

**CENTER FOR DRUG EVALUATION AND  
RESEARCH**

*APPLICATION NUMBER:*  
**21-782**

**STATISTICAL REVIEW(S)**



U.S. Department of Health and Human Services  
Food and Drug Administration  
Center for Drug Evaluation and Research  
Office of Pharmacoeconomics and Statistical Science  
Office of Biostatistics

## STATISTICAL REVIEW AND EVALUATION

### CLINICAL STUDIES

**NDA/Serial Number:** 21-782  
**Drug Name:** Ramelteon  
**Indication(s):** Treatment of insomnia  
**Applicant:** Takeda Global Research and Development Center, Incorporated  
**Date(s):** Received 09/21/04; user fee (10 months) 07/22/05  
**Review Priority:** Standard  
**Biometrics Division:** Division of Biometrics II  
**Statistical Reviewer:** Dionne L. Price, Ph.D.  
**Concurring Reviewers:** Thomas J. Permutt, Ph.D.  
S. Edward Nevius, Ph.D.  
  
**Medical Division:** Division of Anesthesia, Analgesia, and Rheumatology Products  
**Clinical Team:** D. Elizabeth McNeil, M.D.  
Rigoberto Roca, M.D.  
**Project Manager:** Sara Stradley

**Keywords:** NDA review, clinical studies

# Table of Contents

<b>1. EXECUTIVE SUMMARY .....</b>	<b>3</b>
1.1 CONCLUSIONS AND RECOMMENDATIONS .....	3
1.2 BRIEF OVERVIEW OF CLINICAL STUDIES .....	3
1.3 STATISTICAL ISSUES AND FINDINGS .....	4
<b>2. INTRODUCTION .....</b>	<b>6</b>
2.1 OVERVIEW.....	6
2.2 DATA SOURCES .....	6
<b>3. STATISTICAL EVALUATION .....</b>	<b>8</b>
3.1 EVALUATION OF EFFICACY .....	8
3.1.1 Study TL021 .....	8
3.1.2 Study TL017 .....	15
3.1.3 Study TL025 .....	17
3.1.4 Study TL020 .....	22
3.1.5 Study TL023 .....	24
3.2 EVALUATION OF SAFETY .....	26
<b>4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS .....</b>	<b>27</b>
4.1 GENDER, RACE AND AGE .....	27
4.1.1 Study TL021 .....	27
4.1.2 Study TL017 .....	28
4.1.3 Study TL025 .....	29
4.1.3 Study TL023 .....	30
<b>5. SUMMARY AND CONCLUSIONS .....</b>	<b>32</b>
5.1 STATISTICAL ISSUES AND COLLECTIVE EVIDENCE .....	32
5.2 CONCLUSIONS AND RECOMMENDATIONS .....	33
5.2.1 Labeling .....	34
<b>APPENDICES.....</b>	<b>36</b>

# **1. EXECUTIVE SUMMARY**

## **1.1 Conclusions and Recommendations**

Takeda Global Research and Development Center, Incorporated proposes ramelteon 8 mg for the treatment of insomnia. The applicant claims that the drug reduces the latency to persistent sleep and that the treatment effect is maintained through five weeks. Furthermore, the applicant claims that ramelteon increases the overall sleep time. My review of the statistical evidence suggests that ramelteon promotes sleep onset (as measured by sleep latency) during the initial week of treatment. The applicant has also provided some evidence that the treatment effect is maintained over a prolonged period of time. However based on my review, I disagree with the applicant's assessment of the total sleep time and the assessment of supportive analyses. The total sleep time did not consistently demonstrate efficacy across studies and was not a clinically appropriate endpoint to ascertain the effectiveness of the drug. Additionally, the applicant did not garner support for the overall conclusions via pre-specified, supplemental analyses in the study conducted in a real world setting. Moreover various safety concerns, such as the genotoxic and carcinogenic potential of the product and the effect of the product on the endocrine system, have arisen during the course of the clinical and pharmacology/toxicology reviews; therefore, the team will need to collectively evaluate the risks and benefits of ramelteon.

## **1.2 Brief Overview of Clinical Studies**

Currently, most agents approved for the treatment of insomnia are benzodiazepine receptor agonists. In contrast, ramelteon is selective for the melatonin-1 and melatonin-2 receptors and therefore has a novel mechanism of action relative to current available treatments. Ramelteon was introduced to the Division of Neuropharmacological Drug Products via IND 58,136. The IND was subsequently transferred to the Division of Anesthesia, Analgesia, and Rheumatology Products. The clinical development plan was discussed during several meetings between the applicant and the divisions. Issues addressed included the appropriateness of the analysis endpoints and the need for efficacy to be demonstrated in both outpatient and inpatient settings.

Four double-blind, placebo-controlled, multi-center studies were submitted to support the effectiveness of ramelteon. In Study 21, patients with chronic insomnia were randomized to placebo, ramelteon 8 mg, or ramelteon 16 mg for 35 days. The latency to persistent sleep was assessed via polysomnography in a sleep laboratory during the first two nights of weeks 1, 3, and 5, respectively. An analysis of covariance model was used to assess treatment group differences. A categorical or responder analysis was conducted to support the primary analysis. Study 25 was an outpatient study conducted in elderly patients with chronic insomnia. Patients were randomized to placebo, ramelteon 4 mg, or ramelteon 8 mg for 35 days. The primary measure of

efficacy in Study 25 was the mean subjective sleep latency over the initial seven nights of treatment. The analysis of the endpoint mimicked that of Study 21.

Study 17 was a crossover study in elderly patients with chronic insomnia. Study participants were randomized into one of six treatment sequences such that each patient received placebo, ramelteon 4 mg, and ramelteon 8 mg during the course of the study. The mean latency to persistent sleep, calculated from polysomnographic recordings in the sleep laboratory, was the primary measure of efficacy. The applicant employed a mixed linear model for the primary analysis.

A single study was conducted to assess the effects of ramelteon on transient insomnia. Study 23 was conducted in healthy adults that were naïve to a sleep laboratory environment. Participants were randomized to placebo, ramelteon 8 mg, or ramelteon 16 mg. The objective polysomnographic measurement of latency to persistent sleep was the primary measure of efficacy. The statistical methodologies employed in the analysis mimicked those used in the previous parallel group studies.

An additional study, Study 20, was conducted by the applicant; however the applicant did not view the study as being reflective of the efficacy of the drug. I disagreed with the applicant's assessment and included the study in my review. With the exception of the study populations, the study design and analysis plan for Study 20 mimicked that of Study 25. The former study was conducted in an adult population while the latter was conducted in an elderly population.

### **1.3 Statistical Issues and Findings**

During interactions between the agency and Takeda prior to the submission of the NDA, various statistical concerns were expressed. The agency was concerned that the testing of the primary measure of efficacy at multiple time points would inflate the type I error (i.e. falsely concluding that groups differed when in reality, they did not) and that the last observation carried forward (LOCF) imputation strategy was inappropriate for a potentially fast-acting drug. In addition to the statistical concerns, the agency expressed interest in two important facets of the drug, namely, the ability to promote sleep onset and the ability to maintain the treatment effect over time. The agency also commented that the drug would need to demonstrate efficacy in a real world setting.

I initially considered the ability of ramelteon to promote sleep onset. The latency to persistent sleep was significantly reduced for the initial week of treatment in three of the four studies conducted in patients with chronic insomnia. The applicant acknowledged the lack of positive findings in Study 20. In Study 25, the analysis of means using a LOCF imputation strategy resulted in a treatment effect at week 1. Based on the agency's previously expressed concern with the LOCF procedure, I reanalyzed the data using a baseline observation carried forward (BOCF) scheme. The results of both imputation methods were similar and yielded a treatment effect at week 1. The applicant additionally conducted a supportive responder analysis on the population derived via a LOCF strategy. As a result of the applicant's methodology, a patient

withdrawing for an unfavorable treatment-related reason could have been considered a responder. I disagreed with the methodology and thus reanalyzed the data considering a population whereby all withdrawals were treated as non-responders. In general, a responder analysis is less powerful to detect a difference among treatments; however, I anticipated that the analysis would lend some support to the claim that ramelteon reduced the latency to persistent sleep. In contrast, the treatment effect completely disappeared when utilizing a responder analysis. To further elucidate the findings, I examined the cumulative distribution functions (CDFs) of each treatment arm at week 1. The CDFs provided insight into the probability that the latency to persistent sleep was achieved within a certain amount of time. The reduction in mean latency for the active groups compared to the placebo group was mainly attributable to a reduction of large values to slightly less large values. There was little difference in the proportion of values less than 30 min.

Subsequent to my review of sleep onset, I evaluated the  $H_0$  of the treatment effect. Three of the studies were conducted for 35 days. During meetings prior to the submission of the NDA, the agency suggested that the applicant sequentially test the treatment effects at the multiple time points to offset the multiplicity concern. The applicant used the recommended methodology thereby alleviating my concern regarding the inflation of the type I error. The applicant claimed that two of the studies demonstrated evidence of the  $H_1$  of the effect. Although the effect was demonstrated via the analyses of means in both studies, the responder analysis in the outpatient study was not supportive of the  $H_1$  claim.

For completion, the applicant conducted a study in a healthy population to evaluate the effectiveness of ramelteon in individuals with transient insomnia. The applicant concluded that 8 mg of ramelteon was effective in reducing the latency to persistent sleep; however, the conclusion was not supported by the additional categorical analysis.

Two “key” secondary variables were identified by the applicant and assessed in the studies. The agency previously recommended that the applicant consider other secondary variables to support the desired claims. Specifically the agency stated, “We noted that other proposed secondary outcomes, i.e., TST and sleep efficiency, were not ideal and would likely not be acceptable for supporting a  $H_1$  claim.” Upon review of the total sleep time and sleep efficiency, I found that the measures did not provide significant support across studies. Additional secondary variables were assessed; however, the variables also failed to provide consistent support for the proposed claims. Moreover, the study protocols did not provide an explanation of the relative importance of the variables or an explanation of the role of the variables in the interpretation of the results.

## 2. INTRODUCTION

### 2.1 Overview

Takeda Global Research and Development Center, Incorporated proposes ramelteon for the treatment of insomnia. Ramelteon is selective for the melatonin-1 and melatonin-2 receptors and therefore has a novel mechanism of action relative to current available treatments. The drug was initially introduced to the Division of Neuropharmacological Drug Products via IND 58,136. During the development process, Takeda submitted several study protocols for division comment. Moreover, the product was discussed during an End of Phase 1 meeting on 8 November 2001 and a teleconference on 20 November 2002. The clinical development plan, study populations, and analysis endpoints were discussed during an end of phase 2 (EOP2) meeting on 16 July 2002. Subsequent to the EOP2 meeting, Takeda submitted study protocols amended based on recommendations from the meeting. The statistical reviewer, Dr. Yeh-Fong Chen, expressed concern regarding the appropriateness of the proposed methodologies for multiplicity and missing data. On 17 September 2003, the IND was transferred to the Division of Anesthesia, Analgesia, and Rheumatology Products (DAARP). Interactions between Takeda and DAARP included a Chemistry, Manufacturing, and Controls meeting (15 December 2003), a teleconference (11 February 2004), and a pre-NDA meeting (22 June 2004). A germane issue discussed during the teleconference was the need for efficacy to be demonstrated in both inpatient and outpatient settings. Currently, the applicant has submitted NDA 21-782. The submission investigates the safety and efficacy of ramelteon for the treatment of insomnia.

### 2.2 Data Sources

Seven double-blind, placebo-controlled studies were conducted to establish the efficacy and safety of ramelteon. The study reports and data for the completely electronic submission were archived in the Food and Drug Administration internal document room under the network path location \\CDSESUBI\N21782\N-000\2004-09-12 . A summary of the studies is provided in Table 1.

**Table 1: Summary of Studies**

(Source: Adapted from the Summary of Clinical Efficacy, Appendix 6.1, Table 6.a)

Study Number Number of centers (n)	Study Design	Treatment Arms and Number of randomized patients (n)	Primary measure of efficacy
PNFP002 Multi-center (14)	Phase 2, randomized, double-blind, placebo-controlled, single-dose study in adults unfamiliar with a sleep laboratory environment	•Ramelteon 16 mg (126) •Ramelteon 64 mg (126) •Placebo (123)	Mean latency to persistent sleep (LPS) as determined by polysomnography (PSG) recordings
TL023 Multi-center (15)	Phase 3, randomized, double-blind, placebo-controlled, single-dose study in adults unfamiliar with a sleep laboratory environment	•Ramelteon 8 mg (98) •Ramelteon 16 mg (94) •Placebo (97)	LPS from Night 1 of PSG recording.
TL005 Multi-center (13)	Phase 2, randomized, double-blind, placebo-controlled, crossover study in adults (conducted in a sleep laboratory)	•Ramelteon 4 mg (107) •Ramelteon 8 mg •Ramelteon 16 mg •Ramelteon 32 mg •Placebo	Mean LPS as determined by PSG recordings
TL017 Multi-center (17)	Phase 3, randomized, double-blind, placebo-controlled, crossover study in elderly adults with chronic insomnia	•Ramelteon 4 mg (100) •Ramelteon 8 mg •Placebo	Mean LPS as determined by PSG recordings from Nights 1 and 2 of each period
TL021 Multi-center (29)	Phase 3, randomized, double-blind, placebo-controlled study in adults with chronic insomnia	•Ramelteon 8 mg (139) •Ramelteon 16 mg (135) •Placebo (131)	Mean LPS as determined by PSG recordings from Nights 1 and 2 of the double-blind treatment phase
TL020 Multi-center (79)	Phase 3, randomized, double-blind, placebo-controlled, outpatient study in adults with chronic insomnia	•Ramelteon 8 mg (277) •Ramelteon 16 mg (284) •Placebo (287)	Mean subjective sleep latency from Week 1 of double-blind treatment.
TL025 Multi-center (136)	Phase 3, randomized, double-blind, placebo-controlled, outpatient study in elderly adults with chronic insomnia	•Ramelteon 4 mg (281) •Ramelteon 8 mg (274) •Placebo (274)	Mean subjective sleep latency from Week 1 of double-blind treatment.

According to the sponsor,

The major trials in support of efficacy are TL023, TL017, TL021, and TL025. PNFP002 and TL005 were dose-ranging studies. An additional well-controlled study, TL020 was conducted. Although data from this study are presented for completeness, this study is not considered reflective of the efficacy of ramelteon.

I concur that studies PNFP002 and TL005 are dose-ranging studies and are therefore not of focus in my review. However, I disagree with the applicant regarding study TL020. The study was a phase 3, randomized, double-blind, and placebo-controlled study. With the exception of the study populations, the study design and analysis plan for TL020 mimicked that of TL025. The former study was conducted in an adult population while the latter was conducted in an elderly population. In my opinion, the study does provide relevant information.

### **3. STATISTICAL EVALUATION**

#### **3.1 Evaluation of Efficacy**

The main body of my evaluation of efficacy will discuss each study individually.

##### **3.1.1 Study TL021**

###### *Study Design and Endpoints*

Study 21 was a randomized, double-blind, placebo-controlled, multi-center study in adults with chronic insomnia. Eligible patients initially entered a 7-day, single-blind, placebo lead-in period consisting of two consecutive nights of polysomnographic (PSG) screening followed by a 5-night outpatient period. According to the sponsor, "PSG is regarded as the gold standard for objective sleep measurements for a sleep-promoting agent. A central reader was used to provide consistency in interpretation." All PSG assessments were performed in sleep laboratories.

Four hundred and five individuals met all eligibility criteria throughout the screening period and were subsequently randomized to placebo, 8 mg of ramelteon, or 16 mg of ramelteon. During the double-blind treatment period, patients reported to the sleep laboratory for PSG assessments on nights 1 and 2, nights 15 and 16, and nights 29 and 30. Patients maintained sleep diaries throughout the 35-night duration of treatment. At the conclusion of treatment, patients received placebo for two consecutive days and reported to the sleep laboratory nightly for additional assessments to evaluate rebound insomnia and withdrawal effects.

The primary measure of efficacy was the mean latency to persistent sleep (LPS) computed via PSG assessments from nights 1 and 2. "LPS was defined as the elapsed time from the beginning of the PSG recording to the onset of the first 10 minutes of continuous sleep (i.e., total number of epochs before the first 20 consecutive nonwake epochs, divided by 2)." Mean LPS was additionally calculated at week 3 and week 5. Secondary measures of efficacy included, but were not limited to, total sleep time (TST), wake time after sleep onset (WASO), number of awakenings (NAW), sleep efficiency, subjective ease of falling back to sleep, and subjective sleep quality. The former three secondary measures were measured both objectively and

subjectively. The definitions of the secondary measures are included in a glossary in the appendix.

Rebound insomnia and withdrawal effects were additionally evaluated by the applicant. The applicant stated, "Rebound insomnia was assessed using the change from baseline in LPS on each day of the single-blind placebo run-out period." Additionally, the Benzodiazepine Withdrawal Symptom Questionnaire (BWSQ) was employed to assess withdrawal effects. The BWSQ was comprised of questions related to 20 symptoms, each evaluated on a 3-point scale. Specifically, the change in total BWSQ score from the end of double-blind treatment (i.e. Day 29 and Day 30) to the single-blind placebo run-out period was of interest.

A sample size of 390 was formulated to detect a difference in means of 12 minutes between active and placebo treatments with 90% power. The sample size was calculated using a paired t-test (with Bonferroni adjustment) and assuming a standard deviation of 25 minutes and a dropout rate of 0.15. Study participants were enrolled at 29 centers across the United States.

#### *Patient Disposition, Demographic and Baseline Characteristics*

Descriptive demographics and baseline characteristics were summarized for all randomized patients. The ages of patients were between 18 and 64 with a mean age of 39. In the study, 61% of study participants were Caucasian, 20% were Hispanic, and 16% were African-American. Females comprised 67% of the patient population. The average height of study participants was 168 centimeters and the average weight was 73 kilograms. Significant differences with respect to gender, height, and weight were detected. I attributed the imbalance to the 5% risk of committing a type I error (falsely concluding that groups differ when in reality, they do not). Since the imbalance did not exist across studies, I did not investigate further. A detailed table outlining the composition of the study population with respect to demographic and baseline characteristics is presented in the appendix.

Of the 405 randomized participants, 139 were randomized to ramelteon 8 mg, 135 were in the ramelteon 16 mg group, and 131 were randomized to placebo. Thirty-eight patients withdrew during the course of the study. Four of the thirty-eight withdrew during the single-blind placebo run-out period. The highest percentage of discontinuations (19/38 or 50%) was due to withdrawal of consent. Seven patients discontinued due to an adverse event, and of these seven, four were randomized to the ramelteon 8 mg group. Similarly, seven patients discontinued due to protocol deviations, and of these, four were randomized to the ramelteon 8 mg treatment arm. Although the overall percentage of patient withdrawal was low, the timing of the withdrawals was of some interest. The number of withdrawals per week is depicted in Table 2.

**Table 2: Number of patient withdrawals per week**  
(Source: Response to Information Request, 4 April 2005)

	Placebo	Ramelteon 8 mg	Ramelteon 16 mg
Week 1	3	8	1
Week 2	1	1	0
Week 3	3	3	1
Week 4	2	1	1
Week 5 and beyond	4	5	4

The applicant stated,

The SAP for this study was amended to reflect problems with data collection that were discovered during the study. The protocol specified that the data for study weeks would be classified into visits using specified "windows" of the study days: "Weekly time windows will be defined (i.e. Nights 1-7, 8-14, 15-21, 22-28, 29- last dose of double-blind study medication). The average of the nonmissing data for a weekly time window will be analyzed." Because the dates recorded on the diary CRFs were deemed to be potentially inaccurate, the data recorded on the CRFs were applied to the visit label on the CRF. No recorded dates were checked. With diaries being returned to the clinics on Days 15, 29, and 36, the appropriate labels for the diary data during treatment are: "Weeks 1-2," "Weeks 3-4", and "Week 5". The final SAP for the study, as completed prior to unblinding, included these changes.

As a result, I requested the applicant provide a table identifying and enumerating the patients that withdrew during each week. I specifically requested that the day of withdrawal be classified as the day following the last dose of study medication. Table 2 was provided in response to my request.

### *Statistical Methodologies*

The primary analysis employed an analysis of covariance (ANCOVA) model with treatment and pooled center as main effects and baseline LPS as a covariate. An examination of the consistency of the results across centers was conducted via inclusion of a treatment-by-center interaction in the ANCOVA model. The type I error was controlled via Fisher's protected least significant difference (LSD). Using this procedure, an overall test of the treatment effect was conducted at level  $\alpha$ . If an overall treatment effect was present, pairwise comparisons were tested (each at level  $\alpha$ ). Additionally, an evaluation of the  $\tau_1$  efficacy was conducted via a stepwise application of Fisher's protected LSD at weeks 3 and 5. Analyses were performed on the intent-to-treat (ITT) population including all randomized patients receiving at least one dose of the study medication. A last observation carried forward (LOCF) strategy was employed to handle missing data.

The applicant additionally performed a categorical or responder analysis to support the results of the primary analysis. A "responder" was defined as a participant having latency to persistent sleep less than or equal to 30 minutes. The responder status of a discontinued patient was determined via a LOCF strategy. For example, if a patient's last evaluation prior to withdrawal yielded a latency time of 20 minutes, the participant was considered a responder to treatment. A Cochran-Mantel-Haenszel (CMH) test, stratified by pooled center, was used in the analysis of the categorical data.

Analyses of the total sleep time (TST) and subjective sleep quality mimicked the analysis of the primary efficacy measure. According to the applicant, "Interpretation of significant results for TST at Weeks 1, 3, and 5 was contingent on observing significance from the overall F-test in the primary efficacy analysis, i.e., analysis of LPS at Week 1. Likewise, interpretation of significant results for subjective sleep quality at Weeks 1, 3, and 5 was contingent on observing significance from the overall F-test of TST at Week 1."

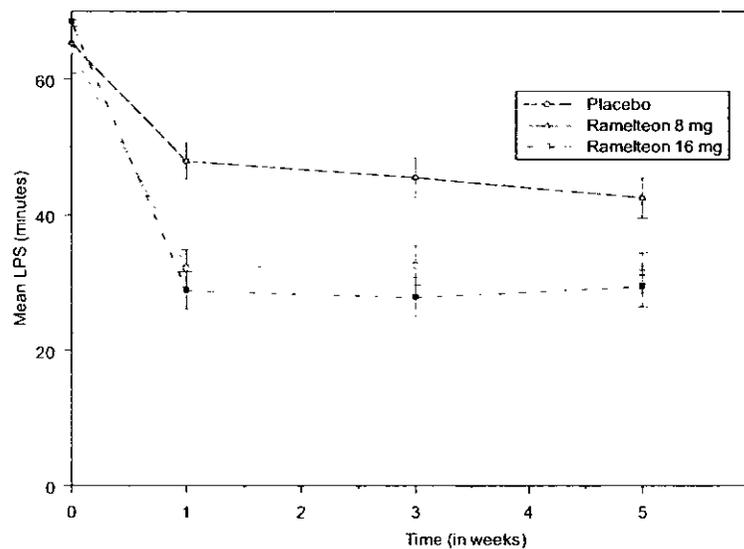
Rebound insomnia was assessed during the 2-day single-blind run-out period using observed data only. Rebound insomnia was analyzed for each day via an ANCOVA model with treatment and pooled center as effects and baseline LPS as a covariate. Summary statistics were additionally calculated. Withdrawal effects were assessed using the total BWSQ score and analyzed using an ANCOVA model with treatment and pooled center as effects and the week 5 total BWSQ score as a covariate.

### Results and Conclusions

The applicant's primary results are graphically depicted in Figure 1 and numerically depicted in Table 3. Based on the results, the applicant concluded that the average latency to persistent sleep over the initial two nights was shorter for patients receiving ramelteon (8 mg and 16 mg) as compared to patients receiving placebo. The applicant further concluded that the treatment effect was maintained throughout the duration of the study. However, a statistically significant difference in means does not necessarily imply clinical meaningfulness. The clinical meaningfulness will be determined by the clinical reviewer, Dr. D. Elizabeth McNeil.

To investigate the robustness of the results to the procedure for handling missing data, I repeated the analysis of covariance using a baseline observation carried forward (BOCF) imputation strategy. The results produced via the BOCF imputation scheme, depicted in the appendix, were similar to those produced by the analysis of covariance employing a LOCF strategy with an exception at week 5. In my reanalysis, the difference in mean LPS between the ramelteon 16 mg group and the placebo group was only borderline significant.

Figure 1: Mean LPS Across Weeks



**Table 3: Summary of LPS (minutes) – LOCF Data: ITT Population**

(Source: Adapted from Final Study Report 01-02-TL-375-021, Table 11.a)

	Placebo (n=131)	Ramelteon 8 mg (n=139)	Ramelteon 16 mg (n=135)	Overall p-value
<b>Baseline</b>				
LS Mean (SE)	65.3 (3.54)	64.3 (3.46)	68.4 (3.54)	
<b>Week 1</b>				
LS Mean (SE)	47.9 (2.72)	32.2 (2.67)	28.9 (2.71)	
LS mean difference from placebo (SE)		-15.7 (3.70)	-18.9 (3.73)	<0.001
95% CI of difference		(-22.9, -8.4)	(-26.3, -11.6)	
Pairwise p-value		<0.001	<0.001	
<b>Week 3</b>				
LS Mean (SE)	45.5 (2.93)	32.6 (2.87)	27.9 (2.92)	
LS mean difference from placebo (SE)		-12.9 (3.98)	-17.6 (4.02)	<0.001
95% CI of difference		(-20.7, -5.1)	(-25.5, -9.7)	
Pairwise p-value		0.001	<0.001	
<b>Week 5</b>				
LS Mean (SE)	42.5 (2.97)	31.5 (2.91)	29.5 (2.96)	
LS mean difference from placebo (SE)		-11.0 (4.03)	-12.9 (4.07)	0.003
95% CI of difference		(-18.9, -3.1)	(-20.9, -4.9)	
Pairwise p-value		0.007	0.002	

Note: L.S. means for baseline are from an ANOVA model with effects for treatment and pooled center. L.S. means for a post-baseline visit are from an ANCOVA model with effects for treatment and pooled center and the baseline value of the variable as a covariate. P-values for pairwise comparisons are obtained using t-tests from the ANCOVA model of the overall treatment comparisons. When the treatment-by-center interaction term was included in the model at Week 1, the p-value for the interaction was 0.533.

According to the applicant, additional support was garnered from results of the responder analysis shown in Table 4. I had some concern regarding the definition of a responder. The applicant defined a responder as a patient having a mean LPS of 30 or less, using a LOCF strategy to handle patient withdrawals. I believe an attractive feature of a responder analysis is the ability of the analysis to handle missing data in a simplistic manner whereby all withdrawals are considered non-responders. This eliminates the concern that potentially favorable outcomes are assigned to patients withdrawing for treatment-related reasons. By conducting the analysis on a population derived via a LOCF strategy, the applicant lost a portion of the utility of the responder analysis. I therefore reanalyzed the data altering the definition of a non-responder to include all drop-outs. My analysis is shown in Table 5.

**Table 4: Latency to Persistent Sleep: Categorical Analysis – LOCF Data**  
 (Source: Final Study Report 01-02-TL-375-021, Table 14.2.1.7)

	Placebo (N=131)	Ramelteon 8 mg* (N=138)	Ramelteon 16 mg (N=135)	Overall p-value**
<b>Baseline</b>				
<= 30	15	20	16	
> 30	116	119	119	
<b>Week 1</b>				<0.001
<= 30	55	87	87	
> 30	76	51	48	
p-value for comparison with placebo		<0.001	<0.001	
<b>Week 3</b>				0.021
<= 30	68	92	86	
> 30	63	46	49	
p-value for comparison with placebo		0.006	0.044	
<b>Week 5</b>				0.028
<= 30	69	90	90	
> 30	62	48	45	
p-value for comparison with placebo		0.017	0.018	

\*Note: N=139 for the baseline measurement in the ramelteon 8 mg group. Patient 211209 dropped out on the first day and was not included in subsequent calculations.

\*\* Overall and pairwise p-values are obtained from the CMH general association test, stratified by pooled center.

Appears This Way  
On Original

**Table 5: Latency to Persistent Sleep: Re-analysis of Categorical Data**

	Placebo (N=131)	Ramelteon 8 mg (N=139)	Ramelteon 16 mg (N=135)	Overall p-value*
<b>Baseline</b>				
<= 30	15	20	16	
> 30	116	119	119	
<b>Week 1</b>				<0.001
<= 30	55	88	87	
> 30	76	51	48	
p-value for comparison with placebo		<0.001	<0.001	
<b>Week 3</b>				0.065
<= 30	67	87	85	
> 30	64	52	50	
p-value for comparison with placebo		0.039	0.043	
<b>Week 5</b>				0.068
<= 30	66	82	87	
> 30	65	57	48	
p-value for comparison with placebo		0.106	0.018	

\* Overall and pairwise p-values are obtained from the CMH general association test, stratified by pooled center.

The responder or categorical analysis collapsed the data into clinically meaningful categories and sought to provide evidence of an association between the percentage of participants achieving persistent sleep within 30 minutes and the treatment. In general, a responder analysis is advantageous in that it allows an ease of interpretation; however, the analysis is less powerful to detect a difference among treatments. In study 21, I concluded that the results from my reanalysis did lend support to the applicant's claim of the onset [ ] of the treatment effect.

The applicant additionally evaluated the total sleep time and the sleep quality. The total sleep time during week 1 was significantly increased for participants randomized to ramelteon. Participants receiving ramelteon 8 mg slept an average of 19 minutes longer than participants randomized to placebo. Similarly, individuals receiving ramelteon 16 mg slept an average of 22 minutes longer than individuals receiving placebo. The effect was not maintained throughout the study. Moreover, patients in the study did not report significant differences in the quality of sleep.

The applicant's results from the analyses of rebound insomnia and withdrawal effects are in the appendix. Based on the results, the applicant concluded that there was no evidence of rebound insomnia or withdrawal effects.

### 3.1.2 Study TL017

#### *Study Design and Endpoints*

Study 17 was a randomized, double-blind, placebo-controlled, three-period, multi-center crossover study in elderly patients with chronic insomnia. One hundred individuals met eligibility criteria and were randomized into one of six treatment sequences illustrated in Table 6.

**Table 6: Treatment Sequences**  
(Source: Adapted from Final Study Report 01-02-TL-375-017, Table 10.a)

Treatment Sequence	N	Treatment Period 1	Treatment Period 2	Treatment Period 3
I	19	Placebo	8 mg	4 mg
II	16	4 mg	Placebo	8 mg
III	15	8 mg	4 mg	Placebo
IV	22	4 mg	8 mg	Placebo
V	15	8 mg	Placebo	4 mg
VI	13	placebo	4 mg	8 mg

During each of three crossover treatment periods, patients entered the sleep laboratory approximately 2 – 2.5 hours before bedtime on two consecutive nights. Patients were administered study medication 30 minutes prior to “lights out”, and PSG measurements were recorded for 8 hours. A five- to twelve-day washout period followed periods 1 and 2.

The mean LPS was the primary measure of efficacy. LPS was defined in the same manner as in Study 21 and was calculated from the PSG recordings from two consecutive nights (per period). Several secondary endpoints of interest were identified and included: total sleep time, sleep efficiency, wake time after sleep onset, and number of awakenings. With the exception of sleep efficiency, these endpoints were measured via PSG and subjective assessments. Additional subjective measures included ease of falling back to sleep after awakening and sleep quality.

The sample size of 100 was formulated to detect a difference in means of 12 minutes between active and placebo treatments with 90% power. The sample size was calculated using a paired t-test (with Bonferroni adjustment) and assuming a standard deviation of 25 minutes and a dropout rate of 0.15. Study participants were enrolled at 17 centers across the United States.

#### *Patient Disposition, Demographic and Baseline Characteristics*

Descriptive demographic and baseline characteristics were summarized for the 100 randomized patients and did not differ among the treatment sequences. The ages of patients were between 65 and 83 with a mean age of 71. Females comprised 63% of the patient population, and approximately 95% of the study participants were Caucasian. Baseline measurements included

weight, height, and body mass index. A table outlining the composition of the study population is presented in the appendix. All patients completed the study.

*Statistical Methodologies*

The primary efficacy variable was analyzed using a mixed linear model. In general, mixed models are often employed when the data exhibits correlation and nonconstant variability. In the cross-over trial design, patients have observations from each treatment period and these observations are correlated as they originate from the same person. In Study 17, the applicant analyzed the mean LPS using a mixed model with sequence, period, treatment, and carryover as fixed effects and subject within sequence as a random effect. According to the applicant, “The carryover effect was removed from the analysis model for the primary efficacy variable if it was not significant at the 0.100 level.” The applicant also anticipated that periods closer together in time would be more closely correlated than periods farther apart. Statistically, this assumption resulted in the specification of an auto-regressive structure, AR(1), for the residual covariance matrix.

The applicant also performed a categorical or responder analysis to support the results of the primary analysis. The definition of a responder mimicked that of Study 21; however in the crossover design, every patient was a responder or non-responder for each treatment. In addition, the analyses of the secondary measures of efficacy, TST and sleep quality, were similar to those of the primary efficacy variable.

All analyses were conducted on the ITT population including all randomized patients receiving at least one dose of the study medication. When assessing the primary efficacy measure, Fisher’s protected least significant difference (LSD) was used to address multiplicity induced by the pairwise comparisons of interest. Similar to Study 21, a stepwise application of Fisher’s protected LSD was used to address the multiplicity induced by the testing of the secondary variables, TST and sleep quality.

*Results and Conclusions*

Based on the results depicted in Table 7, the applicant concluded that the LPS was shorter for patients receiving ramelteon 4 mg and 8 mg compared to patients receiving placebo.

**Table 7: Summary of LPS (minutes): ITT Population**  
(Source: Adapted from Final Study Report 01-02-TL-375-017, Table 11.a)

	<b>Placebo (n=100)</b>	<b>Ramelteon 4 mg (n=100)</b>	<b>Ramelteon 8 mg (n=100)</b>	<b>Overall p-value</b>
LS Mean (SE)	38.4 (2.49)	28.7 (2.49)	30.8 (2.52)	
LS Mean difference from placebo (SE)		-9.7 (2.64)	-7.6 (2.68)	<0.001
95% CI of difference		(-14.9, -4.5)	(-12.9, -2.3)	
Pairwise P-value		<0.001	0.005	

Note: LS means are from a mixed model with effects for sequence, subject (sequence), period and treatment, with subject within sequence as a random effect. In a model with effects for sequence, subject (sequence), period, treatment and carryover, with subject within sequence as a random effect, the p-value for the carryover effect was 0.583.

In general, the mixed model approach is a useful statistical tool in crossover studies where missing data are present and/or an adjustment for a carry-effect is desirable. However, the current study did not have any missing observations and the carry-over effect was not significant; therefore, the benefits of employing the model used by the applicant were theoretically minimal. Moreover, I had some concern regarding the applicant's assumption regarding the correlation among periods. To support the proposed primary analysis and to alleviate my concern, I reanalyzed the primary efficacy variable using a model with patient, period, and treatment as factors. My results were very similar to the results produced by the applicant. In addition, the results of the responder analysis, shown in Table 8, supported the primary analysis.

**Table 8: Latency to Persistent Sleep: Categorical Analysis**  
(Source: Adapted from Final Study Report 01-02-TL-275-017, Table 14.2.1.4)

	Placebo	Ramelteon 4 mg	Ramelteon 8 mg
Latency to Persistent Sleep (minutes)			
<= 30	45	65	65
> 30	55	35	35

Note: The average of two nights' data per dose is used for each subject. Overall and pairwise comparisons were obtained from the chi-square tests and were significant ( $p=0.004$ ). One hundred patients contributed data to each treatment group.

The applicant additionally evaluated the total sleep time and the sleep quality. Both doses of ramelteon resulted in a significant increase in the total sleep time. In comparison to placebo, the ramelteon 4 mg dose resulted in an average increase of 9 minutes of sleep time, and the ramelteon 8 mg dose resulted in an average increase of 12 minutes of sleep time. Patients in the study did not report significant differences in the quality of sleep across treatments.

### 3.1.3 Study TL025

#### *Study Design and Endpoints.*

Study 25 was a randomized, double-blind, placebo-controlled, multi-center, outpatient study in elderly patients with chronic insomnia. Following a 7-night, single-blind, placebo lead-in period, eligible patients were randomized to placebo, 4 mg of ramelteon, or 8 mg of ramelteon. During the 35-night double-blind treatment period, patients maintained a sleep diary and visited a study center weekly. At the conclusion of the treatment period, patients received placebo for seven nights. This placebo run-out phase was used to evaluate rebound insomnia and withdrawal effects.

The primary measure of efficacy was the mean subjective sleep latency (sSL) over the initial seven nights of double-blind treatment. Subjective sleep latency was defined as the patient's perception of the time taken to fall asleep. Secondary measures of efficacy included the sSL at weeks 3 and 5, the subjective total sleep time, sleep quality, and the clinical global impression (CGI) of the change of condition. Of note, the CGI was a subjective evaluation completed by investigators for each patient. Components of the CGI included the severity of illness, the

therapeutic effect, side effects, and the global rating of change of condition. The clinical perception of the change of condition was rated using a 7-point scale ranging from 1 (very much improved) to 7 (very much worse).

Similar to Study 21, rebound insomnia was evaluated via the change from baseline in sSL on each day of the run-out period. Withdrawal effects were evaluated via an assessment of the change in total BWSQ score from the end of double-blind treatment (i.e. week 5) to the single-blind placebo run-out period.

A sample size of 810 was formulated to detect a difference in mean sSL of 12 minutes between active and placebo treatments with 90% power. The sample size was calculated using a paired t-test (with Bonferroni adjustment) and assuming a standard deviation of 35 minutes and a dropout rate of 0.20. Study participants were enrolled at 136 centers across the United States.

*Patient Disposition, Demographic, and Baseline Characteristics*

Descriptive demographics and baseline characteristics were summarized using all randomized patients and did not differ between treatment groups. The ages of patients were between 64 and 93 with a mean age of 72. In the study, 90% of study participants were Caucasian. Females comprised 59% of the patient population. The average height of study participants was 167 centimeters and the average weight was 74 kilograms. A table outlining the composition of the study population is presented in the appendix.

Of the 829 randomized participants, 281 were randomized to ramelteon 4 mg, 274 were in the ramelteon 8 mg group, and 274 were randomized to placebo. During the course of the study, 136 (16%) patients discontinued the study. Eight of these patients withdrew during the single-blind placebo run-out period. Of the 128 patients that discontinued during the double-blind treatment period, 40 patients discontinued because of a lack of efficacy, 35 patients discontinued because of a protocol violation, and 23 patients discontinued because of an adverse event. Of the patients discontinuing because of adverse events, 8 were randomized to the ramelteon 4 mg group, 7 were randomized to ramelteon 8 mg and 8 were randomized to placebo. The number of withdrawals per week is depicted in Table 9 below.

**Table 9: Number of patient withdrawals per week**  
(Source: Response to Information request, 17 April 2005 )

	Placebo	Ramelteon 8 mg	Ramelteon 16 mg
Week 1	7	9	5
Week 2	15	8	5
Week 3	9	14	10
Week 4	8	7	11
Week 5 and beyond	10	12	6

### *Statistical Methodologies*

The statistical methodologies used in Study 25 mimicked the methodologies used in Study 21. The mean subjective sleep latency (sSL) over the initial 7 nights of double blind treatment was analyzed via an analysis of covariance (ANCOVA) model with treatment and center as main effects and baseline sleep latency as a covariate. An examination of the consistency of the results across centers was conducted via inclusion of a treatment-by-center interaction in the ANCOVA model. The type I error was controlled via Fisher's protected LSD. Additionally, an evaluation of the  $\Delta$  efficacy was conducted via an application of Fisher's protected LSD at weeks 3 and 5. Analyses were performed on the ITT population including all randomized patients receiving at least one dose of the study medication. A LOCF strategy was employed to handle missing data.

The applicant also performed a categorical or responder analysis to support the results of the primary analysis. A "responder" was defined as a participant having sSL less than or equal to 30 minutes. The responder status of a discontinued patient was determined via a LOCF strategy. The responder analysis employed a CMH test, stratified by pooled center, to evaluate the association between treatment and the percentage of patients achieving sSL less than or equal to 30 minutes.

Analyses of the subjective total sleep time (sTST), subjective sleep quality, and global rating of change of condition were similar to the analysis of the primary efficacy measure. According to the applicant, interpretation of significant results for sTST was contingent on observing statistically significant results for the primary measure of efficacy at week 1. Likewise, interpretation of results for sleep quality was contingent on results of sTST, and interpretation of results for the CGI was contingent on results of sleep quality at week 1. According to the applicant, "Global rating of change condition was analyzed with the baseline CGI of severity of illness as a covariate."

Rebound insomnia was assessed during the seven day single-blind run-out period using observed data only. Rebound insomnia was analyzed for each day via an ANCOVA model with treatment and pooled center as effects and baseline LPS as a covariate. Summary statistics were additionally calculated. Withdrawal effects were assessed using the total BWSQ score and analyzed using an ANCOVA model with treatment and pooled center as effects and the week 5 total BWSQ score as a covariate.

### *Results and Conclusions*

The applicant's results are numerically depicted in Table 10 and graphically depicted in Figure 2. The applicant concluded that the average sSL as perceived by the patient was shorter for the ramelteon (4 mg and 8 mg) groups as compared to the placebo group. The applicant further concluded that the treatment effect was maintained throughout the duration of the study. Of note, two patients did not have any post-baseline data and were not included in the primary analysis.

Similar to study 21, I reanalyzed the data using a baseline observation carried forward (BOCF) strategy. The overall results of my reanalysis (see appendix) were consistent with the results produced employing a LOCF imputation scheme.

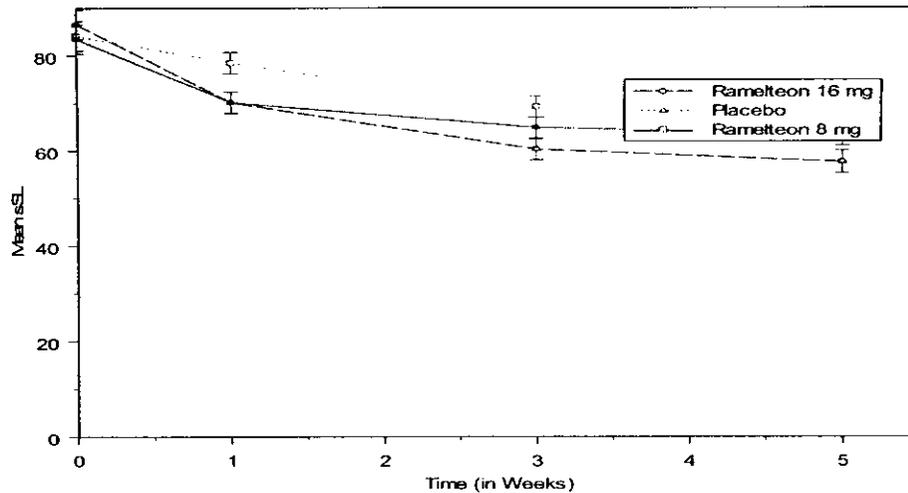
**Table 10: Summary of sSL (minutes) – LOCF Data: ITT Population**

(Source: Adapted from Final Study Report 01-02-TL-375-025, Table 11.a)

	Placebo (N=274)	Ramelteon 4 mg (N=280)	Ramelteon 8 mg (N=272)	Overall p-value
<b>Baseline</b>				
LS Mean (SE)	84.2 (3.13)	83.5 (3.07)	86.6 (3.12)	
<b>Week 1</b>				
LS Mean (SE)	78.5 (2.24)	70.2 (2.21)	70.2 (2.24)	
LS mean difference from placebo (SE)		-8.3 (3.10)	-8.3 (3.12)	0.0009
95% CI of difference		(-14.4, -2.2)	(-14.5, -2.2)	
Pairwise p-value		0.008	0.008	
<b>Week 3</b>				
LS Mean (SE)	69.3 (2.19)	64.9 (2.16)	60.3 (2.19)	
LS mean difference from placebo (SE)		-4.5 (3.03)	-9.0 (3.05)	0.013
95% CI of difference		(-10.4, 1.5)	(-15.0, -3.0)	
Pairwise p-value		0.142	0.003	
<b>Week 5</b>				
LS Mean (SE)	70.6 (2.36)	63.4 (2.32)	57.7 (2.36)	
LS mean difference from placebo (SE)		-7.1 (3.25)	-12.8 (3.28)	<0.001
95% CI of difference		(-13.5, -0.8)	(-19.3, -6.4)	
Pairwise p-value		0.028	<0.001	

Note: L.S. means for baseline are from an ANOVA model with effects for treatment and pooled center. L.S. means for a post-baseline visit are from an ANCOVA model with effects for treatment and pooled center and the baseline value of the variable as a covariate. P-values for pairwise comparisons are obtained using t-tests from the ANCOVA model of the overall treatment comparisons. When the treatment-by-center interaction term was included in the model at Week 1, the p-value for the interaction was 0.281.

Figure 2: Mean sSL Across Weeks



The applicant additionally performed a responder analysis. The results of the responder analysis are shown in Table 11 and did not support the primary analysis at week 1. There was not an association between treatment and responder status at week 1. While evidence of a difference in the proportions of responders versus non-responders between treatments did exist at week 3, the association was not evident at week 5. I, again, had some concern regarding the definition of responder used in the categorical analysis. As in Study 21, I reanalyzed the data altering the definition of a responder to only include patients who completed the study. The results of my analysis were in agreement with the results of the applicant's responder analysis. My results are included in the appendix.

To elucidate the discrepancy between the primary analysis and the supportive analysis, I examined the cumulative distribution functions (CDFs) of each treatment arm. The CDFs provided insight into the probability that the latency to persistent sleep was achieved within a certain amount of time. The reduction in mean latency for the active groups compared to the placebo group was mainly attributable to a reduction of large values to slightly less large values. There was little difference in the proportion of values less than 30 minutes.

Appears This Way  
On Original

**Table 11: Subjective Sleep Latency: Categorical Analysis – LOCF Data**  
(Source: Adapted from Final Study Report 01-02-TL-375-025, Table 14.2.1.7.)

	Placebo (n=274)	Ramelteon 4 mg* (n=280)	Ramelteon 8 mg (n=273)	Overall p-value
<b>Baseline</b>				
<= 30	9	8	5	
> 30	265	273	268	
<b>Week 1</b>				0.716
<= 30	42	49	43	
> 30	232	231	230	
p-value for comparison with placebo		0.353	0.731	
<b>Week 3</b>				0.042
<= 30	54	71	76	
> 30	220	209	197	
p-value for comparison with placebo		0.072	0.010	
<b>Week 5</b>				0.474
<= 30	71	80	81	
> 30	203	200	192	
p-value for comparison with placebo		0.349	0.225	

\*Note: n=281 for the baseline measurement in the ramelteon 4 mg group. Patient 252453 did not have any post-baseline data.

Note: The subjective measurements were collected in the subject diary. Baseline is the average of all data collected before double-blind treatment. The average of data collected during a study week is used for each subject at each post-baseline visit. Overall and pairwise p-values are obtained from the CMH general association test, stratified by pooled center.

The applicant additionally evaluated the sTST, the sleep quality, and the CGI rating of change of condition. For the latter two variables, there were no significant treatment differences at any of the time points. A significant difference in the total sleep times between the ramelteon 4 mg group and the placebo group was detected at weeks 1 and 3.

The applicant's results from the analyses of rebound insomnia and withdrawal effects are in the appendix. Based on the results, the applicant concluded that there was no evidence of rebound insomnia or withdrawal effects.

### 3.1.4 Study TL020

#### *Study Design and Endpoints*

The design of study 20 mimicked that of Study 25. The studies differed in that Study 20 was conducted in adults with chronic insomnia, and Study 25 was conducted in elderly patients. Additionally, the treatment arms differed in the studies. In Study 20, eligible patients were randomized to placebo, 8 mg of ramelteon, or 16 mg of ramelteon.

The primary measure of efficacy was the mean subjective sleep latency (sSL) over the initial seven nights of double-blind treatment. Several “important” secondary measures of efficacy identified by the applicant included the sSL at weeks 3 and 5, the subjective total sleep time, sleep quality, and the clinical global impression (CGI) of the global rating of change of condition.

#### *Patient Disposition, Demographic, and Baseline Characteristics*

Descriptive demographics and baseline characteristics were summarized for all patients and did not differ among treatment groups. The mean age of patients was 44. In the study, 69% of study participants were Caucasian, 17% were African-American, and 11% were Hispanic. Of the 848 randomized patients, 499 were female. Baseline measurements included weight, height, and body mass index. A table outlining the composition of the study population is presented in the appendix.

Two hundred and seventy-seven study participants were randomized to ramelteon 8 mg, 284 were in the ramelteon 16 mg group, and 287 were randomized to placebo. During the double-blind treatment period, 136 patients discontinued. Twenty-five discontinued due to an adverse event. Of these patients, 6 were randomized to the ramelteon 8 mg group, 12 were randomized to ramelteon 16 mg, and 7 were randomized to placebo.

#### *Statistical Methodologies*

The statistical methodologies used in Study 20 mimicked the methodologies used in the previous parallel group studies (study 21 and study 25). The primary analysis employed an ANCOVA model with treatment and center as main effects and baseline sleep latency as a covariate. The type I error was controlled via Fisher’s protected LSD. An evaluation of [ ] efficacy was conducted via an application of Fisher’s protected LSD at weeks 3 and 5. The applicant also performed a responder analysis to support the results of the primary analysis. A CMH test, stratified by pooled center, was used in the responder analysis. Analyses of the secondary variables followed the same strategy as used for the primary efficacy variable. Analyses were performed on the ITT population including all randomized patients receiving at least one dose of the study medication. A LOCF strategy was employed to handle missing data.

#### *Results and Conclusions*

Table 12 depicts the results of the applicant’s primary analysis at week 1. A treatment effect was not demonstrated at week 1; therefore, subsequent time points were not assessed for efficacy (based on the pre-specified stepwise application of Fisher’s protected LSD). The applicant’s responder analysis also demonstrated a lack of a treatment effect.

**Table 12: Summary of sSL (minutes) – LOCF Data: ITT Population**

(Source: Adapted from Final Study Report 01-02-TL-375-020, Table 11.a)

	Placebo	Ramelteon 4 mg	Ramelteon 8 mg	Overall p-value
<b>Baseline</b>				
n	287	277	284	
LS Mean (SE)	85.5 (2.99)	85.2 (3.03)	92.5 (2.98)	
<b>Week 1</b>				0.602
n	283	270	276	
LS Mean (SE)	74.4 (2.17)	74.8 (2.20)	77.2 (2.17)	
LS mean difference from placebo (SE)		0.4 (3.01)	2.8 (3.00)	
95% CI of difference		(-5.5, 6.3)	(-3.1, 8.7)	
Pairwise p-value		0.888	0.349	

Note: L.S. means for baseline are from an ANOVA model with effects for treatment and pooled center. L.S. means for a post-baseline visit are from an ANCOVA model with effects for treatment and pooled center and the baseline value of the variable as a covariate. P-values for pairwise comparisons are obtained using t-tests from the ANCOVA model of the overall treatment comparisons. When the treatment-by-center interaction term was included in the model at Week 1, the p-value for the interaction was 0.541.

### 3.1.5 Study TL023

#### *Study Design and Endpoints*

Study 23 was a randomized, double-blind, placebo-controlled, multi-center study in healthy adults naïve to a sleep laboratory environment. Following an initial screening period, eligible volunteers were randomized to ramelteon 8 or 16 mg or placebo. Participants were administered a single dose of medication approximately 30 minutes prior to their habitual bedtime. PSG measurements were taken during an uninterrupted 8 hours in sleep laboratories. Participants were discharged on the following day.

The primary measure of efficacy was the objective PSG measurement of latency to persistent sleep. Similar to other studies, secondary measures of interest included the total sleep time and the sleep quality.

A sample size of 270 was formulated to detect a difference in mean LPS of 10 minutes between active and placebo treatments with 90% power. The sample size was calculated using a paired t-test (with Bonferroni adjustment) and assuming a standard deviation of 18 minutes. Study participants were enrolled at 15 centers across the United States.

#### *Patient Disposition, Demographic, and Baseline Characteristics*

Descriptive demographics and baseline characteristics were summarized for all volunteers and did not differ among treatment groups. The ages of volunteers ranged from 18 to 63 with a mean age of 29. In the study, 67% of study participants were Caucasian, 22% were Hispanic, and 6%

were African-American. Of the 289 randomized participants, 161 or 56% were female. A detailed table outlining the composition of the study population is presented in the appendix.

Ninety-seven of the study participants were randomized to placebo, 98 were randomized to ramelteon 8 mg, and 94 patients were randomized to ramelteon 16 mg. One volunteer in the placebo group discontinued because of adverse events of increased sweating and agitation.

### *Statistical Methodologies*

The statistical methodologies mimicked those used in previous parallel group studies. The primary analysis employed an analysis of variance (ANOVA) model with factors for treatment and center. Each dose of ramelteon was compared to placebo using Fisher's protected LSD to control the type I error. The homogeneity of the treatment effect across centers was evaluated via inclusion of a treatment-by-center interaction term.

A categorical analysis of the LPS was conducted to support the primary analysis. The CMH test, stratified by pooled center, was used to compare the proportions of patients with LPS of 30 minutes or less among treatment groups. The two secondary variables of specific interest to the applicant were analyzed using an ANOVA model with treatment and center as effects. Moreover, the stepwise application of Fisher's protected LSD was again applied as in previous studies to control the type I error. All analyses were conducted on the ITT population including all randomized volunteers who received a dose of the study medication.

### *Results and Conclusions*

Table 13 depicts the applicant's results. One participant did not have PSG measurements and was therefore excluded from the applicant's analysis. The applicant concluded that treatment with ramelteon 8 mg significantly shortened the mean latency to persistent sleep. However, the applicant's conclusion was not supported by the additional categorical analysis (shown in Table 14). When the proportion of participants achieving LPS within 30 minutes or less was evaluated, the applicant found that there was no difference in the proportions of individuals achieving or not achieving LPS within 30 minutes. I was able to reproduce the applicant's primary results, and I did not have any disagreement with the statistical findings or methodologies.

Appears This Way  
On Original

**Table 13: Summary of LPS (minutes): ITT Population**  
 (Source: Adapted from Final Study Report 01-02-TL-375-023, Table 11.a)

	Placebo (n=97)	Ramelteon 8 mg (n=98)	Ramelteon 16 mg (n=93)	Overall p-value
LS Mean (SE)	19.7 (1.87)	12.2 (1.88)	14.8 (1.93)	
LS mean difference from placebo (SE)		-7.6 (2.62)	-4.9 (2.65)	0.015
95% CI of difference		(-12.7,-2.4)	(-10.1,0.3)	
Pairwise p-value		0.004	0.065	

Note: L.S. means for treatment comparisons are from an ANOVA model with effects for treatment and pooled center. P-values for pairwise comparisons are obtained using t-tests from the ANOVA model of the overall treatment comparisons. When the treatment-by-center interaction term was included in the model, the p-value for the interaction was 0.540.

**Table 14: Latency to Persistent Sleep: Categorical Analysis (ITT population)**  
 (Source: Adapted from Final Study Report 01-02-TL-375-023, Table 14.2.1.4)

	Placebo (n=97)	Ramelteon 8 mg (n=98)	Ramelteon 16 mg (n=93)	Overall p-value
Latency to Persistent Sleep (minutes)				0.292
<= 30	83	91	83	
> 30	14	7	10	
Pairwise p-value		0.143	0.616	

Note: Overall and pairwise p-values are obtained from the CMH general association test, stratified by pooled center.

The applicant additionally evaluated the total sleep time and the sleep quality. The total sleep time was significantly increased for participants randomized to ramelteon. Participants receiving ramelteon 8 mg slept an average of 17 minutes longer than participants randomized to placebo. Similarly, individuals receiving ramelteon 16 mg slept an average of 14 minutes longer than individuals receiving placebo. Moreover, participants in the study did not report significant differences in the quality of sleep between treatments.

### 3.2 Evaluation of Safety

The evaluation of the safety data was conducted by Dr. D. Elizabeth McNeil. The reader is referred to Dr. McNeil's review for information regarding the adverse event profile.

## 4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

### 4.1 Gender, Race and Age

Analyses were performed with respect to gender, age, and race for the studies conducted in adults with chronic insomnia (i.e. Study 21 and Study 23). For studies 17 and 25 conducted in elderly patients, analyses were performed with respect to gender and race. The applicant did not propose any efficacy claims for any subgroups of patients.

#### 4.1.1 Study TL021

The applicant repeated the primary analyses for each of the subgroup categories. Specifically, age was dichotomized into two groups, namely, age  $\geq 40$  and age  $< 40$ . Race was categorized as Caucasian, African-American, Hispanic, or other.

Among both males and females, a significant difference in mean LPS between the ramelteon groups and the placebo group was demonstrated. A significant difference was also detected in the younger participants. In contrast, a treatment effect was not evident in older patients. The evidence of a treatment effect varied among racial groups. Among Caucasians, a shortened LPS was detected at week 1 only. Among African-Americans and Hispanics, an overall effect was detected at weeks 1 and 3. No effect was seen in other racial groups. The results of the applicant's subgroup analyses are depicted in Table 15.

Appears This Way  
On Original

**Table 15: Analysis of Difference from Placebo in LPS by Subgroups – LOCF Data  
ITT Population**

(Source: Adapted from Final Study Report 01-02-TL-375-021, Tables 14.2.1.11.2-14.2.1.13.2)

	Ramelteon 8 mg		Ramelteon 16 mg					
	<b>By Gender</b>							
	<b>Male</b> (n=57)	<b>Female</b> (n=81)	<b>Male</b> (n=46)	<b>Female</b> (n=89)				
Week 1	-19.4* (7.94)	-15.5* (4.45)	-26.0* (8.18)	-16.6* (4.31)				
Week 3	-21.3* (8.43)	-12.0* (4.74)	-30.1* (8.68)	-13.3* (4.60)				
Week 5	-24.0* (10.00)	-8.8* (4.0)	-22.9* (10.29)	-10.8* (3.88)				
	<b>By Age</b>							
	<b>&lt;40</b> (n=82)	<b>≥40</b> (n=56)	<b>&lt;40</b> (n=69)	<b>≥40</b> (n=66)				
Week 1	-21.5* (4.57)	-8.8 (6.26)	-28.1* (4.80)	-8.0 (5.97)				
Week 3	-19.2* (5.54)	-7.1 (5.78)	-28.4* (5.82)	-5.1 (5.52)				
Week 5	-17.1* (5.14)	-4.5 (6.65)	-21.3* (5.40)	-1.9 (6.35)				
	<b>By Race</b>							
	<b>C</b> (n=87)	<b>A</b> (n=19)	<b>H</b> (n=26)	<b>O</b> (n=6)	<b>C</b> (n=82)	<b>AA</b> (n=23)	<b>H</b> (n=27)	<b>O</b> (n=3)
Week 1	-13.6* (5.10)	-17.9 (10.47)	-21.1* (7.84)	-36.0 (15.00)	-16.3* (5.20)	-25.0* (9.98)	-29.0* (7.61)	-41.5 (16.26)
Week 3	-10.2 (5.81)	-31.2* (9.34)	-12.6 (7.00)	-20.8 (10.59)	-13.8* (5.92)	-31.3* (8.89)	-19.3* (6.81)	-14.9 (11.48)
Week 5	-9.5 (5.42)	-17.6 (12.71)	-10.5 (8.49)	33.9 (55.59)	-8.2 (5.52)	-24.5* (12.10)	-12.9 (8.25)	6.4 (60.25)

By Race Categories: C = Caucasian; A= African-American; H = Hispanic; O = Other

\* Significant at .05 level

#### 4.1.2 Study TL017

The results of the applicant's subgroup analyses for Study 17 are depicted in Table 16. Subpopulations defined by race were categorized as Caucasian or non-Caucasian. Among females, a significant difference in LPS of 10 minutes and 9 minutes was detected for the ramelteon 4 mg and ramelteon 8 mg groups respectively (when compared to placebo). In addition, a significant difference was demonstrated in Caucasians.

**Table 16: Summary of Subgroup Analysis of LPS: ITT Population**  
 (Source: Adapted from Final Study Report 01-02-TL-375-017, Table 11.d)

	Placebo	Ramelteon 8 mg	Ramelteon 16 mg
<b>Gender: Male</b>			
N	37	37	37
LS mean (SE)	35.3 (4.77)	26.2 (4.75)	28.5 (4.83)
LS mean difference from placebo		-9.1 (3.86)	-6.8 (4.03)
<b>Gender: Female</b>			
N	63	63	63
LS mean (SE)	40.1 (3.00)	29.9 (3.00)	31.1 (3.02)
LS mean difference from placebo		-10.1 (3.50)*	-8.8 (3.55)*
<b>Race: Caucasian</b>			
N	95	95	95
LS Mean (SE)	37.5 (2.57)	28.5 (2.57)	30.4 (2.60)
LS mean difference from placebo		-9.0 (2.69)*	-7.1 (2.74)*
<b>Race: non-Caucasian</b>			
N	5	5	5
LS Mean (SE)	51.5 (8.71)	28.9 (8.64)	39.6 (8.71)
LS mean difference from placebo		-22.5 (12.65)	-11.9 (12.80)

\* Significant at the 0.05 level

#### 4.1.3 Study TL025

The applicant investigated the effect of ramelteon measured via subjective sleep latency (sSL) for race and gender subgroups. Race was categorized as Caucasian, Hispanic or Asian, African-American, Native American and other. For simplicity, I labeled the latter category as other. The results of the analyses are depicted in Table 17.

Appears This Way  
 On Original

**Table 17: Analysis of Difference from Placebo in sSL by Subgroups – LOCF Data**  
 (Source: Adapted from Final Study Report 01-02-TL-375-025,  
 Tables 14.2.1.11.2-14.2.1.13.2)

	Ramelteon 4 mg		Ramelteon 8 mg	
	<b>By Gender</b>			
	Male (n=110)	Female (n=170)	Male (n=122)	Female (n=150)
Week 1	-7.3 (5.23)	-10.1* (4.13)	-10.4 (5.09)	-8.2 (4.26)
Week 3	-4.8 (5.59)	-4.5 (3.82)	-10.1 (5.44)	-8.0 (3.94)
Week 5	-0.9 (5.75)	-10.6* (4.22)	-9.6 (5.60)	-12.9* (4.36)
	<b>By Race</b>			
	Caucasian (n=251)	Other (n=29)	Caucasian (n=239)	Other (n=33)
Week 1	-7.5* (3.30)	-10.8 (10.53)	-8.6* (3.35)	-7.8 (10.45)
Week 3	-3.7 (3.20)	-22.8 (11.62)	-7.5 (3.24)	-30.1* (11.54)
Week 5	-6.4 (3.46)	-24.9 (13.02)	-12.4* (3.51)	-29.1 (12.93)

\* Significant at .05 level

#### 4.1.3 Study TL023

The results of the subgroup analysis for Study 23 are depicted in Table 18. Age was dichotomized into age less than 40 and age greater than or equal to 40. Race was categorized as Caucasian, Hispanic or Asian, African-American, Native American and other. For simplicity, I labeled the latter category as other. According to the applicant, “When compared to the placebo group, treatment effect of ramelteon 8 mg was more notable with respect to LPS in the subgroups of men, subjects less than 40 years old, and Caucasian subjects, respectively.” The applicant also commented that the treatment effect of ramelteon 16 mg was more notable in men.

Appears This Way  
 On Original

**Table 18: Summary of Subgroup Analysis of LPS: ITT Population**  
 (Source: Adapted from Final Study Report 01-02-TL-375-023, Table 11.d)

	Placebo	Ramelteon 8 mg	Ramelteon 16 mg
<b>Gender: Male</b>			
N	40	43	44
LS mean (SE)	25.8 (3.79)	12.6 (3.79)	14.1 (3.74)
LS mean difference from placebo		-13.2 (5.22)*	-11.7 (5.16)*
<b>Gender: Female</b>			
N	57	55	49
LS mean (SE)	15.2 (1.70)	12.0 (1.82)	14.3 (1.87)
LS mean difference from placebo		-3.2 (2.48)	-0.9 (2.51)
<b>Age: &lt; 40</b>			
N	77	83	84
LS Mean (SE)	16.7 (1.52)	11.3 (1.50)	13.8 (1.48)
LS mean difference from placebo		-5.4 (2.09)*	-2.9 (2.08)
<b>Age: ≥ 40</b>			
N	20	15	9
LS Mean (SE)	33.7 (8.21)	12.0 (9.75)	17.1 (2.41)
LS mean difference from placebo		-8.7 (3.54)	-5.0 (3.43)
<b>Race: Caucasian</b>			
N	64	60	68
LS Mean (SE)	20.8 (2.52)	12.1 (2.59)	15.7 (2.41)
LS mean difference from placebo		-8.7 (3.54)*	-5.0 (3.43)
<b>Race: Hispanic</b>			
N	21	22	19
LS Mean (SE)	21.5 (4.31)	19.6 (4.20)	17.0 (3.72)
LS mean difference from placebo		-1.9 (3.78)	-4.5 (4.13)
<b>Race: Other</b>			
N	12	16	6
LS Mean (SE)	22.5 (7.37)	6.3 (5.57)	9.3 (10.73)
LS mean difference from placebo		-16.2 (8.91)	-13.1 (11.36)

\* Significant at the 0.05 level

Appears This Way  
 On Original

## 5. SUMMARY AND CONCLUSIONS

### 5.1 Statistical Issues and Collective Evidence

Prior to the applicant's submission of the NDA, the agency expressed concerns regarding multiplicity and the handling of missing data. Specifically, the agency was concerned that the testing of the primary measure of efficacy at multiple time points would inflate the type I error and that the last observation carried forward (LOCF) imputation strategy was inappropriate for a potentially fast-acting drug. An additional issue that arose during my review was the significance of a comparison of means versus a responder analysis.

During the course of my review, I initially considered the ability of ramelteon to promote sleep onset. Studies 17 and 21 evaluated efficacy via the mean latency to persistent sleep measured by polysomnography and provided evidence of a treatment effect for the 8 mg dose of ramelteon. The former study was conducted in an elderly population while Study 21 was conducted in individuals younger than 65 years of age. Studies 20 and 25 were also conducted in individuals with chronic insomnia; however, the studies were conducted in outpatient or real world settings and evaluated efficacy via a subjective assessment of sleep latency. The applicant acknowledged the lack of positive findings in Study 20. In contrast, Study 25 provided some evidence of a treatment effect at week 1; however, the effect was absent when evaluating the data via a categorical or responder analysis. To elucidate the discrepancy between the analysis of means and the responder analysis, I examined the cumulative distribution functions (CDFs) of each treatment arm at week 1. The reduction in mean latency for the active groups compared to the placebo group was mainly attributable to a reduction of large values to slightly less large values. There was little difference in the proportion of values less than 30 minutes.

Subsequent to my review of sleep onset, I evaluated [ ] the treatment effect. Studies 20, 21, and 25 were conducted for 35 days. During meetings prior to the submission of the NDA, the agency suggested that the applicant sequentially test the treatment effects at the multiple time points to offset the multiplicity concern. The applicant used the recommended methodology thereby alleviating my concern regarding the inflation of the type I error. The applicant claimed that both studies 21 and 25 demonstrated evidence of [ ] the effect. In both studies, I repeated the applicant's primary analyses using a baseline observation carried forward (BOCF) imputation strategy. I also used a varying definition of a responder for the categorical analyses. The results from my re-analyses employing a BOCF strategy were similar to the results produced via a LOCF imputation scheme. However, the responder analysis was not supportive of the claim [ ] in Study 25.

For completion, the applicant conducted a study in a healthy population to evaluate the effectiveness of ramelteon in individuals with transient insomnia. The applicant concluded that

8 mg of ramelteon was effective in reducing the latency to persistent sleep. The applicant's conclusions were not supported by the additional categorical analysis.

Two "key" secondary variables were identified by the applicant and assessed in the studies. In a meeting on July 16, 2002, the agency recommended that the applicant consider other secondary variables for support of the desired claims. Specifically the agency stated, "We noted that other proposed secondary outcomes, i.e., TST and sleep efficiency, were not ideal and would likely not be acceptable for supporting a [ ] claim." Upon review of the total sleep time and sleep efficiency, I found that the measures did not provide consistent support across studies. This phenomenon was also evident across additional secondary variables assessed by the applicant. Of note, the study protocols did not provide an explanation of the relative importance of the additional secondary variables or an explanation of the role of the variables in the interpretation of the results.

## **5.2 Conclusions and Recommendations**

The applicant submitted NDA 21-782 to provide evidence of the efficacy and safety of ramelteon 8 mg for the treatment of insomnia. The applicant claims that the drug reduces the latency to persistent sleep and that the treatment effect is maintained through five weeks. The applicant further suggests that ramelteon increases the total sleep time.

Based on my review of the data, I conclude that the applicant has shown evidence that ramelteon promotes sleep onset during an initial week of treatment as measured by the latency to persistent sleep. The applicant has also provided some evidence that the treatment effect is maintained for a prolonged period of time. However, I disagree with the applicant's assessment of the total sleep time and other supportive analyses. Specifically, the total sleep time is not ideal for support of the proposed claims and did not consistently demonstrate efficacy across studies. In addition, the applicant did not garner support for the overall conclusions via pre-specified, supplemental analyses (i.e. responder analyses) in the study conducted in a real world setting. Lastly various safety concerns, such as the genotoxic and carcinogenic potential of the product and the effect of the product on the endocrine system, have arisen during the course of the clinical and pharmacology/toxicology reviews; therefore, the team will need to collectively evaluate the risks and benefits of ramelteon.

Appears This Way  
On Original

### 5.2.1 Labeling

The draft label describes the four clinical studies submitted in support of the effectiveness of ramelteon. The clinical trials section of the label reads as follows:

#### **Controlled Trials Supporting Efficacy**

##### *Chronic Insomnia*

[

]

I have several recommendations based on my evaluation of the submission. First, I suggest the deletion of the phrase [ ]. I believe this deletion throughout the label will not alter the interpretation. Second, I recommend removal of the results pertaining to the total sleep time. The measure of total sleep time is not ideal for support of the proposed claims and does not consistently demonstrate efficacy across studies. Lastly, I recommend the label reflect that the findings pertain to the average latency to persistent sleep. An additional issue that will need to be discussed among the review team will be the inclusion of label claims for doses other

than the proposed 8 mg dose. My recommendations may be included in the label in the following manner:

### **Controlled Trials Supporting Efficacy**

#### *Chronic Insomnia*

Ramelteon was studied in two randomized, double-blind trials in subjects with chronic insomnia employing polysomnography (PSG) [ ]

One study enrolled 405 adults (aged 18 to 64 years, inclusive) with chronic insomnia and employed a parallel design in which subjects received a single nightly dose of Ramelteon 8 mg or 16 mg or matching placebo for 35 days. PSG was performed on the first two nights in each of Weeks 1, 3, and 5 of treatment. Both doses of Ramelteon reduced the average latency to persistent sleep at all time points when compared to placebo.

The second study employing PSG was a three-arm crossover trial performed in [ ] subjects aged 65 years and older with a history of chronic insomnia. Subjects received Ramelteon 4 mg or 8 mg or placebo and underwent PSG assessment in a sleep laboratory for two consecutive nights in each of the three study periods. Both doses of Ramelteon [ ] = latency to persistent sleep.

A randomized, double-blind, parallel group study was conducted in [ ] outpatients aged 65 years and older with chronic insomnia and employed subjective measures of efficacy (sleep diaries). Subjects received Ramelteon 4 mg or 8 mg or placebo for 35 nights. Both doses of Ramelteon [ ]

#### *Transient Insomnia*

In a randomized, double-blind, parallel group trial using a first-night-effect model, [ ] healthy adults received placebo or Ramelteon 8 mg or 16 mg before spending one night in a sleep laboratory and being evaluated with PSG. [ ]

Appears This Way  
On Original

## APPENDICES

### *Overall Glossary*

**Latency to Persistent Sleep (LPS)** – the elapsed time from the beginning of the PSG recording to the onset of the first 10 minutes of continuous sleep (i.e. total number of epochs before the first 20 consecutive nonwake epochs, divided by 2)

**Total sleep time (TST)** – the sum of all of the minutes of Stages 1, 2, 3/4 NREM and REM sleep

**Sleep efficiency** – the total sleep time divided by total time in bed, multiplied by 100 (time-in-bed is the number of minutes from beginning of the PSG recording to the end of the recording)

**Wake time after sleep onset (WASO)** – the number of wake minutes after the onset of persistent sleep prior to the end of the recording

**Number of awakenings (NAW)** – the number of times after onset of persistent sleep that there was a wake entry of at least 2 epochs in duration; each entry must have been separated by Stage 2, 3/4 NREM sleep or REM sleep in order to be counted.

Appears This Way  
On Original

Study 21

**Summary of Demographic and Baseline Characteristics: ITT Population**  
(Source: Final Study Report 01-02-TL-375-021, Table 10.b)

Characteristic	Treatment			Overall (n=405)	p-value
	Placebo (n=131)	Ramelteon 8mg (n=139)	Ramelteon 16 mg (n=135)		
Gender, n(%)					0.007
Male	30 (22.9)	57 (41.0)	46 (34.1)	133 (32.8)	
Female	101 (77.1)	82 (59.0)	89 (65.9)	272 (67.2)	
Mean Age (SD) (yr)	39.7 (11.96)	38.0 (11.53)	40.2 (12.44)	39.3 (11.99)	0.226
Race, n(%)					0.971
Caucasian	79 (60.3)	87 (62.6)	82 (60.7)	248 (61.2)	
Asian	3 (2.3)	3 (2.2)	2 (1.5)	8 (2.0)	
Hispanic	21 (16.0)	19 (13.7)	23 (17.0)	63 (15.6)	
Black	27 (20.6)	27 (19.4)	27 (20.0)	81 (20.0)	
Native American	0 (0.0)	1 (0.7)	0 (0.0)	1 (0.2)	
Other	1 (0.8)	2 (1.4)	1 (0.7)	4 (1.0)	
Mean weight (SD)(kg)	71.16 (14.88)	75.93 (15.04)	72.14 (12.44)	73.12 (14.29)	0.006
Mean height (SD) (cm)	166.39 (9.15)	170.16 (10.28)	168.05 (9.23)	168.24 (9.67)	0.005
Mean BMI (SD) (kg/m <sup>2</sup> )	25.60 (4.41)	26.06 (3.64)	25.53 (3.83)	25.73 (3.96)	0.309

Note: There were no significant differences among treatments. Overall p-value for continuous variables from ANOVA with treatment and pooled center as factors. Overall p-value for categorical variables from CMH general association test, stratified by pooled center.

Appears This Way  
On Original

**Summary of LPS (minutes) – BOCF Data: ITT Population  
(My re-analysis, Study 21)**

	<b>Placebo (n=131)</b>	<b>Ramelteon 8 mg (n=139)</b>	<b>Ramelteon 16 mg (n=135)</b>	<b>Overall p-value</b>
<b>Baseline</b>				
LS Mean (SE)	65.3 (3.54)	64.3 (3.46)	68.4 (3.54)	
<b>Week 1</b>				
LS Mean (SE)	47.8 (2.72)	32.3 (2.66)	28.9 (2.71)	
LS mean difference from placebo (SE)		-15.5 (3.69)	-18.9 (3.73)	<0.001
95% CI of difference		(-22.8, -8.3)	(-26.3, -11.6)	
Pairwise p-value		<0.001	<0.001	
<b>Week 3</b>				
LS Mean (SE)	45.3 (2.88)	34.4 (2.82)	27.6 (2.88)	
LS mean difference from placebo (SE)		-10.9 (3.91)	-17.6 (3.96)	<0.001
95% CI of difference		(-18.6, -3.2)	(-25.4, -9.9)	
Pairwise p-value		0.006	<0.001	
<b>Week 5</b>				
LS Mean (SE)	43.5 (2.98)	35.8 (2.93)	31.1 (2.97)	
LS mean difference from placebo (SE)		-7.70 (4.05)	-12.4 (4.08)	0.010
95% CI of difference		(-15.7, 0.3)	(-20.4, -4.3)	
Pairwise p-value		0.058	0.003	

Note: L.S. means for baseline are from an ANOVA model with effects for treatment and pooled center. L.S. means for a post-baseline visit are from an ANCOVA model with effects for treatment and pooled center and the baseline value of the variable as a covariate.

Appears This Way  
On Original

**Analysis of Rebound Insomnia – Observed Data**

(Source: Adapted from Final Study Report 01-02-TL-275-021, Table 14.2.20.2)

<b>Change from baseline in LPS (minutes)</b>	<b>Placebo</b>	<b>Ramelteon 8 mg</b>	<b>Ramelteon 16 mg</b>	<b>Overall p-value</b>
<b>Baseline</b>				
N	131	139	135	
LS Mean	65.3	64.3	68.4	
SE	3.54	3.46	3.54	
<b>Day 1, Placebo Run-Out</b>				
N	118	124	128	0.025
LS Mean	-19.1	-34.7	-29.1	
SE	4.25	4.20	4.16	
LS Mean Difference from Placebo		-15.7	-10.1	
SE of Difference		5.80	5.75	
95% CI of difference		(-27.1, -4.3)	(-21.4, 1.2)	
p-value for comparison with placebo		0.007	0.081	
<b>Day 2, Placebo Run-Out</b>				
N	116	121	128	0.482
LS Mean	-29.8	-22.8	-28.1	
SE	4.40	4.38	4.27	
LS Mean Difference from Placebo		7.0	1.7	
SE of Difference		6.04	5.93	
95% CI of difference		(-4.9, 18.9)	(-9.9, 13.4)	
p-value for comparison with placebo		0.249	0.771	

Note: LS means for baseline are from an ANOVA model with effects for treatment and pooled center. LS means for Day 1 and Day 2 off-treatment are from an ANCOVA model with effects for treatment and pooled center and the baseline latency to persistent sleep as a covariate. P-values for pairwise comparisons are obtained using t-tests from the ANCOVA model of the overall treatment comparison.

Appears This Way  
On Original

**Analysis of Tyrer BWSQ – Observed Data**

(Source: Adapted from Final Study Report 01-02-TL-275-021, Table 14.2.21.2)

<b>Change from baseline in LPS (minutes)</b>	<b>Placebo</b>	<b>Ramelteon 8 mg</b>	<b>Ramelteon 16 mg</b>	<b>Overall p-value</b>
<b>Week 5</b>				
N	122	124	130	
LS Mean	0.9	0.8	0.6	
SE	0.18	0.18	0.18	
<b>Day 1, Placebo Run-Out</b>				
N	119	123	128	0.090
LS Mean	0.1	-0.1	-0.2	
SE	0.08	0.08	0.08	
LS Mean Difference from Placebo		-0.2	-0.2	
SE of Difference		0.11	0.11	
95% CI of difference		(-0.4,0.0)	(-0.4,0.0)	
p-value for comparison with placebo		0.069	0.047	
<b>Day 2, Placebo Run-Out</b>				
N	118	121	127	0.226
LS Mean	-0.1	-0.2	-0.1	
SE	0.08	0.08	0.08	
LS Mean Difference from Placebo		-0.2	-0.1	
SE of Difference		0.11	0.11	
95% CI of difference		(-0.4,0.0)	(-0.3,0.1)	
p-value for comparison with placebo		0.085	0.376	

Note: LS means for Week 5 are from an ANOVA model with effects for treatment and pooled center. Ls Means for Day 1 and Day 2 off-treatment are from an ANCOVA model with effects for treatment and pooled center and the Week 5 total BWSQ score as a covariate. P-values for pairwise comparisons are obtained using t-tests from the ANCOVA model of the overall treatment comparison.

Appears This Way  
On Original

Study 17

Summary of Demographic and Baseline Characteristics: ITT Population  
(Source: Final Study Report 01-02-TL-375-017, Table 10.b)

Characteristic	Treatment Sequence						Overall n=100
	I n=19	II n=16	III n=15	IV n=22	V n=15	VI n=13	
Gender, n(%)							
Male	6 (31.6)	7 (43.8)	3 (20.0)	9 (40.9)	5 (33.3)	7 (53.8)	37 (37.0)
Female	13 (68.4)	9 (56.3)	12 (80.0)	13(59.1)	10 (66.7)	6 (46.2)	63 (63.0)
Mean Age	70.8	71.3	70.5	71.5	69.9	69.7	70.7
(SD) (yr)	(4.34)	(5.47)	(4.87)	(5.19)	(3.69)	(4.07)	(4.63)
Race, n(%)							
Caucasian	18 (94.7)	15 (93.8)	15 (100.0)	21 (95.5)	14 (93.3)	12 (92.3)	95 (95.0)
Asian	1 (5.3)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	1 (1.0)
Hispanic	0 (0.0)	1 (6.3)	0 (0.0)	1 (4.5)	1 (6.7)	1 (7.7)	4 (4.0)
Black	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Native	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
American							
Other	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Mean weight	66.98	75.68	72.21	72.37	74.26	80.19	73.18
(SD) (kg)	(13.76)	(14.25)	(11.68)	(11.38)	(15.77)	(18.82)	(14.34)
Mean height	163.43	165.81	163.49	166.26	167.14	171.46	166.04
(SD) (cm)	(9.06)	(9.06)	(8.85)	(10.10)	(9.13)	(10.37)	(9.56)
Mean BMI	24.86	27.45	26.91	26.21	26.41	26.97	26.38
(SD) (kg/m <sup>2</sup> )	(3.05)	(3.77)	(2.81)	(3.70)	(4.30)	(4.13)	(3.64)

Note: There were no significant differences among treatments. Overall p-value for continuous variables from one-way ANOVA. Overall p-value for categorical variables from chi-square tests.

Appears This Way  
On Original

Study 25

**Summary of Demographic and Baseline Characteristics: ITT Population**  
(Source: Final Study Report 01-02-TL-375-025, Table 10.b)

Characteristic	Treatment			Overall (n=829)
	Placebo (n=274)	Ramelteon 8mg (n=281)	Ramelteon 16 mg (n=274)	
Gender, n(%)				
Male	108 (39.4)	110 (39.1)	123 (44.9)	341 (41.1)
Female	166 (60.6)	171 (60.9)	151 (55.1)	488 (58.9)
Mean Age (SD) (yr)	72.4 (5.94)	72.1 (6.03)	72.6 (5.88)	72.4 (5.95)
Race, n(%)				
Caucasian	251 (91.6)	252 (89.7)	241 (88.0)	744 (89.7)
Black	9 (3.3)	14 (5.0)	17 (6.2)	40 (4.8)
Hispanic	8 (2.9)	12 (4.3)	11 (4.0)	31 (3.7)
Asian	3 (1.1)	0	3 (1.1)	6 (0.7)
Native American	1 (0.4)	3 (1.1)	0	4 (0.5)
Other	2 (0.7)	0	2 (0.7)	4 (0.5)
Mean weight (SD)(kg)	73.39 (13.90)	73.64 (13.00)	74.90 (14.52)	73.97 (13.81)
Mean height (SD) (cm)	167.04 (9.22)	167.30 (9.34)	167.03 (9.90)	167.12 (9.48)
Mean BMI (SD) (kg/m <sup>2</sup> )	26.18 (3.57)	26.21 (3.51)	26.66 (3.56)	26.35 (3.55)

Note: There were no significant differences among treatments. Overall p-value for continuous variables from ANOVA with treatment and pooled center as factors. Overall p-value for categorical variables from CMH general association test, stratified by pooled center.

**Subjective Sleep Latency – Categorical Data**  
(My re-analysis, Study 25)

	Placebo (n=274)	Ramelteon 4 mg (n=281)	Ramelteon 8 mg (n=274)	Overall p-value
<b>Week 1</b>				0.727
<= 30	42	49	43	
> 30	232	232	231	
p-value for comparison with placebo		<0.359	<0.743	
<b>Week 3</b>				0.013
<= 30	47	68	73	
> 30	227	213	201	
p-value for comparison with placebo		.020	0.003	
<b>Week 5</b>				.399
<= 30	60	73	69	
> 30	214	208	205	
p-value for comparison with placebo		.160	.316	

Note: This table differs from that of the sponsor in that patients who drop out prematurely are considered non-responders. In addition, 829 patients are considered as opposed to 827.

**Summary of sSL (minutes) – BOCF Data: ITT Population**  
(My re-analysis, Study 25)

	<b>Placebo (n=274)</b>	<b>Ramelteon 4 mg (n=281)</b>	<b>Ramelteon 8 mg (n=273)</b>	<b>Overall p-value</b>
<b>Baseline</b>				
LS Mean (SE)	84.2 (3.13)	83.5 (3.07)	86.6 (3.12)	
<b>Week 1</b>				
LS Mean (SE)	78.5 (2.24)	70.2 (2.20)	70.1 (2.24)	
LS mean difference from placebo (SE)		-8.3 (3.09)	-8.3 (3.11)	0.009
95% CI of difference		(-14.3, -2.2)	(-14.4, -2.2)	
Pairwise p-value		0.008	0.008	
<b>Week 3</b>				
LS Mean (SE)	67.8 (1.98)	65.3 (1.94)	60.7 (1.97)	
LS mean difference from placebo (SE)		-2.5 (2.73)	-7.1 (2.75)	0.031
95% CI of difference		(-7.8, 2.9)	(-12.5, -1.75)	
Pairwise p-value		0.365	0.010	
<b>Week 5</b>				
LS Mean (SE)	69.3 (2.14)	64.6 (2.10)	59.1 (2.14)	
LS mean difference from placebo (SE)		-4.7 (2.95)	-10.2 (2.98)	0.003
95% CI of difference		(-10.5, 1.09)	(-16.0, -4.3)	
Pairwise p-value		0.111	0.001	

Note: L.S. means for baseline are from an ANOVA model with effects for treatment and pooled center. L.S. means for a post-baseline visit are from an ANCOVA model with effects for treatment and pooled center and the baseline value of the variable as a covariate.

Appears This Way  
On Original

**Summary of Rebound Insomnia – Observed Data**  
(Source: Final Study Report 01-02-TL-275-025, Table 11e)

<b>Change from baseline in LPS (minutes)</b>	<b>Placebo</b>	<b>Ramelteon 4 mg</b>	<b>Ramelteon 8 mg</b>	<b>Overall p-value</b>
<b>Baseline</b>				
N	274	281	273	
LS Mean(SE)	84.2 (3.13)	83.5 (3.07)	86.6 (3.12)	
<b>Day 1, Placebo Run-Out</b>				
N	222	232	236	0.014
LS Mean (SE)	-16.8 (3.29)	-26.8 (3.18)	-29.3 (3.19)	
LS Mean Difference from Placebo (SE)		-10.0 (4.50)	-12.5 (4.47)	
95% CI of difference		(-18.8,-1.2)	(-21.3, -3.7)	
p-value for comparison with placebo		0.027	0.005	
<b>Day 2, Placebo Run-Out</b>				
N	225	231	236	0.039
LS Mean (SE)	-19.6 (3.33)	-23.1 (3.25)	-31.0 (3.25)	
LS Mean Difference from Placebo (SE)		-3.5 (4.59)	-11.3 (4.55)	
95% CI of difference		(-12.5, 5.5)	(-20.2,-2.4)	
p-value for comparison with placebo		0.450	0.013	
<b>Day 3, Placebo Run-Out</b>				
N	223	231	237	0.922
LS Mean (SE)	-23.0 (3.52)	-25.0 (3.43)	-24.3 (3.42)	
LS Mean Difference from Placebo (SE)		-1.9 (4.84)	-1.3 (4.80)	
95% CI of difference		(-11.4, 7.6)	(-10.7, 8.2)	
p-value for comparison with placebo		0.691	0.792	
<b>Day 4, Placebo Run-Out</b>				
N	221	226	233	0.519
LS Mean (SE)	-23.8 (3.28)	-26.1 (3.22)	-28.9 (3.21)	
LS Mean Difference from Placebo (SE)		-2.3 (4.53)	-5.1 (4.48)	
95% CI of difference		(-11.2, 6.6)	(-13.9, 3.7)	
p-value for comparison with placebo		0.611	0.254	
<b>Day 5, Placebo Run-Out</b>				
N	216	222	229	0.885
LS Mean (SE)	-22.9 (3.71)	-22.3 (3.61)	-24.7 (3.60)	
LS Mean Difference from Placebo (SE)		0.6 (5.10)	-1.8 (5.05)	
95% CI of difference		(-9.4, 10.6)	(-11.7,8.2)	
p-value for comparison with placebo		0.904	0.728	

<b>Change from Baseline in sSL (minutes)</b>	<b>Placebo</b>	<b>Ramelteon 4 mg</b>	<b>Ramelteon 8 mg</b>	<b>Overall p-value</b>
<b>Day 6, Placebo Run-Out</b>				
N	209	213	220	0.091
LS Mean (SE)	-16.1 (3.87)	-21.9 (3.79)	-27.7 (3.79)	
LS Mean Difference from Placebo (SE)		-5.8 (5.34)	-11.6 (5.29)	
95% CI of difference		(-16.3, 4.7)	(-22.0, -1.2)	
p-value for comparison with placebo		0.278	0.029	
<b>Day 7, Placebo Run-Out</b>				
N	180	194	200	0.402
LS Mean (SE)	-22.7 (4.00)	-21.3 (3.78)	-28.1 (3.79)	
LS Mean Difference from Placebo (SE)		1.4 (5.44)	-5.3 (5.37)	
95% CI of difference		(-9.3, 12.1)	(-15.9, 5.2)	
p-value for comparison with placebo		0.797	0.322	

Note: LS means for baseline are from an ANOVA model with effects for treatment and pooled center. LS means for a post-baseline visit are from an ANCOVA model with effects for treatment and pooled center and the baseline value of the variable as a covariate. P-values for pairwise comparisons are obtained using t-tests from the ANCOVA model of the overall treatment comparison.

Appears This Way  
On Original

**Analysis of BWSQ – Observed Data**  
 (Source: Final Study Report 01-02-TL-275-025, Table 11.f)

Change from Week 5 in Total BWSQ	Placebo	Ramelteon 4 mg	Ramelteon 8 mg	Overall p-value
<b>Week 5</b>				
N	228	234	238	
LS Mean (SE)	0.8 (0.12)	0.8 (0.12)	0.9 (0.12)	
<b>Day 7, Placebo Run-Out</b>				
N	228	232	237	0.330
LS Mean (SE)	-0.1 (0.06)	-0.1 (0.06)	-0.2 (0.06)	
LS Mean Difference from Placebo (SE)		0.0 (0.09)	-0.1 (0.09)	
95% CI of difference		(-0.2,0.2)	(-0.3, 0.1)	
Pairwise p-value		0.994	0.199	

Note: LS means for Week 5 are from an ANOVA model with effects for treatment and pooled center. Ls Means for Day 7 off-treatment are from an ANCOVA model with effects for treatment and pooled center and the Week 5 total BWSQ score as a covariate. P-values for pairwise comparisons are obtained using t-tests from the ANCOVA model of the overall treatment comparison.

Appears This Way  
On Original

Study 23

Summary of Demographic and Baseline Characteristics: ITT Population  
(Source: Final Study Report 01-02-TL-375-023, Table 10.b)

Characteristic	Treatment			Overall (n=289)
	Placebo (n=97)	Ramelteon 8mg (n=98)	Ramelteon 16 mg (n=94)	
Gender, n(%)				
Male	40 (41.2)	43 (43.9)	45 (47.9)	128 (44.3)
Female	57 (58.8)	55 (56.1)	49 (52.1)	161 (55.7)
Mean Age (SD) (yr)	29.8 (9.17)	28.5 (9.07)	28.1 (9.40)	28.8 (9.21)
Race, n(%)				
Caucasian	64 (66.0)	60 (61.2)	69 (73.4)	193 (66.8)
Asian	7 (7.2)	5 (5.1)	2 (2.1)	14 (4.8)
Black	4 (4.1)	10 (10.2)	4 (4.3)	18 (6.2)
Hispanic	21 (21.6)	22 (22.4)	19 (20.2)	62 (21.5)
Other	1 (1.0)	1 (1.0)	0 (0.0)	2 (0.7)
Mean weight (SD)(kg)	73.39 (15.83)	71.82 (13.86)	72.33 (14.60)	72.52 (14.75)
Mean height (SD) (cm)	171.01 (9.41)	169.84 (10.42)	170.20 (10.11)	170.35 (9.97)
Mean BMI (SD) (kg/m <sup>2</sup> )	24.91 (3.90)	24.88 (4.18)	24.86 (3.94)	24.88 (3.99)

Note: There were no significant differences among treatments. Overall p-value for continuous variables from ANOVA with treatment and pooled center as factors. Overall p-value for categorical variables from CMH general association test, stratified by pooled center.

Appears This Way  
On Original

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

/s/  
-----

Dionne Price  
6/22/05 10:08:21 AM  
BIOMETRICS

Thomas Permutt  
6/22/05 10:33:30 AM  
BIOMETRICS  
concur

S. Edward Nevius  
6/22/05 11:25:41 AM  
BIOMETRICS  
Concur with review.



DEPARTMENT OF HEALTH AND HUMAN SERVICES  
FOOD AND DRUG ADMINISTRATION  
CENTER FOR DRUG EVALUATION AND RESEARCH  
OFFICE OF BIOSTATISTICS

## Statistical Review and Evaluation

### STABILITY STUDIES

NDA: 21-782/N-000

Name of drug:  $\text{C}$   $\text{J}$  (Ramelteon Tablets TAK-375)

Sponsor: Takeda Pharmaceutical Company Ltd.

Indication: insomnia

Documents reviewed: \\Cdsesub1\p21782\N\_000\2005-04-21\cmc\product\datasets

Project manager: Sara Stradley, M.S.

Chemistry reviewer: Pramoda Maturu, Ph.D.

Dates: Received 09/21/2004; PDUFA 06/21/2004;

Statistical reviewer: Joan Buenconsejo

Statistics team leader: Karl Lin, Ph.D.

Biometrics division director: Edward Nevius, Ph.D.

Keywords: NDA review, stability study

---

## 1 BACKGROUND

---

The sponsor submitted the stability data for ramelteon tablets (TAK-375) to support its proposed shelf life. The stability of ramelteon tablets in each of the four proposed commercial package configurations (i.e. 30-cc bottle, 100-cc bottle, 500-cc bottle, and blister) has been carried under conditions representing both long-term (25°C/60% relative humidity (RH)) and accelerated conditions (40°C/75%RH). Additional studies have been performed under stress conditions on one selected lots.

The primary stability studies are comprised of studies on two groups of tablet lots. The main group contains three lots of 8 mg tablets that were produced at pilot scale (Z515F01, Z515F02, and Z515F03) by Takeda Chemical Industries, Ltd. Osaka, Japan. A second group contains three lots of tablets that were produced at full commercial scale (1237, 1247 and 1287) by Takeda Ireland, Ltd., Kilruddery, Ireland, to provide site-specific data for confirmation of findings obtained with the main group. Both groups of tablets were packaged using identical components for the corresponding package sizes. According to the sponsor, all primary studies are still ongoing, and the currently available data include the 24-month test results for pilot-scale tablets and the results for the commercial-scale tablets. Supportive data from clinical batches in additional types of packaging were also presented in the report. Furthermore, stress testing under severe conditions of temperature and humidity, as well as evaluation of photostability, has been performed on one lot of tablets from the pilot-scale group.

Appears This Way  
On Original

## 2 SPONSOR'S STABILITY SUMMARY AND CONCLUSION

A summary of the studies performed and data submitted by the sponsor for ramelteon tablets, 8 mg is provided in Table 1. The electronic data was submitted as part of the NDA amendment. Meanwhile, Table 2 summarizes the specification used by the Sponsor for stability testing of ramelteon tablets.

The sponsor performed separate statistical analysis for the pilot-scale and commercial-scale lots, based respectively on the 24-month [ ] results, under the long-term storage condition at 25°C/60%RH. A linear estimate with a 1-sided 95% confidence interval (for parameters with a one-sided specification) or a 2-sided 95% confidence interval (for parameters with an upper and a lower specification) was calculated over data for each parameter of interest. Parameters evaluated by the Sponsors at 25°C/60%RH include assay, dissolution at 15 minutes and at — minutes, and related substance to aid in estimation of a suitable expiration dating period.

Summary of statistical parameters obtained in analyses of assay, dissolution, and related substance results is provided in Table 3.

Table 1: Summary of Stability Studies Performed for Ramelteon Tablets, 8 mg

Study Type	Container	Number of Batches	Storage Conditions Evaluated	Time Completed
Primary Pilot-Scale	HDPE bottles (30-count)	3	25°C/60% RH -	
		3	40°C/75% RH -	
	HDPE bottles (100-count)	3	25°C/60% RH -	
		3	40°C/75% RH -	
	HDPE bottles (500-count)	3	25°C/60% RH -	
		3	40°C/75% RH -	
Primary Site-Specific (Commercial-Scale)	blister package	3	25°C/60% RH -	
		3	40°C/75% RH -	
	HDPE bottles (30-count)	3	25°C/60% RH -	
		3	40°C/75% RH -	
	HDPE bottles (100-count)	3	25°C/60% RH -	
		3	40°C/75% RH -	
HDPE bottles (500-count)	3	25°C/60% RH -		
	3	40°C/75% RH -		
Supporting	PVC blister package	3	25°C/60% RH -	
		3	40°C/75% RH -	
Photostability (ICH Option 1)	HDPE bottle (100-count)	3	25°C/60% RH -	
		3	40°C/75% RH -	
Severe Conditions	Glass bottle (closed)	1	50°C	
		1	60°C	
		1	25°C/31% R	
		1	25°C/93% RH	

<sup>1</sup> One lot was packaged in 1-count and 100-count bottles; each of the other two lots was packaged in only 100-count or 100-count bottles.

Test items: Appearance, assay, related substances and dissolution for all; and in some cases also microbial limits, enantiomer, hardness, loss on drying

Table 2: Specification of Ramelteon Tablets, 8 mg

Test	Analytical procedure	Acceptance criteria
Appearance	Visual inspection	Pale orange-yellow film coated tablets with " [ ] " and "8" imprinted on one side
Identification	M-11-00547	
	M-11-00554	
Assay (%)	M-11-00530	[ ] of the labeled amount of $C_{16}H_{21}NO_2$
Related Substances (%)	M-11-00524	Any Individual NMT \
		Total NMT \
Dissolution (%)	M-11-00527.001R	NLT (Q) of the labeled amount of $C_{16}H_{21}NO_2$ is dissolved in — minutes
Content Uniformity (%)	M-11-00550.001R, based on USP <905>	Meets requirements

Table 3A: Summary of Statistical Parameters Obtained in Analyses of Assay Results

Package	Worst Case Lot	Slope	Intercept	Estimated Expiry (Months)	Predicted Assay at L (%)
<b>Pilot-Scale</b>					
HDPE bottle (30 count)	Z515F01 <sup>1</sup>				
HDPE bottle (100 count)	Z515F01 <sup>1</sup>				
HDPE bottle (500 count)	Z515F01 <sup>1</sup>				
— blister	Z515F01 <sup>1</sup>				
<b>Commercial-Scale</b>					
HDPE bottle (30 count)	01247 <sup>2</sup>				
— bottle (100 count)	01247 <sup>3</sup>				
HDPE bottle (500 count)	01247 <sup>4</sup>				
— blister	01247 <sup>5</sup>				

<sup>1</sup> Common slope was obtained for all three lots  
<sup>2</sup> Slope was for lot 01237 and for lot 01287  
<sup>3</sup> Slope was for lot 01237 and for lot 01287  
<sup>4</sup> Slope was for lot 01237 and for lot 01287  
<sup>5</sup> Slope was for lot 01237 and for lot 01287

Table 3B: Summary of Statistical Parameters Obtained in Analyses of Total Related Substance Results

Package	Worst Case Lot	Slope	Intercept	Estimated Expiry (Months)	Predicted Content at C (%)	J
<b>Pilot-Scale</b>						
HDPE bottle (30 count)	Pooled					
HDPE bottle (100 count)	Pooled					
HDPE bottle (500 count)	Pooled					
— blister	Pooled					
<b>Commercial-Scale</b>						
HDPE bottle (30 count)	01237 <sup>1</sup>					
HDPE bottle (100 count)	Pooled					
HDPE bottle (500 count)	01247 <sup>2</sup>					
— blister	Pooled					

<sup>1</sup> Slope was  $\frac{1}{100}$  for lot 01247 and  $\frac{1}{100}$  for lot 01287

<sup>2</sup> Common slope was obtained for all three lots

Table 3C: Summary of Statistical Parameters Obtained in Analyses of Dissolution Results

Package	Worst Case Lot <sup>1</sup>	Slope	Intercept	Estimated Expiry (Months) <sup>2</sup>	Predicted Mean % Dissolved at C (%)	J
<b>15-Minute Sample Interval</b>						
HDPE bottle (30 count)	01247 <sup>3</sup>			NA		
HDPE bottle (100 count)	Pooled			NA		
HDPE bottle (500 count)	01237 <sup>4</sup>			NA		
— blister	Pooled			NA		
<b>30-Minute Sample Interval</b>						
HDPE bottle (30 count)	Pooled					
HDPE bottle (100 count)	Pooled					
HDPE bottle (500 count)	Pooled					
— blister	Pooled					

<sup>1</sup> For 15-minute analyses, the worst case is the lot with the lowest bound (●) months. For ●-minute analyses, the worst case is the lot with the shortest estimated expiry

<sup>2</sup> NA: not applicable (because no specification is set at the 15-minute sample interval on which to base an estimated expiry)

<sup>3</sup> Common slope was obtained for all three lots

<sup>4</sup> Slope was  $\frac{1}{100}$  for lot 01247 and  $\frac{1}{100}$  for lot 01287

In summary, the sponsor reported that ramelteon 8 mg tablets have been found to be very stable stored in each of the four proposed market package configurations under long-term and accelerated storage conditions. Although the drug substance exhibits a moderate sensitivity to light, they found that no significant effects were observed for tablets placed in a glass dish and exposed to fluorescent-light illumination.

According to the sponsor, the main degradation product of ramelteon in tablets is [redacted] It was formed [redacted] Other unidentified impurities have also been observed, primarily for 40°C/75% RH storage, but according to them, none have exceeded the ICH identification threshold. The [redacted] ramelteon remained at non-detectable levels under both storage conditions in all packages.

Dissolution results show a slow decrease in percent drug dissolved over time. However, because of the rapid solubility characteristics of ramelteon, the observed mean amount dissolved after [redacted] minutes was greater than [redacted] for tablets in all packages stored for 24 months at 25°C/60% RH. Extrapolations from linear regression analysis of the data predicted a potential worst-case result of [redacted] for the four packages after [redacted] of storage.

The maximum amount of water in tablets measured by [redacted] Tablet hardness showed only small decreases that were most notable for the 30-count bottles and [redacted] blisters.

Results in supporting studies carried out to [redacted] of long-term storage are consistent with observations in the primary studies.

Meanwhile, in studies under more severe stress conditions of heat and humidity, tablets were more susceptible to degradation from heat than from humidity. Total related substances increased under most of the conditions with only small changes observed for tablets stored in open bottles stored at [redacted] increase) and the largest increases found for tablets in closed bottles [redacted] increase). [redacted] was the major degradation product under all conditions.

Based on the results obtained by the sponsor in the primary and supporting stability studies, and extrapolation suggested in the ICH guidance Q1E for a product showing little or no change over time and little or no variability, **an initial expiration dating period of [redacted] is proposed for ramelteon 8 mg tablets in all packages.** This is derived by [redacted] the available amount of long-term data (24 months).

---

### 3 REVIEWER'S STABILITY ANALYSES

---

The sponsor submitted the electronic data on April 21, 2005. The data set included data up to 24 months from the three pilot-scale lots (lots: Z515F01, Z515F02 and Z515F03), and data up to [ ] from the three commercial-scale lots (lots: 01237, 01247, and 01287). This reviewer analyzed the data in accordance with FDA's "Guidelines for Submitting Documentation for Stability of Human Drugs Biologics."

Tables 4 and 5 summarize the results for all parameters tested using data sets from the pilot-scale and commercial-scale lots stored at 25°C/60%RH. The results from the reviewer's analyses based on two sets of data **do not appear to support [ ] expiration date**. If we only consider pilot-scale lot (Table 4), it seems that the shortest estimated expiration dating period is [ ] months, based on the analysis of dissolution at 15 minutes under [ ] packaging. Disregarding the analyses of dissolution at 15 minutes will support the [ ] expiration date based on the pilot-scale data. Similarly, for the commercial-scale data, the shortest estimated expiration dating period is also [ ] This is again due to the data from dissolution at 15 minutes completed thus far did not meet the proposed specification. Disregarding the analyses of dissolution data at 15 minutes will only support up to [ ] expiration date based on the commercial-scale lot. ]

---

### 4 CONCLUSION

---

The results of this reviewer's analysis using data from pilot-scale and commercial-scale lots stored at 25°C/60%RH suggest that the stability data **do not support [ ] expiration date**. In fact, the data does not support any expiration date, because the dissolution data at 15 minutes completed thus far did not meet the specification.

Appears This Way  
On Original

Table 4: Expiry Date Analysis for Ramelteon Tablets, 8 mg stability data (Pilot Scale Lot)

Test	Package	Batch	Specification	Model	Intercept	Slope	Expiry Date
Assay	30-cc	Z515F01 Z515F02 Z515F03	—	The regression lines have separate slopes and intercepts			
	100-cc	Z515F01 Z515F02 Z515F03		The regression lines have separate slopes and intercepts			
	500-cc	Z515F01 Z515F02 Z515F03		The regression lines have separate slopes and intercepts			
	—	Z515F01 Z515F02 Z515F03		The regression lines have separate slopes and intercepts	/	/	/
Total Related Substance	30-cc	Z515F01 Z515F02 Z515F03	—	Pooled			
	100-cc	Z515F01 Z515F02 Z515F03		Pooled			
	500-cc	Z515F01 Z515F02 Z515F03		Pooled			
	—	Z515F01 Z515F02 Z515F03		Pooled			

Table 4 (Continued)

Test	Package	Batch	Specification	Model	Intercept	Slope	Expiry Date
Dissolution at 15 minutes	30-cc	Z515F01 Z515F02 Z515F03	—	The regression lines have separate slopes and intercepts			
	100-cc	Z515F01 Z515F02 Z515F03		The regression lines have separate slopes and intercepts			
	500-cc	Z515F01 Z515F02 Z515F03		The regression lines have separate slopes and intercepts			
	—	Z515F01 Z515F02 Z515F03		The regression lines have separate slopes and intercepts			
Dissolution at 30 minutes	30-cc	Z515F01 Z515F02 Z515F03	—	Pooled			
	100-cc	Z515F01 Z515F02 Z515F03		The regression lines have separate slopes and intercepts			
	500-cc	Z515F01 Z515F02 Z515F03		Pooled			
	—	Z515F01 Z515F02 Z515F03		Pooled			

Table 5: Expiry Date Analysis for Ramelteon Tablets, 8 mg stability data (Commercial Scale Lot)

Test	Package	Batch	Specification	Model	Intercept	Slope	Expiry Date
Assay	30-cc	Z515F01	—	The regression lines have separate slopes and intercepts			
		Z515F02					
		Z515F03					
	100-cc	Z515F01		The regression lines have separate slopes and intercepts			
		Z515F02					
		Z515F03					
	500-cc	Z515F01		The regression lines have separate slopes and intercepts			
		Z515F02					
		Z515F03					
	—	Z515F01		The regression lines have separate slopes and intercepts			
		Z515F02					
		Z515F03					
Total Related Substance	30-cc	Z515F01	—	The regression lines have separate slopes and intercepts			
		Z515F02					
		Z515F03					
	100-cc	Z515F01		Pooled			
		Z515F02					
		Z515F03					
	500-cc	Z515F01		The regression lines have separate slopes and intercepts			
		Z515F02					
		Z515F03					
	—	Z515F01		Pooled			
		Z515F02					
		Z515F03					

Table 5 (Continued)

Test	Package	Batch	Specification	Model	Intercept	Slope	Expiry Date
Dissolution at 15 minutes	30-cc	Z515F01 Z515F02 Z515F03	—	The regression lines have separate slopes and intercepts			
	100-cc	Z515F01 Z515F02 Z515F03		Pooled			
	500-cc	Z515F01 Z515F02 Z515F03		The regression lines have separate slopes and intercepts			
	—	Z515F01 Z515F02 Z515F03		Pooled	/	/	/
Dissolution at 3	30-cc	Z515F01 Z515F02 Z515F03	✓	Pooled	/	/	
	100-cc	Z515F01 Z515F02 Z515F03		Pooled			
	500-cc	Z515F01 Z515F02 Z515F03		Pooled			
	—	Z515F01 Z515F02 Z515F03		Pooled			

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

/s/

-----  
Joan Buenconsejo  
6/7/05 01:58:55 PM  
BIOMETRICS

Karl Lin  
6/7/05 02:28:57 PM  
BIOMETRICS  
Concur with review



U.S. Department of Health and Human Services  
Food and Drug Administration  
Center for Drug Evaluation and Research  
Office of Pharmacoeconomics and Statistical Science  
Office of Biostatistics

## Statistical Review and Evaluation

### CARCINOGENICITY STUDY

BLA/Serial Number: NDA 21,782  
Drug Name: [ ] (Ramelteon) 8 mg tablet  
Indication(s): insomnia  
Applicant: Takeda Global  
Date(s): Applicant's letter date: September 21, 2004  
Review Priority: Standard  
Biometrics Division: Biometrics Division 2  
Statistical Reviewer: Joan Buenconsejo  
Concurring Reviewers: Karl Lin, Ph.D., Biometrics Division 2  
Medical Division: HFD 170  
Pharmacologist: Adam Wasserman, Ph.D., Pharmacology (HFD-170)  
Project Manager: Sara Stradley, M.S.  
**Keywords:** NDA review, carcinogenicity

Table of Contents

<b>EXECUTIVE SUMMARY .....</b>	<b>5</b>
<b>OVERVIEW OF CARCINOGENICITY STUDIES .....</b>	<b>6</b>
<b>SPONSOR'S EVALUATION OF CARCINOGENICITY STUDIES .....</b>	<b>7</b>
<b>REVIEWER'S EVALUATION OF CARCINOGENICITY STUDIES .....</b>	<b>9</b>
<i>Analysis of Male Rats</i> .....	10
<i>Analysis of Female Rats</i> .....	15
<i>Analysis of Male Mice</i> .....	20
<i>Analysis of Female Mice</i> .....	25
<b>CONCLUSIONS .....</b>	<b>30</b>
<b>REFERENCE .....</b>	<b>32</b>
<b>APPENDIX.....</b>	<b>33</b>
EXPLORATORY ANALYSIS: COMBINING TUMORS WITHIN ORGANS.....	33
<i>Adenoma Hepatocellular + Carcinoma Hepatocellular in the Liver</i> .....	33
<i>Adenoma Hepatocellular + Carcinoma Hepatocellular + Hepatoblastoma in the Liver</i> .....	37
STATISTICAL INTERPRETATION OF SIGNIFICANCE IN EVALUATION OF TUMOR-DATA ANALYSES CURRENTLY ADOPTED BY CDER OFFICE OF BIOSTATISTICS.....	39
NUMBER OF TUMOR-BEARING ANIMALS.....	40
<i>Rat Study</i> .....	40
<i>Mouse Study</i> .....	44

Appears This Way  
On Original

List of Tables

TABLE 1 SUMMARY FINDINGS.....	5
TABLE 2 DOSAGE ASSIGNMENT.....	6
TABLE 3 ANALYSIS OF MORTALITY DATA FOR MALE RATS BY TREATMENT AND TIME.....	10
TABLE 4 ANALYSIS OF DOSE-MORTALITY TREND FOR MALE RATS.....	12
TABLE 5 ANALYSIS OF DOSE-MORTALITY TREND FOR MALE RATS (EXCLUDING 1000 MG/KG DOSE GROUP)13	
TABLE 6 REPORT OF P-VALUES < 0.05 FOR TEST OF POSITIVE LINEAR DOSE-TUMOR TRENDS IN MALE RATS: INCLUDING 1000 MG/KG DOSE GROUP.....	13
TABLE 7 PAIRWISE COMPARISONS OF STATISTICAL SIGNIFICANT TUMOR TRENDS IN MALE RATS.....	14
TABLE 8 STATISTICALLY SIGNIFICANT POSITIVE LINEAR DOSE-TUMOR TREND FOUND IN MALE RATS.....	14
TABLE 9 ANALYSIS OF MORTALITY DATA FOR FEMALE RATS BY TREATMENT AND TIME.....	15
TABLE 10 ANALYSIS OF DOSE-MORTALITY TREND FOR FEMALE RATS.....	17
TABLE 11 ANALYSIS OF DOSE-MORTALITY TREND FOR FEMALE RATS (EXCLUDING 1000 MG/KG DOSE GROUP).....	18
TABLE 12 REPORT OF P-VALUES < 0.05 FOR TEST OF POSITIVE LINEAR DOSE-TUMOR TRENDS IN FEMALE RATS: INCLUDING 1000 MG/KG GROUP IN THE ANALYSIS.....	18
TABLE 13 DECISION RULE FOR STATISTICAL SIGNIFICANCE FOR FEMALE RATS.....	19
TABLE 14 PAIRWISE COMPARISONS OF STATISTICAL SIGNIFICANT TUMOR TRENDS IN FEMALE RATS.....	19
TABLE 15 STATISTICALLY SIGNIFICANT POSITIVE LINEAR DOSE-TUMOR TREND FOUND IN FEMALE RATS.....	19
TABLE 16 ANALYSIS OF MORTALITY DATA FOR MALE MICE BY TREATMENT AND TIME.....	20
TABLE 17 ANALYSIS OF DOSE-MORTALITY TREND FOR MALE MICE.....	22
TABLE 18 ANALYSIS OF DOSE-MORTALITY TREND FOR MALE MICE (EXCLUDING 1000 MG/KG DOSE GROUP) .....	23
TABLE 19 REPORT OF P-VALUES < 0.05 FOR TEST OF POSITIVE LINEAR DOSE-TUMOR TRENDS IN MALE MICE: INCLUDING 1000 MG/KG DOSE GROUP.....	23
TABLE 20 DECISION RULE FOR STATISTICAL SIGNIFICANCE FOR MALE MICE.....	24
TABLE 21 PAIRWISE COMPARISONS OF STATISTICAL SIGNIFICANT TUMOR TRENDS IN MALE MICE.....	24
TABLE 22 STATISTICALLY SIGNIFICANT POSITIVE LINEAR DOSE-TUMOR TREND FOUND IN MALE MICE.....	24
TABLE 23 ANALYSIS OF MORTALITY DATA FOR FEMALE MICE BY TREATMENT AND TIME.....	25
TABLE 24 ANALYSIS OF DOSE-MORTALITY TREND FOR FEMALE MICE.....	27
TABLE 25 ANALYSIS OF DOSE-MORTALITY TREND FOR FEMALE MICE EXCLUDING 1000 MG/KG DOSE GROUP .....	28
TABLE 26 REPORT OF P-VALUES < 0.05 FOR TEST OF POSITIVE LINEAR DOSE-TUMOR TRENDS IN FEMALE MICE: INCLUDING 1000 MG/KG DOSE GROUP.....	28
TABLE 27 DECISION RULE FOR STATISTICAL SIGNIFICANCE FOR FEMALE MICE.....	28
TABLE 21 PAIRWISE COMPARISONS OF STATISTICAL SIGNIFICANT TUMOR TRENDS IN MALE MICE.....	29
TABLE 29 STATISTICALLY SIGNIFICANT POSITIVE LINEAR DOSE-TUMOR TREND FOUND IN FEMALE MICE.....	29
TABLE 30 SUMMARY OF FINDINGS.....	31
TABLE 31 NUMBER OF TUMOR-BEARING MALE RATS.....	40
TABLE 32 NUMBER OF TUMOR-BEARING FEMALE RATS.....	42
TABLE 33 NUMBER OF TUMOR-BEARING MALE MICE.....	44
TABLE 34 NUMBER OF TUMOR-BEARING FEMALE MICE.....	45

List of Figures

FIGURE 1 NUMBER OF MALE RATS DIED DURING STUDY BY TIME.....	11
FIGURE 2 CUMULATIVE PCT. OF DEATH IN MALE RATS.....	11
FIGURE 3 KAPLAN-MEIER SURVIVAL FUNCTIONS FOR MALE RATS.....	12
FIGURE 4 NUMBER OF FEMALE RATS DIED DURING STUDY BY TIME.....	16
FIGURE 5 CUMULATIVE PCT. OF DEATH IN FEMALE RATS.....	16
FIGURE 6 KAPLAN-MEIER SURVIVAL FUNCTIONS FOR FEMALE RATS.....	17
FIGURE 7 NUMBER OF MALE MICE DIED DURING STUDY BY TIME.....	21
FIGURE 8 CUMULATIVE PCT. OF DEATH IN MALE MICE.....	21
FIGURE 9 KAPLAN-MEIER SURVIVAL FUNCTIONS FOR MALE MICE.....	22
FIGURE 10 NUMBER OF FEMALE MICE DIED DURING STUDY BY TIME.....	26
FIGURE 11 CUMULATIVE PCT. OF DEATH IN FEMALE MICE.....	26
FIGURE 12 KAPLAN-MEIER SURVIVAL FUNCTIONS FOR FEMALE MICE.....	27

Appears This Way  
On Original

## EXECUTIVE SUMMARY

The evaluations of Rat Study M-11-00-561 and Mouse Study M-11-00-560 for carcinogenic potential found the following tumor types that indicate a statistically significant dose-tumor positive linear trend.

A positive linear trend was found to be statistically significant for the tumor of ADENOMA AND CARCINOMA, HEPATOCELLULAR in liver in both sexes of rats and mice. Combining these tumors (adenoma hepatocellular and carcinoma hepatocellular) in liver yields statistically significant linear trend in both sexes in the rat as well as in the mouse study.

Furthermore, there is also a significant positive linear trend found for the tumor of HEPATOBLASTOMA in the liver for male mice, and LEYDIG CELL TUMOR in the testis and ADENOMA in the parathyroid gland in male rats. Additional analysis in the mouse study by combining hepatocellular adenoma, hepatocellular carcinoma, and hepatoblastoma in the liver also yield significant linear trend in both sexes.

Although the sponsor concluded in their report of the mouse study that there is an increased incidence of adenoma in the Harderian gland in all male treated groups and females in the 100 mg/kg and higher groups, there is no evidence of positive significant linear trend either by Peto's trend test or Exact Permutation trend test according to my results.

**Table 1 Summary Findings**

		Organ Name	Tumor Name	P-Value
Rat	Male	LIVER	Adenoma, hepatocellular	0.0000
		LIVER	Carcinoma, Hepatocellular	0.0000
		PARATHYROID	Adenoma	0.0218
	Female	TESTIS	Leydig Cell Tumor	0.0000
		LIVER	Adenoma, Hepatocellular	0.0000
		LIVER	Carcinoma, Hepatocellular	0.0000
Mice	Male	Liver	Adenoma, Hepatocellular	0.0000
		Liver	Carcinoma, Hepatocellular	0.0000
		Liver	Hepatoblastoma	0.0001
	Female	Liver	Adenoma, Hepatocellular	0.0000
		Liver	Carcinoma, Hepatocellular	0.0000

Appears This Way  
On Original

## OVERVIEW OF CARCINOGENICITY STUDIES

The objective of this review is to evaluate the carcinogenicity studies in rats and mice by Takeda Global for the carcinogenic potential of [ ] when it was given orally to male and female rats and mice. Here is the dosage assignment among the animals :

**Table 2 Dosage assignment**

Species/Study	Sex	No. Animals	Dosage (mg/kg)
Rat/Study No. M-11-00-561	Male	360 (control = 120)	0, 15, 60, 250, 1000
	Female	360 (control = 120)	0, 15, 60, 250, 1000
Mouse/Study No. M-11-00-560	Male	275	0, 30, 100, 300, 1000
	Female	275	0, 30, 100, 300, 1000

There were two vehicle control groups in each sex group, both containing Methylcellulose in the rat study. In contrast, there was only a single vehicle control group in each sex group in the mice study, containing Methylcellulose as well. The terminal sacrifice started during Week 105 for all animals.

Appears This Way  
On Original

## SPONSOR'S EVALUATION OF CARCINOGENICITY STUDIES

Statistical analyses conducted by the sponsor were done according to the following procedures:

1. **Survival Data**  
Survival curves were estimated and presented as life-tables and graphically using Kaplan-Meier's method, and the trend of the survival rate to dose level was analyzed using Tarone-type method. Log-rank test was used to compare the difference between the control group and each dose group. The lowest level of significance was 5% in two-tailed level.
2. **Body Weight, Food Consumption and Hematological Data**  
The data were first tested using Bartlett's test for homogeneity of variance, followed by William's test to determine the non-toxic dose level. If no significant differences with William's test, Dunnett's test was performed to compare the mean in the control group with that in each dose group. When variances were heterogeneous, the Shirley-Williams test was performed to determine the non-toxic dose level. If no significant differences with Shirley-Williams' test, Steel's test was performed to compare the mean in the control group with that in each dose group. The Bartlett's, the Williams' and the Shirley-Williams' tests were conducted at the two-tailed significance level of 0.05.
3. **Tumor Incidence Data**  
Tumors that occurred with a frequency of more than 5% in either the control or any dose group were assessed by Peto's method to assess dose-dependency for all groups and to compare the incidence between the control group and each dose group (level of significance: 5% one-tailed level). In addition, Fisher's exact test was applied to assess the difference of tumor incidence between the control group and each dose group (level of significance: 5% one-tailed level).

### Summary of Sponsor's Results:

1. **Rats**
  - a. **Survival Rate**  
Lower survival rate was observed in females in the 1000 mg/kg group and it was thought to be attributable to high mortality caused by intoxication and test article-induced liver tumor. On the other hand, higher survival rate was observed in males in the 250 mg/kg group and it was thought to be attributable to suppressed body weight.
  - b. **Tumor Incidence**  
The incidence of hepatocellular adenoma was increased in males in all treated groups and females in the 60 mg/kg and higher groups. The incidence of hepatocellular carcinoma was also increased in both sexes in the 1000 mg/kg group. The increase in the incidence of hepatocellular adenoma noted in males in the 15 and 60 mg/kg groups was marginal and there was no statistical significance when the incidence was compared to that in one (control II) of the control groups. In the testis, increased incidence of Leydig cell tumor was observed in the 250 and 1000 mg/kg groups. The increase in the incidence of Leydig cell tumor was observed in the 250 mg/kg group was marginal and there was no statistical significance when the incidence was compared to that in one (control I) of the control groups.

2. Mice

a. Survival Rate

Lower survival rate was observed in both sexes in the 1000 mg/kg group and it was thought to be attributable to high mortality caused by test article-induced liver tumor.

b. Tumor Incidence

The incidence of adenoma in the Harderian gland was increased in males in all treated groups and females in the 100 mg/kg and higher groups. The incidence of hepatocellular adenoma/carcinoma (including hepatoblastoma) of the liver was also increased in males in the 100 mg/kg and higher groups and females in the 300 and 1000 mg/kg groups.

Appears This Way  
On Original

## REVIEWER'S EVALUATION OF CARCINOGENICITY STUDIES

Statistical analyses conducted by this reviewer were done according to the Food and Drug Administration's Guidance for Industry: Statistical Aspects of the Design, Analysis, and Interpretation of Chronic Rodent Carcinogenicity Studies of Pharmaceuticals (May, 2001). In addition, the reviewers' analyses were primarily conducted using eReview of Animal Carcinogenicity, a review tool developed for and utilized by CDER reviewers.

### Mortality Analysis

Tests for homogeneity and dose mortality trends were conducted using survival analysis methods described by Cox (1972) and the Kruskal-Wallis Test (Gehan, 1965; Breslow, 1970; Thomas, Breslow, and Gart, 1977) where the latter test weights early failures more heavily.

### Tumor Data Analysis (Trend Test)

This reviewer conducted the trend tests on tumor incidence rates using the method described by Peto et. al. (1980) and the method of exact permutation trend test developed by the Division of Biometrics II. The sponsor classified tumors as fatal, possibly fatal, incidental, or possibly incidental, in which case, this reviewer combined fatal and possibly fatal as one group called fatal, and combined incidental and possibly incidental in another group called incidental. Data of incidental and fatal tumors were analyzed via the prevalence and death-rates methods, respectively. A combined test was used to analyze tumors classified as both fatal and incidental. The method of exact permutation trend test was used to counter underestimation of p-values when tumor incidence across the treatment group was small. All tests are performed separately for males and females for both species.

### Multiple Testing Adjustment

A rule proposed by Haseman (1983) could be used to adjust the effect of multiple testing for pairwise comparisons between the control and the high groups. A similar rule proposed by the Division of Biometrics, CDER/FDA was used in this review. For a two-species and two-sex study, the rule for testing positive trends states that in order to keep the overall false-positive rate at the nominal level of approximately ten percent, tumor types with a spontaneous tumor rate of no more than one percent should be tested at 0.025 level, otherwise the level should be set at 0.005. (Lin, 1995, 1997; Lin and Rahman, 1998a, 1998b) The ten percent overall false positive rate is seen by CDER statisticians as appropriate in a new drug regulatory setting. On the other hand, the rule for pairwise comparison tests between the control and the high groups state that tumor types with a spontaneous tumor rate of no more than one percent should be tested at 0.05 level, otherwise the level should be set at 0.01. (See Appendix)

### Evaluation of Validity of the Design of the Study

An evaluation of validity of the study design was not conducted because all studies analyzed showed at least one tumor type having a significant positive trend. Readers are referred to papers by Haseman (1984) and Chu, Cueto and Ward (1981) for further information about evaluating the validity of the study design for negative studies.

## Analysis of Male Rats

### Mortality Analysis

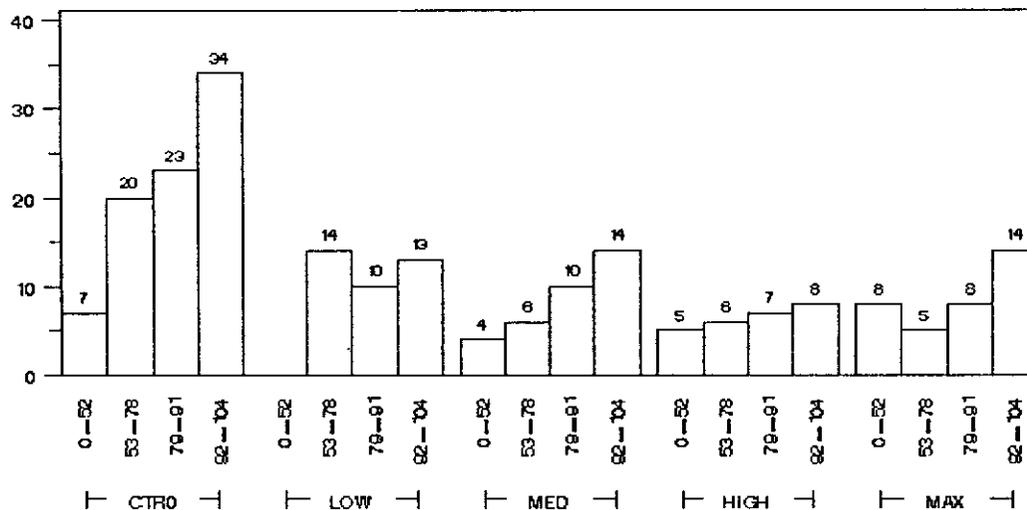
The mortality analysis starts with the display of the animal-mortality statistics by treatment and time interval. The main purpose for these analyses is to discover any statistically significant dose-mortality trend that justifies the age-adjusted test of positive dose-tumor linear trend.

**Table 3 Analysis of Mortality Data for Male Rats by Treatment and Time**

Analysis of Mortality		No. Risk	No. Died	No. Alive	Pct Survival	Pct Mortality
CTRO	0-52	120	7	113	94.2	5.8
	53-78	113	20	93	77.5	22.5
	79-91	93	23	70	58.3	41.7
	92-104	70	34	36	30.0	70.0
	FINALKILL105-105	36	36	0		
LOW	53-78	60	14	46	76.7	23.3
	79-91	46	10	36	60.0	40.0
	92-104	36	13	23	38.3	61.7
	FINALKILL105-105	23	23	0		
MED	0-52	60	4	56	93.3	6.7
	53-78	56	6	50	83.3	16.7
	79-91	50	10	40	66.7	33.3
	92-104	40	14	26	43.3	56.7
	FINALKILL105-105	26	26	0		
HIGH	0-52	60	5	55	91.7	8.3
	53-78	55	6	49	81.7	18.3
	79-91	49	7	42	70.0	30.0
	92-104	42	8	34	56.7	43.3
	FINALKILL105-105	34	34	0		
MAX	0-52	60	8	52	86.7	13.3
	53-78	52	5	47	78.3	21.7
	79-91	47	8	39	65.0	35.0
	92-104	39	14	25	41.7	58.3
	FINALKILL105-105	25	25	0		

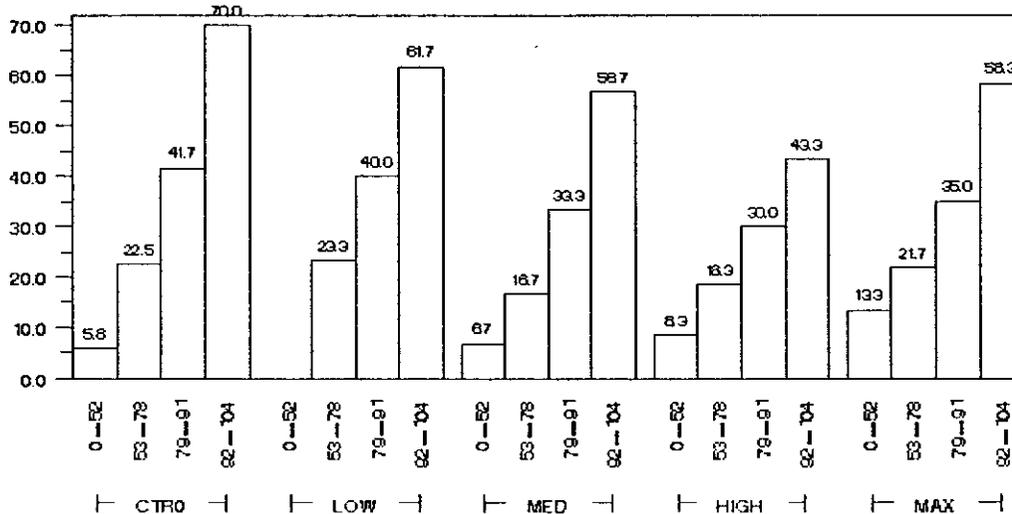
Source data: Analysis data (SAS 9.1.3) R2M21782

Figure 1 Number of Male Rats Died During Study by Time



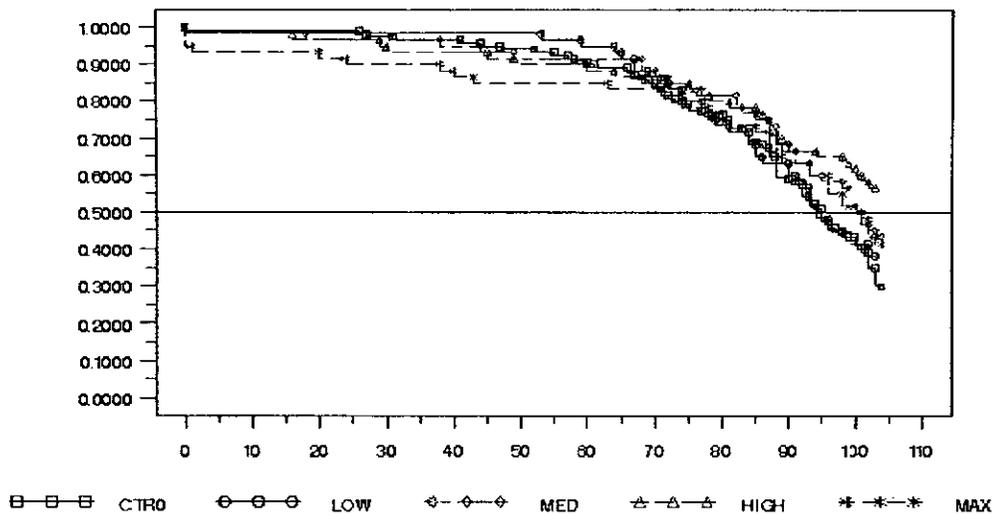
Source data: Analysis data (SAS 9.1.3) R2M21782

Figure 2 Cumulative Pct. of Death in Male Rats



Source data: Analysis data (SAS 9.1.3) R2M21782

Figure 3 Kaplan-Meier Survival Functions for Male Rats



Source data: Analysis data (SAS 9.1.3) R2M21782

The analysis of dose-mortality trend is done using a computer program described in the article "Trend and Homogeneity Analyses of Proportions and Life Table Data," Version 2.1, by Donald G. Thomas, National Cancer Institute. A significant dose-tumor trend gives rise to a statistical justification for the age-adjusted test of positive dose-tumor linear trend.

**Reviewer’s Comment on Mortality Analysis:**

When the maximum dose of 1000 mg/kg is included, the dose-mortality trend was found to be statistically not significant (Table 4). In contrast, exclusion of maximum dose of 1000 mg/kg yields a statistically significant dose- mortality trend (Table 5). The difference could be attributed to the higher survival rate in the 250 mg/kg group and lower survival rate in the 1000 mg/kg group.

Table 4 Analysis of Dose-Mortality Trend for Male Rats

	Method			
	Cox		Kruskal-Wallis	
	Statistics	P-Value	Statistics	P-Value
Time-Adjusted Trend Test	8.7087	0.0334	6.4792	0.0905
Depart from Trend				
Dose-Mortality Trend	0.8930	0.3447	0.4560	0.4995
Homogeneity	9.6017	0.0477	6.9352	0.1394

Source data: Analysis data (SAS 9.1.3) R2M21782

**Table 5 Analysis of Dose-Mortality Trend for Male Rats (excluding 1000 mg/kg dose group)**

	Method			
	Cox		Kruskal-Wallis	
	Statistics	P-Value	Statistics	P-Value
Time-Adjusted Trend Test	1.3422	0.5111	1.1901	0.5515
Depart from Trend				
Dose-Mortality Trend	8.1331	0.0043	5.7265	0.0167
Homogeneity	9.4753	0.0236	6.9166	0.0746

**Trend Analysis**

The test for positive dose-tumor linear trend is the ultimate objective of the evaluation of the carcinogenicity-study. Occasionally, pairwise comparisons are employed, but only under certain condition of the data and are decided on a case-by-case basis. As a cautionary note, blindly impose pairwise comparisons can only undermine the importance of the trend test, inflate the type-1 error, and produce untrustworthy results. The significance of the test is decided based on a decision rule adopted by the Office of Biostatistics. The details of the decision rule can be found in the Appendix of this review.

We only report trend-test results with p-value less than 0.05, which may not imply a statistical significance. Throughout this report, this icon indicates a statistically significant trend:



**Table 6 Report of P-values < 0.05 for Test of Positive Linear Dose-Tumor Trends in Male Rats: including 1000 mg/kg dose group**

Organ Name	Tumor Name	CTR0	LOW	MED	HIGH	MAX	P-Value (Monte Carlo)	P-Value (Asymptotic Method)
Cerebrum	GRANULAR CELL TUMOR	0	0	0	0	1	0.1675	0.0178
Liver	ADENOMA,HEPATOCELLULAR	3	7	7	15	47	0.0000 (🚨)	0.0000
Liver	CARCINOMA,HEPATOCELLULAR	0	0	0	0	38	0.0000 (🚨)	0.0000
Liver	HEMANGIOSARCOMA	0	0	0	0	1	0.1675	0.0178
Liver	HEPATOBLASTOMA	0	0	0	0	1	0.0974	0.0017
Parathyroid	ADENOMA	0	0	0	1	2	0.0218 (🚨)	0.0071
Pituitary	ADENOMA,INTERMEDIATED	0	0	0	0	1	0.0974	0.0017
Skin	MELANOMA,BENIGN,AMELANOTIC	0	0	0	0	1	0.1675	0.0178
Stomach	LEIOMYOMA	0	0	0	0	1	0.1525	0.0120
Stomach	LEIOMYOSARCOMA	0	0	0	0	1	0.1525	0.0120
Testis	LEYDIG CELL TUMOR	3	1	0	7	46	0.0000 (🚨)	0.0000
Urinary bladder	CARCINOMA,TRANSITIONAL CELL	0	0	0	0	2	0.0298	0.0014

Source data: Analysis data (SAS 9.1.3) R2M 21782

**Table 7 Pairwise Comparisons of Statistical Significant Tumor Trends in Male Rats**

Organ Name	Tumor Name	CTR0	LOW	MED	HIGH	MAX	P-Value (EXACT)	P-Value (Asymptotic Method)
Liver	ADENOMA,HEPATOCELLULAR	3	7				0.0164	0.0060
Liver	ADENOMA,HEPATOCELLULAR	3		7			0.0277	0.0114
Liver	ADENOMA,HEPATOCELLULAR	3			15		0.0000	0.0000
Liver	ADENOMA,HEPATOCELLULAR	3				47	0.0000	0.0000
Liver	CARCINOMA,HEPATOCELLULAR	0				38	0.0000	0.0000
Testis	LEYDIG CELL TUMOR	3			7		0.0423	0.0192
Testis	LEYDIG CELL TUMOR	3				46	0.0000	0.0000

**Reviewer's Statistical Findings from the Trend-Test:**

Based on the Office of Biostatistics rules, the positive linear trend of dose-tumor is statistically significant in male rats for hepatocellular adenoma and carcinoma in the liver, adenoma in the parathyroid, and Leydig cell tumor in the testis.

Although hepatocellular carcinoma in liver appeared to show a positive trend, this finding was driven by the high tumor incidence in the 1000 mg/kg (38/60) compared to none in the other control and treated groups. On the other hand, although the trend for the incidence of adenoma in the parathyroid is statistical significant, the incidences were very low to warrant further exploration.

Pairwise comparison between the treated groups and pooled control group suggests that the minimum effective dose for hepatocellular adenoma in the liver in male rats is 250 mg/kg, 1000 mg/kg for Leydig cell tumor in the testis, and 1000 mg/kg for hepatocellular carcinoma in the liver.

**Table 8 Statistically significant positive linear dose-tumor trend found in Male Rats**

Organ Name	Tumor Name	P-Value
LIVER	ADENOMA, HEPATOCELLULAR	0.0000
LIVER	CARCINOMA, HEPATOCELLULAR	0.0000
PARATHYROID	ADENOMA	0.0218
TESTIS	LEYDIG CELL TUMOR	0.0000

## Analysis of Female Rats

### Mortality Analysis

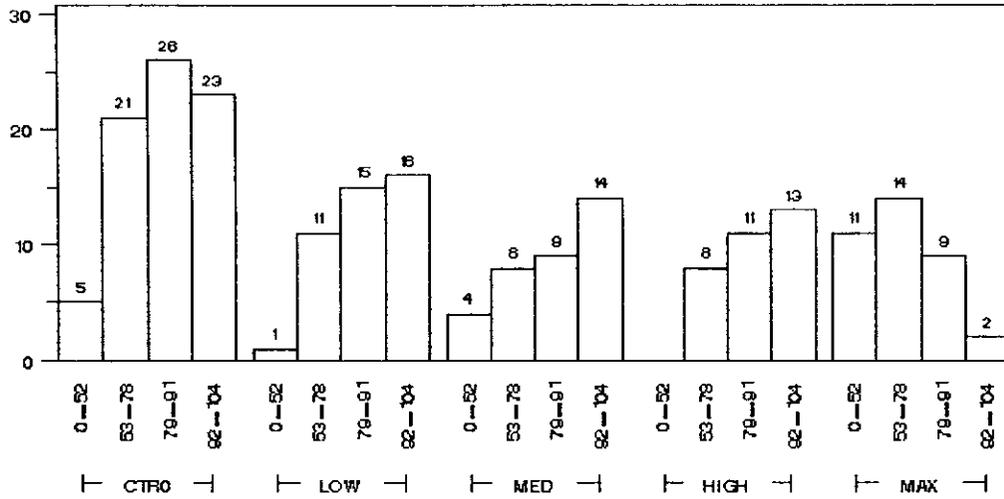
The mortality analysis starts with the display of the animal-mortality statistics by treatment and time interval. The main purpose for these analyses is to discover any statistically significant dose-mortality trend that justifies the age-adjusted test of positive dose-tumor linear trend.

**Table 9 Analysis of Mortality Data for Female Rats by Treatment and Time**

Analysis of Mortality		No. Risk	No. Died	No. Alive	Pct Survival	Pct Mortality
CTRO	0-52	120	5	115	95.8	4.2
	53-78	115	21	94	78.3	21.7
	79-91	94	26	68	56.7	43.3
	92-104	68	23	45	37.5	62.5
	FINALKILL105-105	45	45	0		
LOW	0-52	60	1	59	98.3	1.7
	53-78	59	11	48	80.0	20.0
	79-91	48	15	33	55.0	45.0
	92-104	33	16	17	28.3	71.7
	FINALKILL105-105	17	17	0		
MED	0-52	60	4	56	93.3	6.7
	53-78	56	8	48	80.0	20.0
	79-91	48	9	39	65.0	35.0
	92-104	39	14	25	41.7	58.3
	FINALKILL105-105	25	25	0		
HIGH	53-78	60	8	52	86.7	13.3
	79-91	52	11	41	68.3	31.7
	92-104	41	13	28	46.7	53.3
	FINALKILL105-105	28	28	0		
MAX	0-52	60	11	49	81.7	18.3
	53-78	49	14	35	58.3	41.7
	79-91	35	9	26	43.3	56.7
	92-104	26	2	24	40.0	60.0
	INTERIM KILL		24			

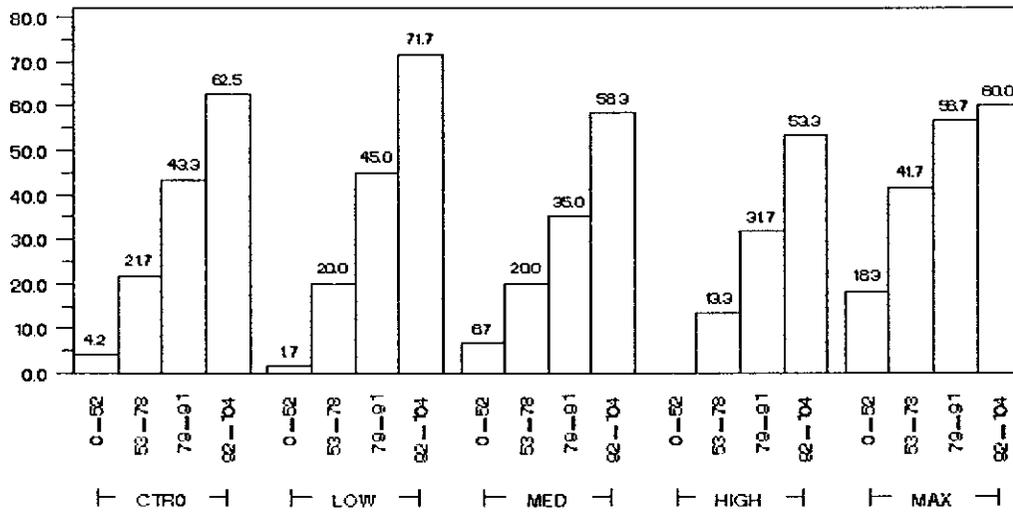
Source data: Analysis data (SAS 9.1.3) R2F21782

Figure 4 Number of Female Rats Died During Study by Time



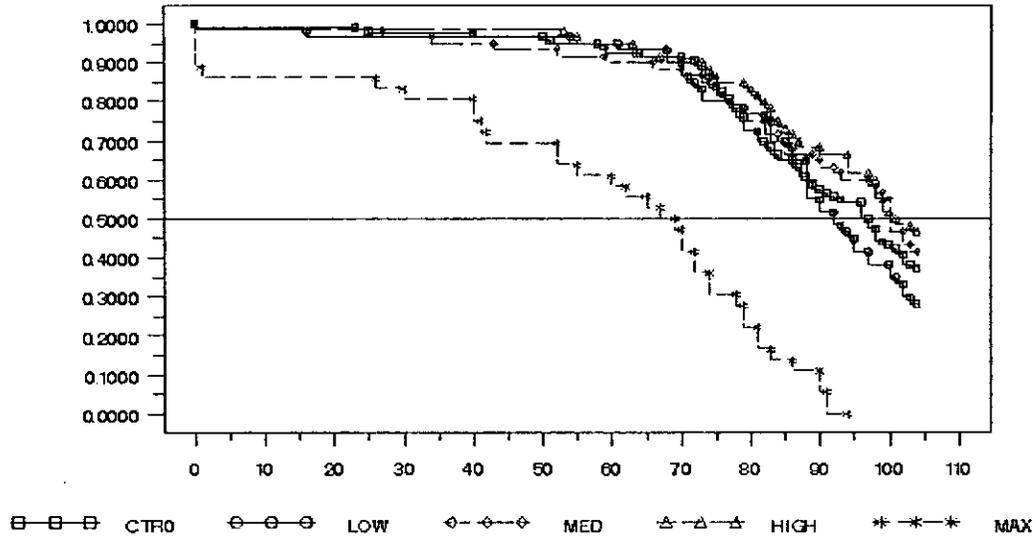
Source data: Analysis data (SAS 9.1.3) R2F21782

Figure 5 Cumulative Pct. of Death in Female Rats



Source data: Analysis data (SAS 9.1.3) R2F21782

Figure 6 Kaplan-Meier Survival Functions for Female Rats



Source data: Analysis data (SAS 9.1.3) R2F21782

The analysis of dose-mortality trend is done using a computer program described in the article "Trend and Homogeneity Analyses of Proportions and Life Table Data," Version 2.1, by Donald G. Thomas, National Cancer Institute. A significant dose-tumor trend gives rise to a statistical justification for the age-adjusted test of positive dose-tumor linear trend.

**Reviewer's Comment on Mortality Analysis:**

When the maximum dose of 1000 mg/kg is included, the dose-mortality trend was found to be significant statistically (Table 10). However, exclusion of maximum dose of 1000 mg/kg yields a non-statistically significant dose-mortality trend (Table 11). The difference could be attributed to the lower survival rate in the 15 mg/kg group compared to other treated groups.

Table 10 Analysis of Dose-Mortality Trend for Female Rats

	Method			
	Cox		Kruskal-Wallis	
	Statistics	P-Value	Statistics	P-Value
Time-Adjusted Trend Test				
Depart from Trend	42.0823	0.0000	32.1680	0.0000
Dose-Mortality Trend	68.7128	0.0000	74.9289	0.0000
Homogeneity	110.7950	0.0000	107.0969	0.0000

Source data: Analysis data (SAS 9.1.3) R2F21782

**Table 11 Analysis of Dose-Mortality Trend for Female Rats (excluding 1000 mg/kg dose group)**

	Method			
	Cox		Kruskal-Wallis	
	Statistics	P-Value	Statistics	P-Value
Time-Adjusted Trend Test	1.6720	0.4334	0.9546	0.6205
Depart from Trend				
Dose-Mortality Trend	2.7299	0.0985	2.6540	0.1033
Homogeneity	4.4019	0.2212	3.6086	0.3069

**Trend Analysis**

The test for positive dose-tumor linear trend is the ultimate objective of the evaluation of the carcinogenicity-study. Occasionally, pairwise comparisons are employed, but only under certain condition of the data and are decided on a case-by-case basis. As a cautionary note, blindly impose pairwise comparisons can only undermine the importance of the trend test, inflate the type-1 error, and produce untrustworthy results. The significance of the test is decided based on a decision rule adopted by the Office of Biostatistics. The details of the decision rule can be found in the Appendix of this review.

We only report trend-test results with p-value less than 0.05, which may not imply a statistical significance. Throughout this report, this icon indicates a statistically significant trend:



**Table 12 Report of  $\chi^2$ -values < 0.05 for Test of Positive Linear Dose-Tumor Trends in Female Rats: including 1000 mg/kg group in the analysis**

Organ Name	Tumor Name	CTR0	LOW	MED	HIGH	MAX	P-Value (Exact Method)	P-Value (Asymptotic Method)
Cerebellum	GRANULAR CELL TUMOR	0	0	0	0	1	0.1286	0.0060
Kidney	ADENOMA,RENAL CELL	0	0	0	2	0	0.0577	0.0074
Large intestine,rectum	GRANULAR CELL TUMOR	0	0	0	0	1	0.2034	0.0262
Liver	ADENOMA,HEPATOCELLULAR	2	0	8	22	49	0.0000 (⚠)	0.0000
Liver	CHOLANGIOMA	0	0	0	1	0	0.2435	0.0433
Liver	HEMANGIOMA	0	0	0	1	1	0.2435	0.0433
Liver	CARCINOMA,HEPATOCELLULAR	0	0	1	1	20	0.0000 (⚠)	0.0000
Ovary	GRANULOSA CELL TUMOR	0	0	0	1	0	0.2435	0.0433
Pancreas	HEMANGIOSARCOMA	0	0	0	0	1	0.0294	0.0000
Spinal cord	GRANULAR CELL TUMOR	0	0	0	1	0	0.2435	0.0433
Small intestine,jejunum	LEIOMYOSARCOMA	0	0	0	1	0	0.2435	0.0433
Uterus	ADENOMA	0	0	0	1	0	0.2435	0.0433
Uterus	GRANULAR CELL TUMOR	3	1	1	3	1	0.0706	0.0442
Uterus	ADENOCARCINOMA	0	0	0	0	1	0.1286	0.0060

Source data: Analysis data (SAS 9.1.3) R2F21782

**Table 13 Decision Rule for Statistical Significance for Female Rats**

Organ name	Tumor name	Overall tumor type	Tumor rate as PCT. in control group	Suggested Interpretation for trend-test
Liver	ADENOMA, HEPATOCELLULAR	Incidental	1.57	Use exact p-value. Use p-value cutoff point of 0.005.
Liver	CARCINOMA, HEPATOCELLULAR	Incidental	0.00	Use exact p-value. Use p-value cutoff point of 0.025.

Source data: Analysis data (SAS 9.1.3) R2F21782

**Table 14 Pairwise Comparisons of Statistical Significant Tumor Trends in Female Rats**

Organ Name	Tumor Name	CTR0	LOW	MED	HIGH	MAX	P-Value (EXACT)	P-Value (Asymptotic Method)
Liver	ADENOMA, HEPATOCELLULAR	2		8			0.0036	0.0010
Liver	ADENOMA, HEPATOCELLULAR	2			22		0.0000	0.0000
Liver	ADENOMA, HEPATOCELLULAR	2				49	0.0000	0.0000
Liver	CARCINOMA, HEPATOCELLULAR	0				20	0.0000	0.0000

**Reviewer's Statistical Findings from the Trend-Test:**

Using Fisher's Exact Test procedure, the positive linear trend of dose-tumor is statistically significant in female rats for hepatocellular adenoma and carcinoma in the liver.

Similar to male rats, although hepatocellular carcinoma in liver appeared to show a positive trend, this finding was driven by the high tumor incidence in the 1000 mg/kg (20/60) compared to single incidence in the 60 mg/kg and 250 mg/kg groups.

Pairwise comparison between the treated groups and pooled control group suggests that the minimum effective dose for hepatocellular adenoma in the liver in female rats is 60 mg/kg, and 1000 mg/kg for hepatocellular carcinoma in the liver.

**Table 15 Statistically significant positive linear dose-tumor trend found in Female Rats**

Organ Name	Tumor Name	P-Value
LIVER	ADENOMA, HEPATOCELLULAR	0.0000
LIVER	CARCINOMA, HEPATOCELLULAR	0.0000

*Appears This Way  
On Original*

## Analysis of Male Mice

### Mortality Analysis

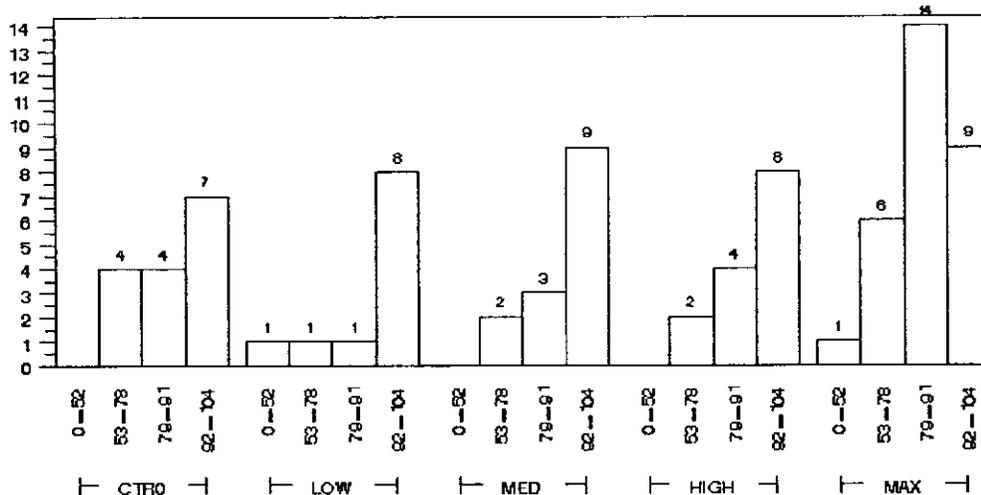
The mortality analysis starts with the display of the animal-mortality statistics by treatment and time interval. The main purpose for these analyses is to discover any statistically significant dose-mortality trend that justifies the age-adjusted test of positive dose-tumor linear trend.

**Table 16 Analysis of Mortality Data for Male Mice by Treatment and Time**

Analysis of Mortality		No. Risk	No. Died	No. Alive	Pct Survival	Pct Mortality
CTRO	53-78	55	4	51	92.7	7.3
	79-91	51	4	47	85.5	14.5
	92-104	47	7	40	72.7	27.3
	FINALKILL105-105	40	40	0		
LOW	0-52	55	1	54	98.2	1.8
	53-78	54	1	53	96.4	3.6
	79-91	53	1	52	94.5	5.5
	92-104	52	8	44	80.0	20.0
	FINALKILL105-105	44	44	0		
MED	53-78	55	2	53	96.4	3.6
	79-91	53	3	50	90.9	9.1
	92-104	50	9	41	74.5	25.5
	FINALKILL105-105	41	41	0		
HIGH	53-78	55	2	53	96.4	3.6
	79-91	53	4	49	89.1	10.9
	92-104	49	8	41	74.5	25.5
	FINALKILL105-105	41	41	0		
MAX	0-52	55	1	54	98.2	1.8
	53-78	54	6	48	87.3	12.7
	79-91	48	14	34	61.8	38.2
	92-104	34	9	25	45.5	54.5
	INTERIM KILL		25			

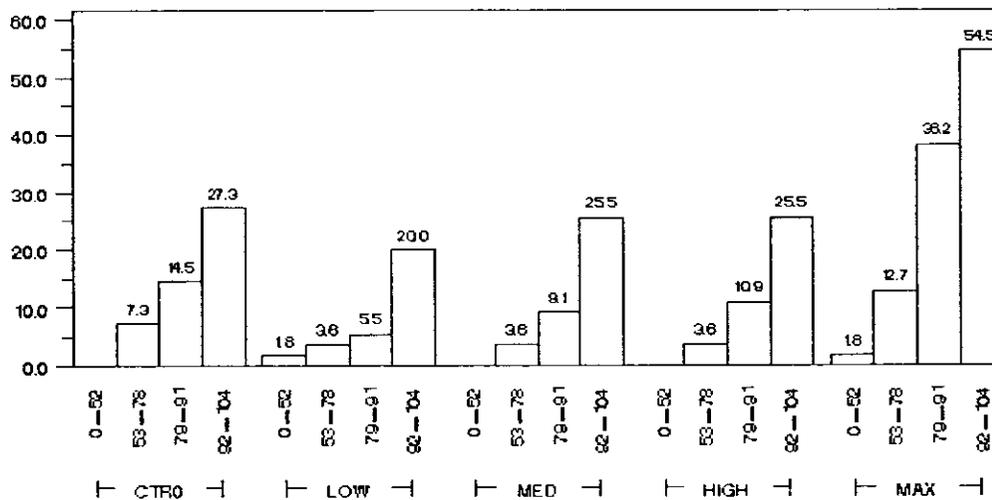
Source data: Analysis data (SAS 9.1.3) M2M21782

Figure 7 Number of Male Mice Died During Study by Time



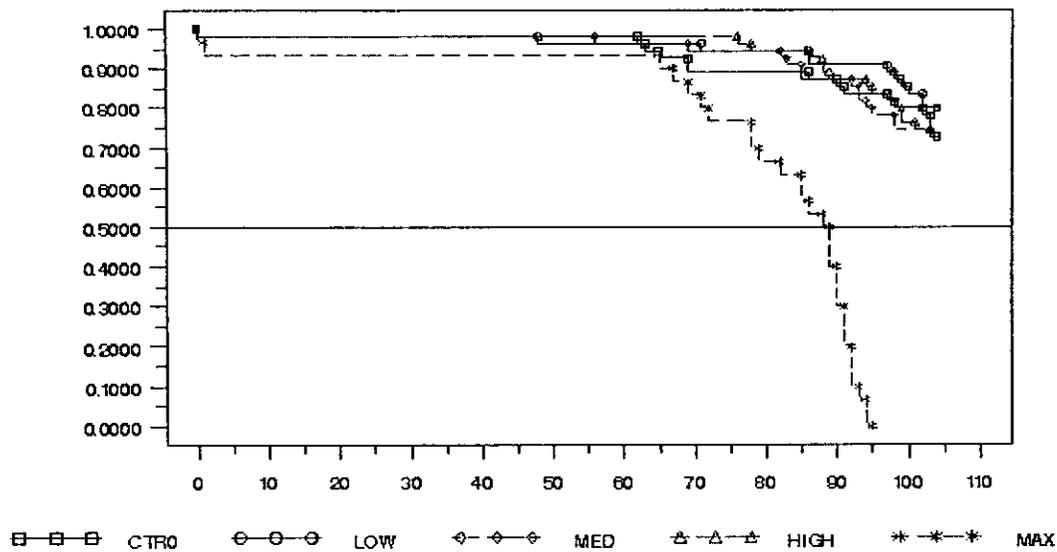
Source data: Analysis data (SAS 9.1.3) M2M21782

Figure 8 Cumulative Pct. of Death in Male Mice



Source data: Analysis data (SAS 9.1.3) M2M21782

Figure 9 Kaplan-Meier Survival Functions for Male Mice



Source data: Analysis data (SAS 9.1.3) M2M21782

The analysis of dose-mortality trend is done using a computer program described in the article "Trend and Homogeneity Analyses of Proportions and Life Table Data," Version 2.1, by Donald G. Thomas, National Cancer Institute. A significant dose-tumor trend gives rise to a statistical justification for the age-adjusted test of positive dose-tumor linear trend.

**Reviewer's Comment on Mortality Analysis:**

When the maximum dose of 1000 mg/kg is included, the dose-mortality trend was found to be significant statistically (Table 17). However, exclusion of maximum dose of 1000 mg/kg yields a non-statistically significant dose-mortality trend (Table 18). The difference could be attributed to the higher survival rate in the 30 mg/kg group and lower survival rate in the 1000 mg/kg group.

**Table 17 Analysis of Dose-Mortality Trend for Male Mice**

	Method			
	Cox		Kruskal-Wallis	
	Statistics	P-Value	Statistics	P-Value
Time-Adjusted Trend Test				
Depart from Trend	41.2398	0.0000	32.8928	0.0000
Dose-Mortality Trend	125.9600	0.0000	118.3775	0.0000
Homogeneity	167.1998	0.0000	151.2704	0.0000

Source data: Analysis data (SAS 9.1.3) M2M21782

**Table 18 Analysis of Dose-Mortality Trend for Male Mice (excluding 1000 mg/kg dose group)**

	Method			
	Cox		Kruskal-Wallis	
	Statistics	P-Value	Statistics	P-Value
Time-Adjusted Trend Test	0.9655	0.6171	1.1114	0.5737
Depart from Trend				
Dose-Mortality Trend	0.0374	0.8466	0.0383	0.8448
Homogeneity	1.0030	0.8005	1.1498	0.7651

**Trend Analysis**

The test for positive dose-tumor linear trend is the ultimate objective of the evaluation of the carcinogenicity-study. Occasionally, pairwise comparisons are employed, but only under certain condition of the data and are decided on a case-by-case basis. As a cautionary note, blindly impose pairwise comparisons can only undermine the importance of the trend test, inflate the type-1 error, and produce untrustworthy results.

The significance of the test is decided based on a decision rule adopted by the Office of Biostatistics. The details of the decision rule can be found in the Appendix of this review.

We only report trend-test results with p-value less than 0.05, which may not imply a statistical significance. Throughout this report, this icon indicates a statistically significant trend:



**Table 19 Report of  $\chi^2$ -values < 0.05 for Test of Positive Linear Dose-Tumor Trends in Male Mice: including 1000 mg/kg dose group**

Organ Name	Tumor Name	CTR0	LOW	MED	HIGH	MAX	P-Value (Exact Method)	P-Value (Asymptotic Method)
Aorta, thoracic	HEMANGIOSARCOMA	0	0	0	1	0	0.2470	0.0488
Harderian gland	ADENOCARCINOMA	1	1	1	2	1	0.0562	0.0266
Kidney	ADENOMA, RENAL CELL	0	0	0	1	0	0.2470	0.0488
Kidney	CARCINOMA, RENAL CELL	0	0	0	1	0	0.2470	0.0488
Kidney	HEMANGIOSARCOMA	0	0	0	1	0	0.2470	0.0488
Liver	ADENOMA, HEPATOCELLULAR	20	24	32	50	54	0.0000	0.0000 (!)
Liver	CARCINOMA, HEPATOCELLULAR	9	6	18	26	50	0.0000	0.0000 (!)
Liver	HEPATOBLASTOMA	0	0	1	3	10	0.0001	0.0003 (!)
Lung(bronchus)	ADENOMA, BRONCHIOLO-ALVEOLAR	7	6	10	15	8	0.0483	0.0437
Small intestine, duodenum	PAPILLOMA	0	0	0	1	0	0.2470	0.0488

Source data: Analysis data (SAS 9.1.3) M2M21782

**Table 20 Decision Rule for Statistical Significance for Male Mice**

Organ name	Tumor name	Overall tumor type	Tumor rate as PCT in control group	Suggested interpretation for trend-test
Liver	ADENOMA, HEPATOCELLULAR	Both fatal and incidental	36.36	Use asymptotic p-value. Use p-value cutoff point of 0.005.
Liver	CARCINOMA, HEPATOCELLULAR	Both fatal and incidental	16.36	Use asymptotic p-value. Use p-value cutoff point of 0.005.
Liver	HEPATOBLASTOMA	Both fatal and incidental	0.00	Use asymptotic p-value. Use p-value cutoff point of 0.025.

Source data: Analysis data (SAS 9.1.3) M2M21782

**Table 21 Pairwise Comparisons of Statistical Significant Tumor Trends in Male Mice**

Organ Name	Tumor Name	CTRL	LOW	MED	HIGH	MAX	P-Value (EXACT)	P-Value (Asymptotic Method)
Liver	ADENOMA, HEPATOCELLULAR	20		32			0.0266	0.0176
Liver	ADENOMA, HEPATOCELLULAR	20			50		0.0000	0.0000
Liver	ADENOMA, HEPATOCELLULAR	20				54	0.0000	0.0000
Liver	CARCINOMA, HEPATOCELLULAR	9		18			0.0397	0.0256
Liver	CARCINOMA, HEPATOCELLULAR	9			26		0.0009	0.0006
Liver	CARCINOMA, HEPATOCELLULAR	9				50	0.0000	0.0000
Liver	HEPATOBLASTOMA	0			3		0.1249	0.0418
Liver	HEPATOBLASTOMA	0				10	0.1215	0.0469

**Reviewer's Statistical Findings from the Trend-Test:**

Using Fisher's Exact Test procedure, the positive linear trend of dose-tumor is statistically significant in male mice for hepatocellular adenoma/carcinoma, and hepatoblastoma in the liver.

Pairwise comparison between the treated groups and the control group suggests that the minimum effective dose for hepatocellular adenoma and hepatocellular carcinoma in the liver in male mice is 300 mg/kg, and 300 mg/kg for hepatoblastoma in the liver using the asymptotic method.

**Table 22 Statistically significant positive linear dose-tumor trend found in Male MICE**

Organ Name	Tumor Name	P-Value
LIVER	ADENOMA, HEPATOCELLULAR	0.0000
LIVER	CARCINOMA, HEPATOCELLULAR	0.0000
LIVER	HEPATOBLASTOMA	0.0001

## Analysis of Female Mice

### Mortality Analysis

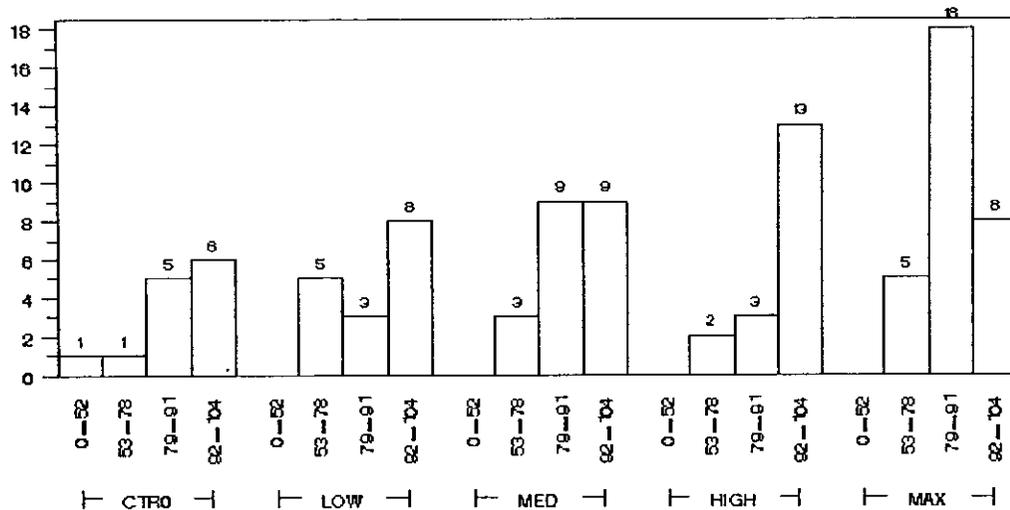
The mortality analysis starts with the display of the animal-mortality statistics by treatment and time interval. The main purpose for these analyses is to discover any statistically significant dose-mortality trend that justifies the age-adjusted test of positive dose-tumor linear trend.

**Table 23 Analysis of Mortality Data for Female Mice by Treatment and Time**

Analysis of Mortality		No. Risk	No. Died	No. Alive	Pct Survival	Pct Mortality
CTRO	0-52	55	1	54	98.2	1.8
	53-78	54	1	53	96.4	3.6
	79-91	53	5	48	87.3	12.7
	92-104	48	6	42	76.4	23.6
	FINALKILL 105-105	42	42	0		
LOW	53-78	55	5	50	90.9	9.1
	79-91	50	3	47	85.5	14.5
	92-104	47	8	39	70.9	29.1
	FINALKILL 105-105	39	39	0		
MED	53-78	55	3	52	94.5	5.5
	79-91	52	9	43	78.2	21.8
	92-104	43	9	34	61.8	38.2
	FINALKILL 105-105	34	34	0		
HIGH	53-78	55	2	53	96.4	3.6
	79-91	53	3	50	90.9	9.1
	92-104	50	13	37	67.3	32.7
	FINALKILL 105-105	37	37	0		
MAX	53-78	55	5	50	90.9	9.1
	79-91	50	18	32	58.2	41.8
	92-104	32	8	24	43.6	56.4
	INTERIM KILL		24			

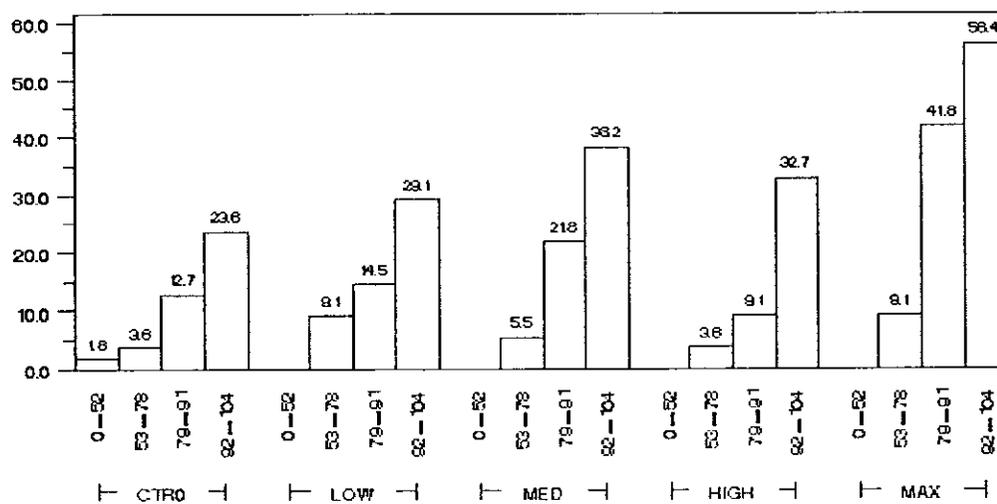
Source data: Analysis data (SAS 9.1.3) M2F21782

Figure 10 Number of Female Mice Died During Study by Time



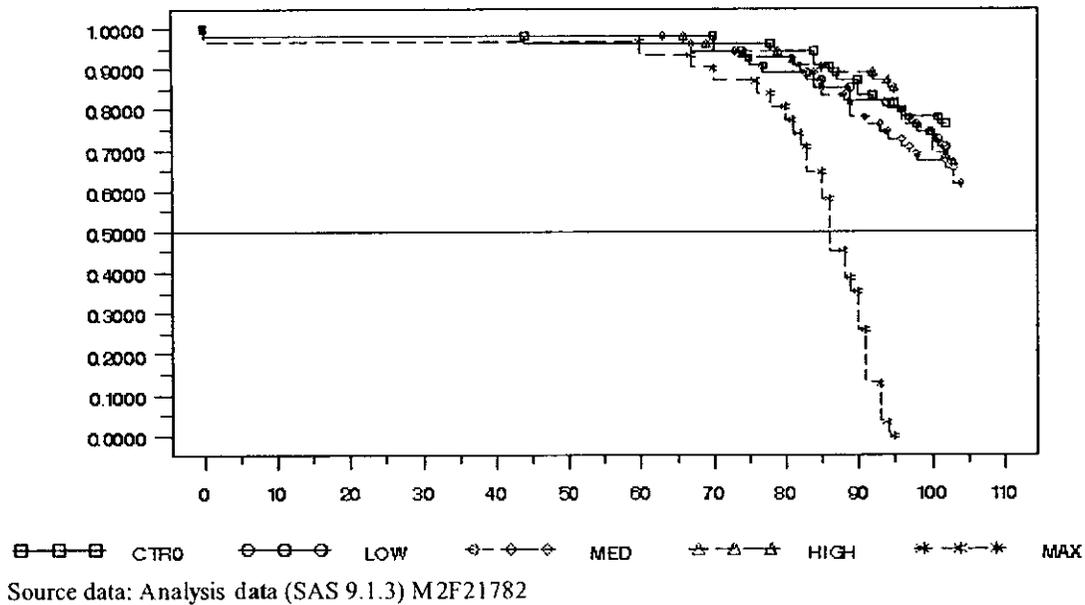
Source data: Analysis data (SAS 9.1.3) M2F21782

Figure 11 Cumulative Pct. of Death in Female Mice



Source data: Analysis data (SAS 9.1.3) M2F21782

**Figure 12 Kaplan-Meier Survival Functions for Female Mice**



The analysis of dose-mortality trend is done using a computer program described in the article "Trend and Homogeneity Analyses of Proportions and Life Table Data," Version 2.1, by Donald G. Thomas, National Cancer Institute. A significant dose-tumor trend gives rise to a statistical justification for the age-adjusted test of positive dose-tumor linear trend.

**Reviewer's Comment on Mortality Analysis:**

When the maximum dose of 1000 mg/kg is included, the dose-mortality trend was found to be significant statistically (Table 24). However, exclusion of maximum dose of 1000 mg/kg yields a non-statistically significant dose-mortality trend (Table 25). The difference could be attributed to the lower survival rate in the 100 mg/kg group and lower survival rate in the 1000 mg/kg group.

**Table 24 Analysis of Dose-Mortality Trend for Female Mice**

	Method			
	Cox		Kruskal-Wallis	
	Statistics	P-Value	Statistics	P-Value
Time-Adjusted Trend Test				
Depart from Trend	88.3232	0.0000	73.6391	0.0000
Dose-Mortality Trend	36.1198	0.0000	34.4769	0.0000
Homogeneity	124.4430	0.0000	108.1160	0.0000

Source data: Analysis data (SAS 9.1.3) M2F21782

**Table 25 Analysis of Dose-Mortality Trend for Female Mice excluding 1000 mg/kg dose group**

	Method			
	Cox		Kruskal-Wallis	
	Statistics	P-Value	Statistics	P-Value
Time-Adjusted Trend Test	1.4161	0.4926	1.5207	0.4675
Depart from Trend				
Dose-Mortality Trend	1.3041	0.2535	1.0284	0.3105
Homogeneity	2.7201	0.4368	2.5491	0.4665

Source data: Analysis data (SAS 9.1.3) M2F21782

**Trend Analysis**

The test for positive dose-tumor linear trend is the ultimate objective of the evaluation of the carcinogenicity-study. Occasionally, pairwise comparisons are employed, but only under certain condition of the data and are decided on a case-by-case basis. As a cautionary note, blindly impose pairwise comparisons can only undermine the importance of the trend test, inflate the type-I error, and produce untrustworthy results.

The significance of the test is decided based on a decision rule adopted by the Office of Biostatistics. The details of the decision rule can be found in the Appendix of this review.

We only report trend-test results with p-value less than 0.05, which may not imply a statistical significance. Throughout this report, this icon indicates a statistically significant trend:

**Table 26 Report of <sup>Ⓢ</sup> P-values < 0.05 for Test of Positive Linear Dose-Tumor Trends in Female Mice: including 1000 mg/kg dose group**

Organ Name	Tumor Name	CTR0	LOW	MED	HIGH	MAX	P-Value (Exact Method)	P-Value (Asymptotic Method)
Adrenal	PHEOCHROMOCYTOMA	0	0	0	2	0	0.0580	0.0087
Adrenal	PHEOCHROMOCYTOMA, MALIGNANT	0	0	0	1	0	0.2434	0.0470
Harderian gland	ADENOMA	3	3	10	10	9	0.0164	0.0139
Liver	ADENOMA, HEPATOCELLULAR	10	14	14	52	55	0.0000	0.0000 <sup>Ⓢ</sup>
Liver	CARCINOMA, HEPATOCELLULAR	4	3	3	11	51	0.0000	0.0000 <sup>Ⓢ</sup>
Liver	HEPATOBLASTOMA	0	0	0	1	3	0.2434	0.0470
Pituitary	CARCINOMA, ANTERIOR	0	0	0	1	0	0.2434	0.0470

Source data: Analysis data (SAS 9.1.3) M2F21782

**Table 27 Decision Rule for Statistical Significance for Female Mice**

Organ name	Tumor name	Overall tumor type	Tumor rate as PCT. in control group	Suggested interpretation for trend-test
Harderian gland	ADENOMA	Incidental	5.45	Use exact p-value. Use p-value cutoff point of 0.005.
Liver	ADENOMA, HEPATOCELLULAR	Both fatal and incidental	18.18	Use asymptotic p-value. Use p-value cutoff point of 0.005.
Liver	CARCINOMA, HEPATOCELLULAR	Both fatal and incidental	7.27	Use asymptotic p-value. Use p-value cutoff point of 0.005.

**Table 28 Pairwise Comparisons of Statistical Significant Tumor Trends in Female Mice**

Organ Name	Tumor Name	CTR	LOW	MED	HIGH	MAX	P-Value (EXACT)	P-Value (Asymptotic Method)
Harderian Gland	ADENOMA	3		10			0.0276	0.0147
Harderian Gland	ADENOMA	3			10		0.0257	0.0134
Liver	ADENOMA, HEPATOCELLULAR	10			52		0.0000	0.0000
Liver	ADENOMA, HEPATOCELLULAR	10				55	0.0000	0.0000
Liver	CARCINOMA, HEPATOCELLULAR	4			11		0.0364	0.0203
Liver	CARCINOMA, HEPATOCELLULAR	4				51	0.0000	0.0000

**Reviewer's Statistical Findings from the Trend-Test:**

Using Fisher's Exact Test procedure, the positive linear trend of dose-tumor is statistically significant in female mice for hepatocellular adenoma/carcinoma in the liver.

Pairwise comparison between the treated groups and the control group suggests that the minimum effective dose for hepatocellular adenoma in the liver in female mice is 300 mg/kg and 1000 mg/kg for hepatocellular carcinoma in the liver.

**Table 29 Statistically significant positive linear dose-tumor trend found in FEMALE MICE**

Organ Name	Tumor Name	P-Value
LIVER	ADENOMA, HEPATOCELLULAR	0.0000
LIVER	CARCINOMA, HEPATOCELLULAR	0.0000

Appears This Way  
On Original

## Conclusions

The evaluations of Rat Study M-11-00-561 and Mouse Study M-11-00-560 for carcinogenic potential found the following tumor types that indicate a statistically significant dose-tumor positive linear trend.

A positive linear trend was found to be statistically significant for the tumor of ADENOMA AND CARCINOMA, HEPATOCELLULAR in liver in both sexes of rats and mice. Combining these tumors (adenoma hepatocellular and carcinoma hepatocellular) in liver yields statistically significant linear trend in both sexes in the rat as well as in the mouse study (see Appendix).

Furthermore, there is also a significant positive linear trend found for the tumor of HEPATOBLASTOMA in the liver for male mice, and LEYDIG CELL TUMOR in the testis and ADENOMA in the parathyroid gland in male rats. Additional analysis in the mouse study by combining hepatocellular adenoma, hepatocellular carcinoma, and hepatoblastoma in the liver also yield significant linear trend in both sexes (see Appendix).

Although the sponsor concluded in their report of the mouse study that there is an increased incidence of adenoma in the Harderian gland in all male treated groups and females in the 100 mg/kg and higher groups, there is no evidence of positive significant linear trend either by Peto's trend test or Exact Permutation trend test according to my results.

Appears This Way  
On Original

Table 30 Summary of Findings

Species / Gender		Organ Name	Tumor Name	CTR0	LOW	MED	HIGH	MAX	P-Value (Exact Method)	P-Value (Asymptotic Method)
RAT MALE	Trend	Liver	ADENOMA, HEPATOCELLULAR	3	7	7	15	47	0.0000 (⚠)	0.0000
		Liver	CARCINOMA, HEPATOCELLULAR	0	0	0	0	38	0.0000 (⚠)	0.0000
		Parathyroid	ADENOMA	0	0	0	1	2	0.0218 (⚠)	0.0071
		Testis	LEYDIG CELL TUMOR	3	1	0	7	46	0.0000 (⚠)	0.0000
	Pairwise	Liver	ADENOMA, HEPATOCELLULAR	3	7				0.0164	0.0060
		Liver	ADENOMA, HEPATOCELLULAR	3		7			0.0277	0.0114
		Liver	ADENOMA, HEPATOCELLULAR	3			15		0.0000	0.0000
		Liver	ADENOMA, HEPATOCELLULAR	3				47	0.0000	0.0000
		Liver	CARCINOMA, HEPATOCELLULAR	0				38	0.0000	0.0000
		Testis	LEYDIG CELL TUMOR	3			7		0.0423	0.0192
		Testis	LEYDIG CELL TUMOR	3				46	0.0000	0.0000
RAT FEMALE	Trend	Liver	ADENOMA, HEPATOCELLULAR	2	0	8	22	49	0.0000 (⚠)	0.0000
		Liver	CARCINOMA, HEPATOCELLULAR	0	0	1	1	20	0.0000 (⚠)	0.0000
	Pairwise	Liver	ADENOMA, HEPATOCELLULAR	2		8			0.0036	0.0010
		Liver	ADENOMA, HEPATOCELLULAR	2			22		0.0000	0.0000
		Liver	ADENOMA, HEPATOCELLULAR	2				49	0.0000	0.0000
		Liver	CARCINOMA, HEPATOCELLULAR	0				20	0.0000	0.0000
MICE MALE	Trend	Liver	ADENOMA, HEPATOCELLULAR	20	24	32	50	54	0.0000	0.0000 (⚠)
		Liver	CARCINOMA, HEPATOCELLULAR	9	6	18	26	50	0.0000	0.0000 (⚠)
		Liver	HEPATOBLASTOMA	0	0	1	3	10	0.0001	0.0003 (⚠)
	Pairwise	Liver	ADENOMA, HEPATOCELLULAR	20		32			0.0266	0.0176
		Liver	ADENOMA, HEPATOCELLULAR	20			50		0.0000	0.0000
		Liver	ADENOMA, HEPATOCELLULAR	20				54	0.0000	0.0000
		Liver	CARCINOMA, HEPATOCELLULAR	9		18			0.0397	0.0256
		Liver	CARCINOMA, HEPATOCELLULAR	9			26		0.0009	0.0006
		Liver	CARCINOMA, HEPATOCELLULAR	9				50	0.0000	0.0000
Trend	Liver	CARCINOMA, HEPATOCELLULAR	0			3		0.1249	0.0418	
	Liver	HEPATOBLASTOMA	0				10	0.1215	0.0469	
MICE FEMALE	Trend	Liver	ADENOMA, HEPATOCELLULAR	10	14	14	52	55	0.0000	0.0000 (⚠)
		Liver	CARCINOMA, HEPATOCELLULAR	4	3	3	11	51	0.0000	0.0000 (⚠)
	Pairwise	Liver	ADENOMA, HEPATOCELLULAR	10			52		0.0000	0.0000
		Liver	ADENOMA, HEPATOCELLULAR	10				55	0.0000	0.0000
		Liver	CARCINOMA, HEPATOCELLULAR	4			11		0.0364	0.0203
		Liver	CARCINOMA, HEPATOCELLULAR	4				51	0.0000	0.0000

## Reference

- Food and Drug Administration (FDA), (2001), Guidance for Industry: Statistical Aspects of the Design, Analysis, and Interpretation of Chronic Rodent Carcinogenicity Studies of Pharmaceuticals
- D.R. Cox (1972), "Regression Models and Life Tables (with discussion)," Journal of Royal Statistical Society, Series B, 34, 187-220
- E.A. Gehan (1965), "A Generalized Wilcoxon Test for Comparing K Samples Subject to Unequal Patterns of Censorship," Biometrika, 52, 203-223
- N.E. Breslow (1970), "A generalized Kruskal-Wallis test for comparing K samples subject to unequal patterns of censorship," Biometrika, 57, 579-594
- D.G. Thomas, N.E. Breslow, and J.J. Gart (1977), "Trend and Homogeneity Analyses of Proportions and Life Table Data," Computer and Biomedical Research, 10, 373-381.
- R. Peto, M.C. Pike, N.E. Day, et. al., (1980), "Guidelines for Simple, Sensitive Significance Tests for Carcinogenic Effects in Long-term Animal Experiments," in Long-term and Short-term Screening Assays for Carcinogens: A Critical Appraisal, World Health Organization.
- J.K. Haseman (1983), "A Reexamination of False-Positive Rates for Carcinogenesis Studies," Fundamental and Applied Toxicology, 3, 334-339.
- K.K. Lin (1995), "A Regulatory Perspective on Statistical Methods for Analyzing New Drug Carcinogenicity Study Data," Bio/Pharm Quarterly, 1(2), 18-20
- K.K. Lin (1997), "Control of Overall False Positive Rates in Animal Carcinogenicity Studies of Pharmaceuticals," presented at 1997 FDA Forum on Regulatory Sciences, December 8-9, 1997, Bethesda, MD.
- K.K. Lin and M.A. Rahman (1998a), "Overall False Positive Rates in Tests for Linear Trend in Tumor Incidence in Animal Carcinogenicity Studies of New Drugs," Journal of Pharmaceutical Statistics, with discussions, 8(1), 1-22
- K.K. Lin and M.A. Rahman (1998b), "False Positive Rates in Tests for Trend and Differences in Tumor Incidence in Animal Carcinogenicity Studies of Pharmaceuticals under ICH Guidance S1B," unpublished report, Division of Biometrics 2, Center for Drug Evaluation and Research, Food and Drug Administration
- J.K. Haseman (1984), "Statistical Issues in the Design, Analysis and Interpretation of Animal Carcinogenicity Studies," Environmental Health Perspective, 58, 385-392.
- J.K. Haseman (1984), "Issues in Carcinogenicity Testing: Dose Selection," Fundamental and Applied Toxicology, 5, 66-78.
- J.K. Haseman (1999), personal communication to K.K. Lin
- K.C. Chu, C. Cueto, and J.M. Ward (1981), "Factors in the Evaluation of 200 National Cancer Institute Carcinogen Bioassays," Journal of Toxicology and Environmental Health, 8, 251-280
- R.E. Tarone (1982), "The Use of Historical Control Information in Testing for a Trend in Proportions", Biometrics, 38, 215-220

## Appendix

### Exploratory Analysis: Combining Tumors within organs

#### Adenoma Hepatocellular + Carcinoma Hepatocellular in the Liver

##### A. MALE MICE

Organ Name	Tumor Name	CTR0	LOW	MED	HIGH	MAX	P-Value (Exact Method)	P-Value (Asymptotic Method)
Aorta, thoracic	HEMANGIOSARCOMA	0	0	0	1	0	0.2470	0.0488
Harderian gland	ADENOCARCINOMA	1	1	1	2	1	0.0562	0.0266
Kidney	ADENOMA, RENAL CELL	0	0	0	1	0	0.2470	0.0488
Kidney	CARCINOMA, RENAL CELL	0	0	0	1	0	0.2470	0.0488
Kidney	HEMANGIOSARCOMA	0	0	0	1	0	0.2470	0.0488
[C]Liver	ADENOMA + CARCINOMA, HEPATOCELLULAR	25	28	40	54	54	0.0000	0.0000 (!)
Liver	HEPATOBLASTOMA	0	0	1	3	10	0.0001	0.0003 (!)
Lung(bronchus)	ADENOMA, BRONCHIOLO-ALVEOLAR	7	6	10	15	8	0.0483	0.0437
Small intestine, duodenum	PAPILLOMA	0	0	0	1	0	0.2470	0.0488

##### Decision Rule:

Organ name	Tumor name	Overall tumor type	Tumor rate as PCT. in control group	Suggested interpretation for trend-test
[C]Liver	ADENOMA + CARCINOMA, HEPATOCELLULAR	Both fatal and incidental	45.45	Use asymptotic p-value. Use p-value cutoff point of 0.005.
Liver	HEPATOBLASTOMA	Both fatal and incidental	0.00	Use asymptotic p-value. Use p-value cutoff point of 0.025.

##### Pairwise Comparisons:

Organ Name	Tumor Name	CTR0	LOW	MED	HIGH	MAX	P-Value (Exact Method)	P-Value (Asymptotic Method)
[C]Liver	ADENOMA + CARCINOMA, HEPATOCELLULAR	25		40			0.0041	0.0025 (!)
[C]Liver	ADENOMA + CARCINOMA, HEPATOCELLULAR	25			54		0.0000	0.0000 (!)
[C]Liver	ADENOMA + CARCINOMA, HEPATOCELLULAR	25				54	0.0000	0.0000 (!)
Liver	HEPATOBLASTOMA	0			3		0.1249	0.0418
Liver	HEPATOBLASTOMA	0				10	0.1215	0.0469

##### Reviewer's Statistical Findings from the Trend-Test:

Based on the exact method, the minimum effective dose on hepatocellular (adenoma + carcinoma) of the liver is 100 mg/kg (compared to 300 mg/kg when adenoma and carcinoma are done separately). Meanwhile, based on asymptotic method, the minimum effective dose on hepatoblastoma of the liver is 300 mg/kg.

**B. FEMALE MICE**

Organ Name	Tumor Name	CTR0	LOW	MED	HIGH	MAX	Decision Rule	ZXC Table	P-Value (Exact Method)	P-Value (Asymptotic Method)
Adrenal	PHEOCHROMOCYTOMA	0	0	0	2	0	Go	Go	0.0580	0.0087
Adrenal	PHEOCHROMOCYTOMA, MALIGNANT	0	0	0	1	0	Go	Go	0.2434	0.0470
Harderian gland	ADENOMA	3	3	10	10	9	Go	Go	0.0164	0.0139
[C]Liver	ADENOMA+CARCINOMA, HEPATOCELLULAR	13	16	17	53	55	Go	Go	0.0000	0.0000 (!)
Liver	HEPATOBLASTOMA	0	0	0	1	3	Go	Go	0.2434	0.0470
Pituitary	CARCINOMA, ANTERIOR	0	0	0	1	0	Go	Go	0.2434	0.0470

Decision Rule:

Organ name	Tumor name	Overall tumor type	Tumor rate as PCT. in control group	Suggested interpretation for trend-test
[C]Liver	ADENOMA + CARCINOMA, HEPATOCELLULAR	Both fatal and incidental	23.64	Use asymptotic p-value. Use p-value cutoff point of 0.005.

Pairwise Comparisons:

Organ Name	Tumor Name	CTR0	LOW	MED	HIGH	MAX	P-Value (Exact Method)	P-Value (Asymptotic Method)
[C]Liver	ADENOMA + CARCINOMA, HEPATOCELLULAR	13			53		0.0000	0.0000 (!)
[C]Liver	ADENOMA + CARCINOMA, HEPATOCELLULAR	13				55	0.0000	0.0000 (!)

**Reviewer's Statistical Findings from the Trend-Test:**

Pairwise comparison between the treated groups and the control group suggests that the minimum effective dose for hepatocellular (adenoma +carcinoma) in the liver in female mice is 300 mg/kg. When these were tested separately, adenoma in the liver is 300 mg/kg and 1000 mg/kg for hepatocellular carcinoma in the liver.

Appears This Way  
On Original

C. MALE RAT

Organ Name	Tumor Name	CTRO	LOW	MED	HIGH	MAX	P-Value (Monte Carlo)	P-Value (Asymptotic Method)
Cerebrum	GRANULAR CELL TUMOR	0	0	0	0	1	0.1701	0.0178
[C]Liver	ADENOMA+CARCINOMA, HEPATOCELLULAR	3	7	7	15	49	0.0000 (!)	0.0000
Liver	HEMANGIOSARCOMA	0	0	0	0	1	0.1701	0.0178
Liver	HEPATOBLASTOMA	0	0	0	0	1	0.0967	0.0017
Parathyroid	ADENOMA	0	0	0	1	2	0.0225 (!)	0.0071
Pituitary	ADENOMA,INTERMEDIATED	0	0	0	0	1	0.0967	0.0017
Skin	MELANOMA,BENIGN,AMELANOTIC	0	0	0	0	1	0.1701	0.0178
Stomach	LEIOMYOMA	0	0	0	0	1	0.1564	0.0120
Stomach	LEIOMYOSARCOMA	0	0	0	0	1	0.1564	0.0120
Testis	LEYDIG CELL TUMOR	3	1	0	7	46	0.0000 (!)	0.0000
Urinary bladder	CARCINOMA,TRANSITIONAL CELL	0	0	0	0	2	0.0295	0.0014

Pairwise Comparisons:

Organ Name	Tumor Name	CTRO	LOW	MED	HIGH	MAX	P-Value (Exact Method)	P-Value (Asymptotic Method)
[C]Liver	ADENOMA + CARCINOMA, HEPATOCELLULAR	3	7				0.0178	0.0060
[C]Liver	ADENOMA + CARCINOMA, HEPATOCELLULAR	3		7			0.0262	0.0114
[C]Liver	ADENOMA + CARCINOMA, HEPATOCELLULAR	3			15		0.0000	0.0000 (!)
[C]Liver	ADENOMA + CARCINOMA, HEPATOCELLULAR	3				49	0.0000	0.0000 (!)
Testis	LEYDIG CELL TUMOR	3			7		0.0396	0.0192
Testis	LEYDIG CELL TUMOR	3				46	0.0000	0.0000 (!)

Reviewer's Statistical Findings from the Trend-Test:

Pairwise comparison between the treated groups and pooled control group suggests that the minimum effective dose for hepatocellular (adenoma + carcinoma) in the liver is 250 mg/kg. Meanwhile, when these were tested separately, the minimum effective dose for hepatocellular adenoma in the liver in male rats is 250 mg/kg, 1000 mg/kg for Leydig cell tumor in the testis, and 1000 mg/kg for hepatocellular carcinoma in the liver.

**D. FEMALE RAT**

Organ Name	Tumor Name	CTR0	LOW	MED	HIGH	MAX	P-Value (Exact Method)	P-Value (Asymptotic Method)
Cerebellum	GRANULAR CELL TUMOR	0	0	0	0	1	0.1286	0.0060
Kidney	ADENOMA,RENAL CELL	0	0	0	2	0	0.0577	0.0074
Large intestine,rectum	GRANULAR CELL TUMOR	0	0	0	0	1	0.2034	0.0262
[C]Liver	ADENOMA+CARCINOMA,HEPATOCELLULAR	2	0	9	22	50	0.0000 (!)	0.0000
Liver	CHOLANGIOMA	0	0	0	1	0	0.2435	0.0433
Liver	HEMANGIOMA	0	0	0	1	1	0.2435	0.0433
Ovary	GRANULOSA CELL TUMOR	0	0	0	1	0	0.2435	0.0433
Pancreas	HEMANGIOSARCOMA	0	0	0	0	1	0.0294	0.0000
Spinal cord	GRANULAR CELL TUMOR	0	0	0	1	0	0.2435	0.0433
Small intestine,jejunum	LEIOMYOSARCOMA	0	0	0	1	0	0.2435	0.0433
Uterus	ADENOMA	0	0	0	1	0	0.2435	0.0433
Uterus	GRANULAR CELL TUMOR	3	1	1	3	1	0.0706	0.0442
Uterus	ADENOCARCINOMA	0	0	0	0	1	0.1286	0.0060

**Decision Rule:**

Organ name	Tumor name	Overall tumor type	Tumor rate as PCT. in control group	Suggested interpretation for trend-test
[C]Liver	ADENOMA +CARCINOMA, HEPATOCELLULAR	incidental	1.67	Use asymptotic p-value. Use p-value cutoff point of 0.005.

**Pairwise Comparisons:**

Organ Name	Tumor Name	CTR0	LOW	MED	HIGH	MAX	P-Value (Exact Method)	P-Value (Asymptotic Method)
[C]Liver	ADENOMA + CARCINOMA, HEPATOCELLULAR	2		9			0.0015	0.0004 (!)
[C]Liver	ADENOMA + CARCINOMA, HEPATOCELLULAR	2			22		0.0000	0.0000 (!)
[C]Liver	ADENOMA + CARCINOMA, HEPATOCELLULAR	2				50	0.0000	0.0000 (!)

**Reviewer's Statistical Findings from the Trend-Test:**

Pairwise comparison between the treated groups and pooled control group suggests that the minimum effective dose for hepatocellular (adenoma + carcinoma) in the liver is 60 mg/kg. When these were done separately, the minimum effective for hepatocellular adenoma in the liver in female rats is 60 mg/kg, and 1000 mg/kg for hepatocellular carcinoma in the liver.

**Adenoma Hepatocellular + Carcinoma Hepatocellular + Hepatoblastoma in the Liver**

**A. MALE MICE**

Organ Name	Tumor Name	CTR0	LOW	MED	HIGH	MAX	P-Value (Exact Method)	P-Value (Asymptotic Method)
Aorta, thoracic	HEMANGIOSARCOMA	0	0	0	1	0	0.2470	0.0488
Harderian gland	ADENOCARCINOMA	1	1	1	2	1	0.0562	0.0266
Kidney	ADENOMA, RENAL CELL	0	0	0	1	0	0.2470	0.0488
Kidney	CARCINOMA, RENAL CELL	0	0	0	1	0	0.2470	0.0488
Kidney	HEMANGIOSARCOMA	0	0	0	1	0	0.2470	0.0488
[C]Liver	ADENOMA+CARCINOMA+HEPATOBLASTOMA, HEPATOCELLULAR	25	28	40	54	54	0.0000	0.0000 (!)
Lung(bronchus)	ADENOMA, BRONCHIOLO-ALVEOLAR	7	6	10	15	8	0.0483	0.0437
Small intestine, duodenum	PAPILLOMA	0	0	0	1	0	0.2470	0.0488

**Decision Rule:**

Organ name	Tumor name	Overall tumor type	Tumor rate as PCT. in control group	Suggested interpretation for trend-test
[C]Liver	ADENOMA+CARCINOMA+HEPATOBLASTOMA, HEPATOCELLULAR	Both fatal and incidental	45.45	Use asymptotic p-value. Use p-value cutoff point of 0.005.

**Pairwise Comparisons:**

Organ Name	Tumor Name	CTR0	LOW	MED	HIGH	MAX	P-Value (Exact Method)	P-Value (Asymptotic Method)
[C]Liver	ADENOMA+CARCINOMA+HEPATOBLASTOMA, HEPATOCELLULAR	25		40			0.0041	0.0025 (!)
[C]Liver	ADENOMA+CARCINOMA+HEPATOBLASTOMA, HEPATOCELLULAR	25			54		0.0000	0.0000 (!)
[C]Liver	ADENOMA+CARCINOMA+HEPATOBLASTOMA, HEPATOCELLULAR	25				54	0.0000	0.0000 (!)

**Reviewer's Statistical Findings from the Trend-Test:**

Based on exact method, the minimum effective dose on hepatocellular (adenoma + carcinoma + hepatoblastoma) of the liver is 100 mg/kg

**A. FEMALE MICE**

Organ Name	Tumor Name	CTR0	LOW	MED	HIGH	MAX	P-Value (Exact Method)	P-Value (Asymptotic Method)
Adrenal	PHEOCHROMOCYTOMA	0	0	0	2	0	0.0580	0.0087
Adrenal	PHEOCHROMOCYTOMA, MALIGNANT	0	0	0	1	0	0.2434	0.0470
Harderian gland	ADENOMA	3	3	10	10	9	0.0164	0.0139
[C]Liver	ADENOMA+CARCINOMA+HEPATOBLASTOMA, HEPATOCELLULAR	13	16	17	53	55	0.0000	0.0000 (!)
Pituitary	CARCINOMA, ANTERIOR	0	0	0	1	0	0.2434	0.0470

**Decision Rule:**

Organ name	Tumor name	Overall tumor type	Tumor rate as PCT. in control group	Suggested interpretation for trend-test
[C]Liver	ADENOMA+CARCINOMA+HEPATOBLASTOMA, HEPATOCELLULAR	Both fatal and incidental	23.64	Use asymptotic p-value. Use p-value cutoff point of 0.005.

**Pairwise Comparisons:**

Organ Name	Tumor Name	CTR0	LOW	MED	HIGH	MAX	P-Value (Exact Method)	P-Value (Asymptotic Method)
[C]Liver	ADENOMA+CARCINOMA+HEPATOBLASTOMA, HEPATOCELLULAR	13			53		0.0000	0.0000 (!)
[C]Liver	ADENOMA+CARCINOMA+HEPATOBLASTOMA, HEPATOCELLULAR	13				55	0.0000	0.0000 (!)

**Reviewer's Statistical Findings from the Trend-Test:**

Based on exact method, the minimum effective dose on hepatocellular (adenoma + carcinoma + hepatoblastoma) of the liver is 300 mg/kg

Appears This Way  
On Original

## ***Statistical Interpretation of Significance in Evaluation of Tumor-Data Analyses Currently Adopted by CDER Office of Biostatistics***

### **Exact Test**

The statistical interpretation of significance is based on the exact test, if one of the two following situation applies:

The tumor is found either fatal to all the animals or non-fatal to all the Animals.

The tumor is fatal only to some but not to all animals, and time intervals for both analyses of fatal and of incidental tumors do not overlap.

The p-value of the exact test is calculated by using the Permutation test with scores that are the actual doses used.

### **Asymptotic Test**

The asymptotic p-value is calculated based on the normal approximation.

### **Decision Rules**

For the trend test, to adjust for the effect of multiple testing, the decision rules proposed by the Divisions of Biometrics, CDER/FDA are applied to the trend tests in the review. In order to keep the overall Type-I error at the level of about 10%, the rules for trend tests are as follows:

For a carcinogenicity study including two species, tumors with spontaneous tumor rates of 1% or less is tested at the 0.025 significance level. Otherwise, the 0.005 significance level is used.

For a carcinogenicity study including only one species, tumors with spontaneous tumor rates of 1% or less is tested at the 0.05 significance level. Otherwise, the 0.01 significance level is used

For the pairwise comparison test, the decision rules are as follows:

Tumor with spontaneous tumor rates of 1% or less is tested at the 0.05 significance level. Otherwise, the 0.01 significance level is used

Appears This Way  
On Original

## Number of Tumor-Bearing Animals

### Rat Study

Table 31 Number of Tumor-Bearing Male Rats

Organ name	Tumor code	Tumor name	DOSE0	DOSE1	DOSE2	DOSE3	DOSE4
Adrenal	707	ADENOMA,CORTICAL CELL	2	1	2	0	2
Adrenal	773	PHEOCHROMOCYTOMA	11	9	6	2	1
Adrenal	862	PHEOCHROMOCYTOMA,MALIGNANT	0	1	0	1	0
Cerebellum	739	GRANULAR CELL TUMOR	0	0	0	1	0
Cerebellum	755	MENINGIOMA	1	0	0	0	0
Cerebellum	804	ASTROCYTOMA,MALIGNANT	0	1	0	0	0
Cerebrum	721	ASTROCYTOMA	1	0	0	0	0
Cerebrum	739	GRANULAR CELL TUMOR	0	0	0	0	1
Cerebrum	804	ASTROCYTOMA,MALIGNANT	2	0	0	1	0
Hemolymphoreticular(all sites)	845	LYMPHOMA,MALIGNANT	2	0	1	0	0
Hemolymphoreticular(all sites)	871	SARCOMA,HISTIOCYTIC	3	0	2	1	3
Hemolymphoreticular(all sites)	891	LEUKEMIA,MYELOID	1	1	0	1	0
Heart	851	MESOTHELIOMA,MALIGNANT,ATRIOCAVAL	1	0	0	0	0
Kidney	703	ADENOMA,RENAL CELL	0	1	0	0	1
Kidney	750	LIPOMA	2	0	0	0	0
Kidney	821	CARCINOMA,RENAL CELL	0	0	1	0	0
Kidney	874	SCHWANNOMA,MALIGNANT	1	0	0	0	0
Lymph node,mesenteric	743	HEMANGIOMA	3	0	1	0	0
Lymph node,mesenteric	752	LYMPHANGIOMA	1	0	0	0	0
Large intestine,rectum	819	CARCINOMA,SQUAMOUS CELL	0	0	0	1	0
Liver	710	ADENOMA,HEPATOCELLULAR	3	7	7	15	47
Liver	724	CHOLANGIOMA	0	0	0	1	1
Liver	815	CARCINOMA,HEPATOCELLULAR	0	0	0	0	38
Liver	836	HEMANGIOSARCOMA	0	0	0	0	1
Liver	837	HEPATOBLASTOMA	0	0	0	0	1
Lung(bronchus)	705	ADENOMA,BRONCHIOLO-ALVEOLAR	0	0	0	1	1
Lung(bronchus)	819	CARCINOMA,SQUAMOUS CELL	1	0	0	0	0
Mammary gland	700	ADENOMA	1	0	0	0	0
Mammary gland	732	FIBROADENOMA	1	2	0	0	0
Parathyroid	700	ADENOMA	0	0	0	1	2
Pancreas	704	ADENOMA,ACINAR CELL	6	7	7	7	2
Pancreas	713	ADENOMA,ISLET CELL	47	16	23	28	15
Pancreas	818	CARCINOMA,ISLET CELL	3	0	0	1	0
Pituitary	719	ADENOMA,ANTERIOR	84	42	42	40	22
Pituitary	720	ADENOMA,INTERMEDIATED	0	0	0	0	1
Pituitary	865	CARCINOMA,ANTERIOR	0	0	1	1	0
Prostate	801	ADENOCARCINOMA	1	1	0	0	0
Spinal cord	739	GRANULAR CELL TUMOR	0	0	1	0	0
Spinal cord	804	ASTROCYTOMA,MALIGNANT	1	0	0	0	0

Carcinogenicity Review of NDA 21,782

Skin	717	ADENOMA,SEBACEOUS CELL	1	0	1	0	0
Skin	733	FIBROMA	7	3	3	2	1
Skin	748	KERATOACANTHOMA	1	4	2	4	0
Skin	750	LIPOMA	2	1	0	0	0
Skin	761	MELANOMA,BENIGN,AMELANOTIC	0	0	0	0	1
Skin	770	PAPILLOMA,SQUAMOUS CELL	2	5	2	4	1
Skin	809	TUMOR,BASAL CELL,MALIGNANT	0	1	1	0	0
Skin	819	CARCINOMA,SQUAMOUS CELL	1	0	0	0	0
Skin	828	FIBROSARCOMA	4	2	0	2	1
Skin	835	HEMANGIOPERICYTOMA,MALIGNANT	1	0	0	0	0
Skin	854	MELANOMA,AMELANOTIC	0	0	0	1	0
Skin	860	OSTEOSARCOMA	1	0	0	0	0
Skin	868	RHABDOMYOSARCOMA	0	1	0	0	0
Spleen	869	SARCOMA,NOS	0	0	1	0	0
Stomach	749	LEIOMYOMA	0	0	0	0	1
Stomach	819	CARCINOMA,SQUAMOUS CELL	1	0	0	0	1
Stomach	841	LEIOMYOSARCOMA	0	0	0	0	1
Testis	746	LEYDIG CELL TUMOR	3	1	0	7	46
Thyroid	706	ADENOMA,C CELL	8	2	4	2	3
Thyroid	709	ADENOMA,FOLLICULAR CELL	1	2	0	4	1
Thyroid	811	CARCINOMA,C CELL	1	0	0	1	1
Thyroid	814	CARCINOMA,FOLLICULAR CELL	0	0	1	0	1
Urinary bladder	750	LIPOMA	0	0	1	0	0
Urinary bladder	771	PAPILLOMA,TRANSITIONAL CELL	0	0	1	0	0
Urinary bladder	820	CARCINOMA,TRANSITIONAL CELL	0	0	0	0	

Appears This Way  
On Original

Table 32 Number of Tumor-Bearing Female Rats

Organ name	Tumor code	Tumor name	DOSE0	DOSE1	DOSE2	DOSE3	DOSE4
Adrenal	707	ADENOMA,CORTICAL CELL	4	1	1	4	2
Adrenal	769	PARANGLIOMA	0	0	0	1	0
Adrenal	773	PHEOCHROMOCYTOMA	6	0	3	1	1
Adrenal	862	PHEOCHROMOCYTOMA,MALIGNANT	2	0	1	0	0
Adrenal	881	CARCINOMA,CORTICAL CELL	1	1	0	0	0
Cerebellum	739	GRANULAR CELL TUMOR	0	0	0	0	1
Cerebrum	739	GRANULAR CELL TUMOR	0	0	0	0	1
Cerebrum	800	RETICULOSIS,MALIGNANT	0	1	0	0	0
Cerebrum	804	ASTROCYTOMA,MALIGNANT	0	1	0	2	0
Hemolymphoreticular(all sites)	845	LYMPHOMA,MALIGNANT	3	0	1	0	0
Hemolymphoreticular(all sites)	871	SARCOMA,HISTIOCYTIC	1	0	0	2	1
Hemolymphoreticular(all sites)	889	LEUKEMIA,LARGE GRANULAR LYMPHOCYT	0	1	2	0	1
Hemolymphoreticular(all sites)	891	LEUKEMIA,MYELOID	0	0	0	1	0
Harderian gland	801	ADENOCARCINOMA	0	0	1	0	0
Kidney	703	ADENOMA,RENAL CELL	0	0	0	2	0
Kidney	858	NEPHROBLASTOMA	0	0	1	0	0
Large intestine,rectum	739	GRANULAR CELL TUMOR	0	0	0	0	1
Liver	710	ADENOMA,HEPATOCELLULAR	2	0	8	22	49
Liver	724	CHOLANGIOMA	0	0	0	1	0
Liver	743	HEMANGIOMA	0	0	0	1	1
Liver	815	CARCINOMA,HEPATOCELLULAR	0	0	1	1	20
Liver	836	HEMANGIOSARCOMA	0	0	2	1	1
Lung(bronchus)	731	EPITHELIOMA,CYSTIC,KERATINIZING	0	0	0	0	1
Lung(bronchus)	801	ADENOCARCINOMA	0	0	0	1	0
Mammary gland	700	ADENOMA	2	4	3	3	0
Mammary gland	732	FIBROADENOMA	56	21	25	20	4
Mammary gland	801	ADENOCARCINOMA	29	14	17	11	0
Mammary gland	836	HEMANGIOSARCOMA	1	0	0	0	0
Ovary	757	GRANULOSA CELL TUMOR	0	0	0	1	0
Ovary	795	ADENOMA,SERTOLIFORM	0	0	0	1	0
Parathyroid	700	ADENOMA	1	0	0	0	0
Pancreas	704	ADENOMA,ACINAR CELL	2	0	2	3	0
Pancreas	713	ADENOMA,ISLET CELL	10	8	6	4	0
Pancreas	818	CARCINOMA,ISLET CELL	1	0	0	0	0
Pancreas	836	HEMANGIOSARCOMA	0	0	0	0	1
Pituitary	719	ADENOMA,ANTERIOR	97	48	44	45	22
Pituitary	865	CARCINOMA,ANTERIOR	5	4	1	3	0
Spinal cord	739	GRANULAR CELL TUMOR	0	0	0	1	0
Small intestine,ileum	749	LEIOMYOMA	1	1	0	0	0
Small intestine,jejunum	841	LEIOMYOSARCOMA	0	0	0	1	0
Skin	717	ADENOMA,SEBACEOUS CELL	1	0	0	0	0
Skin	748	KERATOACANTHOMA	1	0	1	0	0
Skin	750	LIPOMA	2	1	0	0	0
Skin	799	TRICHOEPITHELIOMA	0	0	1	0	0
Skin	828	FIBROSARCOMA	1	0	0	0	0
Skin	869	SARCOMA,NOS	2	1	0	0	0

Carcinogenicity Review of NDA 21,782

Stomach	819	CARCINOMA,SQUAMOUS CELL	0	1	0	0	0
Thyroid	706	ADENOMA,C CELL	5	7	5	1	0
Thyroid	709	ADENOMA,FOLLICULAR CELL	0	0	1	0	0
Thyroid	811	CARCINOMA,C CELL	3	0	2	1	0
Thyroid	814	CARCINOMA,FOLLICULAR CELL	0	1	0	0	0
Urinary bladder	771	PAPILLOMA,TRANSITIONAL CELL	1	0	1	1	0
Uterus	700	ADENOMA	0	0	0	1	0
Uterus	739	GRANULAR CELL TUMOR	3	1	1	3	1
Uterus	743	HEMANGIOMA	0	1	0	0	0
Uterus	790	POLYP,ADENOMATOUS	2	0	0	0	0
Uterus	791	POLYP,ENDOMETRIAL STROMAL	8	5	2	7	2
Uterus	801	ADENOCARCINOMA	0	0	0	0	1
Uterus	819	CARCINOMA,SQUAMOUS CELL	0	0	0	1	0
Uterus	874	SCHWANNOMA,MALIGNANT	0	0	0	1	0
Vagina	739	GRANULAR CELL TUMOR	6	2	1	5	0
Vagina	841	LEIOMYOSARCOMA	1	0	0	0	0
Vagina	874	SCHWANNOMA,MALIGNANT	0	1	1	0	0

Appears This Way  
On Original

**Mouse Study****Table 33 Number of Tumor-Bearing Male Mice**

Organ name	Tumor code	Tumor name	DOSE0	DOSE1	DOSE2	DOSE3	DOSE4
Adrenal	707	ADENOMA, CORTICAL CELL	0	0	1	0	0
Aorta, thoracic	836	HEMANGIOSARCOMA	0	0	0	1	0
Femur + marrow	743	HEMANGIOMA	0	2	0	0	0
Femur + marrow	836	HEMANGIOSARCOMA	0	0	1	0	0
Hemolymphoreticular(all sites)	845	LYMPHOMA, MALIGNANT	4	5	5	6	2
Hemolymphoreticular(all sites)	846	MASTOCYTOMA, MALIGNANT	0	1	0	0	0
Hemolymphoreticular(all sites)	871	SARCOMA, HISTIOCYTIC	4	4	5	1	3
Harderian gland	700	ADENOMA	6	15	14	16	11
Harderian gland	801	ADENOCARCINOMA	1	1	1	2	1
Kidney	703	ADENOMA, RENAL CELL	0	0	0	1	0
Kidney	821	CARCINOMA, RENAL CELL	0	0	0	1	0
Kidney	836	HEMANGIOSARCOMA	0	0	0	1	0
Liver	710	ADENOMA, HEPATOCELLULAR	20	24	32	50	54
Liver	743	HEMANGIOMA	2	4	1	0	0
Liver	815	CARCINOMA, HEPATOCELLULAR	9	6	18	26	50
Liver	836	HEMANGIOSARCOMA	0	4	4	0	0
Liver	837	HEPATOBLASTOMA	0	0	1	3	10
Lung(bronchus)	705	ADENOMA, BRONCHIOLO-ALVEOLAR	7	6	10	15	8
Lung(bronchus)	801	ADENOCARCINOMA	5	2	1	0	0
Pituitary	720	ADENOMA, INTERMEDIATED	0	0	1	2	0
Small intestine, duodenum	770	PAPILLOMA	0	0	0	1	0
Small intestine, duodenum	789	POLYP	0	1	0	0	0
Small intestine, ileum	743	HEMANGIOMA	0	0	0	0	1
Skin	836	HEMANGIOSARCOMA	2	0	0	0	0
Spleen	743	HEMANGIOMA	2	0	2	1	1
Spleen	836	HEMANGIOSARCOMA	0	4	0	0	0
Stemum + marrow	743	HEMANGIOMA	0	1	0	0	0
Testis	746	LEYDIG CELL TUMOR	1	1	1	0	0
Tongue	819	CARCINOMA, SQUAMOUS CELL	1	0	1	0	0
Thyroid	709	ADENOMA, FOLLICULAR CELL	0	1	0	2	0
Urinary bladder	836	HEMANGIOSARCOMA	0	0	1	0	0

Table 34 Number of Tumor-Bearing Female Mice

Organ Name	Tumor ID	Tumor Name	DOSE0	DOSE1	DOSE2	DOSE3	DOSE4
Adrenal	773	PHEOCHROMOCYTOMA	0	0	0	2	0
Adrenal	862	PHEOCHROMOCYTOMA, MALIGNANT	0	0	0	1	0
Femur + marrow	743	HEMANGIOMA	0	2	0	0	0
Femur + marrow	836	HEMANGIOSARCOMA	0	1	0	1	0
Hemolymphoreticular(all sites)	845	LYMPHOMA, MALIGNANT	13	17	19	21	4
Hemolymphoreticular(all sites)	871	SARCOMA, HISTIOCYTIC	5	9	3	5	6
Harderian gland	700	ADENOMA	3	3	10	10	9
Harderian gland	801	ADENOCARCINOMA	1	2	0	0	0
Heart	836	HEMANGIOSARCOMA	1	0	0	0	0
Lymph node, mesenteric	743	HEMANGIOMA	0	0	1	0	0
Lymph node, mesenteric	836	HEMANGIOSARCOMA	1	0	0	0	0
Large intestine, cecum	749	LEIOMYOMA	0	1	0	0	0
Liver	710	ADENOMA, HEPATOCELLULAR	10	14	14	52	55
Liver	743	HEMANGIOMA	0	1	1	1	0
Liver	747	ITO CELL TUMOR	1	0	1	0	0
Liver	815	CARCINOMA, HEPATOCELLULAR	4	3	3	11	51
Liver	836	HEMANGIOSARCOMA	3	3	1	0	0
Liver	837	HEPATOBLASTOMA	0	0	0	1	3
Lung(bronchus)	705	ADENOMA, BRONCHIOLO-ALVEOLAR	2	4	5	2	3
Lung(bronchus)	801	ADENOCARCINOMA	2	0	5	2	2
Mammary gland	700	ADENOMA	1	0	0	0	0
Ovary	727	CYSTADENOMA	0	1	1	1	0
Ovary	833	GRANULOSA CELL TUMOR, MALIGNANT	0	1	0	0	0
Pituitary	719	ADENOMA, ANTERIOR	3	3	5	2	0
Pituitary	720	ADENOMA, INTERMEDIATED	1	0	0	0	0
Pituitary	805	CARCINOMA, INTERMEDIATED	0	0	0	0	1
Pituitary	865	CARCINOMA, ANTERIOR	0	0	0	1	0
Skin	819	CARCINOMA, SQUAMOUS CELL	0	1	0	0	0
Skin	860	OSTEOSARCOMA	0	0	1	0	0
Skin	869	SARCOMA, SPINDLE CELL	0	1	4	1	1
Spleen	743	HEMANGIOMA	2	1	0	1	0
Spleen	836	HEMANGIOSARCOMA	2	4	1	0	0
Stomach	743	HEMANGIOMA	0	1	0	0	0
Stomach	770	PAPILLOMA	1	1	0	2	1
Stomach	790	POLYP, ADENOMATOUS	1	0	0	0	1
Thyroid	709	ADENOMA, FOLLICULAR CELL	0	1	0	0	0
Thyroid	814	CARCINOMA, FOLLICULAR CELL	0	0	0	0	1
Uterus	791	POLYP, ENDOMETRIAL STROMAL	1	3	0	0	0
Uterus	836	HEMANGIOSARCOMA	0	0	0	1	0
Vagina	792	POLYP, VAGINAL STROMAL	0	1	0	0	0

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

/s/

-----  
Joan Buenconsejo  
3/1/05 03:32:52 PM  
BIOMETRICS

Karl Lin  
3/1/05 03:36:57 PM  
BIOMETRICS  
Concur with review