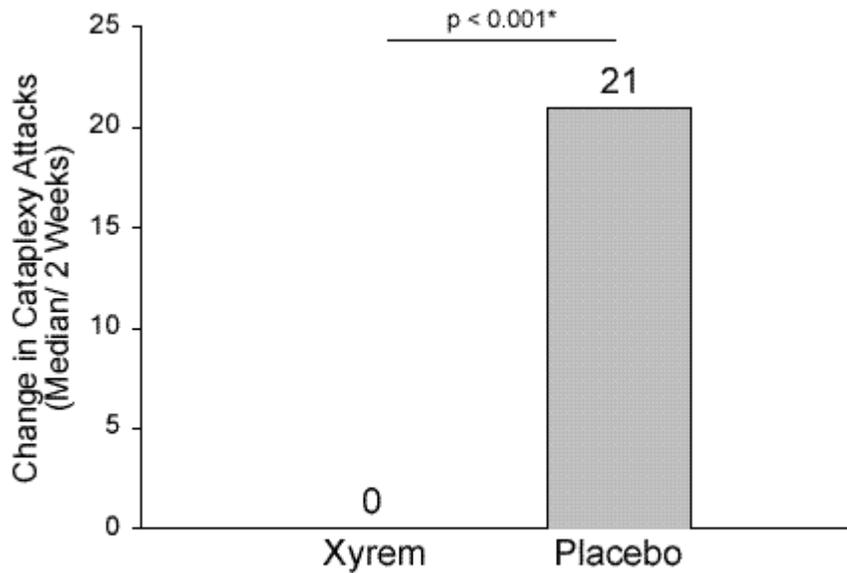
The results of the primary efficacy analysis are outlined in the table and figure below, which I have copied from the submission. For those receiving Xyrem® during the double-blind withdrawal phase there was no median change from baseline in the number of cataplexy attacks over the 2 week period of withdrawal. For those receiving placebo during the withdrawal phase the median change in the number of cataplexy attacks during as compared with baseline showed an increase. The difference was statistically significant (p < 0.001). Note that the table and figure below depict median change

	Xyrem (N=26)			Placebo (N=29)		
	Phase II	Phase III	Change	Phase II*	Phase III	Change
Number of cataplexy attacks (per 2 weeks)						
Mean ± SD	9.0 ± 19.25	12.6 ± 30.34	3.6 ± 20.73	15.7 ± 39.88	50.4 ± 81.09	34.6 ± 55.72
Median	1.9	1.1	0.0	4.0	21.0	21.0
Minimum	0.0	0.0	-24.3	0.0	0.0	-15.0
Maximum	86.8	138.3	87.2	197.0	269.2	206.2
Rank change						
Mean ± SD			18.1 ± 12.65			36.9 ± 13.31*
Median			16.5			39.0
Minimum			1.0			3.0
Maximum			52.0			55.0

SD = standard deviation.

* Placebo group patients received Xyrem during Phase II.

* p < 0.001, from ANCOVA model containing rank baseline, treatment group, and baseline-by-treatment group interaction.

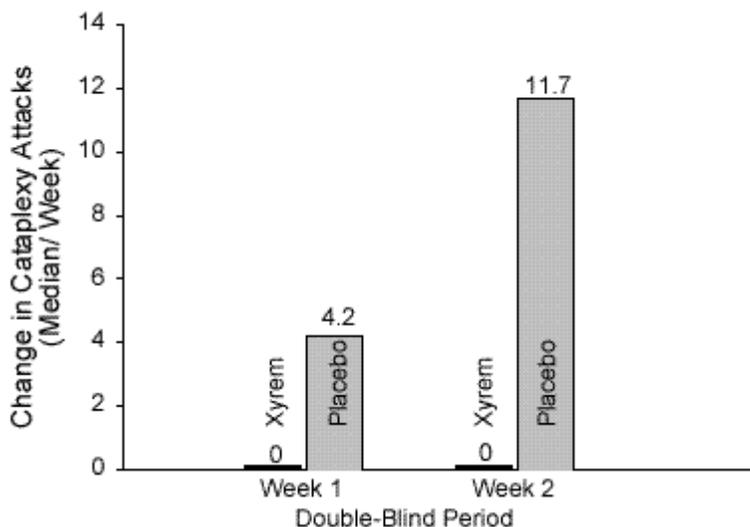


* p < 0.001, from ANCOVA model containing rank baseline, treatment group, and baseline-by-treatment group interaction.

As the next table and figure, copied from the submission, indicate the median change from baseline by week in the number of cataplexy attacks mirrors that for the primary efficacy analysis above

Number of Cataplexy Attacks	Xyrem			Placebo		
	Phase II*	Phase III	Change	Phase II*	Phase III	Change
Week 1						
Number of Patients	26	26	26	29	29	29
Mean ± SD	4.5 ± 9.62	5.3 ± 11.94	0.8 ± 7.48	7.9 ± 19.94	21.1 ± 35.13	13.2 ± 22.02
Median	0.9	1.0	0.0	2.0	7.0	4.2
Minimum	0.0	0.0	-15.4	0.0	0.0	-7.5
Maximum	43.4	50.8	25.2	98.5	126.0	87.5
Week 2						
Number of Patients	26	26	26	29	29	29
Mean ± SD	4.5 ± 9.62	7.2 ± 18.66	2.7 ± 13.74	7.9 ± 19.94	29.7 ± 47.30	21.8 ± 35.16
Median	0.9	0.5	0.0	2.0	13.0	11.7
Minimum	0.0	0.0	-10.7	0.0	0.0	-7.5
Maximum	43.4	87.5	62.0	98.5	168.0	143.5

* Baseline (Phase II) was determined by normalizing the total number of cataplexy attacks during the 2-week Phase II period to 7 days.



No formal analyses were carried out to evaluate differential effects at study sites, or to evaluate drug-drug or drug-disease interactions.

10.15.6 Analysis Of Secondary Efficacy Measures

This study had no secondary efficacy measures

10.16 Safety Results

These are summarized in the NDA Safety Review.

10.17 Sponsor's Conclusions Regarding Efficacy

Xyrem® is effective as a long-term treatment for cataplexy

10.18 Reviewer's Comments

- The design and analysis plan for this study were discussed at length a priori with this Division
- The Division had agreed earlier that the randomized withdrawal paradigm used in this study would be appropriate for demonstrating the long-term efficacy of Xyrem®. Based on that agreement and the study results, I would agree with the sponsor's conclusion that this study provides evidence for the long-term efficacy of Xyrem® in the treatment of cataplexy.
- Dr Sharon Yan, Agency statistical reviewer, has informed me that she also agrees with the sponsor's conclusion that the study provides evidence for the long-term efficacy of GHB in treating cataplexy.
- An effective daily dose of Xyrem® is difficult to determine from this study since patients were not randomized to separate Xyrem® dose groups prior to withdrawal. However, it is noteworthy that although the dose of Xyrem® used in this study ranged from 3 to 9 g/day, 80% of those enrolled were receiving doses of 6-9 g/day at study entry.

11. Overall Comments Regarding Efficacy

- The sponsor is seeking a claim for Xyrem® as a treatment for cataplexy and daytime sleepiness accompanying narcolepsy
- The evidence for the efficacy of GHB in treating cataplexy may be summarized as follows
 - There does appear to be evidence that GHB is effective in treating cataplexy
 - Efficacy is supported by the results of Studies OMC-GHB-2 and OMC-SXB-21, and to a lesser extent by the Scrima study which has a number of deficiencies. In all 3 studies the same outcome measure, the frequency of cataplexy attacks based on patient diaries, was used. The Lammers study must be considered a “negative” one at this time.
 - The effective dose of GHB in treating cataplexy can be best defined from the OMC-GHB-2 study in which patients were randomized to specific doses of GHB and clear evidence of efficacy was seen only at a dose of 9 g/day (and not at 3 g/day and 6 g/day). In the OMC-SXB-21 and Scrima studies, there was no randomization by GHB dose: in the OMC-SXB-21 study 80% of patients had been taking GHB doses \geq 6 g/day prior to randomized withdrawal. In the less than optimal Scrima study the protocol-specified dose was 50 mg/kg/day, and the mean daily dose estimated using body weight data was 4.5 g/day. Thus the most clearly effective dose in treating cataplexy was 9 g/day with less clear and consistent evidence of efficacy at lower doses up to about 4.5 g/day. There is no evidence for efficacy at a dose of 3 g/day
- The evidence for efficacy of GHB in treating daytime sleepiness accompanying narcolepsy may be summarized as follows
 - Efficacy is supported by analysis of the Epworth Sleepiness Scale in the OMC-GHB-2 study, and to a lesser extent by the analysis of the frequency of daytime sleep attacks in the Lammers study which had a number of inadequacies
 - However the analysis of a number of other measures of daytime sleepiness in 3 efficacy studies could not be considered to show a statistically significant superiority of GHB over placebo. These included the following
 - The frequency of daytime sleep attacks and the duration of such attacks in the OMC-GHB-2 study
 - The Sleepiness Index (of the Multiple Sleep Latency Test), which was a primary efficacy measure, and the frequency of daytime sleep attacks, in the Scrima study
 - The severity of daytime sleep attacks in the Lammers study
 - In the OMC-GHB-2 study the only seemingly effective dose in treating daytime sleepiness was 9 g/day. In the Lammers study the mean daily dose used was 4.75 g/day
 - It is unclear to what extent the analysis of the Epworth Sleepiness Scale data in OMC-GHB-2 was confounded by the concurrent use of stimulant medication (it is unclear to what extent the treatment groups were matched in this regard)
 - The lack of replication of the effect of GHB on daytime sleepiness as assessed by a specific measure in more than one study is unsatisfactory, quite apart from the other deficiencies noted in the efficacy studies
 - Currently, there does NOT appear to be adequate evidence that GHB is effective in treating daytime sleepiness accompanying narcolepsy.
- In summary,
 - Evidence has been provided in this application that Xyrem® is effective in treating cataplexy. The evidence is best at a Xyrem® dose of 9 g/day

- The application does not however provide adequate evidence that Xyrem® is effective at treating daytime sleepiness accompanying narcolepsy

12. Labeling Review

Pending as a separate document

13. Overall Comments Regarding Safety Of Xyrem®

See NDA Safety Review for full details

13.1 Clinical Safety

- When GHB is used to treat narcolepsy in doses of 3-9 g/day the most common, dose-related, and seemingly drug-related, adverse events have included the following: headache, unspecified pain, nausea and dizziness. Urinary incontinence is slightly less common, but apparently dose and drug-related as well. More serious, but much less common, adverse events seen at the same dose range, and that could be attributed to Xyrem®, have included vomiting, confusion, restlessness, agitation, paranoia, hallucinations, somnolence and generalized weakness. No deaths that could be attributed to study drug have been reported at therapeutic doses of GHB
- One healthy 39 year old woman participating in pharmacokinetic trials developed dizziness, nausea, vomiting, respiratory depression and fecal incontinence, after a single (and initial) oral dose of 4.5 g of GHB.
- A single older narcoleptic patient who had been taking GHB for approximately 1 ½ years was hospitalized after an overdose of GHB 18 g. At the time of hospitalization he was comatose and unresponsive. He needed intubation and artificial ventilation, and awoke 6 hours later. This incident suggests that the safety margin between therapeutic and toxic doses may not be very wide
- At therapeutic doses of GHB all adverse events appear to be reversible
- While currently there is no strong evidence that GHB in therapeutic doses is epileptogenic or that episodes of urinary and fecal incontinence due to GHB are due to seizures, there is insufficient data at present to rule out either possibility.
- “Recreational” use of GHB generally at doses presumed or known to be higher than the therapeutic has been associated with adverse events that included fatalities attributable to the depressant effects of this drug on the nervous system. However concurrent use of alcohol and of other drugs with effects on the central nervous system has been reported in many of these instances
- There is no evidence that GHB is toxic to any major organ other than the nervous system.

13.2 Withdrawal Phenomena And Abuse Potential

- There is no evidence from a small formal study with a randomized withdrawal paradigm (OMC-SXB-21) that the abrupt discontinuation of therapeutic doses of GHB used for 6 months to 3 ½ years leads to more than mild and

infrequent withdrawal symptoms, except for a significantly increased frequency of cataplexy.

- There are however a number of anecdotal reports of an actual withdrawal syndrome and, possibly, addiction in illicit “recreational” users of GHB, GBL or 1-4 BD. In all these individuals, high doses of GHB or related drugs were believed to have been used at frequent intervals around-the-clock.

13.3 Risk Management Program

Final comments pending

14. Risk-Benefit Equation

Final comments pending

15. Study Site Inspections

Comments pending

16. Financial Disclosure Certification

Financial disclosure certification has been submitted with this application.

16.1 Components Of Certification

This certification has 2 components

16.1.1 Certification Pertinent To Dr Lawrence Scrima

The sponsor has supplied required financial disclosure information for Dr Scrima.

Orphan Medical, Inc, entered into a financial contract with Dr Scrima on 11/10/99. The contract allowed Orphan Medical to access documentation associated with the double-blind, placebo-controlled, cross-over trial in 20 narcoleptic patients. The trial was conducted from April 5, 1986 to December 14, 1987.

The sponsor states that payments to Dr Scrima were made over 10 years after completion of the trial. While the payment was financially disclosable it did not have any impact on data collection, interpretation or analysis

16.1.2 Certification Pertinent To Other Investigators

The sponsor has supplied a list of 32 Investigators who conducted clinical trials on behalf of Orphan Medical, Inc. In regard to this list the sponsor has

- Certified that it has not entered into any financial agreement with the clinical investigators listed in the application whereby the compensation to the investigator could be affected by the outcome of the study in which the investigator was a participant, as defined by 21 CFR 54.2 (a)
- Certified that each listed clinical investigator required to disclose to the sponsor that whether the investigator had a proprietary interest in this product or a significant equity in the sponsor as defined in 21 CFR 54.2 (b) did not disclose any such arrangements

- Certified that no listed investigator was the recipient of significant payments of other sorts as defined in 21 CFR 54.2 (f)

16.2 Reviewer's Comment

It appears unlikely that the financial arrangement disclosed above introduced significant bias into the results of studies carried out with Xyrem®, and submitted with this NDA.

17. Conclusions

Final conclusions pending

18. Recommendations

Pending

Ranjit B. Mani, M.D.
Medical Reviewer

J. Feeney, M.D. _____

rbm 2/14/01
cc:
HFD-120
NDA 21196
Homonnay