APPENDIX XXIV

Statistical Analysis of Histopathology – Females and Males

EFFECT OF OXYBENZONE ON FERTILITY AND EARLY EMBRYONIC DEVELOPMENT IN SPRAGUE-DAWLEY RATS (SEGMENT I)

PI: AMY INSELMAN

NCTR PROTOCOL NUMBER 2186.01

STUDY E0218602

STATISTICAL ANALYSIS OF THE TREATMENT EFFECTS OF OXYBENZONE ON PATHOLOGY

STATISTICAL REPORT

PREPARED BY

BETH E. JULIAR
DIVISION OF BIOINFORMATICS AND BIOSTATISTICS

FOR

NATIONAL CENTER FOR TOXICOLOGICAL RESEARCH
3900 NCTR ROAD
JEFFERSON, ARKANSAS 72079

Signatures and Dates

The report was prepared by: *Date:*

[redacted]

Beth E. Juliar, M.A., M.S., Division of Bioinformatics and Biostatistics/NCTR/FDA

The report was subsequently reviewed and edited by: *Date:*

[redacted]

Paul Felton, M.S., Division of Bioinformatics and Biostatistics /NCTR/FDA

Statistical Team Leader: *Date:*

[redacted]

Paul Felton, M.S., Division of Bioinformatics and Biostatistics /NCTR/FDA

Table of Contents

1. Objectives of the Analysis

1.1 Project Objectives

The objective of the study is to examine the reproductive toxicity of oxybenzone in male and female rats and is designed to focus specifically on fertility and early embryonic development to implantation [ICH Guideline S5(R2) 4.1.1]. An additional objective is to compare the results of a typical Segment I, II, III study design with results from a modified one-generation study proposed by the NTP.

1.2 Analysis Objectives

The goal of this analysis is to determine the effects of oxybenzone on male and female pathology.

2. Experimental Design

A total of 262 rats were to be requested for this study. Of this number, 125 male rats were to be requested along with 125 female rats. Males were to be approximately 5-7 weeks old when delivered to the NCTR, and females were to be approximately 9-11 weeks of age when delivered. All males were to be delivered in one shipment, and all females were to be delivered in a separate shipment. After a two week quarantine period the animals were to be weighed and allocated to the study.

The test article in this study is 2-hydroxy-4-methoxybenzophenone (synonyms: HMB, benzophenone-3, oxybenzone). The animals were to be divided into five treatment groups with 25 male and 25 female rats assigned to each group. The treatment groups were to be four oxybenzone dose levels 0 ppm (control), 3000 ppm, 10,000 ppm, and 30,000 ppm and one estrogen ethinyl estradiol (EE2) 0.05 ppm treatment.

Males were to be dosed for 10 weeks and females for approximately 2 weeks prior to mating. Dosing was to continue until gestational day (GD) 6 for all animals. From GD 6 to GD 15, dams were to receive control chow. All dams were to be sacrificed on GD 15; males were to be sacrificed soon after breeding (approximately GD 6).

All animals were to be housed in pairs in cages prior to breeding. For breeding, males and females were to be housed one male: one female for up to 15 days or until animals have mated. Males and females were to be housed individually upon indication of mating (GD 0) until the time of sacrifice.

For the males, histopathology analysis was to be conducted on the right testis, right epididymis, dorsolateral prostate, ventral prostate, seminal vesicles (with coagulating glands), retained nipples (if noted), mammary glands, adrenal, pituitary, and thyroid glands. For females, histopathology analysis was to be conducted on the mammary glands as well as the adrenal, pituitary, and thyroid glands.

3. Statistical Methods

The exact Cochran-Armitage trend test was used to test for a trend in nonneoplastic incidence with dose, and Fisher's exact test was used to compare incidences between treated groups and the control group (*CAFE*). Tests for trend and comparisons for treated groups to control were performed as one-sided tests.

For the analysis of lesion severity scores, the Jonckheere-Terpstra (*JT*) test was used to test for a trend in lesion severity with dose. Shirley's method, modified by Williams, (*JT-SW*) was used to compare lesion severity for treated groups to the control group.

Tests of trend were performed for the oxybenzone and control treatment groups.

References

Agresti, A. *Categorical Data Analysis*, Second Edition, New York: John Wiley & Sons (2002).

Jonckheere, AR. "A distribution-free k-sample test against ordered alternatives." *Biometrika* 41 (1954): 133-145.

Shirley, Eryl. "A non-parametric equivalent of Williams' test for contrasting increasing dose levels of a treatment." *Biometrics* 33 (1977): 386-389.

Terpstra, TJ. "The asymptotic normality and consistency of Kendall's test against trend when ties are present in one ranking." *Indagationes Mathematicæ* 14 (1952): 327-333.

Williams, DA. "A note on Shirley's nonparametric test for comparing several dose levels with a zero-dose control." *Biometrics* 42 (1986): 183-186.

4. Results

All female and male animals have been included in the analyses of pathology. Inclusion did not depend on dosing durations or on pregnancy. Since only mammary gland aveolus hyperplasia was analyzed for females (1 animal in the the control group and 2 animals in the 30,000 oxybenzone ppm group), the statistical conclusions are the same with or without exclusions. The impact of protocol deviations and pregnancy on male pathology data was deemed negligible by the Principle Investigator.

There was one female in the control group that had a mammary gland adenocarcinoma. There were no other neoplasms.

For each nonneoplastic lesion with incidence of at least two in a targeted dose group, analyses are presented for incidence and severity, including simple incidence and severity profile, defined as counts at each severity level within a treatment group (levels: 1=minimal; 2=mild; 3=moderate; and 4=marked). Severity is not presented for cases in which the lesion status is either present (level=5) or absent (level=0).

Results for nonneoplasms are presented for oxybenzone comparisons to control in Table 1 for females by body system. There were no statistically significant results for female nonneoplasms.

Results of analyses for nonneoplasms are presented for males by body system in Table 2a including all doses of oxybenzone. For incidence of nonneoplasms, there were no statistically significant results in analyses for the oxybenzone groups compared to control. There was a significant difference in severity (p=0.045) of mammary gland aveolus hyperplasia for oxybenzone 3000 ppm compared to the control group. However, there were no significant comparisons to control for higher oxybenzone doses. Results of analyses for nonneoplasms are presented for males by body system

in Table 2b for oxybenzone 30,000 ppm compared to control. Severity differed significantly for thyroid gland ultimobranchial cyst (p=0.030) for the high oxybenzone dose compared to control.

Results for nonneoplasms are presented for ethinyl EE_2 comparisons to control for males by body system in Table 2c. There was a significant difference in severity (p=0.032) of thyroid gland ultimobranchial cyst for EE_2 0.05 ppm compared to the control group. There were significant differences in incidence and severity (p=0.009 and =0.005, respectively) of mammary gland aveolus hyperplasia for EE_2 0.05 ppm relative to the control group (40.0% incidence in the EE_2 group compared to 0.0% in control).

5. Conclusions

In pairwise comparisons of nonneoplasm incidence and severity of dosed groups to control, there were no statistically significant results for females.

For males, there were significant differences relative to the control group in severity of mammary gland aveolus hyperplasia for oxybenzone 3000 ppm, and in incidence and severity for EE_2 0.05 ppm. There were significant differences in severity of thyroid gland ultimobranchial cyst for oxybenzone 30,000 ppm and EE_2 0.05 ppm compared to the control treatment.

Appendices

A1. Statistical Tables

Table 1. Nonneoplastms by Body System for Females (Comparisons to Oxybenzone)

		Oxybenzone (ppm)
	Control	30,000
Integumentary System		
Mammary Gland, Alveolus: Hyperplasia		
Simple Incidence	1/25 (4.0%)	2/25 (8.0%)
CAFE P-Value	0.498	0.500
Severity Profile	0|1|0|0	0|1|0|1
JT/SW P-Value	0.270	0.270

Table 2a. Nonneoplasms by Body System for Males (Comparisons to Oxybenzone)

		Oxybenzone (ppm)		
	Control	3000	10000	30000
Integumentary System				
Mammary Gland, Alveolus: Hyperplasia				
Simple Incidence	2/25 (8.0%)	5/25 (20.0%)	7/25 (28.0%)	2/25 (8.0%)
CAFE P-Value	0.283N	0.209	0.069	0.695N
Severity Profile	1|1|0|0	4|1|0|0	4|3|0|0	1|1|0|0
JT/SW P-Value	0.390	0.045 * [1]	1.000	0.202

1. Indicates a violation in the monotonicity assumption of Shirley's test.
2. P-value is significant due to pooling the adjacent lower dose to enforce monotonicity in the Shirley-Williams' test; the extreme result implies that the monotonicity enforcement is inappropriate in this case.
3. Statistically significant one-sided differences with higher incidence or severity in dosed groups relative to control are indicated by asterisks: *** P<0.001; ** P<0.01; and * P<0.05 (N indicates a negative comparison to the control group; P-value for trend is shown below the control group).

Table 2b. Nonneoplasms by Body System for Males (Comparisons to Oxybenzone)

		Oxybenzone (ppm)
	Control	30,000
Endocrine System		
Pituitary Gland, Pars Distalis: Cyst		
Simple Incidence	2/25 (8.0%)	1/25 (4.0%)
CAFE P-Value	0.505N	0.500N
Severity Profile	2|0|0|0	1|0|0|0
JT/SW P-Value	0.278N	0.278N
Thyroid Gland: Ultimobranchial Cyst		
Simple Incidence	2/25 (8.0%)	7/25 (28.0%)
CAFE P-Value	0.071	0.069
Severity Profile	2|0|0|0	5|2|0|0
JT/SW P-Value	0.030 *	0.030 *

1. Statistically significant one-sided differences with higher incidence or severity in dosed groups relative to control are indicated by asterisks: *** P<0.001; ** P<0.01; and * P<0.05 (N indicates a negative comparison to the control group; P-value for trend is shown below the control group).

Table 2c. Nonneoplasms by Body System for Males (Comparisons to EE_2)

	Control	Ethinyl Estradiol (ppm) 0.05
Endocrine System		
Pituitary Gland, Pars Distalis: Cyst		
Simple Incidence	2/25 (8.0%)	0/25 (0.0%)
CAFE P-Value	0.241N	0.245N
Severity Profile	2|0|0|0	---
JT/SW P-Value	0.077N	0.077N
Thyroid Gland: Ultimobra Cyst		
Simple Incidence	2/25 (8.0%)	7/25 (28.0%)
CAFE P-Value	0.070	0.069
Severity Profile	2|0|0|0	6|1|0|0
JT/SW P-Value	0.032 *	0.032 *
Genital System		
Pros- Ven Lobe: Infiltrat Cell, Lymphocyte		
Simple Incidence	1/25 (4.0%)	3/25 (12.0%)
CAFE P-Value	0.300	0.305
Severity Profile	1|0|0|0	3|0|0|0
JT/SW P-Value	0.151	0.151
Integumentary System		
Mammary Gland, Alveolus: Hyperplasia		
Simple Incidence	2/25 (8.0%)	10/25 (40.0%)
CAFE P-Value	0.010 *	0.009 **
Severity Profile	1|1|0|0	5|5|0|0
JT/SW P-Value	0.005 **	0.005 **

1. Statistically significant one-sided differences with higher incidence or severity in dosed groups relative to control are indicated by asterisks: *** P<0.001; ** P<0.01; and * P<0.05 (N indicates a negative comparison to the control group; P-value for trend is shown below the control group).

A2. Data

The histopathology data were obtained from NARSS AD HOC PROD on NCTR Citrix using Brett Thorn's program PTHEXTRT in Natural to extract data and save to ASC files. These exported data files were read into SAS via the import macros written by the NTP support group within the Division of Bioinformatics and Biostatistics. The imported data were checked against the Pathology Group Report and found to match.

A3. Quality Control

1. Data Verification

The extraction of the data into ASC files and SAS was verified by the reviewer, Paul Felton, by review of the Natural and SAS code used to extract and verify the data.

2. Computer Program Verification

Programs in Natural and SAS were used to extract the data; SAS code was used to explore the distributional properties of the data and perform the statistical analysis.

The Natural program was verified by matching of data in the Pathology Statistical Report to the Pathology Group Report. The SAS programs were verified by detailed review of the program code, the program log, and the program output.

3. Statistical Report Review

3.1 Statistical Report Text

The statistical report was reviewed for logic, internal completeness, technical appropriateness, technical accuracy, and grammar. Technical appropriateness was reviewed based on statistical expertise.

Comments and questions were provided from the reviewer to the statistician. The statistician made appropriate changes and returned the report to the reviewer for final verification.

The text of the final statistical report was considered by the reviewer to be logical, internally complete, and technically appropriate and accurate. The statistical results stated in the text accurately presented those presented in the tables.

3.2 Table Verification

Analysis results were output from SAS to an .rtf file using PROC REPORT, which were then copied into the statistical report.

Statistical report tables were verified by checking the procedure used to create the tables and, additionally, by checking numbers sufficiently to conclude that the tables are correct.

4. Conclusions

The final statistical report has been fully reviewed and is considered by the reviewer to be logical, internally complete, and technically appropriate and accurate