CENTER FOR DRUG EVALUATION AND RESEARCH

APPROVAL PACKAGE FOR:

APPLICATION NUMBER

21-487

Statistical Review(s)

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
FOOD AND DRUG ADMINISTRATION
CENTER FOR DRUG EVALUATION AND RESEARCH
OFFICE OF PHARMACOEPIDEMIOLOGY AND STATISTICAL SCIENCE
OFFICE OF BIOSTATISTICS

STATISTICAL REVIEW AND EVALUATION

Clinical Studies

NDA/Serial Number:	21-487
Drug Name:	Memantine
Indication:	Moderate to Severe Alzheimer's Disease
Sponsor:	Forest Laboratories
Date:	12/19/2002
Review Priority:	Standard
Biometrics Division:	Biometrics I (HFD 710)
Statistical Reviewer:	Tristan Massie
Concurring Reviewers:	Kun Jin, Ph.D.,Team Leader George Chi, Ph.D., Director
Medical Division:	Neuropharmacological Drug Products (HFD 120)
Clinical Team:	Ranjit Mani, M.D., Clinical Reviewer Armando Oliva, M.D., Team Leader Russell Katz, M.D., Director
Project Manager:	Melina Griffis, R. Ph.

Keywords: ANCOVA, Alzheimer's, Memantine

TABLE OF CONTENTS

Statistical Review and Evaluation

1. Executive Summary

1.1. Conclusions and Recommendations

Except for a marginally significant result on one of the two primary endpoints in one of the two main studies, the data supports the sponsor's efficacy claim.

1.2. Brief Overview of Clinical Studies

There are three studies of primary interest for the evaluation of efficacy in this application:

MEM-MD-02 was a 24 week, randomized, double-blind, placebo controlled study conducted in the U.S. It compared the therapeutic effects of the combination of Memantine and Donepezil versus the combination of Placebo and Donepezil for patients with moderately severe to severe Alzheimer's disease on stable doses of Donepezil.

Study 9605 was a 28-week, multicenter, randomized, double-blind, placebo controlled trial designed to assess the efficacy and tolerability of Memantine in patients with moderate to severe Alzheimer's disease. It was conducted among 32 centers in the US and was designed to enroll approximately 250 patients at least 50 years of age.

Study 9403 was a 12-week, multicenter, randomized, double-blind, placebo controlled trial designed to assess the efficacy and tolerability of Memantine in patients with moderate to severe Alzheimer's disease. It was conducted among 7 centers in Latvia and was designed to enroll approximately 150 patients at least 60 years of age.

In each of studies MEM-MD-02 and 9605, approximately 90% of the patients were Caucasian, about 2/3 were female, and the average age was about 76. In study 9403, information on race was not collected, but 58% were female, and the average age was 72.

1.3. Statistical Issues and Findings

Both studies 9605 and MEM-MD-02 exhibited statistical significance for the changes from baseline in the Severe Impairment Battery (SIB) and the modified Alzheimer's Disease Cooperative Study - Activities of Daily Living (ADCS-ADL) Total scores. These were co-primary endpoints for study MEM-MD-02, but the ADCS-ADL and CIBIC-Plus were co-primaries in study 9605, where the SIB was a secondary endpoint. In study 9605 the pre-specified primary analysis of the CIBIC-Plus was only marginally significant (p=0.062). So, technically, the study did not meet the protocol specified criteria for a win (treatment differences significant at 0.05 for both primary endpoints). The observed

cases population did show a significant treatment effect on the CIBIC-Plus, but dropouts seem to have fared worse than completers, particularly, in the Memantine group. Thus, the Observed Cases population does not give the complete picture and may be slightly biased in favor of Memantine. Also, further investigation showed that a center stratified analysis (indicated in the original protocol but later abandoned in an amendment) for the CIBIC-Plus based on the ITT-LOCF population resulted in an even larger p-value (0.0942).
It was also observed that in each of studies MEM-MD-02 and 9605 there was a "best" center that had an outlying treatment effect on the change in the ADCS-ADL in favor of Memantine. In MEM-MD-02 exclusion of the outlying center led to a loss of significance: the p-value went from 0.0277 to 0.0766. The effect of the outlying center in MEM-MD-02 may not be too worrisome since no justification was found for excluding it and significance was restored when the "worst" center, i.e., the center which was least favorable for memantine, was also excluded (p=0.0254). In addition, the p-value was small for the treatment effect on the other primary endpoint, change in SIB Total ($p<0.001$). In study 9605 exclusion of the outlying center nearly led to a loss of significance: the p-value changed from 0.0178 to 0.0493. (Note: the sponsor's p-values for study 9605 differ slightly from ours since the sponsor included patients with no post-baseline efficacy measures in the "ITT" population by carrying the baseline value forward. Our ITT population consists of patients with baseline and at least one post-baseline primary efficacy measure). So, although the treatment effect on the change in the ADCS-ADL Total score seems to be modest, it is probably real. Yet, the treatment effect on the CIBIC-Plus, the other primary endpoint in study 9605, was not quite significant ($p=0.062$). In addition, although there was a more impressive treatment effect ($p<0.001$) on the cognitive measure, the Severe Impairment Battery, the effects on 3 of the 4 other secondary endpoints were not significant at the 0.05 level. For these reasons, the monotherapy study 9605 seems less convincing than the add-on study MEM-MD-02.

Study 9403 was positive but different from MEM-MD-02 and 9605 in several important ways. Study 9403 was conducted in assisted living facilities in Latvia, whereas MEM-MD-02 and 9605 were conducted in the U.S. and were not restricted to assisted living facilities. The sample size was smaller in 9403 (166 total compared to 252 and 403) and it included patients with Vascular dementia (slightly more than 50% of all patients). In addition, the length of observation was only 12 weeks compared to 24 and 28 in the other studies and the daily dose was smaller (10 mg vs. 20 mg). Finally, the primary endpoints in 9403, the Care Dependency subscale of the Behavior Rating scale for Geriatric patients and the CGI-Change, were different from those in the other studies and did not contain a cognitive measure. Keeping these differences in mind we note that the 9403 results were significant for both primary endpoints, even in the subgroup of Alzheimer's patients.

2. Introduction

2.1. Overview

Treatment of Alzheimer's disease with cholinesterase inhibitors can produce measurable improvements in cognition and global performance. However, these symptomatic improvements do not tend to be dramatic or long lasting. In addition, these agents demonstrated efficacy only in mild to moderate forms of the disease so there is a need for different and/or complimentary treatments which might benefit more severely affected patients.

Memantine has a different mechanism of action and is cleared from the body differently than acetylcholinesterase inhibitors so it might be useful to combine it with these treatments. For this reason, study MEM-MD-02, described below, was designed to evaluate the safety and efficacy of memantine relative to placebo, in patients with moderately severe to severe Alzheimer's who are also receiving concurrent treatment with the acetylcholinesterase inhibitor donepezil. Two earlier double blind studies, 9605 and 9403, evaluated memantine as monotherapy in patients with moderately severe to severe dementia.

2.2. Data Sources

The data for studies 9605, 9403, 9202, and 9408 is located at
\Cdsesub1\N21487\N_000\2002-12-19\crt\datasets\
The data for study MEM-MD-02 is located at
\\Cdsesub1\n21487\N_000\2003-01-10\crt\datasets\MEM-MD-02

Table 2.1 Summary of Study Features and Key Efficacy Measures

Study	N	MMSE	Age	Length (weeks)	Dose	Functional	Cognitive	Global
MEMMD02 (Memantine added on to Donepezil)	403	5-16 *	50-93	24	20	ADCS-ADL	SIB	CIBIC-Plus
9605	252	1-14 *	50-93	28	20	ADCS-ADL	SIB	CIBIC-Plus
9403	166	0-9	60-81	12	10	BGP care dependency	BGP cognitive	CGI-C
9202	581	10-25	54-97	28	20	-	ADAS-Cog	-
9408	321	11-20	59-96	28	20	-	ADAS-Cog	-

* some patients were outside permitted range: inclusion criteria were 5-14 for MEM-MD-02 and 3-14 for 9605

Note that the two primary endpoints measure different aspects of the disease and in each study to demonstrate effectiveness the treatment effects should be significant on both primary endpoints at the 0.05 level of significance.

Study MEM-MD-02 was a study of the therapeutic effects of the combination of Memantine and Donepezil versus the combination of Placebo and Donepezil for patients currently treated with stable doses of Donepezil. Patients were randomized to receive Memantine 20 mg/day or placebo in addition to Donepezil for 24 weeks. Initially patients were given 5 mg/day. Doses were increased by 5 mg/day each week until the 20 mg/day target was reached. The primary endpoints were the changes from Baseline in the SIB total and the modified ADCS-ADL Total at 24 weeks (LOCF).

Study 9605 was a 28-week, multicenter, randomized, double-blind, placebo controlled trial designed to assess the efficacy and tolerability of Memantine in patients with moderate to severe Alzheimer's disease. It was conducted among 32 centers in the US and was designed to enroll approximately 250 patients at least 50 years of age. Diagnosis of moderate to severe dementia was based on DSM-IV and NINCDS-ADRDA criteria. Patients were required to have an MMSE score of 3 to 14, GDS Stage 5 to 6, FAST score ≥ 6, HIS ≤ 4, and a CT or MRI brain scan compatible with the diagnosis of DAT. Patients randomized to Memantine received 5 mg once daily during Week 1, 10 mg once daily during Week 2, 15 mg as a divided daily dose during Week 3, and 10 mg bid (20 mg/day) from Week 4 to 28. Follow-up visits were conducted at the end of Week 4, Week 12, and Week 28. Primary endpoints were the CIBIC-Plus and the change from baseline in the ADCS-ADL at week 28. The change from baseline in the severe impairment battery (SIB) at Week 28 was a secondary endpoint to assess cognitive function.

Study 9403 was entirely conducted in nursing homes in Latvia. It was a randomized, placebo controlled, double-blind trial in which 166 patients among 7 centers received either placebo or 10 mg Memantine for a 12 week period (in the second week patients were titrated up from 5 mg). It included patients between the ages of 60 and 80 with moderate to severe dementia. Both vascular and primary degenerative dementia were admissible and the randomization was not stratified according to dementia type. Follow-up evaluations were conducted at the end of Week 1, Week 4, Week 8, and Week 12. Primary endpoints were the clinical global impression of change (CGI-C) and the change in the care dependence domain of the geriatric behavior rating scale (BGP) at the end of treatment.

Studies 9202 and 9408 were 28 week randomized, double-blind, placebo-controlled trials, conducted in Europe, designed to assess efficacy and tolerability of Memantine versus placebo in patients suffering from probable vascular dementia. Study 9202 was conducted among 57 centers in the United Kingdom in patients ≥ 50 years of age with a minimum duration of dementia of 12 months and a baseline MMSE score of 10 to 22. Study 9408 was conducted among 50 centers in France, Belgium, and Switzerland in patients ≥ 60 years of age with a minimum duration of dementia of at least 6 months and a baseline MMSE score of 12 to 20. In each study, patients

randomized to Memantine received 5 mg once daily during Week 1, 10 mg once daily during Week 2, 15 mg as a divided daily dose during Week 3, and 10 mg bid (20 mg/day) from Week 4 to 28. Follow-up visits were conducted at the end of Week 4, Week 12, and Week 28. Primary efficacy measures were the change from baseline in the Alzheimer's Disease Assessment Scale (ADAS-Cog) and the Clinical Global Impression of Change (CGI-C) after 6 months. These studies were not reviewed because they targeted less severely affected patients (MMSE >=10) with vascular dementia.

3. Statistical Evaluation

3.1. Evaluation of Efficacy

3.1.1. Study MEM-MD-02

3.1.1.1. Objective

The objective of this study is to evaluate the safety and efficacy of memantine versus placebo in the treatment of moderate to severe dementia of the Alzheimer's type.

3.1.1.2. Study Design

The study will be conducted as a randomized, double-blind, placebo-controlled, multicenter, parallel-group study comparing memantine to placebo in outpatients diagnosed with probable Alzheimer's disease (NINCDS-ADRDA criteria). The study will consist of 1-2 weeks of single-blind placebo treatment followed by 24 weeks of double-blind treatment. This study will involve a total of seven clinic visits: Screening, Baseline, and at the end of Weeks 4, 8, 12, 18, and 24. Approximately 340 patients will be enrolled into this study with each of the double-blind treatment groups containing approximately 170 patients.

The study population will consist of outpatients who are at least 50 years of age and who have been diagnosed with probable Alzheimer's disease using NINCDS-ADRDA criteria. Alzheimer's disease severity will range from moderate to severe assessed on the basis of MMSE scores (≥ 5 and ≤ 14). Eligible patients will have been receiving donepezil therapy for at least 6 months, and must have been at a stable dose (5-10 mg/day) for the last 3 months. All patients must continue to receive donepezil therapy for the duration of the study.

The following titration scheme will be used. During the first week of treatment, patients randomized to memantine will receive 5 mg/day, followed by 10 mg/day during the second week, and 15 mg/day during the third week. The target dose of 20 mg/day will be administered starting with the fourth week of double-blind treatment and will continue throughout the study.

3.1.1.3. Efficacy Endpoints

Primary Efficacy Assessments

Alzheimer's Disease Cooperative Study-Activities of Daily Living Inventory (ADCS-ADL)

The Modified ADCS-ADL Inventory consists of 19 items, appropriate for patients with moderate to severe dementia, selected from the full 42-item inventory. This battery of ADL questions is used to measure the functional capabilities of patients with dementia.

Each ADL item comprises a series of hierarchical subquestions, ranging from the highest level of independent performance of each ADL to complete loss. The inventory is administered as an interview to a close informant of the patient and covers the patient's most usual and consistent performance of each ADL during the previous 4 weeks. The range of the sum score is 0 to 54.

Severe Impairment Battery

The Severe Impairment Battery (SIB) was developed for the assessment of cognitive dysfunction in patients with advanced AD. It is structured along the usual lines of cognitive testing in AD, covering the areas of memory, language and praxis as well as attention and orientation. The test contains 51 items and the range of possible scores is 0-100 (with 100 being the best result).

Secondary Efficacy Assessments

Clinician's Interview-Based Impression of Change-Plus.

The CIBIC-Plus is a global change measure which is based on information collected by a physician, familiar with the manifestations of dementia, during an interview with the patient and caregiver. The physician assesses disease severity at baseline and is barred from knowledge of all other psychometric test scores conducted as part of this protocol. Using the results from baseline as a reference, the clinician interviews the patient and caregiver at the end of weeks 4, 8, 12, 18, and 24 and determines an "Impression of Change" rating. The format for scoring is a 7 point scale, which provides for symmetrical improvement or worsening (1,2,3 Improved; 4=no change; 5,6,7 Deteriorated).

3.1.1.4. Statistical Analysis Plan

All efficacy analyses will be based upon the randomized patients who took at least one dose of study medication and who had at least one post-baseline primary efficacy assessment, i.e., the intent-to-treat (ITT) population. All statistical tests will be two-sided, and a p-value ≤ 0.05 will be considered statistically significant. Primary analyses will be performed on the ITT population at week 24 using the last observation carried forward (LOCF) approach. This approach consists of using the last observed value before a missing value to impute the missing value.

Primary Efficacy Parameters

For the change from baseline in the total SIB and ADCS-ADL scores at Week 24, the comparison between memantine and placebo will be performed using two-way analysis of covariance (ANCOVA) with treatment group and center as the two factors and the baseline score as covariate.

Missing Data Handling
Missing visits will be replaced using the last observation carried forward approach. If more than 4 of the 19 items which comprise the ADCS-ADL total score are missing then the total score will be set to missing. Otherwise, single missing items will be replaced by 0, the worst value. If more than 11 of the 51 items which comprise the SIB total score are missing, then the total score will be set to missing. Otherwise, single missing items will be replaced by 0, the worst value.

Secondary Efficacy Parameters
The CIBIC-plus rating will be analyzed using the Cochran-Mantel-Haenszel test, controlling for study center. For other measures, comparison between memantine and placebo will be performed using the same approach as for the primary efficacy parameter. Results of the CIBIC-Plus will be included in labeling if the treatment group differences are significant at 0.05 for both primary variables and the CIBIC-Plus.

Sample size Considerations
The primary efficacy variables are the change from baseline in SIB and ADCS-ADL scores. Assuming an effect size (treatment group difference relative to pooled standard deviation) of 0.35, a sample size of 170 patients in each treatment group will provide 90% power at an alpha level of 0.05 (two-sided), based upon a two-sample t-test.

3.1.1.5. Study Population

A total of 404 patients (201 placebo/donepezil and 203 memantine/donepezil patients) were randomized. The ITT population consisted of the 395 of these (197 P/D and 198 M/D) who received at least one dose of double-blind study medication and had a baseline and at least one post-baseline efficacy assessment. More Memantine/Donepezil patients completed the study [172 (85%) for M/D compared to 150 (75%) for P/D]. Adverse events and withdrawal of consent were the most frequent reasons given for discontinuation.

Table 3.1 Reasons for Discontinuation

	Placebo/ Donepezil	Memantine/ Donepezil	Total
Patients who completed	150 (74.6)	172 (85.1)	322 (79.9)
Patients who discontinued*	51 (25.4)	30 (14.9)	81 (20.1)
REASONS FOR DISCONTINUATION			
Adverse Event	25 (12.4)	15 (7.4)	40 (9.9)
Insufficient Response	3 (1.5)	1 (0.5)	4 (1.0)
Protocol Violation	5 (2.5)	1 (0.5)	6 (1.5)
Consent Withdrawn	16 (8.0)	8 (4.0)	24 (6.0)
Lost to Follow-up	0	1 (0.5)	1 (0.2)
Other reasons	2 (1.0)	4 (2.0)	6 (1.5)

* Patient may have had one or more reasons for discontinuation

Table 3.2 Demographic Characteristics

	Placebo/ Donepezil	Memantine/ Donepezil	Total
MEAN AGE (SD)	75.54 (8.73)	75.53 (8.43)	75.54 (8.57)
≤ 64 n (%)	28 (13.9)	26 (12.8)	54 (13.4)
65-74 n (%)	49 (24.4)	54 (26.6)	103 (25.5)
75-84 n (%)	96 (47.8)	99 (48.8)	195 (48.3)
≥ 85 n (%)	28 (13.9)	24 (11.8)	52 (12.9)
SEX			
Male	67 (33.3)	74 (36.5)	141 (34.9)
Female	134 (66.7)	129 (63.5)	263 (65.1)
RACE			
Caucasian	186 (92.5)	183 (90.1)	369 (91.3)
Non-Caucasian	15 (7.5)	20 (9.9)	35 (8.7)
WEIGHT (LB) mean (SD)	146 (31.07)	155.5 (31.49)	150.8 (31.60)

The memantine/donepezil and placebo/donepezil treatment groups were well-matched with respect to demographic characteristics at baseline. Overall, 65% of patients were female, 91 % were Caucasian, and 61% were at least 75 years of age. The mean patient age was 76 years and the mean body weight was 151 pounds. There was a statistically significant difference in the mean body weights for the treatment groups (156 lbs. M/D; 146 lbs. P/D p=0.003).

Baseline assessments of disease severity were comparable between the groups with the exception of the Hachinski scores for which a small but statistically significant difference (p=0.028) was observed. Placebo patients averaged 2 points higher on the SIB Total but this difference was not significant (p=0.21).

Baseline Assessment	Placebo/Donepezil N=197	Memantine/Donepezil N=198
Hachinski	0.6 (0.7)	0.7 (0.9)
MMSE	10.2 (3.0)	9.9 (3.1)
SIB	79.8 (15.5)	77.8 (15.5)
ADCS-ADL	36.2 (9.3)	35.9 (9.8)
NPI	13.8 (12.8)	13.7 (14.1)
BGP Total	13.5 (7.7)	13.3 (7.8)

Baseline Assessment	Placebo/Donepezil N=197	Memantine/Donepezil N=198
BGP Care Dependency	9.2 (6.0)	8.9 (5.8)
BGP Cognitive	1.4 (1.5)	1.3 (1.5)

Donepezil Treatment History

Mean duration of treatment with donepezil at baseline was 129 and 126 weeks for the placebo/donepezil and memantine/donepezil treatment groups respectively. The mean dose of donepezil was 9.49 and 9.25 for the placebo/donepezil and memantine/donepezil treatment groups respectively. The majority of patients (85%) were administered a daily dose of 10 mg of donepezil. 98% placebo/donepezil and memantine/donepezil patients received concomitant medication other than donepezil during the study. The most common concomitant medication taken by patients in both treatment groups was tocopherol (60% of placebo/donepezil patients and 65% of memantine/donepezil patients). Other medications taken were acetylsalicylic acid (38% and 36 %) and multivitamins (39% and 40%). There were no important differences between the treatment groups in the percentage of patients receiving concomitant medications or the types of concomitant medications taken.

3.1.1.6. Sponsor's Efficacy Results

Primary Efficacy Parameters

Severe Impairment Battery

At week 24 (LOCF analysis), the mean change in the SIB from baseline for memantine/donepezil patients was 0.9 compared to a mean change in the placebo/donepezil group of –2.5. The least square mean treatment difference of 3.4 between the two groups was statistically significant in favor of memantine/donepezil ($p<0.001$). Results from the Observed Cases (OC) analysis of the SIB were consistent with the LOCF analysis. The memantine/donepezil-placebo/donepezil least square mean treatment group difference of 3.4, favoring memantine/donepezil, was statistically significant ($p<0.001$) at Week 24.

Table 3.3 SIB Results

	Placebo		Memantine		
	N	Mean	N	Mean	p-value*
Endpoint (LOCF)	196	-2.5	198	0.9	<0.001
Week 28 (OC)	153	-2.4	171	1.0	<0.001

* based on an ANCOVA model for SIB change with treatment group, center, and baseline score effects.

Modified ADCS-ADL Inventory

At Week 24 (LOCF analysis), the least square mean change from baseline in the ADCS-ADL for

the memantine/donepezil treatment group was –2.0 compared to a mean in the placebo/donepezil group of –3.4. The mean difference of 1.4 between the two groups in favor of memantine was statistically significant (p=0.028). Results from the OC analysis of the ADCS-ADL were consistent with the LOCF analysis. The memantine/donepezil-placebo/donepezil least square mean treatment group difference of 1.6, favoring memantine/donepezil, was statistically significant (p=0.020) at Week 24.

Table 3.4 ADCS-ADL Results

	Placebo		Memantine		
	N	Mean	N	Mean	p-value*
Endpoint (LOCF)	197	-3.4	198	-2.0	0.028
Week 28 (OC)	152	-3.3	172	-1.7	0.020

* based on an ANCOVA model for ADCS-ADL change with treatment group, center, and baseline score effects.

Secondary Efficacy Parameters

The mean CIBIC-Plus rating for memantine/donepezil patients was 4.41 at Week 24 (LOCF analysis) compared to 4.66 for patients treated with placebo/donepezil. The difference between treatment groups was statistically significant in favor of memantine (p=0.027) at Week 24. The results of the observed cases analysis were consistent with those of the LOCF analysis at Week 24.

Table 3.5 Mean CIBIC-Plus Rating

	Placebo		Memantine		
	N	Mean	N	Mean	p-value*
Endpoint (LOCF)	196	4.66	198	4.41	0.027
Week 28 (OC)	152	4.64	172	4.38	0.028

* based on a Van Elteren test (i.e., a center stratified rank sum test).

3.1.1.7. Reviewer's Comments

In the site inspection of center 13 discrepancies were found between the case report form and the data listings for patients 0139211 and 0139214. However, this reviewer believes that this is not an issue because these two patients did not complete the study and the discrepancies appear to be a result of the last observation carried forward imputation rule. The data on which the analysis is based agrees with the case report forms.

This reviewer verified the sponsor's primary analyses. The primary analysis method was ANCOVA of the mean change from baseline with treatment and center effects and baseline score

as the covariate. The primary analysis was based on the ITT population and the last observation carried forward method. The mean change from baseline for Memantine was found to be significantly better than placebo for both the SIB and ADCS-ADL Total scores. It is noteworthy that if we ignore baseline scores there was no group difference in mean SIB or ADCS-ADL Total scores at 24 weeks using LOCF. Furthermore, the baseline difference in mean SIB scores (1.88) was numerically greater than the difference in mean SIB Totals at 24 weeks (-1.47) using LOCF. However, the mean changes were significant because of the reduction in variability obtained by incorporating baseline scores and the fact that placebo started out slightly better and ended slightly worse.

This reviewer also noticed that the assumption of normality upon which the p values for the ANCOVA model are based was violated. In particular, a Shapiro-Wilks test for normality of the residuals was significant ($p<0.0001$) suggesting a lack of normality. Other standard tests of normality led to the same conclusion. This means that the ANCOVA based p-values may not be correct. The statistical analysis plan did not propose an alternative method to be used in the case of non-normality. This reviewer found that the p-values were still significant if a non-parametric method such as the Wilcoxon-Mann-Whitney test or the center stratified Wilcoxon (CMH) test was used instead of ANCOVA:

Table 3.6 P-values for test of treatment effect using different methods

	Analysis Method		
	ANCOVA	Van Elteren	Wilcoxon
Endpoint			
Change in SIB Total	0.001	0.001	0.001
Change in ADL Total	0.028	0.005	0.021

Therefore, the following refers to the protocol specified ANCOVA based analyses, despite the fact that the assumption of normality is questionable.

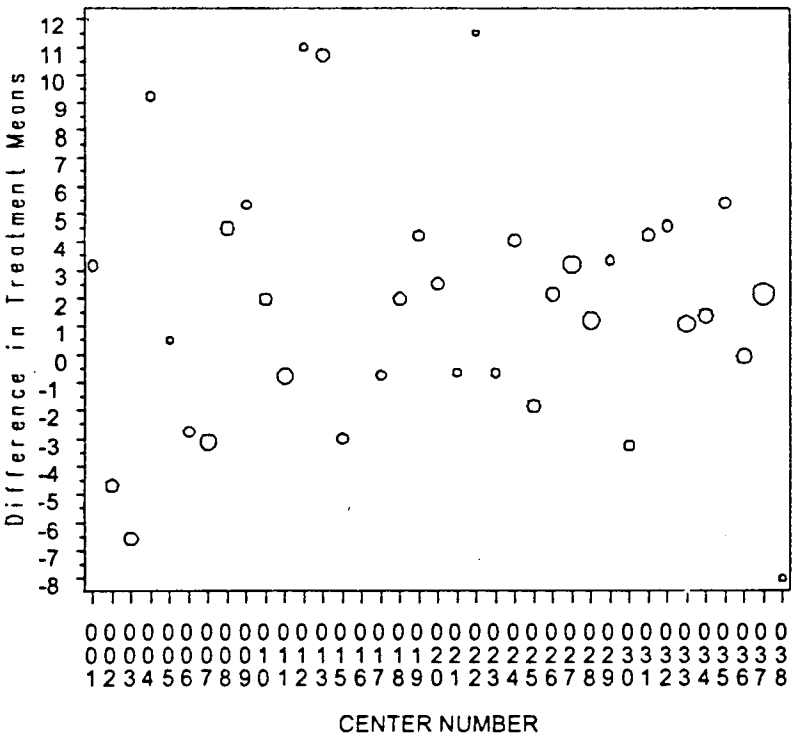
SIB Total

It is notable that on average Memantine patients had slightly improved at the end of the study in terms of the SIB. There was, however, a suggestion of a downward trend in the SIB scores in the last 12 weeks (last 2 visits). The treatment effect was fairly consistent across centers: the mean SIB score was numerically better for Memantine than placebo in 27 of the 36 centers that had patients in both arms.

ADCS-ADL Total

Among the 36 centers with patients in both arms the mean change in ADL for Memantine was numerically better in 23 and worse in 13. This reviewer observed that for center 013 the difference in treatment mean changes from baseline for the ADL Total score was considerably larger than the center average. The differences are shown in Figure 3.1. Note that the size of the plotting symbol is proportional to the size of the center.

Figure 3.1 Difference in Treatment Mean Changes in ADCS-ADL Total by Center



There were 11 patients in center 013 (6 placebo and 5 memantine). The average change for placebo patients in this center was -11.33 compared to -0.60 for memantine patients.

The difference in treatment group mean changes (+10.73 ± 3.77 S.E.) in this center was considerably larger than the average (+1.34). This center's results are somewhat atypical and if we discard this center from the ITT population we find that the treatment effect on the change in

ADL adjusted for baseline and center is no longer significant (p=0.0766). The difference in group mean changes drops from 1.34 to 1.05. On the other hand, center 003 had 14 patients and a large negative difference (-6.57), so exclusion of this center would increase the significance of the treatment effect. Nonetheless, the fact that the removal of center 013 can alter the conclusions calls the strength of this evidence into question.

There were also several patients whose changes were noticeably larger than the rest:

Patient	Treatment	Baseline ADL	Change in ADL
0119218	Placebo	40	-33*
0139214	Placebo	36	-29*
0039228	Memantine	43	-31
0129201	Memantine	9	16

* based on an early termination visit

Both of these placebo patients dropped out of the study because of adverse events. Patient 0119218 suffered a hip fracture and an elbow fracture which might help to explain the observed decline in the activities of daily living. The following table shows the progression of their scores.

ID	ActualWeek	SIB Total	ADL Total
0119218	0	92	40
0119218	5	91	40
0119218	8	97	38
0119218	12	.	.
0119218	18	82	7
0119218	24	.	.
0139214	0	83	36
0139214	4	90	38
0139214	7	85	35
0139214	12	.	7
0139214	18	.	.
0139214	24	.	.

The SIB value at week 12 of patient 0139214 was discarded (as directed by the protocol) since 42 out of the 52 SIB items were missing, yet none of the ADCS-ADL items were missing at week 12. Considering that the average change in ADL Total for placebo patients was -3.18 +/- 6.03 (S.D.) it is not surprising that these two patients have a strong influence on the ADCS-ADL results: if they are removed the difference in treatment group mean changes goes from 1.34 to 1.05 and the p-value increases from 0.028 to 0.0599. It should be pointed out though that the significance of the treatment effect is restored if the two Memantine outliers are also removed. The preceding arguments should be regarded as sensitivity analyses since no justification for removing center 13 or the two placebo patients was found.

Effect of Dropouts

Inspection of the following table shows that placebo dropouts worsened numerically less than memantine dropouts for each dropout time. This is the reverse of the observed cases population where Memantine was significantly better than placebo. The implication is that the Observed Cases analysis may be biased in favor of Memantine. Our current practice is to evaluate the effect of dropouts by comparing the Observed Cases and ITT results and since they agree in this case we are relatively satisfied that there is not a problem here.

Table 3.7 Mean Change in ADL by Visit for Dropouts and Observed Cases

last week	Treatment	n	Base ADL	change 4	change 8	change 12	change 18	n_24	change 24
0	Placebo	12	36.8 (9.4)	.	.	.	.	8	0.3 (3.2)
0	Memantine	8	29.8 (8.1)	.	.	.	.	3	-2.3 (4.5)
4	Placebo	11	35.1 (9.9)	-0.5 (3.7)	.	.	.	7	-3.7 (4.0)
4	Memantine	9	33.9 (8.8)	-0.8 (6.3)	.	.	.	6	-5.8 (6.5)
8	Placebo	13	35.5 (6.9)	-0.5 (3.0)	-1.5 (3.4)	.	.	9	-10.2 (12.7)
8	Memantine	3	31.7 (11.8)	0.7 (5.0)	-3.3 (6.7)	.	.	1	0.0 (.)
12	Placebo	9	40.2 (7.7)	-2.1 (2.9)	-0.7 (1.9)	-2.7 (3.4)	.	4	-7.3 (3.9)
12	Memantine	6	34.3 (8.2)	-1.7 (4.4)	-0.3 (4.9)	-3.0 (7.0)	.	4	-5.0 (12.3)
18	Placebo	3	34.0 (5.2)	-0.3 (0.6)	-1.7 (0.6)	-2.3 (2.3)	-1.0 (3.0)	0	.
18	Memantine	5	32.0 (17.1)	-1.0 (3.4)	1.0 (4.1)	-1.4 (3.6)	-4.4 (7.5)	0	.
24	Placebo	153	36.2 (9.6)	-0.8 (3.8)	-0.9 (4.2)	-1.5 (4.6)	-2.1 (5.2)	153	-3.1 (5.6)
24	Memantine	172	36.3 (9.7)	0.2 (3.5)	0.3 (4.4)	-0.3 (5.0)	-0.8 (5.3)	172	-1.4 (6.3)

* includes retrieved dropouts

For the SIB, on average, placebo dropouts worsened numerically less than or equal to Memantine dropouts for each dropout time, except week 12. The 6 Memantine dropouts for week 12 had an average change of +3.7, while the 7 placebo dropouts had an average change of –3.4 . On the other hand, the 9 Memantine dropouts in week 4 had an average change of –5.8, whereas the 11 placebo dropouts had an average change of +0.3. Thus, the Observed cases analysis of the change in the SIB has no apparent bias. The LOCF analysis of the change in the SIB from baseline also leads to the conclusion that Memantine is significantly better than placebo.

CIBIC-Plus (Secondary Endpoint)

Memantine was found to be significantly better than placebo in terms of the CIBIC-Plus at week 24 using the center stratified Cochran-Mantel-Haenszel test. Among the 36 centers that had patients in both arms, the mean score for Memantine was better than placebo in 22, the same in 3, and worse in 11. Despite the significance of the CMH test the difference in mean scores was small: the mean was 4.66 for placebo and 4.41 for memantine. The clinical relevance of this effect may also be questioned because the difference in percent not worse (CIBIC-Plus <=4) was not significant (p=0.056). The percentage was 55% for Memantine and 45% for placebo. The percentages were also 55% and 45%, respectively, for the observed cases and the difference was not significant (p=0.059).

Other Secondary Endpoints

Significant treatment effects were also observed for the NPI Total, the BGP Total and the BGP care dependency, but not for the FAST.

3.1.2. Study 9605

The study was conducted between 8/21/1998 and 10/04/1999.

3.1.2.1.Objective

The objective was to demonstrate superiority of Memantine treatment versus placebo for moderately severe to severe Alzheimer's disease as assessed by clinical global and functional endpoints.

3.1.2.2. Study Design

The trial is designed as a prospective, randomized, placebo-controlled, double-blind, multicenter trial in patients suffering from moderately severe to severe Alzheimer's disease.

After an initial screening period of two to four weeks, eligible patients will be randomly allocated to two parallel groups. The treatment duration is 28 weeks. The maintenance dose of

Memantine of 20 mg/day p.o. will be reached over four weeks with a weekly dose increment of 5 mg of Memantine. Test sessions for the efficacy parameters are scheduled at Baseline, after 4 (partial), after 12, and 28 weeks of treatment.

3.1.2.3.Efficacy Measures

Statistical evaluation of drug effects is planned with two primary efficacy variables: 1.) a clinical global endpoint (CIBIC-Plus independent rater) and 2.) a functional endpoint (modified ADCS-ADL Inventory; change in sum scores).

The CIBIC-Plus is a global change measure which is based on information collected by a physician, familiar with the manifestations of dementia, during an interview with the patient and caregiver. The format for scoring is a 7 point scale, which provides for symmetrical improvement or worsening (1,2,3 Improved; 4=no change; 5,6,7 Deteriorated).

Modified ADCS-ADL Inventory
Functional Assessment of AD patients should focus on their performance of activities of daily living (ADL). The ADCS-ADL Inventory is a comprehensive battery of ADL/ instrumental ADL questions aimed to measure functional ability of AD patients over a broad range of dementia severity. Each ADL item comprises a series of hierarchical subquestions, ranging from the highest level of independent performance of each ADL to complete loss. The inventory is administered as an interview to a close informant of the patient and covers the patient's most usual and consistent performance of each ADL during the previous 4 weeks. For the purpose of this trial a subset of 19 items was selected to fit the characteristics of the trial population of moderately severe to severe AD patients (MMSE range between 3 and 14). The range of the sum score is 0 to 54.

Secondary Endpoints
The Severe Impairment Battery (SIB) was developed for the assessment of cognitive dysfunction in patients with advanced AD. It is structured along the usual lines of cognitive testing in AD, covering the areas of memory, language and praxis as well as attention and orientation. Out of 40 items, the range of possible scores is 0-100.

3.1.2.4.Statistical Analysis Plan

Handling of missing values
In the ITT analysis two different strategies for replacement of missing values for primary efficacy variables will be used. These are:

1. If available, the endpoint assessment (after 28 weeks under treatment) will be used. If this is unavailable, then the last available observation on the patient (a scheduled or an unscheduled assessment or Baseline values) will be used. This kind of replacement of the missing values will be performed for the confirmatory analysis.
2. If available, the endpoint assessment (after 28 weeks under treatment) will be used. The missing values concerning the primary efficacy variables of discontinued patients (e.g. withdrawals, losses to follow-up) will be replaced by means of the retrieved dropout assessments. If this is unavailable, then the last available observation on the patient (a scheduled or an unscheduled assessment or Baseline values) will be used. This kind of replacement for missing data will be performed in addition to the previous mentioned strategy and for descriptive purposes only.

In case of intermediate values, average values of the nearest pre-values and post-values will be calculated for replacement.

In case of missing data for CIBIC-Plus over the entire study (interview based data only at Baseline available), for the ITT analyses the score 4 (unchanged) will be used.

For the ADL, MMSE, and SIB efficacy scale scores, which are computed by summing items at a visit, wherever possible these sum scores will be calculated from single values by computer programs. If single values are missing, they will be replaced by scores that represent the lowest level of functioning or "worst case" for that scale. For each of these scales higher values represent higher levels of functioning. Therefore, if a patient has at least one non-missing item the missing items will be set to 0 and then the total score will be computed by summing over all the items. If all items are missing then the total score will be treated as missing.

Analysis Methods

Both primary and secondary efficacy outcomes will be analyzed using the Wilcoxon-Mann-Whitney test for independent samples. For all measures, the outcome of interest is the change from baseline in the patient's condition at 28 weeks. The primary efficacy analysis will be performed on the change from baseline at 28 weeks (LOCF) for the ITT population. The trial will be considered positive if memantine is found to be significantly better than placebo at the 0.05 level for both the primary endpoints, the CIBIC-Plus and the change from baseline in the modified ADCS-ADL.

Because the trial will be conducted in more than 30 centers, pooling of centers with ≤ 5 randomized patients will be necessary. Therefore, center effects and treatment by center interactions will only be examined in an exploratory fashion.

Sample Size

Given $\alpha=0.05$ and $\beta=0.05$, in order to show a difference of 20% between the treatments

(improvements) at the end of the double-blind phase (10% placebo 30% Memantine) regarding the trichotomized CIBIC-Plus, 118 patients are needed in each group.

3.1.2.5. Study Population

Patient disposition is presented in Table 3.8. Half of the 252 total patients were randomized to each group. The discontinuation rate was larger for the placebo group than the Memantine group (33 % vs. 23 %). About half of the patients who discontinued did so because of adverse events.

Table 3.8 Patient Disposition

	Placebo N (%)	Memantine N (%)	Total N (%)
Randomized	126	126	252
Completed	84 (67)	97 (77)	181 (72)
Discontinued	42 (33)	29 (23)	71 (28)
Reasons for Discontinuation:			
Adverse Events	24 (19.0)	14 (11.1)	38 (15.1)
Insufficient Therapeutic Response	0 (0.0)	1 (0.8)	1 (0.4)
Protocol Violation	6 (4.8)	4 (3.2)	10 (4.0)
Withdrawal of Consent	10 (7.9)	8 (6.3)	18 (7.1)
Lost to Follow-up	1 (0.8)	2 (1.6)	3 (1.2)
Other reasons	1 (0.8)	0 (0.0)	1 (0.4)

Table 3.9 Patient Demographics – Study 9605

Demographic Parameter	Placebo (N=126)	Memantine (N=126)
AGE (YEARS)		
Mean ± SD	76.3 ± 7.8	75.9 ± 8.4
Range	53, 93	50, 92
≤ 64, n (%)	10 (8%)	12 (10%)
65-74, n(%)	41 (33%)	38 (30%)
75-84, n(%)	60 (48%)	60 (48%)
≥ 85, n(%)	15 (12%)	16 (13%)
SEX, N(%)		
Male	47 (37%)	35 (28%)
Female	79 (63%)	91 (72%)
RACE, N(%)		
Caucasian	115 (91%)	112 (89%)
Non-Caucasian	11 (9%)	14 (11%)
WEIGHT (KG)		
Mean ± SD	66.1 ± 14.1	64.5 ± 12.4
Range	39, 98	31, 104
MMSE		
Mean ± SD	8.1 ± 3.6	7.7 ± 3.7
Range	1, 14	2, 14
HIS		
Mean ± SD	0.6 ± 0.8	0.5 ± 0.7
Range	0, 4	0, 3
BASELINE ADCS-ADL		
Mean ± SD	27.4 ± 10.9	26.8 ± 9.2
BASELINE SIB		
Mean ± SD	68.3 ± 20.8	65.9 ± 22.5

At baseline, the treatment groups were comparable with respect to age, race, weight, baseline MMSE, HIS, ADCS-ADL, and SIB. One noticeable difference was that there were 9% more females in the Memantine group than in the Placebo group but this difference is only marginally significant (p=0.11).

3.1.2.6. Sponsor's Efficacy Results

Functional Assessment: ADCS-ADL

The following table presents the ADCS-ADL mean change from baseline score at endpoint (LOCF) and after 28 weeks of treatment (OC). When daily functioning was evaluated using the ADCS-ADL,

memantine treatment resulted in significantly less deterioration over time compared with placebo. Mean change scores for placebo reflected continuous deterioration, while for Memantine there was evidence of slight improvement at week 4, but the mean scores deteriorated thereafter.

	Placebo		Memantine		
	N	Mean	N	Mean	p-value*
Endpoint (LOCF)	126	-5.08	126	-3.02	0.02
Week 28 (OC)	84	-5.86	97	-2.49	<0.01

* based on a Wilcoxon Rank Sum test

Global Assessment: CIBIC-Plus

	Placebo		Memantine		
	N	Mean	N	Mean	p-value*
Endpoint (LOCF)	126	4.73	126	4.48	0.06
Week 28 (OC)	84	4.74	97	4.38	0.03

* based on a Wilcoxon Rank Sum test

For the CIBIC-Plus a mean difference of 0.25 (0.36) points was observed in favor of memantine in the LOCF (OC) analysis. The LOCF result was marginally significant (p=0.06), while the Observed Cases result was significant (p=0.03). Because of this discrepancy in the results between the LOCF and OC analyses and the observed difference in the rate of premature discontinuations between treatment groups (23% memantine vs. 33% placebo), the effect of missing data on the LOCF analysis was examined in several exploratory analyses. In these alternative LOCF analyses, missing Week 28 CIBIC-Plus ratings were replaced with the worst case (a score of 7), the group mean, or the group median. Each of these imputation rules yielded a more significant result than the Observed cases analysis.

COMMENT: *It is not surprising that these post-hoc sensitivity analyses produced more significant results than the observed cases analysis. In the worst cases imputation procedure, since 13 (10%) more placebo patients discontinued, a higher proportion of worst cases are added for placebo which benefits the memantine group. Likewise, imputation by the mean or median tends to reduce the variability in the scores while having little effect on the mean scores. The result is that the significance of the effect is inflated.*

Furthermore, although the primary analysis method for the CIBIC-Plus was not center adjusted, the center adjusted Wilcoxon test has a p-value of 0.094. The discrepancy between the p-value unadjusted for center and the one adjusted for center is likely due to the presence of negative treatment effects in several centers, including one large negative effect in center 30. Finally, the Memantine dropouts seem to have fared worse than the Memantine completers while the placebo dropouts were more comparable to the placebo completers yet the sponsor's imputation procedures do not reflect this.

Cognitive Assessment – SIB

The mean change scores under placebo treatment provided evidence of continuous deterioration of cognitive performance during the study. Mean change scores for Memantine provide evidence of maintenance of cognitive abilities over the first 12 weeks of treatment. Mean cognitive performance deteriorated after 12 weeks of treatment with memantine, but it remained higher than mean cognitive performance for placebo.

	Placebo		Memantine		
	N	Mean	N	Mean	p-value*
Endpoint (LOCF)	126	-9.84	126	-3.93	<0.01
Week 28 (OC)	83	-10.16	96	-4.46	<0.01

* based on a Wilcoxon Rank Sum test

3.1.2.7. Reviewer's Analysis

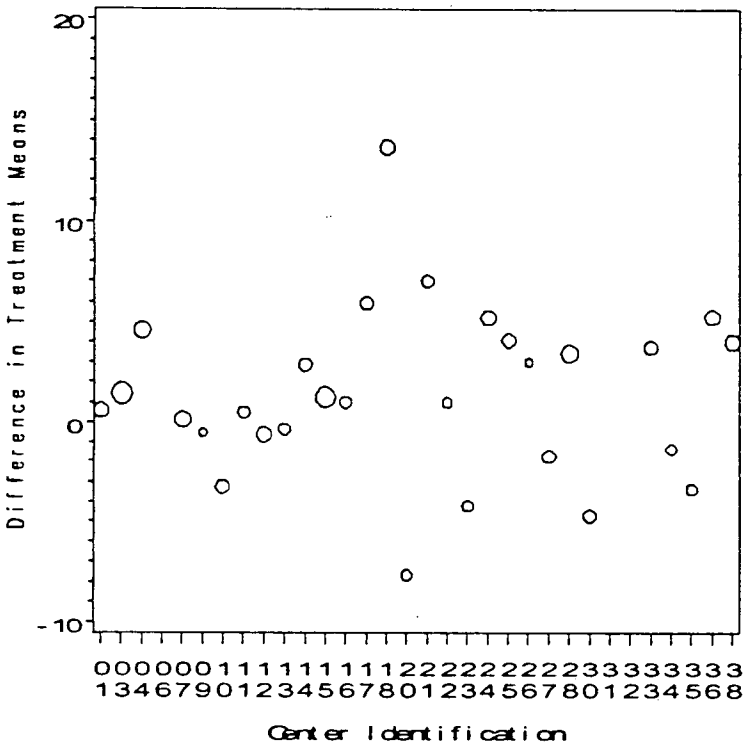
Note: The sponsor's p-values for study 9605 differ slightly from ours since the sponsor included patients in the "ITT" population with no post-baseline efficacy measures by carrying the baseline value forward. Our ITT population consists of patients with baseline and at least one post-baseline primary efficacy measure. There were five patients (3 placebo and 2 memantine) with no post-baseline ADL or SIB measures and 16 (8 in each group) with no post-baseline CIBIC-Plus measures.

ADCS-ADL (Co-Primary)

A significant treatment effect was found on the change in ADL Total score from Baseline to week 28 for both the ITT(LOCF) population (p=0.017) and the Observed Cases (p=0.003) using the Wilcoxon Mann Whitney test.

Among the 29 centers that had patients in both arms, 19 were numerically better for Memantine and 10 were numerically worse in terms of mean change. This reviewer observed that for center 18 the difference in treatment mean changes from baseline for the ADL Total score was considerably larger than the average over all centers. The differences are shown in Figure 3.2. Note that the size of the plotting symbol is proportional to the number of patients in the center.

Figure 3.2 Difference in Treatment Mean Changes in ADCS-ADL Total by Center



Center 18 had a total of 9 patients, 4 placebo and 5 memantine. In this center, the ADL average change from baseline for placebo was -11.25 compared to 2.40 for memantine so the estimated treatment effect in this center is 13.65 ± 4.39 SE. This is about 2.7 standard errors larger than the average, 1.41. The most striking patient, (ID=1800011), from this center was from the placebo group and had a baseline of 32 and a final score of 3 at week 28. Three different patients from this center (2 placebo and 1 memantine) did not complete the study. One placebo patient dropped out at week 12 despite an improvement between weeks 4 and 12 that left them unchanged from baseline. The mean change from baseline is 2.13 (± 0.84) points higher for Memantine than placebo with center 18 included and 1.70 (±0.83) points higher with center 18 removed. Of course, the primary analysis is based on the ranks of the changes, but center 18 also had the largest deviation from the average in terms of the difference in the treatment mean ranks. If we perform the LOCF analysis on the remainder of the ITT population after removing this center we find that the treatment difference is barely significant (p=0.0493). This might be construed as

evidence that the treatment effect is modest. In fact, if one were to use the sponsor's approach where baseline is carried forward when no post-baseline measures are available, as specified in the protocol, this p-value would be 0.0594. However, our practice is to exclude patients with no post-baseline efficacy measures from the ITT population.

Effect of Dropouts

As seen in the following table, Memantine dropouts at their last visit were worse than Memantine completers at the same time, in terms of average change in the ADL. The most striking difference was for the week 12 dropouts. The average change at week 12 for the ten Memantine patients who dropped out at week 12 was –5.4 compared to –0.1 for the completers at week 12. Placebo dropouts were also slightly worse in terms of average change than placebo completers at the time of dropout but to a lesser extent. Therefore, the Observed Cases analysis could be slightly biased in favor of Memantine. However, this doesn't seem to be a cause for alarm since the results were significant for both the Observed Cases and the LOCF analyses.

Table 3.10 Mean (SD) Change in ADL by Last Available Visit

Last week	Treatment	n	adltot 0	change 4	change 12	change 28	n_ et	change et	n_ rd	change rd
0	Placebo	9	22.2 (7.8)	NA	NA	NA	6	-7.0 (7.0)	0	NA
0	Memantine	7	27.7 (8.5)	NA	NA	NA	5	-2.2 (5.3)	1	-2.0 (.)
4	Placebo	11	29.4 (6.1)	-0.7 (4.6)	NA	NA	7	-2.6 (2.9)	1	-10.0 (.)
4	Memantine	12	23.8 (8.1)	-0.1 (4.2)	NA	NA	9	-4.6 (8.9)	2	-8.5 (2.1)
12	Placebo	22	21.7 (10.0)	-0.5 (3.8)	-3.2 (4.3)	NA	6	-4.5 (4.0)	0	NA
12	Memantine	10	30.3 (9.1)	-1.6 (1.7)	-5.4 (7.2)	NA	3	-14.3 (11.1)	1	-10.0 (.)
28	Placebo	84	29.2 (11.3)	-0.7 (4.4)	-1.7 (5.6)	-5.9 (6.8)	.	.	.	.
28	Memantine	97	26.8 (9.3)	0.7 (4.4)	-0.1 (5.6)	-2.5 (6.3)	.	.	.	.

* et = early termination visit; rd = retrieved dropout at week 28

CIBIC-Plus (Co-Primary)

Recall that the CIBIC-Plus ranges from 1='Very Much Improved' to 7='Very Much Worse' with 4='No Change'. This reviewer verified the sponsor's result based on the Wilcoxon Mann Whitney test unadjusted for center, which was indicated as the primary analysis for the CIBIC-Plus in the final analysis plan (written before unblinding). The sponsor's p-value was 0.064 for the ITT population, using LOCF, and 0.025 for the Observed Cases population. There were 16 patients (8 in each group) who had no post-baseline assessment and were assigned a value of 4 (no change) for the CIBIC-Plus at endpoint. Excluding these 16 patients changed the LOCF p-value only slightly to 0.062. This reviewer found that the ITT-LOCF p-value was 0.094 for a center stratified version of the Wilcoxon test (Cochran-Mantel-Haenszel test). In the original protocol the plan was to use this center stratified method. The larger p-value in the center stratified analysis may be due to the existence of several centers where the mean score was numerically better for placebo. Among the 29 centers with patients in both arms, the mean score for Memantine was numerically better than placebo in 19, equal in 2, and worse in 8. Center 30 where the 3 Memantine patients had an average of 6.0 and the 4 placebo patients had an average of 4.5 was the most striking of these.

The sponsor used several imputation methods which produced a significant result for the ITT population, but these methods ignore the observed treatment differences between the completers and dropouts. The sponsor's worst case imputation method also favors Memantine because of the higher number of dropouts in the placebo arm, which means that more worst values are assigned to placebo. As seen in the following table at week 28 (LOCF) Memantine dropouts did more poorly than Memantine completers in terms of percent not worse (CIBIC<=4) and mean score, while placebo dropouts did about the same as placebo completers. Thus, the Observed Cases analysis and imputation analyses for the CIBIC-Plus may be slightly biased in favor of Memantine.

Table 3.11 CIBIC-Plus at Week 28 (LOCF) for Completers and Non-Completers

OC	Treatment	CIBIC-Plus Score N (%)						N	Mean (SD)
		2	3	4	5	6	7		
Yes	Placebo	3 (3.6)	7 (8.3)	24 (28.6)	29 (34.5)	17 (20.2)	4 (4.8)	84	4.74 (1.1)
Yes	Memantine	4 (4.1)	15 (15.5)	38 (39.2)	22 (22.7)	16 (16.5)	2 (2.1)	97	4.38 (1.1)
No	Placebo		1 (2.9)	13 (38.2)	11 (32.4)	7 (20.6)	2 (5.9)	34	4.88 (1.0)
No	Memantine			4 (19.0)	12 (57.1)	3 (14.3)	2 (9.5)	21	5.14 (0.9)

The difference in the percentage of patients not worse (CIBIC-Plus <=4) was not significant for the LOCF analysis (p=0.090). The percentages were 52 for memantine and 41 for placebo. The difference was significant for the observed cases analysis (p=0.014), but as noted above Memantine dropouts did worse than Memantine completers, while placebo dropouts and completers were more similar, so the observed cases analysis may be biased in favor of Memantine.

Secondary Endpoints

Severe Impairment Battery (Secondary)

For the Change in the SIB from Baseline to Week 28, this reviewer verified that the ITT (LOCF) and Observed Cases analyses both yielded p-values < 0.01. Since the results were very similar for the LOCF and Observed Cases analyses no further comparison of dropouts and completers was made. In terms of the mean change Memantine was numerically better in 22 of the 29 centers that had patients in both arms. Furthermore, the treatment effect on the change in the SIB also seems robust with respect to deletion of individual centers.

Table 3.12 shows the results for other secondary endpoints.

Table 3.12 Other Secondary Endpoints

Secondary Endpoints	Placebo Mean (SD)	Memantine Mean (SD)	Wilcoxon p value
NPI	3.63 (15.6)	0.44 (15.4)	0.371
MMSE	-1.14 (3.00)	-0.516 (2.38)	0.191
GDS	0.191 (0.468)	0.095 (0.464)	0.123
FAST	0.524 (1.35)	0.198 (1.22)	0.020

All of the other four secondary endpoints were numerically better for Memantine but three of the four were not significant at the 0.05 level. Thus, the treatment effect was significant for 3 of the 7 endpoints considered here (1 out of 2 primaries and 2 out of 5 secondaries).

3.1.3. Study 9403

3.1.3.1. Study Design

This was a randomized, placebo controlled, double-blind, multicenter trial. The trial was designed to enroll 150 care dependent patients between the ages of 60 and 80 who suffer from moderate to severe primary dementia. The study included patients with either primary degenerative or vascular dementia. Patients were to be treated for 12 weeks. Patients randomized to receive Memantine would receive 5 mg in the morning for the first week and 10 mg in the morning thereafter.

The primary efficacy variables are the CGI-C (clinical global impression of change) and the BGP - Care Dependence (Behavior Rating scale for geriatric patients). The BGP is an observer-rated scale for the assessment of functional disturbances of geriatric patients by the nursing staff. It assesses a patient's performance on the physical, psychological, and social level. The scale consists of 35 items which are divided into 4 subscales: Care dependence; Aggressiveness; Physical disability, Depression, Mental disability; and Inactivity. The CGI-C scores will be obtained at the end of Weeks 4 and 12 and are relative to the severity score (CGI-S) determined at baseline.

The BGP scores will be assessed at baseline and at the end of Weeks 1, 4, 8, and 12.

3.1.3.2. Statistical Analysis Plan

Sample Size

In case of a difference of delta=0.30 in the responder rate (CGI-C), starting from a success rate of ca. p=0.30 under placebo, a sample size of n1=n2=68 patients is required at alpha=0.025 and beta=0.10.

For a medication difference of 0.8 points with regard to the baseline difference of the BGP-Dimension "care dependence" a sample size of n1=n2=23 patients is required at alpha=0.025 and beta=0.10.

Thus, a total number of patients of N=136 is required to define the above medication differences at the 90% confidence level. Taking into account a 10% dropout rate, N=150 patients will have to be recruited for the trial.

Analysis Methods

The primary endpoints are the responder rate based on the CGI-C and the change from baseline in the BGP care dependence at the end of treatment. Patients with CGI-C scores between 1 and 3 will be considered responders. Fisher's Exact test will be used to check for treatment differences in the responder rates (dichotomized CGI-C). Changes from baseline in the BGP care dependence will be checked for treatment differences using the Wilcoxon-Mann-Whitney U tests.

According to Statistical Analysis Plan:

The stratified Wilcoxon rank-sum test [stratified by centers] will be carried out using SAS (Proc Freq, using the CMH test with modified ridit scores. P value will be obtained as the p-value for the row mean scores difference).

The primary analysis on CGI-C will be both the 7-point scale and the response rate at the end of the study (day 84, missing value imputed using LOCF method). Fisher's exact test and stratified Wilcoxon rank-sum test (stratified by centers) will be used to analyze the dichotomized CGI-C. The original 7-point CGI-C scale will be analyzed using the stratified (by center) Wilcoxon rank-sum test.

The primary analysis on BGP care-dependency and BGP cognitive will be the change from baseline at end of study (week 12, missing value imputed using LOCF method). Stratified Wilcoxon rank-sum test [stratified by trial centers] will be used.

3.1.3.3. Study Population

A total of 166 patients were randomized, 82 in the memantine group and 84 in the placebo group. A total of 158 patients (95%) completed the study. The dropout rate was 5% in both treatment groups. All randomized patients were included in the ITT population.

Table 3.13 Patient Disposition

	Placebo	Memantine	Total
Randomized	84	82	166
Completed	80 (95%)	78 (95%)	158 (95%)
Discontinued	4 (5%)	4 (5%)	8 (5%)

Patient demographics are given in Table 3.14. The average memantine patient was 71 years old and weighed 68 kg; 60% of memantine patients were female. Demographic characteristics for placebo patients were similar. The mean baseline MMSE score was 6.5 (range 0 to 9) in the memantine group and 6.1 (range 0 to 9) in the placebo group. Mean baseline scores on the BGP

care dependency subscale and the BGP cognitive subscale were also similar in the two treatment groups, reflecting a similar degree of functional and cognitive impairment.

Table 3.14 Patient Demographics – Study 9403

Demographic Parameter	Placebo (N=84)	Memantine (N=82)
AGE (YEARS)		
Mean ± SD	71.9 ± 6.1	71.2 ± 6.2
Range	60, 80	60, 81
< 65, n (%)	12 (14%)	15 (18%)
65-74, n(%)	40 (48%)	36 (44%)
≥ 75, n(%)	32 (38%)	31 (38%)
SEX, N(%)		
Male	37 (44%)	33 (40%)
Female	47 (56%)	49 (60%)
WEIGHT (KG)		
Mean ± SD	67.4 ± 11.4	67.9 ± 13.6
Range	48, 95	36, 100
MMSE		
Mean ± SD	6.1 ± 2.8	6.5 ± 2.6
Range	0, 9	0, 9
HIS		
Mean ± SD	5.7 ± 3.2	5.2 ± 2.9
Range	1, 12	1, 12
BASELINE BGP CARE		
Mean ± SD	21.8 ± 7.7	21.3 ± 7.6
BASELINE BGP COG		
Mean ± SD	5.4 ± 2.5	5.5 ± 2.6

A total of 79 DAT(Alzheimer's) patients (HIS score ≤ 4) were included in this study; 38 were treated with placebo and 41 were treated with memantine. Of these 79 patients, 76 (96%) completed the study. The discontinuation rate was 3% (1/38) in the placebo group and 5% (2/41)

in the memantine group. Demographic and Baseline characteristics of the DAT population were similar to those of the total population. The treatment groups were similar with respect to age, sex, weight, baseline MMSE score, baseline BGP care dependency score, and baseline BGP cognitive score.

3.1.3.4.Sponsor's Efficacy Results

3.1.3.4.1. Primary Efficacy

The mean change from baseline to endpoint (LOCF) and from baseline to week 12 (OC) on the BGP care dependency subscale is presented in the following table. Memantine was significantly superior to placebo (p=0.01) on the BGP care dependency subscale at endpoint (LOCF). In the memantine group, the mean BGP Care Dependency score decreased by 5.3 points from baseline. In the placebo group, the corresponding values decreased by 3.3 points from baseline. A similar statistically significant difference (p=0.01) favoring memantine was observed at week 12 (OC). The Memantine group showed significantly more improvement overall and in the subgroup of DAT patients.

Table 3.15 Change from Baseline in BGP Care Dependency

		Placebo		Memantine		
		N	Mean	N	Mean	p-value*
All Patients	Endpoint (LOCF)	84	-3.3	82	-5.3	0.01
	Week 28 (OC)	80	-3.5	78	-5.6	0.01
DAT Patients only	Endpoint (LOCF)	38	-2.8	41	-5.8	<0.01
	Week 28 (OC)	37	-2.9	39	-6.1	<0.01

* Wilcoxon Rank Sum Test (stratified by center)

Both groups improved over time as measured by the BGP care dependency. Numerically greater mean improvement was observed in the memantine group relative to the placebo group beginning at Week 4, and a statistically significant difference was demonstrated by Week 12.

Table 3.16 shows that the CGI-C scores were statistically significantly lower in the Memantine group than in the placebo group both overall and in the subgroup of DAT patients. In addition, a significantly greater proportion of patients treated with Memantine (77%) than those treated with placebo (48%) were classified as improved after 12 weeks. A significantly higher response rate

was also observed on the CGI-C at Week 4 and at study endpoint (LOCF). This suggests that there is a therapeutic benefit of memantine over placebo in the clinical global status of patients with dementia.

Table 3.16 Mean CGI-C

		Placebo		Memantine		
		N	Mean	N	Mean	p-value*
All Patients	Endpoint (LOCF)	84	3.5	82	3.1	<0.01
	Week 28 (OC)	80	3.5	78	3.0	<0.01
DAT Patients	Endpoint (LOCF)	38	3.5	41	3.2	<0.01
	Week 28 (OC)	37	3.5	39	3.1	<0.01

* Wilcoxon Rank Sum Test (stratified by center)

Secondary Endpoints

The BGP Cognitive was retrospectively identified as a third key measure of efficacy. Table 3.17 shows that Memantine patients fared better than placebo patients in terms of change from baseline in the BGP Cognitive.

Table 3.17 Change from Baseline in BGP Cognitive

		Placebo		Memantine		
		N	Mean	N	Mean	p-value*
All Patients	Endpoint (LOCF)	84	-1.1	82	-1.9	<0.01
	Week 28 (OC)	80	-1.2	78	-2.0	<0.01
DAT Patients only	Endpoint (LOCF)	38	-1.0	41	-2.0	<0.01
	Week 28 (OC)	37	-1.1	39	-2.1	<0.01

* Wilcoxon Rank Sum Test (stratified by center)

A correlation analysis was conducted to assess the extent to which changes in the BGP care dependency were attributable to the BGP cognitive subscale, which is a subset of the care dependency subscale. It was found that the correlation between the change in BGP-CD and the change in BGP-COG was 0.824. This analysis also revealed that the change in the cognitive subscale accounted for at least 65% of the variance in the change from baseline to Week 12 (LOCF) on the BGP Care Dependency. This finding suggests that the significantly greater improvement on the BGP care dependency subscale in memantine patients relative to placebo patients was largely dependent upon improvement in their cognitive abilities.

COMMENT: *This reviewer found that there were other subscales with the same number of items as the cognitive subscale (but consisting only of care dependency items not part of the cognitive subscale) for which the correlation between the changes was higher, the % of the variance of change in BGP-CD explained was higher and for which there was a more significant treatment difference. Thus, the sponsor's statement that the improvement on the care dependency subscale was largely dependent on improvement in cognitive abilities is suspect. For example, if we form a subscale by summing the responses to questions 4 (incontinent during day), 20 (able to socialize), 23 (cooperative), 25 (often repeats same movements), and 31 (needs assistance dressing) then the change in this subscale has a correlation coefficient of 0.850 with the change in the care dependency, explains 72 % of the variance and the p value for the test of the treatment effect on this subscale is 0.0032. The real issue is the relative importance of the various items or subscales which the sponsor's correlation analysis did not address. Towards this end a stepwise regression of change in BGP-CD with the changes in these two subscales as potential covariates was carried out. The subscale composed of items (4, 20, 23, 25, and 31) was included in the model before the cognitive subscale and adding the cognitive scale to the model only explained an additional 13% of the variance. Therefore, it is not clear that the improvement on the BGP care dependency was largely due to improvement on items contained in the BGP cognitive subscale.*

3.1.3.5. Reviewer's Comments

In the other studies patients tended to have worsened by the end of the study, according to the ADCS-ADL, SIB, and CIBIC-Plus scales. The results from this study are notable in that both Memantine and Placebo patients tended to improve over time according to the BGP and CGI-C scales. This could be explained by the study's shorter duration (12 weeks as compared to 24 or 28). Other notable differences of this study are a lower dose, higher minimum age, assisted living facility setting, a smaller sample size, and inclusion of patients with vascular dementia. This study had a total of 166 patients (79 DAT) as compared to 252 for study 9605 and 403 for MEM-MD-02. This reviewer verified the sponsor's primary analyses of the BGP-Care Dependency and the CGI-C. Center effects were important for both endpoints, i.e., ratings tended to be significantly higher (or lower) irrespective of treatment group assignment in some centers than others. For the BGP Care Dependency center average changes ranged from –6.04 to

−1.25 with a mean of −3.54 ± 1.92. Differences in treatment group means within centers ranged from −5.50 to 0.66 with a mean of −1.83 ± 2.31. For the CGI-C, differences in treatment group mean values at Week 12 within centers ranged from -1.00 to -0.04. Center average values at Week 12 ranged from 2.98 to 3.57.

Center 00005 had the largest difference between the treatment groups on all of the measures. Curiously, in center 00005 all 6 of the memantine patients had a final CGI-C score of 3 and all 6 of the placebo patients had a final score of 4, and no patients changed from baseline in terms of the CGI-Severity.

This reviewer also verified the result for the BGP-Cognitive which was retrospectively designated as a key endpoint. The BGP-Cognitive is a subset of the items in the BGP-Care Dependency. The sponsor's claim that the treatment difference observed for the Care Dependency was largely due to improvement in cognitive abilities was disputed in the comment at the end of the previous section because there were other non-cognitive subscales of the Care Dependency which one could make the same claim about. The real issue is the relative importance of the various subscales which the sponsor's correlation analysis did not address. Therefore, it is not clear that the cognitive items are most responsible for the treatment effect on the change in the BGP-CD.

3.2. Evaluation of Safety

See Clinical Review by Dr. Ranjit Mani.

4. Findings in Special/Subgroup Populations

4.1. Gender, Race, and Age

4.1.1. Gender

About 66% of all the patients studied were female. Overall, there was no consistent evidence that the treatment effect depended on gender.

About 65% of the patients in MEM-MD-02 were female. There were no significant differences in the gender specific treatment effects.

Table 4.1 MEM-MD-02: Mean Outcome Measures (LOCF) by Gender and Treatment

Variable	Group	Treatment Code	n	Baseline	Endpoint	Treatment Effect p value *
Primary						
ADL	Male	Placebo	63	35.0 (10.0)	-3.0 (5.8)	0.2038
	Male	Memantine	74	37.0 (9.0)	-1.1 (6.4)	
	Female	Placebo	134	36.9 (8.8)	-3.3 (6.2)	0.2100
	Female	Memantine	124	35.0 (10.2)	-2.3 (6.5)	
SIB	Male	Placebo	63	77.8 (14.1)	-2.3 (9.2)	0.1264
	Male	Memantine	74	76.1 (15.7)	1.1 (9.1)	
	Female	Placebo	133	80.7 (14.1)	-2.3 (8.9)	0.0009
	Female	Memantine	124	78.5 (15.5)	1.0 (7.2)	
Secondary						
CIBIC+	Male	Placebo	63		4.7 (1.1)	0.0679
	Male	Memantine	74		4.5 (1.0)	
	Female	Placebo	133		4.6 (1.1)	0.1107
	Female	Memantine	124		4.4 (1.1)	

* based on ANCOVA model containing effects for Treatment, Center, and Baseline Score

About 67% of all patients were female in study 9605. The treatment effect on the change in the ADL was slightly larger for males. On the other hand, for the CIBIC-Plus the difference in treatment group means was only 0.03 for males compared to 0.39 for females. There was virtually no gender difference in the treatment means for the SIB.

Table 4.2 Study 9605: Mean Outcome Measures (LOCF) by Gender and Treatment

Variable	Group	Treatment Code	n	Baseline	Endpoint	Treatment Effect p value *
Primary						
ADL	Male	Placebo	46	29.0 (10.6)	-5.9 (6.4)	0.0984
	Male	Memantine	35	25.8 (10.1)	-2.9 (5.9)	
	Female	Placebo	77	26.7 (11.0)	-4.8 (6.3)	0.0952
	Female	Memantine	89	27.1 (8.8)	-3.1 (7.2)	
CIBIC+	Male	Placebo	45		4.6 (1.1)	0.9015
	Male	Memantine	34		4.6 (1.1)	
	Female	Placebo	73		4.9 (1.1)	0.0236
	Female	Memantine	84		4.5 (1.2)	
Secondary						
SIB	Male	Placebo	46	70.5 (17.5)	-7.5 (9.2)	0.0164
	Male	Memantine	35	61.9 (24.3)	-1.1 (12.3)	
	Female	Placebo	77	67.2 (21.9)	-11.6 (15.3)	0.0018
	Female	Memantine	89	67.4 (22.0)	-5.1 (10.8)	

* Based on Wilcoxon Rank Sum Test

4.1.2. Race

Since more than 90% of the patients were white, no separate analyses on race were performed.

4.1.3. Age

About 60% of the patients were 75 years of age or older. In MEM-MD-02 the treatment effect does not seem to be linear as a function of age. The largest difference between treatment group means occurred in the 65-74 age group for the two primary endpoints, ADL and SIB, and the secondary endpoint, CIBIC-Plus. For the ADL there was a significant interaction between treatment and age whether age was treated as continuous or classified into groups. Two different classifications were explored: 50-64, 65-74, 75-84, 85-93 and 50-74 and 75-93. In the small subgroup (N=50) of patients who were 85 years of age or older the mean change in ADL was 3 points worse for memantine than for placebo. For the 38 completers aged 85 and older the mean change for placebo was 2.25 points better. This subgroup is small but memantine was essentially no better than placebo (-2.33 ± 6.78 compared to –2.39 ± 5.66) for the larger subgroup of patients aged 75 and older. This latter group constitutes more than half of the total population. However, for the other primary, SIB total, and the secondary CIBIC-Plus the mean in the 85+ subgroup was not numerically worse for Memantine nor was there a significantly lesser effect

for the 75+ subgroup.

Table 4.3 MEM-MD-02: Mean Outcome Measures (LOCF) by Age Group

Variable	Group	Treatment Code	n	Baseline	Endpoint	Treatment Effect p value *
Primary						
ADL	<=64	Placebo	28	38.0 (10.2)	-2.0 (4.6)	0.8497
	<=64	Memantine	26	37.7 (10.6)	-1.2 (5.7)	
	65-74	Placebo	47	37.1 (8.5)	-6.0 (6.9)	0.0099
	65-74	Memantine	53	35.5 (10.1)	-1.1 (6.0)	
	75-84	Placebo	95	36.0 (9.0)	-2.6 (6.0)	0.4367
	75-84	Memantine	96	35.8 (8.9)	-1.8 (6.2)	
	>=85	Placebo	27	34.1 (10.2)	-1.7 (4.4)	0.1945
	>=85	Memantine	23	33.4 (11.5)	-4.7 (8.5)	
SIB	<=64	Placebo	28	75.1 (17.2)	-3.2 (8.0)	0.4088
	<=64	Memantine	26	71.2 (21.0)	-0.5 (8.9)	
	65-74	Placebo	47	78.0 (14.3)	-5.4 (11.6)	0.0006
	65-74	Memantine	53	72.5 (17.1)	2.4 (8.2)	
	75-84	Placebo	94	80.4 (13.9)	-0.8 (7.8)	0.0543
	75-84	Memantine	96	81.2 (12.8)	0.4 (7.5)	
	>=85	Placebo	27	85.3 (8.5)	-1.2 (7.4)	0.6517
	>=85	Memantine	23	81.1 (10.9)	2.4 (8.1)	
Secondary						
CIBIC+	<=64	Placebo	28		4.5 (1.2)	0.4862
	<=64	Memantine	26		4.5 (1.1)	
	65-74	Placebo	47		5.1 (1.0)	0.0569
	65-74	Memantine	53		4.6 (0.9)	
	75-84	Placebo	94		4.6 (1.0)	0.1006
	75-84	Memantine	96		4.4 (1.1)	
	>=85	Placebo	27		4.3 (0.9)	0.5069
	>=85	Memantine	23		4.1 (1.1)	

* based on ANCOVA model containing effects for Treatment, Center, and Baseline Score

In study 9605 the mean age was 76 and 60% of the patients were 75 years of age or older. There were no consistent or striking differences in treatment effect between the age groups.

Table 4.4 Study 9605: Mean Outcome Measures (LOCF) by Age Group

Variable	Group	Treatment Code	n	Baseline	Endpoint	Treatment Effect p value *
Primary						
ADL	50-74	Placebo	50	30.9 (10.3)	-6.1 (6.5)	0.0253
	50-74	Memantine	49	26.3 (9.7)	-2.9 (6.9)	
	75-93	Placebo	73	25.3 (10.7)	-4.6 (6.1)	0.2064
	75-93	Memantine	75	27.0 (8.8)	-3.2 (6.8)	
					)	
CIBIC+	50-74	Placebo	47		4.8 (1.1)	0.5526
	50-74	Memantine	47		4.7 (1.2)	
	75-93	Placebo	71		4.8 (1.1)	0.0531
	75-93	Memantine	71		4.4 (1.1)	
Secondary						
SIB	50-74	Placebo	50	69.7 (19.3)	-11.4 (15.4)	0.0440
	50-74	Memantine	49	60.1 (24.7)	-4.7 (11.1)	
	75-93	Placebo	73	67.6 (21.2)	-9.2 (12.1)	0.0022
	75-93	Memantine	75	69.6 (20.6)	-3.6 (11.6)	

* based on Wilcoxon Rank Sum test

4.2. Other Special/Subgroup Populations

The baseline MMSE < 10 subgroup did statistically worse than the MMSE >= 10 subgroup on both the ADL and the SIB in studies 9605 and MEM-MD-02. As seen in the following two tables there is no consistent evidence of a MMSE dependent treatment effect though. In MEM-MD-02 where MMSE ranged from 5 to 16, the differences in group means for the SIB and ADL were numerically larger for the MMSE < 10 group while for the CIBIC-Plus the difference was numerically larger for the MMSE ≥ 10 group. The treatment effect was significant at the 0.05

level in the MMSE < 10 subgroup and the MMSE >= 10 subgroup for the SIB. On the other hand, in study 9605, where MMSE ranged from 3 to 14, the difference in treatment means was numerically larger in the MMSE >=10 group for both primary variables, ADL and CIBIC-Plus, and the secondary variable, SIB. The treatment effect was significant at the 0.05 level in the MMSE >= 10 subgroup for the ADL, the CIBIC-Plus, and the SIB, but in the MMSE < 10 subgroup only the SIB was significant. For the CIBIC-Plus the memantine group actually had a lower mean than placebo for MMSE <=5 but was higher for MMSE > 5. In particular, the 43 Memantine patients with MMSE <=5 had a mean CIBIC+ of 4.77 compared to 4.51 for the 41 placebo patients. For MMSE > 5 the 75 memantine patients had a mean of 4.37 while the 77 placebo patients had a mean of 4.92.

When comparing across studies we should remember that baseline scores were higher in MEM-MD-02. In particular, average ADCS-ADL baseline scores were about 7-8 points higher in MEM-MD-02 for both the MMSE < 10 and MMSE ≥ 10 groups. Average baseline SIB scores were about 12 points higher for the MMSE < 10 group and 2 points higher for the MMSE ≥ 10 group.

Table 4.5 MEM-MD-02: Mean Outcome Measures (LOCF) by MMSE and Treatment

Variable	Group	Treatment Code	n	Baseline	Endpoint	Treatment Effect p value *
Primary						
ADL	< 10	Placebo	72	32.4 (9.3)	-4.6 (6.1)	0.1682
	< 10	Memantine	89	33.0 (10.7)	-2.8 (7.6)	
	>= 10	Placebo	125	38.5 (8.5)	-2.4 (5.9)	0.0821
	>= 10	Memantine	109	37.9 (8.4)	-1.1 (5.3)	
SIB	< 10	Placebo	72	69.1 (14.5)	-6.2 (9.9)	0.0023
	< 10	Memantine	89	67.4 (15.4)	0.1 (9.8)	
	>= 10	Placebo	124	86.0 (9.3)	0.0 (7.6)	0.0450
	>= 10	Memantine	109	86.0 (9.7)	1.8 (6.0)	
Secondary						
CIBIC+	< 10	Placebo	72		4.9 (1.1)	0.0353
	< 10	Memantine	89		4.7 (1.0)	
	>= 10	Placebo	124		4.5 (1.0)	0.1209
	>= 10	Memantine	109		4.2 (1.0)	

* based on ANCOVA model containing effects for Treatment, Center, and Baseline Score

Table 4.6 Study 9605: Mean Outcome Measures (LOCF) by MMSE and Treatment

Variable	Group	Treatment Code	n	Baseline	Endpoint	Treatment Effect p value *
Primary						
ADL	<10	Placebo	73	25.5 (11.9)	-5.6 (6.5)	0.2643
	<10	Memantine	79	24.3 (9.0)	-4.5 (6.7)	
	>= 10	Placebo	50	30.7 (8.4)	-4.6 (6.1)	0.0080
	>= 10	Memantine	45	31.0 (7.8)	-0.6 (6.4)	
CIBIC+	<10	Placebo	70		4.8 (1.1)	0.5341
	<10	Memantine	75		4.7 (1.1)	
	>= 10	Placebo	48		4.8 (1.1)	0.0206
	>= 10	Memantine	43		4.2 (1.1)	
Secondary						
SIB	< 10	Placebo	73	58.0 (19.4)	-11.8 (14.0)	0.0082
	< 10	Memantine	79	55.0 (20.4)	-5.8 (12.6)	
	>= 10	Placebo	50	83.7 (8.8)	-7.6 (12.5)	0.0073
	>= 10	Memantine	45	84.8 (11.3)	-0.8 (7.9)	

* based on Wilcoxon Rank Sum Test

5. Summary and Conclusions

5.1 Statistical Issues and Collective Evidence

For each of the MMSE, ADL Total, and SIB Total scales, higher scores indicate less impairment. From the following table we see that baseline values for MMSE, ADL Total, and SIB Total were larger in MEM-MD-02 than Study 9605. This may be attributable to the fact that MEM-MD-02 patients were on stable doses of Donepezil prior to the study and during the study. It is also noteworthy that, despite the suggestion of a downward trend in the last 12 weeks, Memantine patients improved slightly over the course of the study in terms of the SIB. Also, note that the average baseline BGP-Care Dependency was about 10 points higher, indicating more care dependence, in study 9403 than MEM-MD-02.

Table 5.1

	Endpoint	Treatment Code	n	MMSE	Baseline	Change
MD02	ADL Total	Placebo	193	10.2 (3.0)	36.2 (9.3)	-3.2 (6.1)
MD02	ADL Total	Memantine	198	9.9 (3.1)	35.9 (9.8)	-1.8 (6.5)
MD02	SIB Total	Placebo	192	10.2 (3.0)	79.7 (14.2)	-2.4 (9.0)
MD02	SIB Total	Memantine	198	9.9 (3.1)	77.8 (15.5)	1.1 (7.9)
MD02	BGP-CD	Placebo	179	10.2 (3.0)	9.1 (6.0)	2.2 (4.8)
MD02	BGP-CD	Memantine	185	9.9 (3.1)	8.8 (5.8)	0.8 (4.4)
9605	ADL Total	Placebo	123	8.0 (3.5)	27.6 (10.9)	-5.2 (6.3)
9605	ADL Total	Memantine	124	7.7 (3.8)	26.8 (9.1)	-3.1 (6.8)
9605	SIB Total	Placebo	123	8.0 (3.5)	68.4 (20.4)	-10.1 (13.5)
9605	SIB Total	Memantine	124	7.7 (3.8)	65.8 (22.7)	-4.0 (11.3)
9403	BGP-CD	Placebo	84	6.8 (2.4)	21.8 (7.7)	-3.3 (5.2)
9403	BGP-CD	Memantine	82	6.7 (2.6)	21.3 (7.6)	-5.3 (5.1)

Both studies 9605 and MEM-MD-02 exhibited statistical significance for the changes from baseline in the Severe Impairment Battery and the modified Alzheimer's Disease Cooperative Study - Activities of Daily Living Total scores. These were co-primaries for study MEM-MD-02 but the ADCS-ADL and CIBIC-Plus were co-primaries in study 9605, where the SIB was a secondary endpoint. In study 9605 the pre-specified primary analysis of the CIBIC-Plus was not quite significant (p=0.06). So, technically, the study did not meet the criteria for a win. The observed cases population did show a significant treatment effect on the CIBIC-Plus, but dropouts seem to have fared worse than completers, particularly, in the Memantine group. Thus, the Observed Cases population does not give the complete picture and may be slightly biased in favor of Memantine. Further investigation showed that a center stratified analysis (not protocol specified) for the CIBIC-Plus resulted in an even larger p-value (0.095).

Study 9403 was positive but different from MEM-MD-02 and 9605 in several important ways. First, the sample size was smaller (166 total compared to 252 and 403) and it included patients with Vascular dementia (slightly more than 50% of all patients). Second, the length of observation was only 12 weeks compared to 24 and 28 in the other studies. Finally, the primary endpoints in 9403, the Care Dependency subscale of the Behavior Rating scale for Geriatric patients and the CGI-Change, were different from those in the other studies. The BGP care dependency was collected in MEM-MD-02 and the treatment effect was significant however, unlike the shorter study 9403 the scores had worsened rather than improved by the end of the study. Keeping these study differences in mind we note that the results in 9403 were significant for both endpoints, even in the subgroup of Alzheimer's patients.

5.2 Conclusions and Recommendations

There was a significant treatment effect ($p<0.001$) on the change in the Severe Impairment Battery Total score, a cognitive measure, in both the add-on study (MEM-MD-02) and the monotherapy study (9605). The difference in treatment mean changes was 3.5 for MD-02 and 6.1 for 9605. Although the effects were smaller on the change in the ADCS-ADL Total Score (a functional scale), 1.4 and 2.1 respectively, they were also significant. The CIBIC-Plus was not quite significant for the ITT population in study 9605, where it had been designated as a primary endpoint instead of the SIB, but was significant for the Observed Cases population. Study 9403 appears very positive, but was of shorter duration, did not have a pre-specified cognitive measure, and enrolled both Alzheimer's and Vascular dementia patients. Thus, overall, except for a marginally significant result on one of the two primary endpoints in one of the two main studies, the data supports the sponsor's efficacy claim.

/s/

Tristan Massie, Ph.D.
Statistical Reviewer

Kun Jin, Ph.D.
Team Leader

Kooros Mahjoob, Ph.D.
Deputy Director, Division of Biometrics

I

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

/s/

Tristan Massie
9/29/03 12:08:32 PM
BIOMETRICS

Kun Jin
9/29/03 01:57:11 PM
BIOMETRICS

Kooros Mahjoob
9/29/03 02:33:34 PM
BIOMETRICS

Statistical Review and Evaluation

Review of Mouse and Rat Carcinogenicity Studies

NDA #:	21-487
APPLICANT:	Forest Laboratories Inc.
NAME OF DRUG:	Memantine HCl Tablets
INDICATION:	Moderate to Severe Dementia of the Alzheimer's Type
STUDIES REVIEWED:	Study 7279, carcinogenicity study in mice, and Study 7280, carcinogenicity study in rats; Data were submitted electronically
STATISTICAL REVIEWER:	Rajeshwari Sridhara, Ph.D. (HFD-710)
PHARMACOLOGY REVIEWER:	Kathleen Haberry, Ph.D. (HFD-120)

Note on Levels of Statistical Significance:

Trends in inter-current mortality are tested for statistical significance at $\alpha = 0.05$, two-sided. Trends in tumor incidences rates are tested one-sided for statistical significance at $\alpha = 0.025$ and 0.005 for rare and common tumors, respectively. These levels of significance ensure an overall false positive rate of about 10 % in the two-year, two-species, two-gender bioassay despite the multiplicity of testing. If pair-wise comparisons of tumor incidences are performed as well, they are tested one-sided at $\alpha = 0.05$ and 0.01 for rare and common tumors, respectively. The definition of rare ($\leq$ 1.0%) or common (> 1.0%) is based on the occurrence rate among the concurrent controls. It is possible that some rare tumor findings loose their statistical significance if the tumors are re-classified as common based on historical evidence.

1.0 The Mouse Study (Study 7279)

1.1 Sponsor's Results

Fifty male and fifty female B6C3F1 mice each received Memantine-HCL at dose levels of 0, 0, 2.5, 10.0 and 40.0 mg/kg b.w. per day. The actual dose levels were within a range of ± 10% of the nominal values. Duration of dosing was 113 weeks in both females and males.

The type and incidence of neoplastic lesions did not differ between treated and control groups. There was no evidence for a carcinogenic effect of Memantine-HCL up to 40.0 mg/kg b.w., for either sex. The sponsor observed no toxicologically significant effect on mortality or bodyweights.

Reviewer's Comments:

1. This reviewer agrees with the sponsor that the administration of Memantine HCL at the dose levels studied did not affect survival of the mice of either gender (Tables 1, 2, 4, 5, and Figures 1 and 2).
2. This reviewer confirms that there are no statistically significant differences in the incidence of neoplastic lesions between the treated and control groups (Tables 3 and 6) of either sex.
3. During the preparation of the carcinogenicity study electronic database, the sponsor noted some inconsistencies in the final report. A peer review of the initial histopathology findings was conducted by a pathologist from [redacted] where the study had been conducted. The pathology peer review revealed major differences in the incidence of malignant lymphoma and lymphoid lesions in each of the control and treated groups for both sexes of mice. At the completion of this re-evaluation, all lymphoid neoplasms diagnosed by the re-evaluating pathologist were further peer reviewed. Where differences existed, a consensus diagnosis was reached between the re-evaluating and the peer reviewing pathologists. The data in this report represent the consensus diagnoses between these

pathologists. Furthermore, the differences in the pathology reviews would not alter the overall conclusions.

2.0 Validity of the Mouse Study

As there were no statistically significant (positive) tumor findings among the female and male mice, the validity of this study needs to be evaluated. In order to address this issue, two questions need to be answered (Haseman, Statistical Issues in the Design, Analysis and Interpretation of Animal Carcinogenicity Studies, Environmental Health Perspectives, Vol. 58, pp. 385-392, 1984):

(I) Were enough animals exposed for a sufficient length of time to allow for late developing tumors?
(II) Were the dose levels high enough to pose a reasonable tumor challenge in the animals?

The following rules of thumb are suggested by experts in the field: Haseman (Issues in Carcinogenicity Testing: Dose Selection, Fundamental and Applied Toxicology, Vol. 5, pp. 66-78, 1985) had found that on the average, approximately 50% of the animals in the high dose group survived a two-year study. In a personal communication with Dr. Karl Lin (HFD-715), he suggested that 50% survival of the usual 50 initial animals in the high dose group between weeks 80-90 would be considered a sufficient number and adequate exposure. Chu, Cueto, and Ward (Factors in the evaluation of 200 National Cancer Institute Carcinogen Bioassays, Journal of Toxicology and environmental Health, Vol. 8, pp. 251-280, 1981) proposed that " to be considered adequate, an experiment that has not shown a chemical to be carcinogenic should have groups of animals with greater than 50% survival at one year'. From these sources, it appears that the proportions of survival at weeks 52, 80-90, and at two years are of interest in determining the adequacy of exposure and number of animals at risk.

In determining the adequacy of the chosen dose levels, it is generally accepted that the high dose should be close to the MTD. Chu, Cueto, and Ward (Factors in the evaluation of 200 National Cancer Institute Carcinogen Bioassays, Journal of Toxicology and environmental Health, Vol. 8, pp. 251-280, 1981) suggest:

(i) 'A dose is considered adequate if there is a detectable weight loss of up to 10% in a dosed group relative to the controls'.
(ii) 'The administered dose is also considered an MTD if dosed animals exhibit clinical signs or severe histopathologic toxic effects attributed to the chemical'.
(iii) 'In addition, doses are considered adequate if the dosed animals show a slightly increased mortality compared to the controls'.

In another paper, Bart, Chu and Tarone (Statistical Issues in Interpretation of Chronic Bioassay Tests for Carcinogenicity, Journal of the National Cancer Institute Vol. 62, pp.

957-974, 1979), stated that the mean body weight curves over the entire study period should be taken into consideration with the survival curves, when adequacy of dose levels is to be examined. In particular, 'Usually, the comparison should be limited to the early weeks of a study when no or little mortality has yet occurred in any of the groups. Here a depression of the mean weight in the treated groups is an indication that the treatment has been tested on levels at or approaching the MTD'.

Reviewer's Comments:

1. There were more than 50% of the mice of both sexes alive at 112 weeks (Tables 1 and 4) and therefore there were sufficient number of animals exposed long enough.
2. The reviewer agrees with the sponsor that the administration of Memantine HCL at the dose levels studied did not affect survival of mice of either gender (Tables 1, 2, 4, 5, and Figures 1 and 2).
3. In both female and male mice, body weights were not influenced (< 10% differential) by treatment up to highest dose of 40 mg/kg b.w. at 1 and 2 years.

Based on these criteria, it appears that neither the male nor the female mice were challenged at a level of the maximum tolerated dose.

3.0 The Rat Study (Study 7280)

3.1 Sponsor's Results

Fifty male and fifty female Sprague-Dawley rats each received Memantine-HCL at dose levels of 0, 0, 2.5, 10.0 and 40.0 mg/kg b.w. per day. Because of markedly reduced body weights compared to the controls, the high dose of both sexes was reduced to 20.0 mg/kg from test week 71 on. The actual dose levels were within a range of ± 10% of the nominal values. Duration of dosing was 128 weeks in females and 129 weeks in males.

The type and incidence of neoplastic lesions did not differ between treated and control groups. There was no evidence for a carcinogenic effect of Memantine-HCL up to 40.0 mg/kg b.w., for either sex. The sponsor observed no toxicologically significant effect on mortality.

Reviewer's Comments:

1. This reviewer used an adjusted dose of 32.5 mg/kg as the high dose in the statistical analyses.
2. This reviewer agrees with the sponsor that the administration of Memantine-HCL at the dose levels studied did not affect survival of the rats of either gender (Tables 7, 8, 10, 11, and Figures 3 and 4).
3. This reviewer confirms that there are no statistically significant differences in the incidence of neoplastic lesions between the treated and control groups of rats (Tables 9 and 12).

4.0 Validity of the Rat Study

As there were no statistically significant (positive) tumor findings among the female and male rats, the validity of this study needs to be also evaluated. In order to address this issue, two questions need to be answered:

(III) Were enough animals exposed for a sufficient length of time to allow for late developing tumors?
(IV) Were the dose levels high enough to pose a reasonable tumor challenge in the animals?

These questions were addressed by the criteria presented in section 2.0.

Reviewer's Comments:

1. There were more than 50% of female rats alive at 112 weeks and more than 50% of male rats alive at 128 weeks (Tables 7 and 10), therefore satisfying the criteria of sufficient number of animals and sufficient length of exposure.
2. The reviewer agrees with the sponsor that the administration of Memantine-HCL at the dose levels studied did not affect survival of the rats of either gender (Tables 7, 8, 10, 11, and Figures 3 and 4).
3. In both female and male rats, body weights of the high-dosed animals (40.0/20.0 mg/kg b.w.) were below the values of the control animals from the start (approximately 5%). At 51 weeks, the percent differences in average body weight between the high dose and control groups were 13.2% in males and 16.1% in females. At 103 weeks, the percent differences in average weight between the high dose and control groups were 7% in males and 17.1% in females.

Based on these criteria and given the differences at the start of the study, it appears that the male rats were not challenged up to a maximum tolerated dose. The reduction of the high dose level from 40.0 mg/kg to 20.0 mg/kg may have been too great for this gender.

5.0 Conclusion

Memantine-HCL did not show any neoplastic properties at any of the tested dose levels in either mice or rats. However, it appears that the high dose did not reach the MTD for either gender of the mice and for the male rats. Otherwise, the studies appear adequate and lasted an unusually long time.

Table 1: Analysis of Mortality

Species: Mouse, Sex: Female, NDA 21487

Analysis of Mortality		No. Risk	No. Died	No. Alive	Pct Survival	Pct Mortality
CTR1	0-52	50	1	49	98.0	2.0
	79-91	49	1	48	96.0	4.0
	92-112	48	9	39	78.0	22.0
	FINALKILL113-113	39	39	0		
CTR2	79-91	50	4	46	92.0	8.0
	92-112	46	10	36	72.0	28.0
	FINALKILL113-113	36	36	0		
LOW	53-78	50	2	48	96.0	4.0
	79-91	48	5	43	86.0	14.0
	92-112	43	8	35	70.0	30.0
	FINALKILL113-113	35	35	0		
MED	0-52	50	1	49	98.0	2.0
	53-78	49	3	46	92.0	8.0
	79-91	46	3	43	86.0	14.0
	92-112	43	7	36	72.0	28.0
	FINALKILL113-113	36	36	0		
HIGH	79-91	50	1	49	98.0	2.0
	92-112	49	14	35	70.0	30.0
	FINALKILL113-113	35	35	0		

Figure 1: Kaplan-Meier Graph of Survival

Species: Mouse, Sex: Female, NDA 21487

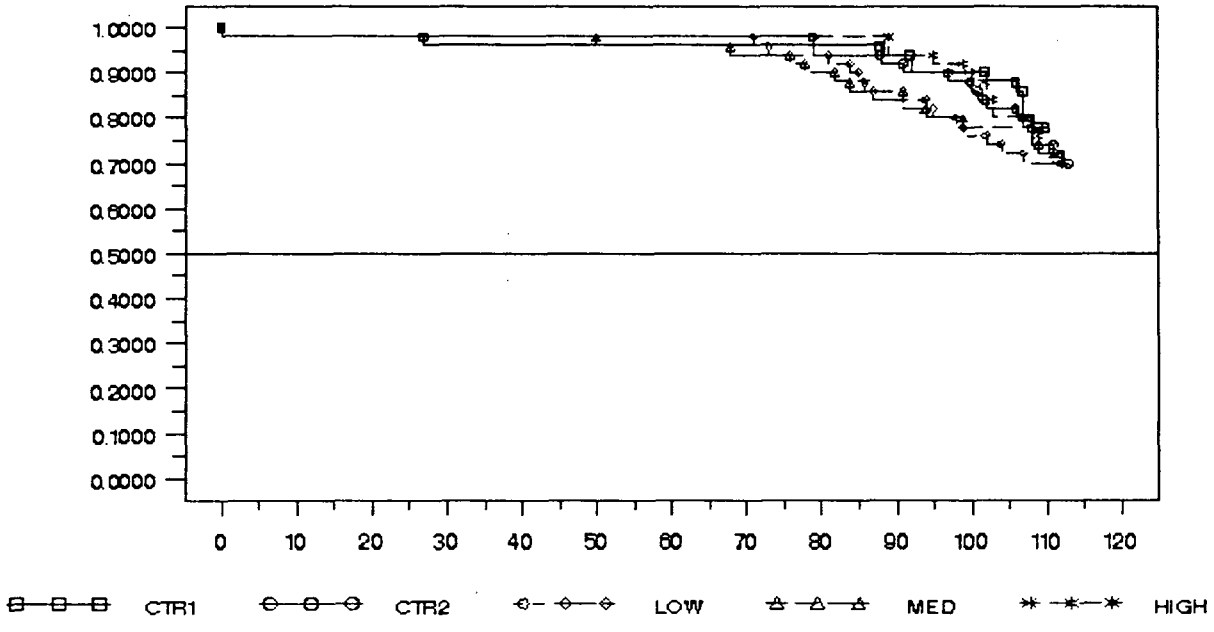


Table 2: Analysis of Dose-Mortality Trend

Species: Mouse, Sex: Female, NDA 21487

This test is run using Trend and Homogeneity Analyses of Proportions and Life Table Data Version 2.1, by Donald G. Thomas, National Cancer Institute

	Method			
	Cox		Kruskal-Wallis	
	Statistics	P-Value	Statistics	P-Value
Time-Adjusted Trend Test				
Depart from Trend	1.1616	0.7622	1.6036	0.6586
Dose-Mortality Trend	0.1043	0.7467	0.0280	0.8671
Homogeneity	1.2659	0.8671	1.6316	0.8031

Table 3: Report on Trend Test

Detecting significant positive dose-tumor linear trend

Species: Mouse, Sex: Female, NDA 21487

Organ Code	Organ Name	Tumor Code	Tumor Name	CT R1	CT R2	LO W	ME D	HIG H	P-Value (Exact Method)	P-Value (Asymptotic Method)	Link 2XC Table >
1	adrenals	23	pheochromocytoma-b	0	1	0	0	0	1.0000	0.7625	1
1	adrenals	4	adenoma-b, cortical (type B)	1	0	0	0	0	1.0000	0.7625	2
1	adrenals	42	pheochromocytoma-m	0	0	0	0	1	0.2979	0.0696	3 *
15	harderian gland	1	adenocarcinoma-m	0	0	1	0	0	0.5957	0.7404	4 *
15	harderian gland	2	adenoma-b	1	2	1	3	3	0.1904	0.1960	5 *
17	hematopoietic system	12	lymphoma-m, lymphocytic	3	1	1	2	2	0.4090	0.4241	6 *
17	hematopoietic system	14	lymphoma-m, pleomorphic	14	18	10	15	7	0.9738	0.9695	7 *
17	hematopoietic system	29	sarcoma-m, histiocytic	0	0	4	0	0	0.8672	0.8533	8 *
17	hematopoietic system	48	leukemia-m, granulocytic	0	0	1	0	0	0.5867	0.7166	9 *
22	liver	10	adenoma-b, hepatocellular	4	0	3	2	5	0.1152	0.1079	10 *
22	liver	22	fibrosarcoma-m	0	0	0	1	0	0.3923	0.5187	11 *
22	liver	8	carcinoma-m, hepatocellular	1	1	2	0	2	0.4027	0.3866	12 *
26	lungs	1	adenocarcinoma-m	1	0	1	3	0	0.6303	0.7028	13 *
26	lungs	3	adenoma-b, bronchial-alveolar	1	1	0	1	2	0.1640	0.1238	14 *
27	mammary gland	1	adenocarcinoma-m	1	1	2	3	0	0.7993	0.8232	15 *
33	ovaries	2	adenoma-b	0	1	0	1	0	0.6346	0.6976	16 *
33	ovaries	24	fibroma-b	0	0	1	0	0	0.6154	0.6336	17 *
33	ovaries	25	granulosa cell tumor-b	0	1	1	1	1	0.3259	0.3565	18 *
33	ovaries	45	adenoma-b, tubulostromal	0	0	0	0	1	0.1944	0.0262	19 *
34	pancreas	20	adenoma-b, islet cell	0	3	0	0	0	1.0000	0.9077	20
36	pituitary	1	adenocarcinoma-m	0	1	0	0	0	1.0000	0.7504	21
36	pituitary	2	adenoma-b	9	9	8	6	2	0.9820	0.9758	22 *
44	skin	22	fibrosarcoma-m	1	0	0	1	1	0.2995	0.2850	23 *
44	skin	24	fibroma-b	0	0	1	0	0	0.6042	0.7478	24 *
44	skin	31	carcinoma-m, squamous cell	0	0	2	0	1	0.2747	0.2767	25 *
48	stomach	1	adenocarcinoma-m	1	0	0	0	0	1.0000	0.7617	26
51	thyroids	5	adenoma-b, follicular cell	0	0	0	1	0	0.3923	0.5187	27 *
55	uterus (incl. cervix)	1	adenocarcinoma-m	0	2	0	0	2	0.1824	0.1089	28 *
55	uterus (incl. cervix)	18	sarcoma-m	1	1	1	1	2	0.2234	0.2039	29 *
55	uterus (incl. cervix)	22	fibrosarcoma-m	1	0	0	0	0	1.0000	0.7917	30
55	uterus (incl. cervix)	24	fibroma-b	0	0	0	0	1	0.1934	0.0258	31 *

55	uterus (incl. cervix)	46	leiomyoma-b	0	0	0	2	0	0.3889	0.5172	32 *
56	vagina	22	fibrosarcoma-m	0	0	0	0	1	0.1955	0.0266	33 *
56	vagina	24	fibroma-b	0	1	0	0	1	0.3537	0.1875	34 *
62	skull	27	osteosarcoma-m	1	0	0	0	0	1.0000	0.7917	35
71	systemic tumor	6	hemangioma-b	0	1	0	1	0	0.5982	0.6898	36 *
71	systemic tumor	7	hemangiosarcoma-m	2	1	1	2	1	0.5732	0.6175	37 *
73	clitoral gland	47	carcinoma-m	0	0	0	1	0	0.3923	0.5187	38 *
8	ears	18	sarcoma-m	0	0	1	0	0	0.5856	0.7073	39 *

Note: The symbol "*" indicates that the p-values fall in (0, 1).
The check mark indicates statistically significant test results,
based on the decision rule of **FDA.CDER.Divisions of Biometrics.**

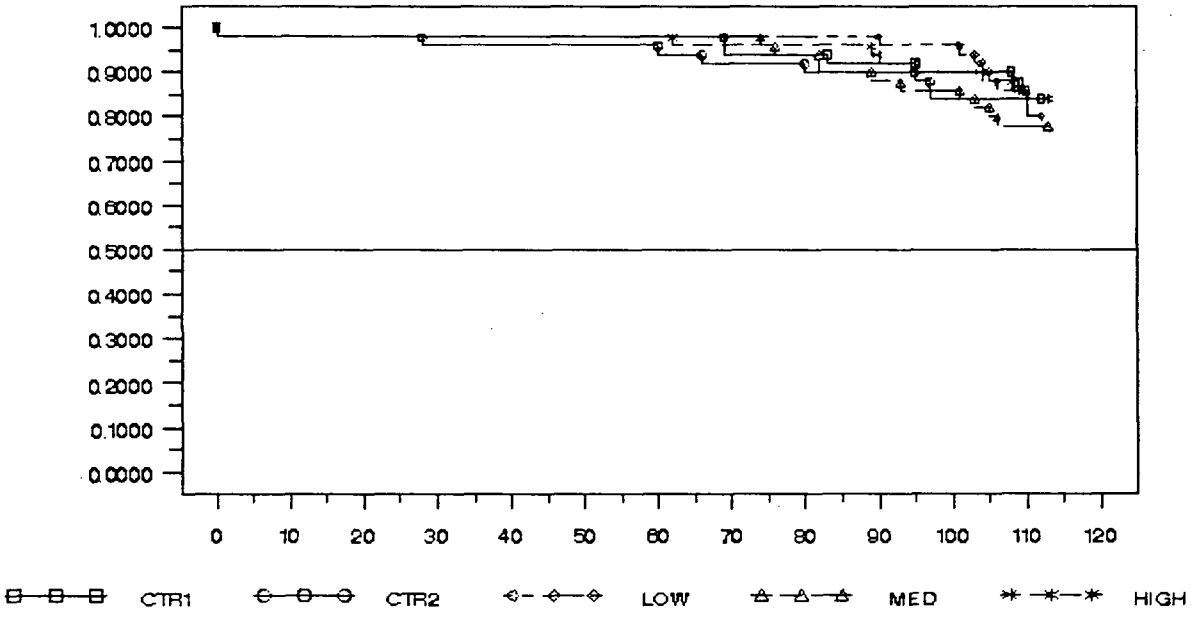
Table 4: Analysis of Mortality

Species: Mouse, Sex: Male, NDA 21487

Analysis of Mortality		No. Risk	No. Died	No. Alive	Pct Survival	Pct Mortality
CTR1	53-78	50	1	49	98.0	2.0
	79-91	49	2	47	94.0	6.0
	92-112	47	5	42	84.0	16.0
	FINALKILL113-113	42	42	0		
CTR2	0-52	50	1	49	98.0	2.0
	53-78	49	2	47	94.0	6.0
	79-91	47	1	46	92.0	8.0
	92-112	46	4	42	84.0	16.0
	FINALKILL113-113	42	42	0		
LOW	79-91	50	1	49	98.0	2.0
	92-112	49	9	40	80.0	20.0
	FINALKILL113-113	40	40	0		
MED	53-78	50	2	48	96.0	4.0
	79-91	48	3	45	90.0	10.0
	92-112	45	5	40	80.0	20.0
	FINALKILL113-113	40	40	0		
HIGH	53-78	50	1	49	98.0	2.0
	79-91	49	2	47	94.0	6.0
	92-112	47	4	43	86.0	14.0
	FINALKILL113-113	43	43	0		

Figure 2: Kaplan-Meier Graph of Survival

Species: Mouse, Sex: Male, NDA 21487



APPEARS THIS WAY
ON ORIGINAL

Table 5: Analysis of Dose-Mortality Trend

Species: Mouse, Sex: Male, NDA 21487

This test is run using Trend and Homogeneity Analyses of Proportions and Life Table Data Version 2.1, by Donald G. Thomas, National Cancer Institute

	Method			
	Cox		Kruskal-Wallis	
	Statistics	P-Value	Statistics	P-Value
Time-Adjusted Trend Test				
Depart from Trend	0.7210	0.8683	0.7243	0.8675
Dose-Mortality Trend	0.2244	0.6357	0.2002	0.6546
Homogeneity	0.9454	0.9180	0.9245	0.9210

Table 6: Report on Trend Test

Detecting significant positive dose-tumor linear trend

Species: Mouse, Sex: Male, NDA 21487

Organ Code	Organ Name	Tumor Code	Tumor Name	CT R1	CT R2	LO W	ME D	HIG H	P-Value (Exact Method)	P-Value (Asymptotic Method)	Link 2XC Table >
1	adrenals	16	adenoma-b, cortical (type A)	0	1	0	0	0	1.0000	0.7670	1
1	adrenals	23	pheochromocytoma-b	0	0	0	1	0	0.4029	0.5326	2 *
1	adrenals	4	adenoma-b, cortical (type B)	3	4	3	2	0	0.9871	0.9751	3 *
15	harderian gland	1	adenocarcinoma-m	0	0	0	1	0	0.4010	0.5316	4 *
15	harderian gland	2	adenoma-b	4	5	5	4	4	0.6231	0.6342	5 *
17	hematopoietic system	12	lymphoma-m, lymphocytic	2	2	0	1	0	0.9281	0.9018	6 *
17	hematopoietic system	14	lymphoma-m, pleomorphic	5	5	3	7	5	0.4390	0.4484	7 *
17	hematopoietic system	29	sarcoma-m, histiocytic	1	1	1	0	0	0.9362	0.8690	8 *
22	liver	10	adenoma-b, hepatocellular	12	12	12	11	8	0.8953	0.8925	9 *
22	liver	8	carcinoma-m, hepatocellular	11	8	8	13	8	0.6344	0.6406	10 *
26	lungs	1	adenocarcinoma-m	5	4	3	3	1	0.9472	0.9359	11 *
26	lungs	3	adenoma-b, bronchial-alveolar	5	4	6	1	0	0.9989	0.9921	12 *
34	pancreas	20	adenoma-b, islet cell	0	0	1	0	0	0.6923	0.6939	13 *
34	pancreas	22	fibrosarcoma-m	0	0	1	0	0	0.6923	0.6939	14 *
4	cecum	44	leiomyosarcoma-m	1	0	0	0	0	1.0000	0.7662	15
44	skin	11	lipoma-b	0	0	0	0	1	0.1481	0.0113	16 *
48	stomach	19	papilloma-b, squamous cell	1	0	0	0	0	1.0000	0.7666	17
49	testes	17	Leydig cell tumor-b	1	1	0	0	0	1.0000	0.8448	18
49	testes	32	seminoma-b	0	0	0	1	0	0.4010	0.5316	19 *
51	thyroids	5	adenoma-b, follicular cell	1	0	0	0	0	1.0000	0.7666	20
7	duodenum	35	polyp-b, tubular adenomatous	0	0	0	0	1	0.1000	0.0023	21 *
71	systemic tumor	6	hemangioma-b	2	1	0	1	0	0.8808	0.8517	22 *
71	systemic tumor	7	hemangiosarcoma-m	3	3	3	1	2	0.7085	0.7131	23 *

Note: The symbol "*" indicates that the p-values fall in (0, 1).
The check mark indicates statistically significant test results,
based on the decision rule of **FDA.CDER.Divisions of Biometrics**.

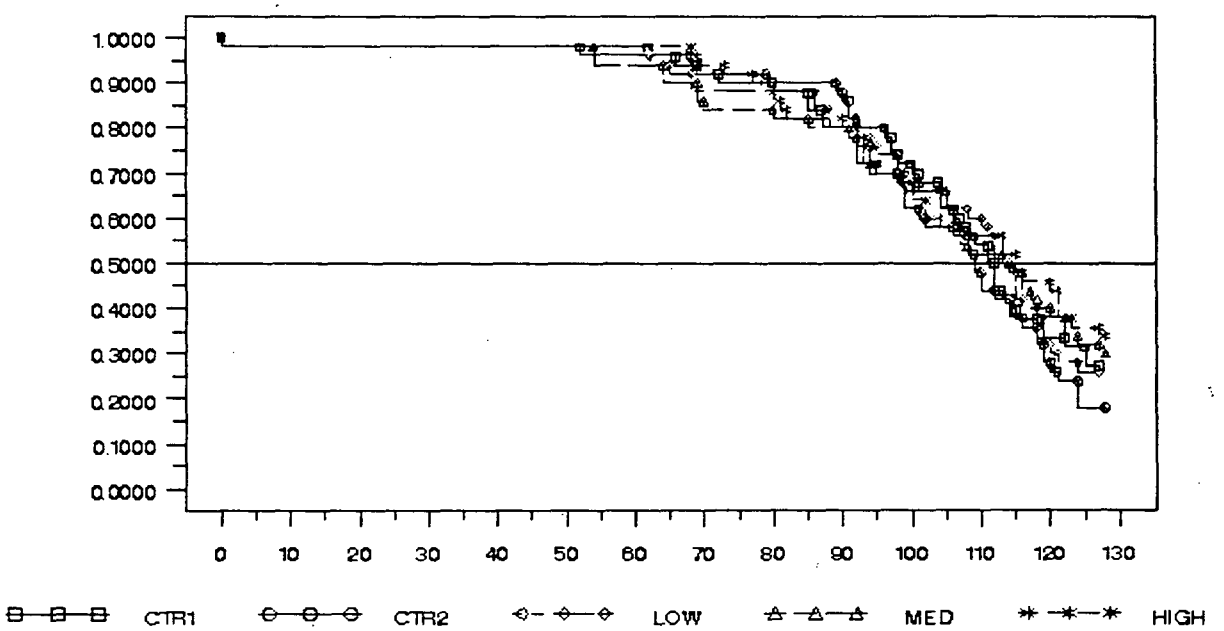
Table 7: Analysis of Mortality

Species: Rat, Sex: Female, NDA 21487

Analysis of Mortality		No. Risk	No. Died	No. Alive	Pct Survival	Pct Mortality
CTR1	53-78	50	4	46	92.0	8.0
	79-91	46	4	42	84.0	16.0
	92-112	42	17	25	50.0	50.0
	FINALKILL113-128	25	25	0		
CTR2	0-52	50	1	49	98.0	2.0
	53-78	49	2	47	94.0	6.0
	79-91	47	4	43	86.0	14.0
	92-112	43	21	22	44.0	56.0
	FINALKILL113-128	22	22	0		
LOW	53-78	50	4	46	92.0	8.0
	79-91	46	4	42	84.0	16.0
	92-112	42	14	28	56.0	44.0
	FINALKILL113-128	28	28	0		
MED	53-78	50	7	43	86.0	14.0
	79-91	43	3	40	80.0	20.0
	92-112	40	13	27	54.0	46.0
	FINALKILL113-128	27	27	0		
HIGH	53-78	50	4	46	92.0	8.0
	79-91	46	5	41	82.0	18.0
	92-112	41	12	29	58.0	42.0
	FINALKILL113-128	29	29	0		

Figure 3: Kaplan-Meier Graph of Survival

Species: Rat, Sex: Female, NDA 21487



APPEARS THIS WAY
ON ORIGINAL

Table 8: Analysis of Dose-Mortality Trend

Species: Rat, Sex: Female, NDA 21487

This test is run using Trend and Homogeneity Analyses of Proportions and Life Table Data Version 2.1, by Donald G. Thomas, National Cancer Institute

	Method			
	Cox		Kruskal-Wallis	
	Statistics	P-Value	Statistics	P-Value
Time-Adjusted Trend Test				
Depart from Trend	2.2599	0.5202	1.0412	0.7913
Dose-Mortality Trend	0.4063	0.5238	0.1937	0.6598
Homogeneity	2.6662	0.6151	1.2350	0.8723

APPEARS THIS WAY
ON ORIGINAL

Table 9: Report on Trend Test

Detecting significant positive dose-tumor linear trend

Species: Rat, Sex: Female, NDA 21487

Organ Code	Organ Name	Tumor Code	Tumor Name	CT R1	CT R2	LO W	ME D	HIG H	P-Value (Exact Method)	P-Value (Asymptotic Method)	Link 2XC Table >
1	adrenals	23	pheochromocytoma-m	1	0	0	0	0	1.0000	0.7947	1
1	adrenals	41	carcinoma-m, cortical	0	0	0	1	0	0.4381	0.5164	2 *
1	adrenals	45	adenoma-b, cortical	1	3	2	2	0	0.9392	0.9306	3 *
1	adrenals	72	pheochromocytoma-b	0	2	1	3	1	0.4516	0.4534	4 *
110	systemic tumor	56	mesothelioma-m, fibrosarcoma.	0	0	0	1	0	0.4041	0.4878	5 *
110	systemic tumor	6	hemangioma-b	1	2	1	1	0	0.8790	0.8726	6 *
110	systemic tumor	7	hemangiosarcoma-m	0	2	0	1	0	0.8156	0.8197	7 *
110	systemic tumor	78	schwannoma-m	2	3	1	0	2	0.5490	0.5651	8 *
112	zymbal gland	52	carcinoma-m	0	1	0	1	0	0.6283	0.6810	9 *
114	bone, other	27	osteosarcoma-m	0	0	0	0	1	0.2214	0.0395	10 *
15	harderian gland	1	adenocarcinoma-m	0	0	0	0	1	0.2214	0.0395	11 *
17	hematopoietic system	29	histiocytic sarcoma-m	0	1	2	1	1	0.4640	0.5151	12 *
17	hematopoietic system	88	lymphoma-m, lymphoblastic	0	0	0	0	1	0.2353	0.0455	13 *
17	hematopoietic system	89	lymphoma-m, lymphocytic	0	0	0	0	1	0.2588	0.0564	14 *
20	kidneys	1	adenocarcinoma-m	0	1	0	0	0	1.0000	0.7811	15
22	liver	10	adenoma-b, hepatocellular	0	0	2	0	0	0.7185	0.8014	16 *
22	liver	58	cholangioma-b	1	0	0	0	0	1.0000	0.7921	17
22	liver	9	adenocarcinoma-m, hepatocellul	0	0	0	1	0	0.3247	0.4205	18 *
26	lungs	1	adenocarcinoma-m	0	0	0	1	1	0.1247	0.0906	19 *
27	mammary gland	1	adenocarcinoma-m	13	10	13	11	11	0.7331	0.7352	20 *
27	mammary gland	2	adenoma-b	7	11	9	7	10	0.2700	0.2701	21 *
27	mammary gland	24	fibroma-b	2	0	5	1	1	0.7983	0.8016	22 *
27	mammary gland	33	fibroadenoma-b	29	34	28	26	29	0.7665	0.7643	23 *
27	mammary gland	87	carcinosarcoma-m	0	0	0	0	1	0.2010	0.0309	24 *
30	mesometrium	11	lipoma-b	1	0	0	0	0	1.0000	0.8643	25
33	ovary	15	cystadenoma-b	0	1	0	0	0	1.0000	0.7921	26
33	ovary	25	granulosa cell tumor-b	0	4	2	5	1	0.7227	0.7277	27 *
33	ovary	59	sertoli cell tumor-b	1	0	0	1	1	0.3113	0.2806	28 *
33	ovary	61	luteoma-b	0	1	1	0	0	0.8730	0.8381	29 *
33	ovary	73	granulosa cell tumor-m	0	2	0	3	0	0.7535	0.7622	30 *
33	ovary	90	granulosa thecal cell tumor-b	0	0	0	1	0	0.4275	0.5092	31 *
34	pancreas	20	adenoma-b, islet cell	2	2	2	0	0	0.9744	0.9384	32 *
36	pituitary	1	adenocarcinoma-m	1	1	0	2	0	0.7367	0.7797	33 *
36	pituitary	2	adenoma-b	36	31	30	31	38	0.1561	0.1538	34 *
40	salivary gland	1	adenocarcinoma-m	1	0	0	0	0	1.0000	0.7499	35
40	salivary gland	2	adenoma-b	0	0	0	0	1	0.2214	0.0395	36 *
44	skin	22	fibrosarcoma-m	0	1	1	0	0	0.8277	0.8126	37 *
44	skin	24	fibroma-b	0	2	2	1	0	0.8402	0.8636	38 *
44	skin	31	carcinoma-m, squamous cell	1	0	1	0	0	0.8766	0.8398	39 *

44	skin	80	neurofibrosarcoma-m	1	0	0	0	0	1.0000	0.7934	40
50	thymus	85	thymoma-b	0	0	0	0	1	0.2295	0.0426	41 *
51	thyroid	39	adenoma-b, c-cell	14	10	9	9	16	0.1010	0.0974	42 *
51	thyroid	5	adenoma-b, follicular	0	0	0	0	1	0.2214	0.0395	43 *
51	thyroid	77	carcinoma-m, c-cell	4	1	1	2	0	0.9410	0.9303	44 *
55	uterus	1	adenocarcinoma-m	2	0	0	0	2	0.2123	0.1554	45 *
55	uterus	18	sarcoma-m	0	0	0	0	1	0.2214	0.0395	46 *
55	uterus	22	fibrosarcoma-m	0	0	0	0	1	0.2214	0.0395	47 *
55	uterus	24	fibroma-b	0	1	0	0	0	1.0000	0.7499	48
55	uterus	81	endometrial stromal polyp-b	6	5	8	4	4	0.8512	0.8490	49 *
55	uterus	82	sarcoma-m, endometrial stromal	1	0	0	0	1	0.3951	0.2429	50 *
56	vagina	24	fibroma-b	1	0	2	0	0	0.8380	0.8538	51 *
56	vagina	51	papilloma-b	2	0	0	0	0	1.0000	0.8702	52
56	vagina	82	sarcoma-m, endometrial stromal	0	0	1	0	0	0.6412	0.7311	53 *
64	abdominal cavity/region	18	sarcoma-m	0	0	0	1	0	0.3881	0.4753	54 *
66	mesenterium	86	mast cell tumor-m	0	0	1	0	0	0.6412	0.7311	55 *
73	cerebellum	91	granular cell tumor-b	0	0	1	0	0	0.6412	0.7311	56 *
74	cerebrum	49	astrocytoma-m	0	0	0	1	1	0.1183	0.0831	57 *
74	cerebrum	71	ependymoma-b	0	0	0	0	1	0.2500	0.0504	58 *
74	cerebrum	83	meningioma-m	1	0	0	0	0	1.0000	0.7811	59
97	cervix	11	lipoma-b	0	0	0	1	0	0.8095	0.5171	60 *
97	cervix	57	fibroma-b, polyploid	1	0	0	0	0	1.0000	0.7872	61

Note: The symbol "*" indicates that the p-values fall in (0, 1).
The check mark indicates statistically significant test results,
based on the decision rule of **FDA.CDER.Divisions of Biometrics**.

APPEARS THIS WAY
ON ORIGINAL

Table 10: Analysis of Mortality

Species: Rat, Sex: Male, NDA 21487

Analysis of Mortality		No. Risk	No. Died	No. Alive	Pct Survival	Pct Mortality
CTR1	53-78	50	1	49	98.0	2.0
	79-91	49	2	47	94.0	6.0
	92-128	47	30	17	34.0	66.0
	FINALKILL129-129	17	17	0		
CTR2	0-52	50	1	49	98.0	2.0
	53-78	49	3	46	92.0	8.0
	79-91	46	4	42	84.0	16.0
	92-128	42	29	13	26.0	74.0
	FINALKILL129-129	13	13	0		
LOW	53-78	50	2	48	96.0	4.0
	79-91	48	8	40	80.0	20.0
	92-128	40	30	10	20.0	80.0
	FINALKILL129-129	10	10	0		
MED	0-52	50	1	49	98.0	2.0
	53-78	49	4	45	90.0	10.0
	79-91	45	5	40	80.0	20.0
	92-128	40	23	17	34.0	66.0
	FINALKILL129-129	17	17	0		
HIGH	53-78	50	3	47	94.0	6.0
	79-91	47	1	46	92.0	8.0
	92-128	46	27	19	38.0	62.0
	FINALKILL129-129	19	19	0		

Figure 4: Kaplan-Meier Graph of Survival

Species: Rat, Sex: Male, NDA 21487

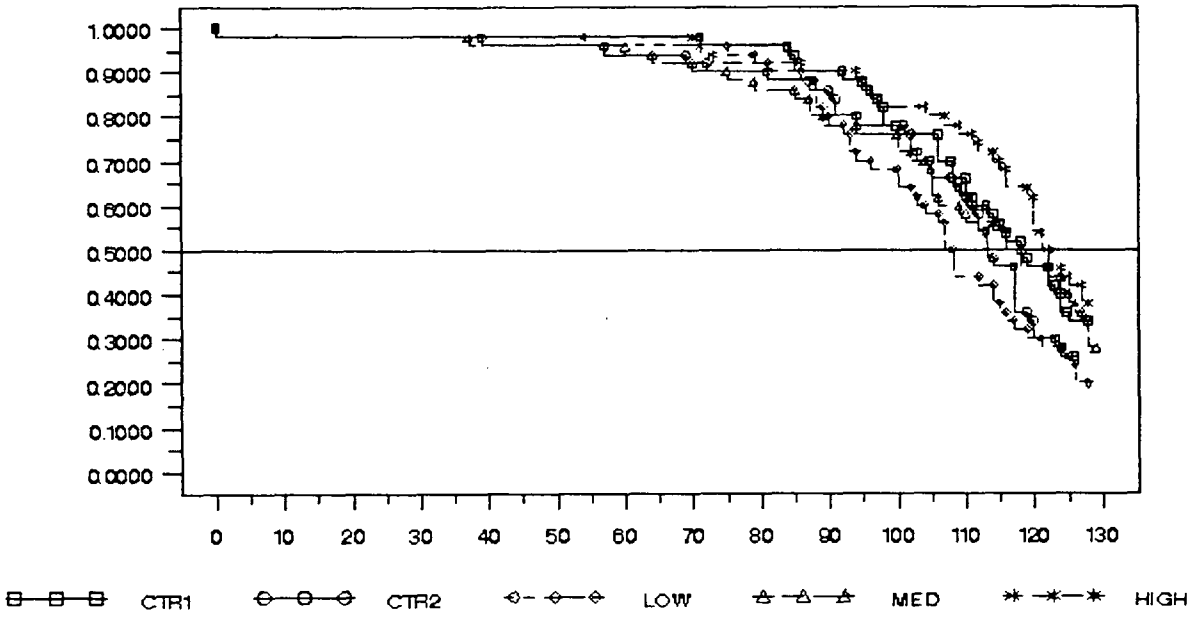


Table 11: Analysis of Dose-Mortality Trend

Species: Rat, Sex: Male, NDA 21487

This test is run using Trend and Homogeneity Analyses of Proportions and Life Table Data Version 2.1, by Donald G. Thomas, National Cancer Institute

	Method			
	Cox		Kruskal-Wallis	
	Statistics	P-Value	Statistics	P-Value
Time-Adjusted Trend Test				
Depart from Trend	4.7111	0.1942	5.0745	0.1664
Dose-Mortality Trend	3.5592	0.0592	4.3730	0.0365
Homogeneity	8.2703	0.0822	9.4475	0.0508

APPEARS THIS WAY
ON ORIGINAL

Table 12: Report on Trend Test

Detecting significant positive dose-tumor linear trend

Species: Rat, Sex: Male, NDA 21487

Organ Code	Organ Name	Tumor Code	Tumor Name	CT R1	CT R2	LOW	MED	HIGH	P-Value (Exact Method)	P-Value (Asymptotic Method)	Link 2XC Table >
1	adrenals	23	pheochromocytoma-m	1	0	2	1	1	0.4233	0.4874	1 *
1	adrenals	41	carcinoma-m, cortical	0	0	1	0	0	0.5756	0.7247	2 *
1	adrenals	45	adenoma-b, cortical	0	2	2	3	0	0.8533	0.8437	3 *
1	adrenals	72	pheochromocytoma-b	4	4	2	0	3	0.6914	0.7040	4 *
11	eyes	18	sarcoma-m	0	0	1	0	1	0.2486	0.2480	5 *
11	eyes	24	fibroma-b	0	0	1	0	0	0.5755	0.7020	6 *
110	systemic tumor	56	mesothelioma-m, fibrosarcoma.	0	1	0	0	0	1.0000	0.7961	7
110	systemic tumor	6	hemangioma-b	5	3	2	4	2	0.7654	0.7701	8 *
110	systemic tumor	7	hemangiosarcoma-m	2	0	1	1	1	0.5572	0.5757	9 *
110	systemic tumor	78	schwannoma-m	0	0	2	0	0	0.6964	0.7958	10 *
112	zymbal gland	52	carcinoma-m	1	0	0	0	0	1.0000	0.7684	11
114	bone, other	27	osteosarcoma-m	0	0	1	0	0	0.5888	0.7203	12 *
114	bone, other	46	osteoma-b	1	1	0	1	1	0.4178	0.4080	13 *
12	bone (femur)	46	osteoma-b	1	0	0	0	0	1.0000	0.7930	14
15	harderian gland	2	adenoma-b	0	0	1	0	0	0.5797	0.7034	15 *
17	hematopoietic system	29	histiocytic sarcoma-m	1	1	0	1	1	0.5277	0.5062	16 *
17	hematopoietic system	88	lymphoma-m, lymphoblastic	1	3	0	0	3	0.2252	0.2096	17 *
17	hematopoietic system	89	lymphoma-m, lymphocytic	0	0	0	2	0	0.3687	0.3990	18 *
20	kidneys	1	adenocarcinoma-m	0	1	0	0	0	1.0000	0.7684	19
22	liver	10	adenoma-b, hepatocellular	0	3	2	1	0	0.9212	0.9187	20 *
22	liver	9	adenocarcinoma-m, hepatocellul	0	0	2	0	0	0.7081	0.8029	21 *
26	lungs	3	adenoma-b, bronchial-alveolar	1	0	0	0	0	1.0000	0.8026	22
27	mammary gland	1	adenocarcinoma-m	1	1	0	2	2	0.2065	0.2010	23 *
27	mammary gland	18	sarcoma-m	0	0	1	0	0	0.5946	0.7136	24 *
27	mammary gland	2	adenoma-b	0	0	0	1	0	0.4667	0.5276	25 *
27	mammary gland	24	fibroma-b	2	2	3	1	0	0.9591	0.9475	26 *
27	mammary gland	31	carcinoma-m, squamous cell	0	0	1	0	0	0.6048	0.7485	27 *
27	mammary gland	33	fibroadenoma-b	1	1	1	0	1	0.5246	0.4923	28 *
32	nasal cavity	1	adenocarcinoma-m	0	0	1	0	0	0.6043	0.7146	29 *
32	nasal cavity	31	carcinoma-m, squamous cell	0	1	0	0	0	1.0000	0.7998	30
34	pancreas	11	lipoma-b	0	1	0	0	0	1.0000	0.7684	31
34	pancreas	20	adenoma-b, islet cell	8	7	4	8	2	0.9655	0.9588	32 *
34	pancreas	44	adenoma-b, acinar cell	2	1	0	1	0	0.8520	0.8454	33 *
34	pancreas	50	adenocarcinoma-m, islet cell	1	5	0	2	0	0.9613	0.9494	34 *
35	parathyroid	2	adenoma-b	2	2	0	0	1	0.7358	0.7181	35 *

36	pituitary	1	adenocarcinoma-m	2	1	1	1	2	0.4547	0.4558	36 *
36	pituitary	2	adenoma-b	28	29	24	30	33	0.1719	0.1696	37 *
38	prostate	1	adenocarcinoma-m	0	1	0	0	0	1.0000	0.7787	38
40	salivary gland	18	sarcoma-m	0	0	1	0	0	0.6013	0.7369	39 *
42	seminal vesicle	31	carcinoma-m, squamous cell	0	0	0	1	0	0.3551	0.4617	40 *
44	skin	11	lipoma-b	0	0	0	0	1	0.2500	0.0526	41 *
44	skin	18	sarcoma-m	1	0	1	0	0	0.8358	0.8318	42 *
44	skin	22	fibrosarcoma-m	0	0	0	1	1	0.1449	0.1005	43 *
44	skin	24	fibroma-b	2	4	2	3	1	0.8659	0.8631	44 *
44	skin	31	carcinoma-m, squamous cell	1	2	0	0	1	0.5580	0.5345	45 *
44	skin	36	adenoma-b, sebaceous	0	0	1	0	0	0.5839	0.7048	46 *
44	skin	43	myxoma-b	1	0	0	0	0	1.0000	0.7707	47
44	skin	47	keratoacanthoma-b	3	0	0	0	0	1.0000	0.9105	48
44	skin	51	papilloma-b	0	1	0	1	0	0.6658	0.7071	49 *
44	skin	79	basal cell tumor-b	0	0	0	1	0	0.3650	0.4719	50 *
49	testes	17	leydig cell tumor-b	1	2	0	1	1	0.6111	0.5987	51 *
50	thymus	76	thymoma-m	0	1	0	0	0	1.0000	0.7837	52
51	thyroid	38	adenocarcinoma-m, follicular	1	2	0	1	0	0.8899	0.8740	53 *
51	thyroid	39	adenoma-b, c-cell	11	11	9	5	12	0.4221	0.4246	54 *
51	thyroid	5	adenoma-b, follicular	3	2	1	0	0	0.9969	0.9501	55 *
51	thyroid	77	carcinoma-m, c-cell	0	1	3	4	2	0.3610	0.3662	56 *
61	retroperitoneum	18	sarcoma-m	0	0	0	1	0	0.4286	0.5189	57 *
73	cerebellum	49	astrocytoma-m	0	1	0	0	0	1.0000	0.7961	58
74	cerebrum	49	astrocytoma-m	1	0	0	1	1	0.3027	0.2772	59 *
74	cerebrum	74	glioblastoma-m	0	0	0	1	0	0.3915	0.4814	60 *

Note: The symbol "*" indicates that the p-values fall in (0, 1).
The check mark indicates statistically significant test results,
based on the decision rule of **FDA.CDER.Divisions of Biometrics**.

Rajeshwari Sridhara, Ph.D.
Mathematical Statistician
Date:

Concur: Roswitha Kelly, M.S.
Team Leader

George Chi, Ph.D.
Division Director, DBI

Cc:
HFD-120/ Dr. Haberry
HFD-120/ Dr. Rosloff
HFD-710/ Dr. Jin
HFD-710/ Dr. Sridhara
HFD-710/ Ms. Kelly
HFD-710/ Dr. Chi
HFD-700/ Dr. Dubey
HFD-700/ Dr. Anello

This review consists of 25 pages of text
C:\NDA\neropharm\21487\21487_carcin.doc

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

/s/

Rajeshwari Sridhara
8/18/03 01:27:41 PM
BIOMETRICS

Roswitha Kelly
8/21/03 10:36:09 AM
BIOMETRICS

George Chi
8/25/03 12:38:53 PM
BIOMETRICS

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

/s/

Janusz Rzeszotarski
10/15/03 01:30:41 PM
CHEMIST

Maryla Guzewska
10/15/03 01:31:46 PM
CHEMIST