CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

203108Orig1s000

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

CARCINOGENICITY STUDIES

NDA:	**203108 (Re-analysis of data under IND 76362)**
IND:	**76362**
Drug Name:	Olodaterol Resplamat® Inhalation Spray
Indication:	**Chronic obstructive pulmonary disease**
Applicants:	**Boehringer Ingelheim Pharmaceuticals Inc.** **900 Ridgebury Road** **Ridgefield, CT 06877**
	CRO: [redacted] (b) (4)
Date:	**Amended reports submitted 10/03/2012**
Review Priority:	**Standard**
Biometrics Division:	**Division 6**
Statistical Reviewer:	**Steve Thomson**
Concurring Reviewers:	**Karl Lin, Ph.D.**
Medical Division:	Pulmonary, Allergy and Rheumatology **Products**
Toxicologist Team:	Marcie L. Wood, Ph.D
	Carol Rivera-Lopez, Ph.D.
Project Manager:	Christine Chung, **Ph.D.**

Keywords: Carcinogenicity, Cox regresson, Kaplan-Meier product limit, Survival analysis, Trend test

Table of Contents

1. EXECUTIVE SUMMARY 3

1.1. CONCLUSIONS AND RECOMMENDATIONS 3
1.2. BRIEF OVERVIEW OF THE STUDIES 8
1.3. STATISTICAL ISSUES AND FINDINGS 8
1.3.1. Statistical Issues 8
1.3.2. Statistical Findings 13

2. INTRODUCTION 13

2.1. OVERVIEW 13
2.2. DATA SOURCES 13

3. STATISTICAL EVALUATION 14

3.1. EVALUATION OF EFFICACY 14
3.2. EVALUATION OF SAFETY 14

4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS 32

5. SUMMARY AND CONCLUSIONS 32

5.1. STATISTICAL ISSUES AND COLLECTIVE EVIDENCE 32
5.2. CONCLUSIONS AND RECOMMENDATIONS 32

APPENDICES 33

APPENDIX 1. FDA SURVIVAL ANALYSIS 33
APPENDIX 2. BAYESIAN SURVIVAL ANALYSIS ..... **ERROR! BOOKMARK NOT DEFINED.**
APPENDIX 3. FDA POLY-K TUMORIGENICITY ANALYSIS 38
APPENDIX 4. REFERENCES 53

Reference ID: 3230218

1. EXECUTIVE SUMMARY

The Sponsor states that Olodaterol Resplamat, in these reports labeled as compound BI 1744 CL (described as a long acting β_2-agonist), is being developed by the Sponsor as an inhalation drug for the treatment of respiratory disorders. The objective of these studies was to assess the oncogenic potential in rats and mice following daily inhalation administration of a nebulised aerosol for 104 weeks. Both studies were conducted at [redacted] (b) (4) After the data were analyzed with the submission for IND 76362, the Sponsor noted that some records that indicated when an organ was not analyzed were missing. The Sponsor sent updated data sets and reports incorporating these data updates. This report is based on updated data. However, the bottom line is that in terms of the interpreted results of statistical tests there were no particularly significant differences between the previous analysis and the analysis of the updated data. Thus, the actual results of the analysis sections are basically unchanged. However, for completeness, the overall analysis is repeated in this report.

1.1. Conclusions and Recommendations

The product is described as "Dosing solutions of BI 1744 CL at the appropriate concentrations were prepared by adding BI 1744 CL, benzalkonium chloride [redacted] (b) (4) disodium edetate [redacted] (b) (4) to sterile water [redacted] (b) (4) and using citric acid [redacted] (b) (4)

"Vehicle formulation was prepared by adding benzalkonium chloride [redacted] (b) (4) disodium edetate [redacted] (b) (4) to sterile water [redacted] (b) (4) and using citric acid [redacted] (b) (4) " (page 18 of rat report)

Gross aspects of the basic study designs for the main study animals in both species are summarized below:

Table 1. Design of Rat Study (dosed at µg/kg/day)

Treatment Group	# Animals	Nominal (Target) Dosage	Estimated (Actual) Dosage
1. Air Control	55	0	0
2. Vehicle Control	55	0	0
3. Low	55	25	25.8
4. Medium	55	75	75.9
5. High	55	250	270

Reference ID: 3230218

Table 2. Design of Mouse Study (dosed at µg/kg/day)

Treatment Group	# Animals	Nominal (Target) Dosage	Estimated (Actual) Dosage
1. Air Control	60	0	0
2. Vehicle Control	60	0	0
3. Low	60	25	26.1
4. Medium	60	75	76.9
5. High	60	250	255

More detailed descriptions of the studies are provided in Sections 3.2.1 and 3.2.2 below. Simple summary life tables are presented in these sections of the report. In Appendix 1, Figures A.1.1 and A.1.2 for rats, and Figures A.1.3 and A.1.4 for mice, display Kaplan-Meier estimated survival curves for each study group for each species and gender combination. These provide estimates of the proportion surviving at time points. The results of the tests of trend and differences in survival are displayed in Tables 3 and 4 below:

Table 3. Statistical Significances of Tests of Homogeneity and Trend in Survival in the Rat Study

Hypothesis Tested	Males		Females	
	Log rank	Wilcoxon	Log rank	Wilcoxon
Rats Homogeneity over Groups 1-5	0.0396	0.0565	0.0659	0.0681
Homogeneity over Groups 2-5	0.0297	0.0439	0.0404	0.0435
No trend over Groups 2-5	0.5781	0.5876	0.0499	0.0750
No Difference Between Groups 1 vs 5	0.3832	0.3972	0.9915	0.8552
No Difference Between Groups 2 vs 5	0.0659	0.0751	0.0136	0.0188
No Difference Between Groups 1 vs 2	0.3201	0.3346	0.0178	0.0187

From the Kaplan-Meier curve depicted in Figure A.1.1 of Appendix 1, in male rats, roughly after day 500 there appears to be a tendency for a separation of the treatment groups in that the vehicle has the highest survival, followed by the air control, largely intertwined with the high and medium groups, and the low group having the greatest mortality (i.e. lowest survival). This is consistent with the results above, i.e., particularly later in the study there is some, possibly equivocal, evidence of herterogeneity across dose groups 1-5 (Logrank p = 0.0396, Wilcoxon p = 0.0565) and in groups 2-5 (Logrank p = 0.0297, Wilcoxon p = 0.0439), but there is no apparent evidence of a dose related trend over groups 2-5 (Logrank p = 0.5781, Wilcoxon p = 0.5876). There is no evidence of a pairwise difference between Groups 1 and 5 (Logrank p = 0.3832, Wilcoxon p = 0.3972), nor between the controls, i.e. Groups 1 and 2 (Logrank p = 0.3201, Wilcoxon p = 0.3346). However, the differences between the Vehicle and the High dose are close to the usual level of statistical significance, 0.05, i.e., Groups 2 and 5, (Logrank p = 0.0751, Wilcoxon p = 0.0659). But recall the principle that no strong proof for an effect is not necessarily proof of no effect.

Reference ID: 3230218

Similarly, it seems from Figure A.1.2, in female rats, after day 400+ or so there also appears to be a tendency for a separation of the treatment groups in that the vehicle control has the highest survival, followed by the high and low dose groups, although these cross. Finally, the highest mortality (i.e. lowest survival) is in the intertwined medium dose group and air controls (i.e. Groups 4 and 1). As with male rats, this is consistent with the results of the statistical tests above, i.e., there is weak evidence of herterogeneity across dose groups 1-5 (Logrank p = 0.0659, Wilcoxon p = 0.0681) and somewhat stronger among groups 2-5 (Logrank p = 0.0404, Wilcoxon p = 0.0435). However, unlike male rats, in females there is some evidence of a dose related trend over groups 2-5 (Logrank p = 0.0499, Wilcoxon p = 0.0750). Also, in this case, there is evidence of a pairwise difference between Group 2 and Group 1 (Logrank p = 0.0178, Wilcoxon p = 0.0187) and Group 5 (Logrank p = 0.0136, Wilcoxon p = 0.0188).

Table 4. Statistical Significances of Tests of Homogeneity and Trend in Survival in the Mouse Study

Hypothesis Tested	Males		Females	
	Log rank	Wilcoxon	Log rank	Wilcoxon
Mice Homogeneity over Groups 1-5	0.4608	0.3984	0.0693	0.2209
Homogeneity over Groups 2-5	0.5370	0.4547	0.0431	0.1531
No trend over Groups 2-5	0.3794	0.3798	0.3396	0.3582
No Difference Between Groups 1 vs 5	0.6823	0.6525	0.4731	0.4888
No Difference Between Groups 2 vs 5	0.2292	0.2141	0.2211	0.2707
No Difference Between Groups 1 vs 2	0.0992	0.0748	0.5908	0.6308

From Figure A.1.3 in Appendix 1, for male mice, from day 300, the vehicle group had the lowest survival, with all other groups fairly closely intertwined. This is consistent with the results of the tests of homogeneity and trend anove (all six $p \geq 0.3794$). As noted above, no strong evidence of differences may not be strong evidence of no differences, but it is consistent with such a conclusion. Similarly there is no statistically significant evidence of differences between the high dose and the two controls (all four $p \geq 0.2141$). However, the tests of differences between Groups 1 and 2 are somewhat close to statistical significance (Logrank p = 0.0992, Wilcoxon p = 0.0748).

In female mice, from Figure A.1.4 below, after day 500 or so. it seems that the medium dose group has the lowest survival (highest mortality) with the vehicle group having the highest survival. The survival curves of the other dose groups seem to be intertwined between these limits. Thus there is weak, somewhat equivocal evidence of overall treatment group differences at the end of the study (Groups 1-5: Logrank p = 0.0693, Groups 2-5: Logrank p = 0.0431). The corresponding Wilcoxon tests do not come close to the usual significance levels (Groups 1-5: Wilcoxon p = 0.2209, Groups 2-5: Wilcoxon p = 0.1531). There is no statistically significant test of trend (Logrank p = 0.3396, Wilcoxon p = 0.3582), plus, none of the six pairwise tests was close to statistical significance (all six $p \geq 0.2211$).

Reference ID: 3230218

The IND (b) (4) report includes a Bayesian analysis of survival that may be of interest, but it is not reproduced here.

.

Of course in a carcinogenicity study, primary interest is on the occurrence of cancers. Statistical analysis compares tumor incidence over dose groups. Table 5, below, displays those organ tumor combinations that had at least one test of trend or pairwise difference from control that was statistically significant at the usual, nominal 0.05 level. For each species by gender by organ combination, the number of animals analyzed and used in the statistical tests is presented first. The tumor incidence for each organ is presented next, with the significance levels of the tests of trend, and the results of pairwise tests between the Air alone, Low, Medium, and High dose groups against the Vehicle group (i.e., Groups 1, 3, 4, and 5 versus 2). These statistical tests are conditional upon the animals actually evaluated, ignoring those not analyzed. Complete tumor incidence tables for each organ are provided in Tables A.2.2 through A.2.5 in Appendix 3. The period '.' in these tables denotes the p-values of tests of dose groups with no tumors in any group involved in the test.

Table 5. Potentially Statistically Significant Neoplasms Rats and Mice

Organ/ Tumor	Air	Veh	Low	Mid	Hi	trend	Hi vs Veh	Med vs Veh	Low vs veh	Air vs Veh
Male Rats										
PITUITARY GLAND										
# Evaluated	53	54	55	55	55					
Adjusted # at risk	50.6	52.0	52.9	53.3	52.1					
ADENOMA, ANTERIOR LOBE[B]	16	15	25	16	18	.5779	.3369	.5253	.0345	.7127
Female Rats										
OVARY										
Adjusted # at risk	45.5	51.2	48.9	44.6	47.9					
LEIOMYOMA [B]	0	0	1	0	4	.0092	.0494	.	.4848	.
Male Mice										
SKIN AND SUBCUT										
# Evaluated	60	60	60	60	60					
Adjusted # at risk	51.6	47.6	51.6	51.0	49.9					
Fibrosarc./Sarc./Rhab-domyosarc	1	6	3	3	2	.8979	.9744	.9341	.9380	.0441
Female Mice										
Systemic										
# Evaluated	60	60	60	60	60					
Adjusted # at risk	44.3	45.1	43.6	39.3	42.6					
HAEMANGIOSARCOMA [M]	1	1	0	0	3	.0479	.2824	1	1	.7584
UTERUS										
# Evaluated	60	60	60	60	60					
Adjusted # at risk	44.3	46.3	47.2	44.7	46.7					
LEIOMYOMA [B]	5	9	18	19	22	.0126	.0038	.0138	.0385	.2177
Adjusted # at risk	43.9	45.0	45.1	41.4	42.6					
LEIOMYOSARCOMA [M]	1	2	5	10	4	.4039	.3160	.0092	.2263	.5087
Adjusted # at risk	44.3	46.3	48.7	46.7	47.5					
Leiomyoma/Leiomyosarcoma	6	11	23	28	26	.0141	.0018	.0003	.0133	.1648
Adjusted # at risk	44.5	45.6	44.4	42.0	44.4					
POLYP [B]	7	4	8	8	12	.0248	.0227	.1339	.1655	.9086

Reference ID: 3230218

Note that this table is almost identical to that included in the original analysis of IND 76362, and comments on the statistically significant results below are identical to those in that report . For readability the order of the pairwise results are reversed from those in the IND 76362 report.

To adjust for the multiplicity of tests the so-called Haseman-Lin-Rahman rules discussed in Section 1.3.1.5, below, are often applied. That is, when testing for trend over dose and the difference between the highest dose group with a control group, to control the overall Type I error rate to roughly 10% for a standard two species, two sex study, one compares the unadjusted significance level of the trend test to 0.005 for common tumors (incidence > 1%) and 0.025 for rare tumors, and the pairwise test to 0.01 for common tumors and 0.05 for rare tumors. As also discussed in section 1.3.1.4, using these adjustments for other tests can be expected to increase the overall type I error rate to some value above the nominal rough 10% level, possibly considerably higher than the nominal 10% rate.

Using the incidence in the Air dose group to define whether a tumor is rare or not, we note that leiomyoma of the ovary would be classified as a rare tumor in female rats, and thus the test of trend would be considered to be highly statistically significant ($p=0.0092 \leq 0.025$) while the comparison between the High dose and the Vehicle dose group would be considered barely statistically significant ($p = 0.0494 \leq 0.05$). Since it uses all dose groups we would expect that the test of trend is a more powerful test than the pairwise test. Even accepting the increase in Type I error, the comparison between the Low dose and the Vehicle in anterior lobe adenoma of the pituitary in male rats would not be classified as statistically significant ($p = 0.0345 > 0.01$). Even ignoring the multiplicity adjustment, no other test results in either rat gender even achieved the nominal 0.05 level.

In male mice the comparison between the Vehicle dose group and the Air dose group in pooled Fibrosarcarcoma/Sarcoma/Rhabdomyoma of the skin and subcutis achieved the 0.05 level, but since it would be classified as a common tumor (incidence in the Air group > 1%), when adjusting for the multiplicity of tests it would not be statistical significamt ($p=0.0441 > 0.01$). In female mice the pairwise comparisons between the Vehicle dose group and the High dose group in benign leiomyoma and pooled leiomyoma/leiomyosarcoma would be considered statistically significant ($p = 0.0038, 0.0018 \leq 0.01$, respectively). The difference between Vehicle and the Medium dose groups in malignant leiomyosarcoma, and pooled leiomyoma/ leiomyosarcoma would also be classified as statistically significant ($p = 0.0092, 0.0003 \leq 0.01$). But that including these latter tests can raise the overall Type I error to some level above the nominal 10%. No other comparisons in female mice achieved the multiplicity adjusted level of significance.

Again, note that the results summarized above are verbatim identical to those in the original analysis for the data submitted with IND 76362

Reference ID: 3230218

1.2. Brief Overview of the Studies

Again, the studies were conducted at (b) (4)

(b) (4) **Study 667930: BI 1774CL: 24 Month Inhalation Carcinogenicity Study in Rats,**

and a similar study:

(b) (4) **Study 668138: BI 1774CL: 24 Month Inhalation Carcinogenicity Study in Mice.**

1.3. Statistical Issues and Findings

1.3.1. Statistical Issues

In this section several issues, typical of statistical analyses of these studies, are considered. These issues include details on the survival analyses, tests on tumorigenicity, multiplicity of tests on neoplasms, and the validity of the designs.

1.3.1.1. Multiple Housing and Dosing of Animals:

The Sponsor's report indicates that unless reduced by mortality, in the rat study animals were housed 4-5 per cage while male mice were housed singly, but female mice were housed in groups of up to 3 animals. Social interaction is important for the welfare of the animals. However, an argument could be made that housing animals together should be treated as part of the treatment of the experiment, and thus the appropriate unit of analysis would be the cage of the animals, not the individual animal. This would reduce the study sample size and probably dramatically decrease power.

Further, carcinogenicity tendencies that could be communicated across animals in that competition for food, fighting, or other within cage effects could cause positive or negative correlations in response. When animals in a cage are dosed together or at the same time, variations in dosing across occasions might also induce positive or negative correlations. Thus, it is possible that within treatment estimated variances may be too large or too small, resulting in conservative or anti-conservative tests (in terms of Type I error). Unless it has been clearly shown that tumor incidence is independent of cage or group in chamber, from a purely statistical analysis point of view, this reviewer would generally recommend single housing and dosing of animals. On the other hand, it would seem that with aerosol dosing the correlation of the within cage treatment effects should be small. So a case can be made that the effect of multiple housing may be close to ignorable.

1.3.1.2. Survival Analysis:

The survival analyses presented here are based on both the log rank test and the Wilcoxon test comparing survival curves. Log rank tests tend to put higher weight on later events, while the Wilcoxon test tends to weight events more equally, and thus is more sensitive to earlier differences in survival. The logrank test is most powerful when the survival curves

Reference ID: 3230218

track each other, and thus the hazards, i.e., the conditional probability of the event in the next infinitesimal interval, would be roughly proportional. This is the test used by the Sponsor. In the FDA analysis, both tests were used to test both homogeneity of survival among the treatment groups and the effect of dose on trend in survival. Appendix 1 reviews the specific animal survival analyses in more detail. The results of the Sponsor's analysis are summarized in Sections 3.2.1.1 and 3.2.2.1.

1.3.1.3. Multiplicity of Tests on Survival:

Using the logrank and Wilcoxon tests, for each gender in each species there are 12 tests of survival differences. If we were to assume that any set of tests are independent across comparisons, which clearly they are not, and assume that there is absolutely no difference in survival, the probability of at least one statistically significant result in each gender in each species is about 0.46. Such is the possible price paid for the multiplicity of hypothesis tests in the frequentist paradigm.

1.3.1.4. Tests on Neoplasms:

The Sponsor's analyses use Cochran-Armitage tests of trend and Peto analyses of neoplasms. The Cochran-Armitage test is designed to test for trend or pairwise differences in event incidence, weighted by dose. However, it ignores differential mortality across dose groups. Inspecting a large number of studies, Bailer and Portier (1988) noted that survival time seemed to fit a Weibull distribution, generally with a shape parameter of between 1 and 5, with 3 a typical value. With t_{max} denoting the maximal time to terminal sacrifice and t_{obs} the time to death of the animal, they propose weighting the animal by $(t_{obs}/t_{max})^k$, so that an animal that survives for say 52 weeks in 104 week study without the tumor being analyzed is counted as $(1/2)^k$ of an animal. For k = 3, that means that animal would count as 1/8 of an animal. Further, the k = 3 specification seems to represent tumor incidence where some animals are perhaps more sensitive and respond earlier to the insult than the remining animals. Under this structure time to incidence would tend to follow a cubic expression. Thus an animal with the specific tumor being studied or who survives to terminal sacrifice without the tumor will be given a weight of 1 when counting the number of animals at risk. However, animals that die early without the tumor are down weighted when counting the number of animals in the risk set for that specific tumor. With differential mortality, as observed in male rats, this can mean a substantial reduction in the size of that risk set. Note this seems to be an appropriate adjustment for dose groups that are terminated early as in the rats study. This weighting is the basis of the poly-k test. In the tumor incidence tables in this report the effective size of the risk set for each tumor is listed in the row labeled "Adjusted # at risk", and seems to be a more appropriate denominator when comparing incidence rates than the simple unadjusted number evaluated. It was noted in the report of the Society of Toxicological Pathology "town hall" meeting in June 2001 that the poly-k modification of the Cochran-Armitage tests of trend has been recommended over the corresponding Peto tests.

In general, the test of a dose related trend uses more data and is thus more powerful than the individual pairwise tests. Further, if there actually is a linear trend, this test is more powerful than the overall tests of homogeneity. Note that both the Air alone and the Vehicle

Reference ID: 3230218

dose groups can be described as 0 doses for testing trend. The aerosol is formed from the drug and vehicle. If the dosage is based on only the aerosol concentration for a fixed dosing period (i.e., the amount of drug plus vehicle product put in the air per unit volume) then the vehicle dosage is proportional to the drug dosage, so the natural zero or intercept is the air only dose. Then the test of trend should be based on this with the combined drug product. If the drug dosage used to create the different aerosols differs across treatment groups and vehicle volume remains more or less constant (as is typical of most forms of dosing, e.g. dietary, gavage, injection, etc.) then the natural zero dose group is the vehicle only dose group. In such a circumstance, the test of trend is based on only the measure of the actual drug dose, assuming the vehicle is constant, not the combined drug product. It seems that the latter specification is appropriate for these studies.

1.3.1.5. Multiplicity of Tests on Neoplasms:

Frequentist hypothesis testing involves accepting or rejecting hypotheses about the parameters of interest on the basis of the values of some statistic. If one does not provide some sort of multiplicity adjustment to the significance level, the chances of rejecting one or more true null hypothesis increases as the number of such tests increases. To avoid this it is common to adjust for multiplicity in hypothesis testing resulting in an adjustment in experiment-wise Type I error (i.e., the probability of rejecting a true null hypothesis). Based on his extensive experience with such carcinogenicity analyses in standard laboratory rodents, for pairwise tests between the high dose group and controls in two species, Haseman (1983) claimed that for a roughly 0.10 (10%) overall false positive error rate, rare tumors should be tested at a 0.05 (5%) level, and common tumors (with a historical control incidence greater than 1%) at a 0.01 level. Similarly, Lin and Rahman (1998) showed that tests of trend should be tested at a 0.025 level for rare tumors and 0.005 for common tumors. This approach is intended to balance both Type I error and Type II error (i.e., the error of concluding there is no evidence of a relation to tumorgenicity when there actually is such a relation).

But note these comments apply only to the test of trend and a single comparison between a control and high dose groups. Significance levels of the pairwise tests between the vehicle groups and the low and medium dose groups are also provided in the tests reported in the FDA analysis. Including these tests can be expected to increase the overall type I error rate to some level above the rough 10% level. Even if one uses the Haseman-Lin-Rahman rules above, the overall type I error associated with including the tests between the vehicle and the air, low, and medium dose groups may be considerably larger than the rough 10% error rate.

1.3.1.6. Validity of the Designs:

When determining the validity of designs there are two key points:

1) adequate drug exposure,
2) tumor challenge to the tested animals.

1) is related to whether or not sufficient animals survived long enough to be at risk of forming late-developing tumors and 2) is related to the Maximum Tolerated Dose (MTD), designed to achieve the greatest likelihood of tumorigenicity.

Reference ID: 3230218

Lin and Ali (2006), quoting work by Haseman, have suggested that in standard laboratory rodent species, a survival rate of about 25 animals, out of 50 or more animals, between weeks 80-90 of a two-year study may be considered a sufficient number of survivors as well as one measure of adequate exposure. Note that as a percentage of animals that survived to week 91, this criterion is met or exceded in the high dose group in either gender in either species. (Please see Tables 16 and 17, below, in rats, and Tables 27 and 28 in mice). Like the other comments in this section this requires the expertise of the toxicologist, but it may suggest that the MTD was not exceeded.

The mean weight values and derived differences and ratios in the following table were taken directly from the Sponsor's reports (Rat Tables 5 and 6, pages 76-83, Mouse Tables 5 and 6, pages 81-88). The change from baseline in the table below is the simple difference between the means at the specified dates, and thus animals that die in the interim are only counted at the study initiation, not at the end of the study.

Table 6. Mean Weights and Changes (in g) in Rats

Dose Group	Est. Dose µg/kg/ day	Males				Females			
		Week		Change from baseline	% change relative to control	Week		Change from baseline	% change relative to control
		0	104			0	104		
1. Air Control	0	151	544	393		126	324	198	
2. Vehicle Control	0	154	561	407	104%	126	328	202	102%
3. Low	25.8	151	512	361	92%	127	322	195	98%
4. Medium	75.9	150	543	393	100%	126	320	194	98%
5. High	270	150	504	354	90%	127	317	190	96%

Table 7. Mean Weights and Changes (in g) in Mice

Dose Group	Est. Dose µg/kg/ day	Males				Females			
		Week		Change from baseline	% change relative to control	Week		Change from baseline	% change relative to control
		0	104			0	101		
1. Air Control	0	32.1	39.2	7.1		24.4	34.7	10.3	
2. Vehicle Control	0	31.7	39.3	7.6	107%	26.0	33.9	7.9	77%
3. Low	26.1	31.1	39.0	7.9	111%	25.1	35.3	10.2	99%
4. Medium	76.9	31.1	39.4	8.3	117%	25.2	37.9	12.7	123%
5. High	255	31.6	41.1	9.5	134%	25.4	35.6	10.2	99%

Chu, Ceuto, and Ward (1981), citing earlier work by Sontag *et al* (1976) recommend that the MTD "is taken as 'the highest dose that causes no more than a 10% weight decrement as compared to the appropriate control groups, and does not produce mortality, clinical signs of toxicity, or pathologic lesions (other than those that may be related to a neoplastic response) that would be predicted to shorten male animal's natural life span' " From Table 6 above, this weight decrement criterion was satisfied in rats. In male mice there seems to be an actual weight gain in the Vvehicle and treated dose groups. In female mice the only dose group that exceded the lower limit criterion was the Vehicle control, which suggests this may only be an artifactual result.

Reference ID: 3230218

Although this requires the expertise of the toxicologist, this may be evidence that the MTD was not exceeded in either species.

The Sponsor summarizes food consumption during the rat study as follows: "Although sometimes variable, the BI 1744 CL treated rats generally consumed more food compared to control group animals for both males and females, which attained statistical significance throughout most of the study period. The difference reduced towards the end of study. A generally dose-related increase in food consumption was noted in treated females, but not in males. (page 48 of rat report)

In mice, food consumption was summarized as follows: "In males, food consumption was generally higher in BI 1744 CL treated groups compared with control groups throughout the study. The difference attained statistical significance across the majority of the treatment period when compared to Vehicle Control. Although there was a lack of notable difference among treatment groups, generally males dosed at 255 µg/kg/day [,i.e., the High dose,] had slightly higher food consumption compared with groups dosed at 26.1 µg/kg/day and 76.9 µg/kg/day [i.e., the Low and Medium doses].

"In females, groups dosed at 76.9 µg/kg/day and 255 µg/kg/day generally consumed more food than other groups throughout the study period. A statistically significant difference between groups treated at 76.9 µg/kg/day or higher and the Vehicle Control group was observed intermittently and was more persistent towards the end of the study." (page 50 of mice report)

Again from 2) above, excess mortality in the high or higher dose groups not associated with any tumor or sacrifice in the higher dose groups might suggest that the MTD was exceeded. This suggests that a useful way to assess whether or not the MTD was achieved is to measure early mortality not associated with any identified tumor. If this is high in the higher dose groups it suggests that animals tend to die before having time to develop tumors. Tables 8 and 9, below, display the number of animals in each dose group that died of a natural death or moribund sacrifice, but did not show any tumors (i.e., the "Event"):

Table 8. Natural Death with No Identified Tumor in Rats (Male/Female)

	1. Air Control	2. Vehicle Control	3. Low	4.Medium	5. High
Males Event	1	0	2	3	4
No event	54	55	53	52	51
Females Event	2	0	2	3	3
No event	53	55	53	52	52

In both rat genders the incidence of the event is quite low. In male rats the results are somewhat equivocal in that there is an apparent trend in the occurrence of the event over increasing dose, but, again incidence is quite low. The time adjusted test of differences when comparing time to event in the high dose group in male rats to the vehicle is barely statistically significant (Logrank p = 0.0356, Wilcoxon p=0.0363), while the comparison to the air control

Reference ID: 3230218

group is not (Logrank p = 0.1601, Wilcoxon p=0.1849). That would seem to suggest that the evidence that the MTD was exceded is present but weak. In female rats there is no evidence of a difference between the high dose and the air control, while the comparison to the vehicle control is, at best, equivocal (Logrank p = 0.0624, Wilcoxon p=0.0667).

Table 9. Natural Death with No Identified Tumor in Mice (Male/Female)

	1. Air Control	2. Vehicle Control	3. Low	4.Medium	5. High
Males Event	7	8	7	7	10
No event	53	52	53	53	50
Females Event	10	12	6	7	4
No event	50	48	54	53	56

In male mice there is no strong evidence of such early deaths in the high dose, whether adjusted for time of death or not. Strictly speaking, such tests do not show there is no dose related effect, only that there is no strong evidence of such an effect. In female mice there is evidence of a trend, but it is negative and is associated with decreasing non-tumor related from the controls to the actual dose groups. Thus it does not seem to be associated with exceeding the MTD. Note again, interpreting these results requires the expertise of the toxicologist.

1.3.2. Statistical Findings

Please see Section 1.1 above.

2. INTRODUCTION

2.1. Overview

This submission summarizes the results of two year rat and mouse inhalation studies to assess the carcinogenic potential carcinogenic potential of aerosoled BI 1744 CL when dosed daily for about 104 consecutive weeks. Both studies were conducted at (b) (4)

2.2. Data Sources

As with the original submission for IND 76362, the Sponsor provided two SAS® formatted transport files, both labeled tumor.xpt, each containing a SAS tumor data set named tumor.sas7bdat, both seemingly fitting the FDA requested format for such studies. Note that the SAS carcinogenicity data sets requested by the FDA for analysis have three primary types of records. The first type reports each tumor in each animal with its associated organ,. The second provides a record for each organ in each animal that was NOT microscopically examined. The third type is only needed if no organs were checked or if all organs were checked but no tumors were found. The original data sets supplied by the Sponsor were apparently missing a number of the records of the second type above. After noting this apparent deficiency, the Sponsor sent corrected data sets as above. This re-analysis uses these updated data sets. However, the sctual impact of the changes in these records is minimal.

Reference ID: 3230218

3. STATISTICAL EVALUATION

3.1. Evaluation of Efficacy

NA

3.2. Evaluation of Safety

3.2.1. (b) (4) Study 667930: BI 1774CL: 24 Month Inhalation Carcinogenicity Study in Rats,

STUDY DURATION: 104 Weeks (planned)
EXPERIMENTAL (DOSING) START DATE: 29 January 2008
FINAL DOSING DATE: 31 January 2010
TREATMENT DURATION: Males 99 to 104 weeks depending upon dose group
Females 97 to 104 weeks depending upon dose group
RAT STRAIN: Sprague Dawley Crl:CD(SD) Rats
ROUTE: Daily nose-only inhalation

The Sponsor indicates that 345 male and 345 female Wistar Han (Crl:WI(Han)) rats were assigned to 5 dose groups and treated as shown in the table below:

Table 10. Design of Rat Study (dosed at µg/kg/day)

Treatment Group	# Main Study Animals	Nominal (Target) Dosage	Estimated (Actual) Dosage	# Toxicity Animals[1]
1. Air Control	55	0	0	10
2. Vehicle Control	55	0	0	10
3. Low	55	25	25.8	10
4. Medium	55	75	75.9	10
5. High	55	250	270	10

[1] There were an additional 4 animals/group labeled as toxicity spares

The Sponsor states that "Test aerosols were generated using Heart airjet nebulisers. The animals were exposed to test aerosol for a target 35 min period daily for at least 104 weeks using a snout-only inhalation exposure technique." (page 14 of rat report)

Dosing was justified as follows: "The inhalation route of administration was selected for this study as this route has been defined by the Sponsor for clinical use. Snout only exposure was used since this is the most applicable method of exposure for the test model.

"The dose levels used on this study were selected by the Sponsor, in discussion with the Study Director based on the results obtained from a 13 Week Inhalation MTD Study in Rats with a 4 Week Recovery Period ((b) (4) Study No. 666413, Report No. 26689 . . .). In this study histological investigations revealed local irritation of the respiratory tract including squamous

Reference ID: 3230218

metaplasia of the laryngeal epithelium in the ventral region and necrosis of the U shaped cartilage. These effects were attributable to the inclusion of benzalkonium chloride in the vehicle. An increase in the incidence and severity was observed in animals treated with BI 1744 BS at 200 µg BI 1744 BS/kg/day and above. The overall no systemic effect level defined by histological examination was considered to be 50 µg BI 1744 BS/kg/day." (page 22 of report)

"All the animals were weighed and the required number of males and females were randomly assigned to 5 treatment groups and the group mean body weights compared to ensure homogeneity. Males and females were randornised separately." (page 20 of report)

Animals of the same gender were housed 4-5 together, unless reduced by mortality. Although it is probably kinder to the animals, as discussed in Section 1.3.1.1, multiple housing may cause problems with the analysis of study results, although this would be particularly true with other forms of dosing. Apparently food and water were available *ad libitum*.

3.2.1.1. Sponsor's Results and Conclusions

This section will present a summary of the Sponsor's analysis on survivability and tumorigenicity in rats.

Sponsor's Survival Analysis:

"In total 153 out of 550 Toxicity animals were found dead, killed in extremis or died incidentally during the minimum 104 week treatment period. In addition, 41 out of 140 Toxicokinetic animals were killed in extremis or died prematurely during the period." (page 46)

"The overall mortality was generally higher in females compared with males." (page 46)

"In both genders, mortality was reduced in the Vehicle Control group when compared to Clean Air Control, reaching statistical significance in females (p=0.018). This difference could not be based on a specific cause of death identified in pathology and was, thus, considered to be a chance observation. Mortality was, in BI 1744 CL treated groups, generally comparable to Clean Air Control, but marginally higher when compared to Vehicle Control. The logrank test for pairwise comparison revealed statistical significance for 25.8 µg/kg/day males (p=0.003), for 75.9 µg/kg/day females (p=0.010) and 270 µg/kg/day females (p=0.014). These differences were due to the low mortality in Vehicle Control groups and, thus, interpreted as incidental. Only in 25.8 µg /kg/day males, mortality was slightly higher also when compared to Clean Air Control." (page 46 of rat report)

The following table presents a slightly reformatted version of the unnumbered table on page 46 of the rat report:

Reference ID: 3230218

Table 11. The incidence of premature deaths in each group is summarised in the table below

Gender	Dose Group/Treatment				
	Clean Air Control	Vehicle Control	25.8 µg/kg/day	day 75.9 µg/kg/day	270 µg/kg/day
Male	11/55	7/55	21/55	12/55	15/55
Female	20/55	9/55	15/55	21/55	21/55

Results are presented as premature deaths/total number of toxicity animals in group"

Sponsor's Tumorigenicity analysis:

The Sponsor summarizes their analysis by stating that: "There was no drug-associated trend in single, multiple, benign, malignant or metastasizing tumours." (pge 53 of rat report)

Further: "Treatment-related neoplastic fmdings were confined to the smooth muscles of the mesovarium.

"Leiomyomas ofthe mesovarian tissue were observed in 1155 females ofthe 25.8 µg/kg/day dose group, and 4/55 females of the 270µg/kg/day dose group . . . . The incidence of mesovarian leiomyomas [B] was significantly higher in the 270 µg/kg/day dose group when compared with the Vehicle Control group (p=0.042). The test for a linear trend with dose was also statistically significant (p=0.008) for this rare tumour. This tumour type was not observed in animals from any of the other groups. Minimal or moderate hyperplasia ofmesovarian smooth muscle . . . was observed in 3/55 females of the 75.9 µg/kg/day dose group, and 3/55 females of the 270 µg/kg/day dose group. This fmding was not observed in animals from other groups.

"The incidence of benign leiomyomas in the 270 µg/kg/day dose group was outside the range of historical control data of the laboratory which conducted the study and of RITA historical control data . . . and considered test item related. The isolated case of a leiomyoma in the group treated at 25.8 µg/kg/day was considered incidental as no benign leiomyomas were observed in the group treated at 75.9 µg/kg/day, indicating the lack of a dose-response relationship. In the 25.8 µg/kg/day dose group, the incidence of benign leiomyomas was within the range of historical control data." (pages 53-54 of rat report)

The following table is copied from the Sponsor's unnumbered table on page 54 of the rat report:

Reference ID: 3230218

Table 12. Un-numbered Sponsor Table: "Selected benign proliferative smooth muscle lesions in the female genital tract of rats"

		Clean Air	Vehicle	25.8 µg/kg	75.9 µg/kg	270 µg/kg	Historical Control Data Mean (range) [%]	
							RITA**	(b) (4)
N=		55	55	55	55	55	5519	1214
Mesovarium, leiomyoma	N	0	0	1	0	4*		
	%	0	0	1.8	0	7.3	< 0.1 (0-2)	0
Mesovarium, hyperplasia smooth muscle	N	0	0	0	3	3		
	%	0	0	0	5.5	5.5	0	N.A.
Uterus, leiomyoma	N	0	0	0	0	0	5519	1218
	%	0	0	0	0	0	0.2(0.0-3.0)	1.0 (0.0-3.9)

* Peto analysis: Trend: p<0.025; pairwise difference p<0.05
**RITA: Registry of Industrial Toxicology Animal-data
N.A. Not applicable

The Sponsor's report provided the following discussions under the heading **"Incidental neoplastic findings."** Note that the Sponsor seems to be using the "Incidental" label to denote tumors they consider unrelated to the test drug, instead of the usual specification for a Peto Analysis where the term denotes a non-fatal tumor discovered when the animal dies or is sacrificed.

Pituitary gland

"The incidence of Pituitary Gland Adenoma, Anterior Lobe [B] was higher in male rats of the 25.8 µg/kg/day dose group when compared with the Vehicle Control group, leading also to a higher incidence of cause of death due to this lesion. Statistical analysis of the tumour incidence revealed a significant pairwise difference to Vehicle Control (p=0.003), but the test for a linear trend with dose was not statistically significant. The incidence of hyperplasia of the pituitary gland pars distalis, representing the preneoplastic precursor lesion of adenomas, was similar in all BI 17 44 CL treated groups and the Vehicle Control group. It was even slightly lower in these groups when compared to the Clean Air Control group.

"This minor inter-group differences in tumour incidence and the related difference in cause of death due to this tumour were considered fortuitous because there was no dose-response relationship, the lack of a corroborative increase of pars distalis hyperplasia, and because the incidences were in all groups within the historical control data of the laboratory which conducted the study." (page 54 of report)

The tables below were copied from the unnumbered tables in the Sponsor's report (pages 55-56), with minor editing changes, in particular numbers (in curly braces seem to be correct):

Reference ID: 3230218

Table 13. Un-numbered Sponsor Table: "Incidence of pituitary gland adenomas in male rats"

		Clean Air	Vehicle	25.8 µg/kg	75.9 µg/kg	270 µg/kg	Historical Control Data Mean (range) [%]	
							RITA**	(b) (4)
N=		55	55	55	55	55	5519	1216
Adenoma anterior lobe	N	15	16	25*	16	18		
	%	25.5 {27.7}	25.5 {29.1}	45.6	29.1	34.5	27.3* # (5.0-56.3)*#	40.7 (18,6-66.0)
Deaths caused by Adenoma anterior lobe	N	4	1	12	6	5		N.A.

* Peto analysis: Ppairwise difference p<0.01
**RITA: Registry of Industrial Toxicology Animal-data
N.A. Not applicable

The report notes that RITA data are not considered as reliable historical control data (page 55 of rat report).

Brain

"In males a malignant astrocytoma was recorded in the Clean Air Control group and the 25.8 µg/kg/day dose group; a benign astrocytomas was recorded in the 270 µg/kg/day dose group. The only glial cell tumour found in the spinal cord was a malignant astrocytoma in a Clean Air Control male. In females, two benign astrocytomas and one malignant glioma were recorded in the 25.8 µg/kg/day dose group. The incidence of astrocytomas was statistically not significant and largely within the range of historical control data. The slightly higher incidence of this tumour was considered incidental due to the lack of this tumour in females dosed at 75.9 and 270 µg/kg/day." (page 55 of rat report)

Table 14. Un-numbered Sponsor Table: "Incidence of brain tumours of glial cell origin in male rats"

		Clean Air	Vehicle	25.8 µg/kg	75.9 µg/kg	270 µg/kg	Historical Control Data Mean (range) [%]	
							RITA**	(b) (4)
N=		55	55	55	55	55	5624	1219
Astrocytoma [B] & [M]	N	0	0	1	0	1		
	%	0	0	1.8	0	1.8	0.3 (0.0 -4.0)	0.8 (0.0-3.3)
Glioma[M]	N	0	0	0	0	0		
	%	0	0	0	0	0	0.1 (0.0 -2.0)	0.5 (0.0-4.1)

**RITA: Registry of Industrial Toxicology Animal-data

Reference ID: 3230218

Table 15. Un-numbered Sponsor Table: "Incidence of brain tumours of glial cell origin in female rats"

		Clean Air	Vehicle	25.8 µg/kg	75.9 µg/kg	270 µg/kg	Historical Control Data Mean (range) [%]	
							RITA**	(b) (4)
N=		55	55	55	55	55	5624	1219
Astrocytoma [B] & [M]	N	0	0	2	0	0		
	%	0	0	3.6	0	0	0.1 (0.0 -2.0)	0.3 (0.0-3.1
Glioma[M]	N	0	0	1	0	0		
	%	0	0	1.8	0	0	< 0.1 (0.0 -2.0)	0.0

**RITA: Registry of Industrial Toxicology Animal-data

3.2.1.2. FDA Reviewer's Results

This section will present the Agency findings on survival and tumorigenicity in male and female rats.

Survival analysis:

The following tables (Table 16 for male rats, Table 17 for females) summarize the mortality results for the study groups. The data were grouped for the specified time period, and present the number of deaths during the time interval over the number at risk at the beginning of the interval. The percentage cited is the percent that survived at the end of the interval. In these tables the terminal period only includes those animals were sacrificed. Animals that died of other causes during the terminal period are included in the preceding, but overlapping time period. The Kaplan-Meier survival plots in Appendix 1 provide a more detailed picture of the profile of mortality losses.

Table 16. Summary of Male Rats Survival (estimated dose in µg/kg/day)

Period (Weeks)	Air Control 0	Vehicle Control 0	Low 25.8	Medium 75.9	High 270
1-52	0/55[1] 100%[2]	1/55 98.2%	1/55 98.2%	3/55 94.5%	0/55 100%
53-78	2/55 96.4%	1/54 96.4%	2/54 94.5%	2/52 90.9%	4/55 92.7%
79-91	4/53 89.1%	2/52 92.7%	5/52 85.5%	4/50 83.6%	3/51 87.3%
92-104	5/49 80%	3/50 87.3%	13/47 61.8%	3/46 78.2%	8/48 72.7%
Terminal[3] 105	44	48	34	43	40

[1] number of deaths / number at risk
[2] overall per cent survival to end of period.
[3] number of animals that survived to terminal sacrifice

Reference ID: 3230218

Table 17. Summary of Female Rats Survival (estimated dose in µg/kg/day)

Period (Weeks)	Air Control 0	Vehicle Control 0	Low 25.8	Medium 75.9	High 270
1-52	1/55[1] 98.2[2]	0/55 100%	1/55 98.2%	2/55 96.4%	1/55 98.2%
53-78	5/54 89.1%	1/55 98.2%	5/54 89.1%	6/53 85.5%	2/54 94.5%
79-91	8/49 74.5%	6/54 87.3%	1/49 87.3%	7/47 72.7%	7/52 81.8%
92-104	6/41 63.6%	2/48 83.6%	8/48 72.7%	6/40 61.8%	11/45 61.8%
Terminal [3] 105	35	46	40	34	34

[1] number of deaths / number at risk
[2] overall per cent survival to end of period.
[3] number of animals that survived to terminal sacrifice

Table 18 below provides the significance levels of the tests of homogeneity and trend over dose groups as proposed in Section 1.3.1.1, above.

Table 18. Statistical Significances of Tests of Homogeneity and Trend in Survival in the Rat Study

Hypothesis Tested	Males		Females	
	Log rank	Wilcoxon	Log rank	Wilcoxon
Rats Homogeneity over Groups 1-5	0.0396	0.0565	0.0659	0.0681
Homogeneity over Groups 2-5	0.0297	0.0439	0.0404	0.0435
No trend over Groups 2-5	0.5781	0.5876	0.0499	0.0750
No Difference Between Groups 1 vs 5	0.3832	0.3972	0.9915	0.8552
No Difference Between Groups 2 vs 5	0.0659	0.0751	0.0136	0.0188
No Difference Between Groups 1 vs 2	0.3201	0.3346	0.0178	0.0187

From the Kaplan-Meier curve in Figure A.1.1 in Appendix 1, it seems that in male rats, roughly after day 500 there appears to be a tendency for a separation of the treatment groups in that the Vehicle control has the highest survival, followed by the Air control, largely intertwined with the High and Medium dose groups. Eventually, the Low dose group has the lowest survival (i.e. highest mortality). This is consistent with the results above, i.e., there is some evidence of herterogeneity across dose groups 1-5 (Logrank p = 0.0396, Wilcoxon p = 0.0565) and groups 2-5 (Logrank p = 0.0297, Wilcoxon p = 0.0439), but no evidence of a dose related trend over groups 2-5 (Logrank p = 0.5781, Wilcoxon p = 0.5876). There is no evidence of a pairwise difference between Groups 1 and 5 (i.e., Air and High doses: Logrank p = 0.3832, Wilcoxon p = 0.3972), nor between the controls, i.e., Groups 1 and 2 (Logrank p = 0.3201, Wilcoxon p = 0.3346). However, at the usual 0.05 level, the differences between the vehicle and high dose, i.e. Groups 2 and 5, are close to significance (Logrank p = 0.0751, Wilcoxon p = 0.0659).

Reference ID: 3230218

Similarly from Figure A.1.2, in female rats, after day 400+ there also appears to be a tendency for a separation of the treatment groups in that the vehicle control has the highest survival, followed by the high and low dose groups, although both of the latter survival curves cross. Finally, the highest mortality (i.e., lowest survival) is in the Medium, High, Air Control dose groups (i.e., Groups 1, 4 and 5). As with male rats, this is consistent with the results of the statistical tests above, i.e., there is some evidence of herterogeneity across dose groups 1-5 (Logrank p = 0.0659, Wilcoxon p = 0.0681) and groups 2-5 (Logrank p = 0.0404, Wilcoxon p = 0.0435). However in this gender there is some evidence of a dose related trend over groups 2-5 (Logrank p = 0.0499, Wilcoxon p = 0.0750). In this case, there is also evidence of a pairwise difference between Group 2 against both Group 1 (Logrank p = 0.0178, Wilcoxon p = 0.0187) and Group 5 (Logrank p = 0.0136, Wilcoxon p = 0.0188).

Results from a supporting experimental Bayesian nonparametric analysis of survival are provided in Appendix 2.

Tumorigenicity analysis:

As discussed in Section 1.3.1.5, the Haseman-Lin-Rahman rules for adjusting for multiplicity in a two species study specify that for a roughly 0.10 (10%) overall false positive error rate, we would reject the hypothesis of no trend if the observed p-vlaue is below 0.025 (2.5%) in rare tumors, and 0.005 (0.5%) in common tumors, plus the pairwise tests between the high dose group and controls should be tested at a 0.05 (5%) level in rare tumors and 0.01 (1%) in common tumors. This approach is intended to balance both Type I error and Type II error (i.e., the error of concluding there is no evidence of a relation to tumorgenicity when there actually is such a relation). The following, table 19 below, is a repeat of the rat section in table 5 above, and display those organ-tumor combinations that were associated with at least one test with a unadjusted significance level of 0.05 or less. In this table the treatment groups are denoted by “Air” for the air control, “Veh” for vehicle control, “Low” for the low dose, “Med” for the middle dose, and “Hi” for the dose.

Table 19. Potentially Statistically Significant Neoplasms Rats

Organ/ Tumor	Air	Veh	Low	Mid	Hi	trend	Hi vs Veh	Med vs Veh	Low vs veh	Air vs Veh
Male Rats										
PITUITARY GLAND										
# Evaluated	53	54	55	55	55					
Adjusted # at risk	50.6	52.0	52.9	53.3	52.1					
ADENOMA, ANTERIOR LOBE[B]	16	15	25	16	18	.5779	.3369	.5253	.0345	.7127
Female Rats										
OVARY										
Adjusted # at risk	45.5	51.2	48.9	44.6	47.9					
LEIOMYOMA [B]	0	0	1	0	4	.0092	.0494	.	.4848	.

Using the incidence in the Air dose group to define whether a tumor is rare or not, we note that leiomyoma of the ovary would be classified as a rare tumor in female rats, and thus the test of trend would be considered to be highly statistically significant (p=0.0092 ≤ 0.025) while the comparison between the High dose and the Vehicle dose group would be considered barely statistically significant (p = 0.0494 ≤ 0.05). Note that since it uses all dose groups the test of

Reference ID: 3230218

trend is a more powerful test than the pairwise test. Even accepting the increase in Type I error, the comparison between the Low dose and the Vehicle in anterior lobe adenoma of the pituitary in male rats would not be classified as statistically significant ($p = 0.0345 > 0.01$). Even ignoring the multiplicity adjustment, no other test results in either rat gender even achieved the nominal 0.05 level.

Complete incidence table in male rats and female rats are presented in Tables A.2.4 and A.2.5 in Appendix 3.

3.2.2. (b) (4) Study 668138: BI 1774CL: 24 Month Inhalation Carcinogenicity Study in Mice,

STUDY DURATION: 104 Weeks (planned)
EXPERIMENTAL START DATE (INITIATING DOSING): 13 December 2007
END OF TREATMENT: Males 15 November - 6 December 2009 (Days 704-724)
Females Group 5 2- 6 December 2009 (Days 721-725)
1-4, 6 13-16 December 2009 (Days 732-735)
MOUSE STRAIN: Charles River Crl:CD-1® Mice
ROUTE: Daily nose only inhalation for 60 minutes

The Sponsor summarized study conduct in the following table (mostly a copy of Table 2) as follows:

Table 20. Design of Mouse Study (dosed at µg/kg/day)

Treatment Group	# Main Study Animals	Nominal (Target) Dosage	Estimated (Actual) Dosage	# Toxicity Animals[1]
1. Air Control	60	0	0	20
2. Vehicle Control	60	0	0	20
3. Low	60	25	26.1	36
4. Medium	60	75	76.9	36
5. High	60	250	255	36

[1] There were an additional 4 animals/group labeled as toxicity spares

"Test aerosols were generated using airjet nebulisers. The animals were exposed to test aerosol for a target 40 min[ute] period daily for at least 104 weeks using a snout-only exposure technique except for Group 4 females for which the dosing was terminated in Week 103 due to surviving animals dropping to 25%.

"Two batches of young adult CD1 mice were ordered. From the first batch supplied, 448 males and 448 females were allocated to 5 dose groups including toxicokinetic spares. From the second batch supplied, 20 males and 20 females were allocated to 5 dose groups for the purpose of Day 1 toxicokinetic bleeding (4 animals/sex/group)." (page 14 of mouse report)

Reference ID: 3230218

"The overall group mean aerosol concentrations of total formulation collected were 0.03, 0.04, 0.04 and 0.04 mg/L for Groups 2, 3, 4 and 5, respectively. The overall group mean aerosol concentration values based on analytical determination of BI 1744 BS were 0.526, 1.56 and 5.23 µg/L respectively for Groups 3, 4 and 5, respectively." (page 14 of mouse report)

"The overall mean achieved doses of BI 1744 BS, for males and females combined, were 26.1, 76.9 and 255 µg/kg/day for Groups 3, 4 and 5, respectively. Overall doses for males were 25.4, 75.5 and 246 µg BI 1744 BS/kg/day and for females were 26.8, 78.5 and 264 µg BI 1744 BS/kg/day for Groups 3, 4 and 5, respectively. The difference in achieved doses between males and females were due to the differences in body weights." (page 47 of mouse report)

" The exposure to BI 1744 BS was moderate to high and AUC-data increased almost propotionally to dose. Exposure levels to BI 1744 BS were 483-/477-fold, 1294-/1201-fold, and 5142-/3596-fold the anticipated human exposure level (determined at 11.9 pmol/h/L) for male/female animals receiving 26.1, 76.9, 255 µg/kg/day, respectively. No relevant change of the BI 1744 BS plasma concentrations was observed after repeated dosing. The BI 1744 BS plasma concentrations were similar in males and females study (Week 103). (page 15 of mouse report)

"Toxicokinetic investigations revealed that overall, the variability of the BI 1744 BS plasma concentrations was moderate." (page 15)

3.2.1.1. Sponsor's Results and Conclusions

This section will present a summary of the Sponsor's analysis on survivability and tumorigenicity in mice.

Survival analysis:

The Sponsor summarizes results as follows: "In total 293 out of 600 Toxicity animals were killed *in extremis* or died prematurely during the treatment period. In addition, 47 out of 336 Toxicokinetic animals were killed *in extremis* or died prematurely during the period. The details of premature decedents are presented in the table below." (page 48 or mouse report)

Table 21. Sponsor Un-numbered Table of Premature Deaths

	Sex/Group No.									
	Male					Female				
	1	2	3	4	5	1	2	3	4*	5
Toxicity Animals										
Total in Group	60	60	60	60	60	60	60	60	60	60
Premature Deaths	19	27	23	27	22	32	30	32	45	36
Toxicokinetic Animals										
Total in Group	24	24	40	40	40	24	24	40	40	40
Premature Deaths	4	5	4	5	4	4	6	2	7	6

* Group 4 females terminated on Day 718, Week 103 due to the number of surviving animals dropping to 25%, incidence calculated up to this point.

Reference ID: 3230218

The Sponor notes that: "Slightly high mortality in the female 76.9 µg/kg/day dose group was attributable to a higher number of incidental deaths in this group, in particular death due to technical procedures, haemorrhagic cysts of the ovary, and haemorrhages at different sites.

"Due to mortality reaching 75%, the dosing of 76.9 µg/kg/day females was ceased on Day 718 of study (Week 103). The decision to terminate this dose group in Week 103 was made after receipt of a corresponding recommendation from the US-FDA Carcinogenicity Advisory Committee upon consultation by the Sponsor.

"Low mortalities were noted in Toxicokinetic animals. This was considered to be due to most Toxicokinetic animals being sacrificed in the first year of study." (page 48 of mouse report)

"Kaplan-Meier estimated survival functions were calculated for the Clean Air Control, Vehicle Control and each of 3 dose groups treated with BI 1744 CL. Based on pairwise comparisons between the Vehicle Control and all other groups using the logrank test modified for censored survival data for males and females separately, mortality in the Intermediate Dose group females was found to be significantly higher than that seen in the Vehicle Control females (p=0.030). There were no other statistically significant differences in mortality between the Vehicle Control and any other groups." (page 48 of mouse report)

Tumorigenicity analysis:

The Sponsor summarizes statistical tumorigenicity results as follows:

Neoplastic findings attributable to dosing with BI 1744 CL

"The incidence of uterine leiomyomas and leiomyosarcomas was increased in BI 1744 CL treated groups reaching a level of statistically significant pairwise difference at 76.9 µg/kg/day for leiomyosarcomas (p=0.006) and at 255 µg/kg/day for leiomyomas (p=0.003). The incidence of both findings exceeded the range of internal laboratory and RITA historical control data not only in the BI 1744 CL treated groups, but also in the Vehicle Control group and in case of the uterine leiomyomas also in the Clean Air Control group. One case of a leiomyosarcoma of the oviduct was present at 255 µg/kg/day in addition." (pages 53-54 of rat report)

Reference ID: 3230218

Table 22. Un-numbered Sponsor Table: "Incidence of smooth muscle tumours in oviduct and uterus of female mice"

	Vehicle	Clean Air	26.1 µg/kg	76.9 µg/kg	255 µg/kg	Historical Control Data Mean (range) [%]	
						(b) (4)	RITA[a]
N=	60	60	60	60	60	707	1921
Oviduct, leiomyosarcoma (rare) Number	0	0	0	0	1		
%	0.0	0.0	0.0	0.0	1.7	0.0[b]	0.1 (0.0-1.7)[b]
N=	60	60	60	60	60	709	(≤1920)[c]
Uterus, leiomyosarcoma (common) Number	2	1	5	10*	4		
%	3.3	1.7	8.3	16.7	6.7	1.4 (0.0-2.7)	1.6 (0.0-3.6)
Uterus, leiomyoma (common) Number	9	5	18	19	22*		
%	15.0	8.3	30.0	31.7	36.7	4.2 (2.0-6.0)	2.9 (0.0-8.0)

RITA: Registry of Industrial Toxicology Animal data
* Peto analysis: pairwise difference $p < 0.01$
[a] expert report prepared by BI
[b] Incidences for the ovary
[c] N = 1920 for uterus leiomyosarcoma; 560 for uterus leiomyoma

Sponsor labeled Incidental Neoplastic Findings

"The incidence of benign mixed sex cord stromal tumours was slightly high in the 76.9 and 255 µg/kg/day dose groups, lacking statistical significance. A difference in incidence between these BI 1744 CL treated groups and the Vehicle Control group was no longer apparent, when all sex cord stromal tumours were combined. The incidences of all types of sex cord stromal tumours were outside the range of historical controls in one or more groups including control groups: benign mixed sex cord stromal tumours (in the 76.9 and 255 µg/kg/day groups), benign granulosa cell tumours (in the Vehicle Control group and the 255 µg/kg/day group), benign luteomas (in the Clean Air Control group), and the combination of these 3 sex cord stromal tumours (in the Vehicle Control group as well as the 76.9 and 255 µg/kg/day dose groups). This indicates an unexpectedly high spontaneous incidence of sex cord stromal tumours in the present study, which needs to be taken into account when assessing the relevance of incidences exceeding historical control data in BI 1744 CL treated groups.

"It was concluded that the slightly high incidence of benign mixed sex cord stromal tumours in 76.9 and 255 µg/kg/day females was incidental due to the minor and statistically nonsignificant difference to the Vehicle Control group, the high spontaneous background incidence of sex cord

Reference ID: 3230218

stromal tumours in this study, and the lack of a difference to the Vehicle Control group when all sex cord stromal tumours were combined." (page 58 of rat report)

Table 23. Un-numbered Sponsor Table: "Incidence of ovarian sex cord/stromal tumours in female mice"

	Vehicle	Clean Air	26.1 µg/kg	76.9 µg/kg	255 µg/kg	Historical Control Data Mean (range) [%]	
						(b) (4)	RITA[a]
N=	59	60	60	60	60	707	1921
Ovary, sex cord/stromal tumor, mixed (rare) Number	1	0	1	3	3		
%	1.7	0.0	1.7	5.0	5.0	0.1 (0.0-0.8)	0.0
Ovary, granulose cell tumor (common) Number	3	0	1	1	3		
%	5.1	0.0	1.7	1.7	5.0	1.4 (0.0-3.3)	0.9 (0.0-2.0)
Ovary luteoma Number	2	3	0	0	1		
%	3.4	5.0	0.0	0.0	1.7	1.6 (0.0-2.7)	0.6 (0.0-4.0)
Ovary: total all sex cord/ /stromal tumors Number	6	3	2	4	7		
%	10.2	5.0	3.3	6.7	11.7	3.1 (2.0-5.8)	Not Applicable

RITA: Registry of Industrial Toxicology Animal data
[a] expert report prepared by BI

"A low incidence of carcinomas or adenocarcinomas of the uterus occurred in BI 1744 CL treated groups, lacking statistical significance or a dose response relationship and being within the range of RITA historical control data for this strain and breeder but slightly exceeding the range of historical control data of the laboratory that conducted the study. It was concluded that these tumours were unrelated to treatment with BI 1744 CL.

"The incidence of uterine polyps was slightly high in 255 µg/kg/day females, without reaching statistical significance. It was within the range of internal laboratory and RITA historical control data. Neither the incidence of uterine stromal tumours, adenomas and adenocarcinomas, nor any type of hyperplasia was increased. It was concluded that the observed inter-group difference was unrelated to treatment with BI 1744 CL." (page 59)

"The incidence of uterine cystic endometrial hyperplasia was slightly high in 255 µg/kg/day females when compared to Vehicle Control ($p<0.05$). However, it was similar to the incidence in the Clean Air Control group, indicating the absence of a BI 1744 CL mediated effect." (page 60)

Reference ID: 3230218

Table 24. Un-numbered Sponsor Table: "Incidence of selected uterine tumours and proliferative findings in female mice"

	Vehicle	Clean Air	26.1 µg/kg	76.9 µg/kg	255 µg/kg	Historical Control Data Mean (range) [%]	
						(b) (4)	RITA[a]
N=	60	60	60	60	60	709	(≤ 1921)[b]
Uterus, adenocarcinoma (rare) Number	0	0	2	1	2		
%	0.0	0.0	3.3	1.7	3.3	0.7 (0.0-2.0)	0.8 (0.0-4.6)
Uterus, polyp (common) Number	4	7	8	8	12		
%	6.7	11.7	13.3	13.3	20.0	7.8 (0.9- 22.0)	7.2 (0.0-21.6)
Uterus, cystic endometrial hyperplasia Number	27	37	29	27	39*		
%	45	62	48	45	65	46 (0.0-2.7)	Not Applicable

RITA: Registry of Industrial Toxicology Animal data
Fisher exact test * $p<0.05$
[a] expert report prepared by BI
[b] N=1870 for adenocarcinoma, 1920 for polyps

"The incidence of bronchiolo-alveolar adenomas and adenocarcinomas of the lung in male mice was similar in all groups. In females, the incidence of bronchiolo-alveolar adenomas was slightly high at 76.9 µg/kg/day, lacking statistical significance. At 26.1 and 255 µg/kg/day, incidences were similar to the Vehicle Control group. The incidence of bronchiolo-alveolar adenocarcinomas was similar in all groups of female mice, including both control groups. Also the incidence of bronchiolo-alveolar hyperplasia was not increased in groups of BI 1744 CL treated mice. It was concluded that the slightly high incidence of bronchiolo-alveolar adenomas in 76.9 µg/kg/day females was not BI 1744 CL related." (page 60 of mouse report)

Reference ID: 3230218

Table 25. Un-numbered Sponsor Table: "Incidence of bronchiolo-alveolar adenomas in female mice"

	Vehicle	Clean Air	26.1 µg/kg	76.9 µg/kg	255 µg/kg	Historical Control Data Mean (range) [%]	
						(b) (4)	RITA[a]
N=	59	60	60	59	60	709	(≤ 1921)[b]
Lung, bronchio-alveolar, adenoma (common) Number	6	3	9	12	8		
%	10.2	5.0	15.0	20.3	13.3	9.1 (2.0-14.0)	Not Applicable#

RITA: Registry of Industrial Toxicology Animal data
[a] expert report prepared by BI
due to marked difference between historical control data from RITA and the (b) (4) compilation

"Histiocytic sarcomas occurred at a low incidence in male mice, one in the Vehicle Control group and one in the 26.1 µg/kg/day dose group. In female mice, a slightly high incidence of histiocytic sarcomas was noted at 76.9 µg/kg/day, without statistical significance. No histiocytic sarcomas were noted in female mice dosed at 255 µg/kg/day. The incidence of this tumour type in all dose groups was well within the range of internal laboratory and RITA historical control data. It was concluded that the slightly high incidence of histiocytic sarcomas in 76.9 µg/kg/day females was not BI 1744 CL related." (page 61)

Table 26. Un-numbered Sponsor Table: "Incidence of histiocytic sarcoma of the haemopoietic system in female mice "

	Vehicle	Clean Air	26.1 µg/kg	76.9 µg/kg	255 µg/kg	Historical Control Data Mean (range) [%]	
						(b) (4)	RITA[a]
N=	60	60	60	60	60	709	560
Haemopoietic system, histiocytic sarcoma (common) Number	2	3	2	5	0		
%	3.3	5.0	3.3	8.3	0.0	6.2 (0.0-10.1)	7.9 (3.3-16.7)

RITA: Registry of Industrial Toxicology Animal data
[a] expert report prepared by BI

Reference ID: 3230218

3.2.1.2. FDA Reviewer's Results

This section will present the Agency findings on survival and tumorigenicity in male and female rats.

Survival analysis:

The following tables (Table 27 for male mice, Table 28 for females) summarize the mortality results for the study groups. The data were grouped for the specified time period, and present the number of deaths during the time interval over the number at risk at the beginning of the interval. The percentage cited is the percent that survived at the end of the interval. In these tables the terminal period only includes those animals were sacrificed. Animals that died of other causes during the terminal period are included in the preceding, but overlapping time period. The Kaplan-Meier survival plots in Appendix 1 provide a more detailed picture of the profile of mortality losses.

Table 27. Summary of Male Mice Survival (estimated dose in µg/kg/day)

Period (Weeks)	Air Control 0	Vehicle Control 0	Low 26.1	Medium 76.9	High 255
1-52	1/60[1] 98.3%[2]	6/60 90.0%	0/60 100%	3/60 95.0%	5/60 91.7%
53-78	3/59 93.3%	9/54 75.0%	7/60 88.3%	3/57 90.0%	4/55 85.0%
79-91	11/56 75.0%	5/48 66.7%	7/53 76.7%	10/54 73,3%	4/51 78.3%
92-104	4/45 68.3%	7/40 55.0%	8/34 63.3%	10/44 56.7%	8/47 65.0%
Terminal [3] 105	41	33	38	34	39

[1] number of deaths / number at risk
[2] overall per cent survival to end of period.
[3] number of animals that survived to terminal sacrifice

Note that unlike the situation in rats, in three of the six dose groups above, in male Groups 3-5, the data set provided by the Sponsor indicates has one more animal coded as being in the terminal sacrifice group than is reflected in their survival table (reproduced in Table 21 above). This may be due to a definition of premature death that could include an animal that was categorized in the data set as being in the terminal sacrifice group. A similar comment applies to Group 2 in female mice in Table 28 below.

Reference ID: 3230218

Table 28. Summary of Female Mice Survival (estimated dose in µg/kg/day)

Period (Weeks)	Air Control 0	Vehicle Control 0	Low 26.1	Medium 76.9	High 255
1-52	3/60[1] 95.0%[2]	6/60 90.0%	3/60 95.0%	3/60 95.0%	4/60 93.3%
53-78	13/57 73.3%	6/54 80.0%	9/57 80.0%	13/57 73.3%	10/56 76.7%
79-91	7/44 61.7%	8/48 66.7%	14/48 56.7%	11/44 55.0%	11/46 58.3%
92-104	9/37 46.7%	9/40 51.7%	6/34 46.7%	18/33 25.0%	11/35 40%
Terminal[3] 105	28	31	28	15	24

[1] number of deaths / number at risk
[2] overall per cent survival to end of period.
[3] number of animals that survived to terminal sacrifice

Table 29. Statistical Significances of Tests of Homogeneity and Trend in Survival in the Mice Study

Hypothesis Tested	Males		Females	
	Log rank	Wilcoxon	Log rank	Wilcoxon
Mice Homogeneity over Groups 1-5	0.4608	0.3984	0.0693	0.2209
Homogeneity over Groups 2-5	0.5370	0.4547	0.0431	0.1531
No trend over Groups 2-5	0.3794	0.3798	0.3396	0.3582
No Difference Between Groups 1 vs 5	0.6823	0.6525	0.4731	0.4888
No Difference Between Groups 2 vs 5	0.2292	0.2141	0.2211	0.2707
No Difference Between Groups 1 vs 2	0.0992	0.0748	0.5908	0.6308

From the Kaplan-Meier curve depicted in Figure A.1.3 of Appendix 1, it appears that in male mice, from day 300, the vehicle group had the lowest survival, with all other groups fairly closely intertwined. This is consistent with the results of the tests of homogeneity and trend above (all six $p \geq 0.3794$). Similarly there is no statistically significant evidence of differences between the high dose and the controls (all four $p \geq 0.2141$). However, the test of differences between Groups 1 and 2 are some close to statistical significance (Logrank $p = 0.0992$, Wilcoxon $p = 0.0748$).

In female mice, from Figure A.1.4 below, after day 500 or so, it seems that the medium dose group has the lowest survival (highest mortality) with the vehicle group having the highest survival. The survival of the other dose groups are intertwined between these limits. Thus there is weak, somewhat equivocal evidence of treatment group differences at the end of the study (Group 1-5 Logrank $p = 0.0693$, Group 2-5 Logrank $p = 0.0431$). The corresponding Wilcoxon tests do not come close to the usual significance levels (Group 1-5 Wilcoxon $p = 0.2209$, Group 2-5 Wilcoxon $p = 0.1531$). There is statistically significant test of trend (Logrank $p = 0.3396$,

Reference ID: 3230218

Wilcoxon p = 0.3582). None of the three pairwise tests was close to statistical significance (all six p ≥ 0.2211).

Tumorigenicity analysis:

As discussed in Section 1.3.1.5, for common tumors, the Haseman-Lin-Rahman rules adjusting for multiplicity in a two species study specify that for a very rough 0.10 (10%) overall false positive error rate, overall trend should be tested at a 0.025 (2.5%) level in rare tumors and at 0.005 (0.5%) level in common tumors, while the pairwise test between the High dose and appropriate control should be tested at a 0.05 (5%) level in rare tumors and at 0.01 (1%) level in common tumors. Those organ-tumor combinations with at least a nominally statistically significant test result (p ≤ 0.05) in mice are summarized below:

Table 30. Potentially Statistically Significant Neoplasms in Mice

Organ/ Tumor	Air	Veh	Low	Mid	Hi	trend	Hi vs Veh	Med vs Veh	Low vs veh	Air vs Veh
Male Mice										
SKIN AND SUBCUT										
# Evaluated	60	60	60	60	60					
Adjusted # at risk	51.6	47.6	51.6	51.0	49.9					
Fibrosarc./Sarc./Rhab-domyosarc	1	6	3	3	2	.8979	.9744	.9341	.9380	.0441
Female Mice										
Systemic										
# Evaluated	60	60	60	60	60					
Adjusted # at risk	44.3	45.1	43.6	39.3	42.6					
HAEMANGIOSARCOMA [M]	1	1	0	0	3	.0479	.2824	1	1	.7584
UTERUS										
# Evaluated	60	60	60	60	60					
Adjusted # at risk	44.3	46.3	47.2	44.7	46.7					
LEIOMYOMA [B]	5	9	18	19	22	.0126	.0038	.0138	.0385	.2177
Adjusted # at risk	43.9	45.0	45.1	41.4	42.6					
LEIOMYOSARCOMA [M]	1	2	5	10	4	.4039	.3160	.0092	.2263	.5087
Adjusted # at risk	44.3	46.3	48.7	46.7	47.5					
Leiomyoma/Leiomyosarcoma	6	11	23	28	26	.0141	.0018	.0003	.0133	.1648
Adjusted # at risk	44.5	45.6	44.4	42.0	44.4					
POLYP [B]	7	4	8	8	12	.0248	.0227	.1339	.1655	.9086

In male mice the comparison between the Vehicle dose group and the Air dose group in Fibrosarcarcoma/Sarcoma/Rhabdomyoma achieved the 0.05 level, but since it would be classified as a common tumor (incidence in the Air group > 1%), when adjusting for the multiplicity of tests it would not be statistical significamt (p=0.0441 > 0.01). In female mice the pairwise comparisons between the Vehicle dose group and the high dose group in benign leiomyoma and pooled leiomyoma/leiomyosarcoma would be considered statistically significant (p = 0.0038, 0.0018 ≤ 0.01, respectively). The difference between vehicle and the Medium dose group in malignant leiomyosarcoma, and pooled leiomyoma/leiomyosarcoma would also be classified as statistically significant (p = 0.0092, 0.0003 ≤ 0.01). But that including these latter tests can raise the overall Type I error to some level above the nominal 10%. No other comparisons in female mice achieved the multiplicity adjusted level of significance. .

Reference ID: 3230218

Further details on these tests and complete incidence tables in both genders are provided in Appendix 2.

4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

NA

5. SUMMARY AND CONCLUSIONS

5.1. Statistical Issues and Collective Evidence

Please see Section 1.3 above.

5.2. Conclusions and Recommendations

Please see Section 1.1 above.

Reference ID: 3230218

Appendices

Appendix 1. FDA Survival Analysis

Simple summary life tables in mortality are presented in the report (Tables 16, 17, 23, and 24, above). Kaplan-Meier estimated survival curves across study groups for each gender are displayed below in Figures A.1.1 and A.1.2 for rats and Figures A.1.3 and A.1.4 for mice. These plots include 95% confidence intervals around each survival curve (colored area around each curve). These plots are also supported by tests of homogeneity over the five treatment groups including the air control group, tests of homogeneity and trend over the four groups excluding the air control, and pairwise comparisons among the controls and the high dose. Group. The statistical significance levels (i.e., p-values) are provided in Tables A.1.1. and A.1.2., below. One might note that the log rank tests places greater weight on later events, while the Wilcoxon test tends to weight them more equally, and thus places more weight on earlier events than does the log rank test.

Table A.1.1 Statistical Significances of Tests of Homogeneity and Trend in Survival in the Rat Study

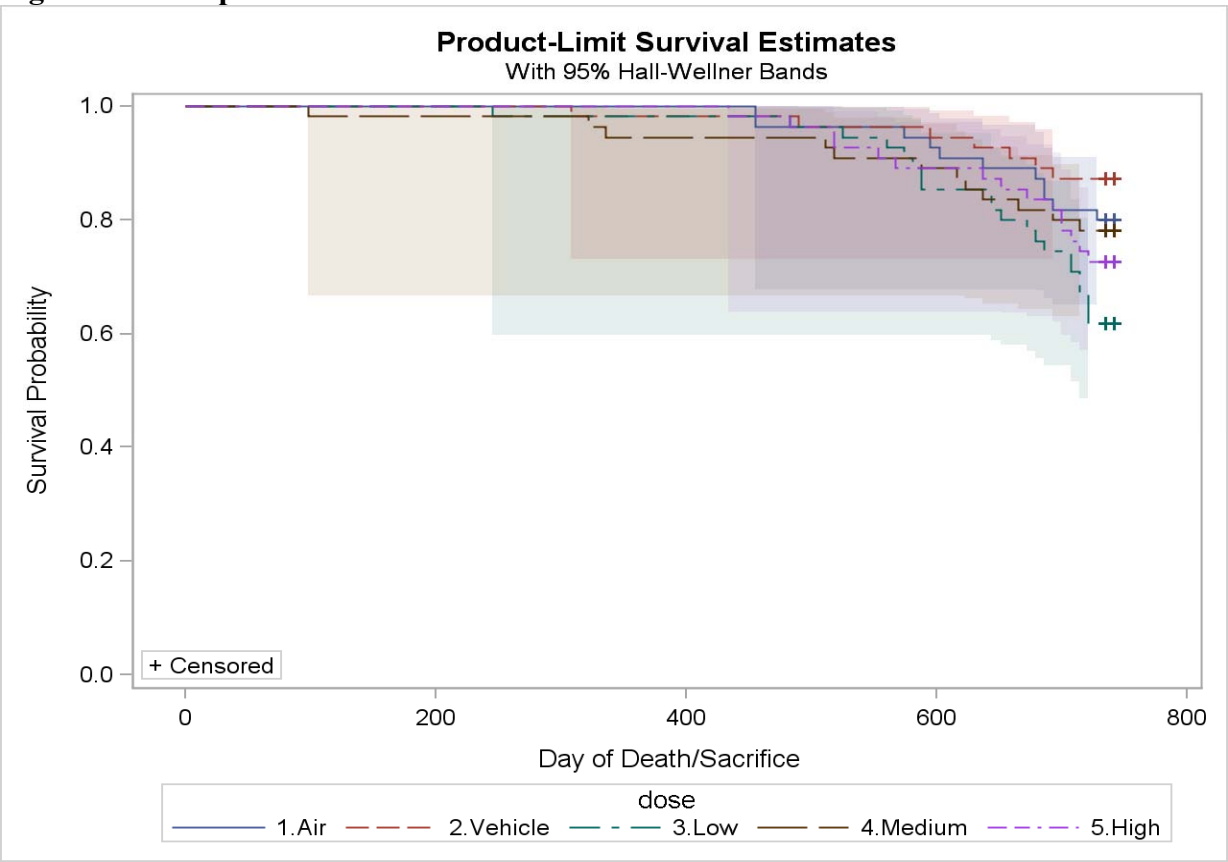
Hypothesis Tested	Males		Females	
	Log rank	Wilcoxon	Log rank	Wilcoxon
Rats Homogeneity over Groups 1-5	0.0396	0.0565	0.0659	0.0681
Homogeneity over Groups 2-5	0.0297	0.0439	0.0404	0.0435
No trend over Groups 2-5	0.5781	0.5876	0.0499	0.0750
No Difference Between Groups 1 vs 5	0.3832	0.3972	0.9915	0.8552
No Difference Between Groups 2 vs 5	0.0659	0.0751	0.0136	0.0188
No Difference Between Groups 1 vs 2	0.3201	0.3346	0.0178	0.0187

From Figure A.1.1, in male rats, roughly after day 500 there appears to be a tendency for a separation of the treatment groups in that the Vehicle control has the highest survival, followed by the Air control, intertwined with the High and Medium dose groups, and the highest mortality (lowest survival) in the Low dose group. This is consistent with the results above, i.e. there is some evidence of herterogeneity across dose groups 1-5 (Logrank p = 0.0396, Wilcoxon p = 0.0565) and groups 2-5 (Logrank p = 0.0297, Wilcoxon p = 0.0439), but no evidence of a dose related trend over groups 2-5 (Logrank p = 0.5781, Wilcoxon p = 0.5876). There is no evidence of a pairwise difference between Groups 1 and 5 (Logrank p = 0.3972, Wilcoxon p = 0.3832), nor between the controls, i.e. Groups 1 and 2 (Logrank p = 0.3201, Wilcoxon p = 0.3346). Somewhat inconsistently at the usual 0.05 level, the slight differences between the vehicle and high dose are close to significance us, i.e. Groups 2 and 5, (Logrank p = 0.0659, Wilcoxon p = 0.0751).

Recall if the time of natural death is greater than the last noted time the survival time is described as "censored." Since we would expect an animal to live past the time it was sacrificed, the time of sacrifice is treated as a time of censoring.

Reference ID: 3230218

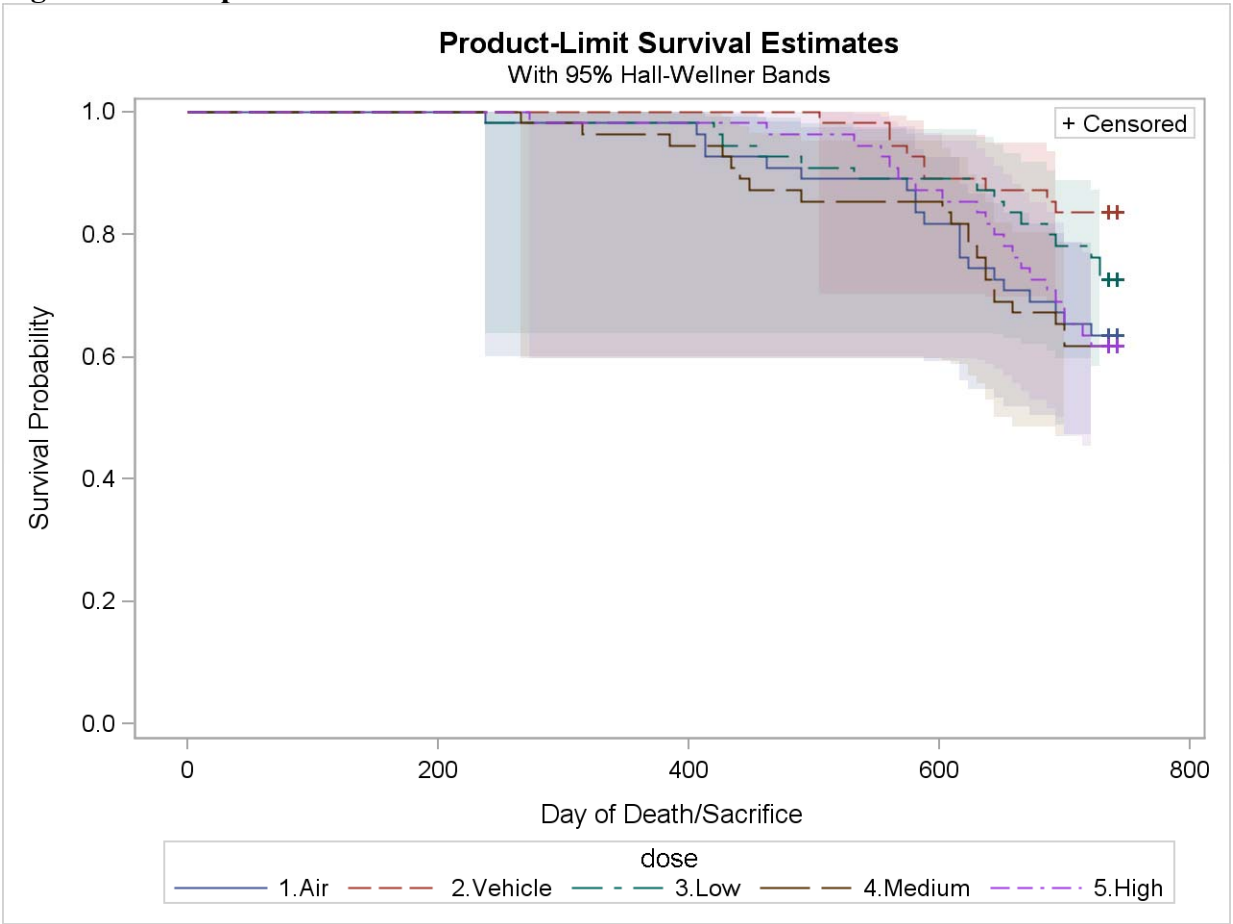
Figure A.1.1 Kaplan-Meier Survival Curves for Male Rats



Similarly from Figure A.1.2, in female rats, after day 400+ there also appears to be a tendency for a separation of the treatment groups in that the vehicle control has the highest survival, followed by the high and low dose groups, although both of the latter seem to intertwined. Finally, the highest mortality (i.e. lowest survival) are in the Medium and High dose groups with the Air controls (i.e. Groups 4 and 1). As with male rats, this is consistent with the results of the statistical tests above, i.e., there is weak some evidence of herterogeneity across dose groups 1-5 (Logrank p = 0.0659, Wilcoxon p = 0.0681) and groups 2-5 (Logrank p = 0.0404, Wilcoxon p = 0.0435). However in this gender there is some evidence of a dose related trend over groups 2-5 (Logrank p = 0.0499, Wilcoxon p = 0.0750). In this case, there is evidence of a pairwise difference between Group 2 and Group 1 (Logrank p = 0.0172, Wilcoxon p = 0.0261) and Group 5 (Logrank p = 0.0188, Wilcoxon p = 0.0136).

Reference ID: 3230218

Figure A.1.2 Kaplan-Meier Survival Curves for Female Rats



Figures A.1.3 through A.1.4, below, provide similar survival curves for each mouse gender, while Table A.1.2 provides a similat tabulation of p-values for the mouse study.

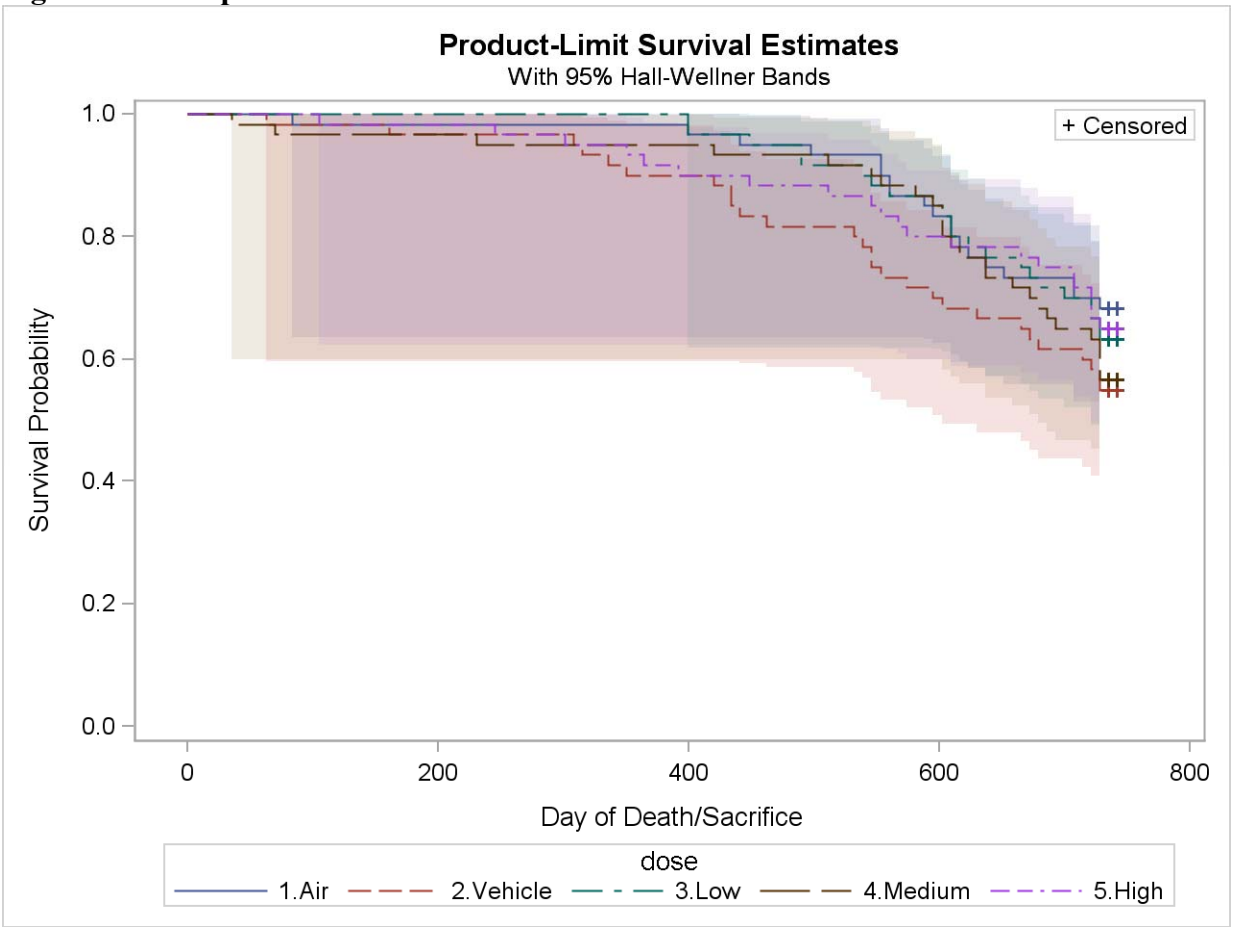
Table A.1.2. Statistical Significances of Tests of Homogeneity and Trend in Survival in the Mouse Study

Hypothesis Tested	Males		Females	
	Log rank	Wilcoxon	Log rank	Wilcoxon
Mice Homogeneity over Groups 1-5	0.4608	0.3984	0.0693	0.2209
Homogeneity over Groups 2-5	0.5370	0.4547	0.0431	0.1531
No trend over Groups 2-5	0.3794	0.3798	0.3396	0.3582
No Difference Between Groups 1 vs 5	0.6823	0.6525	0.4731	0.4888
No Difference Between Groups 2 vs 5	0.2292	0.2141	0.2211	0.2707
No Difference Between Groups 1 vs 2	0.0992	0.0748	0.5908	0.6308

Reference ID: 3230218

From Figure A.1.3, in male mice, from day 300, the Vehicle group had the lowest survival, with all other groups fairly closely intertwined. This is consistent with the results of the tests of homogeneity and trend anove (all six $p \geq 0.3794$). No evidence of differences is not evidence of no differences, but it is consistent with such a conclusion. Similarly there is no statistically significant evidence of differences between the high dose and the controls (all four $p \geq 0.2141$). However, the test of differences between Groups 1 and 2 are some close to statistical significance (Logrank $p = 0.0992$, Wilcoxon $p = 0.0748$).

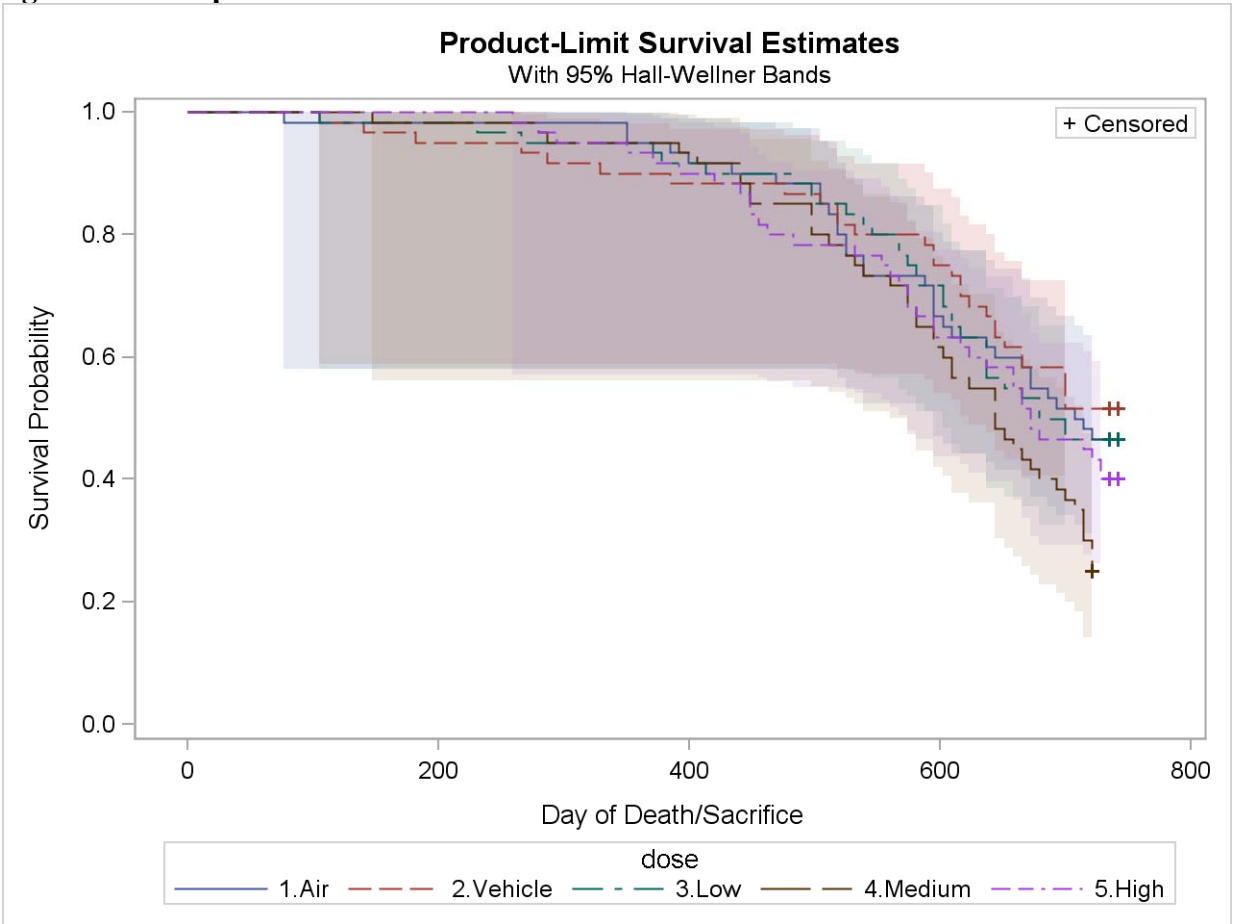
Figure A.1.3 Kaplan-Meier Survival Curves for Male Mice



Reference ID: 3230218

In female mice, from Figure A.1.4, below, after day 500+. it seems that the Medium dose group has the lowest survival (highest mortality) with the Vehicle group having the highest survival. The survival of the other dose groups are intertwined between these limits. Thus there is weak, somewhat equivocal evidence of treatment group differences at the end of the study (Group 1-5 Logrank p = 0.0693, Group 2-5 Logrank p = 0.0431). The corresponding Wilcoxon tests do not come close to the usual significance levels (Group 1-5 Wilcoxon p = 0.2209, Group 2-5 Wilcoxon p = 0.1531). There is statistically significant test of trend (Logrank p = 0.3396, Wilcoxon p = 0.3582). None of the three pairwise tests was close to statistical significance (all six p ≥ 0.2211).

Figure A.1.4 Kaplan-Meier Survival Curves for Female Mice



Reference ID: 3230218

Appendix 2. FDA Poly-k Tumorigenicity Analysis

The poly-k test, here with k=3, modifies the original Cochran-Armitage test to adjust for differences in mortality (please see Bailer & Portier, 1988, Bieler & Williams, 1993). The tests used here are small sample exact permutation tests of tumor incidence. These do assume all marginal totals are fixed, a debatable assumption. This assumption implies that in the pairwise tests when one dose group has no tumors of the specific type and the other does, there is only one permutation of this pattern. Since that means that the only permutation of the data is the one observed, that means that all possible permutations are as extreme as the pattern observed, and thus the significance level of the observed pattern can be logically expressed as 1.0. One could use the same sort of argument when there were no tumors of the specific type being analyzed in either column of the 2x2 table corresponding to a pairwise comparison. Then an argument could be made that the p-value for this test should also be 1.0. However, largely for readability, in the tables below these p-values are considered as missing (i.e., corresponding to a null test), denoted by a period ".". Note that StatXact adjusts for the variance, which would be 0. Then the significance levels of the test statistics are based on the result of a division by 0, i.e., undefined, and hence StatXact codes these p-values as missing.

For each species by gender by organ combination the number of animals analyzed and used in the statistical tests is presented first. The entry for each tumor is preceded by the adjusted number of animals at risk for that endpoint. It seems clear that an animal that dies early without having displaying that endpoint reduces the size of the risk set for that getting that particular endpoint. The poly-k test down weights such animals, and as discussed in Section 1.3.1.3, above, the sum of these poly-k weights seems to be a better estimate of the number of animals at risk of getting that tumor. This sum is given in the row labeled "Adjusted # at risk". Tumor incidence is presented next, with the significance levels of the tests of trend, and the results of pairwise tests between the high, , medium, low, and air alone dose groups versus vehicle.

To adjust for the multiplicity of tests the so-called Haseman-Lin-Rahman rules discussed in Section 1.3.1.4 are often applied. That is, when testing for trend over dose and the difference between the highest dose group with a control group, to control the overall Type I error rate to roughly 10% for a standard two species, two sex study, one compares the unadjusted significance level of the trend test to 0.005 for common tumors and 0.025 for rare tumors, and the pairwise test to 0.01 for common tumors and 0.05 for rare tumors. Incidence in the vehicle group is used to assess background tumor incidence, and thus whether a tumor is considered to be rare (background incidence <1%) or common. As also discussed in section 1.3.1.4, using these tests for the pairwise comparisons between the vehicle and the Air control, the low dose, and the medium dose can be expected to increase the overall type I error rate to some value above the nominal rough 10% level.

Tables A.2.1 in both rats and mice, shows the tumors that had at least one mortality adjusted test whose nominal statistical significance was at least 0.05 (or less). Note that when

Reference ID: 3230218

one adjusts for multiplicity using the rules above these nominally significant comparisons may not be statistically significant. Tables A.2.2 and A.2.3 display all incidences and statistical test results for male and female rats, respectively, while Tables A.2.4 and A.2.5 present similar results in male and female mice. The p-values of the poly-k test are based on exact tests from StatXact as discussed above. As also noted above, the period '.' denotes the p-values of tests of dose groups with no tumors in any group.

Table A.2.1 Potentially Statistically Significant Neoplasms Rats and Mice

Organ/ Tumor	Air	Veh	Low	Mid	Hi	trend	Air vs Veh	Hi vs Veh	Med vs veh	Low vs Veh
Male Rats										
PITUITARY GLAND										
# Evaluated	53	54	55	55	55					
Adjusted # at risk	50.6	52.0	52.9	53.3	52.1					
ADENOMA, ANTERIOR LOBE[B]	16	15	25	16	18	.5779	.3369	.5253	.0345	.7127
Female Rats										
OVARY										
Adjusted # at risk	45.5	51.2	48.9	44.6	47.9					
LEIOMYOMA [B]	0	0	1	0	4	.0092	.0494	.	.4848	.
Male Mice										
SKIN AND SUBCUT										
# Evaluated	60	60	60	60	60					
Adjusted # at risk	51.6	47.6	51.6	51.0	49.9					
Fibrosarc./Sarc./Rhab-domyosarc	1	6	3	3	2	.8979	.9744	.9341	.9380	.0441
Female Mice										
Systemic										
# Evaluated	60	60	60	60	60					
Adjusted # at risk	44.3	45.1	43.6	39.3	42.6					
HAEMANGIOSARCOMA [M]	1	1	0	0	3	.0479	.2824	1	1	.7584
UTERUS										
# Evaluated	60	60	60	60	60					
Adjusted # at risk	44.3	46.3	47.2	44.7	46.7					
LEIOMYOMA [B]	5	9	18	19	22	.0126	.0038	.0138	.0385	.2177
Adjusted # at risk	43.9	45.0	45.1	41.4	42.6					
LEIOMYOSARCOMA [M]	1	2	5	10	4	.4039	.3160	.0092	.2263	.5087
Adjusted # at risk	44.3	46.3	48.7	46.7	47.5					
Leiomyoma/Leiomyosarcoma	6	11	23	28	26	.0141	.0018	.0003	.0133	.1648
Adjusted # at risk	44.5	45.6	44.4	42.0	44.4					
POLYP [B]	7	4	8	8	12	.0248	.0227	.1339	.1655	.9086

Using the incidence in the Air dose group to define whether a tumor is rare or not, we note that leiomyoma of the ovary would be classified as a rare tumor in female rats, and thus the test of trend would be considered to be highly statistically significant ($p=0.0092 \leq 0.025$) while the comparison between the high dose and the Vehicle dose group would be considered barely statistically significant ($p = 0.0494 \leq 0.05$). Note that since it uses all dose groups the test of trend is a more powerful test than the pairwise test. Even accepting the increase in Type I error, the comparison between the Low dose and the Vehicle in anterior lobe adenoma of the pituitary of male rats would not be classified as statistically significant ($p = 0.0345 > 0.01$). Even

Reference ID: 3230218

ignoring the multiplicity adjustment, no other test results in either rat gender even achieved the nominal 0.05 level.

In male mice the comparison between the Vehicle dose group and the Air dose group in pooled Fibrosarcarcoma/Sarcoma/Rhabdomyoma achieved the 0.05 level, but since it would be classified as a common tumor (incidence in the Air group > 1%), when adjusting for the multiplicity of tests it would not be statistical significamt (p=0.0441 > 0.01). In female mice the pairwise comparisons between the Vehicle dose group and the high dose group in benign leiomyoma and pooled leiomyoma/leiomyosarcoma would be considered statistically significant (p = 0.0038, 0.0018 ≤ 0.01, respectively). The difference between vehicle and the Medium dose group in malignant leiomyosarcoma, and pooled leiomyoma/leiomyosarcoma would also be classified as statistically significant (p = 0.0092, 0.0003 ≤ 0.01). But that including these latter tests can raise the overall Type I error to some level above the nominal 10%. No other comparisons in female mice achieved the multiplicity adjusted level of significance. .

Complete incidence tables are presented below.

Table A.2.2 Overall Tumorigenicity Results in Male Rats

Organ/Tumor	Incidence					Significance Levels				
	Air	Veh	Low	Med	High	Trend	Hi vs Veh	Med vs Veh	Low vs Veh	Air vs Veh
ABDOMINAL CAV										
# Evaluated	0	3	0	1	0					
Adjusted # at risk	0.0	2.4	0.0	1.0	0.0					
FIBROSARCOMA [M]	0	1	0	0	0	1	.	1	.	.
ADRENAL GLAND										
# Evaluated	55	55	55	55	54					
Adjusted # at risk	51.1	52.4	48.4	48.9	48.9					
CORTICAL ADENOMA [B]	0	1	0	2	1	.3567	.7321	.4697	1	.5049
Adjusted # at risk	51.1	52.0	48.4	48.9	48.9					
CORTICAL CARCINOMA [M]	0	1	0	0	1	.4327	.7372	1	1	.5000
Adjusted # at risk	51.1	52.4	48.4	48.9	48.9					
Cortical Carc./Adenoma	0	2	0	2	2	.2873	.6604	.6604	1	.2524
Adjusted # at risk	51.1	52.0	48.4	48.9	48.9					
PHAEOCHROMOCYTOMA [B]	0	1	1	0	2	.2295	.4770	1	.7372	.5000
Adjusted # at risk	51.1	52.0	48.9	48.9	48.9					
PHAEOCHROMOCYTOMA [M]	0	0	1	0	0	.7385	.	.	.4848	.
Adjusted # at risk	51.1	52.0	48.9	48.9	48.9					
Phaeochromocytoma B&M	0	1	2	0	2	.3175	.4770	1	.4770	.5000
BRAIN										
# Evaluated	55	55	55	55	55					
Adjusted # at risk	51.1	52.0	48.4	48.9	49.8					
ASTROCYTOMA [B]	0	0	0	0	1	.2500	.4900	.	.	.
Adjusted # at risk	51.3	52.3	48.4	48.9	49.8					
GRANULAR CELL TUMOUR [B]	2	2	1	0	1	.6701	.8674	1	.8633	.6985
Adjusted # at risk	51.6	52.0	48.4	48.9	49.8					
MALIGNANT ASTROCYTOMA [M]	1	0	1	0	0	.7398	.	.	.4848	1
CAECUM										
# Evaluated	53	55	54	55	53					
Adjusted # at risk	50.1	52.0	48.2	48.9	48.5					
STROMAL SARCOMA [M]	0	0	1	0	0	.7385	.	.	.4848	.
CRANIUM										
# Evaluated	0	1	0	0	0					
Adjusted # at risk	0.0	1.0	0.0	0.0	0.0					
MALIGNANT SCHWANNOMA [M]	0	1	0	0	0	1	.	.	.	.

Reference ID: 3230218

Table A.2.2 (cont.) Overall Tumorigenicity Results in Male Rats

Organ/Tumor	Incidence					Significance Levels				
	Air	Veh	Low	Med	High	Trend	Hi vs Veh	Med vs Veh	Low vs Veh	Air vs Veh
EAR										
# Evaluated	0	0	1	3	0					
Adjusted # at risk	0.0	0.0	1.0	2.0	0.0					
? SEBACEOUS CELL ADENOMA [B]	0	0	1	0	0	1	.	.	.	.
EPIDIDYMIS										
# Evaluated	54	55	55	55	55					
Adjusted # at risk	50.9	52.1	48.4	48.9	49.8					
MALIGNANT SCHWANNOMA [M]	0	1	0	0	0	1	1	1	1	.5098
FEMUR										
# Evaluated	54	55	55	55	55					
Adjusted # at risk	50.1	52.1	48.4	48.9	49.8					
OSTEOMA [B]	0	1	0	0	0	1	1	1	1	.5098
FOOT/LEG										
# Evaluated	0	0	0	1	1					
Adjusted # at risk	0.0	0.0	0.0	0.9	1.0					
SYNOVIAL SARCOMA [M]	0	0	0	0	1	1	.	.	.	.
HARDERIAN GLAND										
# Evaluated	55	55	55	55	55					
Adjusted # at risk	51.1	52.0	48.4	48.9	49.9					
ADENOCARCINOMA [M]	0	0	0	0	1	.2500	.4900	.	.	.
HEART										
# Evaluated	55	55	54	55	55					
Adjusted # at risk	51.1	52.0	47.9	48.9	49.8					
SCHWANNOMA [B]	1	0	0	0	0	.	.	.	.	1
HEMOPOIETIC SYS										
# Evaluated	1	1	1	1	1					
Adjusted # at risk	1.0	1.0	0.0	0.7	0.4					
HISTIOCYTIC SARCOMA [M]	0	1	0	0	0	1	.	.	.	.5000
Adjusted # at risk	1.0	0.8	1.0	1.0	1.0					
LYMPHOMA, LYMPHOBLASTIC [M]	1	0	1	1	1	1	.	.	.	.
KIDNEY										
# Evaluated	55	55	55	55	55					
Adjusted # at risk	51.1	52.0	48.7	48.9	49.8					
LIPOMA [B]	0	1	1	0	0	.9333	1	1	.7372	.5000
Adjusted # at risk	51.3	52.4	48.4	48.9	49.8					
LIPOSARCOMA [M]	1	1	0	0	0	1	1	1	1	.7573
LIVER										
# Evaluated	55	55	55	55	55					
Adjusted # at risk	51.1	52.0	48.4	48.9	49.9					
HEPATOCELLULAR CARCINOMA [M]	0	0	0	1	1	.1846	.4900	.4848	.	.
LN MESENTERIC										
# Evaluated	55	55	54	55	55					
Adjusted # at risk	51.7	52.0	47.5	48.9	49.8					
HAEMANGIOMA [B]	2	4	1	2	0	.9729	1	.8833	.9654	.3390
Adjusted # at risk	51.1	52.0	47.5	48.9	49.8					
HAEMANGIOSARCOMA [M]	1	0	0	0	0	.	.	.	.	1
Adjusted # at risk	51.7	52.0	47.5	48.9	49.8					
Hemangioma/Hemangiosarcoma	3	4	1	2	0	.9729	1	.8833	.9654	.5000
MAMMARY GLAND										
# Evaluated	51	45	46	45	51					
Adjusted # at risk	47.5	43.4	41.7	40.8	45.9					
ADENOMA [B]	0	1	0	0	0	1	1	1	1	.4778
ORAL CAVITY										
# Evaluated	0	0	0	1	3					
Adjusted # at risk	0.0	0.0	0.0	1.0	3.0					
SQUAMOUS-CELL CARCINOMA [M]	0	0	0	0	1	.7500	.	.	.	.

Reference ID: 3230218

Table A.2.2 (cont.) Overall Tumorigenicity Results in Male Rats

Organ/Tumor	Incidence					Significance Levels				
	Air	Veh	Low	Med	High	Trend	Hi vs Veh	Med vs Veh	Low vs Veh	Air vs Veh
PANCREAS END										
# Evaluated	55	54	55	55	55					
Adjusted # at risk	51.1	51.2	48.4	48.9	49.8					
ISLET CELL ADENOMA [B]	1	0	0	2	0	.4975	.	.2325	.	1
Adjusted # at risk	51.1	51.2	48.4	48.9	49.8					
ISLET CELL CARCINOMA [M]	1	0	0	0	0	.	.	.	.	1
Adjusted # at risk	51.1	51.2	48.4	48.9	49.8					
Islet Cell Carc./Adenoma	2	0	0	2	0	.4975	.	.2325	.	1
PARATHYROID GLN										
# Evaluated	52	53	49	55	52					
Adjusted # at risk	48.1	50.2	43.1	48.9	47.3					
ADENOMA [B]	0	0	1	1	0	.6201	.	.4898	.4624	.
PITUITARY GLAND										
# Evaluated	53	54	55	55	55					
Adjusted # at risk	50.6	52.0	52.9	53.3	52.1					
ADENOMA, ANTERIOR LOBE [B]	16	15	25	16	18	.5779	.3369	.5253	.0345	.7127
Adjusted # at risk	49.1	51.6	48.4	48.9	50.1					
ADENOMA, INTERMEDIATE LOBE [B]	0	1	1	0	2	.2414	.4925	1	.7372	.5100
SALV GLND SUBMX										
# Evaluated	55	55	55	55	55					
Adjusted # at risk	51.1	52.0	48.4	48.9	49.8					
MALIGNANT SCHWANNOMA [M]	0	0	0	0	1	.2500	.4900	.	.	.
SKIN AND SUBCUT										
# Evaluated	55	55	55	55	55					
Adjusted # at risk	51.1	52.0	48.5	48.9	49.8					
BASAL CELL CARCINOMA [M]	0	0	1	0	0	.7398	.	.	.4848	.
Adjusted # at risk	51.4	52.2	48.5	48.9	49.8					
FIBROMA [B]	3	3	1	3	0	.9271	1	.6221	.9310	.6702
Adjusted # at risk	51.1	52.0	48.4	48.9	50.1					
FIBROSARCOMA [M]	0	0	0	0	1	.2538	.4950	.	.	.
Adjusted # at risk	51.1	52.3	48.4	49.5	49.8					
KERATOACANTHOMA [B]	3	5	3	5	2	.8340	.9342	.5909	.8380	.3687
Adjusted # at risk	51.1	52.3	48.4	49.5	49.8					
Keratocanth./Sq. Cell Pap.	4	7	3	6	2	.9185	.9806	.6829	.9398	.2741
Adjusted # at risk	51.1	52.0	48.4	48.9	49.8					
LIPOMA [B]	1	0	0	0	0	.	.	.	.	1
Adjusted # at risk	51.1	52.0	48.4	48.9	50.3					
NEURAL CREST TUMOUR [M]	0	0	0	0	1	.2538	.4950	.	.	.
Adjusted # at risk	51.1	52.0	48.4	48.9	49.8					
SQUAMOUS-CELL PAPILLOMA [B]	1	2	0	1	0	.8871	1	.8672	1	.5000
SPINAL CORD										
# Evaluated	55	55	55	55	55					
Adjusted # at risk	51.1	52.0	48.4	48.9	49.8					
MALIGNANT ASTROCYTOMA [M]	1	0	0	0	0	.	.	.	.	1
SPLEEN										
# Evaluated	55	55	55	55	55					
Adjusted # at risk	51.1	52.0	48.4	49.1	49.8					
HAEMANGIOMA [B]	0	2	0	1	0	.8888	1	.8712	1	.2475
Adjusted # at risk	51.1	52.0	48.4	49.1	49.8					
Hemangioma/Hemangiosarcoma	0	2	0	1	0	.8888	1	.8712	1	.2475
STOMACH										
# Evaluated	55	55	55	55	55					
Adjusted # at risk	51.9	52.0	48.4	48.9	49.8					
ADENOCARCINOMA [M]	1	0	0	0	0	.	.	.	.	1
Systemic										
# Evaluated	55	55	55	55	55					
Adjusted # at risk	51.7	52.0	48.4	49.1	49.8					
Hemangioma/Hemangiosarcoma	3	5	1	3	0	.9808	1	.8523	.9839	.3576

Reference ID: 3230218

Table A.2.2 (cont.) Overall Tumorigenicity Results in Male Rats

Organ/Tumor	Incidence					Significance Levels				
	Air	Veh	Low	Med	High	Trend	Hi vs Veh	Med vs Veh	Low vs Veh	Air vs Veh
TESTIS										
# Evaluated	54	55	55	55	55					
Adjusted # at risk	50.9	52.0	48.4	48.9	49.8					
INTERSTITIAL CELL ADENOMA [B]	3	1	0	2	1	.3666	.7424	.4770	1	.9436
Adjusted # at risk	50.9	52.0	49.1	48.9	49.8					
MESOTHELIOMA [B]	0	0	1	0	0	.7411	.	.	.4900	.
Adjusted # at risk	50.9	52.0	48.4	48.9	50.9					
MESOTHELIOMA [M]	1	1	0	3	3	.0887	.3012	.2860	1	.7574
Adjusted # at risk	50.9	52.0	49.1	48.9	50.9					
Mesothelioma B&M	1	1	1	3	3	.1456	.3012	.2860	.7424	.7574
THYMUS										
# Evaluated	55	55	53	53	51					
Adjusted # at risk	51.1	52.0	46.6	46.9	47.0					
THYMOMA [B]	0	0	2	0	0	.7983	.	.	.2223	.
THYROID GLAND										
# Evaluated	54	55	55	55	55					
Adjusted # at risk	50.3	52.2	49.3	48.9	49.8					
C-CELL ADENOMA [B]	3	6	7	9	3	.8944	.9052	.2332	.4537	.2638
Adjusted # at risk	50.3	52.0	48.4	48.9	50.4					
C-CELL CARCINOMA [M]	0	0	0	0	1	.2538	.4950	.	.	.
Adjusted # at risk	50.3	52.2	49.3	48.9	50.4					
C-cell Carcinoma/Adenoma	3	6	7	9	4	.8157	.8242	.2332	.4537	.2638
Adjusted # at risk	50.3	52.0	48.7	48.9	49.8					
FOLLICULAR CELL ADENOMA [B]	0	4	2	1	3	.4797	.7651	.9672	.8833	.0612
Adjusted # at risk	50.3	52.0	48.4	48.9	49.8					
FOLLICULAR CELL CARCINOMA [M]	0	0	0	1	0	.4949	.	.4848	.	.
Adjusted # at risk	50.3	52.0	48.7	48.9	49.8					
Foll. Carcinoma/Adenoma	0	4	2	2	3	.5046	.7651	.8833	.8833	.0612
VERTEBRAE										
# Evaluated	2	1	6	1	2					
Adjusted # at risk	2.0	1.0	4.3	0.9	1.4					
OSTEOMA [B]	0	0	1	0	0	.8333	.	.	.8000	.
ZYMBAL'S GLAND										
# Evaluated	43	40	41	36	44					
Adjusted # at risk	40.7	38.2	36.9	32.4	40.3					
CARCINOMA [M]	0	0	0	0	1	.2740	.5128	.	.	.

Table A.2.3. Overall Tumorigenicity Results in Female Rats

Organ/Tumor	Incidence					Significance Levels				
	Air	Veh	Low	Med	High	Trend	Hi vs Veh	Med vs Veh	Low vs Veh	Air vs Veh
ADRENAL GLAND										
# Evaluated	55	55	55	55	55					
Adjusted # at risk	45.5	51.8	48.9	44.6	47.5					
CORTICAL ADENOMA [B]	0	1	0	0	0	1	1	1	1	.5312
Adjusted # at risk	45.5	51.2	49.2	44.6	47.5					
PHAEOCHROMOCYTOMA [B]	1	1	1	1	1	.4897	.7317	.7144	.7424	.7829
Adjusted # at risk	45.5	51.2	48.9	44.6	47.5					
PHAEOCHROMOCYTOMA [M]	0	1	0	0	0	1	1	1	1	.5312
Adjusted # at risk	45.5	51.2	49.2	44.6	47.5					
Phaeochromocytoma B&M	1	2	1	1	1	.6530	.8631	.8495	.8712	.5473
BRAIN										
# Evaluated	55	55	55	55	55					
Adjusted # at risk	45.5	51.2	48.9	44.6	47.5					
ASTROCYTOMA [B]	0	0	2	0	0	.7926	.	.	.2325	.
Adjusted # at risk	45.5	51.2	48.9	44.6	47.5					
GLIOMA [M]	0	0	1	0	0	.7316	.	.	.4848	.

Reference ID: 3230218

Table A.2.3 (cont.) Overall Tumorigenicity Results in Female Rats

Organ/Tumor	Incidence					Significance Levels				
	Air	Veh	Low	Med	High	Trend	Hi vs Veh	Med vs Veh	Low vs Veh	Air vs Veh
DUODENUM										
# Evaluated	54	55	55	55	53					
Adjusted # at risk	44.8	51.2	48.9	44.6	46.1					
LEIOMYOMA [B]	1	0	0	0	0	.	.	.	.	1
HEMOPOIETIC SYS										
# Evaluated	1	1	0	1	1					
Adjusted # at risk	0.2	1.0	0.0	1.0	0.8					
LEUKAEMIA [M]	0	0	0	1	0	.5000	.	.5000	.	.
Adjusted # at risk	0.2	1.0	0.0	1.0	0.8					
LYMPHOMA [M]	0	1	0	0	0	1	.	1	.	.
Adjusted # at risk	1.0	1.0	0.0	1.0	1.0					
LYMPHOMA, LYMPHOBLASTIC [M]	1	0	0	0	1	.3333	.5000	.	.	1
Adjusted # at risk	1.0	1.0	0.0	1.0	1.0					
Lymphoma Any	1	1	0	0	1	.6667	.	1	.	.
LIVER										
# Evaluated	55	55	55	55	55					
Adjusted # at risk	45.5	51.2	48.9	44.6	47.5					
HEPATOCELLULAR ADENOMA [B]	0	1	0	0	0	1	1	1	1	.5312
LN MESENTERIC										
# Evaluated	55	54	55	54	55					
Adjusted # at risk	45.6	50.2	48.9	44.5	47.5					
HAEMANGIOMA [B]	1	1	0	0	1	.4365	.7369	1	1	.7783
Adjusted # at risk	45.6	50.2	48.9	44.5	47.5					
Hemangioma/Hemangiosarcoma	1	1	0	0	1	.4365	.7369	1	1	.7783
LUNG										
# Evaluated	55	55	55	55	55					
Adjusted # at risk	45.5	51.2	48.9	44.6	47.8					
BRONCHIOLO-ALVEOLAR ADENOMA	0	0	0	0	1	.2474	.4796	.	.	.
MAMMARY GLAND										
# Evaluated	54	55	55	55	55					
Adjusted # at risk	45.5	52.1	48.9	44.6	47.5					
ADENOCARCINOMA [M]	1	3	1	2	1	.7460	.9281	.7626	.9310	.3652
Adjusted # at risk	45.5	51.2	48.9	44.6	47.5					
ADENOMA [B]	0	2	0	1	1	.5578	.8631	.8495	1	.2796
Adjusted # at risk	45.5	52.1	48.9	44.6	47.5					
Adenocarc./Adenoma	1	5	1	3	2	.7196	.9267	.8047	.9829	.1390
Adjusted # at risk	46.6	51.2	50.6	46.0	48.8					
FIBROADENOMA [B]	8	6	6	5	7	.3096	.4527	.6748	.6059	.8591
Adjusted # at risk	45.5	51.2	48.9	45.4	47.5					
MIXED MALIGNANT TUMOUR [M]	0	0	0	1	0	.4817	.	.4687	.	.
ORAL CAVITY										
# Evaluated	1	0	0	1	1					
Adjusted # at risk	0.8	0.0	0.0	1.0	1.0					
SQUAMOUS-CELL CARCINOMA [M]	0	0	0	0	1	.5000	.	.	.	.

Reference ID: 3230218

Table A.2.3 (cont.) Overall Tumorigenicity Results in Female Rats

Organ/Tumor	Incidence					Significance Levels				
	Air	Veh	Low	Med	High	Trend	Hi vs Veh	Med vs Veh	Low vs Veh	Air vs Veh
OVARY										
# Evaluated	55	55	55	55	55					
Adjusted # at risk	45.5	51.2	48.9	44.9	47.5					
CYSTADENOMA [B]	0	0	0	2	0	.4872	.	.2119	.	.
Adjusted # at risk	45.5	51.2	48.9	44.6	47.7					
GRANULOSA CELL TUMOUR [M]	0	0	0	0	1	.2474	.4796	.	.	.
Adjusted # at risk	45.5	51.2	49.5	44.6	47.5					
GRANULOSA/THECAL CELL TUMOUR	0	0	1	0	0	.7330	.	.	.4900	.
Adjusted # at risk	45.5	51.2	48.9	44.6	47.9					
LEIOMYOMA [B]	0	0	1	0	4	.0092	.0494	.	.4848	.
Adjusted # at risk	45.5	51.2	48.9	44.6	47.5					
SEX CORD/STROMAL TUMOUR [B]	0	0	0	0	1	.2474	.4796	.	.	.
Adjusted # at risk	45.7	51.2	48.9	44.6	47.5					
SEX CORD/STROMAL TUMOUR [M]	1	0	0	0	0	.	.	.	.	1
Adjusted # at risk	45.5	51.2	48.9	44.6	47.5					
THECAL CELL TUMOUR [B]	0	1	1	0	0	.9290	1	1	.7372	.5312
Adjusted # at risk	45.5	51.2	48.9	44.6	47.5					
TUBULOSTROMAL ADENOMA [B]	0	0	0	0	1	.2474	.4796	.	.	.
Adjusted # at risk	45.5	51.2	49.5	44.6	47.5					
Thecal Cell Tumor(Any)	0	1	2	0	0	.9271	1	1	.4848	.5312
PITUITARY GLAND										
# Evaluated	55	55	55	55	55					
Adjusted # at risk	49.7	54.3	52.1	48.3	51.2					
ADENOMA, ANTERIOR LOBE [B]	26	38	27	29	33	.4033	.7963	.8972	.9844	.0541
Adjusted # at risk	45.5	51.2	48.9	44.6	47.5					
ADENOMA, INTERMEDIATE LOBE [B]	0	1	1	0	1	.5226	.7317	1	.7372	.5312
Adjusted # at risk	45.5	51.2	48.9	44.6	47.7					
GANGLIONEUROMA [B]	0	0	0	0	1	.2474	.4796	.	.	.
SALV GLND PAROT										
# Evaluated	53	55	55	52	52					
Adjusted # at risk	44.0	51.2	48.9	42.4	45.0					
ADENOMA [B]	1	0	0	0	0	.	.	.	.	1
SKIN AND SUBCUT										
# Evaluated	54	55	55	55	55					
Adjusted # at risk	44.5	51.2	48.9	44.6	47.5					
FIBROMA [B]	0	1	0	0	0	1	1	1	1	.5368
Adjusted # at risk	44.5	51.2	48.9	44.7	47.5					
FIBROSARCOMA [M]	0	0	0	1	0	.4789	.	.4632	.	.
Adjusted # at risk	44.5	51.2	48.9	44.6	47.5					
KERATOACANTHOMA [B]	0	1	0	0	0	1	1	1	1	.5368
Adjusted # at risk	44.5	51.3	48.9	44.6	47.5					
NEURAL CREST TUMOUR [M]	0	1	0	0	0	1	1	1	1	.5368
SPLEEN										
# Evaluated	55	55	55	55	55					
Adjusted # at risk	45.5	51.2	48.9	44.6	48.1					
HAEMANGIOSARCOMA [M]	0	0	0	0	1	.2513	.4848	.	.	.
Adjusted # at risk	45.5	51.2	48.9	44.6	48.1					
Hemangioma/Hemangiosarcoma	0	0	0	0	1	.2513	.4848	.	.	.
STOMACH										
# Evaluated	55	55	55	55	55					
Adjusted # at risk	45.5	51.2	48.9	44.7	47.5					
SCHWANNOMA [B]	0	0	0	1	0	.4789	.	.4632	.	.
Systemic										
# Evaluated	55	55	55	55	55					
Adjusted # at risk	45.6	51.2	48.9	44.6	48.1					
Hemangioma/Hemangiosarcoma	1	1	0	0	2	.1562	.4770	1	1	.7829

Reference ID: 3230218

Table A.2.3 (cont.) Overall Tumorigenicity Results in Female Rats

Organ/Tumor	Incidence					Significance Levels				
	Air	Veh	Low	Med	High	Trend	Hi vs Veh	Med vs Veh	Low vs Veh	Air vs Veh
THYMUS										
# Evaluated	55	55	54	50	54					
Adjusted # at risk	46.1	51.7	48.1	40.7	46.9					
THYMOMA [B]	6	5	2	0	3	.5282	.8298	1	.9342	.7946
THYROID GLAND										
# Evaluated	55	55	55	55	53					
Adjusted # at risk	45.6	52.2	48.9	44.6	46.5					
C-CELL ADENOMA [B]	4	6	6	4	2	.9303	.9561	.7641	.5620	.4658
Adjusted # at risk	45.5	51.2	48.9	44.6	46.5					
C-CELL CARCINOMA [M]	1	1	0	0	0	1	1	1	1	.7829
Adjusted # at risk	45.6	52.2	48.9	44.6	46.5					
C-cell Carcinoma/Adenoma	5	7	6	4	2	.9531	.9757	.8389	.6686	.4859
Adjusted # at risk	45.5	51.2	48.9	44.6	46.5					
FOLLICULAR CELL ADENOMA [B]	3	1	0	0	1	.4285	.7262	1	1	.9551
UTERUS										
# Evaluated	55	55	55	55	55					
Adjusted # at risk	46.2	51.7	49.1	46.4	48.9					
ADENOCARCINOMA [M]	4	4	5	7	7	.1756	.2281	.2054	.4743	.6991
Adjusted # at risk	46.0	51.2	48.9	44.6	47.5					
ADENOSQUAMOUS CARCINOMA [M]	1	0	0	0	0	.	.	.	.	1
Adjusted # at risk	46.4	51.7	49.1	46.4	48.9					
Adenocarc./Endo. Adenoma	5	5	7	7	7	.3212	.3373	.3083	.3516	.6940
Adjusted # at risk	45.7	51.2	48.9	44.6	47.5					
ENDOMETRIAL ADENOMA [B]	1	1	2	0	0	.9271	1	1	.4770	.7829
Adjusted # at risk	45.5	51.2	48.9	44.6	47.8					
ENDOMETRIAL CARCINOMA [M]	0	0	0	0	1	.2474	.4796	.	.	.
Adjusted # at risk	45.5	51.2	48.9	45.6	47.5					
ENDOMETRIAL STROMAL SARCOMA	0	0	0	1	0	.4817	.	.4687	.	.
Adjusted # at risk	45.5	51.2	48.9	44.6	47.5					
LEIOMYOMA [B]	0	0	1	0	0	.7316	.	.	.4848	.
Adjusted # at risk	45.5	51.2	48.9	45.0	47.5					
MALIGNANT SCHWANNOMA [M]	0	0	0	1	0	.4789	.	.4632	.	.
Adjusted # at risk	45.5	51.2	48.9	44.6	47.5					
POLYP [B]	1	0	0	0	0	.	.	.	.	1
Adjusted # at risk	48.2	53.3	50.4	46.4	49.5					
STROMAL POLYP [B]	24	32	24	22	22	.8774	.9613	.9270	.9278	.1984
Adjusted # at risk	48.2	53.3	50.4	47.3	49.5					
Strml Polyp/Endo.Strml Sarc.	24	32	24	23	22	.8795	.9613	.9115	.9278	.1984

Reference ID: 3230218

Table A.2.4. Overall Tumorigenicity Results in Male Mice

Organ/Tumor	Incidence					Significance Levels				
	Air	Veh	Low	Med	High	Trend	Hi vs Veh	Med vs Veh	Low vs Veh	Air vs Veh
ADRENAL GLAND										
# Evaluated	60	60	60	58	60					
Adjusted # at risk	51.2	45.1	51.3	47.7	49.8					
PHAEOCHROMOCYTOMA [B]	0	0	2	0	0	.8208	.	.	.2796	.
Adjusted # at risk	51.2	45.1	51.2	47.7	49.8					
PHAEOCHROMOCYTOMA [M]	0	0	0	1	0	.5000	.	.5109	.	.
Adjusted # at risk	51.2	45.1	51.3	47.7	49.8					
Phaeochromocytoma B&M	0	0	2	1	0	.7440	.	.5109	.2796	.
Adjusted # at risk	51.9	45.1	51.6	48.4	50.3					
SUBCAPSULAR CELL TUMOUR [B]	2	2	3	1	3	.3793	.5503	.8906	.5597	.6427
CAECUM										
# Evaluated	57	55	59	59	58					
Adjusted # at risk	49.1	43.3	51.5	49.0	48.4					
ADENOCARCINOMA [M]	0	0	2	0	0	.8294	.	.	.2917	.
EYE										
# Evaluated	59	58	58	57	59					
Adjusted # at risk	50.2	44.2	49.7	47.5	48.8					
RHABDOMYOSARCOMA [M]	0	0	0	0	1	.2553	.5217	.	.	.
FEMUR										
# Evaluated	60	60	60	60	59					
Adjusted # at risk	51.2	45.1	51.2	49.7	49.2					
MAST CELL TUMOUR [B]	1	0	0	0	0	.	.	.	.	1
HARDERIAN GLAND										
# Evaluated	60	60	60	60	60					
Adjusted # at risk	51.6	45.9	51.2	49.7	49.8					
ADENOMA [B]	2	4	3	2	3	.6072	.8162	.9165	.8306	.2809
HEMOPOIETIC SYS										
# Evaluated	60	60	60	60	60					
Adjusted # at risk	51.2	45.4	51.9	49.7	49.8					
HISTIOCYTIC SARCOMA [M]	0	1	1	0	0	.9471	1	1	.7829	.4687
Adjusted # at risk	51.2	46.0	51.2	49.7	49.8					
LEUKAEMIA, GRANULOCYTIC [M]	0	1	0	0	0	1	1	1	1	.4687
Adjusted # at risk	51.2	45.3	51.2	49.7	49.8					
LYMPHOMA [M]	0	1	0	0	0	1	1	1	1	.4687
Adjusted # at risk	52.0	47.6	53.1	51.8	51.7					
LYMPHOMA, FOLLIC CENTRE CELL	3	4	6	5	7	.2273	.3116	.5523	.4495	.4431
Adjusted # at risk	52.2	47.8	51.2	49.7	49.8					
LYMPHOMA, LYMPHOCYTIC [M]	2	3	1	0	2	.4958	.8319	1	.9506	.4517
Adjusted # at risk	51.2	45.1	51.2	49.7	49.8					
LYMPHOMA, PLASMACYTIC [M]	0	0	1	0	0	.7680	.	.	.5312	.
Adjusted # at risk	51.2	45.1	51.6	49.7	49.8					
MAST CELL TUMOUR [M]	1	1	2	1	0	.8865	1	.7735	.5473	.7204
KIDNEY										
# Evaluated	60	60	60	60	60					
Adjusted # at risk	51.2	45.1	51.2	49.7	49.8					
TUBULAR CELL ADENOMA [B]	1	0	0	0	0	.	.	.	.	1
LIVER										
# Evaluated	60	60	60	60	60					
Adjusted # at risk	51.2	45.1	51.8	50.0	50.3					
HAEMANGIOSARCOMA [M]	0	0	1	1	1	.3133	.5263	.5263	.5312	.
Adjusted # at risk	52.1	46.4	51.2	50.4	49.8					
HEPATOCELLULAR ADENOMA [B]	6	6	3	8	6	.3660	.6649	.4534	.9422	.5305
Adjusted # at risk	51.2	45.1	51.2	49.7	50.3					
HEPATOCELLULAR CARCINOMA [M]	4	1	1	3	3	.1560	.3494	.3414	.7829	.9616
LN MESENTERIC										
# Evaluated	60	60	57	60	59					
Adjusted # at risk	51.2	45.1	49.1	49.9	48.8					
HAEMANGIOMA [B]	0	0	0	1	0	.5079	.	.5213	.	.

Reference ID: 3230218

Table A.2.4 (cont.) Overall Tumorigenicity Results in Male Mice

Organ/Tumor	Incidence					Significance Levels				
	Air	Veh	Low	Med	High	Trend	Hi vs Veh	Med vs Veh	Low vs Veh	Air vs Veh
LUNG										
# Evaluated	60	60	60	60	60					
Adjusted # at risk	51.8	46.5	52.1	50.5	51.3					
BRONCHIOLO-ALVEOLAR ADENOMA	15	11	7	18	14	.1812	.4352	.1432	.9449	.7992
Adjusted # at risk	51.7	45.5	51.2	50.6	50.0					
BRONCHIOLO-ALVEOLAR CARCINOMA	3	1	2	4	2	.4154	.5322	.2154	.5473	.9248
Adjusted # at risk	52.4	47.0	52.1	51.4	51.5					
Bronch.-Alv. Carc./Adenoma	18	12	9	22	15	.2414	.4459	.0608	.9038	.8717
PITUITARY GLAND										
# Evaluated	55	56	56	59	56					
Adjusted # at risk	46.6	41.5	47.8	49.7	46.2					
ADENOMA, INTERMEDIATE LOBE [B]	1	0	0	0	0	.	.	.	.	1
Adjusted # at risk	46.6	42.0	47.8	49.7	46.2					
Adenoma/Carcinoma	1	1	0	0	0	1	1	1	1	.7233
Adjusted # at risk	46.6	42.0	47.8	49.7	46.2					
CARCINOMA, INTERMEDIATE LOBE	0	1	0	0	0	1	1	1	1	.4713
RIB										
# Evaluated	60	60	60	60	60					
Adjusted # at risk	51.2	45.1	51.2	49.7	49.8					
SARCOMA (NOT OTHERWISE SPEC)	0	0	0	1	0	.5052	.	.5213	.	.
SEMINAL VESICLE										
# Evaluated	60	60	57	59	60					
Adjusted # at risk	51.2	45.1	49.1	49.0	49.8					
ADENOCARCINOMA [M]	1	0	0	0	0	.	.	.	.	1
SKELETAL MUSCLE										
# Evaluated	60	60	60	60	60					
Adjusted # at risk	51.2	45.1	51.2	49.7	49.8					
FIBROSARCOMA [M]	1	0	0	0	0	.	.	.	.	1
Adjusted # at risk	51.2	45.1	51.2	49.9	49.8					
RHABDOMYOSARCOMA [M]	0	0	0	1	0	.5052	.	.5213	.	.
SKIN AND SUBCUT										
# Evaluated	60	60	60	60	60					
Adjusted # at risk	51.2	45.1	51.2	49.7	49.8					
FIBROMA [B]	1	0	1	0	0	.7680	.	.	.5312	1
Adjusted # at risk	51.2	45.8	51.2	50.4	49.8					
FIBROSARCOMA [M]	0	2	0	2	0	.8386	1	.7303	1	.2171
Adjusted # at risk	51.6	47.6	51.6	51.0	49.9					
Fibrosarc./Sarc./Rhabdomyosarc	1	6	3	3	2	.8979	.9744	.9341	.9380	.0441
Adjusted # at risk	51.2	45.1	51.2	49.7	49.8					
KERATOACANTHOMA [B]	0	0	1	1	0	.6386	.	.5213	.5312	.
Adjusted # at risk	51.2	45.1	51.2	49.7	49.8					
MALIGNANT FIBROUS HISTIOCYTOMA	0	1	0	0	0	1	1	1	1	.4687
Adjusted # at risk	51.2	45.1	51.2	49.7	49.8					
PAPILLOMA [B]	1	0	0	2	0	.5052	.	.2690	.	1
Adjusted # at risk	51.2	45.1	51.2	49.7	49.8					
PLEOMORPHIC SARCOMA [M]	1	0	0	0	0	.	.	.	.	1
Adjusted # at risk	51.6	46.8	51.6	49.7	49.9					
RHABDOMYOSARCOMA [M]	1	3	3	0	2	.6787	.8383	1	.7092	.2705
Adjusted # at risk	51.2	45.1	51.2	50.3	49.8					
SARCOMA (NOT OTHERWISE SPEC)	0	1	0	1	0	.7589	1	.7783	1	.4687
Adjusted # at risk	51.2	45.1	51.2	49.7	49.8					
SEBACEOUS CELL ADENOMA [B]	0	0	0	1	0	.5052	.	.5213	.	.
Adjusted # at risk	51.2	45.1	52.0	49.7	49.8					
SQUAMOUS-CELL CARCINOMA [M]	0	0	1	0	0	.7680	.	.	.5312	.

Reference ID: 3230218

Table A.2.4 (cont.) Overall Tumorigenicity Results in Male Mice

Organ/Tumor	Incidence					Significance Levels				
	Air	Veh	Low	Med	High	Trend	Hi vs Veh	Med vs Veh	Low vs Veh	Air vs Veh
SPLEEN										
# Evaluated	60	60	60	60	60					
Adjusted # at risk	51.2	45.1	51.2	49.7	49.8					
HAEMANGIOMA [B]	0	0	0	2	1	.2058	.5213	.2690	.	.
Adjusted # at risk	51.2	45.1	51.2	49.9	49.9					
HAEMANGIOSARCOMA [M]	1	1	0	2	1	.3775	.7735	.5322	1	.7204
Adjusted # at risk	51.2	45.1	51.2	49.7	49.8					
STROMAL SARCOMA [M]	0	1	0	0	0	1	1	1	1	.4687
STOMACH										
# Evaluated	60	60	59	60	60					
Adjusted # at risk	51.2	45.1	50.7	49.7	49.8					
PAPILLOMA [B]	0	0	0	0	2	.0635	.2690	.	.	.
Systemic										
# Evaluated	60	60	60	60	60					
Adjusted # at risk	51.2	45.1	51.2	49.9	49.8					
HAEMANGIOMA [B]	0	0	0	3	1	.2808	.5213	.1374	.	.
Adjusted # at risk	51.2	45.1	51.8	50.2	50.4					
HAEMANGIOSARCOMA [M]	1	1	1	3	2	.3085	.5399	.3494	.7829	.7204
Adjusted # at risk	51.2	45.1	51.8	50.4	50.4					
Hemangioma/Hemangiosarcoma	1	1	1	6	3	.2133	.3494	.0738	.7829	.7204
Adjusted # at risk	53.0	50.5	53.1	51.8	51.7					
Lymphoma	5	8	8	5	9	.3398	.5180	.8904	.6551	.2520
TESTIS										
# Evaluated	60	60	60	60	60					
Adjusted # at risk	52.2	45.1	51.3	50.9	50.4					
INTERSTITIAL CELL ADENOMA [B]	5	2	7	4	6	.2403	.1708	.3903	.1128	.9183
Adjusted # at risk	51.2	45.1	51.2	49.7	49.8					
SEMINOMA [B]	0	0	0	1	0	.5052	.	.5213	.	.
THYROID GLAND										
# Evaluated	60	59	59	60	60					
Adjusted # at risk	51.3	44.5	50.2	49.7	49.8					
FOLLICULAR CELL ADENOMA [B]	1	0	0	0	1	.2552	.5269	.	.	1
VERTEBRAE										
# Evaluated	11	18	13	10	10					
Adjusted # at risk	7.0	12.3	9.8	6.1	5.2					
MAST CELL TUMOUR [B]	0	1	0	0	0	1	1	1	1	.6316

Table A.2.5. Overall Tumorigenicity Results in Female Mice

Organ/Tumor	Incidence					Significance Levels				
	Air	Veh	Low	Med	High	Trend	Hi vs Veh	Med vs Veh	Low vs Veh	Air vs Veh
ABDOMINAL CAV										
# Evaluated	1	1	2	1	1					
Adjusted # at risk	0.1	1.0	2.0	0.4	1.0					
ADENOCARCINOMA [M]	0	0	1	0	0	.7500	.	.	.6667	.
Adjusted # at risk	0.1	1.0	1.6	1.0	1.0					
ANGIOLIPOMA [B]	0	0	0	1	0	.5000	.	.5000	.	.
Adjusted # at risk	0.1	1.0	1.6	0.4	1.0					
LEIOMYOSARCOMA [M]	0	1	0	0	0	1	1	.	1	.
ADRENAL GLAND										
# Evaluated	59	60	60	60	60					
Adjusted # at risk	42.9	45.0	43.6	39.3	41.8					
SUBCAPSULAR CELL TUMOUR [B]	1	1	0	0	0	1	1	1	1	.7644
BONE										
# Evaluated	2	0	0	1	1					
Adjusted # at risk	1.0	0.0	0.0	0.8	1.0					
OSTEOSARCOMA [M]	0	0	0	0	1	1	.	.	.	.

Reference ID: 3230218

Table A.2.5 (cont.) Overall Tumorigenicity Results in Female Mice

Organ/Tumor	Incidence					Significance Levels				
	Air	Veh	Low	Med	High	Trend	Hi vs Veh	Med vs Veh	Low vs Veh	Air vs Veh
BRAIN										
# Evaluated	60	60	60	60	60					
Adjusted # at risk	43.9	45.0	43.6	39.3	41.8					
MALIGNANT ASTROCYTOMA [M]	0	0	0	0	1	.2455	.4824	.	.	.
Adjusted # at risk	43.9	45.2	43.6	39.3	41.8					
OLIGODENDROGLIOMA [M]	0	1	0	0	0	1	1	1	1	.5114
CAECUM										
# Evaluated	60	57	58	59	59					
Adjusted # at risk	43.9	43.1	43.0	38.4	41.5					
LEIOMYOMA [B]	0	0	0	0	1	.2485	.4881	.	.	.
HARDERIAN GLAND										
# Evaluated	60	60	60	60	59					
Adjusted # at risk	44.8	46.8	45.9	40.1	42.2					
ADENOMA [B]	7	9	5	3	4	.8501	.9500	.9755	.9214	.4303
Adjusted # at risk	44.8	46.8	45.9	40.1	42.6					
Adenoma/Carcinoma	7	9	5	3	5	.7298	.8993	.9755	.9214	.4303
Adjusted # at risk	43.9	45.0	43.6	39.3	41.2					
CARCINOMA [M]	0	0	0	0	1	.2455	.4824	.	.	.
HEMOPOIETIC SYS										
# Evaluated	60	60	60	60	60					
Adjusted # at risk	45.7	45.0	43.7	40.9	41.8					
HISTIOCYTIC SARCOMA [M]	3	2	2	5	0	.8823	1	.1788	.6832	.8126
Adjusted # at risk	44.7	47.4	46.1	41.6	45.2					
LYMPHOMA, FOLLIC CENTRE CELL	4	8	8	7	11	.1546	.2673	.6074	.5894	.2106
Adjusted # at risk	44.9	46.7	43.8	40.4	43.1					
LYMPHOMA, LYMPHOCYTIC [M]	3	4	4	2	2	.8084	.8824	.8635	.6044	.5252
LIVER										
# Evaluated	60	60	60	60	60					
Adjusted # at risk	43.9	45.0	43.6	39.3	42.3					
HAEMANGIOSARCOMA [M]	0	0	0	0	1	.2500	.4884	.	.	.
Adjusted # at risk	43.9	45.0	43.6	40.0	41.8					
HEPATOCELLULAR ADENOMA [B]	2	0	0	3	1	.2630	.4824	.1037	.	1
LN MESENTERIC										
# Evaluated	59	59	56	60	60					
Adjusted # at risk	44.7	44.1	41.4	39.3	41.8					
HAEMANGIOMA [B]	2	1	0	0	1	.4364	.7350	1	1	.8793
LUNG										
# Evaluated	60	59	60	59	60					
Adjusted # at risk	44.5	44.6	45.3	40.7	42.9					
BRONCHIOLO-ALVEOLAR ADENOMA	3	6	9	12	8	.3866	.3494	.0592	.3028	.2420
Adjusted # at risk	45.2	44.6	45.5	40.0	42.3					
BRONCHIOLO-ALVEOLAR CARCINOMA	3	1	3	2	2	.4223	.4824	.4543	.3167	.9390
Adjusted # at risk	45.2	44.7	47.3	42.2	43.4					
Bronch.-Alv. Carc./Adenoma	5	7	12	14	10	.3706	.2767	.0512	.1923	.3627
Adjusted # at risk	43.9	44.4	43.6	38.8	41.8					
FIBROMA [B]	0	0	0	1	0	.4759	.	.4634	.	.
MAMMARY GLAND										
# Evaluated	60	59	60	59	60					
Adjusted # at risk	43.9	44.3	43.6	38.8	41.8					
ADENOACANTHOMA [M]	1	0	0	1	1	.1736	.4824	.4634	.	1
Adjusted # at risk	45.1	44.8	43.6	39.0	41.8					
ADENOCARCINOMA [M]	5	1	0	1	0	.7301	1	.7220	1	.9860
Adjusted # at risk	43.9	44.3	43.6	38.6	41.8					
ADENOMA [B]	0	0	0	1	0	.4759	.	.4634	.	.
Adjusted # at risk	45.1	44.8	43.6	39.3	41.8					
Adenoma/-carc./-canthoma	6	1	0	3	1	.3929	.7350	.2636	1	.9934

Reference ID: 3230218

Table A.2.5 (cont.) Overall Tumorigenicity Results in Female Mice

Organ/Tumor	Incidence					Significance Levels				
	Air	Veh	Low	Med	High	Trend	Hi vs Veh	Med vs Veh	Low vs Veh	Air vs Veh
NASAL CAVITY										
# Evaluated	60	59	60	60	60					
Adjusted # at risk	43.9	45.0	43.6	39.3	41.8					
PAPILLOMA [B]	0	1	0	0	0	1	1	1	1	.5057
Adjusted # at risk	43.9	45.0	43.6	39.3	41.8					
SCHWANNOMA [B]	1	0	0	0	0	.	.	.	.	1
NERVE										
# Evaluated	1	0	0	0	0					
Adjusted # at risk	1.0	0.0	0.0	0.0	0.0					
MALIGNANT SCHWANNOMA [M]	1	0	0	0	0	.	.	.	.	.
OVARY										
# Evaluated	60	59	60	60	60					
Adjusted # at risk	43.9	45.4	44.1	39.3	41.8					
CYSTADENOMA [B]	1	2	1	0	1	.6595	.8613	1	.8750	.5172
Adjusted # at risk	43.9	44.6	43.6	39.6	41.8					
GRANULOSA CELL TUMOUR [B]	0	3	1	1	3	.2862	.6275	.9261	.9390	.1249
Adjusted # at risk	44.4	44.4	43.6	39.3	41.8					
LUTEOMA [B]	3	2	0	0	1	.5730	.8659	1	1	.8198
Adjusted # at risk	43.9	44.4	43.6	39.7	42.6					
SEX CORD/STROMAL TUMOUR [B]	0	1	1	3	3	.1361	.2906	.2636	.7471	.5057
Adjusted # at risk	43.9	44.4	43.6	39.3	41.8					
TUBULOSTROMAL ADENOMA [B]	1	1	0	0	1	.4319	.7350	1	1	.7586
OVIDUCT										
# Evaluated	0	1	0	1	2					
Adjusted # at risk	0.0	1.0	0.0	1.0	2.0					
LEIOMYOSARCOMA [M]	0	0	0	0	1	.6667	.6667	.	.	.
PITUITARY GLAND										
# Evaluated	56	59	58	58	58					
Adjusted # at risk	40.9	44.0	42.6	38.3	40.0					
ADENOMA, ANTERIOR LOBE [B]	0	1	1	1	0	.8087	1	.7213	.7471	.5181
SKIN AND SUBCUT										
# Evaluated	60	60	60	60	60					
Adjusted # at risk	44.5	45.4	44.0	40.1	42.5					
Adenocarc./Adenoma	2	1	1	3	1	.5069	.7354	.2647	.7414	.8834
Adjusted # at risk	44.5	45.0	43.6	39.4	41.8					
BASOSQUAMOUS CARCINOMA [M]	1	0	0	1	0	.4790	.	.4699	.	1
Adjusted # at risk	43.9	45.0	43.6	39.3	42.4					
KERATOACANTHOMA [B]	0	0	0	0	1	.2500	.4884	.	.	.
Adjusted # at risk	44.5	45.0	43.6	39.3	41.8					
LIPOSARCOMA [M]	1	0	0	0	0	.	.	.	.	1
Adjusted # at risk	43.9	45.0	44.0	39.3	41.8					
MALIGNANT SCHWANNOMA [M]	0	0	1	0	0	.7381	.	.	.5000	.
Adjusted # at risk	43.9	45.0	43.6	39.3	41.8					
PAPILLOMA [B]	0	0	0	1	0	.4790	.	.4699	.	.
Adjusted # at risk	43.9	45.0	43.6	39.6	42.5					
RHABDOMYOSARCOMA [M]	0	0	0	1	1	.1781	.4884	.4699	.	.
Adjusted # at risk	43.9	45.4	44.0	39.8	41.8					
SARCOMA (NOT OTHERWISE SPEC)	1	1	1	2	0	.7706	1	.4459	.7414	.7641
Adjusted # at risk	43.9	45.0	43.6	39.3	41.8					
SQUAMOUS-CELL CARCINOMA [M]	0	0	0	1	0	.4790	.	.4699	.	.
Adjusted # at risk	43.9	45.0	43.6	39.3	42.4					
Strml Polyp/Endo.Strml Sarc.	0	0	0	1	1	.1781	.4884	.4699	.	.
SPINAL CORD										
# Evaluated	60	60	60	60	60					
Adjusted # at risk	43.9	45.0	43.6	39.3	41.8					
MENINGEAL SARCOMA [M]	0	0	0	0	1	.2455	.4824	.	.	.

Reference ID: 3230218

Table A.2.5 (cont.) Overall Tumorigenicity Results in Female Mice

Organ/Tumor	Incidence					Significance Levels				
	Air	Veh	Low	Med	High	Trend	Hi vs Veh	Med vs Veh	Low vs Veh	Air vs Veh
SPLEEN										
# Evaluated	60	60	60	60	60					
Adjusted # at risk	43.9	45.0	43.6	39.3	41.8					
HAEMANGIOMA [B]	0	0	1	0	0	.7365	.	.	.4943	.
Adjusted # at risk	43.9	45.0	43.6	39.3	41.8					
HAEMANGIOSARCOMA [M]	0	0	0	0	1	.2455	.4824	.	.	.
STOMACH										
# Evaluated	59	59	60	60	59					
Adjusted # at risk	43.1	44.1	43.6	39.6	41.3					
CARCINOMA [M]	0	0	0	1	0	.4790	.	.4699	.	.
Systemic										
# Evaluated	60	60	60	60	60					
Adjusted # at risk	45.0	45.5	43.9	40.1	41.8					
HAEMANGIOMA [B]	2	2	2	1	1	.7237	.8613	.8563	.6744	.7003
Adjusted # at risk	44.3	45.1	43.6	39.3	42.6					
HAEMANGIOSARCOMA [M]	1	1	0	0	3	.0479	.2824	1	1	.7584
Adjusted # at risk	45.4	45.6	43.9	40.1	42.6					
Hemangioma/Hemangiosarcoma	3	3	2	1	4	.2016	.4609	.9264	.8053	.6617
Adjusted # at risk	45.8	49.1	46.3	42.7	46.5					
Lymphoma	7	12	12	9	13	.3309	.4267	.7230	.5223	.2065
THYROID GLAND										
# Evaluated	60	60	60	59	59					
Adjusted # at risk	43.9	45.0	43.6	38.6	41.3					
FOLLICULAR CELL CARCINOMA [M]	0	1	0	0	0	1	1	1	1	.5057
URINARY BLADDER										
# Evaluated	60	58	59	60	58					
Adjusted # at risk	43.9	43.4	44.0	39.3	41.3					
MESENCHYMAL PROLIF LESION	0	0	1	0	1	.3054	.4881	.	.5000	.
UTERUS										
# Evaluated	60	60	60	60	60					
Adjusted # at risk	43.9	45.0	43.6	39.3	41.8					
ADENOCARCINOMA [M]	0	0	2	1	1	.3956	.4824	.4699	.2414	.
Adjusted # at risk	43.9	45.0	43.6	39.3	41.8					
CARCINOMA [M]	0	0	0	0	1	.2455	.4824	.	.	.
Adjusted # at risk	43.9	45.0	43.6	39.3	41.8					
DECIDUOMA [B]	1	0	0	0	0	.	.	.	.	1
Adjusted # at risk	43.9	45.4	44.0	39.5	42.6					
ENDOMETRIAL STROMAL SARCOMA	5	3	3	3	2	.6637	.7976	.5911	.6509	.8813
Adjusted # at risk	43.9	45.0	43.6	39.3	41.8					
FIBROMA [B]	0	0	1	0	0	.7365	.	.	.4943	.
Adjusted # at risk	43.9	45.3	43.6	40.1	41.8					
GRANULAR CELL TUMOUR [B]	0	1	0	1	1	.3694	.7291	.7227	1	.5114
Adjusted # at risk	43.9	45.4	43.9	40.1	41.8					
HAEMANGIOMA [B]	0	1	1	1	0	.8110	1	.7227	.7414	.5114
Adjusted # at risk	44.3	45.1	43.6	39.3	42.0					
HAEMANGIOSARCOMA [M]	1	1	0	0	1	.4364	.7354	1	1	.7584
Adjusted # at risk	43.9	45.0	43.6	39.3	41.8					
HISTIOCYTIC SARCOMA [M]	2	1	1	0	1	.5222	.7350	1	.7471	.8836
Adjusted # at risk	44.3	46.3	47.2	44.7	46.7					
LEIOMYOMA [B]	5	9	18	19	22	.0126	.0038	.0138	.0385	.2177
Adjusted # at risk	43.9	45.0	45.1	41.4	42.6					
LEIOMYOSARCOMA [M]	1	2	5	10	4	.4039	.3160	.0092	.2263	.5087
Adjusted # at risk	44.3	46.3	48.7	46.7	47.5					
Leiomyoma/Leiomyosarcoma	6	11	23	28	26	.0141	.0018	.0003	.0133	.1648
Adjusted # at risk	43.9	45.0	43.6	39.3	42.5					
MAST CELL TUMOUR [B]	0	0	0	0	1	.2500	.4884	.	.	.
Adjusted # at risk	44.5	45.6	44.4	42.0	44.4					
POLYP [B]	7	4	8	8	12	.0248	.0227	.1339	.1655	.9086
VERTEBRAE										
# Evaluated	23	23	17	19	18					
Adjusted # at risk	18.7	15.9	12.3	12.2	12.1					
OSTEOSARCOMA [M]	1	0	0	0	0	.	.	.	.	1

Reference ID: 3230218

Appendix 4. References

Bailer, A. and Portier, C. (1988), "Effects of Treatment-Induced Mortality on Tests for Carcinogenicity in Small Samples", *Biometrics*, **44**, 4, 417-431.

Bieler, G.S., and Williams, R.L. (1993), "Ratio Estimates, the Delta Method, and Quantal Response Tests for Increased Carcinogenicity", *Biometrics*, **49**, 4, 793-801.

Chu, K.C., Ceuto, C., and Ward, J.M. (1981), "Factors in the Evaluation of 200 National Cancer Institute Carcinogen Bioassays", *Journal of Toxicology and Environmental Health,* 8, 251-280.

Haseman, J. K. (1983), "A Reexamination of False-positive Rates for Carcinogenicity Studies", *Fundamental and Applied Toxicology,* 3, 334-339.

Lin, K. K. and Ali, M.W. (2006), "Statistical Review and Evaluation of Animal Tumorigenicity Studies", *Statistics in the Pharmaceutical Industry,Third Edition,* edited by C.R. Buncher and J.Y. Tsay, Marcel Dekker, Inc. New York.

Lin, K. K. and Rahman, M.A. (1998), "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.

McConnell, E.E., Solleveld, H.A., Swenberg, J.A., and Boorman, G.A. (1986), "Guidelines for Combining Neoplasms for Evaluation of Rodent Carcinogenesis Studies", *Journal of the National Cancer Institute*. **76**, 283-289.

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

Rahman, M.A. and Lin, K.K. (2008), "A Comparison of False Positive Rates of Peto and Poly-3 Methods for Long Term Carcinogenicity Data Analysis Using Multiple Comparison Adjustment Method Suggested by Lin and Rahman", *Journal of Biopharmaceutical Statistics*. 18, 949-958.

STP Peto Working Group (2002), "Statistical Methods for Carcinogenicity Studies", *Toxicologic Pathology*. **30**, 3, 403-414.

U.S. Department of Health and Human Services (2001), Guidance for Industry Statistical Aspects of the Design, Analysis, and Interpretation of Chronic Rodent Carcinogenicity Studies of Pharmaceuticals (DRAFT GUIDANCE), Center for Drug Evaluation and Research, Food and Drug Administration

Reference ID: 3230218

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

/s/

STEVEN F THOMSON
12/13/2012
Statistical Carcinogenicity Review

KARL K LIN
12/13/2012
Concur with review

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

CLINICAL STUDY

NDA / Sequence Number:	NDA 203108 / Seq 0000
Drug Name:	(b) (4) ® Respimat® Inhalation Spray (olodaterol)
Proposed Indication(s):	Maintenance treatment of airflow obstruction in patients with chronic obstructive pulmonary disease, including chronic bronchitis and/or emphysema
Applicant:	Boehringer Ingelheim
Date(s):	Received: 05-14-2012 PDUFA Due Date: 03-14-2013
Review Priority:	Standard
Biometrics Division:	Division of Biometrics II/Office of Biostatistics
Statistical Reviewer:	Robert Abugov, Ph.D.
Concurring Reviewer:	Joan Buenconsejo, Ph.D.
Statistics Supervisor:	Thomas Permutt, Ph.D. (Division Director)
Medical Division:	Division of Pulmonary, Allergy, and Rheumatology Products
Clinical Team:	Robert Lim, M.D. (Medical Reviewer) Terri Michele, M.D. (Medical Team Leader) Badrul A Chowdhury, M.D., Ph.D. (Medical Division Director)
Project Manager:	Christine Chung

Keywords: NDA review, Clinical Studies

Contents

1 EXECUTIVE SUMMARY 5

2 INTRODUCTION 6

2.1 OVERVIEW 6

2.1.1 Drug Class and Indication 6

2.1.2 History of Drug Development 6

2.1.2 Current Submission 7

2.2 DATA SOURCES 9

3 STATISTICAL EVALUATION 10

3.1 EVALUATION OF EFFICACY 10

3.1.1 Study Design and Endpoints 10

3.1.2 Statistical Methodologies 13

3.1.3 Patient Disposition, Demographic and Baseline Characteristics 17

3.1.4 Results and Conclusions 20

3.1.4.1 Primary Endpoint: Trough FEV_1 (Studies 11, 12, 13, 14) 20

3.1.4.2 Primary Endpoint: FEV_1 $AUC_{0\text{-}3hr}$ (Studies 11, 12, 13, and 14) 26

3.1.4.3 Secondary Endpoint: FEV_1 $AUC_{0\text{-}12hr}$ (Studies 11 and 12) 27

3.1.4.4 24-hour Response Profile: FEV_1 $AUC_{12\text{-}24hr}$ and FEV_1 $AUC_{0\text{-}24hr}$ (Studies 24, 25, 39, and 40) 28

3.1.4.5 Interim Summary: Spirometry Results for FEV_1 (Studies 11, 12, 13, 14, 24, 25, 39, and 40) 37

3.1.4.6 Other Primary Endpoint: Mahler TDI Focal Score (Studies 13 and 14) 37

3.1.4.7 Key Secondary Endpoint: SGRQ (Studies 13 and 14) 38

3.1.4.8 Key Secondary Endpoint: Exacerbations (Studies 11, 12, 13, and 14) 38

3.1.4.9 Exercise Endurance (Studies 37 and 38) 40

3.1.4.10 Inspiratory Capacity (IC) at Isotime (Studies 37 and 38) 40

3.1.4.11 Breathing Discomfort: Borg Category Ratio Scale (Studies 37 and 38) 40

3.1.4.12 Functional Residual Capacity (Studies 37 and 38) 40

4 SAFETY 42

4.1 PATIENT MORTALITY (STUDIES 11, 12, 13, AND 14) 42

5 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS 43

5.1 GENDER, RACE, AGE, AND GEOGRAPHIC REGION 43

6 SUMMARY AND CONCLUSIONS 47

6.1 STATISTICAL ISSUES 47

6.2 COLLECTIVE EVIDENCE 47

6.3 CONCLUSIONS AND RECOMMENDATIONS 47

6.4 LABELING RECOMMENDATIONS 48

Reference ID: 3247160

Tables

Table 1. Randomized Phase III spirometry studies in current submission. Trial numbers cross reference to label 8
Table 2. Randomized Phase III exercise tolerance studies in current submission..... 9
Table 3. Dates of final data collection, SAP modification, and database lock 17
Table 4. Patient disposition, number and percent discontinuing due to AE 18
Table 5. Patient disposition, number and percent discontinuing due to lack of efficacy, protocol non-compliance, loss to follow-up, withdrawal of consent or other causes 19
Table 6. Patient demographics, number and percent. Gender, age, and smoking status 49
Table 7. Patient demographics, number and percent. Race, country and tiotropium use 50
Table 8. Patient demographics, mean and standard deviation. Age, FEV_1, FEV_1 % predicted, and smoking pack years 51
Table 9. Change from baseline trough FEV_1 (liters), pre-planned analyses..... 52
Table 10. Change from baseline trough FEV1, observed data only 53
Table 11. Change from baseline trough FEV_1 by week, study 11 54
Table 12. Change from baseline trough FEV_1 by week, study 12 55
Table 13. Change from baseline trough FEV_1 by week, study 13 56
Table 14. Change from baseline trough FEV_1 by week, study 14 57
Table 15. Hypothesis generating trough FEV_1 analyses, studies 11 and 12, olodaterol effects modeled without tiotropium interactions 58
Table 16. Trough FEV1, stratified by baseline tiotropium or long term anticholinergic 59
Table 17. Change from baseline FEV_1 $AUC_{0\text{-}3hr}$ (liters) 61
Table 18. Change from baseline FEV_1 $AUC_{0\text{-}3hr}$ (liters), 48 week studies, 6 and 12 weeks..... 62
Table 19. Change from baseline FEV_1 $AUC_{0\text{-}12hr}$ (liters)..... 63
Table 20. Change from baseline FEV_1 $AUC_{12\text{-}24hr}$ (liters) 64
Table 21.Change from baseline FEV_1 $AUC_{0\text{-}24hr}$ (liters)..... 64
Table 22. Change from baseline Mahler TDI Focal Score, studies 13 and 14 65
Table 23. SGRQ, studies 13 and 14..... 66
Table 24. Exercise endurance, studies 37 and 38 67
Table 25. Inspiratory capacity (L), studies 37 and 38..... 67
Table 26. Breathing discomfort at exercise isotime, studies 37 and 38..... 68
Table 27. Functional residual capacity (L), studies 37 and 38. Exploratory analysis 68
Table 28. Exacerbations, 48 week studies 69
Table 29. FEV_1 $AUC_{0\text{-}3hr}$ (L), by age..... 70
Table 30. FEV_1 $AUC_{0\text{-}12hr}$ (L), by age 71
Table 31. Trough FEV_1 (L), by age 72

Reference ID: 3247160

Figures

Figure 1. Change from baseline trough FEV_1 at Week 12 (Olodaterol 5 vs placebo) 21
Figure 2. Change from baseline trough FEV_1 at Week 12 (Olodaterol 10 vs olodaterol 5) 21
Figure 3. Change from baseline trough FEV_1 (observed data only) 22
Figure 4. Change from baseline trough FEV_1 by Week (Olodaterol 5 vs placebo) 23
Figure 5. Hypothesis generating analysis, change from baseline trough FEV1 (Studies 11 and 12) 24
Figure 6. Hypothesis generating analysis, change from baseline trough FEV_1 at Week 12, by tiotropium use 25
Figure 7. 'Confirmatory' analyses, change from baseline trough FEV_1 at Week 12 by tiotropium use 25
Figure 8. Change from baseline $AUC_{0\text{-}3hr}$ at Week 12 (Olodaterol 5 vs placebo) 26
Figure 9. Change from baseline FEV_1 $AUC_{0\text{-}3hr}$ at Week 12 (Olodaterol 10 vs olodaterol 5) 27
Figure 10. Change from baseline trough FEV_1 (Crossover Studies at Week 6 and Parallel Arm Studies at Week 12) 29
Figure 11. Change from baseline FEV_1 $AUC_{0\text{-}3hr}$ (Crossover Studies at Week 6 and Parallel Arm Studies at Week 12) 29
Figure 12. Change from baseline FEV_1 $AUC_{0\text{-}12hr}$ (Crossover Studies at Week 6 and Parallel Arm Studies at Week 12) 30
Figure 13. Change from baseline trough FEV_1 (Parallel Arm Trials at Weeks 6 and 12) 30
Figure 14. Change from baseline trough FEV_1 $AUC_{0\text{-}3hr}$ (Parallel Arm Trials at Weeks 6 and 12) 31
Figure 15. Change from baseline $AUC_{12\text{-}24hr}$ at Week 6 (Olodaterol 5 vs placebo) 32
Figure 16. Change from baseline FEV_1 $AUC_{0\text{-}24hr}$ at Week 6 (Olodaterol 5 vs placebo) 32
Figure 17. Change from baseline FEV_1 $AUC_{12\text{-}24hr}$ at Week 6 (Olodaterol 10 vs olodaterol 5) 33
Figure 18. Change from baseline FEV1 $AUC_{0\text{-}24}$ hr at Week 6 (Olodaterol 10 vs olodaterol 5) 33
Figure 19. Diurnal FEV_1 by treatment, Study 24 34
Figure 20. Diurnal FEV_1 by treatment, Study 25 35
Figure 21. Diurnal FEV_1 by treatment, Study 39 35
Figure 22. Diurnal FEV_1 by treatment, Study 40 36
Figure 23. Moderate exacerbations (Olodaterol 5 vs placebo) 39
Figure 24. Exacerbations resulting in hospitalization (Olodaterol 5 vs placebo) 39
Figure 25. Kaplan-Meier percent mortality, to 330 days post-treatment 42
Figure 26. Effect of age on change from baseline FEV_1 $AUC_{0\text{-}3hr}$ at Week 12 (Olodaterol 5 vs placebo) in Studies 11, 12, 13 and 14 44
Figure 27. Effect of age on change from baseline FEV_1 $AUC_{0\text{-}3hr}$ at Week 6 (Olodaterol 5 vs placebo) in Studies 24, 25, 39, and 40 44
Figure 28. Effect of age on change from baseline FEV_1 $AUC_{0\text{-}12hr}$ at Week 12 (Olodaterol 5 vs placebo) in Studies 11 and 12 45
Figure 29. Effect of age on change from baseline FEV_1 $AUC_{0\text{-}12hr}$ at Week 6 (Olodaterol 5 vs placebo) in Studies 24, 25, 39, and 40 45
Figure 30. Effect of age on change from baseline trough FEV_1 at Week 12 (Olodaterol 5 vs placebo) in Studies 11, 12, 13 and 14 46
Figure 31. Effect of age on change from baseline FEV_1 at Week 6 (Olodaterol 5 vs placebo) in Studies 24, 25, 39, and 40 46

Reference ID: 3247160

1 EXECUTIVE SUMMARY

Boehringer Ingelheim Pharmaceuticals, Inc. has proposed Striverdi Respimat Inhalation Spray, olodaterol 5 mcg provided in two actuations once daily (qd), for the maintenance treatment of airflow obstruction in patients with chronic obstructive pulmonary disease (COPD), including chronic bronchitis and/or emphysema. Efficacy and safety of this putative long acting beta agonist (LABA) were examined in ten phase 3 clinical trials. Primary efficacy variables addressed pulmonary function in eight of the phase 3 studies and exercise endurance in two of the phase 3 studies.

The submission demonstrated benefits of olodaterol to pulmonary function. Four randomized parallel arm trials showed that olodaterol 5 mcg administered via Respimat inhaler qd (olodaterol 5) as an add-on to standard of care without other LABA provided statistically significant benefits to the primary endpoints: change from baseline in trough one second forced expiratory volume (FEV_1), with an average benefit of 65 mL at week 12, and FEV_1 area under curve from 0 to 3 hours after treatment administration ($AUC_{0\text{-}3hr}$), with an average benefit of 155 mL at week 12. Statistically significant benefits to the secondary endpoint, change from baseline FEV_1 $AUC_{0\text{-}12hr}$ (122 mL average at week 12) were also demonstrated, in two randomized parallel arm studies. There was no statistically significant difference in benefit between olodaterol 5 and a higher dose, olodaterol 10 mcg administered via Respimat inhaler qd (olodaterol 10). Statistically significant improvements in FEV_1 $AUC_{12\text{-}24hr}$ and FEV_1 $AUC_{0\text{-}24hr}$ at week 6 were demonstrated in four crossover trials. Due to gaps in the data, quantitative estimates of improvement were not available for these variables.

A statistically significant difference in St. George's Respiratory Questionnaire (SGRQ) between olodaterol 5 and placebo was seen in only one of two parallel arm studies. The biggest mean reduction compared to placebo in any single study was 3.15, which is less than the 4.0 threshold considered clinically significant. No effect of olodaterol 5 was seen for COPD exacerbation rate or change from baseline Mahler transition dyspnea index (TDI) score.

Statistically significant benefits of olodaterol 5 were seen in two crossover studies for week 6 inspiratory capacity (IC) at isotime (130 mL average) and exercise endurance (47 seconds average). Hypothesis generating analyses on functional residual capacity (FRC) suggested that, compared to placebo, olodaterol 5 improves post-dose (-212 mL average), but not pre-dose FRC.

Reference ID: 3247160

2 INTRODUCTION

2.1 Overview

2.1.1 Drug Class and Indication

Boehringer Ingelheim Pharmaceuticals, Inc proposes Striverdi Respimat Inhalation Spray, olodaterol 2.5 mcg ex-mouthpiece, two actuations once daily, as a LABA for the maintenance treatment of airflow obstruction in patients with COPD, including chronic bronchitis and/or emphysema.

2.1.2 History of Drug Development

The clinical development program for olodaterol was introduced to the Division of Pulmonary, Allergy, and Rheumatology Products in 2007 under IND 076362. The applicant had several interactions with the Agency and pertinent parts of the statistical portion of those communications and interactions are summarized herein.

The design and analysis of the phase 3 studies (Table 1) was discussed at the End-of-Phase 2 meeting held on July 17, 2008. The division informally agreed with the applicant's proposal for the primary efficacy endpoints and to the 12-week time point for analysis in Studies 1222.11 and 1222.12, as well as to the primary efficacy endpoints in studies 1222.24 and 1222.25. While the division informally agreed to the statistical analysis plan for the primary efficacy endpoints in studies 1222.11 and 1222.12, we also suggested that the applicant consider additional imputation methods to handle missing data to ensure that the treatment difference observed is not being influenced by missing data. The division recommended that a closed testing procedure be done within the dose levels for the primary lung function endpoints. The division also recommended inclusion of serial spirometry testing at additional time points beyond three hours post-dose in studies 1222.11 and 1222.12 to characterize the 24 hour profile of olodaterol. The division agreed that inclusion and stratification of patients by treatment with tiotropium were appropriate in studies 1222.11, 1222.12, 1222.13 and 1222.14. In addition, they recommended that stratification of patient by country should be performed in studies 1222.11 and 1222.12. The division agreed that the 5 mcg and 10 mcg doses olodaterol were reasonable to evaluate in the Phase 3 studies. However, the division also recommended an evaluation of different dosing frequencies for olodaterol to support the proposed dosing regimen.

Reference ID: 3247160

In a communication dated June 25, 2010, the Division recommended that the applicant conduct additional dose-ranging and dosing frequency studies in a bronchodilator-responsive population (asthma patients) before proceeding with the development program in a COPD population. Finally, a written response prior to the pre-NDA meeting was provided to the applicant on September 28, 2011 generally agreeing to the proposed content and format of the NDA. In addition, the Division agreed with the proposed subgroup analyses for the pulmonary function tests in the ten studies, and additionally recommended that the applicant include subgroup analyses based on beta-agonist reversal.

2.1.2 Current Submission

The applicant's proposed indication for maintenance treatment of airflow obstruction in patients with COPD is based on four 48-week placebo-controlled studies (studies 1222.11, 1222.12, 1222.13 and 1222.14, hereafter referred to as 11, 12, 13 and 14). Other evidence includes four 6-week, placebo- and active-controlled, cross-over studies to characterize bronchodilating profile of olodaterol (studies 1222.24, 1222.25, 1222.39 and 1222.40, hereafter referred to as 24, 25, 39 and 40), and two 6-week placebo-controlled cross-over exercise tolerance studies (studies 1222.37 and 1222.38, hereafter referred to as 37 and 38), see Table 1 and Table 2. All 10 studies are included in the statistics review. Olodaterol will sometimes be referred to as ‘Olo’ in the remainder of the review.

Reference ID: 3247160

Table 1. Randomized Phase III spirometry studies in current submission. Trial numbers cross reference to label

Study	Design	Population	Endpoints
1222.11 (Trial 1) 1222.12 (Trial 2)	Olo 5 Olo 10 Placebo +ICS, SABA, AC as needed Parallel arm 48 Weeks DB	COPD Age ≥ 40 years Pack years > 10 Post-bronchodilator FEV_1 ≤ 80% pred Post-bronchodilator FEV_1/FVC ≤ 70% N(Study 11)=208:208:209 N(Study 12)=209:217:216 47% USA	*Primary:* ΔFEV_1 $AUC_{0\text{-}3hr}$ W12 ΔTrough FEV_1 W12 *Secondary:* ΔFEV_1 $AUC_{0\text{-}12hr}$ W12 Time to mod exacerbation
1222.13 (Trial 3) 1222.14 (Trial 4)	Olo 5 Olo 10 Placebo F 12 +ICS, SABA, AC as needed Parallel arm 48 Weeks DB, DD	COPD Age ≥ 40 years Pack years > 10 Post-bronchodilator FEV_1 ≤ 80% pred Post-bronchodilator FEV_1/FVC ≤ 70% N(Study 13)=227:225:225:227 N(Study 14)=232:234:235:233 No USA	*Primary:* ΔFEV_1 $AUC_{0\text{-}3hr}$ W24 ΔTrough FEV_1 W24 Mahler TDI focal score W24 *Secondary:* *SGRQ W24* Time to mod exacerbation
1222.24 (Trial 5) 1222.25 (Trial 6)	Olo 5 Olo 10 Placebo F 12 +ICS, SABA, AC as needed Crossover W4S 6 Weeks 14 day washout DB, DD	COPD Age ≥ 40 years Pack years > 10 Post-bronchodilator FEV_1 ≤ 80% pred Post-bronchodilator FEV_1/FVC ≤ 70% N(Study 24)=95:92:93:96 N(Study 25)=93:95:94:93 100% USA	*Primary:* ΔFEV_1 $AUC_{0\text{-}12hr}$ W6 ΔFEV_1 $AUC_{12\text{-}24hr}$ W6 *Secondary* ΔFEV_1 $AUC_{0\text{-}24hr}$ W6

ICS: Inhaled corticosteroids, SABA: Short acting beta-agonist, AC: anticholinergic, F 12: Foradil 12 mcg twice daily (bid), DB: double blind, DD: double dummy, FVC: forced vital capacity.

Reference ID: 3247160

Table 2. Randomized Phase III exercise tolerance studies in current submission

Study	Design	Population	Endpoints
1222.39 (Trial 7) 1222.40 (Trial 8)	Olo 5 Olo 10 Placebo Tio 18 +ICS, SABA as needed Crossover W4S4P 6 Weeks 21 day washout DB, DD	COPD 40 ≤ age ≤75 Pack years > 10 Post-bronchodilator FEV_1 ≤ 80% pred Post-bronchodilator FEV_1/FVC ≤ 70% Discontinue oral beta adrenergics, short acting anticholinergics N(Study 39)=101:101:102:101 N(Study 40)=116:113:110:113 No USA	*Primary:* ΔFEV_1 $AUC_{0\text{-}12hr}$ W6 ΔFEV_1 $AUC_{12\text{-}24hr}$ W6 *Secondary* Δ FEV_1 $AUC_{0\text{-}24hr}$ W6
1222.37 (Trial 9) 1222.38 (Trial 10)	Olo 5 Olo 10 Placebo +ICS, SABA, AC as needed Crossover W6S3P 6 Weeks 14 day washout DB	COPD Age ≥ 40 years Pack years > 10 Post-bronchodilator FEV_1 ≤ 80% pred Post-bronchodilator FEV_1/FVC ≤ 70% N(Study 37)=147:143:143 N(Study 38)=150:147:149 No USA	*Primary:* Exercise endurance W6 *Secondary* IC at isotime W6 Breathing discomfort intensity Isot W6

Tio 18: tiotropium 18 mcg qd, Isot: Isotime

2.2 Data Sources

NDA 203-108 was submitted on May 14, 2012. The study reports including protocols, statistical analysis plan, and all referenced literature were submitted by the applicant to the Agency.

Reference ID: 3247160

3 STATISTICAL EVALUATION

3.1 Evaluation of Efficacy

3.1.1 Study Design and Endpoints

Studies 11, 12, 13 and 14 were of similar design, except for the time of analysis of the primary endpoints (12 weeks in studies 11 and 12, and 24 weeks in studies 13 and 14) and the addition of EU marketed Foradil Aerolyzer (F), formoterol fumarate 12 mcg bid arm in studies 13 and 14. They were all 48-week, randomized, international, double blind, placebo controlled, parallel arm studies designed to evaluate the efficacy (based on lung function in studies 11, 12, 13 and 14 and symptomatic benefit in studies 13 and 14) and safety of once daily treatment of orally inhaled olodaterol delivered by the Respimat inhaler, in patients with COPD. The studies randomized COPD patients at least 40 years of age with a minimum of ten pack years of smoking, post-bronchodilator $FEV_1 \leq 80\%$ predicted, and post-bronchodilator $FEV_1/FVC \leq 70\%$, to placebo, olodaterol 5, olodaterol 10, or to EU marketed Foradil Aerolyzer (F), formoterol fumarate 12 mcg bid (in studies 13 and 14 only). Randomization was stratified by tiotropium use (Yes/No) at screening.

In these studies, the primary endpoints were:

1. FEV_1 $AUC_{0\text{-}3hr}$ response (at Day 85 or week 12 in studies 11 and 12, and at Day 169 or week 24 in studies 13 and 14)
2. Trough FEV_1 response (at week 12 in studies 11 and 12, and at week 24 in studies 13 and 14)
3. Mahler TDI focal score in studies 13 and 14.

The secondary endpoints included:

1. FEV_1 $AUC_{0\text{-}12hr}$ response at week 12 (in a subset of 241 patients in study 11 and 321 patients in study 12)
2. SGRQ score at week 24 in studies 13 and 14
3. Time to exacerbations.

FEV_1 AUC was defined as the area under the FEV_1 curve normalized for time. FEV_1 $AUC_{x\text{-}yhr}$ was calculated as the area under the FEV_1 time curve from 0 to 3 hours post-dose using the trapezoidal rule, divided by the duration (y-x hours) and reported in liters.

Trough FEV_1 was defined as the FEV_1 value at the end of the dosing interval, 24 hours post-drug administration, recorded as the mean of the 10 minute and 1 hour values prior to treatment at weeks 0, 6, and 12, and recorded only 10 minutes prior to treatment on weeks 2, 18, 24, 32, 40, and 48. Trough FEV_1 response was defined as the change from baseline in trough FEV_1.

Reference ID: 3247160

Baseline for all clinical spirometry endpoints was calculated as the mean of the two pre-dose pulmonary function tests at week 0, conducted 1 hour prior to and 10 minutes prior to administration of the first dose of randomized treatment.

Of note, only a subset of patients performed 12-hour post-dose pulmonary function tests at week 12.

The applicant evaluated other secondary endpoints including FEV_1 $AUC_{0\text{-}3hr}$, FEV_1 $AUC_{0\text{-}12hr}$, and trough FEV_1 responses at different timepoints, other spirometry endpoints (FVC, peak expiratory flow or PEF), patient global rating, rescue use, and COPD exacerbations, as well as TDI and SGRQ at different timepoints (in studies 13 and 14 only). However, only the results from the analyses of the primary and key secondary endpoints, including COPD exacerbation will be presented in this review.

Studies 24, 25, 39 and 40 provided pulmonary function tests conducted earlier than week 12. They were of similar design except for the active comparator (US marketed formoterol 12 mcg bid in studies 24 and 25, and tiotropium 18 mcg qd in studies 39 and 40) and the clinical sites (US sites in studies 24 and 25, and international sites in studies 39 and 40). All were six-week randomized, multicenter, double blind, placebo controlled, four sequence, four period Williams square design[1] crossover studies designed to evaluate the efficacy and safety of six weeks of once daily treatment of orally inhaled olodaterol delivered by the Respimat inhaler, in patients with COPD.

The studies randomized COPD patients who were at least 40 years of age and not over 75 years of age with a minimum of ten pack years of smoking, post-bronchodilator $FEV_1 \leq 80\%$ predicted, and post-bronchodilator $FEV_1/FVC \leq 70\%$, to one of four treatment sequences:

Studies 24 and 25

(a) Placebo, Olo 5 mcg, Olo 10 mcg , Formoterol;
(b) Olo 5, Formoterol, Placebo, Olo 10;
(c) Olo 10, Placebo, Formoterol, Olo 5; or
(d) Formoterol, Olo 10, Olo 5, Placebo.

Studies 39 and 40

(a) Olo 5, Olo 10, Tiotropium, Placebo;
(b) Olo 10, Placebo Olo 5, Tiotropium;
(c) Tiotropium, Olo 5, Placebo, Olo 10; or
(d) Placebo, Tiotropium, Olo 10, Olo 5.

[1] Williams crossover designs provide estimates of treatment differences which are not aliased with sequence or period effects, but which are aliased with carryover effects. The designs are balanced and uniform, i.e., equal numbers of patients receive each treatment within a period and the sequence for each patient includes every treatment.

Reference ID: 3247160

In these studies, the primary endpoints were:

1. FEV_1 $AUC_{0\text{-}12hr}$ response after 6 weeks of treatment
2. FEV_1 $AUC_{12\text{-}24hr}$ response after 6 weeks of treatment.

The key secondary endpoint was:

1. FEV_1 $AUC_{0\text{-}24hr}$ after week 6 of treatment.

The primary comparison was between olodaterol and placebo followed by olodaterol and the active comparator (formoterol or tiotropium).

The applicant also evaluated several other endpoints including FEV_1 $AUC_{0\text{-}3hr}$ responses, FVC responses after 1st dose and after 6 weeks of treatment, and peak FEV_1 responses.

Studies 37 and 38 were both six-week randomized, international, double blind, placebo controlled, six sequence, three period Williams square crossover designs to evaluate the efficacy and safety of once daily treatment of orally inhaled olodaterol delivered by the Respimat inhaler, in patients with COPD. The studies randomized COPD patients who were at least 40 years of age with a minimum of ten pack years of smoking, post-bronchodilator $FEV_1 \leq 80\%$ predicted, and post-bronchodilator $FEV_1/FVC \leq 70\%$, to one of six treatment sequences denoted, using A=Olo 5, B=Olo 10, and C=P, ABC, ACB, BAC, BCA, CAB, and CBA. Randomization was blocked by treatment sequence.

The primary efficacy endpoint was

1. logarithm of endurance time[2] during constant work rate cycle ergometry to symptom limitation at 75% Wcap[3].

The secondary endpoints were

1. IC at isotime[4] during constant work rate cycle ergometry to symptom limitation at 75% Wcap
2. intensity of breathing discomfort (Borg Category-Ratio Scale) at isotime during constant work rate cycle ergometry to symptom limitation at 75% Wcap.

The applicant also evaluated other secondary endpoints including functional residual capacity (FRC), total lung capacity (TLC), FVC, PEF and FEV_1.

---

[2] Endurance time during constant work rate cycle ergometry to symptom limitation at 75% Wcap after 6 weeks of treatment (Day 43).
[3] Maximal work rate achieved.
[4] For each individual subject, isotime was defined as the endurance time of the constant work rate exercise test of shortest duration from study baseline or at W6 of any treatment period.

Reference ID: 3247160

3.1.2 Statistical Methodologies

Analyses for primary and key secondary variables varied between study pairs.

In studies 11 and 12, the pre-planned analyses for primary efficacy endpoints compared treatment groups using a restricted maximum likelihood repeated measures model, with fixed effects of treatment, baseline, stratum (tiotropium use), time, treatment by time interaction, baseline by time interaction, treatment by stratum interaction, stratum by time interaction, and stratum by time by treatment interaction.

In studies 13 and 14, the pre-planned analyses for primary efficacy endpoints compared treatment groups using a restricted maximum likelihood mixed effect repeated measures model (MMRM), with fixed effects treatment, baseline, tiotropium use (stratum), time, treatment by time interaction, and baseline by time interaction.

Analyses for studies 11, 12, 13, and 14 all used a spatial power covariance structure with Kenward-Rogers degrees of freedom.

In studies 11, 12, 13, and 14 at weeks 0 (Day 1 of administration), 2, 6, 24, and 48, pulmonary function testing (PFT) was conducted 10 minutes prior to administration of randomized treatment, and 5 minutes, 15 minutes, 30 minutes, and 1, 2, and 3 hours following administration of randomized treatment. Additional PFT, 1 hour prior to administration of randomized treatment, was provided at weeks 0 and 6.

In studies 11 and 12, week 12 PFT was conducted 10 minutes and 1 hour prior to administration of randomized treatment, and 5 minutes, 15 minutes, 30 minutes, and 1, 2, 3, 4, 5, 6, 8, 10, and 12 hours following administration of randomized treatment. At weeks 18, 32, and 40, PFT was conducted at a single timepoint, 10 minutes prior to administration of randomized treatment.

In studies 13 and 14, the week 12 PFT corresponded to those conducted at weeks 2, 6, 24, and 48.

Pulmonary function testing for studies 11, 12, 13, and 14 at week -2 was conducted between 7:00 am and 10:00 am. At week 0 and in subsequent visits the start of all PFT measurements was relative to the administration time of trial medication, with the time of the -10 minute PFT always within 30 minutes of the time of the -10 minute PFT performed at week 0.

Efficacy analyses were conducted on the full analysis set (FAS) population, consisting for each variable of individuals who have both baseline measurements and at least one measurement after initiation of randomized treatment. The primary treatment comparisons were contrasts between treatments at week 12 (in studies 11 and 12) and at week 24 (in studies 13 and 14). Differences were tested at the two sided 0.05 level of significance.

Reference ID: 3247160

As only a single measurement (at week 12) was taken to evaluate the secondary efficacy variable FEV1 $AUC_{0\text{-}12hr}$ response in studies 11 and 12, the ANCOVA model was used for analysis of this endpoint. For the analysis of time to exacerbation, Cox regression and log-rank tests were used. The Cox model contained a stratification term for tiotropium use stratum and a covariate for treatment group. For analyses of exacerbation counts, negative binomial models accounting for exposure were used.

Least favorable observation carried forward (LFOCF) was used in studies 11, 12, 13, and 14 for patients who discontinued the study due to worsening of symptoms. Imputation of missing data was applied in the analyses of the primary and key secondary endpoints. As discussed in the results section, use of this imputation method may be problematic when conducting MMRM analyses, and this review therefore includes sensitivity analyses conducted without imputation for one of the primary variables, trough FEV_1.

Crossover studies 24 and 25 used MMRM analyses with fixed effects treatment, center, and period, and random effect patient nested within center with compound symmetric covariance matrices. Efficacy analyses were conducted on all randomized patients administered at least one dose of randomized treatment with at least one baseline measurement and at least one measurement for the first co-primary endpoint. Analysis plans were further modified for studies 39 and 40, with baseline measurement as an additional fixed effect and with patient rather than patient nested within center as the random effect. Efficacy analyses were conducted within a period on all randomized patients with baseline data and any evaluable post dose data for either coprimary endpoint.

In studies 24, 25, 39 and 40, PFT measurements from patient visits after treatment discontinuation were considered completely missing. No data imputation was performed on completely missing visits.

In studies 37 and 38, analyses for the primary efficacy endpoint used a mixed effect repeated measures model, with treatment, log baseline endurance time, and period as fixed effects and with patient as a random effect, using a compound symmetric covariance matrix. Sensitivity analyses were conducted using an unstructured covariance matrix. Spirometry measurements for secondary endpoints were analyzed using the statistical model described for primary endpoints, but on a linear scale. Efficacy analyses were conducted on all randomized patients with baseline and any evaluable post dosing data for the primary endpoint.

No imputations were conducted for missing endurance time. Missing IC and Borg scale values were imputed using LOCF. Linear interpolation was used for Borg and IC values at isotime if no value was available at the exact time point.

Reference ID: 3247160

To control Type I error in the face of multiple endpoints and multiple doses, hypotheses were tested in the following hierarchy.

- **In studies 11 and 12: Primary comparison between olodaterol and placebo at Week 12**

  1. Δ FEV_1 $AUC_{0\text{-}3h}$ olodaterol 10 mcg
  2. Δ Trough FEV_1 olodaterol 10 mcg
  3. Δ FEV_1 $AUC_{0\text{-}3h}$ olodaterol 5 mcg
  4. Δ Trough FEV_1 olodaterol 5 mcg

- **In studies 13 and 14: Primary comparison between olodaterol and placebo at Week 24**

  1. Δ FEV_1 $AUC_{0\text{-}3h}$ olodaterol 10 mcg
  2. Δ Trough FEV_1 olodaterol 10 mcg
  3. Δ FEV_1 $AUC_{0\text{-}3h}$ olodaterol 5 mcg
  4. Δ Trough FEV_1 olodaterol 5 mcg
  5. TDI focal score olodaterol 10 mcg
  6. TDI focal score olodaterol 5 μg

- **In studies 24, 25, 39 and 40: Primary comparison between olodaterol and placebo**

  1. Δ FEV_1 $AUC_{0\text{-}12h}$ olodaterol 10 mcg
  2. Δ FEV_1 $AUC_{12\text{-}24h}$ olodaterol 10 mcg
  3. Δ FEV_1 $AUC_{0\text{-}12h}$ olodaterol 5 mcg
  4. Δ FEV_1 $AUC_{12\text{-}24h}$ olodaterol 5 mcg

- **In studies 37 and 38: Primary comparison between olodaterol and placebo**

  1. Log endurance time olodaterol 10 mcg
  2. Log endurance time olodaterol 5 mcg
  3. IC at isotime olodaterol 10 mcg
  4. IC at isotime olodaterol 5 mcg
  5. Intensity of breathing discomfort olodaterol 10 mcg
  6. Intensity of breathing discomfort olodaterol 5 mcg

Reference ID: 3247160

For secondary comparisons, the following analysis hierarchies were used to control Type I error.

- **In studies 11, 12, 13, 14: Secondary comparison between olodaterol and placebo**

  1. Week 24 SGRQ oldaterol 10 studies 13 and 14 combined data
  2. Exacerbations olodaterol 10 studies 11, 12, 13, 14 combined data
  3. Exacerbations olodaterol 5 studies 11, 12, 13, 14 combined data

  Tests of significance for exacerbations were to be conducted only if olodaterol differed significantly from placebo for all three co-primary endpoints in studies 13 and 14 and the key secondary endpoint for these studies, SGRQ.

- **In studies 13 and 14: Secondary comparison at Week 24, data combined from both studies**

  1. SGRQ total score olodaterol 10 vs placebo
  2. SGRQ total score olodaterol 5 vs placebo
  3. FEV_1 $AUC_{0\text{-}3h}$ olodaterol 10 vs formoterol
  4. Trough FEV_1 olodaterol 10 vs formoterol
  5. FEV_1 $AUC_{0\text{-}3h}$ response olodaterol 5 vs formoterol
  6. Trough FEV_1 in patients treated with olodaterol 5 vs formoterol
  7. TDI focal score in patients treated with Olodaterol 10 vs formoterol
  8. TDI focal score in patients treated with Olodaterol 5 vs formoterol

  {formoterol is European formulation not used in United States}

- **In studies 24, 25 : No stated analysis hierarchy for secondary comparisons**

- **In studies 37 and 38: Secondary comparison between olodaterol and placebo**

  1. IC at isotime during constant work rate cycle ergometry to symptom limitation at 75% Wcap Olo 10
  2. IC at isotime during constant work rate cycle ergometry to symptom limitation at 75% Wcap Olo 5
  3. Intensity of breathing discomfort (Borg Category-Ratio Scale) at isotime during constant work rate cycle ergometry to symptom limitation at 75% Wcap Olo
  4. Intensity of breathing discomfort (Borg Category-Ratio Scale) at isotime during constant work rate cycle ergometry to symptom limitation at 75% Wcap Olo 5

Control of Type I error within each study was strong over all variables, with secondary variables tested for potential significance in each study if tests of the primary variables were significant.

Reference ID: 3247160

Analysis plans for each study were finalized after all data had been collected but prior to database lock (Table 3).

Table 3. Dates of final data collection, SAP modification, and database lock

Study	Date		
	Last Data	SAP Finalized	Database Lock
11	9/21/2010	11/2/2010	11/29/2010
12	9/27/2010	11/2/2010	11/29/2010
13	12/2/1010	2/17/2011	2/24/2011
14	12/8/2010	2/17/2011	2/24/2011
24	4/30/2010	6/10/2010	6/11/2010
25	4/28/2010	6/10/2010	7/19/2010
39	1/4/2011	12/10/2010	2/11/2011
40	1/10/2011	12/10/2010	2/11/2011

3.1.3 Patient Disposition, Demographic and Baseline Characteristics

In all 10 studies, there were no obvious imbalances in gender (Table 6), race (Table 7), age (Table 6, Table 8), smoking status at initiation of treatment (Table 6), residence in the United States (Table 7), use of tiotropium at initiation of treatment (Table 7), trough FEV_1 (Table 8), trough FEV_1 percent predicted (Table 8), or smoking pack years (Table 8).

In the 48 week parallel arm trials 11, 12, 13, and 14, percent discontinuation and percent discontinuation due to adverse events associated with COPD were numerically higher among patients administered placebo than among patients administered olodaterol (Table 4). Similarly, in the 48 week parallel arm trials, percent withdrawal due to lack of efficacy was numerically higher in the placebo group than in the olodaterol group (Table 5).

Reference ID: 3247160

Table 4. Patient disposition, number and percent discontinuing due to AE, n (%)

Study	Trtmnt	Treated	Discontinued	Discontinued Due to AE		
				COPD	Other Disease	Other
			48-week placebo –controlled studies			
11	Placebo	209 (100)	50 (24)	13 (6)	2 (1)	6 (3)
	Olo 5	208 (100)	35 (17)	7 (3)	0 (0)	10 (5)
	Olo 10	207 (100)	35 (17)	4 (2)	3 (1)	9 (4)
12	Placebo	216 (100)	41 (19)	8 (4)	3 (1)	9 (4)
	Olo 5	209 (100)	24 (11)	3 (1)	3 (1)	4 (2)
	Olo 10	217 (100)	36 (17)	2 (1)	2 (1)	16 (7)
13	Placebo	225 (100)	57 (25)	7 (3)	0 (0)	11 (5)
	Olo 5	227 (100)	36 (16)	11 (5)	2 (1)	3 (1)
	Olo 10	225 (100)	39 (17)	4 (2)	0 (0)	11 (5)
	For 12	227 (100)	43 (19)	4 (2)	2 (1)	14 (6)
14	Placebo	235 (100)	51 (22)	5 (2)	2 (1)	12 (5)
	Olo 5	232 (100)	37 (16)	6 (3)	1 (0)	9 (4)
	Olo 10	234 (100)	36 (15)	4 (2)	3 (1)	9 (4)
	For 12	233 (100)	40 (17)	9 (4)	1 (0)	6 (3)
			6-week cross-over studies			
24	Placebo	96 (100)	6 (6)	1 (1)	0 (0)	3 (3)
	Olo 5	95 (100)	6 (6)	1 (1)	1 (1)	3 (3)
	Olo 10	92 (100)	3 (3)	1 (1)	0 (0)	0 (0)
	For 12	93 (100)	4 (4)	1 (1)	1 (1)	0 (0)
25	Placebo	94 (100)	10 (11)	4 (4)	1 (1)	2 (2)
	Olo 5	93 (100)	9 (10)	4 (4)	1 (1)	2 (2)
	Olo 10	95 (100)	12 (13)	4 (4)	1 (1)	2 (2)
	For 12	93 (100)	10 (11)	5 (5)	0 (0)	1 (1)
39	Placebo	102 (100)	8 (8)	2 (2)	0 (0)	1 (1)
	Olo 5	101 (100)	7 (7)	3 (3)	0 (0)	1 (1)
	Olo 10	101 (100)	7 (7)	2 (2)	0 (0)	2 (2)
	Tio 18	101 (100)	9 (9)	2 (2)	0 (0)	4 (4)
40	Placebo	110 (100)	14 (13)	5 (5)	0 (0)	4 (4)
	Olo 5	116 (100)	19 (16)	3 (3)	1 (1)	6 (5)
	Olo 10	113 (100)	16 (14)	4 (4)	0 (0)	6 (5)
	Tio 18	113 (100)	16 (14)	4 (4)	1 (1)	4 (4)
			6-week exercise endurance studies			
37	Placebo	143 (100)	9 (6)	2 (1)	0 (0)	3 (2)
	Olo 5	147 (100)	13 (9)	2 (1)	1 (1)	6 (4)
	Olo 10	143 (100)	9 (6)	2 (1)	1 (1)	5 (3)
38	Placebo	149 (100)	16 (11)	3 (2)	2 (1)	6 (4)
	Olo 5	150 (100)	15 (10)	3 (2)	2 (1)	6 (4)
	Olo 10	147 (100)	14 (10)	2 (1)	1 (1)	7 (5)

source: disposition.sas

Reference ID: 3247160

Table 5. Patient disposition, number and percent discontinuing due to lack of efficacy, protocol non-compliance, loss to follow-up, withdrawal of consent or other causes

Study	Trtmnt	Treated	Lack of Efficacy	Protocol Compliance	Lost to Follow-up	Withdrew Consent	Other
			48-week placebo-controlled studies				
11	Placebo	209 (100)	13 (6)	0 (0)	2 (1)	11 (5)	3 (1)
	Olo 5	208 (100)	4 (2)	3 (1)	1 (0)	8 (4)	2 (1)
	Olo 10	207 (100)	1 (0)	2 (1)	2 (1)	11 (5)	3 (1)
12	Placebo	216 (100)	10 (5)	2 (1)	1 (0)	3 (1)	5 (2)
	Olo 5	209 (100)	5 (2)	0 (0)	2 (1)	5 (2)	2 (1)
	Olo 10	217 (100)	2 (1)	0 (0)	3 (1)	8 (4)	3 (1)
13	Placebo	225 (100)	9 (4)	2 (1)	2 (1)	20 (9)	6 (3)
	Olo 5	227 (100)	2 (1)	3 (1)	2 (1)	9 (4)	4 (2)
	Olo 10	225 (100)	1 (0)	2 (1)	4 (2)	11 (5)	6 (3)
	For 12	227 (100)	3 (1)	3 (1)	0 (0)	14 (6)	3 (1)
14	Placebo	235 (100)	8 (3)	2 (1)	2 (1)	16 (7)	4 (2)
	Olo 5	232 (100)	1 (0)	2 (1)	1 (0)	8 (3)	9 (4)
	Olo 10	234 (100)	3 (1)	0 (0)	2 (1)	8 (3)	7 (3)
	For 12	233 (100)	2 (1)	2 (1)	3 (1)	13 (6)	4 (2)
			6-week cross-over studies				
24	Placebo	96 (100)	0 (0)	1 (1)	1 (1)	0 (0)	0 (0)
	Olo 5	95 (100)	0 (0)	0 (0)	1 (1)	0 (0)	0 (0)
	Olo 10	92 (100)	0 (0)	1 (1)	1 (1)	0 (0)	0 (0)
	For 12	93 (100)	0 (0)	1 (1)	0 (0)	1 (1)	0 (0)
25	Placebo	94 (100)	0 (0)	0 (0)	0 (0)	1 (1)	2 (2)
	Olo 5	93 (100)	0 (0)	0 (0)	0 (0)	0 (0)	2 (2)
	Olo 10	95 (100)	0 (0)	0 (0)	0 (0)	3 (3)	2 (2)
	For 12	93 (100)	0 (0)	0 (0)	0 (0)	2 (2)	2 (2)
39	Placebo	102 (100)	1 (1)	0 (0)	0 (0)	2 (2)	2 (2)
	Olo 5	101 (100)	1 (1)	0 (0)	0 (0)	1 (1)	2 (2)
	Olo 10	101 (100)	1 (1)	0 (0)	0 (0)	1 (1)	2 (2)
	Tio 18	101 (100)	0 (0)	0 (0)	0 (0)	1 (1)	2 (2)
40	Placebo	110 (100)	1 (1)	3 (3)	0 (0)	0 (0)	4 (4)
	Olo 5	116 (100)	2 (2)	4 (3)	0 (0)	1 (1)	5 (4)
	Olo 10	113 (100)	1 (1)	2 (2)	0 (0)	1 (1)	4 (4)
	Tio 18	113 (100)	2 (2)	4 (4)	0 (0)	0 (0)	4 (4)
			6-week exercise endurance studies				
37	Placebo	143 (100)	0 (0)	1 (1)	0 (0)	1 (1)	3 (2)
	Olo 5	147 (100)	0 (0)	1 (1)	0 (0)	0 (0)	4 (3)
	Olo 10	143 (100)	0 (0)	0 (0)	0 (0)	0 (0)	2 (1)
38	Placebo	149 (100)	0 (0)	2 (1)	1 (1)	0 (0)	2 (1)
	Olo 5	150 (100)	0 (0)	1 (1)	0 (0)	1 (1)	2 (1)
	Olo 10	147 (100)	0 (0)	1 (1)	0 (0)	1 (1)	2 (1)

Reference ID: 3247160

3.1.4 Results and Conclusions

Because studies 13 and 14 did not randomize any patients from the United States, the applicant designated studies 11 and 12 as the 'pivotal' trials for evaluation of approval with the Agency. However, when evaluating whether to approve a drug, the Agency does not subscribe to the concept of pivotal trials, and instead considers all available data. Nevertheless, to control Type I error, endpoints to be evaluated for approval must be clearly defined before the data is analyzed. Therefore, the Agency considered the pre-planned endpoints of studies 11 and 12 as primary for evaluation of efficacy, applying the ordered analysis hierarchy pre-specified in study 11 and 12 protocols to all available data in the submission.

3.1.4.1 Primary Endpoint: Trough FEV_1 (Studies 11, 12, 13, 14)

3.1.4.1.1 Trough FEV_1: Preplanned Analysis

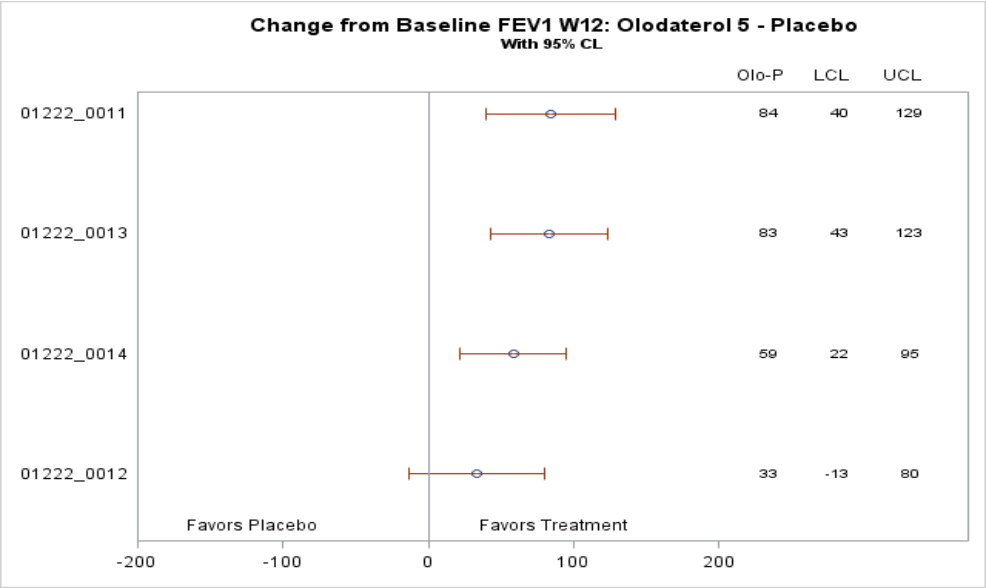
With the exception of study 12, differences between olodaterol 5 and placebo for week 12 change from baseline trough FEV_1 were statistically significant (Figure 1, underlying Table 9). The effect size ranged from 0.033 to 0.084 liters.

The analyses provided in Figure 1 differ for studies 11 and 12 from those presented by the applicant in the efficacy results sections of the clinical trial reports. Those reports provided post-hoc analyses which suggested larger differences between olodaterol 5 and placebo and which were claimed to be 'statistically significant' for study 12. The premises of the applicant's post-hoc analyses will be examined, and refuted, in Section 3.1.4.1.4 below.

The difference between olodaterol 5 and placebo for change from baseline trough FEV_1 at week 24 was statistically significant in studies 13 (78 mL average, 95% CI 37−118) and 14 (average 53 mL 95% CI 15−90).

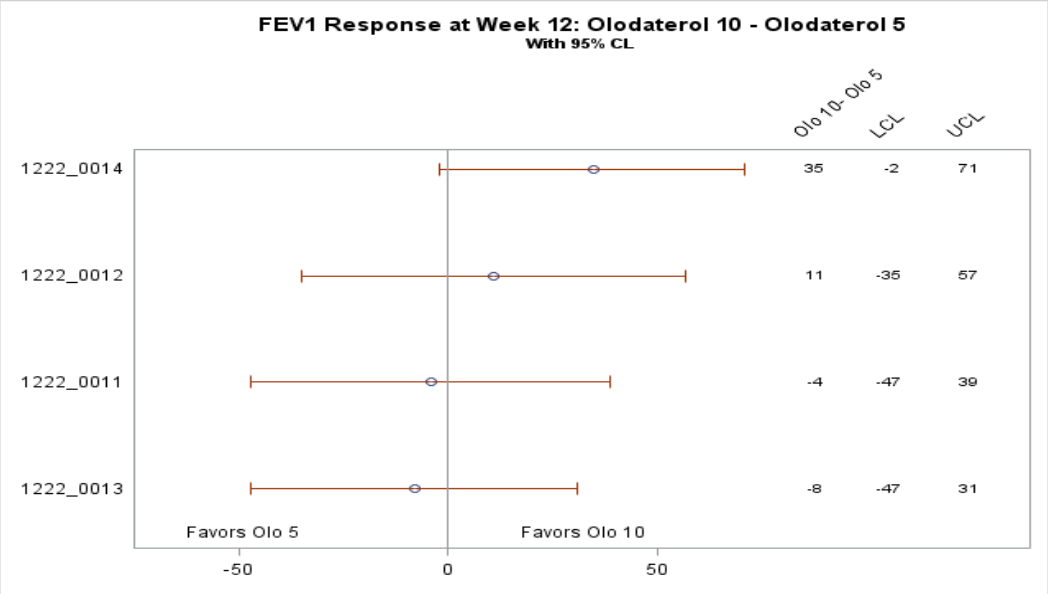
Reference ID: 3247160

Figure 1. Change from baseline trough FEV_1 at Week 12 (Olodaterol 5 vs placebo)



Olodaterol 10 did not have a larger effect on trough FEV_1 than olodaterol 5 (Figure 2, underlying Table 9). This result is in agreement with that presented by the applicant.

Figure 2. Change from baseline trough FEV_1 at Week 12 (Olodaterol 10 vs olodaterol 5)



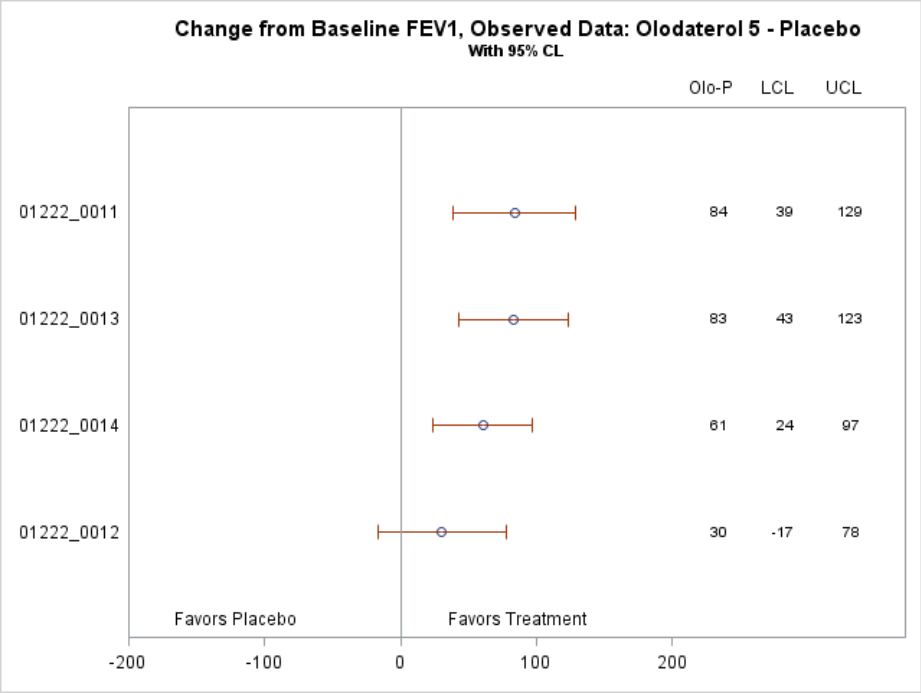
source: Forest Plots 2012 11 26.sas

Reference ID: 3247160

3.1.4.1.2 Trough FEV_1 Sensitivity analyses

The application of mixed effect repeated measures models to data imputed by carrying forward earlier measurements provided in the applicant's prespecified analyses can distort covariance matrices used for the calculation of confidence limits. To examine the effect of the applicant's carry forward imputation on calculated confidence limits, the pre-specified mixed effect repeated measures models were applied to observed data only. Because the widths of the confidence limits calculated using observed data only (Figure 3, underlying Table 10) were within 2 mL of those calculated using carry forward imputed data (Figure 1), the effects of the pre-specified carry forward imputation on the confidence intervals provided in Figure 1 are not an issue of major concern.

Figure 3. Change from baseline trough FEV_1 (observed data only)



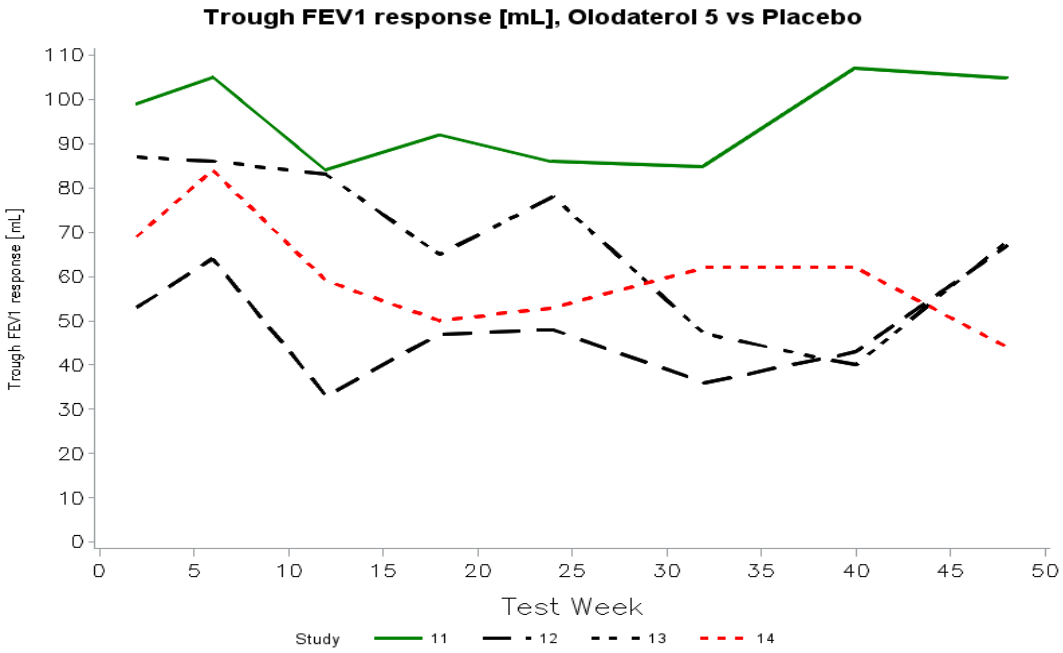
source: Forest Plots 2012 12 01.sas

3.1.4.1.3 Trough FEV_1: Efficacy during first treatment year

Studies 11, 12, 13, and 14 provide no concrete evidence of attenuation of benefit to trough FEV_1 during 48 weeks of treatment (Figure 4, underlying Table 11, Table 12, Table 13, and Table 14). The difference between olodaterol 5 and placebo was numerically greater at week 48 than at week 2 in studies 11 and 13, and was less at week 48 than at week 2 in studies 12 and 14.

Reference ID: 3247160

Figure 4. Change from baseline trough FEV_1 by Week (Olodaterol 5 vs placebo)



source: FEV1 2012 11 29.sas

3.1.4.1.4 Trough FEV_1: Applicant claims from post-hoc analysis

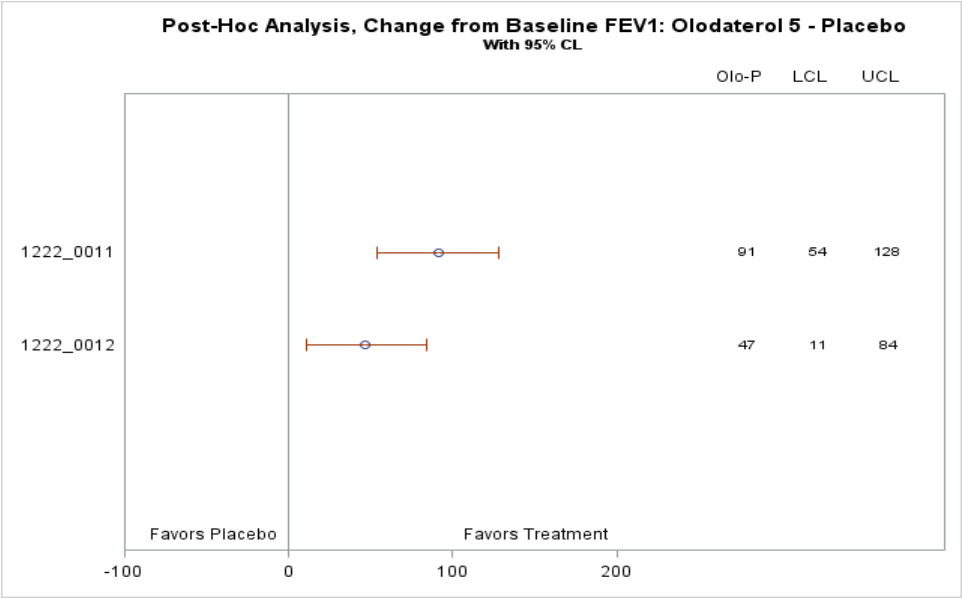
Instead of pre-planned analyses, the submission presented post-hoc analyses which purported to demonstrate that the benefits compared to placebo of olodaterol in studies 11 and 12 were larger than those from the pre-planned analyses shown in Figure 1 (0.91 and 0.47 liters, respectively), and which also purported to demonstrate that the effect in study 12 of olodaterol 5 relative to placebo was statistically significant (Figure 5, underlying Table 15). The applicant argued that, because the benefit of olodaterol was less among patients treated with tiotropium (Figure 6, underlying Table 16), the target population, in which only a minority of patients were taking tiotropium (<25% in these studies), would be more accurately reflected by a model which downweights patients taking tiotropium, which the applicant attempted to accomplish by removing from the statistical model interaction terms between tiotropium stratum and treatment.

The applicant's replacement with post-hoc analyses results from pre-planned analyses of studies 11 and 12 is not acceptable. First, and most important, the use of post-hoc analyses for inference of statistical significance is incorrect, because type I error cannot be controlled (or calculated) when conducting multiple unplanned statistical tests. Second, the applicant's argument is based on a presumption from Figure 6 that tiotropium use reduces the effect of olodaterol. However, analyses of data from studies 13, 14, 24, and 25 show in three of those four studies numerically

Reference ID: 3247160

larger rather than smaller effects of olodaterol among patients taking tiotropium or anticholinergics (Figure 7, underlying Table 16).

Figure 5. Hypothesis generating analysis, change from baseline trough FEV1 (Studies 11 and 12)



source: Forest Plots 2012 12 07.sas

Reference ID: 3247160

Figure 6. Hypothesis generating analysis, change from baseline trough FEV_1 at Week 12, by tiotropium use

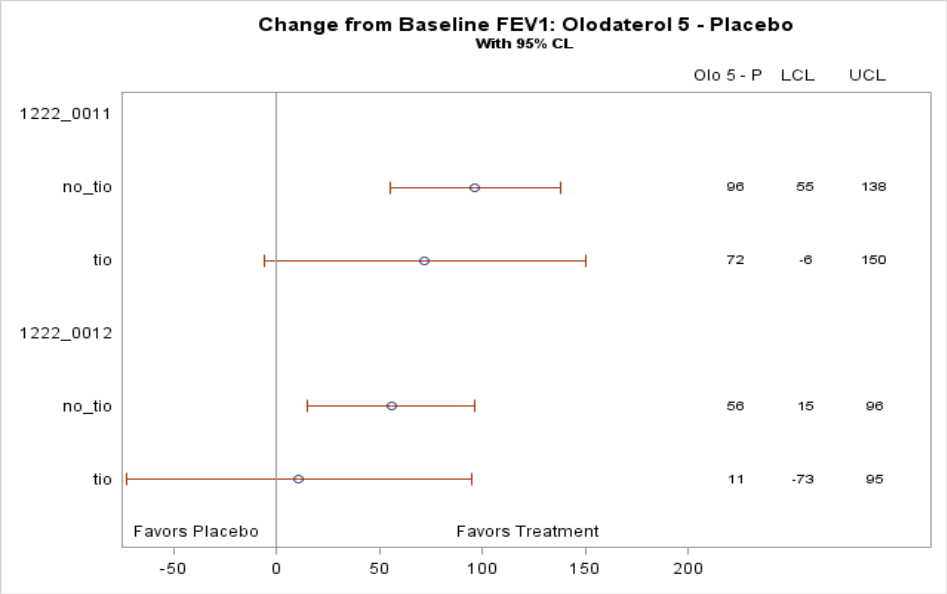
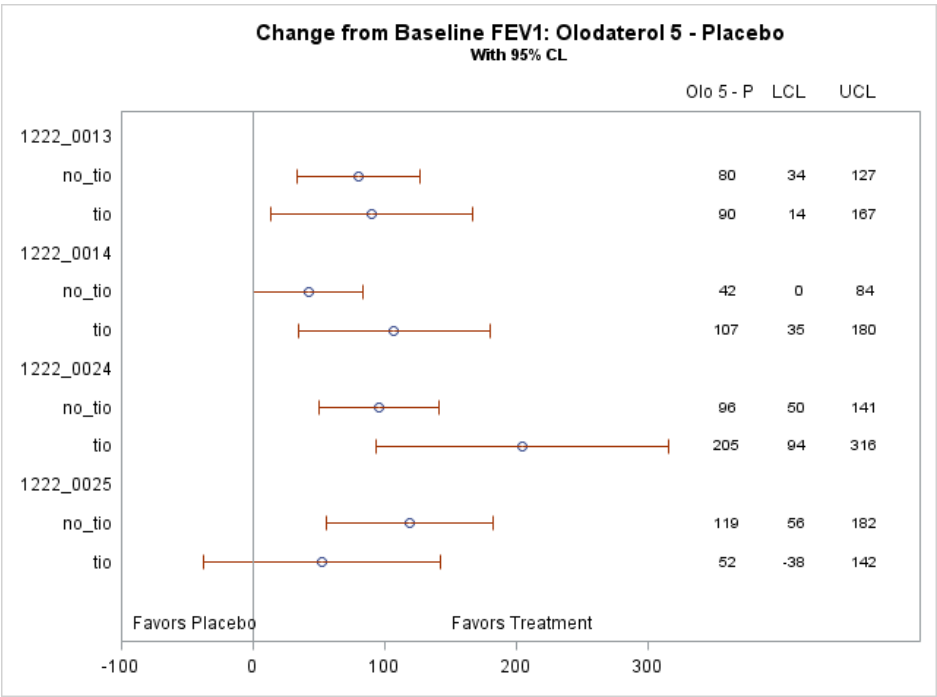


Figure 7. 'Confirmatory' analyses, change from baseline trough FEV_1 at Week 12 by tiotropium use



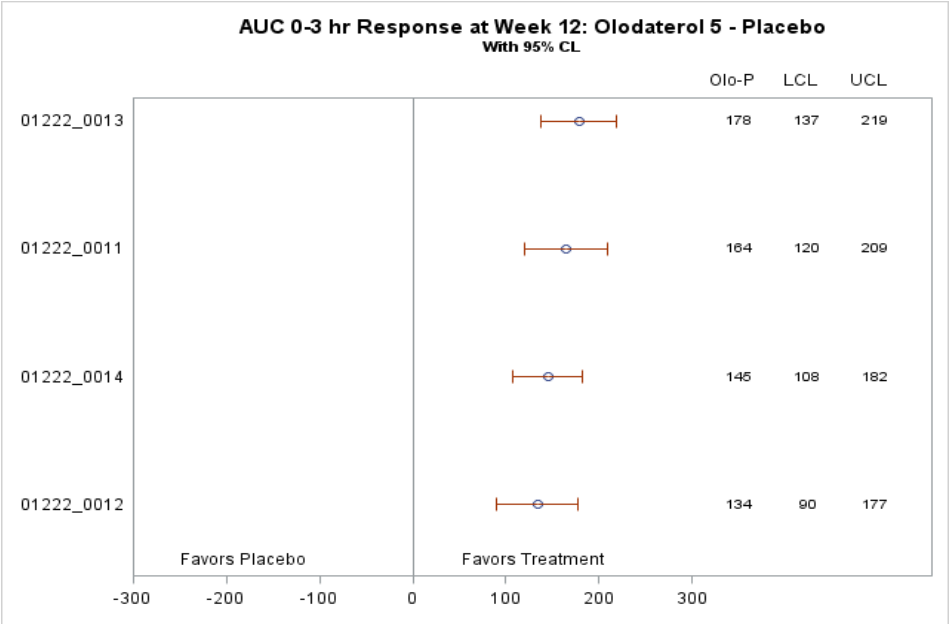
source: Forest Plots 2012 12 07.sas

Reference ID: 3247160

3.1.4.2 *Primary Endpoint: FEV_1 $AUC_{0\text{-}3hr}$ (Studies 11, 12, 13, and 14)*

Change from baseline FEV_1 $AUC_{0\text{-}3hr}$ at week 12, the pre-planned primary endpoints for studies 11 and 12, were consistently higher among patients administered olodaterol 5 than among patients administered placebo (Figure 8, see also underlying Table 17). These results for studies 13 and 14 are consistent with those provided by the applicant. However, those for studies 11 and 12 are not; they incorporate the pre-planned tiotropium by treatment interactions which, as discussed in section 3.1.4.1 above, were not provided by the applicant.

Figure 8. Change from baseline $AUC_{0\text{-}3hr}$ at Week 12 (Olodaterol 5 vs placebo)

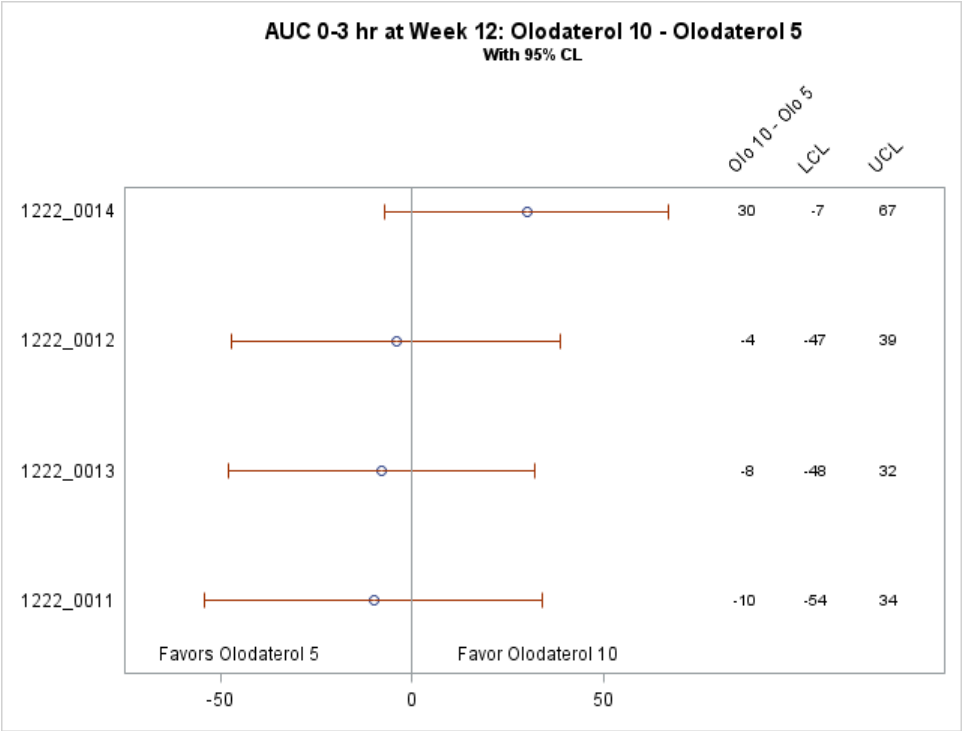


source: Forest Plots 2012 12 07.sas

Oladaterol 10 did not differ significantly from olodaterol 5 in changes from baseline trough FEV_1 and FEV_1 $AUC_{0\text{-}3hr}$ at week 12 (Figure 9 underlying Table 17). This is consistent with the applicant's results.

Reference ID: 3247160

Figure 9. Change from baseline FEV_1 $AUC_{0\text{-}3hr}$ at Week 12 (Olodaterol 10 vs olodaterol 5)



source: Forest Plots 2012 11 26.sas

The difference between olodaterol 5 and placebo for change from baseline FEV_1 $AUC_{0\text{-}3hr}$ at week 24 was statistically significant in studies 13 (151 mL average 95% CI 110−193) and 14 (average 129 mL 95% CI 91−167).

3.1.4.3 Secondary Endpoint: FEV_1 $AUC_{0\text{-}12hr}$ (Studies 11 and 12)

Change from baseline FEV_1 $AUC_{0\text{-}12hr}$ at week 12, the pre-planned key secondary endpoint for studies 11 and 12 was significantly higher among patients administered olodaterol 5 than among patients administered placebo (study 11: mean difference is 0.152 liters (L) 95% CI 0.075−0.230, study 12: mean difference is 0.092 L 95% CI 0.029−0.156, see Table 19 for further details). Olodaterol 10 did not differ significantly from olodaterol 5 for this endpoint (study 11: mean difference is -0.019 L 95% CI -0.087−0.049, study 12: mean difference is -0.015 L 95% CI -0.78−0.49). These results were not provided by the applicant; they incorporate the pre-planned tiotropium by treatment interactions discussed in section 3.1.4.1 above.

Reference ID: 3247160

3.1.4.4 24-hour Response Profile: FEV_1 $AUC_{12\text{-}24hr}$ and FEV_1 $AUC_{0\text{-}24hr}$ (Studies 24, 25, 39, and 40)

Changes from baseline FEV_1 $AUC_{0\text{-}24hr}$ and FEV_1 $AUC_{0\text{-}24hr}$ were not measured in the parallel arm studies 11, 12, 13, and 14, but were instead measured in the six week crossover trials 24, 25, 39, and 40. At the July 17, 2008 EOP2 meeting, the Agency informed the applicant that results for these variables from the crossover trials would be acceptable for on-label time profiles of drug action only if other spirometry measurements from the six week crossover trials were comparable to those from the longer term parallel arm trials.

However, the spirometry results suggest a larger effect of olodaterol after six weeks of treatment in the crossover trials 24, 25, 39, and 40 than after twelve weeks in the parallel arm trials 11, 12, 13, and 14, calling into question comparability of the six and twelve week endpoints. Six week change from baseline trough FEV_1, from crossover trials 24, 25, 39, and 40 ranged from 97 mL to 134 mL, higher than point estimates at twelve weeks from parallel arm trials 11, 12, 13, and 14, which ranged from 33 to 84 mL (Figure 10, underlying Table 9). For FEV_1 $AUC_{0\text{-}3hr}$, point estimates of change from baseline ranged from 164 to 214 mL in the six-week crossover trials compared to 134 to 178 mL after twelve weeks in the parallel arm trials (Figure 11, underlying Table 17) and, for FEV_1 $AUC_{0\text{-}12hr}$, change from baseline ranged from 92 to 152 mL after twelve weeks in the two parallel arm trials in which it was measured, compared to 148 to 197 mL after six weeks of treatment in the four crossover trials (Figure 12, underlying Table 19).

Further, within the parallel arm trials themselves, improvements from baseline tended to be less after twelve weeks than after six weeks of treatment for trough FEV_1 (Figure 13, underlying Table 11, Table 12, Table 13, Table 14) and FEV_1 $AUC_{0\text{-}3hr}$ (Figure 14, underlying Table 18). No similar comparison for FEV_1 $AUC_{0\text{-}12hr}$ was available because six week data for this variable was not collected in the parallel arm trials.

Reference ID: 3247160

Figure 10. Change from baseline trough FEV_1 (Crossover Studies at Week 6 and Parallel Arm Studies at Week 12)

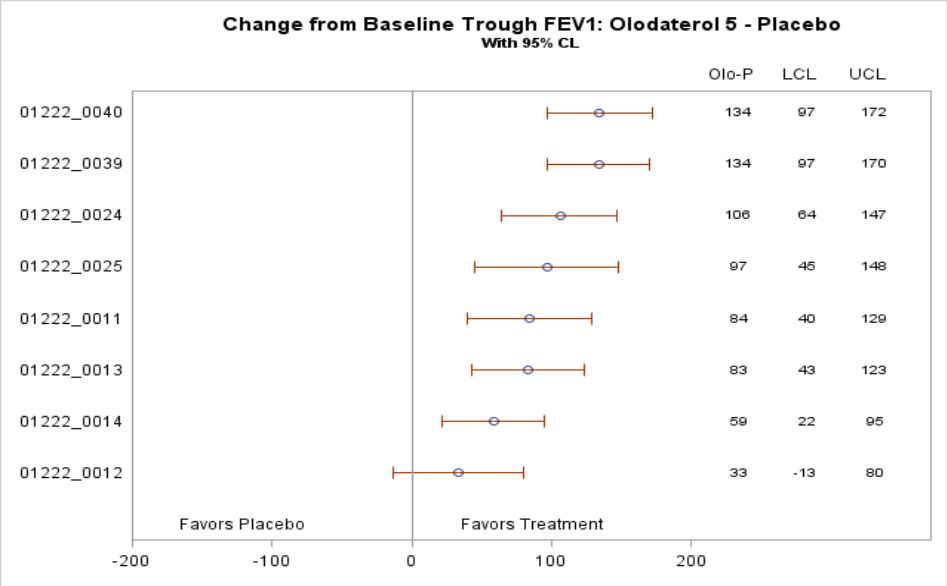
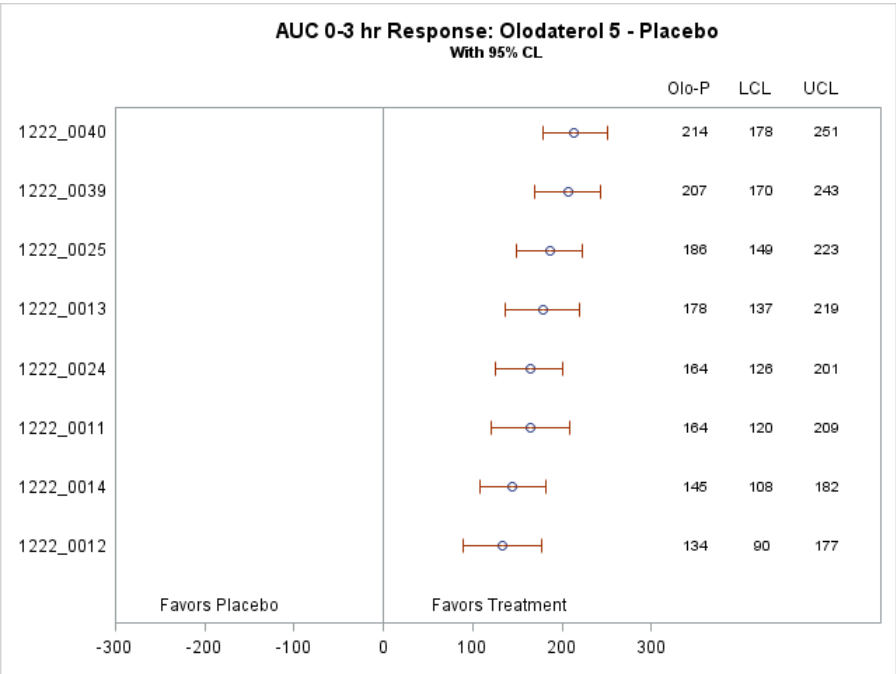


Figure 11. Change from baseline FEV_1 $AUC_{0\text{-}3hr}$ (Crossover Studies at Week 6 and Parallel Arm Studies at Week 12)



source: Forest Plots 2012 11 26.sas

Reference ID: 3247160

Figure 12. Change from baseline FEV_1 $AUC_{0\text{-}12hr}$ (Crossover Studies at Week 6 and Parallel Arm Studies at Week 12)

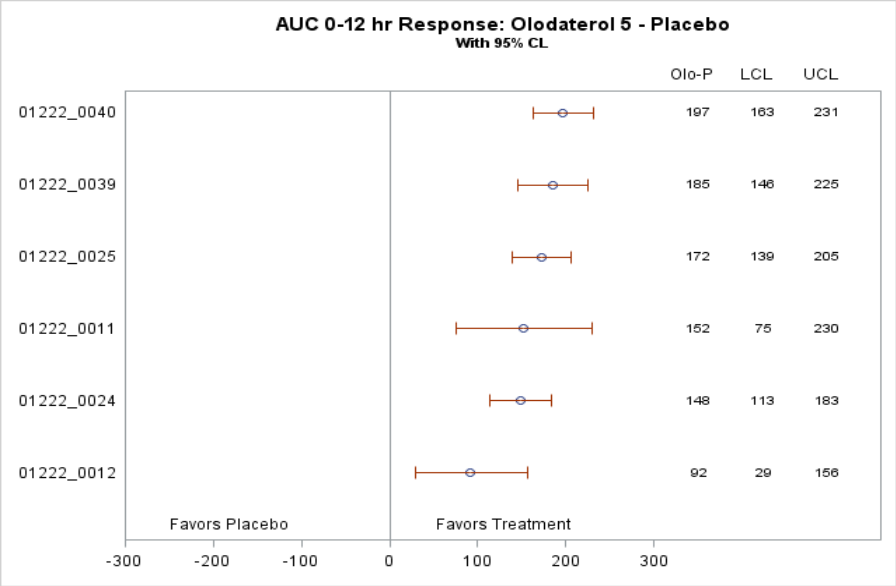
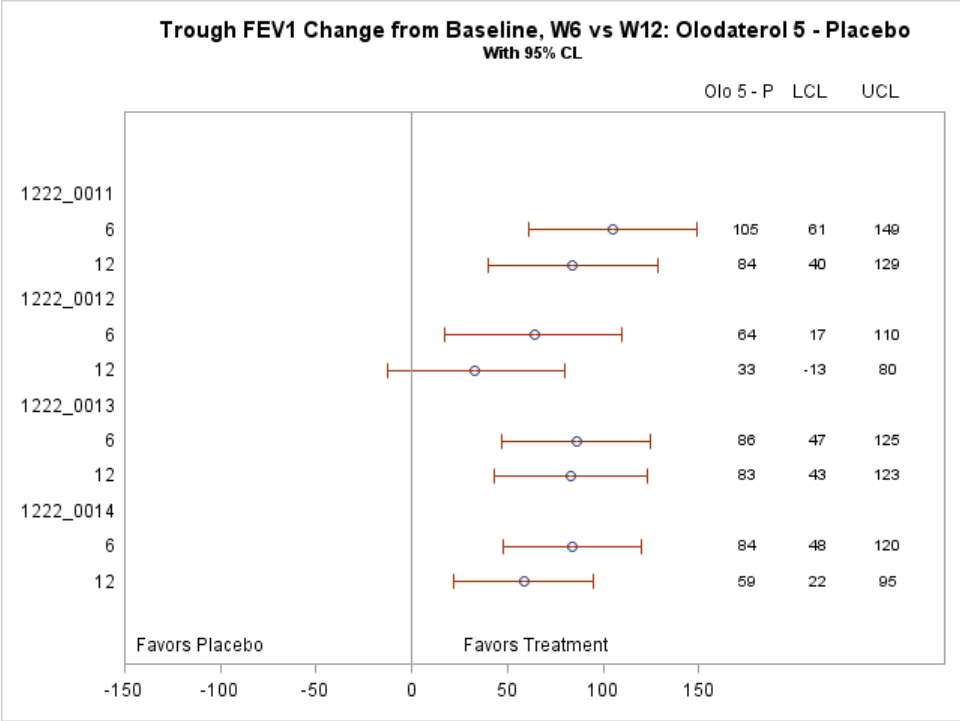


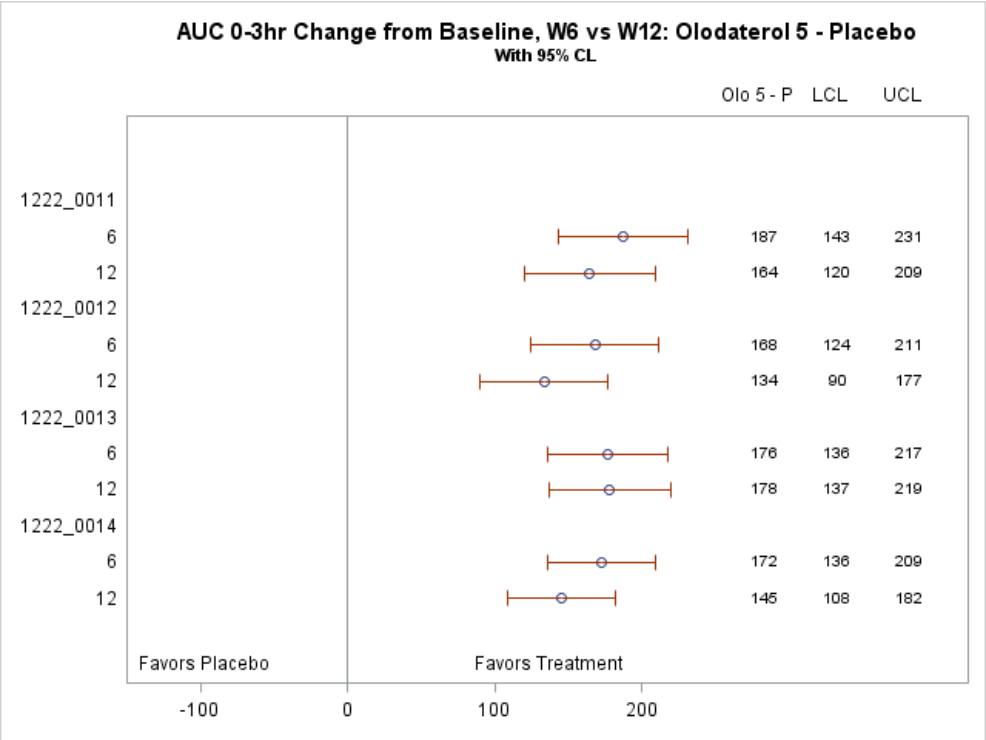
Figure 13. Change from baseline trough FEV_1 (Parallel Arm Trials at Weeks 6 and 12)



source: Forest Plots 2012 11 26.sas

Reference ID: 3247160

Figure 14. Change from baseline trough FEV_1 $AUC_{0\text{-}3hr}$ (Parallel Arm Trials at Weeks 6 and 12)



source: Forest Plots 2012 11 26.sas

With the caveat that effects after six weeks of treatment do not necessarily represent effects after twelve weeks of treatment, olodaterol 5 did have a significant effect after six weeks of treatment on FEV_1 $AUC_{12\text{-}24hr}$ (Figure 15, underlying Table 20) and FEV_1 $AUC_{0\text{-}24\ hr}$ (Figure 16, underlying Table 21). Of eight comparisons between between olodaterol 10 and olodaterol 5, (Figure 17 and Figure 18, underlying Table 20 and Table 21), there was only one nominally significant difference, for $AUC_{12\text{-}24hr}$ in study 39 (Figure 17).

Reference ID: 3247160

Figure 15. Change from baseline $AUC_{12\text{-}24hr}$ at Week 6 (Olodaterol 5 vs placebo)

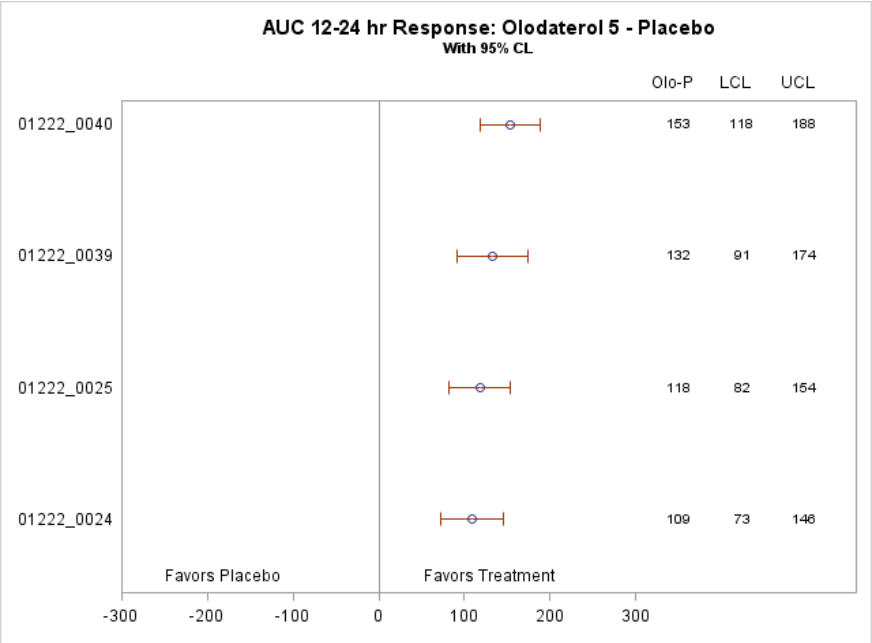
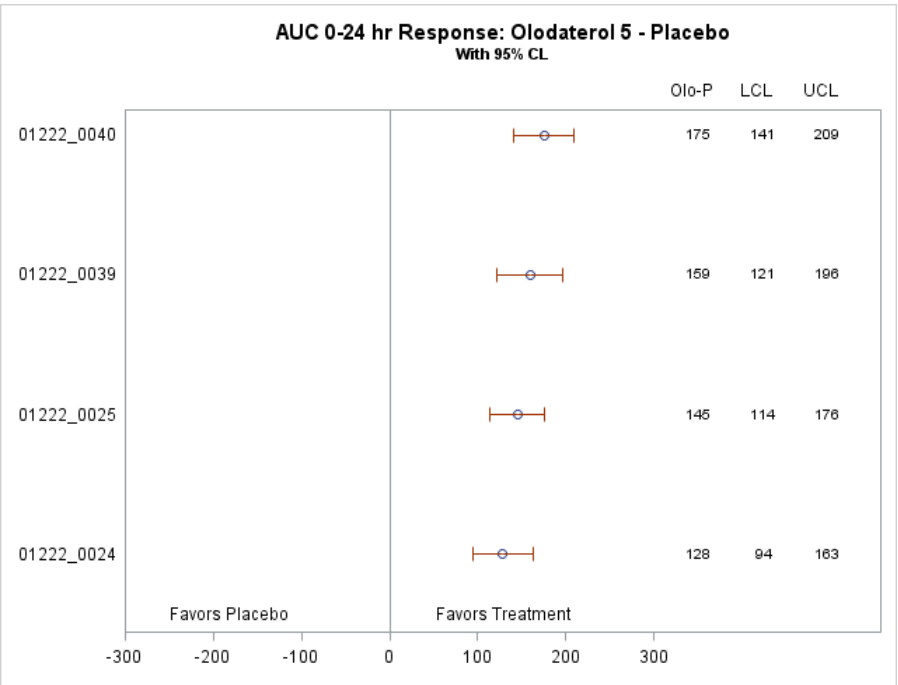


Figure 16. Change from baseline FEV_1 $AUC_{0\text{-}24hr}$ at Week 6 (Olodaterol 5 vs placebo)



source: Forest Plots 2012 11 02.sas

Reference ID: 3247160

Figure 17. Change from baseline FEV_1 $AUC_{12\text{-}24hr}$ at Week 6 (Olodaterol 10 vs olodaterol 5)

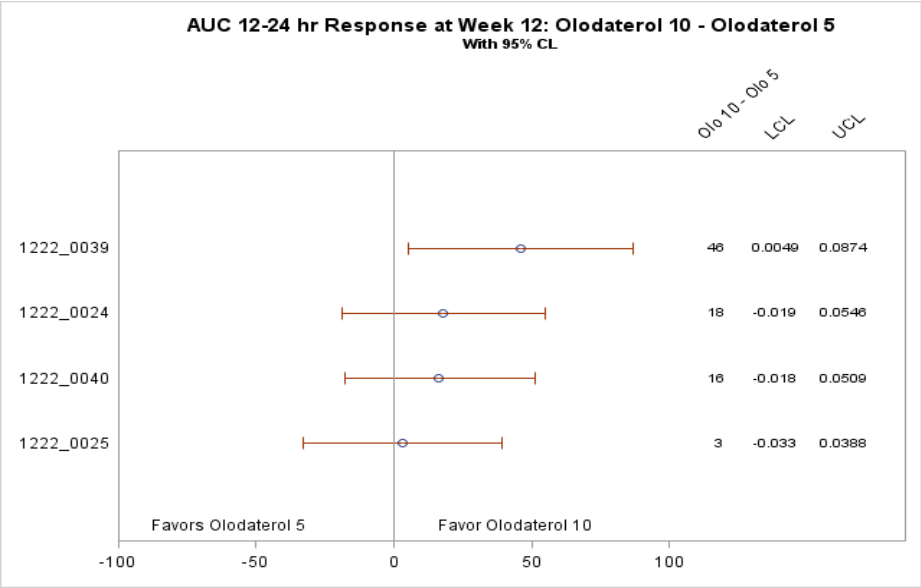
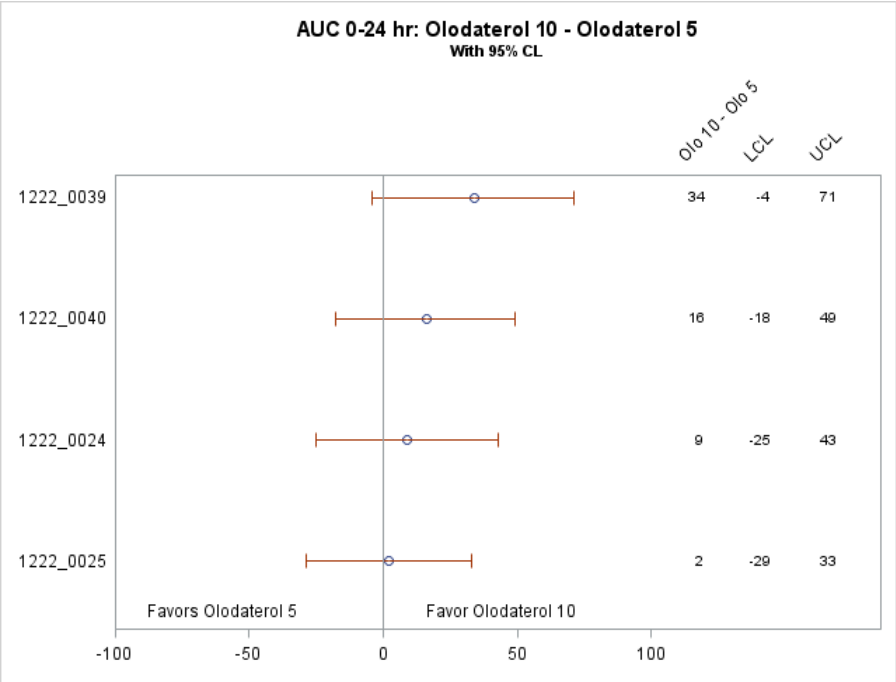


Figure 18. Change from baseline FEV1 $AUC_{0\text{-}24}$ hr at Week 6 (Olodaterol 10 vs olodaterol 5)

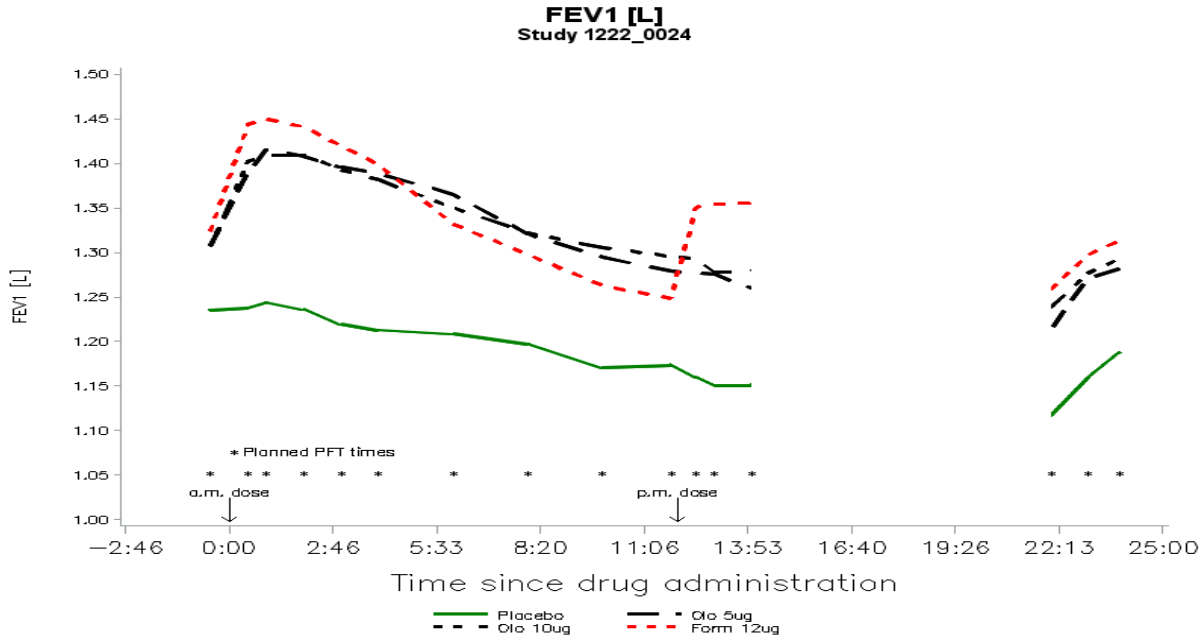


source: Forest Plots 2012 11 02.sas

Reference ID: 3247160

Modeled FEV_1 used for calculation of FEV_1 $AUC_{x\text{-}yhr}$ for each time period x to y hours after treatment administration are directly examined in Figure 19 and Figure 20, which show a spike in FEV_1 among patients treated with formoterol after administration of the second daily formoterol dose. Unlike formoterol, administration of olodaterol is only once per day, and so a spike in FEV_1 seen with administration of the second daily dose of formoterol does not occur among patients randomized to olodaterol. The diurnal pattern of effect on FEV_1 of olodaterol more closely matches that of tiotropium (Figure 21 and Figure 22) which, like olodaterol, is administered once per day.

Figure 19. Diurnal FEV_1 by treatment, Study 24



Reference ID: 3247160

Figure 20. Diurnal FEV_1 by treatment, Study 25

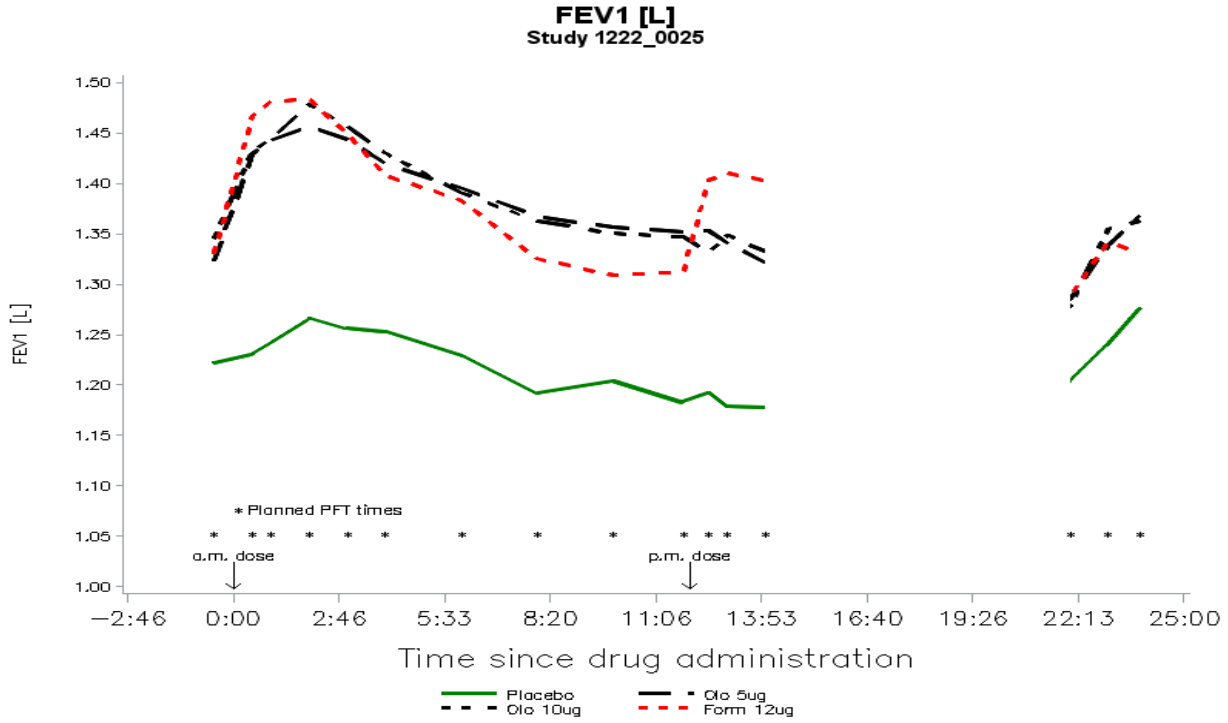
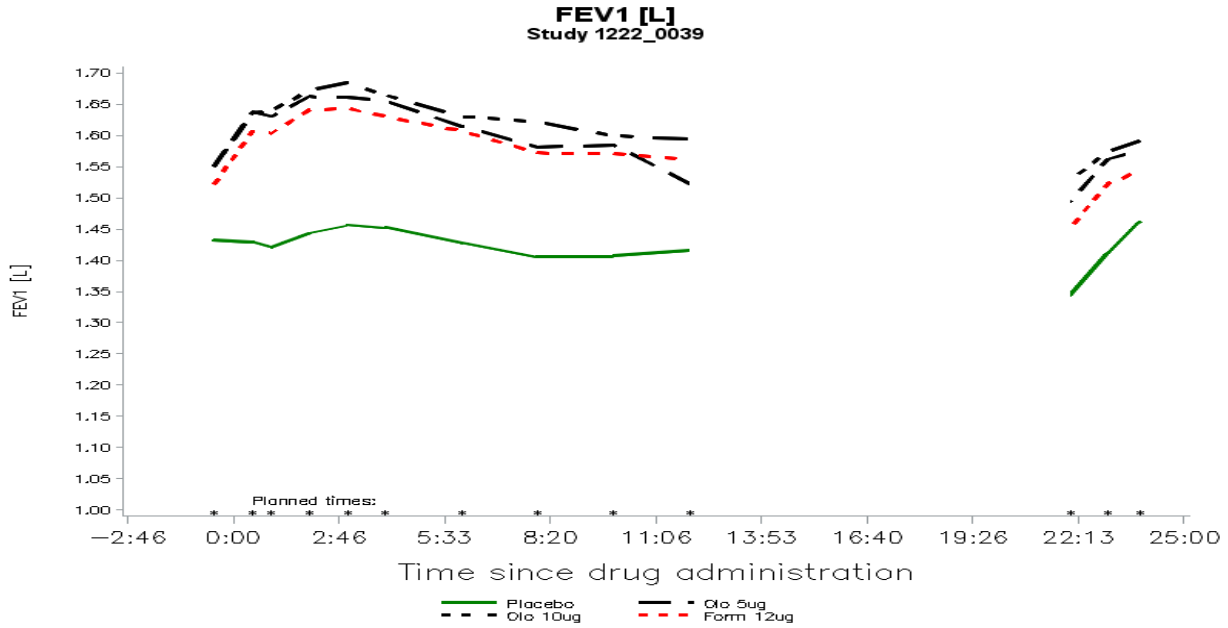


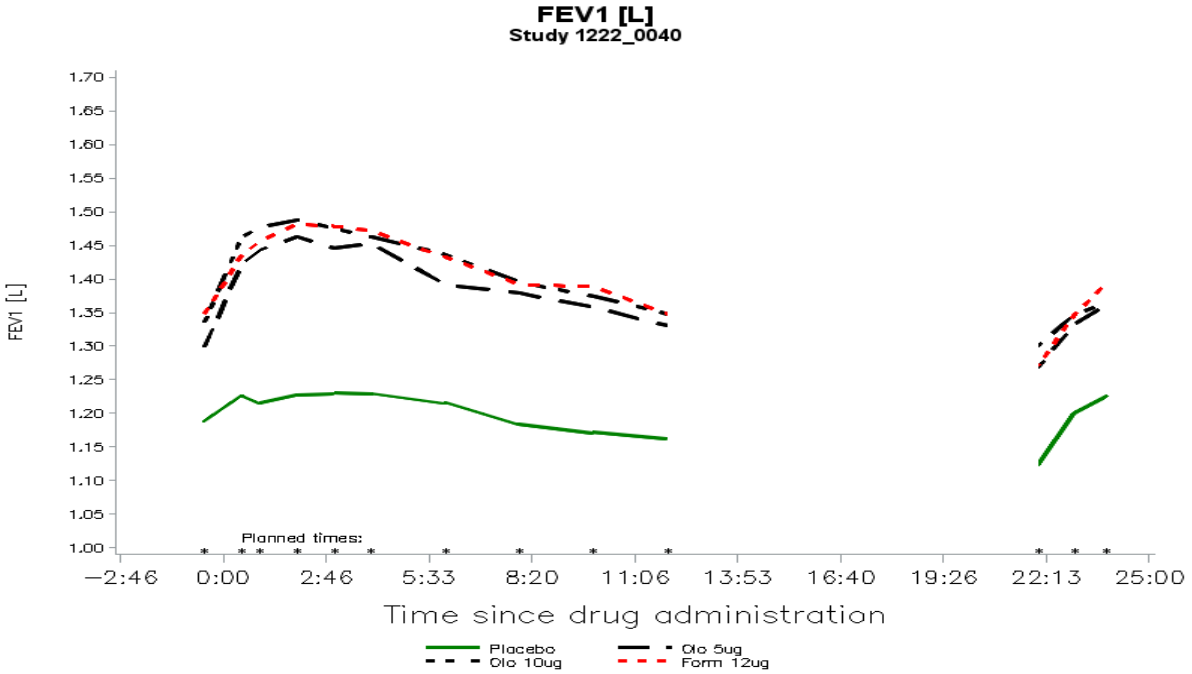
Figure 21. Diurnal FEV_1 by treatment, Study 39



source: FEV1 Time For 2012 08 14.sas

Reference ID: 3247160

Figure 22. Diurnal FEV_1 by treatment, Study 40



source: FEV1 Time For 2012 08 14.sas

Because Figure 19, Figure 20, Figure 21, and Figure 22 all show large gaps in the data during the second 12 hours of each study, quantitative estimates FEV_1 $AUC_{12\text{-}24hr}$ and FEV_1 $AUC_{0\text{-}24hr}$ should not be considered accurate or precise. For example, in studies 24 and 25, data for calculation of FEV_1 $AUC_{12\text{-}24hr}$ was collected 12 ½, 13, 14, 23, and 23 5/6 hours after administration of the morning dose, with a nine hour window during which no spirometric data was collected, between hours 14 and 23.

Despite the lack of quantitatively reliable estimates for FEV_1 $AUC_{12\text{-}24hr}$ and FEV_1 $AUC_{0\text{-}24hr}$, we do consider the effects of olodaterol seen in Figure 15 and Figure 16 non-spurious, i.e., statistically significant. First, the statistical analyses were preplanned, with significant p-values unlikely unless a olodaterol has a real effect on observed values of FEV_1. Second, to argue that, despite a real effect on observed values of FEV_1, olodaterol 5 has no effect on the overall value of FEV_1 $AUC_{12\text{-}24hr}$, one would have to argue that the lines between olodaterol 5 and placebo cross, i.e. that olodaterol 5 consistently becomes less effective than placebo during the data gaps in Figure 19, Figure 20, Figure 21, and Figure 22. However, according to the medical reviewer, such a sudden and temporary reversal of treatment effect is clinically implausible.

Reference ID: 3247160

In summary, compared to placebo, olodaterol 5 showed significant effects compared to placebo on trough FEV_1 and on FEV_1 from 0 to 14 hours after administration. Precise quantification of the effects of olodaterol on FEV_1 $AUC_{12\text{-}24hr}$ and FEV_1 $AUC_{0\text{-}24hr}$ were hampered by scarcity of data; no FEV_1 measurements were recorded from hours 14 to 23 in studies 24 and 25 or from hours 12 to 22 in studies 39 and 40. Further quantitative estimates for the effects of olodaterol 5 compared to placebo for FEV_1 $AUC_{12\text{-}24hr}$ and FEV_1 $AUC_{0\text{-}24hr}$ at week 12 were not provided; they were estimated only in the crossover studies at week 6, which should not be considered representative of quantitative estimates at week 12.

Beyond the 12 hour post-morning dose, exploratory analyses suggest that, compared to formoterol 18 mcg bid, olodaterol qd suffers from an attenuation of effect. During this time period, the effects of olodaterol are more similar those of tiotropium 12 mcg qd.

3.1.4.5 Interim Summary: Spirometry Results for FEV_1 (Studies 11, 12, 13, 14, 24, 25, 39, and 40)

This submission demonstrates statistically significant benefits of olodaterol 5 as an add-on to standard of care (without other LABA) for trough FEV_1, FEV_1 $AUC_{0\text{-}3hr}$, and FEV_1 $AUC_{0\text{-}12hr}$ at week 12 and for FEV $AUC_{12\text{-}24hr}$ and FEV_1 $AUC_{0\text{-}24hr}$ at week 6.

To quantify the effects of olodaterol on FEV $AUC_{12\text{-}24hr}$ and FEV_1 $AUC_{0\text{-}24hr}$ at week 12, measurements at week 6 were analyzed on the assumption that spirometric results at week 6 would match those at week 12. However, spirometric results at week 6 did not match those at week 12; therefore estimates of FEV $AUC_{12\text{-}24hr}$ and FEV_1 $AUC_{0\text{-}24hr}$ at week 12 could not be estimated from the present submission. Further, estimates of olodaterol's effect compared to placebo on FEV $AUC_{12\text{-}24hr}$ and FEV_1 $AUC_{0\text{-}24hr}$ at week 6 should not be considered quantitatively accurate because of gaps in the data from 14 to 22 hours after administration of treatment.

3.1.4.6 Other Primary Endpoint: Mahler TDI Focal Score (Studies 13 and 14)

Differences between olodaterol 5 and placebo for change from baseline Mahler TDI Focal Score at week 24 were not statistically significant (study 13: mean difference 0.19 95% CI -0.49−0.86, study 14: mean difference 0.40 95% CI -0.21−1.02, underlying Table 22 and Table 23, combined data from studies 13 and 14, mean difference 0.31 95% CI -0.14−0.77). Results from the individual studies match those provided by the applicant. However, the applicant did not provide the pre-planned statistical analysis for significance which combined data from studies 13 and 14.

Reference ID: 3247160

3.1.4.7 Key Secondary Endpoint: SGRQ (Studies 13 and 14)

Statistically significant differences between olodaterol 5 and placebo for SGRQ were seen in one of two studies (study 13: mean difference -2.44 95% CI -5.19−0.31, study 14: mean difference -3.15 95% CI -5.8− -0.5, underlying Table 23). The maximum reduction of SGRQ by olodaterol compared to placebo was 3.15.

3.1.4.8 Key Secondary Endpoint: Exacerbations (Studies 11, 12, 13, and 14)

Because the effects of olodaterol on Mahler TDI were not statistically significant, analyses of exacerbations should be considered only for purposes of hypothesis generation. Point estimates of hazard ratio for olodaterol 5 compared to placebo for moderate and hospitalized exacerbations ranged above and below unity (Figure 23 and Figure 24, underlying Table 28), showing no consistent effect of olodaterol on exacerbation rate, with none of the hazard ratios differing significantly from unity. The analyses used 48 weeks of data from each parallel arm study using a Cox regression model stratified by tiotropium use at baseline.

Integrated analysis for time to first COPD exacerbation from all four studies yielded no hazard ratios which indicated differences between olodaterol 5 and placebo (mean moderate 0.922 95% CI 0.766−1.109, mean hospitalized 0.910 95% CI 0.619−1.227, mean any 0.913 95% CI 0.773−1.080). These results match those provided in the applicant's submission.

Reference ID: 3247160

Figure 23. Moderate exacerbations (Olodaterol 5 vs placebo)

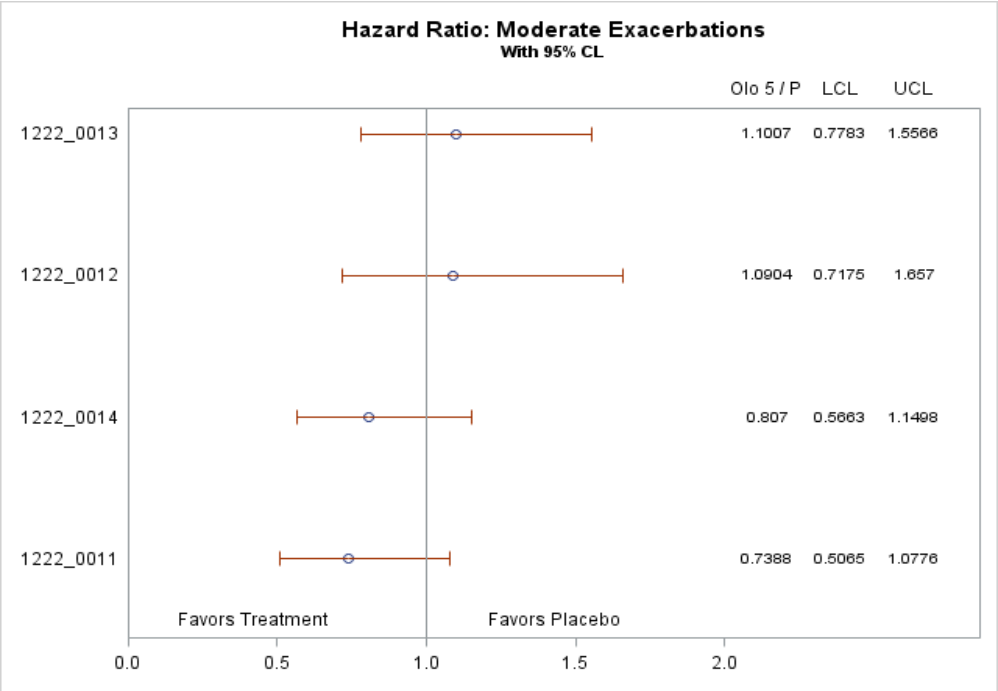
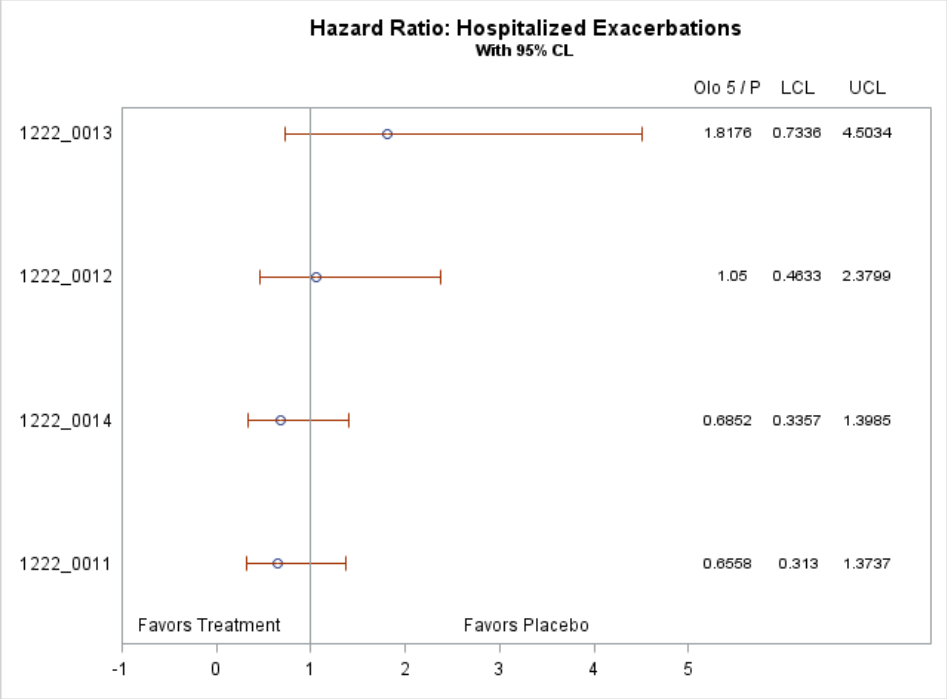


Figure 24. Exacerbations resulting in hospitalization (Olodaterol 5 vs placebo)



source: Forest Plots 2012 11 27.sas

Reference ID: 3247160

3.1.4.9 *Exercise Endurance (Studies 37 and 38)*

The $\log_{10}$ difference between olodaterol 5 and placebo for exercise endurance time differed significantly from 0, with point estimates suggesting that, two hours after administration, olodaterol increased exercise tolerance by 10% to 14% compared to placebo (study 37: mean difference 1.14 95% CI 1.07−1.22, study 38: mean difference 1.12 95% CI 1.04−1.20, underlying Table 24). The mean absolute difference between olodaterol 5 and placebo was 52 seconds in study 37 and 42 seconds in study 37. Analyses by the applicant showed no statistically significant differences between the two olodaterol doses.

3.1.4.10 *Inspiratory Capacity (IC) at Isotime (Studies 37 and 38)*

Olodaterol 5 increased inspiratory capacity at isotime compared to placebo (study 37: mean difference 182 mL 95% CI 112−252, study 38: mean difference 84 mL 95% CI 16−152, underlying Table 25). Increased capacity at isotime was supported by increases both pre- and post- exercise (Table 25).

Despite the fact that pre-exercise IC was not included in the analysis hierarchy to control type 1 error, the applicant proposed to include claims on the label for increased IC at rest.

3.1.4.11 *Breathing Discomfort: Borg Category Ratio Scale (Studies 37 and 38)*

Compared to placebo, olodaterol 5 had statistically significant effects on the Borg Category Ratio Scale at exercise isotime for breathing discomfort in study 37 but not in study 38 (study 37: mean difference -0.77 95% CI -1.21− -0.33, study 38: mean difference -0.34 95% CI -0.76− 0.09, underlying Table 26).

3.1.4.12 *Functional Residual Capacity (Studies 37 and 38)*

Analyses for FRC provided below, including nominal[5] p-values, should be considered exploratory rather than confirmatory, because this endpoint was not included in the analysis hierarchy to control type 1 error among confirmatory endpoints and was planned only as an 'other' secondary endpoint. Even if FRC were included in the hierarchy, it would presumably have been below the Borg Scale at isotime (a key secondary variable high in the analysis hierarchy), for which no statistically significant difference between olodaterol and placebo was seen in study 38.

[5] 'Nominal' confidence intervals, p-values, and significance will be used in this review to denote calculations made without adequate control of Type I error to account for multiple outcomes or analyses.

Reference ID: 3247160

Compared to placebo, nominally statistically significant FRC differences associated with olodaterol 5 were not observed pre-dose (study 37: mean difference -122 mL 95% CI -258−14, study 38: mean difference -86 mL 95% CI -190−18, underlying Table 27) but were nominally observed post-dose (study 37: mean difference -210 95% CI -339− -0.81, study 38: mean difference -213 mL 95% CI -318− -108, underlying Table 27).

The applicant also claimed on the label that improvements in FRC demonstrated that olodaterol 5 ‘reduced lung hyperinflation.’

Reference ID: 3247160

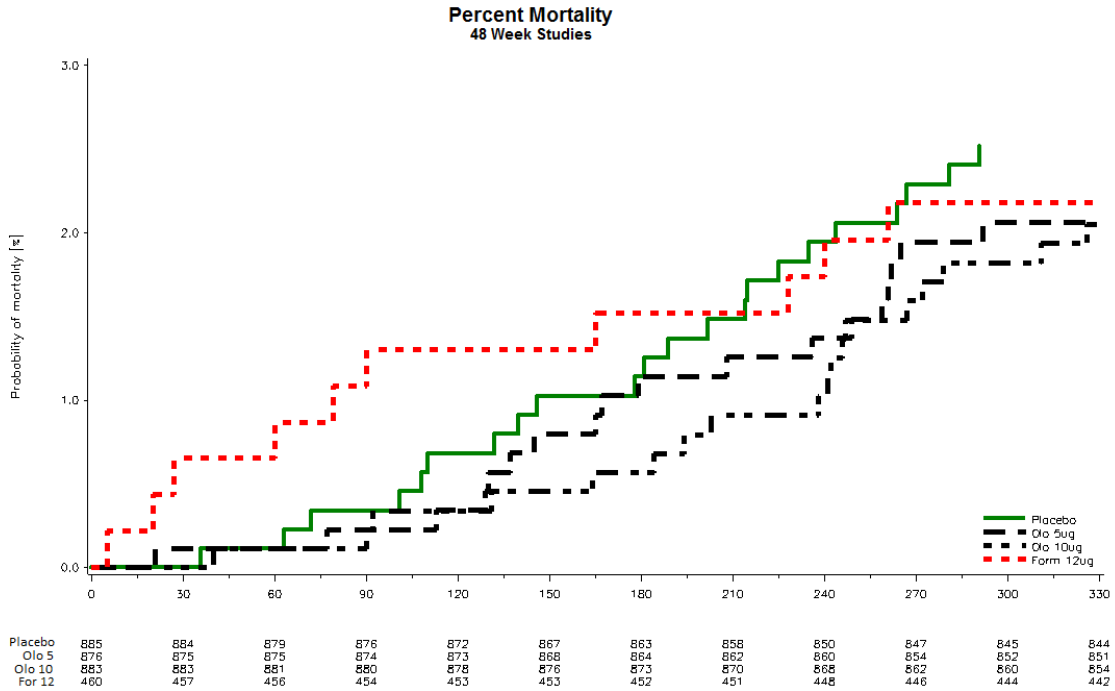
4 Safety

Safety evaluations for this submission will be evaluated by the Medical Reviewer, Robert Lim, M.D. However, because mortality is central to safety, a brief treatment is provided below.

4.1 Patient Mortality (Studies 11, 12, 13, and 14)

Analysis of the 48 week parallel arm trials showed Kaplan-Meier percent cumulative mortality numerically lower among patients randomized to olodaterol 5 and olodaterol 10 than among patients randomized to formoterol or placebo. Cumulative percent mortality of the olodaterol treatments converged with that of formoterol at approximately day 360. At day 240, where difference between olodaterol 10 and placebo is larger than at other times, a log rank test showed no significant difference (p=0.843) in cumulative mortality between the placebo and olodaterol 10 treatment arms.

Figure 25. Kaplan-Meier percent mortality, to 330 days post-treatment



source: mortality.sas

Reference ID: 3247160

5 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

5.1 Gender, Race, Age, and Geographic Region

Subgroup analyses examining the effects on efficacy of gender, age (40-50, 51-65, or >65), race, or geographic region (USA vs non-USA) were conducted on data integrated from similarly designed studies 11/12, 13/14, 24/25, and 39/40.

The analyses were conducted using the same statistical models used for the primary analyses, with the additional inclusion of subgroup and subgroup by treatment interaction terms. Analyses for the parallel arm studies included an additional interaction term for subgroup by treatment by time. Interactions between subgroup and treatment were tested at the 0.05 level of significance, with no adjustment to significance level made to account for testing of multiple endpoints.

For $AUC_{0\text{-}3hr}$, no significant interactions were seen between treatment and gender, race, or geographic region. However, the interaction between treatment and age was nominally significant for studies 13/14 (p=0.0162) and 24/25 (p=.0040), but not for studies 11/12 (p=0.2796) or 39/40 (p=0.4371).

Increased age (40-50 vs >50) consistently decreased the effect of olodaterol 5 on FEV_1 $AUC_{0\text{-}3hr}$ at week 12 in the parallel arm studies (Figure 26, underlying Table 29) and week 6 in the crossover studies (Figure 27, underlying Table 29). Increased age was also associated with decreased effects of olodaterol 5 on FEV_1 $AUC_{0\text{-}12hr}$ at week 12 (Figure 26, underlying Table 30) and week 6 (Figure 27, underlying Table 30). The association between increased age with decreased effects of olodaterol 5 was not completely consistent for change from baseline trough FEV_1 (Figure 30 and Figure 31 underlying Table 31), with the effect of of olodaterol 5 on trough FEV_1 numerically greater among older patients in study 25 (Figure 31, underlying Table 31) . None of the differences between age classes within studies were nominally statistically significant when tested at α=0.05.

Reference ID: 3247160

Figure 26. Effect of age on change from baseline FEV_1 $AUC_{0\text{-}3hr}$ at Week 12 (Olodaterol 5 vs placebo) in Studies 11, 12, 13 and 14

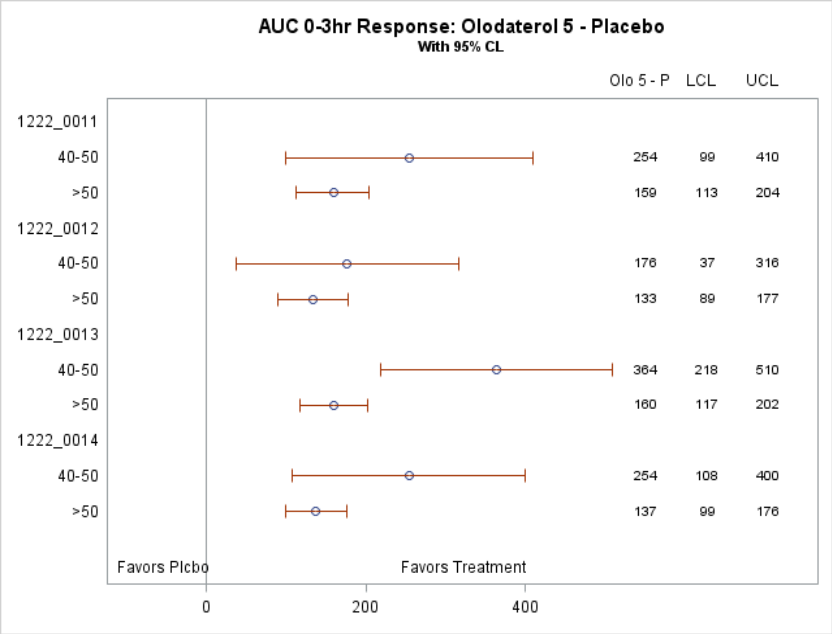
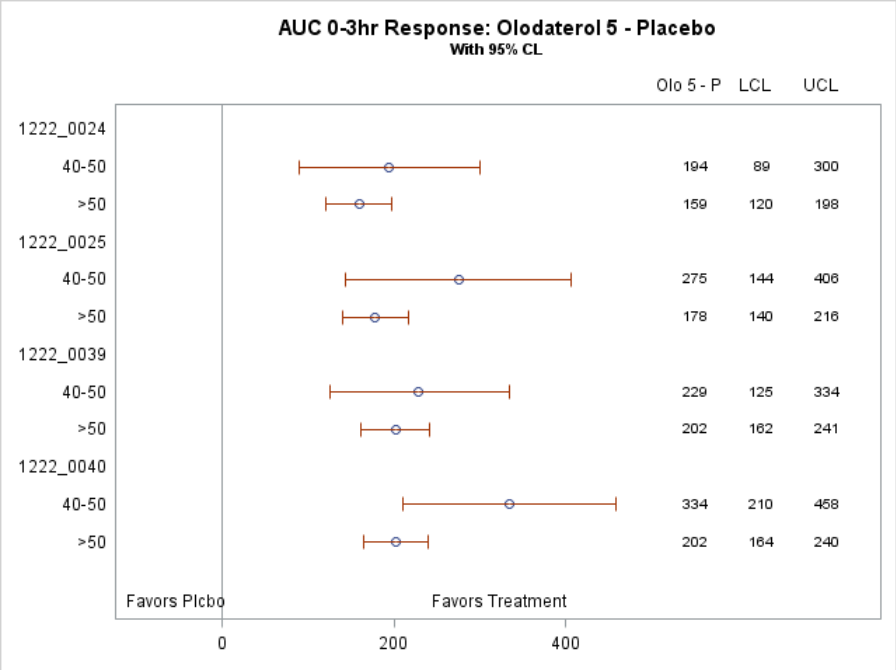


Figure 27. Effect of age on change from baseline FEV_1 $AUC_{0\text{-}3hr}$ at Week 6 (Olodaterol 5 vs placebo) in Studies 24, 25, 39, and 40



source: Forest Plots 2012 12 01.sas

Reference ID: 3247160

Figure 28. Effect of age on change from baseline FEV_1 $AUC_{0\text{-}12hr}$ at Week 12 (Olodaterol 5 vs placebo) in Studies 11 and 12

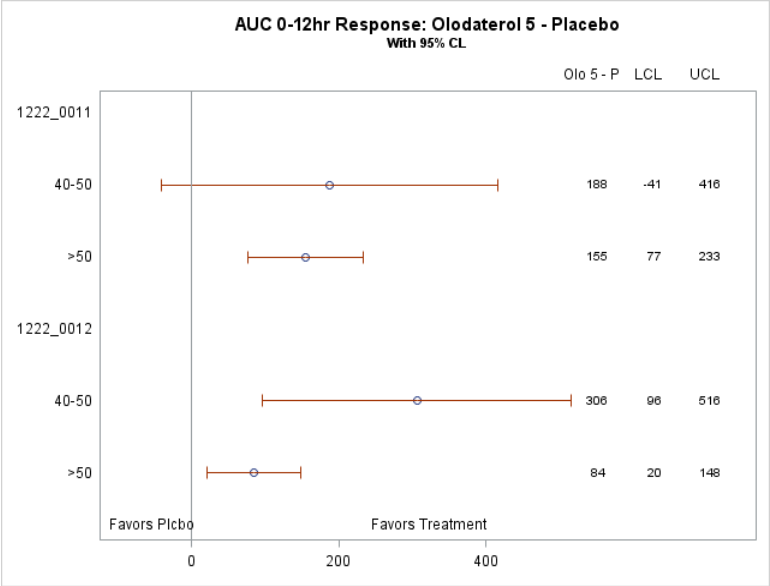
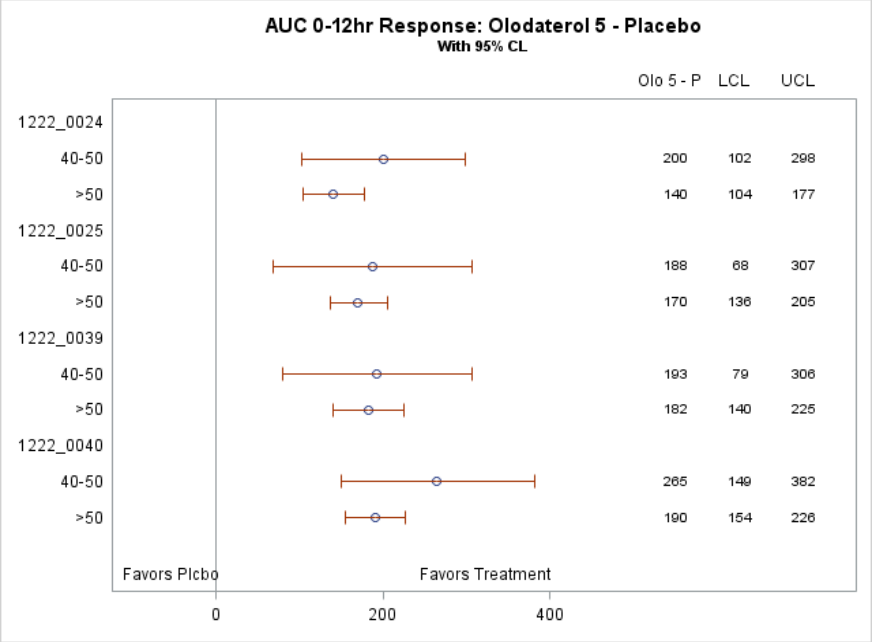


Figure 29. Effect of age on change from baseline FEV_1 $AUC_{0\text{-}12hr}$ at Week 6 (Olodaterol 5 vs placebo) in Studies 24, 25, 39, and 40



source: Forest Plots 2012 12 01.sas

Reference ID: 3247160

Figure 30. Effect of age on change from baseline trough FEV_1 at Week 12 (Olodaterol 5 vs placebo) in Studies 11, 12, 13 and 14

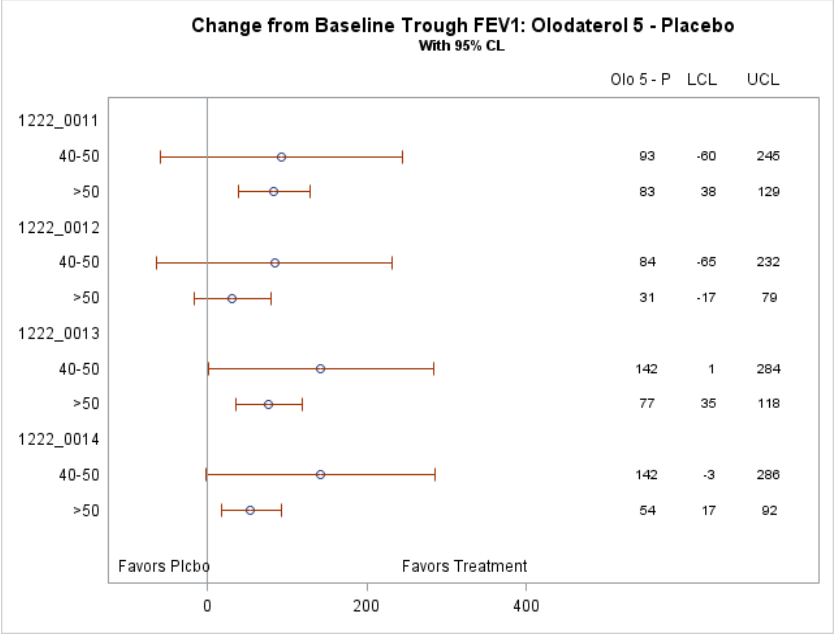
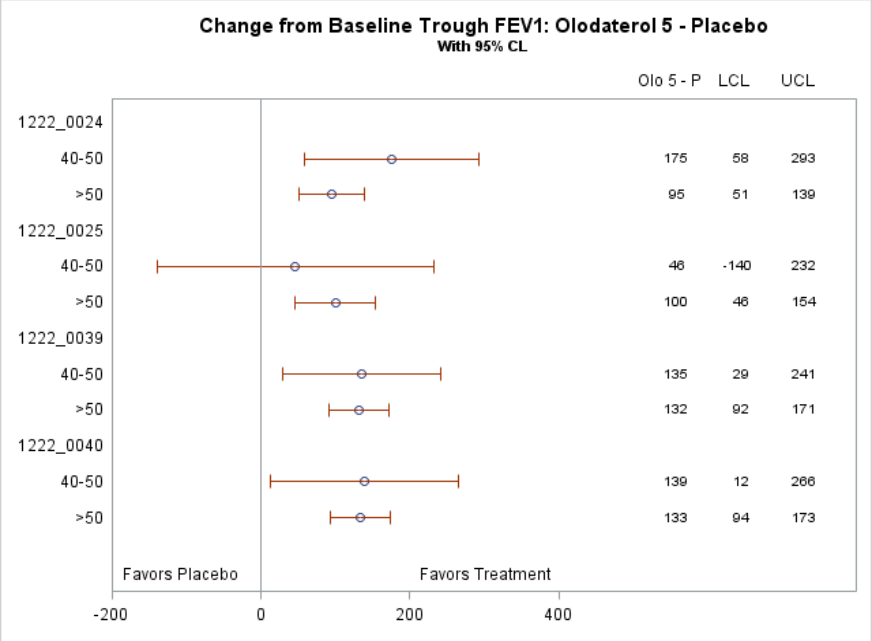


Figure 31. Effect of age on change from baseline trough FEV_1 at Week 6 (Olodaterol 5 vs placebo) in Studies 24, 25, 39, and 40



source: Forest Plots 2012 12 01.sas

Reference ID: 3247160

6 SUMMARY AND CONCLUSIONS

6.1 Statistical issues

For studies 11 and 12, the applicant attempted to replace pre-planned analyses with post-hoc analyses which inflated the estimated benefit of olodaterol, based on a presumptive reduction of olodaterol's effect among patients administered tiotropium as standard of care. However, by using the applicant's post-hoc analyses as hypothesis generating models, additional data available in the submission was used in this review to show that the presumed reduction of olodaterol's effect among patients administered tiotropium as standard of care is not correct. In addition, even if the applicant's underlying hypothesis had proven correct, the applicant's proposed replacement of the pre-planned analysis with a post-hoc analyses would have failed in control of Type I error.

To estimate the effects of olodaterol on FEV_1 $AUC_{12\text{-}24hr}$ and FEV_1 $AUC_{0\text{-}24hr}$ at week 12, measurements at week 6 were analyzed, on the assumption that spirometric results at week 6 would match those at week 12. However, spirometric results at week 6 do not match those at week 12; quantitatively accurate estimates of FEV_1 $AUC_{12\text{-}24hr}$ and FEV_1 $AUC_{0\text{-}24hr}$ at week 12 therefore could not be estimated from the present submission.

6.2 Collective evidence

This submission demonstrates statistically significant benefits of olodaterol 5 as an add-on to standard of care (without other LABA) for week 12 trough FEV_1, FEV_1 $AUC_{0\text{-}3hr}$, and FEV_1 $AUC_{0\text{-}12hr}$, as well as week 6 FEV_1 $AUC_{0\text{-}12hr}$,FEV_1 $AUC_{0\text{-}24hr}$, IC at isotime, and exercise endurance and SGRQ. The submission fails to clearly demonstrate statistically significant benefits to Mahler TDI focal score, pre-dose FRC or reduction of exacerbations. Differences between olodaterol 5 and placebo are only nominally significant for post-dose FRC.

6.3 Conclusions and Recommendations

The submission demonstrated benefits of olodaterol to pulmonary function. Four randomized parallel arm trials showed that olodaterol 5 as an add-on to standard of care without other LABA provided statistically significant benefits to the primary endpoints week 12 change from baseline trough FEV_1 (65 mL average) and $FEV_1AUC_{0\text{-}3hr}$ (155 mL average). Statistically significant benefits to the secondary endpoint week 12 change from baseline FEV_1 $AUC_{0\text{-}12hr}$ (122 mL average) were also demonstrated, in two randomized parallel arm studies. There was no statistically significant difference in benefit between olodaterol 5 and olodaterol 10. Statistically significant improvements in FEV_1 $AUC_{12\text{-}24hr}$ and FEV_1 $AUC_{0\text{-}24hr}$ at week 6 were demonstrated

Reference ID: 3247160

in four crossover trials. Due to gaps in the data, quantitative estimates of improvement were not available for these variables.

Statistically significant difference in SGRQ between olodaterol 5 and placebo was seen in only one of two parallel arm studies. The maximum mean reduction compared to placebo in a single study was 3.15, which is less than the 4.0 threshold considered clinically significant. The submission did not demonstrate statistically significant effects of olodaterol 5 on COPD exacerbation rate or Mahler TDI score.

Statistically significant benefits of olodaterol 5 were seen in two crossover studies for week 6 IC at isotime (130 mL average) and exercise endurance (47 seconds average). Hypothesis generating analyses on FRC suggested that, compared to placebo, olodaterol 5 improves post-dose (-212 mL average), but not pre-dose FRC.

Finally, there was an indication that the benefit of olodaterol 5 to change from baseline trough FEV_1, FEV_1 $AUC_{0\text{-}3hr}$, and change from baseline FEV_1 $AUC_{0\text{-}12hr}$ is greater among younger (≤50 yr) than among older (>50 yr) patients.

6.4 Labeling Recommendations

Labelling recommendations will be deferred until after the Pulmonary Allergy Drug Advisory Committee meets on January 29, 2013 to discuss this submisstion.

Reference ID: 3247160

Appendix: Supplemental Tables

Table 6. Patient demographics, number and percent. Gender, age, and smoking status

Study	Trt	N	Male	Age			Current Smoker
				< 65	65-74	≥ 75	
11	Olo 10	207	155 (75)	95 (46)	85 (41)	27 (13)	80 (39)
	Olo 5	208	150 (72)	105 (51)	84 (40)	19 (9)	75 (36)
	Placebo	209	152 (73)	85 (41)	93 (45)	31 (15)	86 (41)
12	Olo 10	217	152 (70)	95 (44)	81 (37)	41 (19)	92 (42)
	Olo 5	209	152 (73)	100 (48)	81 (39)	28 (13)	98 (47)
	Placebo	216	152 (70)	106 (49)	90 (42)	20 (9)	91 (42)
13	Olo 10	225	170 (76)	125 (56)	82 (36)	18 (8)	78 (35)
	Olo 5	227	177 (78)	110 (49)	88 (39)	29 (13)	68 (30)
	Placebo	225	180 (80)	112 (50)	90 (40)	23 (10)	87 (39)
	For 12	227	179 (79)	111 (49)	86 (38)	30 (13)	83 (37)
14	Olo 10	234	184 (79)	118 (50)	94 (40)	22 (9)	76 (33)
	Olo 5	232	187 (81)	122 (53)	89 (38)	21 (9)	88 (38)
	Placebo	235	195 (83)	121 (52)	94 (40)	20 (9)	72 (31)
	For 12	233	192 (82)	107 (46)	95 (41)	31 (13)	72 (31)
24	Olo 10	92	48 (52)	54 (59)	32 (35)	6 (7)	55 (60)
	Olo 5	95	48 (51)	57 (60)	32 (34)	6 (6)	57 (60)
	Placebo	96	51 (53)	57 (59)	32 (33)	7 (7)	57 (59)
	For 12	93	48 (52)	56 (60)	31 (33)	6 (7)	57 (61)
25	Olo 10	95	51 (54)	53 (56)	32 (34)	10 (11)	41 (43)
	Olo 5	93	49 (53)	53 (57)	32 (34)	8 (9)	40 (43)
	Placebo	94	49 (52)	53 (56)	32 (34)	9 (10)	40 (43)
	For 12	93	50 (54)	52 (56)	31 (33)	10 (11)	41 (44)
37	Olo 10	143	112 (78)	94 (66)	49 (34)	0 (0)	60 (42)
	Olo 5	147	113 (77)	98 (67)	49 (33)	0 (0)	64 (44)
	Placebo	143	112 (78)	95 (66)	48 (34)	0 (0)	61 (43)
38	Olo 10	147	109 (74)	101 (69)	45 (31)	1 (1)	57 (39)
	Olo 5	150	112 (75)	104 (69)	45 (30)	1 (1)	60 (40)
	Placebo	149	111 (75)	100 (67)	48 (32)	1 (1)	60 (40)
39	Olo 10	101	77 (76)	66 (65)	24 (24)	11 (11)	44 (44)
	Olo 5	101	78 (77)	66 (65)	24 (24)	11 (11)	43 (43)
	Placebo	102	79 (78)	66 (65)	25 (25)	11 (11)	44 (43)
	Tio 18	101	79 (78)	66 (65)	25 (25)	10 (10)	42 (42)
40	Olo 10	113	69 (61)	65 (58)	43 (38)	5 (4)	62 (55)
	Olo 5	116	71 (61)	64 (55)	47 (41)	5 (4)	62 (53)
	Placebo	110	68 (61)	63 (57)	43 (39)	4 (4)	60 (55)
	Tio 18	113	69 (61)	63 (56)	46 (41)	4 (4)	62 (55)

source: demographics 2012 08 30.sas

Reference ID: 3247160

Table 7. Patient demographics, number and percent. Race, country and tiotropium use

Study	Trt	N	Race				USA	Tio*
			white	black	asian	other	Y	
11	Olo 10	207	131 (63)	4 (2)	71 (34)	1 (1)	85 (41)	51 (25)
	Olo 5	208	130 (63)	3 (1)	72 (35)	3 (1)	86 (41)	47 (23)
	Placebo	209	130 (62)	5 (2)	70 (34)	4 (2)	86 (41)	49 (23)
12	Olo 10	217	136 (63)	9 (4)	71 (33)	1 (1)	115 (53)	44 (20)
	Olo 5	209	134 (64)	7 (3)	68 (33)	0 (0)	111 (53)	39 (19)
	Placebo	216	139 (64)	3 (1)	73 (34)	1 (1)	112 (52)	41 (19)
13	Olo 10	225	159 (71)	0 (0)	66 (29)	0 (0)	0 (0)	58 (26)
	Olo 5	227	160 (71)	3 (1)	63 (28)	1 (0)	0 (0)	60 (26)
	Placebo	225	159 (71)	2 (1)	64 (28)	0 (0)	0 (0)	57 (25)
	For 12	227	163 (72)	1 (0)	62 (27)	1 (0)	0 (0)	58 (26)
14	Olo 10	234	158 (68)	0 (0)	76 (33)	0 (0)	0 (0)	59 (25)
	Olo 5	232	153 (66)	0 (0)	78 (34)	1 (0)	0 (0)	59 (25)
	Placebo	235	156 (66)	1 (0)	78 (33)	0 (0)	0 (0)	61 (26)
	For 12	233	154 (66)	1 (0)	78 (34)	0 (0)	0 (0)	58 (25)
24	Olo 10	92	86 (94)	6 (7)	0 (0)	0 (0)	92 (100)	14 (16)
	Olo 5	95	89 (94)	6 (6)	0 (0)	0 (0)	95 (100)	15 (16)
	Placebo	96	90 (94)	6 (6)	0 (0)	0 (0)	96 (100)	15 (16)
	For 12	93	87 (94)	6 (7)	0 (0)	0 (0)	93 (100)	14 (16)
25	Olo 10	95	87 (92)	8 (8)	0 (0)	0 (0)	95 (100)	32 (34)
	Olo 5	93	84 (90)	8 (9)	0 (0)	1 (1)	93 (100)	31 (33)
	Placebo	94	86 (92)	8 (9)	0 (0)	0 (0)	94 (100)	31 (33)
	For 12	93	84 (90)	8 (9)	0 (0)	1 (1)	93 (100)	32 (34)
37	Olo 10	143	143 (100)	0 (0)	0 (0)	0 (0)	0 (0)	83 (58)
	Olo 5	147	147 (100)	0 (0)	0 (0)	0 (0)	0 (0)	84 (57)
	Placebo	143	143 (100)	0 (0)	0 (0)	0 (0)	0 (0)	80 (56)
38	Olo 10	147	145 (99)	0 (0)	0 (0)	2 (1)	0 (0)	55 (38)
	Olo 5	150	148 (99)	0 (0)	0 (0)	2 (1)	0 (0)	56 (38)
	Placebo	149	147 (99)	0 (0)	0 (0)	2 (1)	0 (0)	55 (37)
39	Olo 10	101	101 (100)	0 (0)	0 (0)	0 (0)	0 (0)	41 (41)
	Olo 5	101	101 (100)	0 (0)	0 (0)	0 (0)	0 (0)	41 (41)
	Placebo	102	102 (100)	0 (0)	0 (0)	0 (0)	0 (0)	39 (39)
	Tio 18	101	101 (100)	0 (0)	0 (0)	0 (0)	0 (0)	39 (39)
40	Olo 10	113	112 (99)	1 (1)	0 (0)	0 (0)	35 (31)	44 (40)
	Olo 5	116	115 (99)	1 (1)	0 (0)	0 (0)	36 (31)	48 (43)
	Placebo	110	109 (99)	1 (1)	0 (0)	0 (0)	34 (31)	44 (41)
	Tio 18	113	112 (99)	1 (1)	0 (0)	0 (0)	34 (30)	48 (44)

*Tio: Number and percent patients taking long acting anticholinergic at initiation of treatment (studies 11, 12, 13, 14) or at time of informed consent (studies 24, 25, 37, 38, 29, 40)

source: demographics 2012 08 30.sas

Reference ID: 3247160

Table 8. Patient demographics, mean and standard deviation. Age, FEV_1, FEV_1 % predicted, and smoking pack years

Study	Trt	N	Age	FEV_1	FEV_1 % pred	Pack Years
11	Olo 10	207	65 (8.2)	1.2 (0.5)	43.4 (15.4)	49.1 (37.5)
	Olo 5	208	64 (8.6)	1.1 (0.4)	41.3 (14.4)	49 (31.9)
	Placebo	209	65.8 (8.5)	1.2 (0.5)	43 (14.4)	48.4 (26.5)
12	Olo 10	217	65.4 (9.7)	1.2 (0.5)	43 (15.2)	50.8 (31)
	Olo 5	209	64.7 (8.1)	1.2 (0.5)	42.9 (13.8)	52.3 (27.3)
	Placebo	216	63.8 (8.3)	1.2 (0.5)	43.1 (16)	47.8 (23)
13	Olo 10	225	62.6 (8.8)	1.2 (0.5)	44.2 (14.6)	43.4 (31.6)
	Olo 5	227	63.7 (9.1)	1.3 (0.5)	46.4 (14.8)	45.2 (25)
	Placebo	225	64 (8.4)	1.2 (0.5)	44.4 (14.4)	46.2 (26.4)
	For 12	227	64.8 (8.6)	1.3 (0.5)	46.4 (14.5)	45.8 (23.6)
14	Olo 10	234	63.8 (8.5)	1.3 (0.5)	45.4 (14.6)	43.6 (28.6)
	Olo 5	232	63.7 (8.8)	1.3 (0.5)	46.7 (14.8)	40.7 (24.9)
	Placebo	235	63.9 (7.8)	1.3 (0.5)	46.6 (14.9)	41.2 (23.5)
	For 12	233	65 (8.2)	1.2 (0.5)	45.2 (15.3)	44.5 (28.1)
24	Olo 10	92	61.8 (8.6)	1.2 (0.4)	45.2 (13.8)	54.3 (22.9)
	Olo 5	95	61.5 (8.7)	1.2 (0.4)	45.2 (13.6)	53.8 (22.7)
	Placebo	96	62 (8.8)	1.3 (0.4)	45.3 (13.7)	54.6 (24.6)
	For 12	93	61.5 (8.8)	1.2 (0.5)	44.8 (14.1)	54.7 (23.4)
25	Olo 10	95	63.5 (8.3)	1.2 (0.5)	46.2 (14.7)	51.8 (26.8)
	Olo 5	93	63.2 (8)	1.2 (0.5)	46 (14.7)	49.9 (24.9)
	Placebo	94	63.3 (8.2)	1.3 (0.5)	46.5 (14.4)	50.9 (25.4)
	For 12	93	63.5 (8.3)	1.2 (0.5)	45.9 (14.8)	51.5 (27.1)
37	Olo 10	143	60.9 (7.6)	1.5 (0.5)	48.4 (14.6)	45.8 (23)
	Olo 5	147	60.7 (7.7)	1.5 (0.5)	48.3 (14.6)	45.6 (22.8)
	Placebo	143	60.7 (7.7)	1.5 (0.5)	48.8 (14.3)	46.1 (22.8)
38	Olo 10	147	60.7 (7.5)	1.6 (0.5)	51.5 (13.9)	50.3 (29.4)
	Olo 5	150	60.5 (7.6)	1.6 (0.5)	52 (13.8)	49.9 (29.2)
	Placebo	149	60.6 (7.7)	1.6 (0.5)	51.8 (13.8)	50.7 (29.6)
39	Olo 10	101	61.6 (8.7)	1.5 (0.5)	49.6 (15.4)	41.1 (17.9)
	Olo 5	101	61.6 (8.7)	1.5 (0.5)	49.3 (15.2)	40.8 (17.8)
	Placebo	102	61.5 (8.8)	1.5 (0.5)	50.4 (15.4)	41.2 (19.2)
	Tio 18	101	61.6 (8.6)	1.5 (0.5)	50.2 (15.4)	41.2 (19.1)
40	Olo 10	113	62.5 (7.9)	1.2 (0.5)	43.1 (13.8)	47.5 (23.1)
	Olo 5	116	62.8 (7.9)	1.2 (0.5)	43 (13.7)	47.1 (23)
	Placebo	110	62.3 (8)	1.2 (0.5)	42.7 (13.8)	48.1 (23.8)
	Tio 18	113	62.6 (7.9)	1.2 (0.5)	42.7 (13.7)	47.8 (23.7)

source: demographics 2012 08 30.sas

Reference ID: 3247160

Table 9. Change from baseline trough FEV_1 (liters), pre-planned analyses

Study	Wk	Trt	N	Δ FEV1	Trt -P	P-value	95% Conf Int
11	12	Placebo	188	-0.032			
		Olo 5	191	0.052	0.084	0.0002	(0.040, 0.129)
		Olo 10	192	0.048	0.080	0.0003	(0.037, 0.124)
12	12	Placebo	194	0.005			
		Olo 5	196	0.038	0.033	0.1624	(-0.013, 0.080)
		Olo 10	200	0.049	0.045	0.0563	(-0.001, 0.090)
13	12	Placebo	187	-0.027			
		Olo 5	207	0.056	0.083	<.0001	(0.043, 0.123)
		Olo 10	206	0.048	0.075	0.0002	(0.035, 0.114)
		For 12	200	0.033	0.059	0.0037	(0.019, 0.100)
14	12	Placebo	212	-0.041			
		Olo 5	220	0.018	0.059	0.0017	(0.022, 0.095)
		Olo 10	219	0.052	0.093	<.0001	(0.057, 0.130)
		For 12	217	0.024	0.065	0.0005	(0.028, 0.101)
24	6	Placebo	93	-0.093			
		Olo 5	92	0.012	0.106	<.0001	(0.064, 0.147)
		Olo 10	91	0.02	0.113	<.0001	(0.072, 0.155)
		For 12	90	0.04	0.133	<.0001	(0.092, 0.175)
25	6	Placebo	91	0.012			
		Olo 5	92	0.109	0.097	0.0003	(0.045, 0.148)
		Olo 10	90	0.115	0.103	0.0001	(0.051, 0.155)
		For 12	90	0.093	0.080	0.0026	(0.028, 0.132)
39	6	Placebo	99	-0.043			
		Olo 5	100	0.091	0.134	<.0001	(0.097, 0.170)
		Olo 10	99	0.104	0.147	<.0001	(0.110, 0.184)
		Tio 18	99	0.054	0.097	<.0001	(0.060, 0.133)
40	6	Placebo	105	0.004			
		Olo 5	115	0.138	0.134	<.0001	(0.097, 0.172)
		Olo 10	107	0.146	0.143	<.0001	(0.105, 0.181)
		Tio 18	112	0.161	0.158	<.0001	(0.120, 0.195)

Source: FEV1 2012 08 02.sas

Reference ID: 3247160

Table 10. Change from baseline trough FEV1, observed data only

Study	Wk	Trt	N	Δ FEV1	Trt -P	P-value	95% Conf Int
11	12	Placebo	178	-0.030			
		Olo 5	185	0.054	0.084	0.0003	(0.039, 0.129)
		Olo 10	189	0.049	0.079	0.0005	(0.035, 0.123)
12	12	Placebo	184	0.009			
		Olo 5	193	0.039	0.030	0.2077	(-0.017, 0.078)
		Olo 10	197	0.049	0.040	0.0912	(-0.006, 0.086)
13	12	Placebo	180	-0.024			
		Olo 5	200	0.059	0.083	<.0001	(0.043, 0.123)
		Olo 10	200	0.052	0.076	0.0002	(0.036, 0.117)
		For 12	197	0.034	0.059	0.0046	(0.018, 0.099)
14	12	Placebo	206	-0.037			
		Olo 5	216	0.023	0.061	0.0012	(0.024, 0.097)
		Olo 10	216	0.053	0.09	<.0001	(0.054, 0.127)
		For 12	210	0.027	0.065	0.0006	(0.028, 0.102)

Source: FEV1 2012 08 02.sas

Reference ID: 3247160

Table 11. Change from baseline trough FEV_1 by week, study 11

Study	Wk	Trt	N	ΔFEV_1	Trt -P	P-value	95% Conf Int
11	2	Placebo	200	-0.023			
		Olo 5	196	0.076	0.099	<.0001	(0.055, 0.143)
		Olo 10	199	0.082	0.105	<.0001	(0.062, 0.148)
	6	Placebo	194	-0.025			
		Olo 5	195	0.08	0.105	<.0001	(0.061, 0.149)
		Olo 10	195	0.057	0.082	0.0002	(0.039, 0.125)
	12	Placebo	188	-0.032			
		Olo 5	191	0.052	0.084	0.0002	(0.040, 0.129)
		Olo 10	192	0.048	0.080	0.0003	(0.037, 0.124)
	18	Placebo	183	-0.031			
		Olo 5	187	0.061	0.092	<.0001	(0.047, 0.137)
		Olo 10	189	0.05	0.082	0.0003	(0.038, 0.125)
	24	Placebo	180	-0.04			
		Olo 5	186	0.046	0.086	0.0002	(0.041, 0.131)
		Olo 10	179	0.039	0.080	0.0004	(0.036, 0.124)
	32	Placebo	174	-0.036			
		Olo 5	180	0.049	0.085	0.0002	(0.040, 0.130)
		Olo 10	179	0.034	0.070	0.0019	(0.026, 0.114)
	40	Placebo	174	-0.06			
		Olo 5	175	0.047	0.107	<.0001	(0.062, 0.153)
		Olo 10	174	0.042	0.103	<.0001	(0.058, 0.147)
	48	Placebo	168	-0.082			
		Olo 5	176	0.023	0.105	<.0001	(0.060, 0.151)
		Olo 10	173	0.014	0.097	<.0001	(0.052, 0.141)

source: FEV1 2012 11 29.sas

Reference ID: 3247160

Table 12. Change from baseline trough FEV_1 by week, study 12

Study	Wk	Trt	N	ΔFEV_1	Trt -P	P-value	95% Conf Int
12	2	Placebo	200	0.009			
		Olo 5	203	0.062	0.053	0.0261	(0.006, 0.099)
		Olo 10	206	0.071	0.061	0.0081	(0.016, 0.107)
	6	Placebo	197	0.005			
		Olo 5	203	0.069	0.064	0.0070	(0.017, 0.110)
		Olo 10	204	0.081	0.076	0.0011	(0.031, 0.122)
	12	Placebo	194	0.005			
		Olo 5	196	0.038	0.033	0.1624	(-0.013, 0.080)
		Olo 10	200	0.049	0.045	0.0563	(-0.001, 0.090)
	18	Placebo	190	0.008			
		Olo 5	192	0.055	0.047	0.0478	(0.000, 0.094)
		Olo 10	193	0.034	0.026	0.2701	(-0.020, 0.072)
	24	Placebo	183	-0.015			
		Olo 5	192	0.033	0.048	0.0475	(0.001, 0.095)
		Olo 10	192	0.026	0.040	0.0884	(-0.006, 0.087)
	32	Placebo	182	-0.014			
		Olo 5	190	0.022	0.036	0.1380	(-0.012, 0.083)
		Olo 10	186	-0.015	-0.002	0.9491	(-0.048, 0.045)
	40	Placebo	181	-0.019			
		Olo 5	187	0.024	0.043	0.0742	(-0.004, 0.091)
		Olo 10	181	0.035	0.054	0.0235	(0.007, 0.100)
	48	Placebo	181	-0.062			
		Olo 5	182	0.005	0.067	0.0065	(0.019, 0.115)
		Olo 10	181	0.017	0.079	0.0009	(0.033, 0.126)

source: FEV1 2012 11 29.sas

Reference ID: 3247160

Table 13. Change from baseline trough FEV_1 by week, study 13

Study	Wk	Trt	N	ΔFEV_1	Trt -P	P-value	95% Conf Int
13	2	Placebo	201	-0.019			
		Olo 5	217	0.068	0.087	<.0001	(0.048, 0.126)
		Olo 10	214	0.06	0.079	<.0001	(0.040, 0.118)
		For 12	212	0.061	0.08	<.0001	(0.040, 0.119)
	6	Placebo	196	-0.037			
		Olo 5	215	0.049	0.086	<.0001	(0.047, 0.125)
		Olo 10	213	0.041	0.078	0.0001	(0.038, 0.117)
		For 12	208	0.042	0.079	<.0001	(0.039, 0.119)
	12	Placebo	187	-0.027			
		Olo 5	207	0.056	0.083	<.0001	(0.043, 0.123)
		Olo 10	206	0.048	0.075	0.0002	(0.035, 0.114)
		For 12	200	0.033	0.059	0.0037	(0.019, 0.100)
	18	Placebo	183	-0.019			
		Olo 5	203	0.046	0.065	0.0016	(0.025, 0.105)
		Olo 10	200	0.026	0.045	0.0276	(0.005, 0.085)
		For 12	199	0.023	0.042	0.0426	(0.001, 0.082)
	24	Placebo	174	-0.056			
		Olo 5	200	0.021	0.078	0.0002	(0.037, 0.118)
		Olo 10	198	0.028	0.085	<.0001	(0.044, 0.125)
		For 12	194	-0.002	0.054	0.0088	(0.014, 0.095)
	32	Placebo	175	-0.023			
		Olo 5	198	0.023	0.047	0.0237	(0.006, 0.087)
		Olo 10	194	0.026	0.049	0.0175	(0.009, 0.090)
		For 12	190	0.021	0.044	0.0339	(0.003, 0.085)
	40	Placebo	165	-0.02			
		Olo 5	198	0.02	0.040	0.0537	(-0.001, 0.081)
		Olo 10	187	0.017	0.037	0.0808	(-0.004, 0.078)
		For 12	185	0.004	0.024	0.2579	(-0.017, 0.065)
	48	Placebo	169	-0.065			
		Olo 5	194	0.003	0.068	0.0011	(0.027, 0.109)
		Olo 10	189	-0.009	0.057	0.0069	(0.016, 0.098)
		For 12	186	-0.006	0.059	0.0047	(0.018, 0.101)

source: FEV1 2012 11 29.sas

Reference ID: 3247160

Table 14. Change from baseline trough FEV_1 by week, study 14

Study	Wk	Trt	N	ΔFEV_1	Trt -P	P-value	95% Conf Int
14	2	Placebo	230	-0.016			
		Olo 5	226	0.053	0.069	0.0002	(0.033, 0.105)
		Olo 10	227	0.103	0.119	<.0001	(0.083, 0.155)
		For 12	227	0.033	0.049	0.0071	(0.013, 0.085)
	6	Placebo	223	-0.036			
		Olo 5	222	0.047	0.084	<.0001	(0.048, 0.120)
		Olo 10	225	0.068	0.104	<.0001	(0.068, 0.141)
		For 12	223	0.034	0.07	0.0001	(0.034, 0.106)
	12	Placebo	212	-0.041			
		Olo 5	220	0.018	0.059	0.0017	(0.022, 0.095)
		Olo 10	219	0.052	0.093	<.0001	(0.057, 0.130)
		For 12	217	0.024	0.065	0.0005	(0.028, 0.101)
	18	Placebo	203	-0.036			
		Olo 5	214	0.013	0.050	0.0085	(0.013, 0.087)
		Olo 10	219	0.049	0.086	<.0001	(0.049, 0.122)
		For 12	212	0.015	0.051	0.0067	(0.014, 0.088)
	24	Placebo	198	-0.055			
		Olo 5	209	-0.003	0.053	0.0055	(0.015, 0.090)
		Olo 10	211	0.014	0.069	0.0003	(0.032, 0.106)
		For 12	208	-0.013	0.042	0.0270	(0.005, 0.080)
	32	Placebo	191	-0.039			
		Olo 5	204	0.023	0.062	0.0012	(0.025, 0.100)
		Olo 10	204	0.034	0.073	0.0002	(0.035, 0.110)
		For 12	205	0.009	0.048	0.0117	(0.011, 0.086)
	40	Placebo	189	-0.043			
		Olo 5	202	0.019	0.062	0.0014	(0.024, 0.099)
		Olo 10	199	0.041	0.084	<.0001	(0.047, 0.122)
		For 12	201	0.013	0.056	0.0035	(0.019, 0.094)
	48	Placebo	185	-0.06			
		Olo 5	198	-0.016	0.044	0.0228	(0.006, 0.082)
		Olo 10	198	-0.001	0.059	0.0024	(0.021, 0.097)
		For 12	197	-0.024	0.036	0.0664	(-0.002, 0.074)

source: FEV1 2012 11 29.sas

Reference ID: 3247160

Table 15. Hypothesis generating trough FEV_1 analyses, studies 11 and 12, olodaterol effects modeled without tiotropium interactions

Study	Wk	Trt	N	Δ FEV1	Trt -P	P-value	95% Conf Int
11	12	Placebo	188	-0.041			
		Olo 5	191	0.050	0.091	<.0001	(0.054, 0.128)
		Olo 10	192	0.060	0.101	<.0001	(0.064, 0.137)
12	12	Placebo	194	-0.003			
		Olo 5	196	0.044	0.047	0.0116	(0.011, 0.084)
		Olo 10	200	0.045	0.048	0.0095	(0.012, 0.085)

Source: FEV1 2012 08 02.sas

Reference ID: 3247160

Table 16. Trough FEV1, stratified by baseline tiotropium or long term anticholinergic

Study	Tio	Wk	Trt	N	Δ FEV1	Trt -P	PValue	Conf_Int
11	Yes	12	Placebo	41	-0.016			
			Olo 5	42	0.056	0.072	0.0720	(-0.006, 0.150)
			Olo 10	48	0.025	0.041	0.2893	(-0.035, 0.117)
	No		Placebo	147	-0.048			
			Olo 5	149	0.048	0.096	<.0001	(0.055, 0.138)
			Olo 10	144	0.071	0.119	<.0001	(0.077, 0.161)
12	Yes		Placebo	37	0.019			
			Olo 5	37	0.03	0.011	0.7990	(-0.073, 0.095)
			Olo 10	42	0.058	0.039	0.3489	(-0.043, 0.121)
	No		Placebo	157	-0.009			
			Olo 5	159	0.047	0.056	0.0073	(0.015, 0.096)
			Olo 10	158	0.041	0.05	0.0162	(0.009, 0.091)
13	Yes	24	Placebo	51	-0.044			
			Olo 5	57	0.047	0.090	0.0211	(0.014, 0.167)
			Olo 10	53	0.028	0.072	0.0693	(-0.006, 0.150)
			For 12	47	0.043	0.087	0.0338	(0.007, 0.167)
	No		Placebo	136	-0.014			
			Olo 5	150	0.067	0.080	0.0007	(0.034, 0.127)
			Olo 10	153	0.062	0.075	0.0015	(0.029, 0.122)
			For 12	153	0.037	0.05	0.0329	(0.004, 0.097)
14	Yes		Placebo	55	-0.079			
			Olo 5	55	0.029	0.107	0.0036	(0.035, 0.180)
			Olo 10	57	0.032	0.11	0.0027	(0.038, 0.182)
			For 12	54	0.021	0.1	0.0071	(0.027, 0.173)
	No		Placebo	157	-0.02			
			Olo 5	165	0.022	0.042	0.0523	(-0.000, 0.084)
			Olo 10	162	0.068	0.087	<.0001	(0.045, 0.130)
			For 12	163	0.033	0.053	0.0148	(0.010, 0.095)
24	Yes		Placebo	13	-0.118			
			Olo 5	14	0.087	0.205	0.0003	(0.094, 0.316)
			Olo 10	14	0.081	0.199	0.0005	(0.088, 0.310)
			For 12	14	0.084	0.202	0.0004	(0.092, 0.313)
	No	6	Placebo	77	-0.093			
			Olo 5	75	0.003	0.096	<.0001	(0.050, 0.141)
			Olo 10	74	0.008	0.101	<.0001	(0.055, 0.146)
			For 12	73	0.032	0.125	<.0001	(0.079, 0.171)

Reference ID: 3247160

Table 16 (continued)

Study	Tio	Wk	Trt	N	Δ FEV1	Trt -P	PValue	Conf_Int
25	Yes		Placebo	29	0.008			
			Olo 5	31	0.061	0.052	0.2562	(-0.038, 0.142)
			Olo 10	30	0.093	0.084	0.0694	(-0.007, 0.176)
			For 12	32	0.051	0.042	0.3534	(-0.047, 0.132)
25	No		Placebo	62	0.015			
			Olo 5	61	0.134	0.119	0.0002	(0.056, 0.182)
			Olo 10	60	0.127	0.111	0.0006	(0.048, 0.175)
			For 12	58	0.115	0.1	0.0024	(0.036, 0.164)

Source: Subgr FEV1 AUC 2012 11 28.sas

Reference ID: 3247160

Table 17. Change from baseline FEV_1 $AUC_{0\text{-}3hr}$ (liters)

Study	Wk	Trt	N	Δ AUC	Trt -P	P-value	95% Conf Int
11	12	Placebo	188	0.002			
		Olo 5	192	0.167	0.164	<.0001	(0.120, 0.209)
		Olo 10	192	0.157	0.155	<.0001	(0.111, 0.199)
12	12	Placebo	197	0.021			
		Olo 5	196	0.155	0.134	<.0001	(0.090, 0.177)
		Olo 10	201	0.151	0.130	<.0001	(0.087, 0.172)
13	12	Placebo	187	-0.003			
		Olo 5	208	0.176	0.178	<.0001	(0.137, 0.219)
		Olo 10	207	0.167	0.170	<.0001	(0.129, 0.211)
		For 12	200	0.182	0.185	<.0001	(0.144, 0.226)
14	12	Placebo	212	-0.008			
		Olo 5	220	0.138	0.145	<.0001	(0.108, 0.182)
		Olo 10	219	0.167	0.175	<.0001	(0.138, 0.212)
		For 12	217	0.163	0.17	<.0001	(0.133, 0.208)
24	6	Placebo	93	-0.03			
		Olo 5	92	0.134	0.164	<.0001	(0.126, 0.201)
		Olo 10	91	0.135	0.164	<.0001	(0.127, 0.202)
		For 12	90	0.168	0.198	<.0001	(0.160, 0.236)
25	6	Placebo	91	0.004			
		Olo 5ug	92	0.190	0.186	<.0001	(0.149, 0.223)
		Olo 10ug	90	0.202	0.198	<.0001	(0.162, 0.235)
		Form 12ug	90	0.217	0.213	<.0001	(0.176, 0.250)
39	6	Placebo	99	-0.045			
		Olo 5	100	0.162	0.207	<.0001	(0.170, 0.243)
		Olo 10	99	0.170	0.215	<.0001	(0.179, 0.252)
		Tio 18	99	0.137	0.182	<.0001	(0.146, 0.219)
40	6	Placebo	105	0.011			
		Olo 5ug	115	0.225	0.214	<.0001	(0.178, 0.251)
		Olo 10ug	106	0.255	0.245	<.0001	(0.208, 0.281)
		Tio 18ug	112	0.246	0.235	<.0001	(0.199, 0.272)

Source: AUC 2012 08 14.sas

Reference ID: 3247160

Table 18. Change from baseline FEV_1 $AUC_{0\text{-}3hr}$ (liters), 48 week studies, 6 and 12 weeks

Study	Wk	Trt	N	Δ AUC	Trt -P	P-value	95% Conf Int
11	6	Placebo	194	-0.002			
		Olo 5	196	0.185	0.187	<.0001	(0.143, 0.231)
		Olo 10	196	0.151	0.153	<.0001	(0.109, 0.196)
	12	Placebo	188	0.002			
		Olo 5	192	0.167	0.164	<.0001	(0.120, 0.209)
		Olo 10	192	0.157	0.155	<.0001	(0.111, 0.199)
12	6	Placebo	197	0.02			
		Olo 5	203	0.187	0.168	<.0001	(0.124, 0.211)
		Olo 10	205	0.166	0.146	<.0001	(0.104, 0.189)
	12	Placebo	197	0.021			
		Olo 5	196	0.155	0.134	<.0001	(0.090, 0.177)
		Olo 10	201	0.151	0.13	<.0001	(0.087, 0.172)
13	6	Placebo	197	0.001			
		Olo 5	218	0.178	0.176	<.0001	(0.136, 0.217)
		Olo 10	213	0.161	0.16	<.0001	(0.119, 0.200)
		For 12	208	0.194	0.192	<.0001	(0.152, 0.233)
	12	Placebo	187	-0.003			
		Olo 5	208	0.176	0.178	<.0001	(0.137, 0.219)
		Olo 10	207	0.167	0.17	<.0001	(0.129, 0.211)
		For 12	200	0.182	0.185	<.0001	(0.144, 0.226)
14	6	Placebo	223	-0.01			
		Olo 5	223	0.162	0.172	<.0001	(0.136, 0.209)
		Olo 10	225	0.181	0.192	<.0001	(0.155, 0.228)
		For 12	223	0.174	0.184	<.0001	(0.148, 0.221)
	12	Placebo	212	-0.008			
		Olo 5	220	0.138	0.145	<.0001	(0.108, 0.182)
		Olo 10	219	0.167	0.175	<.0001	(0.138, 0.212)
		For 12	217	0.163	0.17	<.0001	(0.133, 0.208)

Source: Subgr FEV1 AUC 2012 12 01 sas

Reference ID: 3247160

Table 19. Change from baseline FEV_1 $AUC_{0\text{-}12hr}$ (liters)

Study	Wk	Trt	N	Δ AUC	Trt -P	P-value	95% Conf Int
11	12	Placebo	71	-0.007			
		Olo 5	85	0.146	0.152	0.0001	(0.075, 0.230)
		Olo 10	85	0.127	0.133	0.0006	(0.058, 0.209)
12	12	Placebo	98	0.019			
		Olo 5	116	0.111	0.092	0.0047	(0.029, 0.156)
		Olo 10	107	0.097	0.078	0.0150	(0.015, 0.140)
24	6	Placebo	93	-0.06			
		Olo 5	92	0.088	0.148	<.0001	(0.113, 0.183)
		Olo 10	91	0.088	0.148	<.0001	(0.113, 0.183)
		For 12	90	0.081	0.141	<.0001	(0.106, 0.177)
25	6	Placebo	91	-0.022			
		Olo 5	92	0.15	0.172	<.0001	(0.139, 0.205)
		Olo 10	90	0.152	0.174	<.0001	(0.140, 0.208)
		For 12	90	0.136	0.158	<.0001	(0.124, 0.191)
39	6	Placebo	99	-0.054			
		Olo 5	100	0.131	0.185	<.0001	(0.146, 0.225)
		Olo 10	99	0.153	0.207	<.0001	(0.167, 0.247)
		Tio 18	99	0.118	0.172	<.0001	(0.132, 0.212)
40	6	Placebo	105	-0.008			
		Olo 5	115	0.189	0.197	<.0001	(0.163, 0.231)
		Olo 10	106	0.213	0.221	<.0001	(0.186, 0.255)
		Tio 18	112	0.214	0.221	<.0001	(0.187, 0.255)

Source: AUC 2012 08 14.sas

Reference ID: 3247160

Table 20. Change from baseline FEV_1 $AUC_{12\text{-}24hr}$ (liters)

Study	Wk	Trt	N	Δ AUC	Trt –P	P-value	95% Conf Int
24	6	Placebo	93	-0.123			
		Olo 5	92	-0.014	0.109	<.0001	(0.073, 0.146)
		Olo 10	91	0.004	0.127	<.0001	(0.091, 0.164)
		For 12	90	0.049	0.172	<.0001	(0.135, 0.209)
25	6	Placebo	91	-0.048			
		Olo 5	92	0.069	0.118	<.0001	(0.082, 0.154)
		Olo 10	90	0.072	0.120	<.0001	(0.084, 0.157)
		For 12	90	0.107	0.155	<.0001	(0.119, 0.191)
39	6	Placebo	99	-0.095			
		Olo 5	100	0.037	0.132	<.0001	(0.091, 0.174)
		Olo 10	99	0.083	0.178	<.0001	(0.137, 0.220)
		Tio 18	99	0.027	0.122	<.0001	(0.081, 0.164)
40	6	Placebo	105	-0.059			
		Olo 5	115	0.094	0.153	<.0001	(0.118, 0.188)
		Olo 10	107	0.111	0.169	<.0001	(0.134, 0.205)
		Tio 18	112	0.105	0.164	<.0001	(0.129, 0.199)

Source: AUC 2012 08 14.sas

Table 21.Change from baseline FEV_1 $AUC_{0\text{-}24hr}$ (liters)

Study	Wk	Trt	N	Δ AUC	Trt –P	P-value	95% Conf Int
24	6	Placebo	93	-0.092			
		Olo 5	92	0.037	0.128	<.0001	(0.094, 0.163)
		Olo 10	91	0.046	0.137	<.0001	(0.103, 0.172)
		For 12	90	0.065	0.156	<.0001	(0.122, 0.191)
25	6	Placebo	91	-0.035			
		Olo 5	92	0.11	0.145	<.0001	(0.114, 0.176)
		Olo 10	90	0.112	0.147	<.0001	(0.116, 0.179)
		For 12	90	0.121	0.156	<.0001	(0.125, 0.187)
39	6	Placebo	99	-0.075			
		Olo 5	100	0.084	0.159	<.0001	(0.121, 0.196)
		Olo 10	99	0.118	0.193	<.0001	(0.155, 0.230)
		Tio 18	99	0.072	0.147	<.0001	(0.110, 0.185)
40	6	Placebo	105	-0.033			
		Olo 5	115	0.142	0.175	<.0001	(0.141, 0.209)
		Olo 10	107	0.158	0.191	<.0001	(0.157, 0.224)
		Tio 18	112	0.159	0.192	<.0001	(0.159, 0.226)

Source: AUC 2012 08 14.sas

Reference ID: 3247160

Table 22. Change from baseline Mahler TDI Focal Score, studies 13 and 14

Study	Wk	Trt	N	ΔTDI	Trt –P	P-value	95% Conf Int
1222_0013	24	Placebo	175	2.046			
		Olo 5	200	2.234	0.189	0.5843	(-0.485, 0.860)
		Olo 10	192	2.068	0.021	0.9494	(-0.656, 0.699)
		Form 12	192	1.818	-0.228	0.5099	(-0.908, 0.451)
1222_0014	24	Placebo	197	1.102			
		Olo 5	208	1.504	0.402	0.1999	(-0.213, 1.018)
		Olo 10	209	1.521	0.42	0.1818	(-0.196, 1.035)
		Form 12	202	1.703	0.603	0.0572	(-0.019, 1.222)
Combined	24	Placebo	372	1.554			
		Olo 5	408	1.867	0.312	0.1767	(-0.141, 0.768)
		Olo 10	401	1.787	0.231	0.3180	(-0.224, 0.689)
		Form 12	394	1.759	0.204	0.3808	(-0.254, 0.664)

Source: Mahler TDI 2012 11 08.sas

7

Reference ID: 3247160

Table 23. SGRQ, studies 13 and 14

Study	Wk	Trt	N	SGRQ	Trt –P	P-value	95% Conf Int
13	12	Placebo	182	42.11			
		Olo 5	199	39.32	-2.78	0.0450	(-5.51, -0.06)
		Olo 10	198	36.96	-5.14	0.0002	(-7.86, -2.43)
		For 12	200	40.35	-1.75	0.2061	(-4.47, 0.97)
	24	Placebo	175	41.07			
		Olo 5	193	38.63	-2.44	0.0816	(-5.19, 0.31)
		Olo 10	192	37.67	-3.39	0.0155	(-6.14, -0.65)
		For 12	198	40.12	-0.95	0.4954	(-3.69, 1.79)
	48	Placebo	165	40.42			
		Olo 5	190	38.54	-1.87	0.1878	(-4.66, 0.91)
		Olo 10	182	36.85	-3.57	0.0126	(-6.36, -0.77)
		For 12	184	40.43	0.02	0.9913	(-2.78, 2.81)
14	12	Placebo	198	42.68			
		Olo 5	214	40.05	-2.63	0.0491	(-5.24, -0.01)
		Olo 10	209	40.19	-2.49	0.0636	(-5.12, 0.14)
		For 12	203	39.52	-3.16	0.0194	(-5.81, -0.51)
	24	Placebo	190	42.12			
		Olo 5	207	38.97	-3.15	0.0197	(-5.80, -0.50)
		Olo 10	205	38.6	-3.52	0.0094	(-6.18, -0.87)
		For 12	197	40.7	-1.42	0.2995	(-4.09, 1.26)
	48	Placebo	176	39.91			
		Olo 5	195	39.56	-0.35	0.7995	(-3.07, 2.36)
		Olo 10	192	38.82	-1.09	0.4336	(-3.82, 1.64)
		For 12	190	40.02	0.11	0.9365	(-2.63, 2.85)

SGRQ 2012 08 21.sas

Reference ID: 3247160

Table 24. Exercise endurance, studies 37 and 38

Study	Trt	N	ExEnd (sec)	Trt /P	P-value	95% Conf Int
37	Placebo	136	369.81			
	Olo 5	140	421.58	1.14	0.0002	(1.065, 1.221)
	Olo 10	136	420.72	1.14	0.0003	(1.062, 1.219)
38	Placebo	146	354.33			
	Olo 5	141	396.31	1.12	0.0018	(1.043, 1.199)
	Olo 10	140	391.45	1.10	0.0052	(1.030, 1.184)

Exercise Endure 2012 08 21.sas

Table 25. Inspiratory capacity (L), studies 37 and 38

Study	Wk	Time	Trt	N	IC	Trt –P	P-value	95% Conf Int
37	6	Pre-exercise	Placebo	134	2.220			
			Olo 5	139	2.478	0.258	<.0001	(0.191, 0.325)
			Olo 10	133	2.513	0.294	<.0001	(0.226, 0.362)
		Isotime	Placebo	136	1.917			
			Olo 5	140	2.099	0.182	<.0001	(0.112, 0.252)
			Olo 10	135	2.091	0.174	<.0001	(0.104, 0.245)
		End-exercise	Placebo	136	1.887			
			Olo 5	140	2.067	0.180	<.0001	(0.107, 0.252)
			Olo 10	135	2.024	0.137	0.0003	(0.064, 0.210)
38	6	Pre-exercise	Placebo	146	2.273			
			Olo 5	141	2.437	0.164	<.0001	(0.094, 0.234)
			Olo 10	138	2.468	0.195	<.0001	(0.125, 0.265)
		Isotime	Placebo	146	2.162			
			Olo 5	141	2.246	0.084	0.0155	(0.016, 0.152)
			Olo 10	139	2.328	0.166	<.0001	(0.098, 0.234)
		End-exercise	Placebo	146	2.158			
			Olo 5	141	2.236	0.078	0.0245	(0.010, 0.146)
			Olo 10	139	2.330	0.172	<.0001	(0.105, 0.240)

Source: IC Isotime 2012 08 21.sas

Reference ID: 3247160

Table 26. Breathing discomfort at exercise isotime, studies 37 and 38

Study	Time	Trt	N	Mean (SE)	Trt - P	P-Value	95% CI
37	Pre-exercise	Placebo	136	0.288 (0.041)			
		Olo 5	140	0.185 (0.040)	-0.103	0.045	(-0.204, -0.002)
		Olo 10	135	0.224 (0.041)	-0.064	0.213	(-0.166, 0.037)
	Isotime	Placebo	136	5.870 (0.185)			
		Olo 5	140	5.104 (0.182)	-0.766	<.001	(-1.205, -0.326)
		Olo 10	136	5.235 (0.185)	-0.634	0.005	(-1.077, -0.192)
	End-exercise	Placebo	136	6.978 (0.143)			
		Olo 5	140	6.890 (0.141)	-0.088	0.570	(-0.391, 0.216)
		Olo 10	136	7.234 (0.143)	0.256	0.100	(-0.050, 0.561)
38	Pre-exercise	Placebo	146	0.389 (0.062)			
		Olo 5	141	0.315 (0.063)	-0.074	0.290	(-0.211, 0.063)
		Olo 10	140	0.364 (0.064)	-0.025	0.720	(-0.162, 0.112)
	Isotime	Placebo	146	5.585 (0.177)			
		Olo 5	141	5.250 (0.180)	-0.336	0.118	(-0.757, 0.085)
		Olo 10	140	5.520 (0.181)	-0.066	0.759	(-0.486, 0.355)
	End-exercise	Placebo	146	7.010 (0.129)			
		Olo 5	141	7.101 (0.131)	0.091	0.531	(-0.196, 0.379)
		Olo 10	140	7.351 (0.132)	0.341	0.020	(0.055, 0.628)

Source: Borg.sas

Table 27. Functional residual capacity (L), studies 37 and 38. Exploratory analysis

Study	Wk	Time	Trt	N	FRC	Trt –P	P-value	95% Conf Int
37	6	-0:30:00	Placebo	134	4.977			
			Olo 5	139	4.855	-0.122	0.0784	(-0.258, 0.014)
			Olo 10	134	4.862	-0.115	0.1013	(-0.252, 0.023)
		1:00:00	Placebo	134	4.950			
			Olo 5	139	4.740	-0.210	0.0015	(-0.339, -0.081)
			Olo 10	134	4.577	-0.373	<.0001	(-0.503, -0.243)
38	6	-0:30:00	Placebo	147	4.842			
			Olo 5	145	4.757	-0.086	0.1048	(-0.190, 0.018)
			Olo 10	141	4.723	-0.120	0.0246	(-0.224, -0.015)
		1:00:00	Placebo	147	4.770			
			Olo 5	145	4.557	-0.213	<.0001	(-0.318, -0.108)
			Olo 10	141	4.583	-0.187	0.0005	(-0.293, -0.082)

Source: FRC 2012 08 21.sas

Reference ID: 3247160

Table 28. Exacerbations, 48 week studies

			Comparison to Placebo		
Study	**Exacerbation**	**Trt**	**HazardRatio (SE)**	**P_Value**	**95% Conf Int**
11	Any	Olo 5	0.728 (0.124)	0.0658	(0.522, 1.016)
		Olo 10	0.799 (0.133)	0.1701	(0.576, 1.107)
	Hospitalization	Olo 5	0.656 (0.247)	0.2754	(0.313, 1.374)
		Olo 10	1.010 (0.342)	0.9792	(0.521, 1.961)
	Moderate	Olo 5	0.739 (0.142)	0.1194	(0.506, 1.078)
		Olo 10	0.811 (0.153)	0.2570	(0.561, 1.174)
12	Any	Olo 5	0.993 (0.187)	0.9853	(0.687, 1.436)
		Olo 10	1.241 (0.222)	0.2344	(0.873, 1.763)
	Hospitalization	Olo 5	1.050 (0.438)	0.9035	(0.463, 2.380)
		Olo 10	0.958 (0.408)	0.9304	(0.415, 2.209)
	Moderate	Olo 5	1.090 (0.233)	0.6690	(0.717, 1.657)
		Olo 10	1.344 (0.273)	0.1437	(0.902, 2.002)
13	Any	Olo 5	1.162 (0.190)	0.3424	(0.843, 1.603)
		Olo 10	1.186 (0.195)	0.3023	(0.859, 1.636)
		For 12	0.854 (0.150)	0.3589	(0.605, 1.205)
	Hospitalization	Olo 5	1.818 (0.841)	0.1910	(0.734, 4.503)
		Olo 10	1.743 (0.817)	0.2274	(0.695, 4.369)
		For 12	1.346 (0.664)	0.5538	(0.512, 3.538)
	Moderate	Olo 5	1.101 (0.195)	0.5577	(0.778, 1.557)
		Olo 10	1.017 (0.184)	0.9423	(0.714, 1.448)
		For 12	0.735 (0.143)	0.1097	(0.502, 1.076)
14	Any	Olo 5	0.802 (0.133)	0.1946	(0.580, 1.111)
		Olo 10	0.872 (0.141)	0.4071	(0.634, 1.198)
		For 12	0.922 (0.150)	0.6404	(0.670, 1.267)
	Hospitalization	Olo 5	0.685 (0.249)	0.2942	(0.336, 1.399)
		Olo 10	0.947 (0.316)	0.8594	(0.493, 1.820)
		For 12	0.803 (0.281)	0.5466	(0.405, 1.593)
	Moderate	Olo 5	0.807 (0.146)	0.2494	(0.566, 1.150)
		Olo 10	0.791 (0.143)	0.2006	(0.555, 1.126)
		For 12	0.903 (0.160)	0.5795	(0.637, 1.279)

Source: Exac.sas

Reference ID: 3247160

Table 29. FEV_1 $AUC_{0\text{-}3hr}$ (L), by age

Study	Wk	Age	Trt	N	AUC	Olo 5 –P	95% Conf Int
11	12	40-50	Placebo	10	-0.093		
			Olo 5ug	14	0.161	0.254	(0.099, 0.410)
		>50	Placebo	178	0.008		
			Olo 5ug	178	0.167	0.159	(0.113, 0.204)
12	12	40-50	Placebo	16	0.068		
			Olo 5ug	10	0.245	0.176	(0.037, 0.316)
		>50	Placebo	181	0.017		
			Olo 5ug	186	0.15	0.133	(0.089, 0.177)
13	12	40-50	Placebo	12	-0.067		
			Olo 5ug	23	0.298	0.364	(0.218, 0.510)
		>50	Placebo	175	0.001		
			Olo 5ug	185	0.16	0.160	(0.117, 0.202)
14	12	40-50	Placebo	11	-0.052		
			Olo 5ug	20	0.202	0.254	(0.108, 0.400)
		>50	Placebo	201	-0.001		
			Olo 5ug	200	0.136	0.137	(0.099, 0.176)
24	6	40-50	Placebo	11	-0.037		
			Olo 5ug	12	0.157	0.194	(0.089, 0.300)
		>50	Placebo	82	-0.025		
			Olo 5ug	80	0.134	0.159	(0.120, 0.198)
25	6	40-50	Placebo	7	-0.06		
			Olo 5ug	7	0.215	0.275	(0.144, 0.406)
		>50	Placebo	84	0.009		
			Olo 5ug	85	0.187	0.178	(0.140, 0.216)
39	6	40-50	Placebo	12	-0.099		
			Olo 5ug	12	0.13	0.229	(0.125, 0.334)
		>50	Placebo	87	-0.04		
			Olo 5ug	88	0.162	0.202	(0.162, 0.241)
40	6	40-50	Placebo	10	-0.028		
			Olo 5ug	9	0.306	0.334	(0.210, 0.458)
		>50	Placebo	95	0.025		
			Olo 5ug	106	0.227	0.202	(0.164, 0.240)

source: Subgr FEV1 AUC 2012 12 14.sas

Reference ID: 3247160

Table 30. FEV_1 $AUC_{0\text{-}12hr}$ (L), by age

Study	Wk	Age	Trt	N	AUC	Olo 5 –P	95% Conf Int
11	12	40-50	Placebo	5	-0.141		
			Olo 5ug	7	0.047	0.188	(-0.041, 0.416)
		>50	Placebo	66	-0.001		
			Olo 5ug	78	0.154	0.155	(0.077, 0.233)
12	12	40-50	Placebo	11	-0.001		
			Olo 5ug	5	0.305	0.306	(0.096, 0.516)
		>50	Placebo	87	0.021		
			Olo 5ug	111	0.105	0.084	(0.020, 0.148)
24	6	40-50	Placebo	11	-0.109		
			Olo 5ug	12	0.091	0.2	(0.102, 0.298)
		>50	Placebo	82	-0.051		
			Olo 5ug	80	0.089	0.14	(0.104, 0.177)
25	6	40-50	Placebo	7	-0.056		
			Olo 5ug	7	0.132	0.188	(0.068, 0.307)
		>50	Placebo	84	-0.019		
			Olo 5ug	85	0.152	0.17	(0.136, 0.205)
39	6	40-50	Placebo	12	-0.098		
			Olo 5ug	12	0.095	0.193	(0.079, 0.306)
		>50	Placebo	87	-0.049		
			Olo 5ug	88	0.133	0.182	(0.140, 0.225)
40	6	40-50	Placebo	10	-0.037		
			Olo 5ug	9	0.228	0.265	(0.149, 0.382)
		>50	Placebo	95	0.008		
			Olo 5ug	106	0.198	0.19	(0.154, 0.226)

source: Subgr FEV1 AUC 2012 12 14.sas

Reference ID: 3247160

Table 31. Trough FEV_1 (L), by age

Study	Wk	Age	Trt	N	AUC	Olo 5 –P	95% Conf Int
11	12	40-50	Placebo	10	-0.132		
			Olo 5ug	14	-0.039	0.093	(-0.060, 0.245)
		>50	Placebo	178	-0.026		
			Olo 5ug	177	0.057	0.083	(0.038, 0.129)
12	12	40-50	Placebo	16	0.025		
			Olo 5ug	10	0.109	0.084	(-0.065, 0.232)
		>50	Placebo	178	0.003		
			Olo 5ug	186	0.034	0.031	(-0.017, 0.079)
13	12	40-50	Placebo	12	-0.037		
			Olo 5ug	23	0.105	0.142	(0.001, 0.284)
		>50	Placebo	175	-0.021		
			Olo 5ug	184	0.056	0.077	(0.035, 0.118)
14	12	40-50	Placebo	11	-0.112		
			Olo 5ug	20	0.03	0.142	(-0.003, 0.286)
		>50	Placebo	201	-0.031		
			Olo 5ug	200	0.023	0.054	(0.017, 0.092)
24	6	40-50	Placebo	11	-0.109		
			Olo 5ug	12	0.067	0.175	(0.058, 0.293)
		>50	Placebo	82	-0.087		
			Olo 5ug	80	0.008	0.095	(0.051, 0.139)
25	6	40-50	Placebo	7	-0.046		
			Olo 5ug	7	0	0.046	(-0.140, 0.232)
		>50	Placebo	84	0.017		
			Olo 5ug	85	0.117	0.1	(0.046, 0.154)
39	6	40-50	Placebo	12	-0.06		
			Olo 5ug	12	0.076	0.135	(0.029, 0.241)
		>50	Placebo	87	-0.028		
			Olo 5ug	88	0.104	0.132	(0.092, 0.171)
40	6	40-50	Placebo	10	0.079		
			Olo 5ug	9	0.218	0.139	(0.012, 0.266)
		>50	Placebo	95	0.011		
			Olo 5ug	106	0.145	0.133	(0.094, 0.173)

source: Subgr FEV1 AUC 2012 12 14.sas

Reference ID: 3247160

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

/s/

ROBERT ABUGOV
01/17/2013

JOAN K BUENCONSEJO
01/17/2013
I concur with Dr. Abugov's statistical review of NDA 203108 (Olodaterol 5 mcg once daily for the maintenance treatment of airflow obstruction in patients with chronic obstructive pulmonary disease (COPD), including chronic bronchitis and/or emphysema.).

STATISTICS FILING CHECKLIST

NDA Number: 203108 **Applicant: Boerhinger Ingelheim** **Stamp Date: 5/14/2012**

Drug Name: Olodaterol **NDA/BLA Type: Standard**

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

	Content Parameter	Yes	No	NA	Comments
1	Index sufficient to locate necessary reports, tables, data, etc.	**y**			
2	Original protocols, statistical analysis plans, and subsequent amendments available.	**y**			
3	Safety and efficacy for gender, racial, and geriatric subgroups investigated (if applicable).	**y**			
4	Data sets in EDR available and conform to applicable guidance.	**y**			

IS THE STATISTICAL SECTION OF THE APPLICATION FILEABLE? Y

{communications to sponsor to be sent independently of filing letter}

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

/s/

ROBERT ABUGOV
07/10/2012

JOAN K BUENCONSEJO
07/10/2012
I concur