CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER: 20-922

STATISTICAL REVIEW(S)

Pre-clinical Statistical Consult
Addendum

NDA/ Drug Class: 20-922 / 1S

Name of Product: Tradename™ Depigmenting Solution (Formulation BMS 181158/181159)

Applicant: Bristol-Myers Squibb

Indication: Skin Pigmentation Disorders

I. Introduction:

It was the opinion of the toxicologist that certain tumors and organs analyzed separately in the original report should be combined. The purpose of this addendum to the original statistical analysis is to present these tables.

II. Study No. 93720, The Topical Mouse Study:

Two hundred and fifty male and two hundred and fifty female CD-1 mice were each randomly divided into five equal sized groups, each group having 50 animals. Treatment groups were as follows:

i) Clipped, untreated
ii) Vehicle (100 µL/day/animal)
iii) Low dose (10 µL/day/animal)
iv) Medium dose (30 µL/day/animal)
v) High dose (100 µL/day/animal)

For groups ii)-v) above, the appropriate dose level of the formulation, or the vehicle, was applied by the means of a dedicated, precalibrated pipette. The administered dose was applied daily for approximately 104 weeks to a pre-shaved area on the dorsal region. The untreated animals were sham dosed using the same procedure with water.

This reviewer performed the positive linear trend test on data of all pooled tumor types or pooled organ systems. Following Peto et al (1980), this reviewer applied the 'death rate method/life table' and the 'prevalence method' for testing positive linear trend in both types of tumors. Results for both males and females are displayed in table 8, page 4. Note that page 3 discusses the interpretation of the table on page 4.

Among the groupings of organs and tumors specified by the toxicologist, only systemic hemangioma's and hemangiosarcomas appeared among the males. Whether one used Haseman's decision rules or not (see the discussion on page 3 of this review), there were no statistically significant trends or pairwise differences among dose groups among males in these tumors. Among female mice there was barely a statistically significant evidence of trend in stromal polyps and sarcomas of the cervix ($p \leq 0.0049$ versus the 0.005 specified by Haseman's rule for tests of trend.). There is weak, but statistically significant evidence of a difference

between the pooled control and vehicle groups versus the low dose group, 10 µL/day/animal, in terms of pooled leiomyomas and leiomyosarcomas in the uterus ($p \leq 0.0340$ which is less than the 0.05 specified by using Haseman's rule for comparing groups). Whether or not these marginally statistically significant events are indications of underlying patterns or are merely artifacts of the experiment is a decision for the toxicologist. For other neoplasms no statistically significant difference was found.

Summary:

1. Using the methods of Peto et al (1980), there was no statistically significant evidence of trends or pairwise differences among dose groups for the male mice. However, these tests require many comparisons. Based on general experience Haseman (1970) proposed a p-value adjustment rule applicable to these comparisons. That is, for a roughly 0.10 overall false positive error rate, rare tumors (with a historical control incidence 1% or below) should be tested at a 0.05 level, and common tumors (with a historical control incidence greater than 1%) at a 0.01 level. For tests of trend, for an overall incidence of approximately 0.10, rare tumors should be tested at a level of 0.025, and common tumors at a level of 0.005.

2. In the cervix of female mice in there was barely a statistically significant evidence of trend in stromal polyps and sarcomas ($p \leq 0.0049$ versus the 0.005 specified by Haseman's rule for tests of trend with common tumors.). There was a weak, but statistically significant evidence of a difference between the pooled control and vehicle groups versus the low dose group, in terms of pooled leiomyomas and leiomyosarcomas in the uterus ($p \leq 0.0340$ versus the 0.05 specified by Haseman's rule for tests of differences with rare tumors.). Whether or not these marginally statistically significant events are indications of underlying pattern of response or are merely artifacts of the experiment is a decision for the toxicologist, not the statistician. For other neoplasms no statistically significant difference was found.

References

Peto, R., Pike, M.C., Day, N.E., Gray, R.G., Lee, P.N., Parrish, S., Peto, J., Richards, S., and Wahrendorf, J. (1980) "Guidelines for sample sensitive significance tests for carcinogenic effects in long-term animal experiments," *IARC Monographs on the Evaluation of Carcinogenic Risk of Chemicals to Humans, supplement 2: Long term and Short term Screening Assays for Carcinogens: A Critical Appraisal*, International Agency for Research Against Cancer, 311-426.

Haseman, J. K. (1983), "A Re-examination of false-positive rates for carcinogenesis studies," *Fundamentals of Applied Toxicology,* 3, 334-339.

APPEARS THIS WAY
ON ORIGINAL

/S/ Nov. 12, 1998

Steve Thomson
Mathematical Statistician, Biometrics IV

/S/ Nov 12, '98

concur: R. Srinivasan, Ph.D.
Team Leader, Biometrics IV

This review has 3 text pages including this signature page, 1 table, and 5 total pages .
cc:
Archival: NDA 20,922
HFD-540/
HFD-540/Dr. Wilkin
HFD-540/Mr. Cross
HFD-540/Dr. Nostrand
HFD-540/Dr. Jacobs
HFD-725/Dr. Huque
HFD-725/Mr. Thomson
HFD-725/Dr. Srinivasan
HFD-344/Dr. Pierce
Chron.

APPEARS THIS WAY
ON ORIGINAL

Interpreting Table 8.

In the following table, for each tumor there is a listing of the frequency of tumor type or group of tumors. This is followed by two rows of p-values. The first row provides a test of dose related trend where the control dose is 0 µL/day/animal, the vehicle is arbitrarily rated as 0.05 µL/day/animal, the low dose is 10 µL/day/animal, the medium dose is 30 µL/day/animal, and the high dose is 100 µL/day/animal. Note the 0.05 level dose for the vehicle is an artifact used to prevent the software used to generate the tests from pooling the control and vehicle groups for the test of trend. However, 0.05 µL/day/animal is so close to 0 relative to the other dose levels that the tests used will treat these as almost coincident, The second p-value is a Cochrane-Armitage test of homogeneity of control and vehicle, versus an alternative of trend.

The second row provides p-values of Cochrane-Armitage tests for pairwise comparisons of dose, comparing the pooled control and vehicle to the low dose group (CV vs L), to the medium dose group (CV vs M), and to the high dose group (CV vs H), followed by comparisons of the low dose group to the medium and high dose groups (L vs M and L vs H, respectively), and finally a comparison of the medium dose group to the high dose group (M vs H). In these tests it may be noted that when only fatal tumors or incidental tumors were observed, the reported p-values are based on an exact test. When both types were observed, the printed p-value is from a continuity corrected approximate pooled test.

One statistical problem with interpreting the outcomes from all these statistical tests is the very large number of statistical tests performed. This leads to the so-called "multiplicity problem" in statistical decision theory. Based on general experience Haseman (1970) proposed a p-value adjustment rule applicable to these comparisons. That is, for a roughly 0.10 overall false positive error rate, rare tumors (with a historical control incidence 1% or below) should be tested at a 0.05 level, and common tumors (with a historical control incidence greater than 1%) at a 0.01 level. For tests of trend, for an overall incidence of approximately 0.10, rare tumors should be tested at a level of 0.025, and common tumors at a level of 0.005. Note that the following tables include the incidence rates in the untreated control, and may be used to help determine if a tumor should be classified as rare or not.

Table 8. Tumorigenicity in Mice
Study No. 93720, The Topical Mouse Study

Mouse Males:

Organ / tissue name And tumor name	Dose Level Number of tumors Cntl	Veh	Low	Med	High	p-values of tests trend CV vs L	C vs V CV vs M	CV vs H	L vs M	L vs H	M vs H
Systemic	6	8	3	4	4	0.2881	0.3276				
hemangioma/-sarcoma						0.9351	0.7413	0.4417	0.3273	0.1823	0.4139
					Spontaneous tumor pct: 12%						

Mouse Females:

Organ / tissue name And tumor name	Dose Level Number of tumors Cntl	Veh	Low	Med	High	p-values of tests trend CV vs L	C vs V CV vs M	CV vs H	L vs M	L vs H	M vs H
CERVIX	2	0	2	2	5	0.0049	1.0000				
polyp/sarcoma-stromal						0.3770	0.4135	0.0189	0.7273	0.1910	0.1450
					Spontaneous tumor pct: 4%						
UTERUS	3	5	4	6	2	0.7422	0.4119				
polyp/sarcoma-endo. strom.						0.6394	0.3232	0.8482	0.4185	0.8288	0.9830
					Spontaneous tumor pct: 6%						
Uterus/Cervix	5	5	6	7	7	0.1564	0.7237				
polyp/sarcoma-endo. strom.						0.4862	0.2875	0.2323	0.4521	0.3662	0.6511
					Spontaneous tumor pct: 10%						
Systemic	6	7	8	10	9	0.1046	0.5715				
hemangioma/-sarcoma						0.3505	0.2209	0.1538	0.4674	0.4916	0.4206
					Spontaneous tumor pct: 12%						
UTERUS	0	0	3	1	1	0.3784	.				
leiomyoma/-sarcoma						0.0340	0.3158	0.3286	0.9438	0.9157	0.8500
					Spontaneous tumor pct: <= 1%						
VAGINA	0	0	1	0	0	0.5766	.				
leiomyoma/-sarcoma						0.2769	.	.	1.0000	1.0000	.
					Spontaneous tumor pct: <= 1%						
CERVIX	3	1	1	0	0	0.9882	0.9468				
leiomyoma/-sarcoma						0.8716	1.0000	1.0000	1.0000	1.0000	.
					Spontaneous tumor pct: 6%						
Uterus/Cervix/Vagina	3	1	4	1	1	0.8234	0.9468				
leiomyoma						0.2485	0.8697	0.8657	0.9726	0.9448	0.8500
					Spontaneous tumor pct: 6%						
Uterus/Cervix/Vagina	0	0	1	0	0	0.5714	.				
leiomyosarcoma						0.2800	.	.	1.0000	1.0000	0.1450
					Spontaneous tumor pct: <= 1%						
Uterus/Cervix/Vagina	3	1	5	1	1	0.8732	0.9450				
leiomyoma/-sarcoma						0.1527	0.9114	0.8984	0.9840	0.9797	0.8431
					Spontaneous tumor pct: 6%						

STATISTICAL/CLINICAL REVIEW AND EVALUATION.

SEP 8 1 1999
151

NDA#: 20-922

Applicant: Bristol-Myers Squibb

Name of Drug: 4-hydroxyanysole 2%, tretinoin 0.01% solution

Documents Reviewed: Phase III studies reports dated February 18, 1998 and data on disks provided by the sponsor

Type of Report: NDA review

Indication: Treatment of solar lentigines [redacted] resulting from chronic sun exposure.

Medical officer: Denise Cook, M.D. (HFD-540)

INTRODUCTION

The applicant has submitted two pivotal studies (Protocols DE132-005 and DE132-010) to support the claim that 2% 4-hydroxyanisole/ 0.01% tretinoin solution administered topically twice daily for up to 24 weeks is safe and effective in the treatment of solar lentigines [redacted] resulting from chronic sun exposure.

Throughout the review, the terms "Study 005" and "Study 010" refer to Protocols DE132-005 and DE132-010, respectively. The treatment names 4HA/tretinoin, 4HA, and tretinoin refer to 2% 4-hydroxyanisole/0.01% tretinoin solution, 2% 4-hydroxyanisole, and 0.01% tretinoin, respectively. The designs of the two studies were the same except for: 1) the follow-up period was 24 weeks in Study 005 and 4 weeks in Study 010; 2) The Patient's Self-Assessment Questionnaire was added to Protocol of Study 005 at request of FDA after initiation of the study. Thus only 43% of evaluable subjects in Study 005 completed the Questionnaire at the end of the treatment.

DESIGN

Objective:

The objective of each study was to evaluate the efficacy and safety of 4HA/tretinoin solution in the treatment of solar lentigines [redacted] when administered topically twice-daily for up to 24 weeks.

Study design and methodology:

Each of the two pivotal studies was a multi-center, randomized, parallel-group, double-blind study of 4HA/tretinoin versus tretinoin alone, 4HA alone and vehicle in the treatment of solar lentigines [redacted] Subjects were randomized at a 4:4:2:1 ratio to the 4HA/tretinoin, tretinoin, 4HA and vehicle groups, respectively. Treatments were applied twice daily for up to 24 weeks, followed by a 24 week no treatment observation phase in Study 005 (4 weeks, followed by a 4 week no treatment observation phase in Study 010). Investigators and subjects were blinded to treatments. In Study 005, clinical observations were performed at Weeks 0, 1, 4, 8, 12, 16, 20, 24, 28, 36, 48. In Study 010, clinical observations were performed at Weeks 0, 1, 4, 8, 12, 16, 20, 24, 28.

Diagnosis and main criteria for inclusion:

Healthy adults, 30 years of age and older, both genders, with clinical diagnosis of solar lentigines at least moderately darker than surrounding skin.

Efficacy populations:

Two populations were defined: the "Intent-To-Treat" population which consisted of every subject randomized into the study; and the "Evaluable" population which comprised all subjects randomized into the study who were without significant protocol violations. Evaluable subjects were defined prior to unblinding. All subjects who were randomized and received at least one dose of study medication were evaluated for safety.

Efficacy variables:

Primary:

- Success rate in Physician's Global Assessment;
- Success rate in Subject's Self-Assessment Questionnaire.

Secondary:

- Target Lesion Pigmentation rating;
- Physician's Assessment of Overall Cosmetic Effect.

A) **Physician's Global Assessment** of improvement/worsening was evaluated on a 7-point ordinal scale:

Score	Characteristic	Description
0	Clear	No evidence of hyperpigmentation, 100% improvement.
1	Almost Clear	Very significant clearance (about 90%). Only minor evidence of hyperpigmentation remains.
2	Marked Improvement	Significant improvement (about 75%); some evidence of hyperpigmentation remains.
3	Moderate Improvement	Intermediate between slight and marked improvement; about 50% improvement in appearance of hyperpigmentation
4	Slight Improvement	Some improvement (about 25%); however, significant evidence of hyperpigmentation remains.
5	No Improvement	Hyperpigmentation condition has not changed.
6	Worse	Hyperpigmentation is worse than at week 0 (visit 1).

The Physician's Global Assessment of improvement/worsening compared with baseline was completed for each visit beginning one week after the start of treatment. **"Success" for a subject was defined as moderate improvement or greater.**

B) A **Subject Self-Assessment Questionnaire** consisted of six questions in which the face, forearms and backs of hands were separately evaluated for improvement in **overall** appearance and improvement in **brown spots** was administered at the end of the treatment phase and again at the end of follow-up.

Each subject evaluated the improvement/worsening of the treated sites **at end of treatment and end of follow-up.** These assessments were done separately for the **face, forearms and backs of hands**. The subjects were instructed to think back to how the areas they treated with the medication (face, forearms, back of hands) looked before they began treatment. The subjects were to respond to two questions:

1) How would you rate the overall appearance of your face, both forearms and backs of your hands compared to when you started treatment?

- 0- completely improved
- 1- mostly improved
- 2- slightly improved
- 3- no improvement
- 4- worse

2) How do you compare the color of the brown spots that you were treating on your face, both of your forearms and the backs of your hands, to when you started treatment?

- 0- completely lightened
- 1- much lighter
- 2- slightly lighter
- 3- no change
- 4- darker.

The Subject Self-Assessment Questionnaire was added at request of FDA after initiation of Study 005. Thus only 256 (43%) subjects in the evaluable population of Study 005 completed the questionnaire at the end of treatment. In Study 010, the Subject Self-Assessment Questionnaire was in the protocol from the initiation of the study.

Success in the Subject's Self Assessment was defined as completely or mostly improved (Overall Appearance) and completely or much lighter (Brown Spots).

C) The **Target Lesion Pigmentation** was used by this reviewer to support the Physician's Global Assessment. The target lesion pigmentation was assessed using the 9-point bipolar ordinal scales:

Score	Description
0	Extremely lighter than pigment of surrounding skin (completely depigmented)
1	Markedly lighter than pigment of surrounding skin
2	Moderately lighter than pigment of surrounding skin
3	Slightly lighter than pigment of surrounding skin
4	Equal with pigment of surrounding skin
5	Slightly darker than pigment of surrounding skin
6	Moderately darker than pigment of surrounding skin
7	Markedly darker than pigment of surrounding skin
8	Extremely darker than pigment of surrounding skin

The target lesion pigmentation characteristic was evaluated by the investigator's examination of the target lesion in **each treatment area** and graded using an integer from 0 to 8. Evaluations were conducted at each visit. Each investigator was instructed to consider the condition at all treated sites at the time of the evaluation in relation to his knowledge of the disease, not in relation to evaluation of the subject at a previous visit.

Statistical methods

To assess baseline comparability of treatment groups, differences among investigational centers and treatment groups with respect to age, height, and weight, were evaluated by a two-way analysis of variance (SAS-PROC GLM) with "Investigator," "treatment" and "treatment-Investigator" interaction as effects in the model.

Differences among treatment groups in gender and race were evaluated by the investigator-adjusted Cochran Mantel-Haenzel test for general association (SAS-PROC FREQ, CMH option) or Fisher's Exact Test. Differences among treatment groups in skin type were evaluated by the investigator-adjusted Kruskal-Wallis test (SAS-PROC FREQ, CMH option, scores=rank.) Baseline comparability in Target Lesion Pigmentation was evaluated by an approximation to the Kruskal-Wallis test using an analysis of variance (SAS-PROC GLM) based on the ranks of the raw scores.

In Study 010, the primary efficacy variables are success rates in the Physician's Global Assessment and the Subject's Self Assessment Questionnaire. In Study 005, the primary efficacy variable is success rate in the Physician's Global Assessment alone, because the Subject's Self-Assessment Questionnaire was added into Protocol of Study 005 later and less than 50% of patients have data for the Subject's Self Assessment. The primary efficacy timepoint is the end of treatment.

The Patient's Self Assessment Questionnaire, measured on a 5-point ordinal scale, assessed the extent of improvement in overall appearance and improvement in brown spots. These two questions separately evaluated the face, forearms and backs of hands, and were administered at the end of treatment and end of follow-up.

Since the test drug is a combination drug, the null hypothesis is the union of the three hypotheses that state that 4HA/tretinoin is equal to its components (i.e., 4HA/tret = tretinoin **OR** 4HA/tret = 4HA **OR** 4HA/tret = vehicle). Contrast statements within GLM (or CATMOD in the binary case) were used to evaluate each of these hypotheses at the 0.05 level of significance. **Rejection of this combination null hypothesis requires rejection of each component hypothesis and adopting the alternative hypotheses that 4HA/tretinoin is more effective than each of its components.** Since a Type I error can only be committed by the joint rejection of the three component null hypotheses, the overall Type I error for the three comparisons is necessarily ≤ 0.05.

For the dichotomized outcome analysis of the Physician's Global Assessment and Patient's Self Assessment (analysis of success rates), the general association CMH test controlling for investigator was used.

All-category analysis was considered as supporting for the dichotomized outcome analysis. The Physician's Global Assessment and the Patient's Self Assessment Questionnaire have ordinal scales. Since the numbers assigned to these response measures have no defined metric, linearity of the scale does not apply. However, the natural ordering of the categories does have meaning and the numbers assigned can be helpful, via a rank transformation, in assisting an ordinal comparison of the associated categories. Therefore, for the all-category analysis, the choice of statistics was limited to the class of nonparametric procedures for ordinal data (such as the investigator-adjusted Kruskal-Wallis test). An approximation to the Kruskal-Wallis test in the analysis of variance (SAS-PROC GLM) is based on the ranks of the raw scores. Included in the model were "Treatment", "Investigator", and their interaction (to evaluate the poolability of the data). If the interaction was not significant then it was dropped from the model and the analysis rerun.

Poolability of Data

The method for evaluating the poolability of the data from the different investigational sites was based on the Physician's Global Assessment at end of treatment. Since sparsity of sample size at the individual investigator level was a concern here (particularly due to the unequal 'n' design) some investigational sites had to be combined to keep the sample size for the vehicle of at least 8 patients at each center. The combining of the 14 investigational sites was based upon geographical proximity, resulting in 5 combined sites.

The analysis for poolability used an analysis of variance based on the ranks of the raw scores with "treatment," "Investigator," and the "Investigator-treatment" interaction as independent variables. The investigator-treatment interaction is associated with a measure of uniformity of treatment response (differences between treatments) among investigators. Two types of interactions are possible - quantitative or qualitative. A quantitative interaction occurs whenever the investigator-treatment interaction is significant at ≤ 0.10 and one treatment is consistently better than the other treatment at all centers (the treatment differences in cure rates for the investigational sites have the same sign but differ in magnitude). The statistical significance of the investigator-treatment interaction is due to the varying magnitude of the cure rate differences relative to experimental error. The data were deemed to be poolable whenever no Investigator-treatment interaction was observed or if a quantitative interaction were observed.

A qualitative interaction results when the investigator-treatment interaction is significant and one treatment is significantly better than the other treatment for some subset of investigators, while the opposite is true for another subset. If such interactions were present the data would

be deemed not poolable for the entire set of investigators due to this lack of consistency.

Subgroup Analysis

Subgroup analyses were performed to determine whether pre-existing characteristics such as gender, age, skin type, and baseline Total Lesion Pigmentation were associated with the Physician's Global Assessment at treatment cessation. Classification variables such as age (dichotomized into ≥65 years and <65 years) and gender were evaluated by an ANOVA (SAS-PROC GLM) with the classification variable, treatment, and their interaction as effects in the model. If the interaction were significant, contrasts would be employed (each at alpha=0.05) to evaluate treatment differences within each level of the classification variable. These contrasts would only be performed if the overall test for interaction were significant, preserving the overall Type I error rate at 0.05 under the complete null hypothesis.

For the quantitative independent variables (skin type and baseline Target Lesion pigmentation), regression analyses (SAS-PROC GLM) were performed regressing the ranked Physician's Global Assessment at treatment cessation upon the ranked quantitative variables. Also included in the regression model were treatment and the interaction of treatment and quantitative variable (to evaluate the homogeneity of slope). If the interaction term were not significant then the effect would be dropped from the model and the analysis rerun.

Multiple comparisons.

In the NDA 20-922, the efficacy of 4HA/tretinoin is evaluated **separately** in two different areas of the body **(the face and arms).** To maintain an overall significance level of 0.05, a p-value adjustment for multiple comparisons is needed. This reviewer applied a Bonferroni adjustment for two pairwise comparisons, using a significance level of 0.05/2=**0.025**.

RESULTS OF STUDY 005

A total of 595 subjects were enrolled at 14 study sites: 217 subjects each in the 4HA/tretinoin and tretinoin treatment groups, 106 in the 4HA group and 55 in the vehicle group. Four hundred eighty-seven (82%) subjects completed the 24 week treatment phase of the study and 476 (80%) completed the full 48 weeks.

Of the 595 subjects enrolled, 594 were evaluable at baseline (Week 0). One Subject in the vehicle group had all visit data excluded for efficacy due to concurrent use of Medrol. The demographic characteristics of subjects enrolled in the study are presented in Table 1. Analysis of the subject demographic characteristics revealed no statistically significant differences ($p>0.18$) among treatment groups for any of the measured parameters. There was also no significant difference between the treatment groups at baseline relative to skin type ($p=0.9$) and target lesion pigmentation ($p=0.3$).

Table 1
Demographic Characteristics
(Evaluable Subjects in Study 005)

Parameter	Treatment			
	4HA/Tretinoin	Tretinoin	4HA	Vehicle
No. of Subjects	217	217	106	54
Sex (%Male/Female)	20/80	20/80	12/88	13/87
Race (%Wh./Bl./Other)	98/0/2	98/0/2	98/0/2	96/0/4
Mean Age (range - years)	62.2 (37-85)	63.5 (34-84)	61.8 (39-85)	61.8 (38-76)

Treatment-by-investigator interactions in the Physician's Global Assessment were not significant ($p>0.1$) at either arms or the face, thus the data were deemed poolable. For the Physician's Global Assessment, subjects treated with 4HA/tretinoin demonstrated statistically significant superiority over each of its active components and vehicle at the end of treatment on both the arm and face ($p\leq 0.0017$ in the all-category analysis and $p\leq 0.006$ in the dichotomized outcome analysis).

Table 2 presents the number and percent of subjects in each treatment group that had at least moderate improvement in the Physician's Global Assessment at the end of treatment (success rate at the end of treatment).

Table 2
Success rate in Physician's Global Assessment at End of Treatment
(Percent of Subjects with Moderate or Greater Improvement at End of Treatment)
Evaluable Subjects in Study 005

Treatment Site	Treatment			
	4HA/Tretinoin	Tretinoin	4HA	Vehicle
Arm	n=212 ~~110 (52%)~~	n=213 75 (35%) p<0.001	n=105 25 (24%) p<0.001	n=53 9 (17%) p<0.001
Face	n=212 119 (56%)	n=212 91 (43%) p=0.006	n=104 34 (33%) p<0.001	n=53 10 (19%) p<0.001

During the 24 week post treatment follow-up phase, 4HA/tretinoin continued to demonstrate statistically significant superiority to each of its active components and vehicle for the

Physician's Global Assessment ($p \leq 0.0033$ in the all-category analysis and $p \leq 0.01$ in the dichotomized outcome analysis). Table 3 presents the number and percent of subjects with moderate or greater improvement in the Physician's Global Assessment at the end of the follow-up (success rate at the end of follow-up).

Table 3
Success Rate in Physician's Global Assessment at End of Follow-up
(Percent of Subjects with Moderate or Greater Improvement at End of Follow-up)
Evaluable Subjects in Study 005

Treatment Site	Treatment			
	4HA/Tretinoin	Tretinoin	4HA	Vehicle
Arm	n=167 64 (38%)	n=164 39 (24%) p=0.002	n=78 18 (23%) p=0.008	n=45 4 (9%) p<0.001
Face	n=170 87 (51%)	n=165 64 (39%) p=0.01	n=78 22 (28%) p<0.001	n=45 8 (18%) p<0.001

The secondary efficacy endpoint, Target Lesion Pigmentation, supports the claim that 4HA/tretinoin was significantly superior ($p \leq 0.0197$) over both its active components and vehicle on both arm and face at the end of treatment.

For another secondary efficacy variable, Physician's Assessment of overall cosmetic effect, 4HA/tretinoin is also significantly ($p < 0.0043$) better than each of its active ingredients and vehicle on the arm and face at the end of treatment.

Success rate in the Subject Self-Assessment was not a primary endpoint in Study 005 because the questionnaire was added to the protocol after initiation of the study. Table 4 presents p-values for 4HA/tretinoin relative to its active components and vehicle for the success rate in the Subject Self-Assessment at the end of treatment. As is seen from Table 4, for the subjects who completed the questionnaire, 4HA/tretinoin is not significantly different from tretinoin.

Table 4
Success rates Subject's Self-Assessment at End of Treatment
P-values for 4-HA/Tretinoin Over Individual Components
(Evaluable Subjects in Study 005)

Parameter	Treatment			
	4HA/Tretinoin	Tretinoin	4HA	Vehicle
Overall Appearance Face	n=97	n=94 p=0.946	n=39 p=0.017	n=26 p=0.024
	40(41%)*	40 (43%)	8 (21%)	4 (15%)
Overall Appearance Forearms	n=92	n=90 p=0.319	n=39 p=0.876	n=25 p=0.712
	19 (21%)	25 (28%)	8 (21%)	4 (16%)
Overall Appearance Hands	n=96	n=92 p=0.908	n=39 p=0.1	n=26 p=0.026
	29 (30%)	28 (30%)	7 (18%)	2 (8%)
Brown Spots on Face	n=97	n=94 p=0.554	n=39 p=0.171	n=26 p=0.093
	38 (39%)	34 (36%)	11 (28%)	5 (19%)
Brown Spots on Forearms	n=92	n=90 p=0.955	n=39 0.426	n=25 p=0.265
	26 (28%)	26 (29%)	9 (23%)	4 (16%)
Brown Spots on Hands	n=96	n=92 p=0.847	n=39 p=0.027	n=26 p=0.17
	29 (30%)	30 (33%)	5 (13%)	4 (15%)

* The number and percent (%) of subjects who rated themselves completely or mostly improved (Overall Appearance) and completely or much lighter (Brown Spots).

Results in the ITT Population

Results in the ITT population were very similar to the results in the Per Protocol population. For the Physician's Global Assessment, subjects treated with 4HA/tretinoin demonstrated statistically significant superiority ($p \leq 0.002$) over each of its active components and vehicle at the end of treatment on both the arm and face.

For the secondary efficacy parameter of Target Lesion Pigmentation, 4HA/tretinoin again showed significant superiority ($p \leq 0.008$) over both its active components and vehicle, on both arm and face

at the end of treatment.

Results of the Subject Self-Assessment Questionnaire at the end of treatment showed that 4HA/tretinoin was rated as significantly superior to 4HA and vehicle for most questions, but not significantly different from tretinoin (P>0.3).

Subgroup Analysis:

The analysis for age differences (≥65 vs <65) and sex differences in the end-of-treatment Physician's Global Assessment found no statistically significant differences (p>0.097) and no significant interactions (p>0.7). Analysis for differences in race were not performed, because only 2% of the subjects were non-White. Relative to the baseline pigmentation, subjects who enter the study with a lower pigmentation grade on the face are likely to end the study with a better Physician's Global Assessment score compared with subjects with higher baseline pigmentation. This is true regardless of treatment group.

Safety Results:

Two hundred ninety one (49%) subjects had adverse events (AE) that were considered related to treatment: 130 of 217 (60%) in the 4HA/tretinoin group, 131 of 217 (60%) in the tretinoin group, 21 of 106 (20%) in the 4HA group, and 9 of 55 (16%) in the vehicle group. The Chi-square test showed that 4HA and vehicle were significantly safer than 4HA/tretinoin or tretinoin relative to drug related AE (p=0.001). Safety profiles of 4HA/tretinoin and tretinoin were similar. Table 5 presents a summary of adverse events by treatment group.

Table 5
Number of Subjects With Adverse Events
(Study 005)

	Treatment				
	4HA/Tretinoin	**Tretinoin**	**4HA**	**Vehicle**	**p-value**
No. Of Subjects	217	217	106	55	
Adverse Events (related) n (%)	130 (60%)	131 (60%)	21 (20%)	9 (16%)	0.001
Adverse Events (unrelated) n (%)	164 (76%)	155 (71%)	83 (78%)	41 (75%)	0.57
Serious Adverse Events (unrelated) n (%)	16 (7%)	17 (8%)	6 (6%)	2 (4%)	0.67
Discontinued for Adverse Events n (%)	17 (8%)	18 (8%)	4 (4%)	3 (5%)	0.45

Reviewer's Conclusions on Study 005:

Study 005 supports the claim that 4HA/tretinoin is statistically significantly superior ($p \leq 0.006$) to each of its active ingredients (tretinoin and 4HA) and vehicle on both arm and face at the end of treatment for Physician's Global Assessment. This finding is supported by Target Lesion Pigmentation and Physician's Assessment of Overall Cosmetic Effect.

The safety profiles of 4HA/tretinoin and tretinoin are similar.

RESULTS OF STUDY 010:

Of the 580 subjects enrolled, 579 were evaluable at baseline (Week 0). One subject in the 4HA treatment group had all visit data excluded for efficacy because the subject refused to have clinical evaluations performed and to complete required protocol procedures after Visit 1. The demographic characteristics of subjects enrolled in the study are presented in Table 6. Analysis of the subject demographic characteristics revealed no statistically significant pairwise differences ($p>0.5$) between treatment groups for any of the measured parameters.

Table 6
Demographic Characteristics
(Evaluable Subjects of Study 010)

Parameter	Treatment			
	4HA/Tretinoin	Tretinoin	4HA	Vehicle
No. of Subjects	212	210	104	53
Sex (%Male/Female)	15/85	22/78	13/87	9/91
Race (%White/Black/Other)	91/0/9	90/0/10	88/0/12	87/2/12
Mean Age (range-years)	63.2 (33-88)	64.2 (40-90)	65.2 (36-82)	62.9 (46-80)

Treatment-by-investigator interactions in the Physician's Global Assessment were not significant ($p>0.1$), thus the data were deemed poolable. **On the arm,** in the Physician's Global Assessment, 4HA/tretinoin demonstrated significant superiority ($p \leq 0.004$) over each of its active components and vehicle at the end of treatment. 4HA/tretinoin also demonstrated statistically significant superiority over 4HA and vehicle ($p<0.001$) on the face at the end of treatment. However, **on the face,** 4HA was not statistically significantly better than tretinoin at end of treatment ($p=0.211$ in the all-category analysis and $p=0.097$ in the dichotomized outcome analysis). Table 7 presents the number and percent of subjects in each treatment group that had at least moderate improvement in the Physician's Global Assessment at the end of treatment (success rate in the Physician's Global Assessment at the end of treatment).

Table 7
Success rate in Physician's Global Assessment at End of Treatment
(Percent of Subjects with Moderate or Greater Improvement at End of Treatment)
Evaluable Subjects of Study 010

Treatment Site	Treatment			
	4HA/Tretinoin	Tretinoin	4HA	Vehicle
Arm	n = 208 111 (53%)	n = 207 81 (39%) p=0.004	n = 103 26 (25%) p<0.001	n = 53 6 (11%) p<0.001
Face	n = 209 118 (57%)	n = 207 100 (48%) p=0.097	n = 103 35 (34%) p<0.001	n = 53 6 (11%) p<0.001

During the 4 week post treatment follow-up phase, 4HA/tretinoin continued to demonstrate statistically significant superiority to 4HA and vehicle ($p<0.001$) for the Physician's Global Assessment on the arms and face. Compared to tretinoin, 4HA/tretinoin demonstrated statistically significant superiority on the arm at the end of follow-up in the all-category analysis ($p=0.0189$) and numerical superiority in the dichotomized analysis ($p=0.067$). However, 4HA was not statistically significantly better than tretinoin on **the face** at end of follow-up ($p=0.0613$ in the all-category analysis and $p=0.109$ in the dichotomized outcome analysis). Table 8 presents the number and percent of subjects with moderate or greater improvement in the Physician's Global Assessment at the end of the follow-up.

Table 8
Success rate in Physician's Global Assessment at End of Follow-up
(Percent of Subjects with Moderate or Greater Improvement at End of Follow-up)
Evaluable Subjects of Study 010, Week 4 of Follow-up

Treatment Site	Treatment			
	4HA/Tretinoin	Tretinoin	4HA	Vehicle
Arm	n = 164 79 (48%)	n = 180 69 (38%) p=0.067	n = 93 25 (27%) p<0.001	n = 46 6 (13%) p<0.001
Face	n = 166 99 (60%)	n = 182 93 (51%) p=0.109	n = 93 30 (32%) p<0.001	n = 47 7 (15%) p<0.001

On the **hands**, the results of the dichotomized analysis of the Subject Self-Assessment Questionnaire showed that 4HA/tretinoin was significantly superior to tretinoin, 4HA, and vehicle

at the end of treatment ($p<0.021$). 4HA/tretinoin was also statistically significantly superior to tretinoin and vehicle on the **arms** at the end of treatment ($p\leq 0.02$). Compared to 4HA on the **arms**, 4HA/tretinoin was marginally significantly better in the dichotomized outcome analysis ($p<0.049$) and significantly better in the all-category analysis ($p<0.003$).

Table 9
Success Rates in the Subject's Self-Assessment Questionnaire at End of Treatment
P-values for 4HA/Tretinoin Over Individual Components
(Evaluable Subjects if Study 010)

Parameter	Treatment			
	4HA/tretinoin	Tretinoin	4HA	Vehicle
Overall Appearance Face	n = 193	n=192 p=0.120	n=94 p<0.001	n=49 p<0.001
	92 (48%)*	76 (40%)	24 (26%)	8 (16%)
Overall Appearance Forearms	n = 193	n=195 p=0.001	n=94 p=0.046	n=49 p<0.001
	70 (36%)	40 (21%)	23 (24%)	6 (12%)
Overall Appearance Hands	n = 193	n=190 p=0.001	n=94 0.018	n=49 p<0.001
	70 (36%)	39 (21%)	21 (22%)	5 (10%)
Brown Spots on Face	n = 193	n=192 p=0.355	n=94 p=0.002	n=49 p=0.001
	88 (46%)	78 (41%)	25 (27%)	9 (18%)
Brown Spots on Forearms	n = 193	n=195 p=0.020	n=94 p=0.049	n=49 p=0.017
	70 (36%)	49 (25%)	23 (24%)	9 (18%)
Brown Spots on Hands	n = 194	n=190 p<0.001	n=94 p=0.021	n=49 p=0.001
	72 (37%)	39 (21%)	22 (23%)	6 (12%)

* The number and percent (%) of subjects who rated themselves completely or mostly improved (Overall Appearance) and completely or much lighter (Brown Spots).

On the face, 4HA/tretinoin was not significantly better than tretinoin relative to both appearance ($p=0.12$) and brown spots ($p=0.355$) at the end of treatment. Table 9 presents p-values for

4HA/tretinoin relative to its active components and vehicle for the Subject Self-Assessments at the end of treatment. The top two responses for Overall Appearance (completely or mostly improved) and for Brown Spots (completely or much lighter) were combined for each response measure and the results are presented in Table 9 as a percentage (success rate).

The results in the dichotomized outcome analysis are supported by the results in the all-category comparisons. For example, on the face at the end of the treatment, 4HA/tretinoin is not significantly better than tretinoin relative to the Physician's Global Assessment ($p=0.211$), the Patient's Assessment of overall appearance of the face ($p=0.16$), and the Patient's Assessment of the brown spots on the face ($p=0.453$).

For the secondary efficacy parameter of Target Lesion Pigmentation, 4HA/tretinoin showed significant superiority over both of its active components and vehicle on the arm ($p\leq 0.0023$) at the end of treatment. **For the face**, 4HA/tretinoin was significantly superior to 4HA and vehicle ($p=0.0001$) at the end of treatment, but not statistically different from tretinoin ($p=0.366$).

For the secondary efficacy parameter of Physician's Assessment of Overall Cosmetic Effect, 4HA/tretinoin was statistically significantly superior ($p\leq 0.0179$) to each of its active components and vehicle on the arm at the end of treatment. 4HA/tretinoin was also significantly superior to 4HA and vehicle ($p\leq 0.0002$) on the face at the end of treatment. However, 4HA/tretinoin was not statistically significantly better than tretinoin **on the face** at the end of treatment ($p=0.0829$).

Results in the ITT Population of Study 010

Results in the ITT population were very similar to the results in the Per Protocol population. For the Physician's Global Assessment, 4HA/tretinoin was at least marginally significantly better ($p\leq 0.0279$) than each of its active components and vehicle **on the arm** at the end of treatment. 4HA/tretinoin also demonstrated statistically significant superiority over 4HA and vehicle ($p<0.001$) on the face at the end of treatment. However, 4HA/tretinoin was not statistically significantly superior to tretinoin **on the face** both at the end of treatment ($p=0.136$) and at the end of follow-up ($p=0.1$).

Results of the Subject Self-Assessment Questionnaire showed that 4HA/tretinoin was rated as significantly superior to 4HA and vehicle across all questions at the end of treatment. However, 4HA/tretinoin was not statistically significantly superior to tretinoin **on the face** for both overall appearance and brown spots ($0.14<p<0.5$).

4HA/tretinoin was not statistically superior to tretinoin **on the face** at the end of treatment relative to Physician's Assessment of Cosmetic effect ($p=0.1$). Results for Target Lesion Pigmentation are not available.

Subgroup Analysis

The analysis for age differences (≥65 vs <65) in end-of-treatment Physician's Global Assessment revealed no statistically significant age differences ($p \geq 0.145$), and no significant age-treatment interaction ($p \geq 0.463$) on either the arms or the face.

Analysis for differences in race (dichotomized into White and non-White) were performed, even though only 10% of the subjects were non-White. No statistically significant race ($p \geq 0.466$) or race-treatment interactions ($p \geq 0.452$) were observed.
For the quantitative measures (skin type and baseline pigmentation), regression analyses were performed regressing the end-of-treatment Physician's Global Assessment on these quantitative measures. The results of the test for equality of slopes were not significant ($p \geq 0.215$), permitting a common slope model to be fit. Upon fitting the common slope model, the results indicate no statistically significant linear relationship between any of the quantitative measures and Physician's Global Assessment on either anatomical site ($p \geq 0.349$).

Safety Results:

Three hundred eighty (66%) subjects had adverse events that were considered related to treatment: 171 of 212 (81%) in the 4HA/tretinoin group, 169 of 210 (80%) in the tretinoin group, 26 of 105 (25%) in the 4HA group, and 14 of 53 (26%) in the vehicle group. The Chi-square test showed that 4HA and vehicle were statistically significantly safer than 4HA /tretinoin or tretinoin ($p=0.001$).

Table 10
Number of Subjects With Adverse Events in Study 010

	Treatment				
	4HA/Tretinoin	**Tretinoin**	**4HA**	**Vehicle**	**p-value**
	n = 212	n = 210	n = 105	n = 53	
Adverse Events (related) - n (%)	171 (81%)	169(80%)	26(25%)	14(26%)	0.001
Adverse Events (unrelated)- n (%)	150 (71%)	152 (72%)	80 (77%)	40(75%)	0.74
Serious Adverse Events - n (%)	14 (7%)	9 (4%)	1 (1%)	1 (2%)	0.097
ontinued for Adv. Events - %)	26 (12%)	6 (3%)	5 (5%)	1 (2%)	0.001
Discontinued for Adv. Events (related) - n (%)	20 (9%)	5(2%)	2 (2%)	1 (2%)	0.001

Twenty eight subjects (5%) prematurely discontinued the study due to adverse events that were considered related to treatment: 20 (9%) from the 4HA/tretinoin group, 5 (2%) from the tretinoin group, 2 (2%) from the 4HA group, and 1(2%) from the vehicle group (p=0.001). Table 10 presents a summary of adverse events by treatment group.

The most frequent adverse event related to treatment for 4HA/tretinoin and tretinoin was erythema: 137 (65%) and 139 (66%) respectively. The most frequent adverse event related to treatment for all four treatment groups was burning/stinging/tingling: 87 (41%) reports for 4HA/tretinoin, 86 (41%) reports for tretinoin, 10 (9.5%) reports for 4HA, and 8 (15.1%) reports for vehicle (p=0.001).

Reviewer's Conclusions on Study 010:

1. Arm (forearm plus hand):

Study 010 supports the claim that, on the arm, 4HA/tretinoin is statistically significantly superior to each of its active components (tretinoin and 4HA) and vehicle relative to the Physician's Global Assessment, Subject's Self Assessment, Target Lesion Pigmentation and the Physician's Assessment of Overall Cosmetic Effect.

2. Face:

On the face, at the end of treatment, the difference between 4HA/tretinoin and its component tretinoin is not statistically significant relative to the primary efficacy variables, success rate in Physician's Global Assessment (p=0.097), and success rate in Subject's Self-Assessment of overall appearance (p=0.12) and brown spots (p=0.355).

The results in the dichotomized outcome analysis are supported by the results in the all-category analyses. On the face at the end of the treatment, 4HA/tretinoin is not significantly better than tretinoin relative to the Physician's Global Assessment (p=0.211), the Patient's Assessment of overall appearance of the face (p=0.16), and the Patient's Assessment of the brown spots on the face (p=0.453).

Results for the primary efficacy variables are supported by the results for the secondary efficacy variables: the difference between 4HA/tretinoin and tretinoin on the face at the end of treatment is not statistically significant relative to both Target Lesion Pigmentation (p=0.37) and Physician's Assessment of Cosmetic Effect (p=0.083).

The results at the end of the treatment are supported by the results at the end of follow-up period: at the end of the follow-up period (Week 4), the difference between 4HA/tretinoin and tretinoin on the face is not statistically significant relative to Physician's Global Assessment (p=0.06), Subject Self-Assessment Questionnaire of overall appearance (p=0.06) and brown spots (p=0.37), and Physician's Assessment of Cosmetic Effect (p=0.11).

The results in the Per Protocol population are supported by the results in the ITT population. On the face, at the end of treatment, in the ITT population, the difference between 4HA/tretinoin and tretinoin is not statistically significant relative to Physician's Global Assessment (p=0.14), Subject's Self-Assessment Questionnaire of overall appearance (p=0.15) and brown spots (p=0.50), and Physician's Assessment of Cosmetic Effect (p=0.1).

Safety:

Relative to safety of 4HA/tretinoin versus tretinoin, the 4HA/tretinoin group had a significantly greater percentage of subjects who discontinued due to adverse events (12% vs 3%, p=0.001), and subjects who discontinued due to drug related adverse events (9% vs 2%, p=0.002).

INTEGRATED SAFETY ANALYSIS (STUDIES 005 AND 010 COMBINED)

In the integrated safety analysis, 429 subjects treated with 4HA/tretinoin were compared with 427 subjects treated with tretinoin. The 4HA/tretinoin group had a significantly greater percentage of subjects who discontinued due to adverse events (10.0% vs 5.6%, p=0.016), and subjects who discontinued due to drug related adverse events (6.0% vs 2.1%, p=0.003). Otherwise, the safety profiles of the two drugs were similar.

REVIEWER'S CONCLUSIONS (which may be conveyed to the sponsor):

The applicant submitted two Studies 005 and 010 as a pivotal evidence to support the claim that a combination drug, 4-HA/tretinoin, is safe and effective in the treatment of solar lentigines [redacted] resulting from chronic sun exposure on the arm and face.

EFFICACY:

In Study 010, the primary efficacy variables are success rates in the Physician's Global Assessment and the Subject's Self Assessment. In Study 005, the primary efficacy variable is the Physician's Global Assessment alone, because the Subject's Self-Assessment Questionnaire was added to Protocol of Study 005 after the study was initiated and only 43% of evaluable patients completed the Self Assessment Questionnaire at the end of treatment. Secondary efficacy variables are Target Lesion Pigmentation and Physician's Assessment of Overall Cosmetic Effect.

The primary efficacy timepoint is the end of treatment, the primary efficacy population is Per Protocol population. In NDA 20-922, the efficacy of 4HA/tretinoin is evaluated **separately** in two different areas of the body **(the face and arms).** To maintain an overall significance level of 0.05, a p-value adjustment for multiple comparisons is needed. This reviewer applied a Bonferroni adjustment for two pairwise comparisons, using a significance level of 0.05/2=**0.025**.

1. Study 005:

Study 005 supports the claim that 4HA/tretinoin is statistically significantly superior ($p \leq 0.006$) to each of its active ingredients (tretinoin and 4HA) and vehicle on both arm and face at the end of treatment for Physician's Global Assessment. This finding is supported by Target Lesion Pigmentation and Physician's Assessment of Overall Cosmetic Effect.

2. Study 010:

a) Arm:

Study 010 supports the claim that, on the arm, 4HA/tretinoin is statistically significantly superior to each of its active components (tretinoin and 4HA) and vehicle relative to the Physician's Global Assessment, Subject's Self Assessment, Target Lesion Pigmentation and the Physician's Assessment of Overall Cosmetic Effect.

b) Face:

On the face, at the end of treatment, the difference between 4HA/tretinoin and tretinoin is not statistically significant relative to the primary efficacy variables, success rate in Physician's Global Assessment ($p=0.097$), and success rate in Subject's Self-Assessment of overall appearance ($p=0.12$) and brown spots ($p=0.355$).

The results in the dichotomized outcome analysis are supported by the results in the all-category analyses. On the face at the end of the treatment, 4HA/tretinoin is not significantly better than tretinoin relative to the Physician's Global Assessment ($p=0.211$), the Patient's Assessment of overall appearance of the face ($p=0.16$), and the Patient's Assessment of the brown spots on the face ($p=0.453$).

Results for the primary efficacy variables are supported by the results for the secondary efficacy variables: the difference between 4HA/tretinoin and tretinoin on the face at the end of treatment is not statistically significant relative to Target Lesion Pigmentation ($p=0.37$), and Physician's Assessment of Cosmetic Effect ($p=0.083$).

The results at the end of the treatment are supported by the results at the end of follow-up period: at the end of the Follow-up period (Week 4), the difference between 4HA/tretinoin and tretinoin on the face is not statistically significant relative to Physician's Global Assessment ($p=0.06$), Subject Self-Assessment Questionnaire of overall appearance ($p=0.06$) and brown spots ($p=0.37$), and Physician's Assessment of Cosmetic Effect ($p=0.11$).

The results in the Per Protocol population are supported by the results in the ITT population. On the face, at the end of treatment, in the ITT population, the difference between 4HA/tretinoin and tretinoin is not statistically significant relative to Physician's Global Assessment ($p=0.14$),

Subject's Self-Assessment Questionnaire of overall appearance (p=0.15) and brown spots (p=0.50), and Physician's Assessment of Cosmetic Effect (p=0.1).

SUBGROUP ANALYSES: *Results of the subgroup analyses in Studies 005 and 010 are inconsistent. Therefore, the results of the subgroup analyses are inconclusive and should not be shown in the label.*

SAFETY:

In the safety analysis of Study 010, compared to the tretinoin group, the 4HA/tretinoin group has a significantly greater percentage of subjects who discontinued due to adverse events (12% vs 3%, p=0.001), and subjects who discontinued due to drug related adverse events (9% vs 2%, p=0.002). This finding is supported by the integrated safety analysis of the combined data from Studies 005 and 010: compared to the tretinoin group, the 4HA/tretinoin group has a significantly greater percentage of subjects who discontinued due to adverse events (10.0% vs 5.6%, p=0.016), and subjects who discontinued due to drug related adverse events (6.0% vs 2.1%, p=0.003). Otherwise, the safety profiles of 4HA/tretinoin and tretinoin are similar.

OVERALL CONCLUSIONS (which may be conveyed to the sponsor):

1. Arm:

Pivotal Studies 005 and 010 support the claim that on the arm 4HA/tretinoin is statistically significantly more effective than each of its active ingredients (tretinoin and 4HA) and vehicle.

In the safety analysis of Study 010, compared to the tretinoin group, the 4HA/tretinoin group has a significantly greater percentage of subjects who discontinued due to adverse events (12% vs 3%, p=0.001), and subjects who discontinued due to drug related adverse events (9% vs 2%, p=0.002). This is supported by the integrated safety analysis of the combined data from Studies 005 and 010: compared to the tretinoin group, the 4HA/tretinoin group has a significantly greater percentage of subjects who discontinued due to adverse events (10% vs 6%, p=0.016), and subjects who discontinued due to drug related adverse events (6% vs 2%, p=0.003). Otherwise, the safety profiles of 4HA/tretinoin and tretinoin are similar.

This is a matter for the clinical judgement of the Medical Division to decide whether 4HA/tretinoin should be approved for the treatment of solar lentigines [redacted] resulting from chronic sun exposure on the arm given these safety issues.

2. Face:

In the pivotal Study 010, on the face, 4HA/tretinoin is not statistically significantly superior to tretinoin relative to: primary efficacy variables (Physician's Global Assessment and Subject's Self Assessment), and secondary efficacy variables (Target Lesion Pigmentation and the Physician's

Assessment of Overall Cosmetic Effect), at the end of treatment and at the end of follow-up, in the Per Protocol and ITT populations.

[redacted]

/S/ 08.07.98

Valeria Freidlin, Ph.D.
Mathematical Statistician, Biometrics IV

/S/ Sept 8, '98

Concur: Rajagopalan Srinivasan, Ph.D.
Team Leader, Biometrics IV

cc:
Archival NDA 20-922
HFD-540
HFD-540/Mr. Cross
HFD-540/Dr. Wilkin
HFD-540/Dr. Cook
HFD-725/Dr. Huque
HFD-725/Dr. Srinivasan
HFD-725/Dr. Freidlin
HFD-344/Dr. Carreras
Chron.
This review contains 21 pages.

[redacted]

APPEARS THIS WAY
ON ORIGINAL

HFD-540 Cross

JUL 17 1998

Pre-clinical Statistical Consult

NDA/ Drug Class: 20-922 / 1S

Name of Product: Tradename™ Depigmenting Solution (Formulation BMS 181158/181159)

Applicant: Bristol-Myers Squibb
Pharmaceutical Research Institute
100 Forest Avenue
Buffalo, New York 14213-1091

Indication: [redacted]

Documents Reviewed: Volumes 28 and 32 of NDA 20-922 dated 20 September 1997, plus supporting data on three diskettes.

I. Background:

According to the sponsor: Formulation "BMS-181158/BMS-181159 is being developed as a topical drug combination for the treatment of [redacted]" Two animal carcinogenicity studies (one in mice and one in rats) were included in this submission. The first study, labeled study number 93720, consisted of a report of one 24-month carcinogenicity study in CD-1 mice, intended to assess the oncogenic potential of the formulation when administered topically. In addition, there was a second study, not reviewed here, labeled 93721, of the oncogenic potential of the solution applied to rats. Dr. Amy Nostrand, HFD-540, Division of Dermatologic and Dental Drug Products, the reviewing toxicologist and pharmacologist for this NDA, determined that the following review and evaluation should be limited to the 93720 study.

II. Study No. 93720, The Topical Mouse Study:

II. a. Design

Two hundred and fifty male and two hundred and fifty female CD-1 mice were each randomly divided into five equal sized groups, each group having 50 animals. Treatment groups were as follows:

I) Clipped, untreated
ii) Vehicle (100 µL/day/animal)
iii) Low dose (10 µL/day/animal)
iv) Medium dose (30 µL/day/animal)
v) High dose (100 µL/day/animal)

For groups ii)-v) above, the appropriate dose level of the formulation, or the vehicle, was applied by the means of a dedicated, precalibrated pipette. The administered dose was applied daily for approximately 104 weeks to a pre-shaved area on the dorsal region. The untreated animals were sham dosed using the same procedure with water.

The sponsor indicated that during the study all animals were housed individually and examined regularly for clinical signs of ill health or reaction to treatment. Detailed clinical observations were made on day 0 and biweekly thereafter. Body weights were recorded on the week prior to dose initiation (the pretest), day 0, weekly for the first 13 weeks, thereafter every two weeks until termination.

II. b. Sponsor's Analysis

Survival and incidence data were analyzed using logrank tests to compare the within treatment group survival curves. For the survival data analysis, in the statistical report the sponsor cites the program and paper of Thomas, Breslow, and Gart. However the presentation of the logrank tests seems to this reviewer to suggest that only the hypothesis of homogeneity of survival across treatment strata was performed. The sponsor found statistically significant evidence for lack of homogeneity in survival across strata among males ($p<0.01$), and more equivocal evidence among females (not statistically significant using the logrank test, statistically significant using the Gehan-Breslow-Wilcoxon weights $p<0.04$). It may be noted that the Thomas, et al, paper and program give a test of dose related trend that is apparently not referenced in the sponsor's analysis.

For tumorigenicity analyses the sponsor cites Peto, et al (1980). Using this method of analyses the males and females were analyzed separately for survivorship, incidental tumors (non-fatal tumors discovered at necropsy) and fatal tumors. The judgement of fatality was made by a pathologist. Peto type analyses use the incidence count from control and treated groups, adjusting for survival, to estimate the expected incidence assuming homogeneity. As with this reviewer's analysis, there was no statistically significant evidence of dose tumorigenicity among the subset of neoplasms chosen by the sponsor. However, the sponsor's analysis did not provide any details of the tests performed, only a bald statement that the results were not statistically nonsignificant.

II. c. Reviewer's Analysis

This reviewer independently performed analyses on the weight differences and the survival/ tumorigenicity data. For the survival data analysis, the methods of Cox (1972) and Gehan (1965) were used. The tumor data were analyzed using the techniques described in the paper of Peto, et al (1980), with p-values computed from an exact permutation test or a pooling of permutation tests. All data used in the reviewer's analysis were provided by the sponsor on three diskettes.

II. c. 1. Body Weights and Food Consumption:

To compare body weights in (for example weights see table 2, page 8 of this report) grams across levels of treatment, simple ANOVAs comparing weight means were computed at each time point where weights were measured. Table 2 displays body mean weights by level of treatment with an estimated pooled standard error at the various points in the study, from beginning to end, plus p-values for the ANOVA test of mean differences among levels.

Although it is not clear from the table, for females, from week 25 to week 83, these body weight group means were usually statistically significantly different. For females the vehicle weight group is generally the lowest in weight, with the other groups roughly equal in weight. For males, after 40 weeks the profiles of body weight treatment group means are generally naturally ranked by dose, with the no treatment control having usually the highest weight, followed by the vehicle group, then the low (10 µ/day), medium (30 µ/day), and with the high (100 µ/day) dose group generally the lowest weight. However these rankings are seldom statistically significant. Plots of these means appear as figures 1 and 2, on pages 9 and 10, for males and females respectively.

II. c. 2. Survival Analysis:

Grouped intercurrent mortality rates are given in table 1, page 7, separately for both male and female mice. In the sponsor's analysis, the control group and the vehicle group were pooled, although results were also computed with either group deleted from the study. While it is true that both groups have none of the supposed active ingredient, this reviewer would not consider them to be equivalent treatments. Hence they are not pooled in this reviewer's analysis of mortality. Instead the dose of the vehicle group is made small, close to zero relative to the other dose groups.

The plots of the Kaplan-Meier, product-limit estimates of the survival distributions for day of death of male and female mice are given in figures 3 and 4, on pages 12 and 13 of this report respectively. These are for time to death. The overall homogeneity of the survival distributions of the five treatment groups (Control, Vehicle, Low, Medium, High) as well as the effect of a dose-related trend were tested separately for male and female mice using the Cox logrank test and the Gehan-Breslow Generalized Kruskal-Wallis test. The p-values of these overall tests are given in table 3 on page 11.

For both genders there is statistically significant evidence of a lack of homogeneity in mortality across treatment groups ($p<0.0001$ for both tests for males, and $p\leq0.0350$ and $p\leq0.0246$, respectively, for females). That this lack of homogeneity in mortality is primarily due to trend in dose is attested by the statistically highly significant tests for dose effect ($p<0.0001$ for both tests for males, and $p\leq0.0046$ and $p\leq0.0034$, respectively, for females) and the non-significant tests for departure from trend ($p\leq0.7045$ and $p\leq0.7015$ for males, and $p\leq0.5109$ and $p\leq0.4620$, respectively, for females).

So, to summarize, both genders show statistically significant evidence of a mortality trend due to increasing dose with no other statistically significant evidence of mortality differences in dose group except those differences attributable to trend in dose. Tables 4 and 5, on pages 15 and 16 respectively, have similar results for all pairwise comparisons. The results in these tables were generated by a program described by Thomas, et al (1977), using VERSION 2.1 of their program.

For males , one way to interpret the results of the pairwise difference tests is to denote the no-treatment control group by 0, the vehicle group with a 1, the low dose group by 2, sequentially up to 4, for the high dose group. These can be ordered as suggested by the pairwise tests as follows:

1V 0C 2L 3M 4H

In this diagram, groups connected by a line are not statistically significantly different. Thus, to interpret this diagram, for males, the survival curves of the two highest dose groups are not statistically significantly different, but are both statistically significantly different from the vehicle and control groups. The control, vehicle, and low dose are not statistically significantly different, as are also the low and medium, and medium and high dose groups. This is the kind of result that one would expect with a strong dose related trend.

For females, a similar plot to the above is:

1V 2L 3M 0C 4H

where, again, 0 represents the no-treatment control, 1 the vehicle group, up to 4, the high dose group. Note that the control, vehicle, low, and medium dose groups are not statistically significantly different. The control, medium, and high dose groups are not statistically significantly different. Here the results are a little more equivocal, but the general pattern is that the first four dose groups differ little, while the high dose group has higher mortality than the other dose groups. (Significantly higher than the low or medium dose, but not the control or vehicle.)

II. c. 3. Tumor Data Analysis:

This reviewer performed the positive linear trend test on data of all recorded tumor types. Following Peto et al (1980), this reviewer applied the 'death rate method/life table' and the 'prevalence method' for testing positive linear trend in both types of tumors. Mortality independent tumors (i.e., observable tumors as on the skin) were tested using a life table method equivalent to the death rate method. Overall results for males are displayed in table 6, pages 17-18, for females in table 7, pages 19-22 .

These tables display the incidence of various neoplasm types, plus whether they were classified as incidental, fatal, or mortality-independent (note this latter category was not used by the sponsor's toxicologist. P-values from tests of dose related trend and homogeneity of control and vehicle appear first. For other tests, the dose and control are pooled. Thus the remaining p-values correspond to tests of pairwise homogeneity among the four groups:

i) Pooled control and vehicle)
ii) Low dose (10 µL/day/animal)
iii) Medium dose (30 µL/day/animal)
iv) High dose (100 µL/day/animal).

One statistical problem with interpreting the outcomes from all these statistical tests is

the large number of statistical tests performed. This leads to the so-called "multiplicity problem" in statistical decision theory. Based on general experience Haseman (1970) proposed a p-value adjustment rule is applicable to these comparisons. That is, for a roughly 0.10 overall false positive error rate, rare tumors (with a historical control incidence 1% or below) should be tested at a .05 level, and common tumors (with a historical control incidence greater than 1%) at a 0.01 level. Using this rule, no tumor differences or trends were statistically significant. For the test of trend, for an overall incidence of approximately 0.10, rare tumors should be tested at a level of 0.01, and common tumors at a level of 0.005. Note the detailed listing in tables 8 and 9 give incidence rates in the untreated control, and rnay be used to help determine if a tumor should be classed as rare or not.

Using these rules, there were no statistically significant trends or pairwise differences among dose groups among males. Among females there was a statistically significant evidence of trend in stromal sarcoma of the cervix ($p \leq 0.0037$ versus 0.01 level). However, this was due to two cases, both in the high dose group. Whether or not this is merely an artifact of the experiment is a decision for the toxicologist. For other neoplasms no statistically significant difference was found.

References

Peto, R., Pike, M.C., Day, N.E., Gray, R.G., Lee, P.N., Parrish, S., Peto, J., Richards, S., and Wahrendorf, J. (1980) "Guidelines for sample sensitive significance tests for carcinogenic effects in long-term animal experiments," *IARC Monographs on the Evaluation of Carcinogenic Risk of Chemicals to Humans, supplement 2: Long term and Short term Screening Assays for Carcinogens: A Critical Appraisal*, International Agency for Research Against Cancer, 311-426.

Thomas, D.G., Breslow, N. and Gart, J.J. (1977), "Trend and Homogeneity Analysis of Proportions and Life Table Data," Computers and Biomedical Research 10, 373-381.

Haseman, J. K. (1983), "A Re-examination of false-positive rates for carcinogenisis studies," *Fundamentals of Applied Toxicology,* 3, 334-339.

Ccx, D.R. (1972), "Regression models and life tables," *Journal of the Royal Statistical Society, Series B*, 34, 187-220.

Gehan, E.A. (1965), "A generalized Wilcoxon test for comparing arbitrarily single censored samples," *Biometrika*, 52, 203-223.

Summary:

1. Although there are a number of significant differences in weight gains, particularly among females, only among male mice do these differences seem to be largely attributable to trend in dose. As reflected in the table 2, page 8, among males these differences were seldom statistically significant. Among females the low weight group was generally the vehicle group, with little difference in weights among the other treatment groups.

2. For both genders there is statistically significant evidence of a lack of homogeneity in mortality across treatment groups ($p<0.0001$ for both tests for males, and $p\leq 0.0350$ and $p\leq 0.0246$, respectively, for females). Note that these results do differ slightly from those reported by the sponsor. That this lack of homogeneity in mortality is primarily due to trend in dose is attested by the statistically highly significant tests for dose effect ($p<0.0001$ for both tests for males, and $p\leq 0.0046$ and $p\leq 0.0034$, respectively, for females) and the non-significant tests for departure from trend ($p\leq 0.7045$ and $p\leq 0.7015$ for males, and $p\leq 0.5109$ and $p\leq 0.4620$, respectively, for females).

3. Using the methods of Peto et al (1980), there was no statistically significant evidence of trends or pairwise differences among dose groups for the male mice. Among female mice there was a statistically significant evidence of trend in stromal sarcoma of the cervix ($p\leq 0.0037$ versus 0.01 level). However, this was due to two cases, both in the high dose group. Whether or not this is merely an artifact of the experiment is a decision for the toxicologist. For other neoplasms no statistically significant difference was found.

/S/ 07/14/98

Steve Thomson
Mathematical Statistician, Biometrics IV

/S/ 07/17/98

concur: R. Srinivasan, Ph.D.
Team Leader, Biometrics IV

This review has 6 text pages, 7 tables, and 34 total pages.
cc:
Archival: NDA 20,922
HFD-540/
HFD-540/Dr. Wilkin
HFD-540/Mr. Cross
HFD-540/Dr. Nostrand
HFD-540/Dr. Jacobs
HFD-725/Dr. Huque
HFD-725/Mr. Thomson
HFD-725/Dr. Srinivasan
HFD-344/Dr. Pierce
Chron.

Table 1. Intercurrent Mortality for Both Genders

Number died / number at risk ratio(%)

Sex	Time (weeks)	Control 0.0	Vehicle (0.05)μ/L†	Low 10μ/L	Medium 30μ/L	High 100μ/L
		Treatment Group / Dose Level				
Male	0-52	4/50 8.0%	4/50 8.0%	3/50 6.0%	6/50 12.0%	12/50 24.0%
	53-80	5/46 10.9%	11/46 23.9%	12/47 25.5%	14/44 31.8%	19/38 50.0%
	81-104	21/41 51.2%	13/35 37.1%	17/35 48.6%	21/30 70.0%	13/19 68.4%
Number at Terminal Sacrifice	105	20	22	18	9	6
Female	0-52	2/50 4.0%	0/50 0.0%	5/50 10.0%	1/50 2.0%	3/50 6.0%
	53-80	7/48 14.6%	6/50 12.0%	6/45 13.3%	6/49 12.2%	14/47 29.8%
	81-104	24/41 58.6%	23/44 52.3%	18/39 46.2%	23/43 53.5%	23/33 69.7%
Number at Terminal Sacrifice	105+	17	21	21	20	10

† - This is a small dose level that is used to distinguish the vehicle group from the no-treatment control group. Note that since 0.05≈ 0.0 relative to the other dose levels, these two groups are almost pooled for the trend tests.

APPEARS THIS WAY
ON ORIGINAL

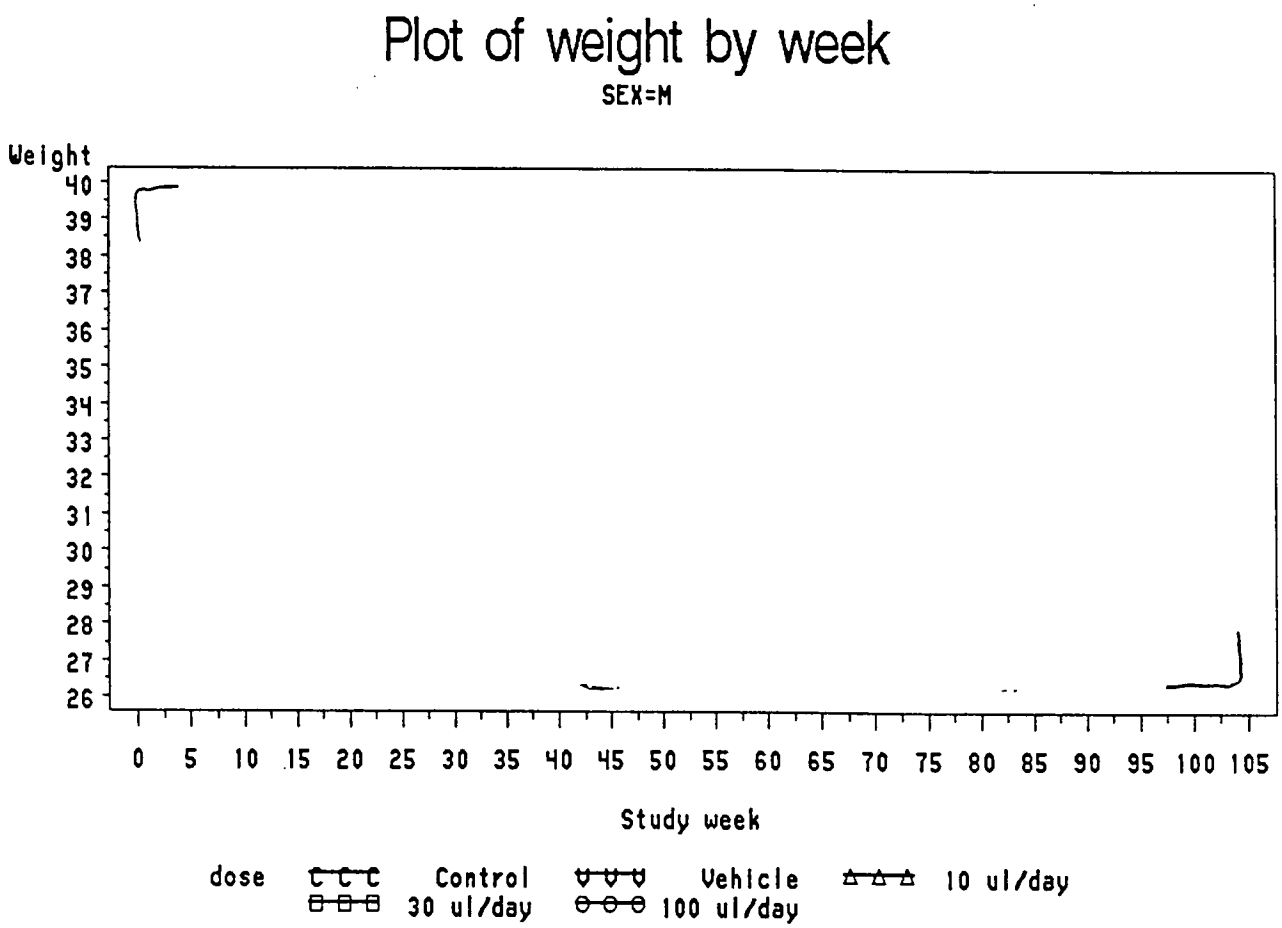
Table 2. Selected Mean Body Weights During Study (grams) CD-1 mice

The following table displays treatment group mean least squares mean weights, standard errors, and group sample size.

	Control LSMean(SE) N	Vehicle LSMean(SE) N	10 ul/day LSMean(SE) N	30 ul/day LSMean(SE) N	100 ul/day LSMean(SE) N	p-value†
Males						
0	26.9 (0.19) 50	26.9 (0.19) 50	26.7 (0.19) 50	26.9 (0.19) 50	26.5 (0.19) 50	0.5500
25	36.8 (0.37) 48	36.6 (0.36) 50	36.8 (0.37) 48	36.7 (0.37) 48	36.2 (0.37) 47	0.7808
51	38.6 (0.40) 46	37.9 (0.40) 46	38.1 (0.39) 47	37.5 (0.41) 44	36.9 (0.44) 38	0.0606
71	39.0 (0.45) 44	38.9 (0.46) 42	38.6 (0.48) 39	38.1 (0.51) 34	37.5 (0.57) 27	0.2405
81	39.5 (0.51) 39	38.4 (0.55) 34	38.6 (0.55) 34	38.4 (0.60) 28	37.0 (0.75) 18	0.1113
91	39.0 (0.55) 32	38.9 (0.60) 27	38.6 (0.63) 25	37.5 (0.70) 20	36.8 (0.90) 12	0.1608
101	38.9 (0.90) 21	38.4 (0.88) 22	38.5 (0.97) 18	37.9 (1.10) 14	37.3 (1.69) 6	0.9083
Females						
0	22.8 (0.17) 50	22.5 (0.17) 50	22.7 (0.17) 50	22.3 (0.17) 50	22.3 (0.17) 50	0.1051
25	32.5 (0.34) 50	31.8 (0.34) 50	33.4 (0.34) 50	32.5 (0.34) 50	32.8 (0.35) 49	0.0294
51	34.7 (0.45) 48	33.3 (0.44) 50	34.8 (0.46) 45	34.6 (0.44) 49	35.3 (0.46) 46	0.0323
71	37.2 (0.55) 44	34.9 (0.53) 47	37.0 (0.56) 42	36.6 (0.53) 46	38.0 (0.60) 37	0.0027
81	37.8 (0.68) 41	35.9 (0.67) 42	37.7 (0.70) 38	37.7 (0.68) 41	38.3 (0.79) 30	0.1189
91	37.1 (0.78) 31	36.0 (0.71) 37	37.1 (0.76) 32	37.8 (0.74) 34	37.0 (0.92) 22	0.5133
101	38.4 (1.04) 20	34.8 (0.97) 23	36.6 (0.97) 23	36.7 (1.02) 21	36.4 (1.29) 13	0.1753

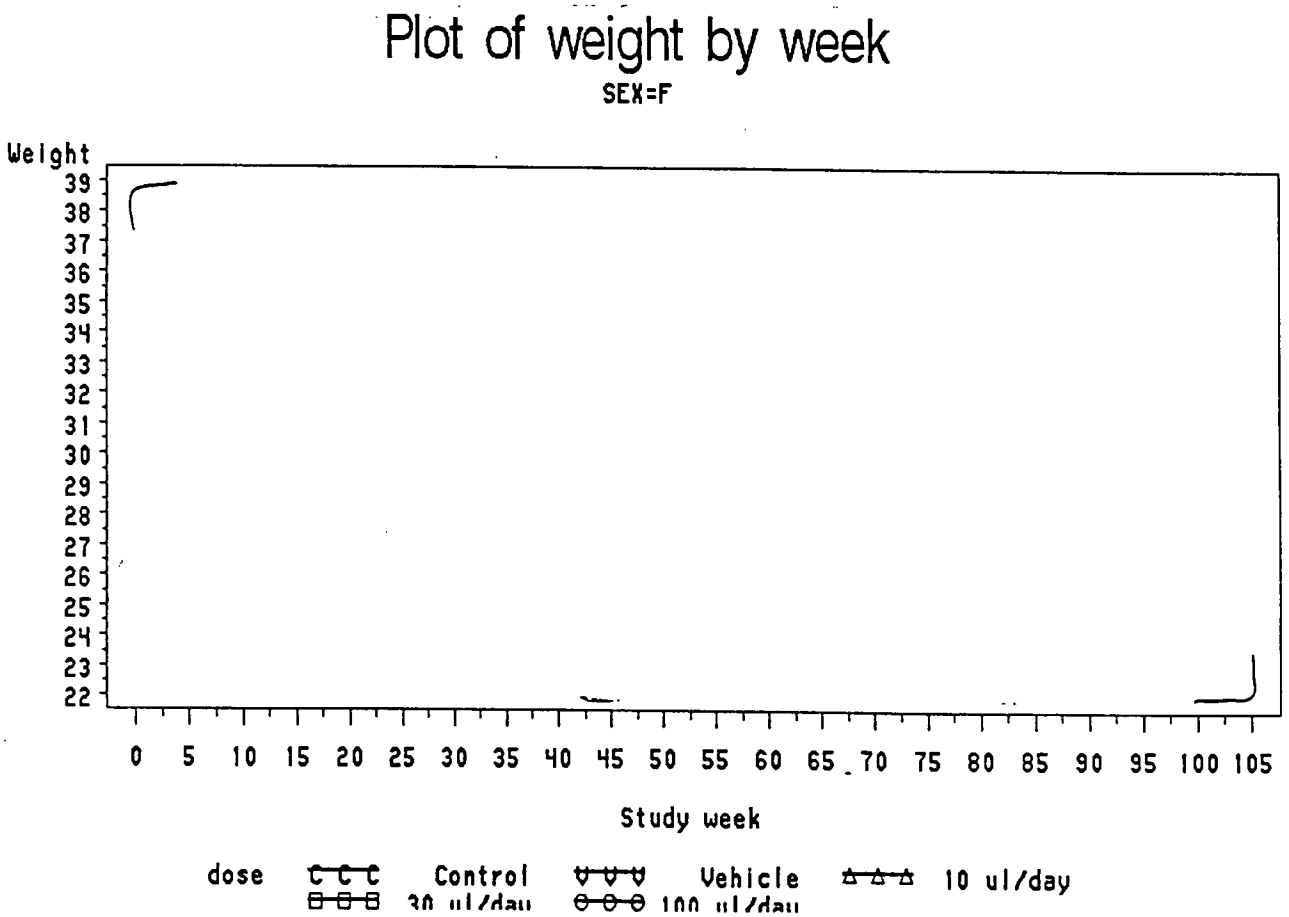
† From an ANOVA of the test of no differences between treatment group means at that point in time.

Figure 1. Males



APPEARS THIS WAY
ON ORIGINAL

Figure 2. Females



APPEARS THIS WAY
ON ORIGINAL

Table 3. Dose Related Trends in Mortality

P-values of tests for positive linear trend, and departure from trend in mortality.

Male:

Method	Time-Adjusted Trend Test	Test Statistic	P-value
Cox (Log-rank)	Dose-Mortality Trend	28.31	0.0000
	Depart from Trend	1.40	0.7045
	Homogeneity	29.71	0.0000
Kruskal-Wallis (Gehan-Breslow-Wilcoxon)	Dose-Mortality Trend	29.59	0.0000
	Depart from Trend	1.42	0.7015
	Homogeneity	31.01	0.0000

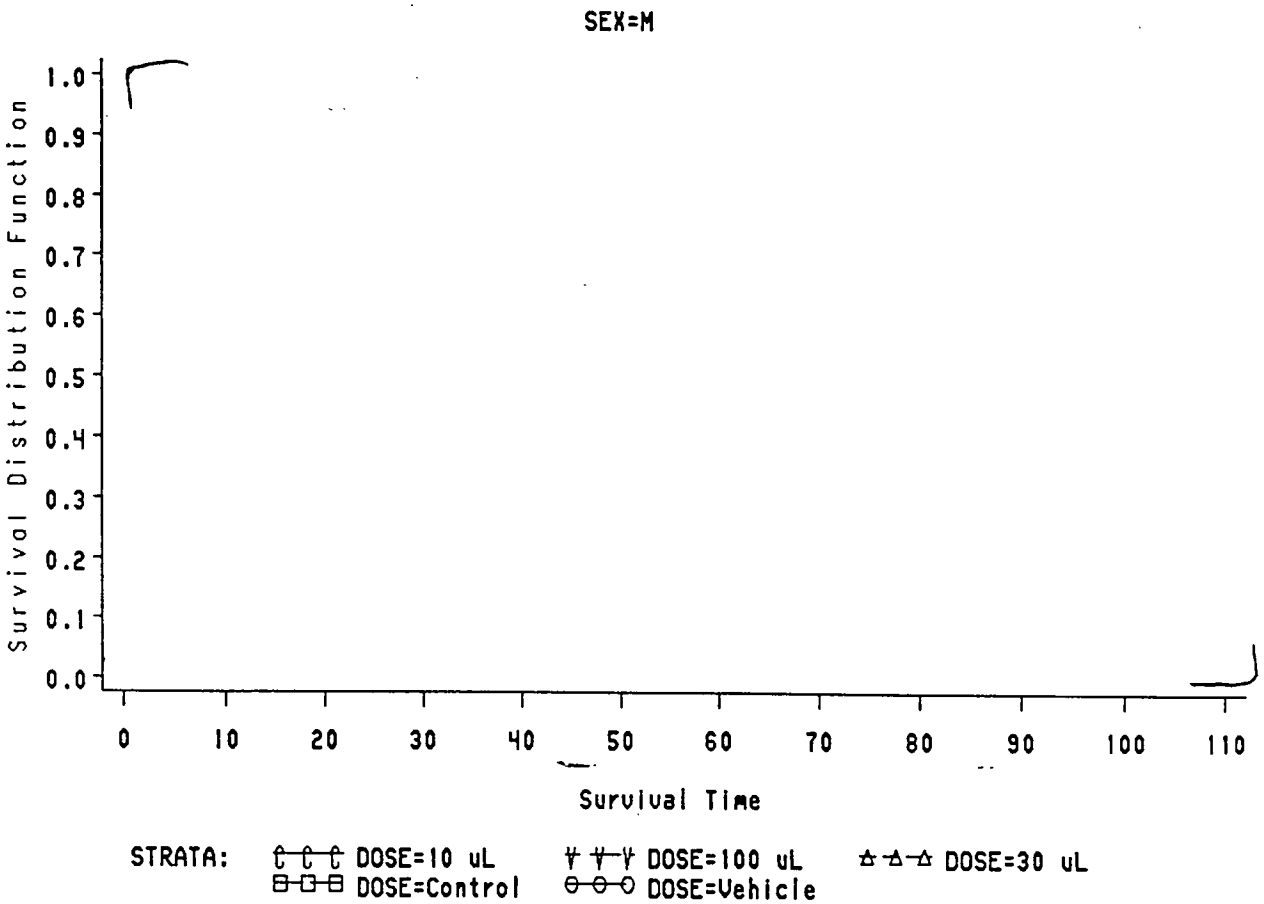
Female:

Method	Time-Adjusted Trend Test	Test Statistic	P-value
Cox (Log-rank)	Dose-Mortality Trend	8.04	0.0046
	Depart from Trend	2.31	0.5109
	Homogeneity	10.35	0.0350
Kruskal-Wallis (Gehan-Breslow-Wilcoxon)	Dose-Mortality Trend	8.60	0.0034
	Depart from Trend	2.57	0.4620
	Homogeneity	11.18	0.0246

Note the Kruskal-Wallis-Gehan-Breslow-Wilcoxon test is more sensitive to discrepancies earlier in the course of the study (when more mice are at risk).

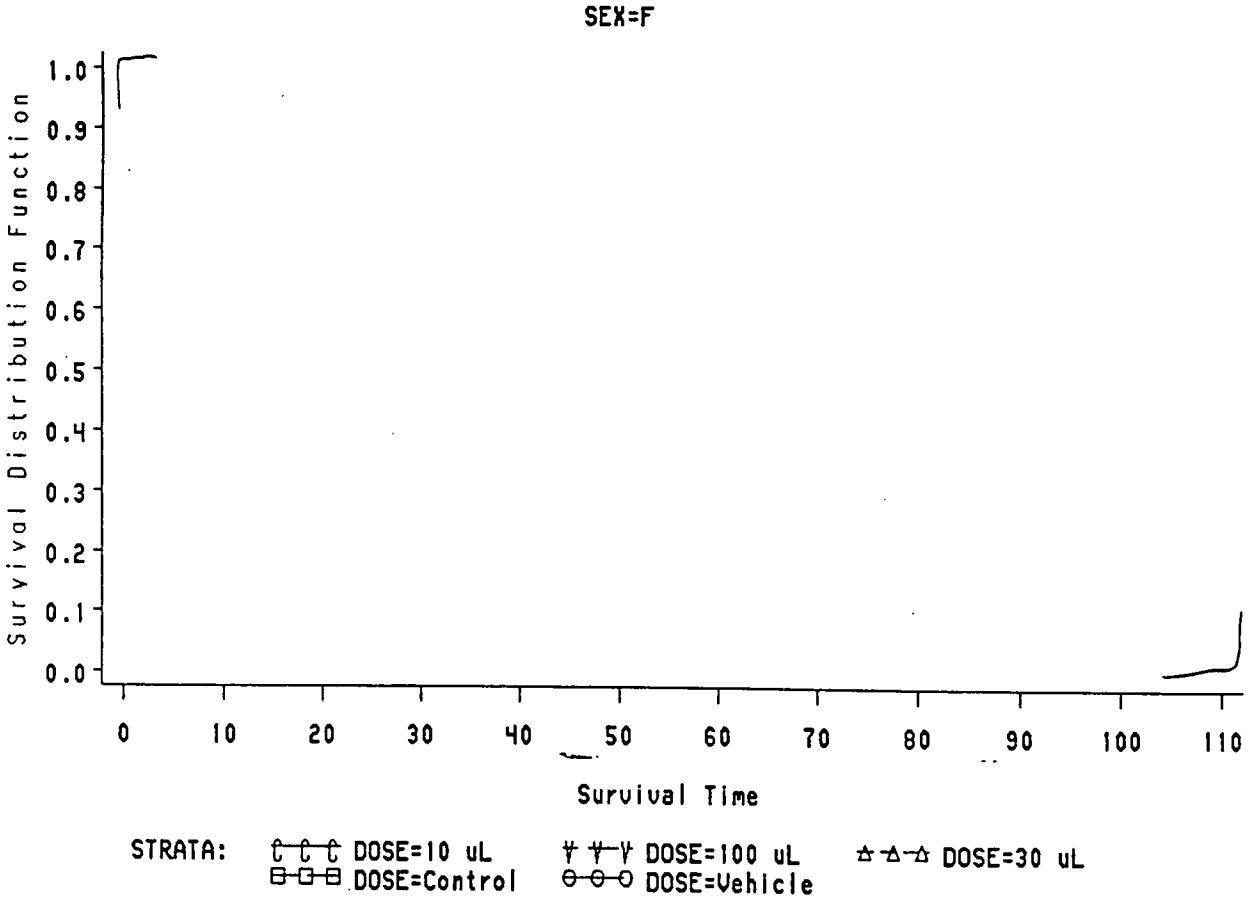
These tests are run using the Trend and Homogeneity Analysis of Proportions and Life Table Data, Version 2.1, by Donald G. Thomas, National Cancer Institute.

Figure 3. Male Estimated Survival



APPEARS THIS WAY
ON ORIGINAL

Figure 4. Female Estimated Survival



APPEARS THIS WAY
ON ORIGINAL

Use of the Tests of Survival Comparing Treatment Groups in Tables 4 and 5.

The following tables 4 and 5 provide tests of treatment group differences in survival separately for each gender. In these tables, group 0 refers to the control group, group 1 to the vehicle, group 2 to the low dose group (10μ/L), group 3 to the medium dose group (30μ/L), and group 4 to the high dose group (100μ/L). To test differences, essentially six different tests are provided, each with a null hypothesis of homogeneity across treatment group:

1) 2x2 Fisher exact test,
2) 2x2 chi-square test of homogeneity,
3) Cox (log-rank) test using an "exact" inverse,
4) An approximate (usually conservative) Cox (log-rank test),
5) Kruskal-Wallis (usually denoted Wilcoxon, or Gehan-Breslow-Wilcoxon), and
6) An approximate (usually conservative) Kruskal-Wallis-Wilcoxon.

Many analysts might question the value of so many tests of essentially the same hypothesis. All these statistics are all provided by the very standard program noted below[1], and apparently there is history in the agency of providing all six tests to the users of these reports. Hence, while this reviewer would be inclined to agree with such a criticism, all six test are available in these tables.

The Fisher exact test and the chi-square test actually ignore time dependence in survival, and merely summarize overall survival. The Cox (log-rank) tests are more sensitive to differences in survival later in the course of the experiment than are the so-called Kruskal-Wallis-Wilcoxon statistics. The versions of the Cox (log-rank) and the Kruskal-Wallis-Wilcoxon statistics labeled "exact inverse" are apparently computed using a matrix inverse in the program (please see computing note below). The approximate tests, labeled "conservative" in 4) and 6) above, avoid this computation, and are, in fact, usually conservative. Unless there is some specific reason for ignoring time dependence in survival, this reviewer would generally recommend use of the exact Cox (log-rank) statistic or the Kruskal-Wallis-Wilcoxon, particularly the former. Again, the Kruskal-Wallis-Wilcoxon test weights by observation and thus emphasizes differences early in the survival curve (where there are more observations).

[1] Thomas, D.G., Breslow, N. and Gart, J.J. (1977), "Trend and Homogeneity Analysis of Proportions and Life Table Data," *Computers and Biomedical Research*, 10, 373-381, program, version 2.1.

Computing Note:

The program cited above apparently implements the calculation of the Cox and Kruskal-Wallis statistics using matrix inversion. Note that even if a matrix A is invertible, computation of A^{-1} is inherently a computationally ill-conditioned problem. In comparison, solution of a system $Ax=b$, i.e., $x=A^{-1}b$, can be implemented as a well conditioned problem (say via Gaussian elimination with full or partial pivots). However, for the inherently 2x2 systems involved in pairwise comparisons, as here, these problems are presumably moot. The approximate test mentioned above (labeled "conservative") avoids computation of this inverse, but to this reviewer, it is still only useful as a possible indicator of numerical instability. Such instability would only obtain when the number of levels tested is moderately large, certainly greater than two. Thus this reviewer would be recommend routinely ignoring tests 4) and 6) above.

Table 4. Males

P-values of pairwise treatment group tests for homogeneity of survival.

In the following table note that group 0 refers to the control group, group 1 to vehicle, group 2 to the low dose group (10μ/L), group 3 to the medium dose group (30μ/L), and group 4 to the high dose group (100μ/L). For an explanation of variables, see the discussion on the preceeding page.

PAIRWISE COMPARISONS (1 D.F. CHI-SQUARES, WITH CONT CORR)

GROUP	EXACT ONE TAIL TEST	2X2 CHI-SQ USING N IN DEN	DIRECTION OF 2X2 CHI-SQ	COX'S TEST EXACT	COX'S TEST INVERSE	COX'S TEST CONSERVATIVE	GENERALIZED K/W EXACT	GENERALIZED K/W INVERSE	GENERALIZED K/W CONSERVATIVE
0 VS.1 CHISQ		.6528	NEG		.0319	.0319		.0046	.0046
PROB	.2096	.4191			.8582	.8583		.9458	.9458
0 VS.2 CHISQ		.1758	POS		.7562	.7546		1.5282	1.5248
PROB	.3377	.6750			.3845	.3850		.2164	.2169
0 VS.3 CHISQ		4.0179	POS		6.5289	6.4734		7.3492	7.2957
PROB	.0220*	.0450*			.0106*	.0110*		.0067**	.0069**
0 VS.4 CHISQ		6.2890	POS		16.5897	15.9883		19.8048	19.3001
PROB	.0056**	.0121*			.0000**	.0001**		.0000**	.0000**
1 VS.2 CHISQ		2.0417	POS		1.1139	1.1130		.9725	.9719
PROB	.0763	.1530			.2912	.2914		.3241	.3242
1 VS.3 CHISQ		8.8648	POS		7.3612	7.3672		6.4111	6.3904
PROB	.0013**	.0029**			.0067**	.0068**		.0113*	.0115*
1 VS.4 CHISQ		11.9682	POS		16.1922	15.8707		17.6319	17.3856
PROB	.0002**	.0005**			.0001**	.0001**		.0000**	.0000**
2 VS.3 CHISQ		1.9200	POS		2.4426	2.4326		2.3171	2.3129
PROB	.0826	.1659			.1181	.1188		.1280	.1283
2 VS.4 CHISQ		3.6138	POS		8.8672	8.7014		10.6769	10.5494
PROB	.0279*	.0573			.0029**	.0032**		.0011**	.0012**
3 VS.4 CHISQ		.0744	POS		2.2898	2.2603		2.2603	3.6193
PROB	.3929	.7850			.1302	.1327		.1327	.0571

* - pvalue ≤ 0.05
** - pvalue ≤ 0.01

THOMAS, D.G., BRESLOW, N. AND GART, J.J. TREND AND HOMOGENEITY ANALYSES OF PROPORTIONS AND LIFE TABLE DATA. COMPUTERS AND BIOMEDICAL RESEARCH 10, 373-381 (1977), VERSION 2.1.

APPEARS THIS WAY
ON ORIGINAL

Table 5. Females

P-values of pairwise treatment group tests for homogeneity of survival.

In the following table note that group 0 refers to the control group, group 1 to vehicle, group 2 to the low dose group (10μ/L), group 3 to the medium dose group (30μ/L), and group 4 to the high dose group (100μ/L). For an explanation of variables, see the discussion on page 14.

PAIRWISE COMPARISONS (1 D.F. CHI-SQUARES, WITH CONT CORR)

GROUP		EXACT ONE TAIL TEST	2X2 CHI-SQ. USING N IN DEN	DIRECTION OF 2X2 CHI-SQ	COX'S TEST EXACT	COX'S TEST INVERSE	COX'S TEST CONSERVATIVE	GENERALIZED K/W EXACT	GENERALIZED K/W INVERSE	GENERALIZED K/W CONSERVATIVE
0 VS.1	CHISQ		1.0263	NEG	1.6308		1.6280	2.7875		2.7830
	PROB	.1555	.3110		.2016		.2020	.0950		.0953
0 VS.2	CHISQ		1.0263	NEG	.6401		.2566	.5718		.5709
	PROB	.1555	.3110		.4237		.6125	.4496		.4499
0 VS.3	CHISQ		.1681	NEG	.2962		.2960	.4734		.4731
	PROB	.3410	.6818		.5863		.5864	.4914		.4916
0 VS.4	CHISQ		.7480	POS	1.6987		1.6917	2.4498		2.4407
	PROB	.1937	.3871		.1925		.1934	.1175		.1182
1 VS.2	CHISQ		.0000	POS	.0432		.0432	.4689		.4685
	PROB	.5793	1.0000		.8354		.8354	.4935		.4937
1 VS.3	CHISQ		.1616	POS	.3317		.3315	.8749		.8744
	PROB	.3440	.6877		.5647		.5648	.3496		.3497
1 VS.4	CHISQ		4.2900	POS	6.9112		6.8690	9.6287		9.5711
	PROB	.0188*	.0383*		.0086**		.0088**	.0019**		.0020**
4										
2 VS.3	CHISQ		.1616	POS	.0259		.0259	.0187		.0187
	PROB	.3440	.6877		.8722		.8722	.8913		.8914
2 VS.4	CHISQ		4.2900	POS	4.7063		4.6661	4.7970		4.7687
	PROB	.0188*	.0383*		.0301*		.0308*	.0285*		.0290*
3 VS.4	CHISQ		2.1836	POS	3.7329		3.7159	4.9044		4.8833
	PROB	.0695	.1395		.0534		.0539	.0268*		.0271*

* - pvalue ≤ 0.05
** - pvalue ≤ 0.01

APPEARS THIS WAY
ON ORIGINAL

Table 6. Male Tumorigenicity

Analysis of Carcinogenic Potential in Male Mouse
Test of Dose-Response (Tumor) Positive Linear Trend and Pairwise Tests

Note: Dose Levels Included: CTRL VEH LOW MED HIGH (0 0.05 10 30 100) denoted C V L M H, respectively.
Tumor Type: IN: Incidental (nonfatal) tumor, FA: Fatal tumor, MI: Mortality independent. For analysis purposes mortality independent tumors are treated as fatal.
Pairwise tests are conducted for Control versus Vehicle (C vs V), and all pairs of pooled control and vehicle (CV), with Low (L), Medium (M), and High (H) dose groups.

ORGAN/TISSUE NAME AND TUMOR NAME	TUMOR TYPES	# Tumors C	V	L	M	H	TREND/ CV vs L	C vs V/ CV vs M	CV vs H	L vs M	L vs H	M vs H
ADRENAL GLANDS	IN	1	1	0	0	0	0.9781	0.8512				
adenoma, cortical							1.0000	1.0000	1.0000	NA	NA	NA
HARDERIAN GLANDS	IN	4	2	2	4	0	0.7899	0.9044				
adenoma							0.7718	0.2324	1.0000	0.2936	1.0000	1.0000
HEART	IN	0	1	0	0	0	0.7529	0.3824				
hemangiosarcoma							1.0000	1.0000	1.0000	NA	NA	NA
JEJUNUM	IN	0	0	1	0	0	0.4459	NA				
cystadenoma							0.3051	NA	NA	1.0000	1.0000	NA
LIVER	FA	0	1	0	0	0	0.2561	0.5561				
adenoma, hepatocellular	IN	9	9	8	14	7	0.7305	0.0557	0.5715	0.0239	0.2738	0.9137
LIVER	FA	2	0	0	1	0	0.4156	0.7842				
carcinoma, hepatocellular	IN	1	2	3	1	2	0.5555	0.6440	0.4346	0.8449	0.7317	0.5325
LIVER	IN	1	1	0	0	0	0.8300	0.5049				
hemangioma	FA	0	1	0	0	0	0.9624	0.9464	0.9386	NA	NA	NA
LIVER	IN	1	2	1	1	1	0.4650	0.7874				
hemangiosarcoma	FA	2	0	0	1	0	0.8868	0.5565	0.7800	0.2885	0.5750	0.7791
LUNG	IN	2	1	1	1	1	0.3608	0.8637				
carcinoma, bronchiolo-alv.	FA	0	0	0	1	0	0.8147	0.3616	0.5198	0.4234	0.7097	0.7996
LUNG	IN	6	6	6	2	0	0.9849	0.3707				
adenoma, bronchiolo-alveolar	FA	0	1	0	0	0	0.6172	0.9549	0.9877	0.9431	1.0000	1.0000

Note in reading these tables, for each tumor there is a listing of the numbers of tumors, and their class (fatal, incidental, or mortality-independent). For each tumor there are a two rows of p-values. The first row provides a test of dose related trend where control dose is 0.0, vehicle dose is 0.05, low dose is 10, medium dose is 30, and high dose is 100. Thus for adenoma in the lung the statistical significance of the test for trend in dose is $p \leq 0.9849$. The significance level of the corresponding test for homogeneity of the control to vehicle is $p \leq 0.3707$. The second row provides p-values of similar tests for comparing the pooled control and vehicle to the low dose group, to the medium dose group, to the high dose group, followed by comparisons of the low dose group to the medium and high dose groups, and finally a comparison of the medium dose group to the high dose group.

Table 6. (cont.) Male Tumorigenicity

Analysis of Carcinogenic Potential in Male Mouse
Test of Dose-Response (Tumor) Positive Linear Trend and Pairwise Tests

ORGAN/TISSUE NAME AND TUMOR NAME	TUMOR TYPES	# Tumors C	V	L	M	H	TREND/ CV vs L	C vs V/ CV vs M	CV vs H	L vs M	L vs H	M vs H
LYMPH NODE, MESENTERIC	IN	0	0	0	0	1	0.1529	NA				
hemangioma							NA	NA	0.2766	NA	0.4333	0.3824
LYMPHORETICULAR SYSTEM	IN	3	0	1	0	0	0.1564	0.6880				
lymphosarcoma, systemic	FA	0	3	2	5	3	0.6150	0.1003	0.3064	0.2703	0.5204	0.6838
LYMPHORETICULAR SYSTEM	IN	0	1	2	1	0	0.3715	0.6778				
sarcoma, histiocytic	FA	2	1	0	0	1	0.5681	0.7239	0.6282	0.7381	0.6793	0.6675
PANCREAS	IN	2	0	0	0	0	1.0000	1.0000				
adenoma, islet-cell							1.0000	1.0000	1.0000	NA	NA	NA
PENIS	IN	0	0	0	1	0	0.5410	NA				
hemangioma							NA	0.4667	NA	0.5385	NA	1.0000
PITUITARY	IN	0	0	1	0	0	0.7377	NA				
adenocarcinoma, pars dist							0.4286	NA	NA	1.0000	1.0000	NA
SKIN (GROSS LESION)	MI	0	0	1	0	0	0.4400	NA				
hemangioma							0.3000	NA	NA	1.0000	1.0000	NA
SKIN (GROSS LESION)	MI	0	1	0	0	0	0.7333	1.0000				
hemangiosarcoma							1.0000	1.0000	1.0000	NA	NA	NA
SPLEEN	IN	0	1	0	0	0	0.7529	0.3824				
hemangioma							1.0000	1.0000	1.0000	NA	NA	NA
SPLEEN	IN	2	1	0	1	0	0.0698	0.8979				
hemangiosarcoma	FA	0	0	1	0	2	0.7846	0.6592	0.1758	0.7742	0.2912	0.3412
TESTES	IN	0	0	0	0	1	0.1529	NA				
hemangioma							NA	NA	0.2766	NA	0.4333	0.3824
TESTES	IN	1	0	1	0	0	0.6897	1.0000				
tumor, interstitial-cell							0.5136	1.0000	1.0000	1.0000	1.0000	NA
THYROID	IN	0	1	1	0	1	0.2165	0.3824				
adenoma, follicular							0.5600	1.0000	0.3670	1.0000	0.5750	0.4000
TRACHEA	IN	0	0	0	0	1	0.0800	NA				
polyp, inflammatory							NA	NA	0.1250	NA	0.2500	0.4000
URINARY BLADDER	IN	0	0	1	0	0	0.6071	NA				
tumor, mesenchymal							0.3400	NA	NA	1.0000	1.0000	NA

APPEARS THIS WAY
ON ORIGINAL

Table 7. Female Tumorigenicity

Analysis of Carcinogenic Potential in Female Mouse
Test of Dose-Response (Tumor) Positive Linear Trend and Pairwise Tests

Note: Dose Levels Included: CTRL VEH LOW MED HIGH (0 0.05 10 30 100) denoted C V L M H respectively.
Tumor Type: IN: Incidental (nonfatal) tumor, FA: Fatal tumor, MI: Mortality Independent.
Pairwise tests are conducted for Control versus Vehicle (C vs V), and all pairs of pooled control and vehicle (CV), Low (L), Medium (M), and High (H) dose groups.

ORGAN/TISSUE NAME AND TUMOR NAME	TUMOR TYPES	# Tumors C	V	L	M	H	TREND/ CV vs L	C vs V/ CV vs M	CV vs H	L vs M	L vs H	M vs H
CECUM	IN	0	0	1	0	0	0.5730	NA				
mastocytoma							0.3559	NA	NA	1.0000	1.0000	NA
CERVIX	IN	0	0	0	1	0	0.4144	NA				
hemangioma							NA	0.3286	NA	0.5610	NA	1.0000
CERVIX	IN	3	1	1	0	0	0.9882	0.9468				
leiomyoma							0.8716	1.0000	1.0000	1.0000	1.0000	NA
CERVIX	IN	2	0	2	2	3	0.0796	1.0000				
polyp, stromal							0.3770	0.4135	0.1179	0.7273	0.4324	0.3722
CERVIX	IN	0	0	0	0	1	0.0037	NA				
sarcoma, stromal	FA	0	0	0	0	1	NA	NA	0.0933	NA	0.2716	0.2448
CERVIX	IN	0	0	1	0	0	0.6667	NA				
schwannoma							0.3158	NA	NA	1.0000	1.0000	NA
CERVIX	IN	0	1	0	0	0	0.7838	0.4894				
tumor, granular-cell							1.0000	1.0000	1.0000	NA	NA	NA
CERVIX	IN	1	0	0	0	0	1.0000	1.0000				
adenoma							1.0000	1.0000	1.0000	NA	NA	NA
ADRENAL GLANDS	IN	0	0	0	2	0	0.2617	NA				
pheochromocytoma							NA	0.1149	NA	0.2317	NA	1.0000
ADRENAL GLANDS	IN	0	0	1	3	1	0.1354	NA				
adenoma, cortical							0.3559	0.0378	0.2083	0.3169	0.5484	0.8596

When reading these tables, for each tumor there is a listing of the numbers of tumors, and their class (fatal, incidental, or mortality-independent). Again, for each tumor there are a two rows of p-values. The first row provides a test of dose related trend where control dose is 0.0, vehicle dose is 0.05, low dose is 10, medium dose is 30, and high dose is 100. Thus for cortical adenoma in the adrenal glands the statistical significance of the test for trend in dose is $p \leq 0.1354$. Since no tumors were identified in both the control and the vehicle groups there is no test of homogeneity. The second row provides p-values of similar tests for comparing the pooled control and vehicle to the low dose group, to the medium dose group, to the high dose group, followed by comparisons of the low dose group to the medium and high dose groups, and finally a comparison of the medium dose group to the high dose group.

Table 7.(cont.) Female Tumorigenicity

Analysis of Carcinogenic Potential in Female Mouse
Test of Dose-Response (Tumor) Positive Linear Trend and Pairwise Tests

ORGAN/TISSUE NAME AND TUMOR NAME	TUMOR TYPES	# Tumors C	V	L	M	H	TREND/ CV vs L	C vs V/ CV vs M	CV vs H	L vs M	L vs H	M vs H
HARDERIAN GLANDS adenocarcinoma	IN	0	0	1	0	0	0.6667	NA				
							0.3158	NA	NA	1.0000	1.0000	NA
HARDERIAN GLANDS adenoma	IN	2	3	0	1	2	0.3760	0.5520				
							1.0000	0.9190	0.5988	0.5610	0.3085	0.5000
LIVER adenoma, hepatocellular	IN	4	0	1	4	2	0.2531	1.0000				
							0.8651	0.2644	0.4884	0.1805	0.3501	0.7647
LIVER carcinoma, hepatocellular	IN	0	2	1	1	1	0.2267	0.2704				
	FA	0	0	0	0	1	0.6981	0.6990	0.4181	0.7805	0.5288	0.5613
LIVER hemangioma	IN	1	0	0	2	1	0.1824	1.0000				
							1.0000	0.2591	0.4519	0.2736	0.3226	0.7816
LIVER hemangiosarcoma	IN	1	2	2	5	1	0.6487	0.5638				
	FA	1	1	0	0	0	0.7231	0.2272	0.8451	0.2499	0.9235	0.9470
LIVER sarcoma, NOS	IN	0	0	0	1	0	0.4144	NA				
							NA	0.3286	NA	0.5610	NA	1.0000
LUNG carcinoma, bron.-alveolar	IN	1	0	2	0	2	0.1857	0.2644				
	FA	0	4	1	0	2	0.5769	0.9803	0.2602	0.9884	0.5123	0.0570
LUNG adenoma, bron.-alveolar	IN	6	7	6	3	3	0.8894	0.7280				
	FA	1	0	0	0	0	0.8425	0.9637	0.9007	0.8377	0.8399	0.6373
LYMPH NODE, MESENTERIC hemangioma	IN	0	0	1	0	1	0.1602	NA				
							0.3554	NA	0.3071	1.0000	0.7026	0.5000
LYMPHORETICULAR SYSTEM lymphoma, thymic	IN	3	0	2	0	0	0.9423	0.9612				
	FA	0	1	1	0	0	0.5813	0.9770	0.9561	0.9817	1.0000	NA
LYMPHORETICULAR SYSTEM lymphosarcoma, systemic	IN	4	5	4	3	2	0.5094	0.8678				
	FA	7	4	3	4	5	0.8535	0.8646	0.6180	0.6227	0.3448	0.4738
LYMPHORETICULAR SYSTEM sarcoma, histiocytic	IN	0	1	0	0	3	0.3426	0.4527				
44444	FA	3	4	4	2	1	0.6043	0.8927	0.4135	0.9020	0.5602	0.2352
MAMMARY GLAND/REGION adenoacanthoma	IN	0	0	0	1	0	0.4144	NA				
							NA	0.3286	NA	0.5610	NA	1.0000
MAMMARY GLAND/REGION adenocarcinoma	IN	3	2	2	1	1	0.8023	0.8637				
							0.6774	0.9173	0.9051	0.9235	0.9235	0.7627
MAMMARY GLAND/REGION adenoma	IN	1	0	0	0	0	1.0000	1.0000				
							1.0000	1.0000	1.0000	NA	NA	NA
OVARIES adenoma, papillary	IN	2	1	1	0	0	0.9621	0.9012				
							0.8124	1.0000	1.0000	1.0000	1.0000	NA
OVARIES cystadenoma	IN	0	2	0	3	0	0.4977	0.2987				
							1.0000	0.2185	1.0000	0.1069	NA	1.0000
OVARIES granulosa-theca-cell tumor	IN	1	0	0	0	0	1.0000	1.0000				
							1.0000	1.0000	1.0000	NA	NA	NA

Table 7.(cont.) Female Tumorigenicity

Analysis of Carcinogenic Potential in Female Mouse
Test of Dose-Response (Tumor) Positive Linear Trend and Pairwise Tests

ORGAN/TISSUE NAME AND TUMOR NAME	TUMOR TYPES	# Tumors C	V	L	M	H	TREND/ CV vs L	C vs V/ CV vs M	CV vs H	L vs M	L vs H	M vs H
OVARIES	IN	0	1	0	2	1	0.2640	0.3052				
hemangioma	FA	0	1	0	0	0	0.9600	0.4424	0.6428	0.2317	0.5489	0.7907
OVARIES	IN	0	0	1	0	0	0.6667	NA				
hemangiosarcoma							0.3158	NA	NA	1.0000	1.0000	NA
OVARIES	IN	0	1	0	0	1	0.2324	0.4894				
luteoma							1.0000	1.0000	0.4519	NA	0.3226	0.3333
OVARIES	IN	0	0	1	0	0	0.5730	NA				
tumor, tubulo-stromal							0.3559	NA	NA	1.0000	1.0000	NA
PANCREAS	IN	1	1	0	1	0	0.7182	0.7716				
adenoma, islet-cell							1.0000	0.7144	1.0000	0.4878	NA	1.0000
PITUITARY	IN	3	1	4	0	2	0.5731	0.9816				
adenoma, pars distalis	FA	1	0	0	0	0	0.3705	0.9833	0.6464	1.0000	0.8890	0.2444
SKIN (GROSS LESION)	MI	0	1	0	0	1	0.2826	0.5325				
fibrosarcoma							1.0000	1.0000	0.4742	NA	0.4462	0.4328
SKIN (GROSS LESION)	MI	0	0	1	0	0	0.8182	NA				
trichoepithelioma							0.3311	NA	NA	1.0000	1.0000	NA
SKIN, TREATED	MI	0	0	0	0	1	0.2072	NA				
keratoacanthoma							NA	NA	0.3077	NA	0.5610	0.5000
SKULL	IN	1	0	0	0	0	1.0000	1.0000				
sarcoma, osteogenic							1.0000	1.0000	1.0000	NA	NA	NA
SPLEEN	IN	0	0	1	0	0	0.2468	NA				
hemangiosarcoma	FA	0	0	1	0	1	0.0795	NA	0.2787	1.0000	0.8669	0.4486
THYMUS/REGION	IN	0	0	0	0	1	0.2072	NA				
hemangioma							NA	NA	0.3077	NA	0.5610	0.5000
THYMUS/REGION	IN	0	0	1	0	0	0.8182	NA				
lymphoma, thymic							0.7143	NA	NA	1.0000	1.0000	NA
THYROID	IN	0	0	1	0	0	0.5766	NA				
cystadenoma, follicular							0.2769	NA	NA	1.0000	1.0000	NA
THYROID	IN	0	0	1	0	1	0.1602	NA				
adenoma, follicular							0.3559	NA	0.3077	1.0000	0.7026	0.5000
TONGUE	IN	1	0	0	0	0	1.0000	1.0000				
carcinoma, squamous-cell							1.0000	1.0000	1.0000	NA	NA	NA
TRACHEA	IN	1	0	0	0	0	1.0000	1.0000				
papilloma							1.0000	1.0000	1.0000	NA	NA	NA
UTERUS	IN	1	1	2	1	3	0.0995	0.9030				
hemangioma	FA	1	0	0	0	0	0.5897	0.8225	0.2531	0.8752	0.3927	0.2489

Table 7.(cont.) Female Tumorigenicity

Analysis of Carcinogenic Potential in Female Mouse
Test of Dose-Response (Tumor) Positive Linear Trend and Pairwise Tests

ORGAN/TISSUE NAME AND TUMOR NAME	TUMOR TYPES	# Tumors C	V	L	M	H	TREND/ CV vs L	C vs V/ CV vs M	CV vs H	L vs M	L vs H	M vs H
UTERUS	IN	1	1	0	0	0	0.9548	0.7447				
hemangiosarcoma							1.0000	1.0000	1.0000	NA	NA	NA
UTERUS	IN	0	0	2	1	1	0.2574	NA				
leiomyoma							0.1207	0.3158	0.5313	0.8720	0.8017	0.8500
UTERUS	IN	0	0	1	0	0	0.5766	NA				
leiomyosarcoma							0.2769	NA	NA	1.0000	1.0000	NA
UTERUS	IN	3	4	4	6	1	0.8581	0.4143				
polyp(s),endometrial, stromal	FA	0	1	0	0	0	0.6394	0.3232	0.9234	0.4185	0.9165	0.9951
UTERUS	IN	0	0	0	0	1	0.3590	NA				
sarcoma, endoemtrial stromal							NA	NA	0.5313	NA	0.7000	0.7000
UTERUS	IN	0	0	0	1	0	0.4144	NA				
sarcoma, osteogenic (extra-osseus)							NA	0.3286	NA	0.5610	NA	1.0000
UTERUS	IN	0	1	1	0	1	0.3498	0.4615				
adenocarcinoma, endometrial							0.5593	1.0000	0.6289	1.0000	0.5484	0.3333
UTERUS	IN	1	0	0	0	1	0.3730	1.0000				
tumor, granular-cell							1.0000	1.0000	0.5240	NA	0.5610	0.5000
VAGINA	IN	0	0	1	0	0	0.5766	NA				
leiomyoma							0.2769	NA	NA	1.0000	1.0000	NA
VAGINA	IN	0	0	0	1	0	0.4144	NA				
schwannoma							NA	0.3286	NA	0.5610	NA	1.0000

APPEARS THIS WAY
ON ORIGINAL

Table 8. Males
Detailed listing of all tumors.

Note that technically, if one follows the Peto et al recommendations, skin tumors should be classed as "mortality independent" not as "incidental," and analyzed using life table methods, as with fatal tumors. Hence, in the analysis below such tumors are labeled as "fataL". This is just an artifact of the program used for the analysis.

Analysis of Carcinogenic Potential in Male Mouse
Test of Dose-Response (Tumor) Positive Linear Trend
Ted Guo, PH.D, CDER/FDA

Note: Dose Levels Included: CTRL VEH LOW MED HIGH (0 0.05 10 30 100)
Missing value in Tumor-Caused Death is treated as tumor not causing death
Tumor Type: IN: Incidental (nonfatal) tumor, FA: Fatal tumor.

ORGAN/TISSUE NAME / AND TUMOR NAME	(ORG#) (TMR#)	TUMOR TYPES	TIME STRATA	ROW NO.	2xC_CONTINGENCY -----TABLE-----	EXACT PROB	ASYMP TOTIC	ASYMP(CONTI NUITY CORR) =PR(STATISTIC.GE.OBSERVED)
ADRENAL GLANDS	(2)	IN	53-80	1	0 1 0 0 0	0.9781	0.8610	0.8632
adenoma, cortical	(8)	IN	53-80	2	5 10 12 14 19			
		IN	105-105	1	1 0 0 0 0			
		IN	105-105	2	19 22 18 9 6			
Spontaneous tumor pct: 2%	in ctrl. -	Total		-	1 1 0 0 0			
HARDERIAN GLANDS	(22)	IN	81-104	1	1 0 1 3 0	0.7899	0.7897	0.7911
adenoma	(6)	IN	81-104	2	20 13 16 18 13			
		IN	105-105	1	3 2 1 1 0			
		IN	105-105	2	17 20 17 8 6			
Spontaneous tumor pct: 8%	in ctrl. -	Total		-	4 2 2 4 0			
HEART	(24)	IN	81-104	1	0 1 0 0 0	0.7529	0.7654	0.7699
hemangiosarcoma	(22)	IN	81-104	2	21 12 17 21 13			
Spontaneous tumor pct: <= 1%	in ctrl. -	Total		-	0 1 0 0 0			
JEJUNUM	(26)	IN	105-105	1	0 0 1 0 0	0.4459	0.5613	0.5685
cystadenoma	(17)	IN	105-105	2	19 22 17 9 6			
Spontaneous tumor pct: <= 1%	in ctrl. -	Total		-	0 0 1 0 0			
LIVER	(28)	IN	53-80	1	0 3 1 2 3	0.2547	0.2553	0.2561
adenoma, hepatocellular	(10)	IN	53-80	2	5 7 11 12 16			
		IN	81-104	1	2 2 0 8 2			
		IN	81-104	2	19 11 17 13 11			
		IN	105-105	1	7 4 7 4 2			
		IN	105-105	2	13 18 11 5 4			
		FA	78	1	0 1 0 0 0			
		FA	78	2	43 39 37 30 21			
Spontaneous tumor pct: 18%	in ctrl. -	Total		-	9 10 8 14 7			
LIVER	(28)	IN	53-80	1	0 0 1 0 0	0.3966	0.4138	0.4156
carcinoma, hepatocellular	(15)	IN	53-80	2	5 11 11 14 19			
		IN	81-104	1	0 1 2 1 1			
		IN	81-104	2	19 12 15 19 12			
		IN	105-105	1	1 1 0 0 1			
		IN	105-105	2	19 21 18 9 5			
		FA	81	1	1 0 0 0 0			
		FA	81	2	40 35 35 30 19			
		FA	82	1	0 0 0 1 0			
		FA	82	2	39 34 34 27 18			
		FA	87	1	1 0 0 0 0			
		FA	87	2	36 31 29 22 15			
Spontaneous tumor pct: 6%	in ctrl. -	Total		-	3 2 3 2 2			

APPEARS THIS WAY
ON ORIGINAL

(Continued)

Analysis of Carcinogenic Potential in Male Mouse
Test of Dose-Response (Tumor) Positive Linear Trend

Note: Dose Levels Included: CTRL VEH LOW MED HIGH (0 0.05 10 30 100)
Missing value in Tumor-Caused Death is treated as tumor not causing death
Tumor Type: IN: Incidental (nonfatal) tumor, FA: Fatal tumor.

ORGAN/TISSUE NAME AND TUMOR NAME	(ORG#) (TMR#)	TUMOR TYPES	TIME STRATA	ROW NO.	2xC_CONTINGENCY -----TABLE-----					EXACT PROB =PR(STATISTIC.GE.OBSERVED)	ASYMP TOTIC	ASYMP(CONTI NUITY CORR)
LIVER	(28	) IN	105-105	1	1	1	0	0	0	0.9232	0.8274	0.8300
hemangioma	(21	) IN	105-105	2	19	21	18	9	6			
		FA	95	1	0	1	0	0	0			
		FA	95	2	29	24	24	19	11			
Spontaneous tumor pct: 2%	in ctrl. -	Total		-	1	2	0	0	0			
LIVER	(28	) IN	81-104	1	0	1	0	0	1	0.4236	0.4628	0.4650
hemangiosarcoma	(22	) IN	81-104	2	20	12	17	20	12			
		IN	105-105	1	1	1	1	1	0			
		IN	105-105	2	19	21	17	8	6			
		FA	69	1	1	0	0	0	0			
		FA	69	2	44	42	41	36	30			
		FA	92	1	1	0	0	0	0			
		FA	92	2	31	27	25	20	13			
		FA	104	1	0	0	0	1	0			
		FA	104	2	21	22	18	10	6			
Spontaneous tumor pct: 6%	in ctrl. -	Total		-	3	2	1	2	1			
LUNG	(29	) IN	53-80	1	0	0	1	0	0	0.3254	0.3585	0.3608
carcinoma, bronchiolo-alv	(14	) IN	53-80	2	5	11	11	14	19			
		IN	81-104	1	0	1	0	0	0			
		IN	81-104	2	21	12	17	20	13			
		IN	105-105	1	2	0	0	1	1			
		IN	105-105	2	18	22	18	8	5			
		FA	86	1	0	0	0	1	0			
		FA	86	2	38	31	30	25	17			
Spontaneous tumor pct: 4%	in ctrl. -	Total		-	2	1	1	2	1			
LUNG	(29	) IN	53-80	1	0	0	1	0	0	0.9944	0.9847	0.9849
adenoma, bronchiolo-alveo	(7	) IN	53-80	2	5	11	11	14	19			
		IN	81-104	1	3	2	1	2	0			
		IN	81-104	2	18	10	16	19	13			
		IN	105-105	1	3	4	4	0	0			
		IN	105-105	2	17	18	14	9	6			
		FA	91	1	0	1	0	0	0			
		FA	91	2	33	28	26	21	13			
Spontaneous tumor pct: 12%	in ctrl. -	Total		-	6	7	6	2	0			
LYMPH NODE, MESENTERIC	(34	) IN	81-104	1	0	0	0	0	1	0.1529	0.0136	0.0141
hemangioma	(21	) IN	81-104	2	21	13	17	21	12			
Spontaneous tumor pct: <= 1%	in ctrl. -	Total		-	0	0	0	0	1			

APPEARS THIS WAY
ON ORIGINAL

(Continued) Analysis of Carcinogenic Potential in Male Mouse
Test of Dose-Response (Tumor) Positive Linear Trend

Note: Dose Levels Included: CTRL VEH LOW MED HIGH (0 0.05 10 30 100)
Missing value in Tumor-Caused Death is treated as tumor not causing death
Tumor Type: IN: Incidental (nonfatal) tumor, FA: Fatal tumor.

ORGAN/TISSUE NAME AND TUMOR NAME	(ORG#) (TMR#)	TUMOR TYPES	TIME STRATA	ROW NO.	2xC_CONTINGENCY -----TABLE-----	EXACT PROB =PR(STATISTIC.GE.OBSERVED)	ASYMP TOTIC	ASYMP(CONTINUITY CORR)
LYMPHORETICULAR SYSTEM	(41	) IN	53-80	1	0 0 1 0 0	0.1621	0.1555	0.1564
lymphosarcoma, systemic	(28	) IN	53-80	2	5 10 10 12 19			
		IN	105-105	1	3 0 0 0 0			
		IN	105-105	2	17 22 18 9 6			
		FA	10	1	0 0 0 0 1			
		FA	10	2	50 50 49 48 48			
		FA	25	1	0 0 0 0 1			
		FA	25	2	49 50 49 48 47			
		FA	38	1	0 1 0 0 0			
		FA	38	2	47 49 48 45 43			
		FA	57	1	0 1 0 0 0			
		FA	57	2	46 44 47 42 37			
		FA	73	1	0 0 0 1 0			
		FA	73	2	44 42 39 32 27			
		FA	76	1	0 0 0 1 0			
		FA	76	2	43 41 38 31 21			
		FA	79	1	0 0 1 0 0			
		FA	79	2	42 38 36 30 20			
		FA	91	1	0 0 0 1 0			
		FA	91	2	33 29 26 20 13			
		FA	98	1	0 0 1 0 1			
		FA	98	2	26 23 21 17 8			
		FA	101	1	0 1 0 0 0			
		FA	101	2	22 22 18 14 6			
		FA	102	1	0 0 0 2 0			
		FA	102	2	21 22 18 12 6			
Spontaneous tumor pct: 6%	in ctrl. -	Total		-	3 3 3 5 3			
LYMPHORETICULAR SYSTEM	(41	) IN	105-105	1	0 1 2 1 0	0.3395	0.3691	0.3715
sarcoma, histiocytic	(37	) IN	105-105	2	20 21 16 8 6			
		FA	80	1	0 0 0 0 1			
		FA	80	2	41 35 36 30 19			
		FA	87	1	0 1 0 0 0			
		FA	87	2	37 30 29 22 15			
		FA	99	1	1 0 0 0 0			
		FA	99	2	25 23 18 15 7			
		FA	104	1	1 0 0 0 0			
		FA	104	2	20 22 18 11 6			
Spontaneous tumor pct: 4%	in ctrl. -	Total		-	2 2 2 1 1			
PANCREAS	(50	) IN	81-104	1	1 0 0 0 0	1.0000	0.8131	0.8161
adenoma, islet-cell	(11	) IN	81-104	2	20 13 17 21 13			
		IN	105-105	1	1 0 0 0 0			
		IN	105-105	2	19 22 18 9 6			
Spontaneous tumor pct: 4%	in ctrl. -	Total		-	2 0 0 0 0			
PENIS	(52	) IN	53-80	1	0 0 0 1 0	0.5410	0.5948	0.5994
hemangioma	(21	) IN	53-80	2	5 11 12 13 19			
Spontaneous tumor pct: <= 1%	in ctrl. -	Total		-	0 0 0 1 0			
PITUITARY	(53	) IN	53-80	1	0 0 1 0 0	0.7377	0.7640	0.7677
adenocarcinoma, pars dist	(5	) IN	53-80	2	5 11 11 14 19			
Spontaneous tumor pct: <= 1%	in ctrl. -	Total		-	0 0 1 0 0			

(Continued)

Analysis of Carcinogenic Potential in Male Mouse
Test of Dose-Response (Tumor) Positive Linear Trend

Note: Dose Levels Included: CTRL VEH LOW MED HIGH (0 0.05 10 30 100)
Missing value in Tumor-Caused Death is treated as tumor not causing death
Tumor Type: IN: Incidental (nonfatal) tumor, FA: Fatal tumor.

ORGAN/TISSUE NAME AND TUMOR NAME	(ORG#) (TMR#)	TUMOR TYPES	TIME STRATA	ROW NO.	2xC_CONTINGENCY -----TABLE-----	EXACT PROB =PR(STATISTIC.GE.OBSERVED)	ASYMP TOTIC	ASYMP(CONTINUITY CORR)
SKIN (GROSS LESION)	(59	) FA	105-105	1	0 0 1 0 0	0.4400	0.5588	0.5661
hemangioma	(21	) FA	105-105	2	20 22 17 9 6			
Spontaneous tumor pct: <= 1% in ctrl. - Total				-	0 0 1 0 0			
SKIN (GROSS LESION)	(59	) FA	105-105	1	0 1 0 0 0	0.7333	0.6967	0.7031
hemangiosarcoma	(22	) FA	105-105	2	20 21 18 9 6			
Spontaneous tumor pct: <= 1% in ctrl. - Total				-	0 1 0 0 0			
SPLEEN	(66	) IN	81-104	1	0 1 0 0 0	0.7529	0.7654	0.7699
hemangioma	(21	) IN	81-104	2	21 12 17 21 13			
Spontaneous tumor pct: <= 1% in ctrl. - Total				-	0 1 0 0 0			
SPLEEN	(66	) IN	81-104	1	0 0 0 1 0	0.0969	0.0689	0.0698
hemangiosarcoma	(22	) IN	81-104	2	21 13 16 20 11			
		IN	105-105	1	2 1 0 0 0			
		IN	105-105	2	18 21 18 9 6			
		FA	81	1	0 0 1 0 0			
		FA	81	2	41 35 34 30 19			
		FA	86	1	0 0 0 0 1			
		FA	86	2	38 31 30 26 16			
		FA	90	1	0 0 0 0 1			
		FA	90	2	35 29 27 21 14			
Spontaneous tumor pct: 4% in ctrl. - Total				-	2 1 1 1 2			
TESTES	(69	) IN	81-104	1	0 0 0 0 1	0.1529	0.0136	0.0141
hemangioma	(21	) IN	81-104	2	21 13 17 21 12			
Spontaneous tumor pct: <= 1% in ctrl. - Total				-	0 0 0 0 1			
TESTES	(69	) IN	105-105	1	1 0 1 0 0	0.6897	0.6820	0.6867
tumor, interstitial-cell	(44	) IN	105-105	2	19 22 17 9 6			
Spontaneous tumor pct: 2% in ctrl. - Total				-	1 0 1 0 0			
THYROID	(72	) IN	81-104	1	0 1 1 0 0	0.2165	0.1986	0.2011
adenoma, follicular	(9	) IN	81-104	2	21 12 16 21 13			
		IN	105-105	1	0 0 0 0 1			
		IN	105-105	2	20 22 18 9 5			
Spontaneous tumor pct: <= 1% in ctrl. - Total				-	0 1 1 0 1			
TRACHEA	(74	) IN	105-105	1	0 0 0 0 1	0.0800	0.0008	0.0008
polyp, inflammatory	(33	) IN	105-105	2	20 22 18 9 5			
Spontaneous tumor pct: <= 1% in ctrl. - Total				-	0 0 0 0 1			
URINARY BLADDER	(75	) IN	81-104	1	0 0 1 0 0	0.6071	0.6697	0.6750
tumor, mesenchymal	(45	) IN	81-104	2	21 12 16 21 13			
Spontaneous tumor pct: <= 1% in ctrl. - Total				-	0 0 1 0 0			

APPEARS THIS WAY
ON ORIGINAL

Table 9. Females
Detailed listing of all tumors.

Analysis of Carcinogenic Potential in Female Mouse
Test of Dose-Response (Tumor) Positive Linear Trend
Ted Guo, PH.D, CDER/FDA

Note: Dose Levels Included: CTRL VEH LOW MED HIGH (0 0.05 10 30 100)
Missing value in Tumor-Caused Death is treated as tumor not causing death
Tumor Type: IN: Incidental (nonfatal) tumor, FA: Fatal tumor.

ORGAN/TISSUE NAME AND TUMOR NAME	(ORG#) (TMR#)	TUMOR TYPES	TIME STRATA	ROW NO.	2xC_CONTINGENCY -----TABLE-----	EXACT PROB =PR(STATISTIC.GE.OBSERVED)	ASYMP TOTIC	ASYMP(CONTINUITY CORR)
CECUM	(10)	IN	105-105	1	0 0 1 0 0	0.5730	0.6325	0.6386
mastocytoma	(29)	IN	105-105	2	17 21 20 20 10			
Spontaneous tumor pct: <= 1% in ctrl. -		Total	-		0 0 1 0 0			
CERVIX	(11)	IN	81-104	1	0 0 0 1 0	0.4144	0.4851	0.4903
hemangioma	(21)	IN	81-104	2	24 23 18 22 23			
Spontaneous tumor pct: <= 1% in ctrl. -		Total	-		0 0 0 1 0			
CERVIX	(11)	IN	53-80	1	1 0 0 0 0	0.9882	0.9392	0.9399
leiomyoma	(24)	IN	53-80	2	6 6 6 6 14			
		IN	81-104	1	0 1 0 0 0			
		IN	81-104	2	24 22 18 23 23			
		IN	105-105	1	2 0 1 0 0			
		IN	105-105	2	15 21 20 20 10			
Spontaneous tumor pct: 6% in ctrl. -		Total	-		3 1 1 0 0			
CERVIX	(11)	IN	81-104	1	1 0 1 1 2	0.0796	0.0655	0.0662
polyp, stromal	(34)	IN	81-104	2	23 23 17 22 21			
		IN	105-105	1	1 0 1 1 1			
		IN	105-105	2	16 21 20 19 9			
Spontaneous tumor pct: 4% in ctrl. -		Total	-		2 0 2 2 3			
CERVIX	(11)	IN	81-104	1	0 0 0 0 1	0.0407	0.0036	0.0037
sarcoma, stromal	(40)	IN	81-104	2	24 23 18 23 22			
		FA	53	1	0 0 0 0 1			
		FA	53	2	48 50 45 49 46			
Spontaneous tumor pct: <= 1% in ctrl. -		Total	-		0 0 0 0 2			(P<0.025)
CERVIX	(11)	IN	53-80	1	0 0 1 0 0	0.6667	0.7647	0.7681
schwannoma	(41)	IN	53-80	2	7 6 5 6 14			
Spontaneous tumor pct: <= 1% in ctrl. -		Total	-		0 0 1 0 0			
CERVIX	(11)	IN	81-104	1	0 1 0 0 0	0.7838	0.7724	0.7763
tumor, granular-cell	(43)	IN	81-104	2	24 22 18 23 23			
Spontaneous tumor pct: <= 1% in ctrl. -		Total	-		0 1 0 0 0			
CERVIX	(11)	IN	81-104	1	1 0 0 0 0	1.0000	0.7728	0.7767
adenoma	(6)	IN	81-104	2	23 23 18 23 23			
Spontaneous tumor pct: 2% in ctrl. -		Total	-		1 0 0 0 0			
ADRENAL GLANDS	(2)	IN	105-105	1	0 0 0 2 0	0.2617	0.3268	0.3310
pheochromocytoma	(31)	IN	105-105	2	17 21 21 18 10			
Spontaneous tumor pct: <= 1% in ctrl. -		Total	-		0 0 0 2 0			
ADRENAL GLANDS	(2)	IN	81-104	1	0 0 0 1 0	0.1354	0.1031	0.1044
adenoma, cortical	(8)	IN	81-104	2	24 23 18 22 23			
		IN	105-105	1	0 0 1 2 1			
		IN	105-105	2	17 21 20 18 9			
Spontaneous tumor pct: <= 1% in ctrl. -		Total	-		0 0 1 3 1			

(Continued) Analysis of Carcinogenic Potential in Female House
Test of Dose-Response (Tumor) Positive Linear Trend

Note: Dose Levels Included: CTRL VEH LOW MED HIGH (0 0.05 10 30 100)
Missing value in Tumor-Caused Death is treated as tumor not causing death
Tumor Type: IN: Incidental (nonfatal) tumor, FA: Fatal tumor.

ORGAN/TISSUE NAME AND TUMOR NAME	(ORG#) (TMR#)	TUMOR TYPES	TIME STRATA	ROW NO.	2xC_CONTINGENCY -----TABLE-----	EXACT PROB =PR(STATISTIC.GE.OBSERVED)	ASYMP TOTIC	ASYMP(CONTINUITY CORR)
HARDERIAN GLANDS	(22)	IN	53-80	1	0 0 1 0 0	0.6667	0.7647	0.7681
adenocarcinoma	(3)	IN	53-80	2	7 6 5 6 14			
Spontaneous tumor pct: <= 1% in ctrl.		-	Total	-	0 0 1 0 0			
HARDERIAN GLANDS	(22)	IN	81-104	1	1 1 0 1 2	0.3760	0.3953	0.3972
adenoma	(6)	IN	81-104	2	23 22 18 22 21			
		IN	105-105	1	1 2 0 0 0			
		IN	105-105	2	16 19 21 20 10			
Spontaneous tumor pct: 4% in ctrl.		-	Total	-	2 3 0 1 2			
LIVER	(28)	IN	81-104	1	2 0 0 1 1	0.2531	0.2497	0.2512
adenoma, hepatocellular	(10)	IN	81-104	2	22 23 18 22 22			
		IN	105-105	1	2 0 1 3 1			
		IN	105-105	2	15 21 20 17 9			
Spontaneous tumor pct: 8% in ctrl.		-	Total	-	4 0 1 4 2			
LIVER	(28)	IN	53-80	1	0 0 0 1 0	0.2082	0.2250	0.2267
carcinoma, hepatocellular	(15)	IN	53-80	2	7 6 6 5 14			
		IN	81-104	1	0 1 1 0 1			
		IN	81-104	2	24 22 17 23 21			
		IN	105-105	1	0 1 0 0 0			
		IN	105-105	2	17 20 21 20 10			
		FA	81	1	0 0 0 0 1			
		FA	81	2	41 44 39 43 32			
Spontaneous tumor pct: <= 1% in ctrl.		-	Total	-	0 2 1 1 2			
LIVER	(28)	IN	81-104	1	1 0 0 1 0	0.1824	0.1833	0.1852
hemangioma	(21)	IN	81-104	2	23 23 18 22 23			
		IN	105-105	1	0 0 0 1 1			
		IN	105-105	2	17 21 21 19 9			
Spontaneous tumor pct: 2% in ctrl.		-	Total	-	1 0 0 2 1			
LIVER	(28)	IN	81-104	1	1 0 2 1 1	0.6306	0.6472	0.6487
hemangiosarcoma	(22)	IN	81-104	2	22 22 16 22 22			
		IN	105-105	1	0 2 0 4 0			
		IN	105-105	2	17 19 21 16 10			
		FA	81	1	0 1 0 0 0			
		FA	81	2	41 43 39 43 33			
		FA	98	1	1 0 0 0 0			
		FA	98	2	22 30 27 26 19			
Spontaneous tumor pct: 4% in ctrl.		-	Total	-	2 3 2 5 1			
LIVER	(28)	IN	81-104	1	0 0 0 1 0	0.4144	0.4851	0.4903
sarcoma, NOS	(35)	IN	81-104	2	24 23 18 22 23			
Spontaneous tumor pct: <= 1% in ctrl.		-	Total	-	0 0 0 1 0			
LUNG	(29)	IN	53-80	1	0 0 1 0 1	0.1776	0.1846	0.1857
carcinoma, bronchiolo-alv	(14)	IN	53-80	2	7 6 5 6 13			
		IN	81-104	1	0 0 0 0 2			
		IN	81-104	2	24 19 17 23 21			
		IN	105-105	1	1 0 1 0 1			
		IN	105-105	2	16 21 20 20 9			
		FA	94	1	0 1 0 0 0			
		FA	94	2	26 34 30 30 20			

APPEARS THIS WAY
ON ORIGINAL

(Continued) Analysis of Carcinogenic Potential in Female Mouse
Test of Dose-Response (Tumor) Positive Linear Trend

Note: Dose Levels Included: CTRL VEH LOW MED HIGH (0 0.05 10 30 100)
Tumor Type: IN: Incidental (nonfatal) tumor, FA: Fatal tumor.

ORGAN/TISSUE NAME AND TUMOR NAME	(ORG#) (TMR#)	TUMOR TYPES	TIME STRATA	ROW NO.	2xC_CONTINGENCY -----TABLE-----					EXACT PROB =PR(STATISTIC.GE.OBSERVED)	ASYMP TOTIC	ASYMP(CONTI NUITY CORR)
		FA	98	1	0	2	0	0	0			
		FA	98	2	23	28	27	26	19			
		FA	101	1	0	1	0	0	0			
		FA	101	2	20	23	25	22	16			
		FA	103	1	0	0	1	0	0			
		FA	103	2	18	22	22	21	12			
Spontaneous tumor pct: 2% in ctrl. - Total				-	1	4	3	0	4			
LUNG	(29)	IN	0-52	1	0	0	2	0	0	0.8902	0.8888	0.8894
adenoma, bronchiolo-alveo	(7)	IN	0-52	2	2	0	3	1	3			
		IN	53-80	1	0	0	0	1	1			
		IN	53-80	2	7	6	6	5	13			
		IN	81-104	1	2	2	1	0	1			
		IN	81-104	2	21	21	17	23	22			
		IN	105-105	1	4	5	3	2	1			
		IN	105-105	2	13	16	18	18	9			
		FA	96	1	1	0	0	0	0			
		FA	96	2	23	31	28	28	20			
Spontaneous tumor pct: 14% in ctrl. - Total				-	7	7	6	3	3			
LYMPH NODE, MESENTERIC	(34)	IN	81-104	1	0	0	0	0	1	0.1602	0.1059	0.1078
hemangioma	(21)	IN	81-104	2	24	23	18	23	22			
		IN	105-105	1	0	0	1	0	0			
		IN	105-105	2	17	21	20	20	10			
Spontaneous tumor pct: <= 1% in ctrl. - Total				-	0	0	1	0	1			
LYMPHORETICULAR SYSTEM	(41)	IN	0-52	1	0	0	1	0	0	0.9809	0.9417	0.9423
lymphoma, thymic	(27)	IN	0-52	2	2	0	3	1	3			
		IN	105-105	1	3	0	1	0	0			
		IN	105-105	2	14	21	20	20	10			
		FA	44	1	0	0	1	0	0			
		FA	44	2	49	50	46	49	47			
		FA	99	1	0	1	0	0	0			
		FA	99	2	22	27	26	25	19			
Spontaneous tumor pct: 6% in ctrl. - Total				-	3	1	3	0	0			
LYMPHORETICLLAR SYSTEM	(41)	IN	0-52	1	0	0	0	0	1	0.4938	0.5085	0.5094
lymphosarcoma, systemic	(28)	IN	0-52	2	2	0	5	1	2			
		IN	81-104	1	1	1	1	2	0			
		IN	81-104	2	17	19	14	18	21			
		IN	105-105	1	3	4	3	1	1			
		IN	105-105	2	14	17	18	19	9			
		FA	57	1	0	0	0	1	0			
		FA	57	2	45	50	45	48	46			
		FA	66	1	0	0	0	0	1			
		FA	66	2	44	48	43	46	42			
		FA	69	1	0	0	0	0	1			
		FA	69	2	44	48	43	46	40			
		FA	72	1	1	0	0	0	1			
		FA	72	2	43	47	42	46	37			
		FA	80	1	0	1	0	0	0			
		FA	80	2	42	45	39	44	34			
		FA	84	1	0	1	0	0	0			
		FA	84	2	39	41	36	39	29			
		FA	87	1	0	0	0	0	1			
		FA	87	2	36	40	36	38	26			

APPEARS THIS WAY
ON ORIGINAL

(Continued) Analysis of Carcinogenic Potential in Female House
Test of Dose-Response (Tumor) Positive Linear Trend

Note: Dose Levels Included: CTRL VEH LOW MED HIGH (0 0.05 10 30 100)
Tumor Type: IN: Incidental (nonfatal) tumor, FA: Fatal tumor.

ORGAN/TISSUE NAME AND TUMOR NAME	(ORG#) (TMR#)	TUMOR TYPES	TIME STRATA	ROW NO.	2xC_CONTINGENCY -----TABLE-----	EXACT PROB	ASYMP TOTIC	ASYMP(CONTINUITY CORR) =PR(STATISTIC.GE.OBSERVED)
		FA	89	1	1 0 0 0 0			
		FA	89	2	33 39 35 36 25			
		FA	93	1	2 1 0 0 0			
		FA	93	2	28 36 30 33 23			
		FA	94	1	1 0 0 0 0			
		FA	94	2	25 35 30 30 20			
		FA	95	1	0 1 0 0 0			
		FA	95	2	25 31 28 29 20			
		FA	96	1	0 0 0 0 1			
		FA	96	2	24 31 28 28 19			
		FA	98	1	0 0 1 0 0			
		FA	98	2	23 30 26 26 19			
		FA	99	1	1 0 1 1 0			
		FA	99	2	21 28 25 24 19			
		FA	101	1	0 0 1 1 0			
		FA	101	2	20 24 24 21 16			
		FA	103	1	0 0 0 1 0			
		FA	103	2	18 22 23 20 12			
		FA	104	1	1 0 0 0 0			
		FA	104	2	17 21 21 20 10			
Spontaneous tumor pct: 22% in ctrl. - Total				-	11 9 7 7 7			
LYMPHORETICULAR SYSTEM	(41)	IN	81-104	1	0 0 0 0 2	0.3225	0.3413	0.3426
sarcoma, histiocytic	(37)	IN	81-104	2	22 19 15 22 20			
		IN	105-105	1	0 1 0 0 1			
		IN	105-105	2	17 20 21 20 9			
		FA	62	1	0 0 0 1 0			
		FA	62	2	45 48 45 46 44			
		FA	65	1	0 0 1 0 0			
		FA	65	2	44 48 43 46 43			
		FA	73	1	1 0 0 0 0			
		FA	73	2	42 47 41 46 37			
		FA	81	1	0 1 0 0 0			
		FA	81	2	41 43 39 43 33			
		FA	82	1	0 0 1 0 0			
		FA	82	2	41 42 37 41 31			
		FA	87	1	0 1 0 0 0			
		FA	87	2	36 39 36 38 27			
		FA	88	1	0 0 0 1 0			
		FA	88	2	36 39 36 37 26			
		FA	89	1	0 0 1 0 0			
		FA	89	2	34 39 34 36 25			
		FA	90	1	1 0 0 0 1			
		FA	90	2	32 39 34 36 24			
		FA	91	1	0 0 1 0 0			
		FA	91	2	32 38 33 36 24			
		FA	94	1	0 1 0 0 0			
		FA	94	2	26 34 30 30 20			
		FA	96	1	0 1 0 0 0			
		FA	96	2	24 30 28 28 20			
		FA	102	1	1 0 0 0 0			
		FA	102	2	19 23 23 21 14			
Spontaneous tumor pct: 6% in ctrl. - Total				-	3 5 4 2 4			

APPEARS THIS WAY
ON ORIGINAL

Analysis of Carcinogenic Potential in Female Mouse
Test of Dose-Response (Tumor) Positive Linear Trend

Note: Dose Levels Included: CTRL VEH LOW MED HIGH (0 0.05 10 30 100)
Tumor Type: IN: Incidental (nonfatal) tumor, FA: Fatal tumor.

ORGAN/TISSUE NAME AND TUMOR NAME	(ORG#) (TMR#)	TUMOR TYPES	TIME STRATA	ROW NO.	2xC_CONTINGENCY -----TABLE-----	EXACT PROB =PR(STATISTIC.GE.OBSERVED)	ASYMP TOTIC	ASYMP(CONTINUITY CORR)
MAMMARY GLAND/REGION	(42)	IN	81-104	1	0 0 0 1 0	0.4144	0.4851	0.4903
adenoacanthoma	(2)	IN	81-104	2	24 23 18 22 23			
Spontaneous tumor pct: <= 1% in ctrl.		-	Total	-	0 0 0 1 0			
MAMMARY GLAND/REGION	(42)	IN	81-104	1	1 2 2 1 1	0.8023	0.8008	0.8021
adenocarcinoma	(3)	IN	81-104	2	23 21 16 22 22			
		IN	105-105	1	2 0 0 0 0			
		IN	105-105	2	15 21 21 20 10			
Spontaneous tumor pct: 6% in ctrl.		-	Total	-	3 2 2 1 1			
MAMMARY GLAND/REGION	(42)	IN	105-105	1	1 0 0 0 0	1.0000	0.7471	0.7523
adenoma	(6)	IN	105-105	2	16 21 21 20 10			
Spontaneous tumor pct: 2% in ctrl.		-	Total	-	1 0 0 0 0			
OVARIES	(49)	IN	81-104	1	0 1 0 0 0	0.9621	0.8902	0.8916
adenoma, papillary	(12)	IN	81-104	2	24 22 18 23 23			
		IN	105-105	1	2 0 1 0 0			
		IN	105-105	2	15 21 20 20 10			
Spontaneous tumor pct: 4% in ctrl.		-	Total	-	2 1 1 0 0			
OVARIES	(49)	IN	105-105	1	0 2 0 3 0	0.4977	0.5692	0.5721
cystadenoma	(17)	IN	105-105	2	17 19 21 17 10			
Spontaneous tumor pct: <= 1% in ctrl.		-	Total	-	0 2 0 3 0			
OVARIES	(49)	IN	105-105	1	1 0 0 0 0	1.0000	0.7471	0.7523
granulosa-theca-cell tumo	(20)	IN	105-105	2	16 21 21 20 10			
Spontaneous tumor pct: 2% in ctrl.		-	Total	-	1 0 0 0 0			
OVARIES	(49)	IN	81-104	1	0 0 0 0 1	0.2215	0.2618	0.2640
hemangioma	(21)	IN	81-104	2	24 22 18 23 22			
		IN	105-105	1	0 1 0 2 0			
		IN	105-105	2	17 20 21 18 10			
		FA	93	1	0 1 0 0 0			
		FA	93	2	30 36 30 33 23			
Spontaneous tumor pct: <= 1% in ctrl.		-	Total	-	0 2 0 2 1			
OVARIES	(49)	IN	53-80	1	0 0 1 0 0	0.6667	0.7647	0.7681
hemangiosarcoma	(22)	IN	53-80	2	7 6 5 6 14			
Spontaneous tumor pct: <= 1% in ctrl.		-	Total	-	0 0 1 0 0			
OVARIES	(49)	IN	81-104	1	0 1 0 0 0	0.2324	0.1480	0.1504
luteoma	(26)	IN	81-104	2	24 22 18 23 23			
		IN	105-105	1	0 0 0 0 1			
		IN	105-105	2	17 21 21 20 9			
Spontaneous tumor pct: <= 1% in ctrl.		-	Total	-	0 1 0 0 1			
OVARIES	(49)	IN	105-105	1	0 0 1 0 0	0.5730	0.6325	0.6386
tumor, tubulo-stromal	(46)	IN	105-105	2	17 21 20 20 10			
Spontaneous tumor pct: <= 1% in ctrl.		-	Total	-	0 0 1 0 0			
PANCREAS	(50)	IN	81-104	1	1 0 0 0 0	0.7182	0.7523	0.7551
adenoma, islet-cell	(11)	IN	81-104	2	23 23 18 23 23			
		IN	105-105	1	0 1 0 1 0			
		IN	105-105	2	17 20 21 19 10			
Spontaneous tumor pct: 2% in ctrl.		-	Total	-	1 1 0 1 0			

Analysis of Carcinogenic Potential in Female Mouse
Test of Dose-Response (Tumor) Positive Linear Trend

Note: Dose Levels Included: CTRL VEH LOW MED HIGH (0 0.05 10 30 100)
Tumor Type: IN: Incidental (nonfatal) tumor, FA: Fatal tumor.

ORGAN/TISSUE NAME AND TUMOR NAME	(ORG#) (TMR#)	TUMOR TYPES	TIME STRATA	ROW NO.	2xC_CONTINGENCY -----TABLE-----	EXACT PROB =PR(STATISTIC.GE.OBSERVED)	ASYMP TOTIC	ASYMP(CONTINUITY CORR)
PITUITARY	(53)	IN	81-104	1	0 0 2 0 2	0.5587	0.5713	0.5731
adenoma, pars distalis	(13)	IN	81-104	2	23 23 16 23 21			
		IN	105-105	1	3 1 2 0 0			
		IN	105-105	2	14 20 19 20 10			
		FA	102	1	1 0 0 0 0			
		FA	102	2	19 23 23 21 14			
Spontaneous tumor pct: 8%	in ctrl. - Total			-	4 1 4 0 2			
SKIN (GROSS LESION)	(59)	FA	81-104	1	0 1 0 0 1	0.2826	0.2124	0.2151
fibrosarcoma	(19)	FA	81-104	2	24 22 18 23 22			
Spontaneous tumor pct: <= 1%	in ctrl. - Total			-	0 1 0 0 1			
SKIN (GROSS LESION)	(59)	FA	0-52	1	0 0 1 0 0	0.8182	0.7265	0.7306
trichoepithelioma	(42)	FA	0-52	2	2 0 4 1 3			
Spontaneous tumor pct: <= 1%	in ctrl. - Total			-	0 0 1 0 0			
SKIN, TREATED	(61)	FA	81-104	1	0 0 0 0 1	0.2072	0.0307	0.0316
keratoacanthoma	(23)	FA	81-104	2	24 23 18 23 22			
Spontaneous tumor pct: <= 1%	in ctrl. - Total			-	0 0 0 0 1			
SKULL	(63)	IN	81-104	1	1 0 0 0 0	1.0000	0.7728	0.7767
sarcoma, osteogenic	(38)	IN	81-104	2	23 23 18 23 23			
Spontaneous tumor pct: 2%	in ctrl. - Total			-	1 0 0 0 0			
SPLEEN	(66)	IN	81-104	1	0 0 1 0 0	0.2586	0.2442	0.2468
hemangiosarcoma	(22)	IN	81-104	2	24 23 16 23 23			
		FA	80	1	0 0 0 0 1			
		FA	80	2	42 46 39 44 33			
		FA	91	1	0 0 1 0 0			
		FA	91	2	32 38 33 36 24			
Spontaneous tumor pct: <= 1%	in ctrl. - Total			-	0 0 2 0 1			
THYMUS/REGION	(71)	IN	81-104	1	0 0 0 0 1	0.2072	0.0307	0.0316
hemangioma	(21)	IN	81-104	2	24 23 18 23 22			
Spontaneous tumor pct: <= 1%	in ctrl. - Total			-	0 0 0 0 1			
THYMUS/REGION	(71)	IN	0-52	1	0 0 1 0 0	0.8182	0.7265	0.7306
lymphoma, thymic	(27)	IN	0-52	2	2 0 4 1 3			
Spontaneous tumor pct: <= 1%	in ctrl. - Total			-	0 0 1 0 0			
THYROID	(72)	IN	81-104	1	0 0 1 0 0	0.5766	0.6866	0.6912
cystadenoma, follicular	(18)	IN	81-104	2	24 23 17 23 23			
Spontaneous tumor pct: <= 1%	in ctrl. - Total			-	0 0 1 0 0			
THYROID	(72)	IN	81-104	1	0 0 0 0 1	0.1602	0.1059	0.1078
adenoma, follicular	(9)	IN	81-104	2	24 23 18 23 22			
		IN	105-105	1	0 0 1 0 0			
		IN	105-105	2	17 21 20 20 10			
Spontaneous tumor pct: <= 1%	in ctrl. - Total			-	0 0 1 0 1			
TONGUE	(73)	IN	53-80	1	1 0 0 0 0	1.0000	0.8281	0.8309
carcinoma, squamous-cell	(16)	IN	53-80	2	6 6 6 6 14			
Spontaneous tumor pct: 2%	in ctrl. - Total			-	1 0 0 0 0			

(Continued) Analysis of Carcinogenic Potential in Female Mouse
Test of Dose-Response (Tumor) Positive Linear Trend

Note: Dose Levels Included: CTRL VEH LOW MED HIGH (0 0.05 10 30 100)
Tumor Type: IN: Incidental (nonfatal) tumor, FA: Fatal tumor.

ORGAN/TISSUE NAME AND TUMOR NAME	(ORG#) (TMR#)	TUMOR TYPES	TIME STRATA	ROW NO.	2xC_CONTINGENCY -----TABLE-----					EXACT PROB =PR(STATISTIC.GE.OBSERVED)	ASYMP TOTIC	ASYMP(CONTI NUITY CORR)
TRACHEA	(74	IN	81-104	1	1	0	0	0	0	1.0000	0.7728	0.7767
papilloma	(30	IN	81-104	2	23	23	18	23	23			
Spontaneous tumor pct: 2% in ctrl. - Total				-	1	0	0	0	0			
UTERUS	(76	IN	53-80	1	0	0	0	0	1	0.1170	0.0987	0.0995
hemangioma	(21	IN	53-80	2	7	6	6	6	13			
		IN	81-104	1	0	0	0	0	1			
		IN	81-104	2	24	23	18	23	22			
		IN	105-105	1	1	1	2	1	1			
		IN	105-105	2	16	20	19	19	9			
		FA	46	1	1	0	0	0	0			
		FA	46	2	48	50	45	49	47			
Spontaneous tumor pct: 4% in ctrl. - Total				-	2	1	2	1	3			
UTERUS	(76	IN	81-104	1	1	1	0	0	0	0.9548	0.8558	0.8579
hemangiosarcoma	(22	IN	81-104	2	23	22	18	23	23			
Spontaneous tumor pct: 2% in ctrl. - Total				-	1	1	0	0	0			
UTERUS	(76	IN	53-80	1	0	0	0	1	0	0.2574	0.2973	0.2997
leiomyoma	(24	IN	53-80	2	7	6	6	5	14			
		IN	81-104	1	0	0	0	0	1			
		IN	81-104	2	24	23	18	23	22			
		IN	105-105	1	0	0	2	0	0			
		IN	105-105	2	17	21	19	20	10			
Spontaneous tumor pct: <= 1% in ctrl. - Total				-	0	0	2	1	1			
UTERUS	(76	IN	81-104	1	0	0	1	0	0	0.5766	0.6866	0.6912
leiomyosarcoma	(25	IN	81-104	2	24	23	17	23	23			
Spontaneous tumor pct: <= 1% in ctrl. - Total				-	0	0	1	0	0			
UTERUS	(76	IN	53-80	1	0	0	0	1	0	0.8581	0.8574	0.8581
polyp(s), endometrial, st	(32	IN	53-80	2	7	6	6	5	14			
		IN	81-104	1	1	2	0	4	1			
		IN	81-104	2	23	20	18	19	22			
		IN	105-105	1	2	2	4	1	0			
		IN	105-105	2	15	19	17	19	10			
		FA	103	1	0	1	0	0	0			
		FA	103	2	18	21	23	21	12			
Spontaneous tumor pct: 6% in ctrl. - Total				-	3	5	4	6	1			
UTERUS	(76	IN	53-80	1	0	0	0	0	1	0.3590	0.0961	0.0981
sarcoma, endoemtrial stro	(36	IN	53-80	2	7	6	6	6	13			
Spontaneous tumor pct: <= 1% in ctrl. - Total				-	0	0	0	0	1			
UTERUS	(76	IN	81-104	1	0	0	0	1	0	0.4144	0.4851	0.4903
sarcoma, osteogenic (extr	(39	IN	81-104	2	24	23	18	22	23			
Spontaneous tumor pct: <= 1% in ctrl. - Total				-	0	0	0	1	0			
UTERUS	(76	IN	53-80	1	0	1	0	0	0	0.3498	0.3295	0.3324
adenocarcinoma, endometri	(4	IN	53-80	2	7	5	6	6	14			
		IN	105-105	1	0	0	1	0	1			
		IN	105-105	2	17	21	20	20	9			
Spontaneous tumor pct: <= 1% in ctrl. - Total				-	0	1	1	0	1			

(Continued) Analysis of Carcinogenic Potential in Female Mouse
Test of Dose-Response (Tumor) Positive Linear Trend

Note: Dose Levels Included: CTRL VEH LOW MED HIGH (0 0.05 10 30 100)
Tumor Type: IN: Incidental (nonfatal) tumor, FA: Fatal tumor.

ORGAN/TISSUE NAME AND TUMOR NAME	(ORG#) (TMR#)	TUMOR TYPES	TIME STRATA	ROW NO.	2xC_CONTINGENCY -----TABLE-----	EXACT PROB	ASYMP TOTIC	ASYMP(CONTINUITY CORR) =PR(STATISTIC.GE.OBSERVED)
UTERUS	(76)	IN	81-104	1	1 0 0 0 1	0.3730	0.2127	0.2154
tumor, granular-cell	(43)	IN	81-104	2	23 23 18 23 22			
Spontaneous tumor pct: 2%	in ctrl. -	Total		-	1 0 0 0 1			
VAGINA	(77)	IN	81-104	1	0 0 1 0 0	0.5766	0.6866	0.6912
leiomyoma	(24)	IN	81-104	2	24 23 17 23 23			
Spontaneous tumor pct: <= 1%	in ctrl. -	Total		-	0 0 1 0 0			
VAGINA	(77)	IN	81-104	1	0 0 0 1 0	0.4144	0.4851	0.4903
schwannoma	(41)	IN	81-104	2	24 23 18 22 23			
Spontaneous tumor pct: <= 1%	in ctrl. -	Total		-	0 0 0 1 0			

APPEARS THIS WAY
ON ORIGINAL