CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

203505Orig1s000

STATISTICAL REVIEW(S)

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Sciences
Office of Biostatistics

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

NDA/BLA #:	NDA 203505/0000, 0005
Drug Name:	Osphena (ospemifene) Oral Tablets, 60 mg
Indication(s):	Treatment of moderate to severe symptoms of vulvar and vaginal atrophy of dyspareunia and/or dryness due to menopause
Applicant:	Shionogi Inc.
Date(s):	Submission date: 04/26/2012
	PDUFA date: 02/26/2013
Review Priority:	Standard
Biometrics Division:	Division of Biometrics III
Statistical Reviewer:	Xin Fang, Ph.D.
Concurring Reviewers:	Mahboob Sobhan, Ph.D.
Medical Division:	Division of Reproductive and Urological Products
Clinical Team:	Theresa H. Van Der Vlugt, M.D., Clinical Reviewer
Project Manager:	George A. Lyght

Keywords: NDA review, analysis of covariance, Cochran-Mantel-Haenszel, mixed models, clinical study

Table of Contents

LIST OF TABLES

1 EXECUTIVE SUMMARY

The Applicant is seeking approval of ospemifene 60 mg for the treatment of moderate to severe symptoms of vulvar and vaginal atrophy (VVA) of dyspareunia and/or dryness due to menopause. To support the above indication, safety and efficacy data from two phase 3 studies (15-50310 and 15-50821) was submitted. This review evaluates from a statistical perspective the adequacy of the submitted efficacy data supporting this claim.

Both studies were randomized, placebo-controlled, double-blind, multicenter, and parallel-arm, placebo-controlled clinical trials conducted in the US in post-menopausal women. Study 15-50310 was a three-arm (Placebo, ospemifene 30 mg and 60 mg) study, while study 15-50821 was a two-arm (placebo and ospemifene 60mg) study. The objectives of the studies were to demonstrate efficacy with respect to the co-primary endpoints: change from baseline to week 12 in the severity of most bothersome symptoms of dyspareunia and/or vaginal dryness, percent superficial and parabasal cells, and vaginal pH.

There were no statistical issues noted in this submission. FDA analysis was based on mITT population, as per Division's preferred analysis population for this indication. The mITT population was defined as all subjects having vaginal pH >5, vaginal superficial cell ≤5%, and with a most bothersome moderate to severe vaginal symptom of dryness or pain related to sexual activity (dyspareunia) at baseline.

Adjusting for multiplicity (for statistical evidence in favor of either dyspareunia or vaginal dryness endpoint), FDA analyses showed that ospemifene 60 mg statistically significantly improved dyspareunia only, compared to placebo with respect to all co-primary endpoints. Ospemifene 60 mg did not show statistically significant improvement in vaginal dryness. Ospemifene 30 mg did not show statistically significant improvement in either dyspareunia or vaginal dryness.

From a statistical perspective, data from the two submitted studies provided statistical evidence in support of ospemifene 60 mg in the treatment of moderate to severe VVA symptom of dyspareunia in post-menopausal women at 12 weeks.

2 INTRODUCTION

This section provides information on ospemifene development, and the studies submitted to support the purported indication.

2.1 Overview

Ospemifene is a selective estrogen receptor modulator (SERM) that binds specifically to estrogen receptors α and β with similar affinity. In rats, ospemifene induces mucification and a beneficial shift of maturation index, and therefore has potential ability to repair vaginal epithelium and to improve symptoms of VVA.

Ospemifene 60 mg oral tablets have been developed for the treatment of VVA. The initial owner, Homos Medical, started the development in 2003 under IND 67,216. On April 23, 2010, the current owner, Shionogi USA Inc, acquired the product and submitted this NDA.

Two placebo-controlled, pivotal phase-3 studies were conducted to demonstrate the safety and efficacy of ospemifene 60 mg as shown in Table 1. The Applicant also submitted four other supportive studies: a phase-2 dose-ranging placebo-controlled study (15-50717), a long-term safety extension study (15-50310X), a stand-alone phase-3 placebo-controlled 52-week safety study of ospemifene 60 mg oral dose, and a 52-week open label extension study (15-50312).

Table 1. Summary of Studies

Study	Phase and Design	# of Subjects per Arm	Study Site [Country (# of sites)]	Treatment Period	Follow-up Period	Study Population
15-50310	*Phase 3, 3-arm, randomized, Double-blind, multi-center in US*	*Planned 795 265 per arm Analyzed 826 30 mg 282 60 mg 276 Placebo 268*	*USA (76)*	*12 weeks*	*4 weeks*	*Postmenopausal women aged 40-80 with a diagnosis of VVA as assessed by vaginal PH, maturation index, and self-reported symptoms*
15-50821	*Phase 3, 2-arm, randomized, Double-blind, multi-center in US*	*Planned 750 375 per arm Analyzed 919 60 mg 463 Placebo 456*	*USA (112)*	*12 weeks*	*4 weeks*	*Postmenopausal women aged 40-80 with a diagnosis of VVA as assessed by vaginal PH>5.0, superficial cell≤5%, and self-reported MBS symptoms of dry or pain*

Source reviewer analysis based on dataset ADSL for the integrated summary of efficacy

2.2 Data Sources

The study reports and additional information were submitted electronically. The data quality of the submission is acceptable. Analysis datasets and the associated definition files used for both pivotal phase-3 studies are listed in Table 2.

Table 2. Data Sources

Study	File	Location
15-50310	*Datasets*	*\\CDSESUB1\EVSPROD\nda203505\0000\m5\datasets\ise\analysis\legacy\datasets* *\\CDSESUB1\EVSPROD\nda203505\0000\m5\datasets\15-50310\tabulations\sdtm*
	Definition	*\\CDSESUB1\EVSPROD\nda203505\0000\m5\datasets\ise\analysis\legacy\datasets\define.pdf* *\\CDSESUB1\EVSPROD\nda203505\0000\m5\datasets\15-50310\tabulations\sdtm\define.xml*
15-50821	*Datasets*	*\\CDSESUB1\EVSPROD\nda203505\0000\m5\datasets\ise\analysis\legacy\datasets* *\\CDSESUB1\EVSPROD\nda203505\0000\m5\datasets\15-50821\tabulations\sdtm*
	Definition	*\\CDSESUB1\EVSPROD\nda203505\0000\m5\datasets\ise\analysis\legacy\datasets\define.pdf* *\\CDSESUB1\EVSPROD\nda203505\0000\m5\datasets\15-50821\tabulations\sdtm\define.xml*

2.3 Indication(s)

Ospemifene 60 mg oral tablet is indicated for the treatment of moderate to severe symptoms of vulvar and vaginal atrophy (VVA) of dyspareunia and/or vaginal dryness.

3 STATISTICAL EVALUATION

This section presents a detailed review of the two pivotal phase-3 studies (15-50310 and 15-50821).

3.1 Data and Analysis Quality

The Applicant submitted efficacy datasets in Module 5.3.5.3 integrated summary of efficacy (ISE). The four co-primary efficacy variables can be recreated from the associated SDTM datasets (QS, CF … etc) for Studies 15-50310 and 15-50821. Non-CDISC legacy datasets were also submitted for both pivotal phase-3 studies, but these legacy datasets did not contain the Agency preferred modified intent-to-treat (mITT) population flag variable. Therefore, this reviewer used both ISE and SDTM datasets for the evaluation of study specific efficacy and safety of ospemifene 60 mg in the treatment of VVA.

The Applicant acquired phase-3 study reports from the previous owner and was unable to clearly define how to identify efficacy analysis populations from the ISE. Specifically, no clear description was provided to identify appropriate analysis datasets for the efficacy evaluation. Subsequent to Division's request, the Applicant provided addendums to the two Phase 3 study reports later on July 9, 2012. This reviewer had to conduct intensive programming to verify the information and was able to replicate the individual study results.

In addition, eight subjects in study 15-50821 had no identification numbers in the planned randomization code and scheme, yet they have received the study drugs. For lack of clarity, this reviewer conducted the efficacy analysis with and without these subjects.

3.2 Evaluation of Efficacy

This section evaluates the study design and the efficacy results of the two phase-3 studies 15-50310 and 15-50821. First, a brief regulatory history is described below.

In the meeting minutes dated 04/12/2007, the Agency agreed with the Applicant that they can claim the treatment efficacy for either one or both of the two most bothersome (MB) symptoms: (1) moderate to severe vaginal dryness, and (2) moderate to severe vaginal pain associated with sexual activity. The Agency also recommended separate analysis for each symptom using the patient self-indentified most bothersome symptom (MBS) of moderate to severe vaginal dryness or moderate to severe vaginal pain associated with sexual activity. The Agency stated in this meeting minutes: "Please note that in order to demonstrate a "win" for the endpoint of MBS, specify which of the following you intend to demonstrate: That both MB symptoms need to be demonstrate statistical significance for the endpoint of MBS to succeed. Or that only one of the two MB symptoms needs to demonstrate statistical significance for the endpoint of MBS to succeed. You will need to appropriately adjust the significance level of this test for multiplicity, taking into account the correlation between the two symptoms."

In the meeting minutes dated 10-29-2009, the Agency told the Applicant that the efficacy would be based on the results of each pivotal Phase-3 study (15-50310, and 15-50821), not on the overall results of the integrated summary of efficacy.

In the meeting minutes dated 05-12-2011, the Agency conveyed to the Applicant that "the primary efficacy analyses should be based on subjects meeting all three of the baseline inclusion criteria: vaginal pH greater than 5, less than 5% superficial cells on vaginal smear, and a most bothersome moderate to severe vaginal symptom."

3.2.1 Study Design and Endpoints

Two submitted phase-3 studies (15-50310, 15-50821) were randomized, placebo-controlled, double-blind, multicenter, and parallel-arm clinical trials conducted in post-menopausal women aged 40-80. Subjects who had superficial cells ≤5%, vaginal pH>5.0, and at least one moderate or severe VVA symptom were eligible for randomization. Study 15-50310 enrolled 826 (planned 795, 265/arm) women across 76 study sites, and Study 15-50821 enrolled 919 (planned 750, 375/arm) across 112 study sites. The primary objectives of both studies were to assess the efficacy, safety, and tolerability of ospemifene (30 mg and 60 mg in study 15-50310 and only 60 mg in study 15-50821), compared with placebo.

The VVA symptom in Study 15-50310 included vaginal dryness, itching, burning and dyspareunia (vaginal pain associated with sexual activity), while the VVA symptom in Study 15-50821 included only vaginal dryness and dyspareunia.

Two common treatment arms (ospemifene 60 mg QD, placebo QD) were used in both studies. Study 15-50310 has an extra treatment arm (ospemifene 30 mg QD). All subjects in both studies received non-hormonal vaginal lubricant for as-needed use during the treatment duration of 12 weeks followed by a 4-week post-treatment follow-up period.

The randomization was stratified by uterine status (intact or hysterectomized) in Study 15-50310, while the randomization was stratified by most bothersome symptoms (MBS: vaginal dryness or dyspareunia) in Study 15-50821.

The four co-primary endpoints were essentially the same in both studies:

- Change from baseline (screening) to Week 12 in the percentage of parabasal cells in the maturation index of the vaginal smear
- Change from baseline (screening) to Week 12 in the percentage of superficial cells in the maturation index of the vaginal smear
- Change from baseline (screening) to Week 12 in vaginal pH
- Change from baseline (randomization) to Week 12 in severity of the most bothersome VVA symptom (vaginal dryness, or vaginal pain associated with sexual activity)

Secondary endpoints included the above four changes from baseline to Week 4.

The severity scale of the most bothersome VVA symptom was coded as 0=None, 1=Mild, 2=Moderate, and 3=Severe. Therefore, the change from baseline of severity score of a VVA symptom had values of -3, -2, -1, 0, and 1 at Week 12.

3.2.2 Statistical Methodologies

In both studies, the intent-to-treat (ITT) population included all randomized subjects who have received at least 1 dose of the study medication. The per-protocol (PP) population included all ITT subjects who completed at least 10 weeks of treatment, and completed the end of study assessments, and took at least 85% of the study drug, and did not have any major protocol violations, and did not have a vaginal infection or any other medical condition that confounded the primary efficacy assessment.

The mITT population was considered as the primary efficacy population by the Agency in the meeting minutes dated 05/12/2011, which must include ITT subjects having vaginal pH >5, vaginal superficial cell ≤5%, and a most bothersome moderate to severe vaginal symptom of dryness or pain related to sexual activity. Subjects with other VVA symptoms in Study 15-50310 should be excluded.

For missing value, a last-observation-carried-forward (LOCF) approach was used to replace missing values in the analysis population. In subjects with no post-treatment value, baseline value was carried forward. If a subject did not have a baseline measurement, but have post-baseline measurement, the change score was set to zero. In Study 15-50310, if the early termination visit occurred ≤35 days from randomization visit, the missing value at Week 4 was replaced with early termination records; otherwise the value was set to missing at Week 4. In Study 15-50821, the Week 4 visit window included treatment Days 2 – 57. Assessments after 14 days from the last dose were not included in the analysis for study 15-50821.

In Study 15-50310, parabasal cells, superficial cells, and vaginal pH were evaluated using an analysis of covariance (ANCOVA) model. The model included fixed effects of treatment, uterus status, study center, and baseline value as covariate. The ANCOVA model in Study 15-50821 was similar except that there was no fixed effect of uterus status. In case of the ANCOVA assumptions were severely violated, a rank-based ANCOVA was used including study center and uterus status. For VVA MBS symptoms, the change from baseline to Week 12 in the severity of vaginal dryness and dyspareunia was analyzed using Cochran-Mantel-Haenszel (CMH) method controlling for study center (in both studies) and uterus status (only in Study 15-50310).

Type –I error rate for between dose comparisons was controlled by step-down approach. Ospemifene 60 mg dose was tested first: if all four co-primary endpoints were statistically significant, then ospemifene 30 mg dose was tested. There was no multiplicity adjustment for evaluating the two VVA symptoms as agreed upon with the Agency in order to claim either or

both symptoms. A Bonferroni multiplicity adjustment was performed by this reviewer between vaginal dryness and dyspareunia.

The sample size was determined with the following assumptions:

- Parabasal cells: 20% difference between treatment and placebo with a standard deviation (SD) of 40.3
- Superficial cells: 5% difference between treatment and placebo with a SD of 12.4
- Vaginal pH: 0.5 difference between treatment and placebo with a SD of 1.1
- 81% power to detect a difference of 0.4 between treatment and placebo with a SD of 0.94
- 40% of the subjects have a MBS VVA symptom of vaginal dryness
- A 2-sided alpha of 5% and a dropout rate of 15%

In the meeting minutes dated 05/12/2012, the Agency also requested the Applicant to perform the following:

- "Analysis for the co-primary MBS moderate to severe symptom endpoints of vaginal dryness and pain associated with sexual activity using the ANOVA model as used for the primary efficacy analysis that includes an indicator for vaginal lubricant use (Y/N) at week 12."
- "The primary efficacy analyses should be based on subjects meeting all three of the baseline inclusion criteria: vaginal pH greater than 5, less than 5% superficial cells on vaginal smear, and a most bothersome moderate to severe vaginal symptom." (mITT population)

This reviewer performed statistical analysis as per mITT population to evaluate the efficacy of ospemifene 60 mg, compared with placebo. To handle missing values, a mixed model repeated measure analysis (MMRM) was performed as a sensitivity analysis to LOCF method since the latter does not assume missing at random.

3.2.3 Patient Disposition, Demographic and Baseline Characteristics

Table 3 shows the subject disposition in both phase-3 studies. A total of 826 and 919 subjects were randomized in Studies 15-50310 and 15-50821, respectively. No single study site dominated the study enrollment. The overall discontinuation rates were 17% and 11% in Studies 15-50310 and 15-50821, respectively. The main reasons for discontinuations were withdrawal and adverse event. The mITT population consisted of 655 subjects (less than planned 795) and 905 subjects (greater than planned 750) in Studies 15-50310 and 15-50821, respectively.

Table 4 shows subject demographic and baseline characteristics of the mITT population. The subject mean age, racial subgroups, and mean body mass index were comparable among the treatment groups in both studies. The mean baseline values of vaginal dry score, dyspareunia score, percent superficial cell, percent parabasal cell, and pH were also similar among the treatment groups in both studies.

Table 3. Subject Disposition (ITT)

	Study 15-50310				Study 15-50821		
	Ospemifene 30 mg N (%)	Ospemifene 60 mg N (%)	Placebo N (%)	Total N (%)	Ospemifene 60 mg N (%)	Placebo N (%)	Total N (%)
Randomized	282 (100.0)	276 (100.0)	268 (100.0)	826 (100.0)	463 (100.0)	456 (100.0)	919 (100.0)
Completed Study	225 (79.8)	234 (84.8)	230 (85.8)	689 (83.4)	416 (89.8)	403 (88.4)	819 (89.1)
Discontinued	57 (20.2)	42 (15.2)	38 (14.2)	137 (16.6)	47 (10.2)	53 (11.6)	100 (10.9)
Subject decision	14 (5.0)	14 (5.1)	12 (4.5)	40 (4.8)	8 (1.7)	19 (4.2)	27 (2.9)
Lost to follow-up	8 (2.8)	6 (2.2)	4 (1.5)	18 (2.2)	9 (1.9)	9 (2.0)	18 (2.0)
Adverse event	15 (5.3)	13 (4.7)	11 (4.1)	39 (4.7)	25 (5.4)	14 (3.1)	39 (4.2)
Protocol violation	14 (5.0)	7 (2.5)	9 (3.4)	30 (3.6)	1 (0.2)	2 (0.4)	3 (0.3)
Lack of efficacy	1 (0.4)	0 (0.0)	0 (0.0)	1 (0.1)	0 (0.0)	0 (0.0)	0 (0.0)
Other	5 (1.8)	2 (0.7)	2 (0.7)	9 (1.1)	4 (0.9)	9 (2.0)	13 (1.4)
MBS							
Vaginal dryness*	95 (33.7)	113 (40.9)	100 (37.3)	308 (37.3)	157 (33.9)	150 (32.9)	307 (33.4)
Vaginal pain with sex*	124 (44.0)	110 (39.9)	113 (42.2)	347 (42.,0)	301 (65.0)	297 (65.1)	598 (65.1)
mITT Population (FDA)	219 (77.7)	223 (80.8)	213 (79.5)	655 (79.3)	458 (98.9)	447 (98.0)	905 (98.5)
mITT Population	257 (91.1)	254 (92.0)	247 (92.2)	758 (91.8)	458 (98.9)	447 (98.0)	905 (98.5)
PP Population	181 (64.2)	177 (64.1)	194 (72.4)	552 (66.8)	382 (82.5)	388 (85.1)	770 (83.8)

Source: reviewer analysis on Dataset ISE.ADDS, QS dataset in each of individual study SDTM dataset
** All subjects reported the pertinent most bothersome VVA symptom at baseline*

Table 4. Subject Demographic and Baseline Characteristics (mITT)

Characteristics	Study 15-50310				Study 15-50821		
	Ospemifene 30 mg (N =219)	Ospemifene 60 mg (N =223)	Placebo (N =213)	Total (N =655)	Ospemifene 60 mg (N =458)	Placebo (N =447)	Total (N =905)
Mean Age (SD)	58 (5.9)	58 (6.2)	59 (5.9)	59 (6.0)	59 (6.6)	58 (6.3)	59 (6.4)
BMI (SD)	26.1 (4.5)	25.7 (4.2)	25.9 (4.1)	25.9 (4.2)	26.1 (4.3)	26.2 (4.3)	26.2 (4.3)
Race: N (%)							
White	196 (89.5)	204 (91.5)	191 (89.6)	591 (90.2)	404 (88.2)	391 (87.5)	795 (87.9)
Black	12 (5.5)	12 (5.4)	12 (5.6)	36 (5.5)	28 (6.1)	33 (7.4)	61 (6.7)
Asian	5 (2.3)	3 (1.4)	5 (2.4)	13 (2.0)	8 (1.8)	3 (0.7)	11 (1.2)
American Indian/Alaska Native	0 (0.0)	2 (0.9)	0 (0.0)	2 (0.3)	0 (0.0)	0 (0.0)	0 (0.0)
Native Hawaiian or Other Pacific Islander	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	2 (0.4)	0 (0.0)	2 (0.2)
Other	6 (2.7)	2 (0.9)	5 (2.3)	13 (2.0)	16 (3.5)	20 (4.5)	36 (4.0)
Uterus Status							
Intact	101 (46.1)	105 (47.1)	101 (47.4)	307 (46.9)	240 (52.4)	240 (53.7)	480 (53.0)
Baseline Value							
MBS dry	2.5 (0.50)	2.5 (0.50)	2.4 (0.49)	2.5 (0.50)	2.5 (0.50)	2.5 (0.50)	2.5 (0.50)
MBS pain with sex	2.6 (0.48)	2.7 (0.44)	2.7 (0.45)	2.7 (0.46)	2.7 (0.47)	2.7 (0.47)	2.7 (0.47)
% Superficial cell	1.0 (1.51)	0.7 (1.33)	0.7 (1.23)	0.8 (1.37)	0.8 (1.36)	0.8 (1.37)	0.8 (1.37)
% Parabasal cell	40.3 (38.04)	41.3 (39.18)	40.7 (37.43)	40.8 (38.18)	49.7 (39.03)	49.3 (40.09)	49.5 (39.54)
Vaginal pH	6.4 (0.72)	6.4 (0.76)	6.4 (0.70)	6.4 (0.73)	6.3 (0.78)	6.3 (0.76)	6.3 (0.77)

Source: reviewer analysis on Dataset ISE.ADSL, ISE.ADMBS, ISE.ADSC, ISE.ADPH, ISE.ADPC

3.2.4 Results and Conclusion

This section presents the efficacy results only for ospemifene 60 mg compared with placebo at 12 weeks in dyspareunia subjects; because ospemifene 30 mg did not statistically significantly improve a VVA symptom at 12 weeks, and neither ospemifene doses statistically significantly improved a VVA symptom at 4 weeks or vaginal dryness at 12 weeks in both studies (Appendix Table 10).

This reviewer verified the Applicant's efficacy results submitted on 04/26/2012 and 11/01/2012 based on the Agency preferred mITT population.

This reviewer also performed several sensitivity analyses to confirm these efficacy results using an ANCOVA model with LOCF without baseline carried forward, and a mixed model repeated measures (MMRM) analysis. Results from the sensitivity analyses (not shown) are comparable to those using the Applicant's pre-specified statistical models.

Our analysis showed that ospemifene 60 mg did not improve vaginal dryness in either study. Therefore, only results for dyspareunia are shown here in more detail. Results for dryness can be found in the Appendix (Table 10).

As shown in Table 5, ospemifene 60 mg statistically significantly improved all four primary efficacy endpoints at week 12 in both studies in dyspareunia subjects. The placebo-adjusted mean change from baseline of the dyspareunia severity score is -0.51 and -0.36 in Studies 15-50310 and 15-50821, respectively. The placebo-adjusted mean change from baseline of vaginal pH is -0.97 and -0.90 in Studies 15-50310 and 15-50821, respectively. The placebo-adjusted mean change from baseline of percent parabasal cell count is -40.3 and -40.0 in Studies 15-50310 and 15-50821, respectively. The placebo-adjusted mean change from baseline of percent superficial cell count is 8.2 and 10.7 in Studies 15-50310 and 15-50821, respectively. Ospemifene 60 mg statistically significantly improved the VVA symptom of dyspareunia in both studies.

Table 5. Mean (SE) Change from Baseline to Week 12 in the Co-primary Endpoints: mITT population, LOCF

Study	Co-primary Endpoint(s)	Ospemifene 60 mg	Placebo	Difference (95% CI)	Nominal P-value
15-50310	**Dyspareunia:**				
	N	110	113		
	Baseline (SD)	2.74 (0.44)	2.73 (0.45)		
	Change from baseline	-1.39 (0.11)	-0.89 (0.11)	-0.51 (-0.81, -0.20)[b]	0.0012[a]
	% Superficial Cells	10.88 (1.27)	2.73 (1.27)	8.16 (4.73, 11.58)	<.0001[b]
	% Parabasal Cells	-34.44 (2.44)	5.84 (2.44)	-40.3 (-46.9, -33.7)	<.0001[b]
	pH	-0.97 (0.09)	-0.002 (0.09)	-0.97 (-1.22, -0.73)	<.0001[b]
15-50821	Dyspareunia:				
	N	301	297		
	Baseline (SD)	2.67 (0.47)	2.67 (0.47)		
	Change from baseline	-1.55 (0.06)	-1.29 (0.07)	-0.36 (-0.53, -0.18)[b]	<.0001[a]
	% Superficial Cells [(LS mean (SE)]	12.35 (0.68)	1.69 (0.69)	10.66 (8.81, 12.52)	<.0001[b]
	% Parabasal Cells [(LS mean (SE)]	-40.57 (1.57)	-0.56 (1.59)	-40.0 (-44.3, -35.7)	<.0001[b]
	pH [(LS mean (SE)]	-0.95 (0.05)	-0.08 (0.05)	-0.87 (-1.01, -0.73)	<.0001[b]

Source: Reviewer's analysis based on ISE analysis datasets ADMBS310, ADMBS821, ADPH, ADPC, and ADSC
a: Test based on CMH stratified by pooled site (both studies), and uterus status (Study 15-50310 only)
b: Test based on ANCOVA model having fixed effect of treatment, uterus status (Study 15-50310 only), pooled site, and baseline

Reviewer Comments on the Efficacy Results

Studies 15-50310 and 15-50821 demonstrate that ospemifene 60 mg is effective in the treatment of moderate to severe dyspareunia as well as superficial cells and pH in postmenopausal women at Week 12. Ospemifene 60 mg, however, did not demonstrate statistically significant improvement in vaginal dryness in both studies.

3.3 Evaluation of Safety

The safety review is referred to the clinical review.

4 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

4.1 Gender, Race, Age, and Geographic Region

Subgroup analyses by gender, age, race, and region were not performed because all study subjects were Caucasian women (88%-90%) in the US, and age subgroup was not identified by the Applicant.

4.2 Other Special/Subgroup Populations

Table 6 and Table 7 show uterus subgroup analysis in dyspareunia subjects. Subjects were dichotomized into two subgroups (having intact uterus or otherwise). The placebo-adjusted mean change from baseline of severity score in subjects having intact uterus appeared to be worse compared to subjects without intact uterus (-0.4 vs. -0.7 in Study 15-50310, and -0.3 vs. -0.5 in Study 15-50821). The improvements in vaginal pH, percent parabasal cell count, and percent superficial cell count are similar between the two subgroups.

Table 8 and Table 9 show lubrication subgroup analysis in dyspareunia subjects. Subjects were dichotomized into two subgroups (used lubrication or otherwise). The placebo-adjusted mean change from baseline of severity score in subjects who used lubrication seems better than that in subjects who did not used lubrication (-0.5 vs. 0.1 in Study 15-50310, and -0.4 vs. -0.3 in Study 15-50821). The improvements in vaginal pH, percent parabasal cell count, and percent superficial cell count are similar between the two subgroups.

Reviewer Comments on the subgroup Analyses

Overall, exploratory subgroup analyses by lubricate use and uterus status showed no significant difference between the treatment groups.

5 SUMMARY AND CONCLUSIONS

5.1 Statistical Issues

There were no statistical issues regarding the design, and statistical analysis methods in this submission. The four co-primary efficacy endpoints were evaluated based on Agency preferred mITT population, which contains subjects having vaginal pH >5, vaginal superficial cell ≤5%, and a most bothersome moderate to severe vaginal symptom of dryness or dyspareunia. This reviewer adjusted the multiplicity of tests between the two VVA symptoms using Bonferroni method because the Agency may approve either of the two treatments: dyspareunia or vaginal dryness. The pre-specified step-down adjustment of multiplicity was applied for the tests between the two doses. Consequently tests were performed at a two-sided alpha of 0.0125 and 0.025 in Study 15-50310 and 15-50821, respectively.

The Applicant's LOCF imputation included baseline value forwarded if there was no post-baseline assessment. This reviewer confirmed the efficacy conclusion using both an ANCOVA model with LOCF (no baseline value carried forward) and an MMRM analysis. Our exploratory analyses also suggested that use of lubricant have some effect in the improvement of symptoms compared to subjects' who did not use lubricant. Eight subject IDs were not in the planned randomization scheme and codes, however efficacy results with or without these eight subjects are similar.

5.2 Collective Evidence

Study 15-50310 and study 15-50821 showed that Ospemifene 60 mg was effective in the treatment of moderate to severe dyspareunia symptom in postmenopausal women at the end of week 12. Ospemifene 60 mg also showed statistically significant (p-value ≤0.0125) improvement in other three co-primary efficacy measurements compared with placebo. Ospemifene 60 mg was not effecting in the treatment of vaginal dryness in both studies.

5.3 Conclusions and Recommendations

From a statistical perspective, the submitted data from studies 15-50310 and 15-50821 provided adequate evidence demonstrating the efficacy of ospemifene 60 mg in the treatment of the VVA symptom of dyspareunia in postmenopausal women aged 40 – 80 at Week 12.

5.4 Labeling Recommendations

The efficacy information in Table 5 is to be used for Section 14.1 of the label.

APPENDICES (Suppotive Efficacy Tables)

Table 6. Dyspareunia Intact-Uterus Subgroup Results: the Mean Change from Baseline to Week 12 (mITT population , LOCF)

Study	Co-primary Endpoints	Ospemifene 60 mg (N=57)	Placebo (N= 53)	Difference (95% CI)	Nominal P-value
15-50310	Pain with Sex [LS mean (SE)]	-1.35 (0.16)	-0.94 (0.17)	-0.42 (-0.85, 0.02) [b]	0.1135[a]
	pH [(LS mean (SE)]	-1.05 (0.13)	-0.10 (0.14)	-0.9 (-1.3,-0.6)	<.0001[b]
	% Parabasal Cells [(LS mean (SE)]	-36.98 (3.60)	3.93 (3.87)	-40.9 (-51.1, -30.7)	<.0001[b]
	% Superficial Cells [(LS mean (SE)]	11.30 (1.94)	2.60 (2.09)	8.7 (3.2, 14.2)	0.0021[b]
	Co-primary Endpoints	**Ospemifene 60 mg (N= 173)**	**Placebo (N= 168)**	**Difference (95% CI)**	**Nominal P-value**
15-50821	Pain with Sex [LS mean (SE)]	-1.47 (0.09)	-1.20 (0.09)	-0.26 (-0.50, -0.03) [b]	0.0361[a]
	pH [(LS mean (SE)]	-1.04 (0.06)	-0.03 (0.06)	-1.0 (-1.2, -0.8)	<.0001[b]
	% Parabasal Cells [(LS mean (SE)]	-43.16 (2.09)	0.84 (2.11)	-44.0 (-49.7, -38.3)	<.0001[b]
	% Superficial Cells [(LS mean (SE)]	13.32 (0.96)	1.99 (0.97)	11.3 (8.7, 14.0)	<.0001[b]

Source: Reviewer's analysis based on ISE analysis datasets ADMBS310, ADMBS821, ADPH, ADPC, and ADSC
a: Test based on CMH stratified by pooled site (both studies)
b: Test based on ANCOVA model having fixed effect of treatment, pooled site, and baseline

Table 7. Dyspareunia No-Intact-Uterus Subgroup Results: the Mean Change from Baseline to Week 12 (mITT population, LOCF)

Study	Co-primary Endpoints	Ospemifene 60 mg (N= 53)	Placebo (N= 60)	Difference (95% CI)	Nominal P-value
15-50310	Pain with Sex [LS mean (SE)]	-1.53 (0.17)	-0.86 (0.16)	-0.7 (-1.1, -0.2) [b]	0.0031[a]
	pH [(LS mean (SE)]	-0.87 (0.14)	0.11 (0.13)	-1.0 (-1.3, -0.6)	<.0001[b]
	% Parabasal Cells [(LS mean (SE)]	-29.60 (3.66)	8.70 (3.41)	-38.3 (-47.5, -29.1)	<.0001[b]
	% Superficial Cells [(LS mean (SE)]	11.26 (1.80)	2.57 (1.70)	8.7 (4.1, 13.2)	0.0002[b]
	Co-primary Endpoints	**Ospemifene 60 mg (N= 128)**	**Placebo (N= 128)**	**Difference (95% CI)**	**Nominal P-value**
15-50821	Pain with Sex [LS mean (SE)]	-1.68 (0.10)	-1.23 (0.10)	-0.5 (-0.7, -0.2) [b]	0.0009 [a]
	pH [(LS mean (SE)]	-0.84 (0.08)	-0.20 (0.08)	-0.6 (-0.9, -0.4)	<.0001[b]
	% Parabasal Cells [(LS mean (SE)]	-36.61 (2.47)	-2.30 (2.51)	-34.3 (-41.0, -27.7)	<.0001[b]
	% Superficial Cells [(LS mean (SE)]	10.72 (0.94)	1.23 (0.95)	9.5 (7.0, 12.0)	<.0001[b]

Source: Reviewer's analysis based on ISE analysis datasets ADMBS310, ADMBS821, ADPH, ADPC, and ADSC
a: Test based on CMH stratified by pooled site (both studies)
b: Test based on ANCOVA model having fixed effect of treatment, pooled site, and baseline

Table 8. Dyspareunia Lubrication-User Subgroup Results: the Mean Change from Baseline to Week 12 (mITT population, LOCF)

Study	Co-primary Endpoints	Ospemifene 60 mg (N=72)	Placebo (N=89)	Difference (95% CI)	Nominal P-value
15-50310	Pain with Sex [LS mean (SE)]	-1.11 (0.14)	-0.62 (0.12)	-0.5 (-0.8, -0.2) [b]	0.0024[a]
	pH [(LS mean (SE)]	-1.16 (0.12)	-0.03 (0.11)	-1.1 (-1.4, -0.9)	<.0001[b]
	% Parabasal Cells [(LS mean (SE)]	-34.66 (3.22)	7.97 (2.90)	-42.6 (-50.3, -34.9)	<.0001[b]
	% Superficial Cells [(LS mean (SE)]	10.85 (1.84)	3.02 (1.65)	7.8 (3.5, 12.2)	0.0005[b]
	Co-primary Endpoints	**Ospemifene 60 mg (N=230)**	**Placebo (N=234)**	**Difference (95% CI)**	**Nominal P-value**
15-50821	Pain with Sex [LS mean (SE)]	-1.48 (0.07)	-1.10 (0.07)	-0.4 (-0.6, -0.2) [b]	0.0003[a]
	pH [(LS mean (SE)]	-0.97 (0.06)	-0.05 (0.06)	-0.9 (-1.1, -0.8)	<.0001[b]
	% Parabasal Cells [(LS mean (SE)]	-39.54 (1.81)	-0.59 (1.78)	-39.0 (-43.8, -34.1)	<.0001[b]
	% Superficial Cells [(LS mean (SE)]	12.55 (0.78)	1.69 (0.77)	10.9 (8.8, 13.0)	<.0001[b]

Source: Reviewer's analysis based on ISE analysis datasets ADMBS310, ADMBS821, ADPH, ADPC, and ADSC
a: Test based on CMH stratified by pooled site (both studies), and uterus status (Study 15-50310 only)
b: Test based on ANCOVA model having fixed effect of treatment, uterus status (Study 15-50310 only), pooled site, and baseline

Table 9. Dyspareunia No-Lubrication-User Subgroup Results: the Mean Change from Baseline to Week 12 (mITT population, LOCF)

Study	Co-primary Endpoints	Ospemifene 60 mg (N=31)	Placebo (N=17)	Difference (95% CI)	Nominal P-value
15-50310	Pain with Sex [LS mean (SE)]	-1.58 (0.26)	-1.70 (0.35)	0.1 (-0.8, 1.0) [b]	0.8256[a]
	pH [(LS mean (SE)]	-0.93 (0.22)	0.28 (0.30)	-1.2 (-2.0, -0.4)	0.0027[b]
	% Parabasal Cells [(LS mean (SE)]	-34.11 (5.69)	8.82 (7.55)	-42.9 (-62.6, -23.2)	<.0001[b]
	% Superficial Cells [(LS mean (SE)]	11.42 (2.86)	2.10 (3.93)	9.3 (-0.7, 19.3)	0.0676[b]
	Co-primary Endpoints	**Ospemifene 60 mg (N=65)**	**Placebo (N=56)**	**Difference (95% CI)**	**Nominal P-value**
15-50821	Pain with Sex [LS mean (SE)]	-1.84 (0.14)	-1.57 (0.16)	-0.3 (-0.7, 0.1) [b]	0.1916[a]
	pH [(LS mean (SE)]	-1.00 (0.12)	-0.24 (0.13)	-0.8 (-1.1, -0.4)	<.0001[b]
	% Parabasal Cells [(LS mean (SE)]	-46.19 (3.51)	-0.80 (3.87)	-45.4 (-55.5, -35.3)	<.0001[b]
	% Superficial Cells [(LS mean (SE)]	12.56 (1.64)	1.72 (1.81)	10.8 (6.1, 15.6)	<.0001[b]

Source: Reviewer's analysis based on ISE analysis datasets ADMBS310, ADMBS821, ADPH, ADPC, and ADSC
a: Test based on CMH stratified by pooled site (both studies), and uterus status (Study 15-50310 only)
b: Test based on ANCOVA model having fixed effect of treatment, uterus status (Study 15-50310 only), pooled site, and baseline

Table 10. Mean (SE) Change from Baseline to Week 12 in the Co-primary Endpoints mITT population, LOCF

Study	Co-primary Endpoint(s)	Ospemifene 60 mg	Placebo	Difference (95% CI)	Nominal P-value
15-50310	**Vaginal Dryness:**				
	N	113	100		
	Baseline (SD)	2.5 (0.50)	2.4 (0.49		
	Change from baseline	-1.29 (0.09)	-0.92 (0.10)	-0.37 (-0.63, -0.11)[b]	0.0136[a]
	% Superficial Cells	11.16 (1.19)	2.33 (1.25)	8.83 (5.48, 12.18)	<.0001[b]
	% Parabasal Cells	-26.65 (2.35)	0.12 (2.47)	-26.76 (-33.40, -20.13)	<.0001[b]
	pH	-0.92 (0.09)	-0.16 (0.09)	-0.75 (-0.99, -0.51)	<.0001[b]
15-50821	**Vaginal Dryness:**				
	N	157	150		
	Baseline (SD)	2.5 (0.50)	2.5 (0.50)		
	Change from baseline	-1.33 (0.08)	-1.11 (0.08)	-0.22 (-0.44, 0.003)[b]	0.0853[a]
	% Superficial Cells	12.32 (1.03)	3.53 (1.06)	8.79 (5.91, 11.67)	<.0001[b]
	% Parabasal Cells	-31.65 (2.13)	-4.11 (2.19)	-27.55 (-33.51, -21.59)	<.0001[b]
	pH	-0.95 (0.07)	-0.25 (0.07)	-0.71 (-0.90, -0.51)	<.0001[b]

Source: Reviewer's analysis based on ISE analysis datasets ADMBS310, ADMBS821, ADPH, ADPC, and ADSC
a: Test based on CMH stratified by pooled site (both studies), and uterus status (Study 15-50310 only)
b: Test based on ANCOVA model having fixed effect of treatment, uterus status (Study 15-50310 only), pooled site, and baseline

---

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

/s/

XIN FANG
02/12/2013

MAHBOOB SOBHAN
02/12/2013

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

Statistical Review and Evaluation

CARCINOGENICITY STUDIES

IND/NDA Number:	NDA 203-505
Drug Name:	Ospemifene
Indication(s):	104 Week Carcinogenicity in Rats and Mice
Applicant:	Sponsor: Hormos Medical Limited, PharmaCity,Itainen Pitkakatu 4, Fl-20520 Turku, Finland [redacted] (b) (4)
Documents Reviewed:	Electronic submission, Dated: April 26, 2012 Electronic data submitted on April 26, 2012
Review Priority:	Standard
Biometrics Division:	Division of Biometrics -6
Statistical Reviewer:	Mohammad Atiar Rahman, Ph.D.
Concurring Reviewer:	Karl Lin, Ph.D.
Medical Division:	Division of Reproductive and Urologic Products
Reviewing Pharmacologist:	Jeffrey Bray, Ph.D.
Project Manager:	George Lyght
Keywords:	Carcinogenicity, Dose response

Table of Contents

1. Background

In this submission the sponsor included reports of two animal carcinogenicity studies, one in rats and one in mice. These studies were intended to assess the carcinogenic potential of Ospemifene in rats and mice when administered orally via gavage at appropriate drug levels. The length of both the rat and the mouse studies were designed for 104 weeks. Results of this review have been discussed with the reviewing pharmacologist Dr. Bray.

In this review, the phrase "dose response relationship" refers to the linear component of the effect of treatment, and not necessarily to a strictly increasing or decreasing mortality or tumor rate as dose increases.

2. Rat Study

Two separate experiments were conducted, one in males and one in females. In each of these two experiments there were three treated groups and two identical vehicle control groups. Two hundred and fifty (b) (4) Crl: WI(Hans) rats of each sex were randomly allocated to treated and control groups in equal size of 50 rats. The dose levels for treated groups were 10, 50, and 300 mg/kg/day. In this review these dose groups would be referred to as the low, medium, and high dose group, respectively. The two controls would be referred to as Control 1 and Control 2. The rats in both control groups received the vehicle (corn oil) by gavage.

During the administration period all rats were observed daily for signs of ill health or overt toxicity. In addition, each rat was given a detailed physical examination at weekly intervals which included palpation for tissue masses. An individual record was maintained of the clinical condition of each rat. Individual body weights were recorded before treatment on the first day of dosing, at weekly intervals for 16 weeks, and once every four weeks thereafter and before the necropsy.

2.1. Sponsor's analyses

2.1.1. Survival analysis

The survival probability functions were estimated by the Kaplan-Meier techniques. The tests for both an increasing and a decreasing dose response in mortality across all groups were performed following the methods suggested by Peto et al. al. (1980) using the dose levels as weighting coefficients. One directional pairwise tests of the treated groups against the combined control groups were also performed. The two control groups were compared using two-sided tests.

Sponsor's findings: Sponsor's calculations showed mortality rates of 32%, 26%, 4%, 14%, and 6% in Control 1, Control 2, low, medium, and high dose groups, respectively in male rats, and 28%. 42%, 8%, 18%, and 12% in control 1, control 2, low, medium, and high dose groups, respectively in female rats. The overall mortality in the combined control was 29% and 35% in male and female rats. The sponsor concluded that for males, there was a significant decreasing dose response in mortality across the groups, and the mortality in groups given 10, 50 or 300 mg/kg/day was significantly lower than that in the controls. For females, there was a significant decreasing dose response in mortality across the groups, and the mortality in females given 10, 50 or 300 mg/kg/day was significantly lower than that in the controls. There was no significant difference in mortality between the two controls in either sex.

2.1.2. Tumor data analysis

The sponsor pooled the two control groups for the analysis of the tumor data and analyzed the numbers of tumor bearing animals for tumor types found in at least three animals of the given sex. Tumors of similar histogenic origin were merged, as requested by the Pathologist. The tests for both an increasing and a decreasing dose response in tumor incidence rates were performed across all groups following the methods suggested by Peto et al. al. (1980) using the dose levels as weighting coefficients. One directional pairwise tests of the treated groups against the combined control groups were also performed. The incidences in the two control groups were compared using two-sided tests. Non-fatal tumors were analyzed using fixed intervals of 1 to 50 weeks, 51 to 80 weeks, 81 weeks to 104 weeks and the terminal kill phase. The fatal and non-fatal results were combined in accordance with the methodologies outlined in Peto et al. (1980). Where the combined analysis was significant ($P<0.05$), separate analyses for fatal and non-fatal tumors were performed.

Adjustment for multiple comparisons: The sponsor used the criterion for adjustment of multiple testing suggested in the FDA guidance for carcinogenicity data analysis. The FDA guidance propose the decision rules which test an overall dose response at the 0.005 level for common tumors, i.e. those with background incidence rates higher than 1%, and at the 0.025 level for rare tumors, i.e. those with background incidence rates less than or equal to 1%. The decision rules test pairwise comparisons at the 0.01 level for common tumors and at the 0.05 level for rare tumors.

Sponsor's findings: In the sponsor's analysis for tests of increasing dose response and or pairwise comparisons the following tumor types gave rise to p-values less than 0.05:

(a) Liver hepatocellular tumor, in both sexes and (b) Thymus epithelial tumor, in both sexes

The tumor type (a) was classified as rare in males and common in females and tumor type (b) was classified as common in both sexes. Using the adjustment for multiple testing described above, the sponsor's noted the following significant finings:

Male liver hepatocellular tumor: Fatal and non-fatal tumors combined, controls vs. animals given 50 mg/kg/day ($P=0.049$) and controls vs. animals given 300 mg/kg/day ($P=0.015$).

Female liver hepatocellular tumor: Fatal and non-fatal tumors combined, controls vs. animals given 50 mg/kg/day ($P=0.003$). Non-fatal tumors only, controls vs. animals given 50 mg/kg/day ($P=0.007$).

Male thymus epithelial tumor: Fatal and non-fatal tumors combined, overall dose response ($P=0.006$), controls vs. animals given 10 mg/kg/day ($P=0.009$), controls vs. animals given 50 mg/kg/day ($P<0.001$) and controls vs. animals given 300 mg/kg/day ($P<0.001$). Non-fatal tumors only, overall dose response ($P=0.004$), controls vs. animals given 10 mg/kg/day ($P=0.009$), controls vs. animals given 50 mg/kg/day ($P<0.001$) and controls vs. animals given 300 mg/kg/day ($P<0.001$).

Female thymus epithelial tumor: Fatal and non-fatal tumors combined, controls vs. animals given 50 mg/kg/day ($P<0.001$) and controls vs. animals given 300 mg/kg/day ($P=0.003$). Non-fatal tumors only, controls vs. animals given 50 mg/kg/day ($P<0.001$) and controls vs. animals given 300 mg/kg/day ($P=0.001$).

In the comparisons between the two control groups, there was no significant difference in the incidence of any tumor type in either sex.

In overall, the sponsor's concluded that the marked increase in thymus tumors in both sexes at all dose levels may be due to the antiestrogenic effect of the study drug in this target tissue, which effect is attenuating the physiological thymic involution (atrophy) process induced by estrogens starting during puberty.

2.2. Reviewer's analyses

To verify sponsor's analyses and to perform additional analysis suggested by the reviewing pharmacologist, this reviewer independently performed survival and tumor data analyses. Data used in this reviewer's analysis were provided by the sponsor electronically.

In the submitted data set the sponsor identified both control 1 and control 2 by zero (0). Therefore, it was not possible to separate the two control groups. For studies with two identical controls the FDA guidance for carcinogenicity data analysis recommends to analyze the data using the combined controls. Due to the way the data were submitted and following the FDA guidance this reviewer analyzed both the survival and the tumor data using the combined control.

2.2.1. Survival analysis

The survival distributions of animals in placebo control, untreated control, low, and medium dose groups were estimated by the Kaplan-Meier product limit method. The dose response relationship was tested using the likelihood ratio test and the homogeneity of survival distributions was tested using the log-rank test. The intercurrent mortality data are given in Tables 1A and 1B in the appendix for male and female mice, respectively. The Kaplan-Meier curves for survival rates are given in Figures 1A and 1B in the appendix for male and female mice, respectively. Results of the tests for dose response relationship and homogeneity of survivals, are given in Tables 2A and 2B in the appendix for male and female mice, respectively.

Reviewer's findings: This reviewer's analysis showed 27%, 4%, 14%, and 6% overall mortality in male rats and 35%, 8%, 18%, and 12% overall mortality in female rats in combined control, low, medium, and high dose groups, respectively. This reviewer's analysis showed no statistically significant positive dose response relationship in mortality across treatment groups in either sex. The pairwise comparisons showed statistically significant increased mortality in combined control compared to any of the treated groups in both sexes.

Reviewer's comment: *In male rat, the sponsor's calculation showed 29% mortality in the combined controls, while this reviewer's calculation showed 27%. This difference is due to the fact that there were two male rats in control groups (#0001and #0223) that died due to natural causes during the terminal sacrifice week. The sponsor did not count them with the terminally sacrificed animals, while this reviewer counted them with the terminally sacrificed animals.*

2.2.2. Tumor data analysis

The tumor data were analyzed for dose response relationships and pairwise comparisons of control group with each of the treated groups. Both the dose response relationship tests and pairwise comparisons were performed using the Poly-K method described in the paper of Bailer and Portier (1988) and Bieler and Williams (1993). In this method an animal that lives the full study period ($w_{\max}$) or dies before the terminal sacrifice but develops the tumor type being tested gets a score of s_h =1. An animal that dies at week w_h without a tumor before the end of

the study gets a score of $s_h = \left(\frac{w_h}{w_{\max}} \right)^k$ <1. The adjusted group size is defined as $N^* = \Sigma s_h$. As an interpretation, an animal with score s_h =1 can be considered as a whole animal while an animal with score s_h <1 can be considered as a partial animal. The adjusted group size N^* is equal to N (the original group size) if all animals live up to the end of the study or if each animal that dies before the terminal sacrifice develops at least one tumor being tested, otherwise the adjusted group size is less than N. These adjusted group sizes are then used for the dose response relationship (or the pairwise) tests using the Cochran-Armitage test. One critical point for Poly-K test is the choice of the appropriate value of K, which depends on the tumor incidence pattern with the increased dose. For long term 104 week standard rat and mouse studies, a value of K=3 is suggested in the literature. Hence, this reviewer used K=3 for the analysis of this data. For the calculation of p-values the exact permutation method was used.

Multiple testing adjustment: For the adjustment of multiple testing of dose response relationship the FDA guidance for the carcinogenicity study design and data analysis suggests the use of test levels of α=0.005 for common tumors and α=0.025 for rare tumors for a submission with two species, and a significance level of α=0.01 for common tumors and α=0.05 for rare tumors for a submission with one species in order to keep the false-positive rate at the nominal level of approximately 10%. A rare tumor is defined as one in which the published spontaneous tumor rate is less than 1%. For multiple pairwise comparisons of treated group with control the FDA guidance suggests the use of test levels α=0.01 for common tumors and α=0.05 for rare tumors, in order to keep the false-positive rate at the nominal level of approximately 10% for both submissions with two or one species.

It should be noted that the FDA guidance for multiple testing for dose response relationship is based on a publication by Lin and Rahman (1998). In this work the authors investigated the use of this rule for Peto analysis. However, in a later work Lin and Rahman (2008) showed that this rule for multiple testing for dose response relationship is also suitable for Poly-K tests.

The tumor rates and the p-values of the tumor types tested for dose response relationship and pairwise comparisons of combined control with treated groups are given in Tables 3A and 3B in the appendix for male and female rats, respectively.

Reviewer's findings: Following tumor types showed p-values less than or equal to 0.05 either for dose response relationship and/or pairwise comparisons of combined control and treated groups.

Tumor Types with P-Values ≤ 0.05 for Dose Response Relationship or Pairwise Comparisons

Sex	Organ Name	Tumor Name	Combined Control N=100	Low N=50	Med N=50	High N=50	P-Values Dose Resp	C vs L	C vs M	C vs H
Male	LIVER	HEPATOCELLULAR ADENO	0	2	3	2	0.1747	0.1247	0.0414*	0.1281
		HEPATO CELL ADEN+CAR	0	2	3	3	0.0692	0.1247	0.0414*	0.0449*
	THYMUS	BENIGN THYMOMA	0	5	8	9	0.0050*	0.0049*	<0.001*	<0.001*
Female	LIVER	HEPATOCELLULAR ADENO	0	1	4	4	0.0208*	0.3712	0.0166*	0.0157*
		HEPATO_CELL_ADEN+CAR	0	1	6	4	0.0527	0.3712	0.0020*	0.0157*
	SPLEEN	HAEMANGIOMA	0	1	2	3	0.0244*	0.3712	0.1325	0.0453*
	THYMUS	BENIGN THYMOMA	5	10	21	11	0.0948	0.0132	<0.001*	0.0047*
		MALIGNANT THYMOMA	0	0	0	2	0.0421	.	.	0.1289

Based on the above criterion, the incidences of benign thymoma in thymus in male rats, hepatocellular adenoma and haemangioma in spleen in female rats were considered to have statistically significant positive dose response relationship. Also, the pairwise comparisons of the treated groups with the combined controls marked by the asterisks were considered to be statistically significant for the increased incidence of the related tumors.

3. Mouse Study

Two separate experiments were conducted, one in males and one in females. In each of these two experiments there were three treated groups and two identical vehicle control groups. Two hundred and fifty (b) (4) Crl:CD1 (ICR) mice of each sex were randomly allocated to treated and control groups in equal size of 50 mice. The dose levels for treated groups were 100, 400, and 1500 mg/kg/day. In this review these dose groups would be referred to as the low, medium, and high dose group, respectively. The two controls would be referred to as Control 1 and Control 2. The mice in both control groups received the vehicle (corn oil) by gavage.

During the administration period all mice were observed daily for signs of ill health or overt toxicity. Each mouse was given a detailed physical examination at weekly intervals which included palpation for tissue masses. In addition, due to the severity of the uro-genital findings, the male mice in low, medium, and high dose groups in were subjected to daily uro-genital checks from Day 3 of Week 12. An individual record was maintained of the clinical condition of each mouse. Individual body weights were recorded before treatment on the first day of dosing, at weekly intervals for 16 weeks, approximately once every four weeks thereafter and before necropsy.

In male mice an unexpected in-life finding - swelling of the uro-genital and/or abdominal areas - was recorded in mice given Ospemifene, the onset of which occurred in Week 3 of dosing. The severity of this finding led to the early sacrifice (prior to Week 13) of several male mice from all Ospemifene-treated groups. All male mice from both control groups and all female mice (both control groups and all Ospemifene-treated groups) were unaffected by this finding. Due to the relatively high number of male sacrifice/deaths, it was concluded that the original study objective (i.e. to determine the effects of the test article on the incidence and morphology of tumors following oral administration of the test article to the mouse for 104 weeks) could not

be achieved for male mice. The decision was subsequently taken to cease dosing of all surviving male mice in Week 14 of the study. Designated male mice from selected groups were killed and necropsied in Week 15. All remaining male mice were subject to a treatment-free period, whereupon designated male mice from selected groups were killed and necropsied in Week 27; male mice not selected for necropsy in Week 27 were killed on Week 27 and discarded without necropsy.

3.1. Sponsor's analyses

3.1.1. Survival analysis

Sponsor used similar methodologies to analyze the mouse survival data as they used to analyze the survival data from the rat study.

Sponsor's findings: Sponsor's did not present any data of naturally dead male mice, however reported that 1/51, 0/51, 7/51, 9/51, and 10/51 male mice were sacrificed due to ill health in control 1, control 2, low, medium, and high dose groups, respectively before Week 27. In female mice the sponsor's analysis showed 62.74%, 54.90%, 49.02%, 49.02%, and 58.82% mortality in control 1, control 2, low, medium, and high dose groups, respectively. The overall mortality in the combined control was 58.82%. The sponsor concluded that for female mice the mortality in both control 1, control 2, and Ospemifene-treated animals were generally consistent with the usual patterns of mortality in mice of this strain, with no evidence of effect on mortality due to Ospemifene.

3.1.2. Tumor data analysis

Sponsor used similar methodologies to analyze the mouse tumor data as they used to analyze the tumor data from the rat study.

Sponsor's findings: In sponsor's analysis for tests of increasing dose response, the following tumor types gave rise to results statistically significant at the 5% level:

(a) adrenal subcapsular cell adenoma, (b) adrenal cortical tumor, (c) liver hepatocellular tumor, (d) ovary adenoma/carcinoma, (e) ovary sex cord/stromal tumor, and (f) pituitary adenoma/carcinoma

With tumor types (c), (d), (e) and (f) being classified as common**,** and tumor types (a) and (b) being classified as rare and using the FDA suggested method of adjustment for multiple testing, the sponsor noted the following significant finings:

Adrenal subcapsular cell adenoma: Non-fatal tumors only, overall dose response (P=0.001) and controls v animals given 1500 mg/kg/day (P=0.007).

Adrenal cortical tumor: Fatal and non-fatal tumors combined, overall dose response (P=0.008) and controls v animals given 1500 mg/kg/day (P=0.042).

Ovary sex cord/stromal tumor: Fatal and non-fatal tumors combined, overall dose response (P<0.001), controls v animals given 100 mg/kg/day (P<0.001), controls v animals given 400 mg/kg/day (P<0.001) and controls v animals given 1500 mg/kg/day (P<0.001). Non-fatal tumors only, overall dose response (P<0.001), controls v animals given 1000 mg/kg/day (P<0.001), controls v animals given 400 mg/kg/day (P<0.001) and controls v animals given 1500 mg/kg/day (P<0.001).

In the pairwise comparison between the two control groups, there was no significant difference in the incidence of any tumor type.

3.2. Reviewer's analyses

This reviewer independently performed survival and tumor data analyses from the mouse study. For the mouse data analyses this reviewer used similar methodologies as he used to analyze the data from the rat study. Data used in this reviewer's analyses were provided by the sponsor electronically.

As pointed out earlier due to ill health, all male mice were sacrificed prior to Week 28. There were no data available from the male mice. The sponsor only submitted data of 10 male mice per group of up to Week 27. Due to the lack of adequacy of data, this reviewer did not perform any analysis on male mouse data.

Similar to the rat data, in the submitted mouse data the sponsor identified both control 1 and control 2 by zero (0). Therefore, it was not possible to separate the two control groups. Also as pointed out earlier, for studies with two identical controls, the FDA guidance for carcinogenicity data analysis recommends to analyze the data using the combined control. Due to the way the data were submitted and following the FDA guidance this reviewer analyzed both the survival and the tumor mouse data using the combined control.

3.2.1. Survival analysis

The intercurrent mortality data for female mice is given in Tables 4 in the appendix. The Kaplan-Meier curves for death rate are given in Figures 2 in the appendix. Results for test of dose response relationship and homogeneity of survivals among treatment groups are given in Tables 5 in the appendix.

Reviewer's findings: The reviewer's analysis showed 58.82%, 49.02%, 49.02%, and 58.82% mortality in combined control, low, medium, and high dose group. The tests showed no statistically significant positive dose response relationship in mortality in female mice. Also in female mice, the pairwise comparisons did not show statistically significant increased deaths in any of the treated dose groups compared to the combined control.

3.2.2. Tumor data analysis

The tumor rates and the p-values of the tumor types tested for dose response relationship and pairwise comparisons of combined control and treated groups for female mice are given in Table 6 in the appendix.

Reviewer's findings: Following tumor types showed p-values less than or equal to 0.05 either for dose response relationship and/or pairwise comparisons of combined control with treated groups.

Tumor Types with P-Values ≤ 0.05 for Dose Response Relationship or Pairwise Comparisons

			Cont	Low	Medium	High	P-Value			
	Organ Name	Tumor Name	N=102	N=51	N=51	N=51	Dose Resp	C vs. L	C vs. M	C vs. H
Female	ADRENAL	CORTICAL CARCINOMA	0	0	0	2	0.0409	.	.	0.1197
		SUBCAPSULAR CELL ADE	0	0	1	4	0.0021*	.	0.3491	0.0143*
	LIVER	HEPATOCELLULAR ADENO	0	0	0	2	0.0391	.	.	0.1154
		HEPATO_CELL_ADEN+CAR	0	0	0	3	0.0080*	.	.	0.0403*
	OVARY	BENIGN GRANULOSA CEL	0	1	0	3	0.0141*	0.3551	.	0.0381*
		BENIGN LUTEOMA	5	10	4	4	0.5780	0.0108	0.3934	0.3616
		BENIGN SEX CORD STRO	2	8	13	13	<0.001*	0.0036*	<0.001*	<0.001*
		MALIGNANT GRANULOSA	0	0	3	1	0.1912	.	0.0425*	0.3491
		MALIGNANT LUTEOMA	0	0	5	1	0.2419	.	0.0052*	0.3429
		TUBULOSTROMAL ADENOM	0	2	6	2	0.1806	0.1282	0.0015*	0.1154

Based on the multiple testing adjustment criterion discussed in the rat data analysis section, in female mice the incidences of subcapsular cell adenoma in adrenal, combined incidences of hepatocellular adenomas and carcinomas, benign granulosa cell tumor, and benign sex cord stromal tumor in ovary were considered to have statistically significant positive dose response relationship. Also, the pairwise comparisons of the treated groups with the combined controls marked by the asterisks were considered to be statistically significant for the increased incidence of the related tumors.

4. Summary

In this submission the sponsor included reports of two animal carcinogenicity studies, one in rats and one in mice. These studies were intended to assess the carcinogenic potential of Ospemifene in rats and mice when administered orally via gavage at appropriate drug levels. The length of both the rat and the mouse studies were designed for 104 weeks.

In this review, the phrase "dose response relationship" refers to the linear component of the effect of treatment, and not necessarily to a strictly increasing or decreasing mortality or tumor rate as dose increases.

Rat Study: Two separate experiments were conducted, one in males and one in females. In each of these two experiments there were three treated groups and two identical vehicle control groups. Two hundred and fifty (b) (4) Crl: WI(Hans) rats of each sex were randomly allocated to treated and control groups in equal size of 50 rats. The dose levels for treated groups were 10, 50, and 300 mg/kg/day. In this review these dose groups would be referred to as the low, medium, and high dose group, respectively. The two controls would be referred to as Control 1 and Control 2. The rats in both control groups received the vehicle (corn oil) by gavage.

During the administration period all rats were observed daily for signs of ill health or overt toxicity. In addition, each rat was given a detailed physical examination at weekly intervals which included palpation for tissue masses. Individual body weights were recorded before treatment on the first day of dosing, at weekly intervals for 16 weeks, and once every four weeks thereafter and before necropsy.

The tests showed no statistically significant positive dose response relationship in mortality across treatment groups in either sex. The pairwise comparisons showed statistically significant increased mortality in combined controls compared to any of the treated groups in both sexes. Tests showed statistically significant positive dose

response relationship in the incidences of the incidences of benign thymoma in thymus in male rats, hepatocellular adenoma and haemangioma in spleen in female rats. Pairwise comparisons of the treated groups with the combined controls showed statistically significant increased incidence in the following tumor types.

Sex	Organ name	Tumor name	Comparisons
Male	Liver	Hepatocellular adenoma	Combine control vs. medium dose group
		Hepatocellular aden+car	Combine control vs. medium and high dose groups
	Thymus	Benign thymoma	Combine control vs. all treated groups
Female	Liver	Hepatocellular adenoma	Combine control vs. medium and high dose groups
		Hepatocellular aden+car	Combine control vs. medium and high dose groups
	Spleen	Haemangioma	Combine control vs. high dose group
	Thymus	Benign thymoma	Combine control vs. medium and high dose groups

Mouse Study: Two separate experiments were conducted, one in males and one in females. In each of these two experiments there were three treated groups and two identical vehicle control groups. Two hundred and fifty (b) (4) Crl:CD1 (ICR) mice of each sex were randomly allocated to treated and control groups in equal size of 50 mice. The dose levels for treated groups were 100, 400, and 1500 mg/kg/day.

During the administration period all mice were observed daily for signs of ill health or overt toxicity. Each mouse was given a detailed physical examination at weekly intervals which included palpation for tissue masses. In addition, due to the severity of the uro-genital findings, the male mice in low, medium, and high dose groups in were subjected to daily uro-genital checks from Day 3 of Week 12. Individual body weights were recorded before treatment on the first day of dosing, at weekly intervals for 16 weeks, approximately once every four weeks thereafter and before necropsy.

In male mice an unexpected in-life finding - swelling of the uro-genital and/or abdominal areas - was recorded in mice given Ospemifene, the onset of which occurred in Week 3 of dosing. The severity of this finding led to the early sacrifice (prior to Week 13) of several male mice from all Ospemifene-treated groups. All male mice from both control groups and all female mice (both control groups and all Ospemifene-treated groups) were unaffected by this finding. Due to the relatively high number of male sacrifice/deaths, it was concluded that the original study objective (i.e. to determine the effects of the test article on the incidence and morphology of tumors following oral administration of the test article to the mouse for 104 weeks) could not be achieved for male mice. The decision was subsequently taken to cease dosing of all surviving male mice in Week 14 of the study. Designated male mice from selected groups were killed and necropsied in Week 15. All remaining male mice were subject to a treatment-free period, whereupon designated male mice from selected groups were killed and necropsied in Week 27; male mice not selected for necropsy in Week 27 were killed on Week 27 and discarded without necropsy.

The tests showed no statistically significant positive dose response relationship in mortality in female mice. Also in female mice, the pairwise comparisons did not show statistically significant increased deaths in any of the treated dose groups compared to the combined control. Test showed statistically significant positive dose response relationship in the incidences of subcapsular cell adenoma in adrenal, combined incidences of hepatocellular adenomas and carcinomas, benign granulosa cell tumor, and benign sex cord stromal tumor in ovary in female mice. Also in female mice, the pairwise comparisons of the treated groups with the combined controls showed statistically significant increased incidence in the following tumor types.

Sex	Organ name	Tumor name	Comparisons
Female	Adrenal	Subcapsular cell adenoma	Combine control vs. High dose group
	Liver	Hepatocellular aden+car	Combine control vs. high dose groups
	Ovary	Benign granulosa cell tumor	Combine control vs. High dose group
		Benign sex cord stromal tumor	Combine control vs. all treated groups
		Malignant granulosa cell tumor	Combine control vs. Medium dose group
		Malignant luteoma	Combine control vs. Medium dose group
		Tubulostromal adenoma	Combine control vs. Medium dose group

Mohammad Atiar Rahman, Ph.D.
Mathematical Statistician

Concur: Karl Lin, Ph.D.
Team Leader, Biometrics-6

cc:
Archival NDA 203-505 Ospemifene
Dr. Jeffrey Bray
Mr. Lyght
Dr. Machado
Dr. Lin
Dr. Rahman
Ms. Patrician

5. Appendix

Table 1A: Intercurrent Mortality Rate
Male Rats

	0 mg|kg|day		10 mg|kg|day		50 mg|kg|day		300 mg|kg|day	
Week	No. of Death	Cum. %	No. of Death	Cum. %	No. of Death	Cum. %	No. of Death	Cum. %
0 - 52	3	3.00	2	4.00	1	2.00	.	.
53 - 78	5	8.00	.	.	2	6.00	1	2.00
79 - 91	11	19.00	.	.	1	8.00	1	4.00
92 - 104	8	27.00	.	.	3	14.00	1	6.00
Ter. Sac.	73	73.00	48	96.00	43	86.00	47	94.00

Table 1B: Intercurrent Mortality Rate
Female Rats

	0 mg|kg|day		10 mg|kg|day		50 mg|kg|day		300 mg|kg|day	
Week	No. of Death	Cum. %	No. of Death	Cum. %	No. of Death	Cum. %	No. of Death	Cum. %
0 - 52	2	2.00	.	.	.	.	.	.
53 - 78	10	12.00	.	.	2	4.00	2	4.00
79 - 91	16	28.00	2	4.00	1	6.00	3	10.00
92 - 104	7	35.00	2	8.00	6	18.00	1	12.00
Ter. Sac.	65	65.00	46	92.00	41	82.00	44	88.00

Table 2A: Intercurrent Mortality Comparison
Male Rats

Test	Statistic	P_Value
Dose-Response	Likelihood Ratio	0.3691
Homogeneity	Log-Rank	0.0001

Table 2B: Intercurrent Mortality Comparison
Female Rats

Test	Statistic	P_Value
Dose-Response	Likelihood Ratio	0.3376
Homogeneity	Log-Rank	0.0001

Table 3A: Tumor Rates and P-Values for Pairwise Comparisons of Control 1 and High Dose Group Male Rats

Organ Name	Tumor Name	0 mg Cont N=100	10 mg Low N=50	50 mg Med N=50	300 mg High N=50	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
ADRENAL	BENIGN PHAEOCHROMOCY	1	1	0	0	0.8592	0.5864	1.0000	1.0000
	CORTICAL ADENOMA	0	1	1	1	0.1953	0.3556	0.3507	0.3603
	CORTICAL CARCINOMA	1	0	0	0	1.0000	1.0000	1.0000	1.0000
BRAIN	BENIGN GRANULAR CELL	2	1	2	0	0.7936	0.7357	0.4387	1.0000
	MALIGNANT MENINGIOMA	1	0	0	0	1.0000	1.0000	1.0000	1.0000
CAECUM	MALIGNANT SCHWANNOMA	1	0	0	0	1.0000	1.0000	1.0000	1.0000
CONNECTIVE TISS	HAEMANGIOMA	1	0	0	0	1.0000	1.0000	1.0000	1.0000
EPIDIDYMIS	BENIGN MESOTHELIOMA	1	0	0	0	1.0000	1.0000	1.0000	1.0000
FOOT/LEG	MALIGNANT BASAL CELL	1	0	0	0	1.0000	1.0000	1.0000	1.0000
HAEMOLYMPHORETI	LYMPHOCYTIC LEUKAEMI	4	0	0	0	1.0000	1.0000	1.0000	1.0000
KIDNEY	MALIGNANT NEPHROBLAS	1	0	0	0	1.0000	1.0000	1.0000	1.0000
	TUBULAR CELL ADENOMA	0	0	0	1	0.2121	.	.	0.3603
LIVER	HEPATOCELLULAR ADENO	0	2	3	2	0.1747	0.1247	0.0414*	0.1281
	HEPATOCELLULAR CARCI	0	0	0	1	0.2121	.	.	0.3603
	HEPATO CELL ADEN+CAR	0	2	3	3	0.0692	0.1247	0.0414*	0.0449*
LUNG	BRONCHIOLO-ALVEOLAR	1	0	0	0	1.0000	1.0000	1.0000	1.0000
MAMMARY GLAND	FIBROADENOMA	0	1	0	1	0.2195	0.3556	.	0.3603
MESENTERIC LYMP	HAEMANGIOMA	11	1	2	1	0.9684	0.9960	0.9752	0.9963
	HAEMANGIOSARCOMA	1	0	0	0	1.0000	1.0000	1.0000	1.0000
ORAL CAVITY	SQUAMOUS CELL PAPILL	1	0	0	0	1.0000	1.0000	1.0000	1.0000
PANCREAS	ACINAR CELL ADENOMA	2	0	0	0	1.0000	1.0000	1.0000	1.0000
	ISLET CELL ADENOMA	4	0	1	0	0.9332	1.0000	0.8870	1.0000
	ISLET CELL CARCINOMA	1	0	0	0	1.0000	1.0000	1.0000	1.0000
PARATHYROID	ADENOMA	1	0	0	0	1.0000	1.0000	1.0000	1.0000
PAROTID SALIVAR	HISTIOCYTIC SARCOMA	1	0	0	0	1.0000	1.0000	1.0000	1.0000
PITUITARY	ADENOMA	23	1	2	2	0.9957	1.0000	0.9999	0.9999
PROSTATE	ADENOMA	2	0	0	0	1.0000	1.0000	1.0000	1.0000
SKIN + SUBCUTIS	BENIGN HAIR FOLLICLE	9	0	0	0	1.0000	1.0000	1.0000	1.0000
	DERMAL FIBROMA	2	0	0	0	1.0000	1.0000	1.0000	1.0000
	FIBROLIPOMA	1	0	0	0	1.0000	1.0000	1.0000	1.0000
	FIBROMA	7	0	0	0	1.0000	1.0000	1.0000	1.0000
	FIBROSARCOMA	1	0	0	0	1.0000	1.0000	1.0000	1.0000
	HISTIOCYTIC SARCOMA	1	0	1	0	0.6574	1.0000	0.5768	1.0000

Table 3A: Tumor Rates and P-Values for Pairwise Comparisons of Control 1 and High Dose Group Male Rats

Organ Name	Tumor Name	0 mg Cont N=100	10 mg Low N=50	50 mg Med N=50	300 mg High N=50	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
SKIN + SUBCUTIS	LIPOMA	1	0	0	0	1.0000	1.0000	1.0000	1.0000
	MALIGNANT BASAL CELL	2	0	0	0	1.0000	1.0000	1.0000	1.0000
	SARCOMA - NOS	0	0	0	1	0.2121	.	.	0.3603
	SQUAMOUS CELL CARCIN	1	0	1	1	0.2931	1.0000	0.5830	0.5891
	SQUAMOUS CELL PAPILL	4	2	0	0	0.9885	0.6967	1.0000	1.0000
SPLEEN	HAEMANGIOMA	2	0	3	2	0.1963	1.0000	0.2287	0.4511
STOMACH	ADENOCARCINOMA	1	0	0	1	0.3785	1.0000	1.0000	0.5891
TESTIS	INTERSTITIAL CELL AD	3	0	0	0	1.0000	1.0000	1.0000	1.0000
THYMUS	BENIGN THYMOMA	0	5	8	9	0.0050*	0.0049*	<0.001*	<0.001*
	HAEMANGIOSARCOMA	0	0	1	0	0.4156	.	0.3507	.
THYROID	C-CELL ADENOMA	11	6	2	7	0.2971	0.5991	0.9752	0.4794
	FOLLICULAR CELL ADEN	6	1	0	0	0.9990	0.9565	1.0000	1.0000
	FOLLICULAR CELL CARC	0	0	1	0	0.4156	.	0.3507	.
TONGUE	SQUAMOUS CELL PAPILL	0	1	0	0	0.6234	0.3556	.	.

Table 3B: Tumor Rates and P-Values for Dose Response Relationship and Pairwise Comparisons Female Rats

Organ Name	Tumor Name	0 mg Cont N=100	10 mg Low N=50	50 mg Med N=50	300 mg High N=50	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
ADRENAL	BENIGN PHAEOCHROMOCY	4	0	2	1	0.6260	1.0000	0.7116	0.8963
	CORTICAL CARCINOMA	1	0	0	0	1.0000	1.0000	1.0000	1.0000
BRAIN	BENIGN GRANULAR CELL	2	2	0	1	0.6281	0.4753	1.0000	0.7432
LIVER	HEPATOCELLULAR ADENO	0	1	4	4	0.0208*	0.3712	0.0166*	0.0157*
	HEPATOCELLULAR CARCI	0	0	2	0	0.4159	.	0.1325	.
	HEPATO_CELL_ADEN+CAR	0	1	6	4	0.0527	0.3712	0.0020*	0.0157*
MAMMARY GLAND	ADENOCARCINOMA	2	0	0	1	0.5097	1.0000	1.0000	0.7491
	ADENOMA	1	0	0	0	1.0000	1.0000	1.0000	1.0000
	FIBROADENOMA	27	0	0	0	1.0000	1.0000	1.0000	1.0000
MESENTERIC LYMP	HAEMANGIOMA	2	1	2	0	0.7934	0.7548	0.4667	1.0000
OVARY	BENIGN SERTOLI CELL	0	1	0	1	0.2199	0.3712	.	0.3615
	BENIGN SEX CORD STRO	0	1	0	0	0.6344	0.3712	.	.
	CYSTADENOMA	0	0	1	1	0.1301	.	0.3664	0.3615
PANCREAS	ISLET CELL ADENOMA	1	0	0	1	0.3720	1.0000	1.0000	0.5942
	ISLET CELL CARCINOMA	1	0	0	0	1.0000	1.0000	1.0000	1.0000
PITUITARY	ADENOMA	50	1	0	0	1.0000	1.0000	1.0000	1.0000
	CARCINOMA	1	0	0	0	1.0000	1.0000	1.0000	1.0000
SKIN + SUBCUTIS	BENIGN BASAL CELL TU	1	1	0	0	0.8673	0.6064	1.0000	1.0000
	BENIGN HAIR FOLLICLE	1	0	0	0	1.0000	1.0000	1.0000	1.0000
	FIBROMA	3	0	0	0	1.0000	1.0000	1.0000	1.0000
	SQUAMOUS CELL CARCIN	1	0	0	0	1.0000	1.0000	1.0000	1.0000
	SQUAMOUS CELL PAPILL	1	0	0	1	0.3720	1.0000	1.0000	0.5942
SPLEEN	HAEMANGIOMA	0	1	2	3	0.0244*	0.3712	0.1325	0.0453*
	HAEMANGIOSARCOMA	0	1	0	0	0.6344	0.3712	.	.
STOMACH	SQUAMOUS CELL CARCIN	0	1	0	0	0.6344	0.3712	.	.
TAIL	SQUAMOUS CELL PAPILL	0	1	0	0	0.6344	0.3712	.	.
THORACIC CAVITY	MALIGNANT SCHWANNOMA	1	0	0	0	1.0000	1.0000	1.0000	1.0000
	PARAGANGLIOMA	0	0	1	0	0.4185	.	0.3664	.
THYMUS	BENIGN THYMOMA	5	10	21	11	0.0948	0.0132	<0.001*	0.0047*
	MALIGNANT THYMOMA	0	0	0	2	0.0421	.	.	0.1289
THYROID	C-CELL ADENOMA	9	8	7	4	0.7755	0.2500	0.3470	0.7563
	FOLLICULAR CELL ADEN	0	0	0	1	0.2070	.	.	0.3615
	FOLLICULAR CELL CARC	1	0	0	1	0.3720	1.0000	1.0000	0.5942
UTERUS	ADENOCARCINOMA	5	1	0	2	0.5149	0.9407	1.0000	0.7882
	HISTIOCYTIC SARCOMA	1	0	0	0	1.0000	1.0000	1.0000	1.0000
	LEIOMYOMA	1	0	0	1	0.3705	1.0000	1.0000	0.5906

Table 3B: Tumor Rates and P-Values for Dose Response Relationship and Pairwise Comparisons Female Rats

Organ Name	Tumor Name	0 mg Cont N=100	10 mg Low N=50	50 mg Med N=50	300 mg High N=50	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
UTERUS	MALIGNANT SCHWANNOMA	0	0	0	1	0.2105	.	.	0.3664
	SQUAMOUS CELL CARCIN	0	0	1	1	0.1301	.	0.3664	0.3615
	STROMAL POLYP	8	1	0	0	0.9999	0.9865	1.0000	1.0000
	STROMAL SARCOMA	2	0	0	0	1.0000	1.0000	1.0000	1.0000
VAGINA	STROMAL POLYP	2	0	0	1	0.5031	1.0000	1.0000	0.7432

Table 4: Intercurrent Mortality Rate in Female Mice

	0 mg|kg|day		100 mg|kg|day		400 mg|kg|day		1500 mg|kg|day	
Week	No. of Death	Cum. %	No. of Death	Cum. %	No. of Death	Cum. %	No. of Death	Cum. %
0 - 52	9	8.82	2	3.92	6	11.76	5	9.80
53 - 78	20	28.43	9	21.57	6	23.53	7	23.53
79 - 91	19	47.06	5	31.37	5	33.33	5	33.33
92 - 104	12	58.82	9	49.02	8	49.02	13	58.82
Ter. Sac.	42	41.18	26	50.98	26	50.98	21	41.18

Table 5: Intercurrent Mortality Comparison Female Mice

Test	Statistic	P_Value
Dose-Response	Likelihood Ratio	0.9579
Homogeneity	Log-Rank	0.3080

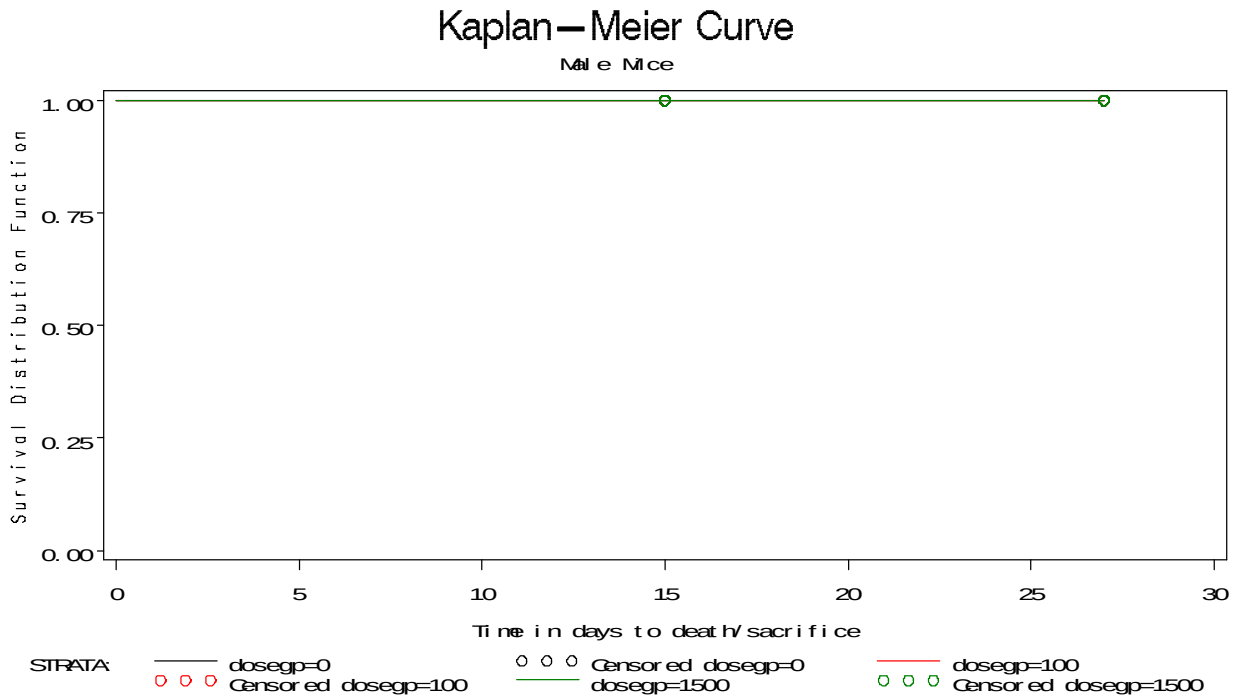
Table 6: Tumor Rates and P-Values for Dose Response Relationship and Pairwise Comparisons Female Mice

Organ Name	Tumor Name	0 mg Cont N=102	100 mg Low N=51	400 mg Medium N=51	1500 mg High N=51	P-Value Dose Resp	P-Value C vs. L	P-Value C vs. M	P-Value C vs. H
ABDOMINAL CAVIT	MALIGNANT MESOTHELIO	0	0	1	0	0.4088	.	0.3551	.
ADRENAL	BENIGN PHAEOCHROMOCY	0	0	0	1	0.2000	.	.	0.3429
	CORTICAL ADENOMA	0	0	0	1	0.2000	.	.	0.3429
	CORTICAL CARCINOMA	0	0	0	2	0.0409	.	.	0.1197
	MALIGNANT PHAEOCHROM	0	0	0	1	0.2044	.	.	0.3491
	SUBCAPSULAR CELL ADE	0	0	1	4	0.0021*	.	0.3491	0.0143*
BONE	FIBROSARCOMA	0	0	0	1	0.2000	.	.	0.3429
	OSTEOMA	1	0	0	1	0.3609	1.0000	1.0000	0.5703
BRAIN	BENIGN MENINGIOMA	0	0	1	0	0.4088	.	0.3551	.
	MALIGNANT MENINGIOMA	1	0	0	0	1.0000	1.0000	1.0000	1.0000
CAECUM	LEIOMYOSARCOMA	0	1	0	0	0.6167	0.3551	.	.
CONNECTIVE TISS	HAEMANGIOSARCOMA	1	0	0	0	1.0000	1.0000	1.0000	1.0000
	LIPOMA	1	0	0	0	1.0000	1.0000	1.0000	1.0000
HAEMOLYMPHORETI	ERYTHROID LEUKAEMIA	1	0	0	0	1.0000	1.0000	1.0000	1.0000
	HISTIOCYTIC SARCOMA	3	0	0	1	0.5997	1.0000	1.0000	0.8187
	LYMPHOCYTIC LEUKAEMI	1	0	0	0	1.0000	1.0000	1.0000	1.0000
	MALIGNANT LYMPHOMA -	12	4	5	3	0.8479	0.8762	0.7722	0.9350
		2	0	0	0	1.0000	1.0000	1.0000	1.0000
		7	6	4	2	0.8544	0.3029	0.6266	0.8803
	MALIGNANT LYMPHOMA-L	0	2	0	1	0.3063	0.1325	.	0.3491
HARDERIAN GLAND	ADENOCARCINOMA	0	0	0	1	0.2000	.	.	0.3429
	ADENOMA	0	1	1	1	0.1855	0.3551	0.3491	0.3491
KIDNEY	TUBULAR CELL ADENOMA	1	0	0	0	1.0000	1.0000	1.0000	1.0000
LIVER	HAEMANGIOMA	0	1	0	0	0.6167	0.3551	.	.
	HEPATOCELLULAR ADENO	0	0	0	2	0.0391	.	.	0.1154
	HEPATOCELLULAR CARCI	0	0	0	1	0.2044	.	.	0.3491
	HEPATO_CELL_ADEN+CAR	0	0	0	3	0.0080*	.	.	0.0403*
LUNG	BRONCHIOLO-ALVEOLAR	1	1	1	3	0.0515	0.5863	0.5784	0.1212
		10	2	7	1	0.9420	0.9649	0.3473	0.9926
MAMMARY GLAND	ADENOCARCINOMA	3	0	0	0	1.0000	1.0000	1.0000	1.0000
	ADENOMA	2	0	0	0	1.0000	1.0000	1.0000	1.0000
MANDIBULAR LYMP	HAEMANGIOMA	0	1	0	0	0.6188	0.3611	.	.
MESENTERIC LYMP	HAEMANGIOMA	1	0	0	0	1.0000	1.0000	1.0000	1.0000
OVARY	BENIGN GRANULOSA CEL	0	1	0	3	0.0141*	0.3551	.	0.0381*
	BENIGN LUTEOMA	5	10	4	4	0.5780	0.0108	0.3934	0.3616
	BENIGN SEX CORD STRO	2	8	13	13	<0.001*	0.0036*	<0.001*	<0.001*

Table 6: Tumor Rates and P-Values for Dose Response Relationship and Pairwise Comparisons Female Mice

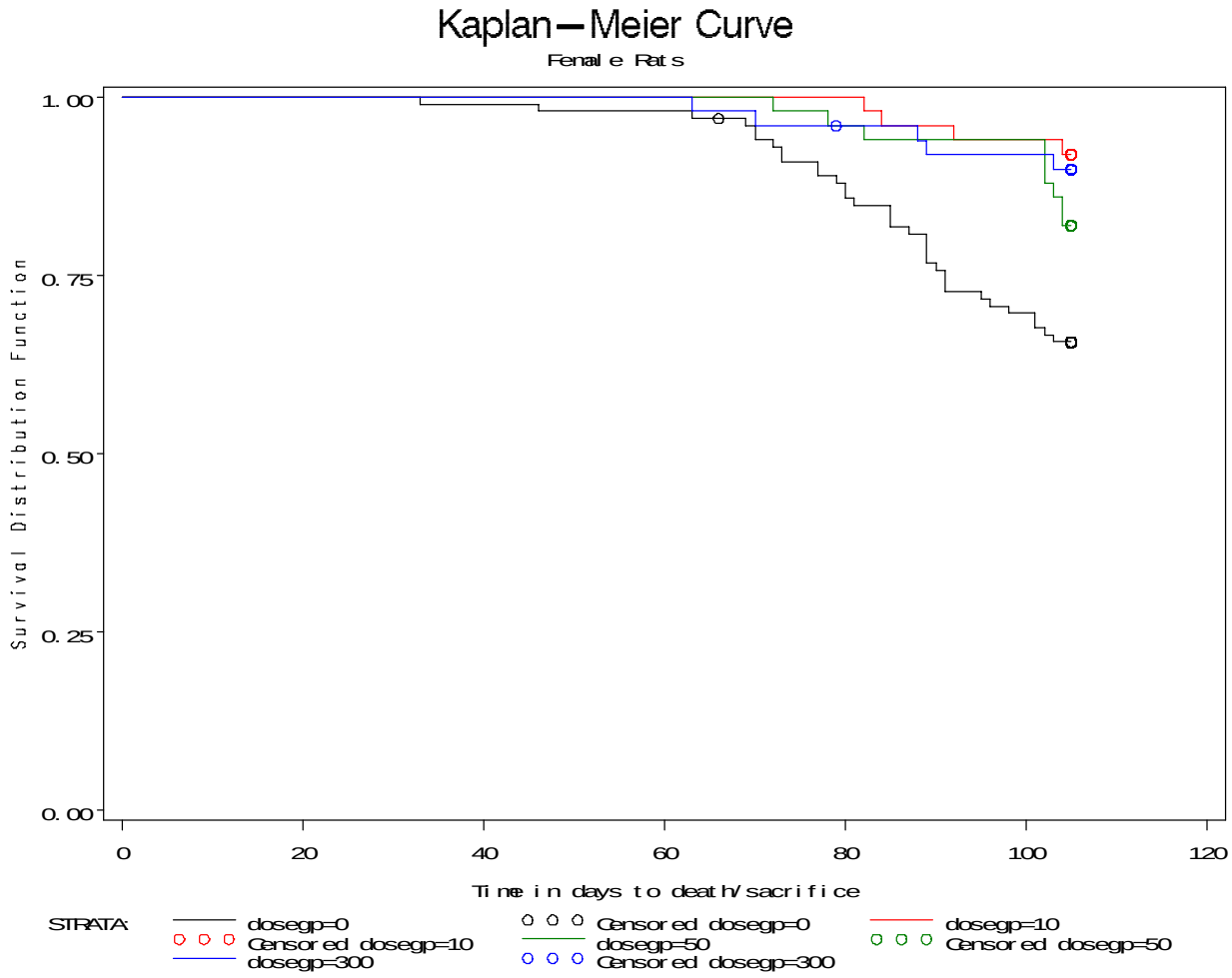
Organ Name	Tumor Name	0 mg Cont N=102	100 mg Low N=51	400 mg Medium N=51	1500 mg High N=51	P-Value Dose Resp	C vs. L	C vs. M	C vs. H
OVARY	BENIGN THECOMA	0	0	2	0	0.4022	.	0.1197	.
	CYSTADENOCARCINOMA	0	1	0	1	0.2067	0.3551	.	0.3429
	CYSTADENOMA	0	1	0	2	0.0594	0.3551	.	0.1197
	MALIG. SEX CORD STRO	0	1	1	2	0.0593	0.3551	0.3551	0.1197
	MALIGNANT GRANULOSA	0	0	3	1	0.1912	.	0.0425*	0.3491
	MALIGNANT LUTEOMA	0	0	5	1	0.2419	.	0.0052*	0.3429
	TUBULOSTROMAL ADENOM	0	2	6	2	0.1806	0.1282	0.0015*	0.1154
	TUBULOSTROMAL CARCIN	0	0	2	0	0.4022	.	0.1197	.
PANCREAS	ISLET CELL ADENOMA	1	0	0	1	0.3609	1.0000	1.0000	0.5703
	ISLET CELL CARCINOMA	0	1	0	0	0.6167	0.3551	.	.
PITUITARY	ADENOMA	1	0	1	2	0.0986	1.0000	0.5784	0.2785
	CARCINOMA	0	0	0	1	0.2044	.	.	0.3491
PREPUTIAL/CLITO	HAEMANGIOSARCOMA	0	1	0	0	0.6167	0.3551	.	.
	SARCOMA NOS	1	0	0	0	1.0000	1.0000	1.0000	1.0000
SKIN + SUBCUTIS	HAEMANGIOSARCOMA	1	0	0	0	1.0000	1.0000	1.0000	1.0000
	MALIGNANT FIBROUS HI	2	0	0	0	1.0000	1.0000	1.0000	1.0000
	SARCOMA - NOS	1	0	0	0	1.0000	1.0000	1.0000	1.0000
	SQUAMOUS CELL PAPILL	0	0	1	0	0.4056	.	0.3491	.
SPINAL CORD	GLIOMA-NOS	1	0	0	0	1.0000	1.0000	1.0000	1.0000
SPLEEN	HAEMANGIOSARCOMA	1	0	0	0	1.0000	1.0000	1.0000	1.0000
TAIL	SARCOMA - NOS	0	0	0	1	0.2000	.	.	0.3429
THYMUS	BENIGN THYMOMA	0	1	0	0	0.6167	0.3551	.	.
UTERUS	BENIGN GRANULAR CELL	1	0	0	0	1.0000	1.0000	1.0000	1.0000
	HAEMANGIOMA	2	0	0	0	1.0000	1.0000	1.0000	1.0000
	HISTIOCYTIC SARCOMA	1	0	0	0	1.0000	1.0000	1.0000	1.0000
	LEIOMYOMA	2	0	0	0	1.0000	1.0000	1.0000	1.0000
	LEIOMYOSARCOMA	3	0	1	0	0.8797	1.0000	0.8260	1.0000
	STROMAL POLYP	12	0	1	2	0.9056	1.0000	0.9974	0.9840
	STROMAL SARCOMA	1	0	1	1	0.2780	1.0000	0.5863	0.5703
VAGINA	MALIGNANT SCHWANNOMA	0	1	0	0	0.6167	0.3551	.	.

Figure 1A: Kaplan-Meier Survival Functions for Male Rats
Male Rats



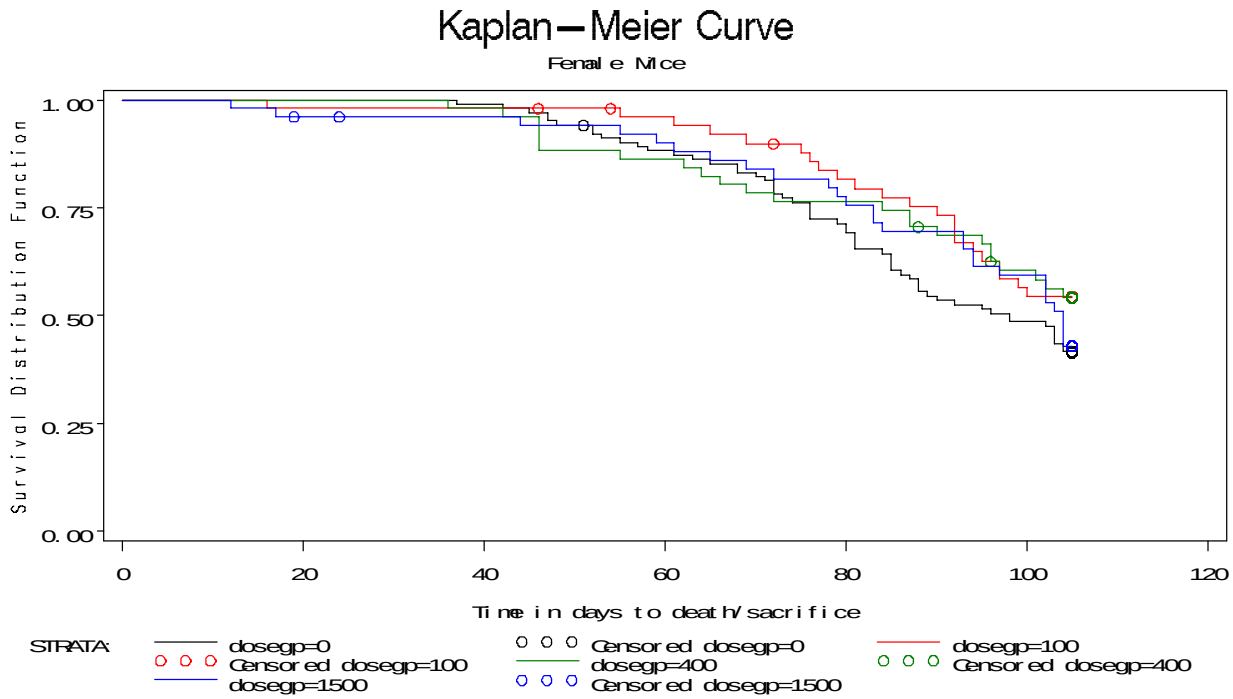
X-Axis: Weeks, Y-Axis: Survival rates

Figure 1B: Kaplan-Meier Survival Functions for Female Rats
Female Rats



X-Axis: Weeks, Y-Axis: Survival rates

Figure 2: Kaplan-Meier Survival Functions for Female Mice
Female Mice



X-Axis: Weeks, Y-Axis: Survival rates

6. References:

1. Peto, R., M.C. Pike, N.E. Day, R.G. Gray, P.N. Lee, S. Parish, J. Peto, Richards, and J.Wahrendorf, "Guidelines for sample sensitive significance test for carcinogenic effects in long-term animal experiments", Long term and short term screening assays for carcinogens: A critical appraisal, International agency for research against cancer monographs, *Annex to supplement, World Health Organization, Geneva,* 311-426, 1980.
2. Bailer AJ, Portier CJ (1988). "Effects of treatment-induced mortality and tumor-induced mortality on tests for carcinogenicity in small samples." *Biometrics*, 44, 417-431.
3. Bieler, G. S. and Williams, R. L. (1993). "Ratio estimates, the delta method, and quantal response tests for increased carcinogenicity". *Biometrics* 49, 793-801.
4. Lin K.K. and Rahman M.A.," Overall false positive rates in tests for linear trend in tumor incidence in animal carcinogenicity studies of new drugs", *Journal of Biopharmaceutical Statistics,* 8(1), 1-15, 1998.
5. Rahman M.A. and Lin K.K, "A comparison of False Positive Rates of Peto and Poly-3 Method For Long Term Carcinogenicity Data Analysis Using Multiple Comparison Adjustment Method Suggested by Lin and Rahman " *Journal of Biopharmaceutical Statistics,* 18: 949-958, 2008.
6. Guidance for Industry. Statistical Aspects of the Design, Analysis, and Interpretation of Chronic Rorent Carcinogenicity Statues of Pharmaceuticals (Draft Guidance). U.S. Department of Health and Human Services, Food and Drug Administation, Center for Drug Evaluation and Research (CDER), May 2001.

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

/s/

MOHAMMAD A RAHMAN
12/03/2012

KARL K LIN
12/03/2012
Concur with review

STATISTICS FILING CHECKLIST FOR A NEW NDA/BLA

NDA Number: 203-505/0000 **Applicant:** Shionogi, Inc **Stamp Date:** 04/26/2012

Drug Name: (b) (4) (Ospemifene Oral Tablets 60 mg) **NDA/BLA Type:** Original/Standard **Indication:** Treatment of moderate to severe symptoms of vulvar and vaginal atrophy due to menopause, such as dyspareunia and/or vaginal dryness

On **initial** overview of the NDA/BLA application for RTF:

	Content Parameter	Yes	No	NA	Comments
1A	Paper Submission: Index is sufficient to locate necessary reports, tables, data, etc.			√	
1B	Electronic Submission: Indexing and reference links within the electronic submission are sufficient to permit navigation through the submission, including access to reports, tables, data, etc.	√			
2	ISS, ISE, and complete study reports are available (including original protocols, subsequent amendments, etc.)	√			
3	Safety and efficacy were investigated for gender, racial, and geriatric subgroups investigated.		√		All subjects are female.
4	Data sets in EDR are accessible and conform to applicable guidances (e.g., existence of define.pdf file for data sets).	√			

IS THE STATISTICAL SECTION OF THE APPLICATION FILEABLE? ___YES_____

Content Parameter (possible review concerns for 74-day letter)	Yes	No	NA	Comment
Designs utilized are appropriate for the indications requested.	√			
Endpoints and methods of analysis are specified in the protocols/statistical analysis plans.	√			
Interim analyses (if present) were pre-specified in the protocol and appropriate adjustments in significance level made. DSMB meeting minutes and data are available.			√	
Appropriate references for novel statistical methodology (if present) are included.			√	
Safety data organized to permit analyses across clinical trials in the NDA/BLA.	√			
Investigation of effect of dropouts on statistical analyses as described by applicant appears adequate.	√			

Information requests for the Applicant: None at this time.

File name: 5_Statistics Filing Checklist for a New NDA_203505

STATISTICS FILING CHECKLIST FOR A NEW NDA/BLA

Background:

Ospemifene Oral Tablets 60 mg for the treatment of vulvar and vaginal atrophy has been developed by previous sponsor (Hormos Medical) since 2003 under IND 67,216. On 04-23-2010, the current sponsor (Shionogi USA, Inc) acquired the product and submitted it to the Agency in this NDA.

In the meeting minutes dated 2007-04-12, the Agency provided the following recommendation to the sponsor: "Please note that in order to demonstrate a "win" for the endpoint of MBS, specify which of the following you intend to demonstrate:

1. That both MB symptoms need to be demonstrate statistical significance for the endpoint of MBS to succeed.

Or

2. That only one of the two MB symptoms needs to demonstrate statistical significance for the endpoint of MBS to succeed. You will need to appropriately adjust the significance level of this test for multiplicity, taking into account the correlation between the two symptoms."

Basically, the Agency want the sponsor to analyze separately the patient self-indentified most bothersome symptom of moderate to severe vaginal dryness or moderate to severe vaginal pain associated with sexual activity.

In the meeting minutes dated 10-29-2009, the Agency told the sponsor that the efficacy would be based on the results of each pivotal Phase 3 study (15-50310, and 15-50821), not on the overall results of the integrated summary of efficacy. In the meeting minutes dated 05-12-2011, the Agency conveyed to the sponsor that "the primary efficacy analyses should be based on subjects meeting all three of the baseline inclusion criteria: vaginal pH greater than 5, less than 5% superficial cells on vaginal smear, and a most bothersome moderate to severe vaginal symptom."

Sponsor's Efficacy Results:

The reported primary efficacy results from the two pivotal phase-3 studies are summarized in Tables 1-2. In both studies, the % parabasal cells, the % superficial cells, the vaginal pH, and the vaginal pain associated with sexual activity are statistically significantly improved in patients on ospemifene 60 mg compared with those on placebo. However, only one study (15-50310) shows vaginal dryness is statistically significantly improved in patients on ospemifene 60 mg compared with those on placebo; the other study (15-50821) does not show vaginal dryness is statistically significantly improved. The results for mITT analysis are similar and are reported in the summary of clinical efficacy instead of in each of the two individual studies.

Table 1. Summary of Primary Efficacy at Week 12 in Study 15-50310 (ITT)

Co-primary endpoint	Ospemifene 30 mg			Ospemifene 60 mg			Placebo	
	N	Mean CFB	p-value	N	Mean CFB	p-value	N	Mean CFB
% Parabasal Cells	274	-21.9	<0.001	272	-30.1	<0.001	261	3.98
% Superficial Cells	274	7.78	<0.001	272	10.8	<0.001	261	2.18
Vaginal pH	282	-0.67	<0.001	276	-1.01	<0.001	268	-0.10
MBS-Vaginal Dryness	102	-1.22	0.04	118	-1.26	0.021	104	-0.84
MBS-Vaginal Pain Associated with Sex	136	-1.02	0.20	120	-1.19	0.023	122	-0.89

CFB: change from baseline
Source: Tables 14.2.4.1, Table 14.2.4.2, Table 14.2.4.3, Table 14.2.4.4, and Table 14.2.4.5

STATISTICS FILING CHECKLIST FOR A NEW NDA/BLA

Table 2. Summary of Primary Efficacy at Week 12 in Study 15-50821 (ITT)
Mean Change from Baseline

Co-primary endpoint	Vaginal Dryness			Vaginal Pain associated with Sex		
	Ospemifene 60 mg (N=160)	Placebo (N=154)	p-value	Ospemifene 60 mg (N=303)	Placebo (N=302)	p-value
% Parabasal Cells	-31.7	-3.9	<0.0001	-40.3	-0.4	<0.0001
% Superficial Cells	7.0	0.0	<0.0001	7.0	0.0	<0.0001
Vaginal pH	-0.95	-0.25	<0.0001	-0.94	-0.07	<0.0001
MBS-Vaginal Dryness	-1.3	-1.1	0.0803	--	--	--
MBS-Vaginal Pain Associated with Sex	--	--	--	-1.5	-1.2	0.0001

Source: Tables 14.2.1.1.1. and 14.2.1.1.2

Xin Fang, Ph.D.	06/19/2012
Reviewing Statistician	Date
Mahboob Sobhan, Ph.D.	06/19/2012
Supervisor/Team Leader	Date

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

/s/

XIN FANG
07/12/2012

MAHBOOB SOBHAN
07/13/2012