

**CENTER FOR DRUG EVALUATION AND  
RESEARCH**

*APPLICATION NUMBER:*

**203975Orig1s000**

**STATISTICAL REVIEW(S)**



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

## STATISTICAL REVIEW AND EVALUATION

### CLINICAL STUDIES

**NDA/BLA #:** NDA 203-975

**Drug Name:** Anoro Ellipta (umeclidinium/vilanterol)

**Indication(s):** Long-term, once-daily, maintenance bronchodilator treatment of airflow obstruction in patients with chronic obstructive pulmonary disease (COPD), including chronic bronchitis and/or emphysema

**Applicant:** GlaxoSmithKline

**Date(s):** NDA submitted: December 18, 2012  
PDUFA due date: December 18, 2013

**Review Priority:** Standard

**Biometrics Division:** Division of Biometrics II

**Statistical Reviewer:** Gregory Levin, PhD

**Concurring Reviewers:** Joan Buenconsejo, PhD  
Thomas Permutt, PhD

**Medical Division:** Division of Pulmonary, Allergy, and Rheumatology Products

**Clinical Team:** Jennifer Pippins, MD, MPH, Medical Reviewer  
Susan Limb, MD, Medical Team Leader

**Project Manager:** Leila Hann

**Keywords:** NDA, clinical studies, missing data, combination rule

## Table of Contents

<b>1</b>	<b>EXECUTIVE SUMMARY .....</b>	<b>6</b>
<b>2</b>	<b>INTRODUCTION .....</b>	<b>7</b>
2.1	OVERVIEW .....	7
2.1.1	<i>Background .....</i>	<i>7</i>
2.1.2	<i>History of Drug Development .....</i>	<i>8</i>
2.1.3	<i>Specific Studies Reviewed .....</i>	<i>9</i>
2.2	DATA SOURCES .....	9
<b>3</b>	<b>STATISTICAL EVALUATION .....</b>	<b>10</b>
3.1	DATA AND ANALYSIS QUALITY .....	10
3.2	EVALUATION OF EFFICACY .....	10
3.2.1	<i>Study Design.....</i>	<i>10</i>
3.2.2	<i>Statistical Methodologies .....</i>	<i>12</i>
3.2.3	<i>Dose Selection.....</i>	<i>14</i>
3.2.4	<i>Patient Disposition, Demographic and Baseline Characteristics.....</i>	<i>19</i>
3.2.5	<i>Results and Conclusions.....</i>	<i>24</i>
3.3	EVALUATION OF SAFETY .....	45
<b>4</b>	<b>FINDINGS IN SPECIAL/SUBGROUP POPULATIONS .....</b>	<b>52</b>
<b>5</b>	<b>SUMMARY AND CONCLUSIONS .....</b>	<b>54</b>
5.1	STATISTICAL ISSUES .....	54
5.2	COLLECTIVE EVIDENCE.....	56
5.3	LABELING RECOMMENDATIONS .....	57
<b>6</b>	<b>REFERENCES .....</b>	<b>57</b>
<b>7</b>	<b>APPENDIX.....</b>	<b>58</b>

## LIST OF TABLES

Table 1. Differences from Placebo in Mean Change from Baseline in Trough FEV <sub>1</sub> for Different Once-Daily Doses of the Umeclidinium Monotherapy in Studies 321, 73, 589, and 408.....	15
Table 2. Differences from Placebo in Mean Change from Baseline in Trough FEV <sub>1</sub> for Different Once-Daily Doses of the Vilanterol Monotherapy in Studies 45, 575, and 310 .....	19
Table 3. Proportion of Patients Failing to Complete Study 361, by Reason for Withdrawal.....	22
Table 4. Proportion of Patients Failing to Complete Study 373, by Reason for Withdrawal.....	22
Table 5. Proportion of Patients Failing to Complete Study 360, by Reason for Withdrawal.....	22
Table 6. Proportion of Patients Failing to Complete Study 374, by Reason for Withdrawal.....	23
Table 7. Proportion of Patients Failing to Complete Study 359, by Reason for Withdrawal.....	23
Table 8. Comparisons of UMEC, VI, and UMEC/VI against Placebo with Respect to the Mean Changes from Baseline in Trough FEV <sub>1</sub> (Primary Endpoint) and 0–6 Hour Weighted Mean FEV <sub>1</sub> (Secondary Endpoint) at 24 Weeks in Studies 361 and 373 .....	25
Table 9. Comparisons against Placebo for Additional Supportive Endpoints in Studies 361 and 373.....	30
Table 10. Contributions of the Umeclidinium and Vilanterol Components to the Efficacy of the Combination Product with Respect to the Mean Changes from Baseline in Trough FEV <sub>1</sub> (Primary Endpoint) and 0–6 Hour Weighted Mean FEV <sub>1</sub> (Secondary Endpoint) at 24 Weeks in Studies 361, 373, 360, and 374 .....	31
Table 11. Contributions of the Umeclidinium and Vilanterol Components to the Efficacy of the Combination Product with Respect to Additional Supportive Endpoints in Studies 361, 373, 360, and 374.....	35
Table 12. Comparisons against Tiotropium with Respect to the Mean Changes from Baseline in Trough FEV <sub>1</sub> (Primary Endpoint) and 0–6 Hour Weighted Mean FEV <sub>1</sub> (Secondary Endpoint) at 24 Weeks in Studies 360 and 374 .....	36
Table 13. Comparisons against Tiotropium for Additional Supportive Endpoints in Studies 360 and 374.....	37
Table 14. Baseline Characteristics, Stratified According to Whether Patients Completed the Study, Based on Integrated Data from Studies 361, 373, 360, and 374 .....	39
Table 15. Exploring the Potential Effect of Missing Data: Results for the Primary Endpoint Trough FEV <sub>1</sub> with the Primary Mixed Effects Analysis as Compared to a Multiple Imputation Sensitivity Analysis in Studies 361 and 373.....	41
Table 16. Exploring the Potential Effect of Missing Data: Results for the Primary Endpoint Trough FEV <sub>1</sub> with the Primary Mixed Effects Analysis as Compared to a Multiple Imputation Sensitivity Analysis in Studies 360 and 374.....	41
Table 17. Exploring the Potential Effect of Missing Data: Results for Secondary and Additional Endpoints with the Primary Mixed Effects Analysis as Compared to a Multiple Imputation Sensitivity Analysis in Studies 361 and 373.....	42
Table 18. Treatment Effects on the Co-Primary Endpoints Exercise Endurance Time (EET) and Trough FEV <sub>1</sub> at 12 Weeks in the Cross-Over Study 417 .....	44
Table 19. Treatment Effects on the Co-Primary Endpoints Exercise Endurance Time (EET) and Trough FEV <sub>1</sub> at 12 Weeks in the Cross-Over Study 418 .....	45
Table 20. Numbers of MACE and Adjudicated Cardiovascular Serious Adverse Events, and Unadjusted Pooled Incidence Rates, by Treatment Group and Source of Data .....	47

Table 21. Baseline Characteristics in Study 361 .....	58
Table 22. Baseline Characteristics in Study 373 .....	59
Table 23. Baseline Characteristics in Study 360 .....	60
Table 24. Baseline Characteristics in Study 374 .....	61

## LIST OF FIGURES

Figure 1. Postdose 6-Hour Serial Mean Change from Baseline in FEV <sub>1</sub> at Day 1 for Different Once-Daily Umeclidinium Doses, and Tiotropium, in Study 321.....	16
Figure 2. Postdose 24-Hour Serial Mean Change from Baseline in FEV <sub>1</sub> at Day 7 for Different Once-Daily Umeclidinium Doses, and Tiotropium, in Study 321.....	16
Figure 3. Difference from Placebo in Mean Change from Baseline in Postdose 0-6 Hour Weighted Mean FEV <sub>1</sub> at Day 1 for Different Once-Daily Umeclidinium Doses, and Tiotropium, in Study 321 .....	17
Figure 4. Postdose 6-Hour Serial Mean Change from Baseline in FEV <sub>1</sub> at Day 1 for Different Once- and Twice-Daily Umeclidinium Doses, and Tiotropium, in Study 321 .....	17
Figure 5. Postdose 24-Hour Serial Mean Change from Baseline in FEV <sub>1</sub> at Day 7 for Different Once- and Twice-Daily Umeclidinium Doses, and Tiotropium, in Study 321 .....	18
Figure 6. Difference from Placebo in Mean Change from Baseline in Trough FEV <sub>1</sub> at Day 8 for Different Once-Daily Umeclidinium Doses, and Tiotropium, in Study 321 .....	18
Figure 7. Proportion of Patients Withdrawing Early over Time in Study 361 .....	20
Figure 8. Proportion of Patients Withdrawing Early over Time in Study 373 .....	20
Figure 9. Proportion of Patients Withdrawing Early over Time in Study 360 .....	21
Figure 10. Proportion of Patients Withdrawing Early over Time in Study 374 .....	21
Figure 11. Mean Change from Baseline in Trough FEV <sub>1</sub> over Time in Study 361 .....	26
Figure 12. Mean Change from Baseline in Trough FEV <sub>1</sub> over Time in Study 373 .....	26
Figure 13. Mean Change from Baseline in Serial FEV <sub>1</sub> over 0 to 24 Hours Postdose at Day 168 in the Subset of Study 361 with 24-Hour measurements.....	27
Figure 14. Mean Change from Baseline in Serial FEV <sub>1</sub> over 0 to 24 Hours Postdose at Day 168 in the Subset of Study 373 with 24-Hour measurements.....	27
Figure 15. Empirical Distribution Function for Change from Baseline in Trough FEV <sub>1</sub> at 24 Weeks in Study 361 .....	28
Figure 16. Empirical Distribution Function for Change from Baseline in Trough FEV <sub>1</sub> at 24 Weeks in Study 373 .....	29
Figure 17. Mean Change from Baseline in Trough FEV <sub>1</sub> over Time in Study 360 .....	32
Figure 18. Mean Change from Baseline in Trough FEV <sub>1</sub> over Time in Study 374 .....	32
Figure 19. Empirical Distribution Function for Change from Baseline in Trough FEV <sub>1</sub> at 24 Weeks in Study 360 .....	33
Figure 20. Empirical Distribution Function for Change from Baseline in Trough FEV <sub>1</sub> at 24 Weeks in Study 374 .....	34

Figure 21. Mean Change from Baseline in Trough FEV <sub>1</sub> by Treatment Group over Time, Stratified by Dropout Pattern, Based on Integrated Data from Studies 361 and 373 .....	40
Figure 22. Proportion Suffering MACE over Time by Treatment Group Based on Data from the Primary Efficacy Studies .....	48
Figure 23. Proportion Suffering MACE over Time by Treatment Group Based on Data from all Phase 3 Studies .....	49
Figure 24. Proportion Suffering Adjudicated Cardiovascular Serious Adverse Events over Time by Treatment Group Based on Data from Primary Efficacy Studies.....	50
Figure 25. Proportion Suffering Adjudicated Cardiovascular Serious Adverse Events over Time by Treatment Group Based on Data from all Phase 3 Studies.....	51
Figure 26. Estimated Treatment Effect of UMEC/VI 62.5/25 on Mean Trough FEV <sub>1</sub> at 24 Weeks, Stratified by Different Subgroups, Based on Integrated Data from Studies 360, 361, 373, and 374 .....	53
Figure 27. Proportion Suffering a COPD Exacerbation over Time in Study 361 .....	62
Figure 28. Proportion Suffering a COPD Exacerbation over Time in Study 373 .....	63
Figure 29. Proportion Suffering a COPD Exacerbation over Time in Study 360 .....	64
Figure 30. Proportion Suffering a COPD Exacerbation over Time in Study 374 .....	65

## 1 EXECUTIVE SUMMARY

This review considers the inhaled umeclidinium and vilanterol combination product for long-term, once-daily, maintenance bronchodilator treatment of airflow obstruction in patients with chronic obstructive pulmonary disease (COPD), including chronic bronchitis and/or emphysema. Umeclidinium (UMEC), a long-acting muscarinic antagonist, and vilanterol (VI), a long-acting  $\beta_2$  agonist, are not currently marketed as single-ingredient products. Umeclidinium is a new molecular entity, and is currently under review as a monotherapy. Breo Ellipta, a related GlaxoSmithKline combination product containing vilanterol and the inhaled corticosteroid fluticasone furoate, was recently approved by the Agency.

We focus on the primary efficacy Studies 361, 373, 360, and 374, which were phase 3, multicenter, randomized, double-blind, parallel-group clinical trials designed to evaluate the 24-week efficacy of UMEC/VI for treatment of airflow obstruction in COPD. Patients in these studies had moderate-to-severe airflow obstruction, an extensive smoking history, and dyspnea. Concomitant use of systemic corticosteroids or additional long-acting bronchodilators was prohibited, but patients were permitted to use a stable dose of inhaled corticosteroids and study-provided salbutamol for as-needed relief medication.

In the only one of the four primary efficacy studies (Study 373) that included both placebo and UMEC/VI 62.5/25 mcg treatment arms, the combination product provided a statistically significant 0.167 L (95% confidence interval: 0.128, 0.207) improvement over placebo in the primary endpoint, 24-week mean change from baseline in trough FEV<sub>1</sub>. There was also independent, supportive evidence of a treatment effect on FEV<sub>1</sub> from a 12-week phase 3 cross-over study, from comparisons against the active comparator tiotropium, and from results for the higher 125/25 mcg dose. Estimated treatment effects for UMEC/VI 62.5/25 mcg were largely consistent across subgroups of interest, including sex, age, race, and geographic region.

We consider FEV<sub>1</sub> to be a surrogate endpoint, because it does not directly measure how a patient functions or feels in daily life, or how long a patient survives. The claim of effectiveness based on the primary analyses thus relies on the conclusion that the treatment effect on FEV<sub>1</sub> will reliably predict a treatment effect on a clinically meaningful endpoint. Therefore, we also gave importance to analyses of the following secondary endpoints that might be considered to directly measure how patients function or feel: COPD exacerbation, daily rescue medication use, St. George's Respiratory Questionnaire (SGRQ) score, and Shortness of Breath with Daily Activities (SOBDA) score. (b) (4)

observed trends toward benefit increase confidence that the treatment effect on the surrogate endpoint FEV<sub>1</sub> is likely to predict clinical benefit (i.e., improvements in how COPD patients function, feel, or survive).

There was evidence of the contribution of UMEC 62.5 mcg to the efficacy of the UMEC/VI 62.5/25 mcg combination product in Studies 360 and 373. There was evidence of the contribution of VI 25 mcg to the efficacy of the combination in only a single study (Study 373), but there was supportive evidence from independent, related study data. Supportive data included results from Study 361 for the higher 125/25 mcg dose, exploratory analyses from two phase 3 cross-over studies, and findings of efficacy relative to placebo for the VI monotherapy. There was also replicate evidence of efficacy, relative to placebo, for the UMEC monotherapy at the proposed 62.5 mcg dose.

There were substantial missing data in the primary efficacy studies, with dropout rates ranging from 15% to 33%. Dropouts on active treatment showed improvement in FEV<sub>1</sub> before withdrawal. Because the primary mixed effects model assumes that patient outcomes after withdrawal are missing at random,

model estimates rely on the assumption that the treatment effect observed before dropout persisted after patients stopped taking the therapy. This assumption is implausible because bronchodilators are generally considered symptomatic and not disease-modifying therapies, and any FEV<sub>1</sub> improvement attributable to a bronchodilator will likely go away within a few days of treatment discontinuation. Therefore, if the estimand of interest is the effectiveness of the assigned treatment in all randomized participants, at real world achievable adherence and tolerability, the mixed effects model used in the primary analysis likely does not provide a reliable estimate of the truth. As a result, we gave importance to a sensitivity analysis that multiply imputed missing data under the assumption that dropouts on all treatment arms would have had outcomes similar to those that were observed among completers in the control group. Statistical significance was maintained for all relevant treatment comparisons using this sensitivity analysis, but estimated magnitudes of treatment effect were approximately 20-30% smaller than those based on the primary analysis. For example, in Study 373, the estimated mean improvement in FEV<sub>1</sub> on UMEC/VI 62.5/25, relative to placebo, was 0.132 L (95% CI: 0.092, 0.173), as compared to 0.167 L (95% CI: 0.128, 0.207) in the primary analysis.

The complete safety evaluation was conducted by Dr. Jennifer Pippins, the Medical Reviewer, but we performed additional analyses to explore potential cardiovascular safety signals. Rates of major adverse cardiac events (MACE) were similar across the treatment arms, but an analysis of cardiovascular-related serious adverse events in the primary efficacy studies suggested a possible trend toward greater risk on the UMEC, VI, and UMEC/VI treatment arms, as compared to placebo and tiotropium. This imbalance in the rates of cardiovascular-related serious adverse events was not evident in analyses that included data from all of the phase 3 studies.

## 2 INTRODUCTION

### 2.1 Overview

#### 2.1.1 Background

Chronic obstructive pulmonary disease (COPD) is a common, progressive disease that causes symptoms such as coughing and shortness of breath, and increases risks of disability and death. Patients with COPD may have chronic bronchitis and/or emphysema. Chronic bronchitis is characterized by inflammation of the lining of bronchial tubes that leads to increased mucus formation and airflow obstruction. In emphysema, the air sacs (alveoli) at the end of the smallest airways (bronchioles) in the lung are damaged and the amount of gas exchange is reduced.

Medications used to treat patients with COPD include bronchodilators and/or steroids. Bronchodilators, usually administered through an inhaler, relax muscles around the airways to improve airflow and relieve symptoms. There are two major types of bronchodilators:  $\beta_2$  agonists, which act on  $\beta_2$  receptors, and muscarinic antagonists, which inhibit the action of cholinergic nerves. Bronchodilators may be short-acting or long-acting, and many have been approved by FDA for treatment of airflow obstruction in COPD. Approved bronchodilators include but are not limited to the short-acting  $\beta_2$  agonist salbutamol, short-acting muscarinic antagonist ipratropium, long-acting  $\beta_2$  agonists (LABAs) salmeterol and formoterol, and long-acting muscarinic antagonists (LAMAs) tiotropium and aclidinium bromide. FDA has also improved inhalers that combine a LABA and inhaled corticosteroid (ICS), such as Advair (salmeterol and fluticasone propionate) and Symbicort (formoterol and budesonide). No LAMA/LABA combination products have been approved by FDA.



This review considers the inhaled umeclidinium plus vilanterol combination product for long-term, once-daily, maintenance bronchodilator treatment of airflow obstruction in patients with COPD, including chronic bronchitis and emphysema. Umeclidinium (UMEC), a long-acting muscarinic antagonist, and vilanterol, a long-acting  $\beta_2$  agonist, are not currently marketed as single-ingredient products (monotherapies). Two doses, UMEC/VI 62.5/25 mcg once daily and 125/25 mcg once daily, were evaluated in the phase 3 clinical development program, but only the lower 62.5/25 mcg dose is proposed for approval. We often omit the mcg unit when referring to doses of UMEC and VI in this review.

## **2.1.2 History of Drug Development**

The applicant submitted the results of seven phase 3 clinical trials to support the regulatory approval of UMEC/VI for treatment of airflow obstruction in patients with COPD. The clinical development program for UMEC/VI was introduced to the Division of Pulmonary, Allergy, and Rheumatology Products under IND 106,616. UMEC and VI were developed under INDs 104,479 and 74,696, respectively, and dose-ranging studies were conducted separately for the UMEC and VI monotherapies.

Neither component is currently marketed as a single-ingredient inhalation product. Umeclidinium is a new molecular entity, and is currently under review as a monotherapy (NDA 205-382). A related GlaxoSmithKline combination product, Breo Ellipta (vilanterol/fluticasone furoate), was recently approved by the Agency (under NDA 204-275). This combination LABA/ICS inhalation powder is indicated for treatment of airway obstruction and reduction of exacerbations in patients with COPD, and the proposed dose of vilanterol (25 mcg) is the same as that in the UMEC/VI combination product. The dose selection, safety, and effectiveness of VI were reviewed as part of the Breo Ellipta program.

We next summarize important meetings and correspondence with the applicant. An end-of-phase 2 meeting was held on October 29, 2010. FDA generally agreed with the two proposed phase 3 placebo-controlled clinical trial designs, but recommended further exploration of UMEC doses lower than 125 mcg. FDA also requested justification of trough forced expiratory volume in 1 second (FEV<sub>1</sub>) as the primary endpoint in the NDA submission, and noted that additional spirometric and non-spirometric outcomes would be evaluated during NDA review. It was also noted that only about 20-25% of the phase 3 study populations would come from North America, so generalizability of results to the United States would be a review issue.

FDA also sent comments to the applicant on December 17, 2010 regarding the proposed phase 3 study designs. The Division noted that replicate evidence of safety and efficacy was needed for each dose of the UMEC monotherapy, but that the proposed designs allowed comparisons of each dose against placebo only once. A preNDA meeting occurred on January 18, 2012. FDA expressed concern about dose selection because the results of the phase 3 trials would be needed to help determine the appropriate dose. The Division also stated that the clinical development program must provide replicate evidence of efficacy for both monotherapies, as well as substantial evidence of the efficacy and safety of UMEC/VI as compared to each monotherapy (and that this typically means replicate positive trials). FDA recommended that the applicant first submit an NDA for the UMEC monotherapy. Finally, it was noted that information regarding an active comparator is typically not included in a product label unless doing so is necessary to support the proposed use in the intended population.

Several meetings occurred between 2006 and 2010 to discuss the applicant's development of the Shortness of Breath with Daily Activities Questionnaire (SOBDA) as a patient-reported outcome measure of dyspnea. The meetings included participants from the FDA Study Endpoints and Label Development (SEALD) team. At these meetings, FDA provided feedback on the development of the questionnaire. In

(b) (4)

(b) (4)

(b) (4)

FDA submitted an information request to the applicant on February 24, 2013 regarding the potential effect of missing data on the reliability of efficacy results. FDA requested additional sensitivity analyses that did not rely on the assumption that observed treatment effects before withdrawal would be preserved after patients stopped taking the therapy. The applicant responded with results based on two additional sensitivity analyses (see 3.2.2 for more details).

### 2.1.3 Specific Studies Reviewed

This review focuses on four phase 3 randomized clinical trials designed to evaluate the 24-week efficacy of UMEC/VI for treatment of airflow obstruction in COPD. Studies DB2113361 (361) and DB2113373 (373) were 24-week, randomized, double-blind, parallel-group, placebo-controlled clinical trials. Studies DB2113360 (360) and DB2113374 (374) were 24-week, randomized, double-blind, parallel-group, active-controlled clinical trials in which the LAMA tiotropium was the active treatment for comparison. Studies 361, 373, 360, and 374 will be referred to as the *primary efficacy studies*.

We also discuss results from three additional phase 3 studies of the combination product. Studies DB2114417 (417) and DB2114418 (418) were 12-week, randomized, double-blind, placebo-controlled, incomplete block, cross-over clinical trials to evaluate the efficacy of UMEC/VI with respect to both exercise endurance and lung function. Study DB2113359 (359) was a 52-week, randomized, double-blind, parallel-group, placebo-controlled trial to evaluate the safety and tolerability of UMEC/VI.

Finally, we briefly comment on several studies used to support the dose selection of umeclidinium and vilanterol. Studies AC4113589 (589), AC4113073 (73), AC4115321 (321), and AC4115408 (408) were randomized, double-blind, placebo-controlled trials that evaluated multiple doses of the UMEC monotherapy. Studies B2C111045 (45), HZA113310 (310), and B2C109575 (575) were randomized, double-blind, placebo-controlled, dose-ranging trials for the VI monotherapy.

## 2.2 Data Sources

Data were submitted by the applicant to the CDER electronic data room in SAS transport format. Protocols, correspondence, data listings, and study reports were accessed under the network path <\\CDSESUB1\EVSPROD\NDA203975\203975.enx>

## 3 STATISTICAL EVALUATION

### 3.1 Data and Analysis Quality

The submitted datasets were of acceptable quality and were adequately documented. We were able to reproduce the results of all key primary and secondary analyses.

### 3.2 Evaluation of Efficacy

#### 3.2.1 Study Design

##### 3.2.1.1 Primary Efficacy Studies

Studies 361, 373, 360, and 374 were designed to evaluate the 24-week efficacy of the once-daily bronchodilator UMEC/VI for treatment of airflow obstruction in COPD. The four studies were largely similar in design, with the exception of the treatment arms included. All were phase 3, multicenter, randomized, double-blind, parallel-group clinical trials in COPD patients with an extensive smoking history ( $\geq 10$  pack-years), moderate-to-severe airflow obstruction (percent predicted  $FEV_1 \leq 70\%$  and  $FEV_1/FVC < 0.7$  post-salbutamol), and dyspnea (score of  $\geq 2$  on the Modified Medical Research Council Dyspnea Scale). Concomitant use of systemic corticosteroids or additional long-acting bronchodilators was prohibited, but patients were permitted to use inhaled corticosteroids at a stable dose  $\leq 1000$ mg/day and study-provided salbutamol for as-needed relief medication.

There was a 1- to 2-week run-in period, followed by a 24-week double-blind treatment period. Visits occurred at Days 1 and 2, Weeks 4, 8, 12, 16, and 24, and 1 day after Week 24 (Day 169). All patients provided serial  $FEV_1$  measurements at 15 and 30 minutes, and 1, 3, 6, 23, and 24 hours after dosing on Day 1 and Week 24, and at 15 and 30 minutes, and 1, 3, and 6 hours after dosing on Weeks 4 and 12. A subset of about 200 patients in each of Studies 361 and 373 provided more comprehensive 24-hour serial spirometry assessments, as well as 24-hour Holter monitoring, at Day 1, and Weeks 12 and 24.

Withdrawal *from the treatment* was equivalent to withdrawal *from the study* because patients who stopped taking the therapy early were not followed up for safety and efficacy assessment for the remainder of the 24-week treatment period. The protocol categorized primary reasons for early withdrawal from the study as follows: adverse event, withdrawal of consent, loss to follow-up, protocol deviation, lack of efficacy (e.g., COPD exacerbation), protocol-defined stopping criteria, and study termination. The many potential reasons for stopping treatment, combined with the fact that the applicant did not continue to collect information on patients who stopped therapy early, led to substantial missing data in efficacy and safety analyses (see 3.2.4.4 for further discussion). Patients who stopped treatment early were scheduled for an early withdrawal visit soon thereafter, but pulmonary function assessments were not performed.

The primary endpoint was change from baseline in predose trough  $FEV_1$  on Day 169, where trough  $FEV_1$  was defined as the mean of values obtained 23 and 24 hours after the dose of study treatment administered on Day 168 (Week 24). The lone secondary endpoint was the weighted mean  $FEV_1$  0–6 hours postdose on Day 168. The weighted mean is time-weighted, calculated by dividing the area under the 0–6 hour postdose  $FEV_1$  curve (using measurements at baseline, 15 and 30 minutes, 1, 3, and 6 hours, and the trapezoidal rule) by the time of observation. Mean SOBDA score on Week 24 was specified as a secondary endpoint in the original protocol, but was later changed to an “Other Efficacy Endpoint.”

Additional endpoints included trough and weighted mean FEV<sub>1</sub> at earlier time points, mean Transition Dyspnea Index (TDI) focal score, rescue salbutamol use, time to first COPD exacerbation, and several other spirometric outcomes.

Studies 361 and 373 were placebo-controlled trials with 3:3:3:2 randomization to UMEC/VI, UMEC, VI, and placebo, respectively. Different doses of UMEC were used in these two studies: 125 mcg in Study 361 and 62.5 mcg in Study 373 (for both the UMEC/VI combination and the UMEC monotherapy). In each study, a total sample size of 1463 patients was planned to provide 90% power to detect a 58 mL mean difference between the combination and either monotherapy, or a 68 mL difference between any active treatment and placebo (presuming about 30% missing data).

Studies 360 and 374 were tiotropium-controlled trials with 1:1:1:1 randomization to one of four treatment arms. In Study 360, the treatment arms were UMEC/VI 125/25 mcg, UMEC 62.5/25 mcg, VI 25 mcg, and tiotropium. In Study 374, the treatment arms were UMEC/VI 125/25 mcg, UMEC 62.5/25 mcg, UMEC 125 mcg, and tiotropium. In each study, a total sample size of 832 patients was planned to provide 98% power to detect a 100 mL mean difference between two of the treatment groups (presuming about 30% missing data).

### 3.2.1.2 Additional Studies

Studies 417 and 418 were phase 3, multicenter, randomized, double-blind, placebo-controlled, 2-period (12 weeks per period), incomplete block, cross-over clinical trials to evaluate the efficacy of UMEC/VI and its components with respect to both exercise endurance and lung function. The studies were identical in design and conducted in COPD patients with an extensive smoking history, moderate-to-severe airflow obstruction and dyspnea, and lung hyperinflation (resting functional residual capacity (FRC)  $\geq 120\%$  of predicted normal). Concomitant use of systemic corticosteroids or additional long-acting bronchodilators was prohibited, but patients were permitted to use inhaled corticosteroids at a stable dose  $\leq 1000\text{mg/day}$  and study-provided salbutamol for as-needed relief medication. A sample size of 312 was planned, and subjects were randomized to receive a sequence consisting of two of the following treatments: UMEC/VI 62.5/25, UMEC/VI 125/25, UMEC 62.5, UMEC 125, VI 25, or placebo. The studies consisted of a 12- to 21-day run-in period, followed by two 12-week treatment periods that were separated by a 2-week washout period. As in the other phase 3 studies, patients who stopped treatment early were also withdrawn from the study, and a number of reasons for early withdrawal were listed in the protocol (e.g., adverse event and lack of efficacy).

The co-primary endpoints were change from period baseline in exercise endurance time (EET) and trough FEV<sub>1</sub> at 12 weeks. The term co-primary indicates that statistical significance (at the typical two-sided 5% level) needed to be achieved on *both* endpoints for the trial to be considered positive. Trough FEV<sub>1</sub> at 12 weeks was defined as the value obtained 24 hours after dosing on Day 84, and EET was measured 3 hours postdose on Day 84 using the endurance shuttle walk test (ESWT). The incremental shuttle walk test (ISWT) was performed during the run-in and washout periods to determine the walking speed at which to conduct the ESWT in each patient during the subsequent treatment period. Secondary efficacy endpoints included measures of lung volume (inspiratory capacity, functional residual capacity, residual volume), and 3-hour postdose FEV<sub>1</sub> at Week 12.

Study 359 was a randomized, double-blind, parallel-group, placebo-controlled, 52-week clinical trial to evaluate the safety and tolerability of UMEC/VI. The plan was for a total sample size of 500 subjects to be randomized 2:2:1 to UMEC/VI 125/25, UMEC 125, or placebo. The primary objective was to evaluate safety, so no primary efficacy endpoints were specified, although spirometry measurements were

obtained at randomization and Months 1, 3, 6, 9, and 12. As in the primary efficacy studies, the protocol listed several reasons for a patient to withdraw from the study. Possible reasons included adverse event, lack of efficacy, and protocol-defined stopping criteria based on electrocardiogram (ECG), Holter, or other laboratory abnormalities.

Studies 589, 73, 321, and 408 were randomized, double-blind, placebo-controlled trials evaluating multiple doses of the UMEC monotherapy in COPD. Study 589 was a 28-day parallel-group trial, Study 73 was an incomplete block, 3-period (14 days per period) cross-over trial, Study 321 was an incomplete block, 3-period (7 days per period) cross-over trial, and Study 408 was a 12-week, phase 3, parallel-group trial. Doses of UMEC in these trials ranged from 15.6 to 1000 mcg once daily, with some intermediate twice-daily doses evaluated as well.

Studies 45, 310, and 575 were randomized, double-blind, placebo-controlled, dose-ranging trials for the VI monotherapy. Study 45 was a 28-day parallel-group trial in COPD, Study 310 was a 5-period (7 days per period) cross-over trial in asthma, and Study 575 was a 28-day parallel-group trial in asthma. Doses of VI ranged from 3 mcg to 50 mcg once daily, with a 6.25 mcg twice-daily dose evaluated as well.

### **3.2.2 Statistical Methodologies**

#### **3.2.2.1 Primary Efficacy Studies**

In Studies 361, 373, 360, and 374, the primary efficacy analysis was based on a mixed effects model for repeated measures (MMRM) to compare treatment groups with respect to the mean change from baseline in trough FEV<sub>1</sub> at Day 169. The model used FEV<sub>1</sub> measurements at Days 2, 28, 56, 84, 112, 168, and 169, and included the following covariates: treatment group, baseline FEV<sub>1</sub>, center group, smoking status, visit (categorical variable), visit-by-baseline FEV<sub>1</sub> interaction, and visit-by-treatment group interaction. Variance estimation was based on an unstructured covariance matrix, which does not presume a particular correlation structure for repeated FEV<sub>1</sub> measurements within patients over time. The MMRM model has important assumptions, including constant variance, normality of errors, and normality of random intercepts. Residuals plots suggested some departures from constant variance and normality. Therefore, we also fit simple linear regression models (using only baseline and Day 169 data) to estimate treatment effects, with adjustment for baseline FEV<sub>1</sub>, center group, and smoking status, and the use of robust Huber-White standard errors. These analyses, which do not rely on assumptions of normality or constant variance, produced nearly identical estimates (results not shown) to the primary analyses.

The analysis of the secondary endpoint, weighted mean FEV<sub>1</sub> 0–6 hours postdose (assessed at Days 1, 28, 84, and 168), was based on the same mixed effects model for repeated measures as the primary analysis. Analyses of other continuous endpoints, such as SOBDA score, SGRQ score, and mean daily rescue medication use, were based on analogous MMRM models. The treatment effect on time to first exacerbation was evaluated using a Cox proportional hazards model adjusting for smoking status and center group, with the exact method to handle ties. Analyses of binary endpoints (e.g. proportion of responders based on some threshold change) were based on logistic regression models adjusting for baseline value, smoking status, and center group.

The applicant used sequential step-down closed testing procedures to control the false positive rate across the multiple comparisons in each study. In Studies 361 and 373, the following treatment comparisons were performed in order: (1) UMEC/VI versus placebo; (2) UMEC versus placebo; (3) VI versus placebo; (4) UMEC/VI versus VI; and (5) UMEC/VI versus UMEC. These analyses were performed first

for the primary endpoint (trough FEV<sub>1</sub>), and then for the secondary endpoint (weighted mean FEV<sub>1</sub>). In Studies 360 and 374, the following treatment comparisons were performed in order: (1) UMEC/VI versus tiotropium; (2) UMEC/VI versus vilanterol (Study 360) or UMEC (Study 374). Analyses were performed first for the high dose (UMEC/VI 125/25) with respect to both the primary and secondary endpoints, and then for the low dose (UMEC/VI 62.5/25). The applicant did not control for multiplicity across other efficacy endpoints (e.g., SGRQ, SOBDA, rescue medication use, exacerbation rate) in any of the studies. Efficacy analyses for Study 360 excluded one study center (with 20 subjects) at which the applicant identified significant deviations from Good Clinical Practice (GCP). Results were similar when the center was included.

The applicant performed a number of prespecified sensitivity analyses based on multiple imputation to explore the potential effect of missing data. The applicant's Missing at Random (MAR) approach assumes that data are missing at random and bases multiple imputation on mean and covariance estimation performed separately within each treatment arm. The Copy Differences from Control (CDC) approach assumes that changes over time in future outcomes in patients who withdraw from all treatment arms are similar to those future changes observed among completers in the control group. The Last Mean Carried Forward (LMCF) approach assumes that a constant mean trend over time (0 mL/year) or constant mean rate of decline (-25 mL/year), starting with the last observed value, would have occurred in all subjects following withdrawal. All imputation models used the same covariates to help estimate missing outcome data as were included in the primary MMRM model.

The underlying assumptions of these three imputation approaches are likely not scientifically plausible. If the estimand of interest is the effectiveness of the assigned treatment in all randomized participants, then the MAR, CDC, and LMCF approaches all essentially assume that any observed treatment effect before dropout would have persisted in patients, even after they stopped taking the therapy. This is unlikely, because bronchodilators are generally considered symptomatic and not disease-modifying therapies, and their effects on FEV<sub>1</sub> likely do not persist more than a few days after patients stop using them.

We find more merit in two additional sensitivity analyses provided by the applicant in response to an information request. Both the Copy Reference (CR) and Jump to Reference (J2R) approaches multiply impute missing data using estimated means in the control group. This is justifiable scientifically under the assumption that patients who stop taking the therapy will no longer benefit from it in the future, and thus will tend to have outcomes similar to those in the control group (in particular, the subset of control patients with similar baseline characteristics). The difference in the two methods is that the CR approach presumes patients who withdraw from the active arm were on the control (rather than the active) treatment before dropout; the resulting positive residuals before withdrawal leads to imputed values that slowly (rather than quickly) trend toward the estimated mean on the control arm. Given that the majority of withdrawals occurred at or before Day 112, and therefore at least two months before the primary efficacy assessment, we would expect any treatment effect observed before dropout to have gone completely away during the time following treatment discontinuation. Therefore, we focus on the Jump to Reference approach in assessing questions about the effectiveness of the treatments in all randomized participants (often called the intention-to-treat or de facto estimand). Although the scientific justification of the Jump to Reference approach seems reasonable, it is important to note that any such sensitivity analysis still relies on untestable assumptions about unobserved data. More information about the different multiple imputation models used by the applicant can be found at [www.missingdata.org.uk](http://www.missingdata.org.uk).

The scoring system used for the SOBDA Questionnaire was different than that proposed during its development. In particular, a response of "I did not do the activity" was scored as missing, and the daily mean score was calculated as the mean of the non-missing response scores (provided at least 7 of 13 scores were non-missing). Analyses of SOBDA were based on weekly mean scores (reported in patient

electronic diaries), which were considered non-missing if at least four of the seven days had non-missing daily mean scores. Also of note - the Week 24 mean score was defined as the mean of the daily scores occurring between Day 163 and either Day 169 or the day before the Day 168 visit, whichever came first. Because many Day 168 visits were scheduled a few days early, several patients did not have four days of SOBDA diary entries during Week 24. Therefore, this definition resulted in substantial missing data in SOBDA analyses.

The mean number of rescue medication puffs per day and percentage of rescue-free days over 24 weeks were considered non-missing if at least half of the daily electronic diary entries between Day 2 and Day 169 (or the day before the Day 169 visit) were non-missing. As a result, patients who completed regular daily diary entries for at least 12 weeks but dropped out early would still contribute data to analyses of rescue medication use over 24 weeks. These analyses will only reliably estimate mean differences in rescue medication use over 24 weeks if a patient's rescue medication use prior to dropout accurately reflects his or her rescue use after study withdrawal (and if data from patients without at least 12 weeks of diary entries are missing at random).

### **3.2.2.2 Additional Studies**

In Studies 417 and 418, the primary efficacy analyses were based on MMRM models to compare treatment groups with respect to the mean change from baseline in EET, and trough FEV<sub>1</sub>, at Week 12. For EET, the model used measurements at Day 2, and Weeks 6 and 12, and included the following covariates: treatment group, period baseline walking speed, mean baseline walking speed (mean of two period baseline speeds), period, center group, smoking status, visit (categorical variable), visit-by-mean walking speed interaction, and visit-by-treatment group interaction. Variance estimation was based on an unstructured covariance matrix. An analogous model was used for FEV<sub>1</sub>.

Comparisons of the two doses of the combination product (UMEC/VI 62.5/25 and UMEC/VI 125/25) against placebo were designated as primary, with a step-down testing procedure starting with the high dose comparisons (for both EET and FEV<sub>1</sub>) to account for multiplicity. Comparisons of the combination against placebo with respect to secondary efficacy endpoints, as well as comparisons of the monotherapies against placebo, and of the combination product against the monotherapies, were also of interest, but multiplicity was not controlled across these additional analyses. We performed additional supportive analyses in the subgroup of patients who completed both 12-week treatment periods.

Study 359 was a safety trial and therefore did not have prespecified primary efficacy analyses. However, exploratory efficacy analyses were conducted for trough FEV<sub>1</sub>, rescue puffs per day, and time to exacerbation using the same methods as in the four 24-week primary efficacy trials.

### **3.2.3 Dose Selection**

Separate dose-ranging studies were conducted for UMEC and VI to support the selection of the dose and dosing interval for both the monotherapies and UMEC/VI combination product in the primary efficacy studies. Data on the efficacy of different doses of UMEC are available from phase 2 Studies 321, 73, and 589, and from phase 3 Study 408 (Table 1). The results suggested no additional improvement in FEV<sub>1</sub> at doses greater than 125 mcg. In addition, adverse events were more common at doses of 250 mcg and above. The results suggest that the 62.5 and 125 mcg doses selected for phase 3 study were reasonable.

Study 321 was the only clinical trial that evaluated doses lower than 62.5 mcg. In the 6 hours postdose at Day 1, there was clear separation in efficacy between all once-daily UMEC doses and placebo (Figure 1). UMEC 15.6 mcg demonstrated the smallest improvement in FEV<sub>1</sub> and UMEC 125 mcg demonstrated the largest improvement, while the intermediate 31.25 and 62.5 mcg time-response profiles were largely overlapping. There was a similar dose-response pattern in the 24 hours postdose at Day 7 (Figure 2). At both Day 1 and Day 7, the time-response profile of the approved LAMA tiotropium was comparable to those of the once-daily 62.5 and 125 mcg UMEC doses selected for phase 3 study. These trends are also evident when examining mean differences in trough FEV<sub>1</sub> at Day 8 (Figure 3). Figure 4 and Figure 5 present time-response profiles at Days 1 and 7, respectively, for both once- and twice-daily doses of UMEC. Average FEV<sub>1</sub> improvements over time on twice-daily UMEC 15.6 and 31.25 mcg were similar to that of once-daily UMEC 62.5 mcg. This trend was also evident in comparisons of trough FEV<sub>1</sub> at Day 8 (Figure 6).

Data on the efficacy of different once-daily doses of VI are available from phase 2 Study 45 in COPD, as well as phase 2 Studies 575 and 310 in asthma (Table 2). These results were reviewed as part of the Breo Ellipta development program and are summarized in the reviews of that application, as well as in the label. In summary, results from Study 45 suggested greater benefit for 25 and 50 mcg VI, as compared to lower doses. Results from the asthma studies did not show a clear separation between the 12.5 and 25 mcg doses with respect to trough FEV<sub>1</sub>, but analyses of a number of secondary efficacy endpoints suggested some added benefit for the 25 mcg dose. Both Study 45 and Study 575 results suggested that a greater improvement in mean trough FEV<sub>1</sub> may be possible with VI 50 mcg, as compared to the selected 25 mcg dose.

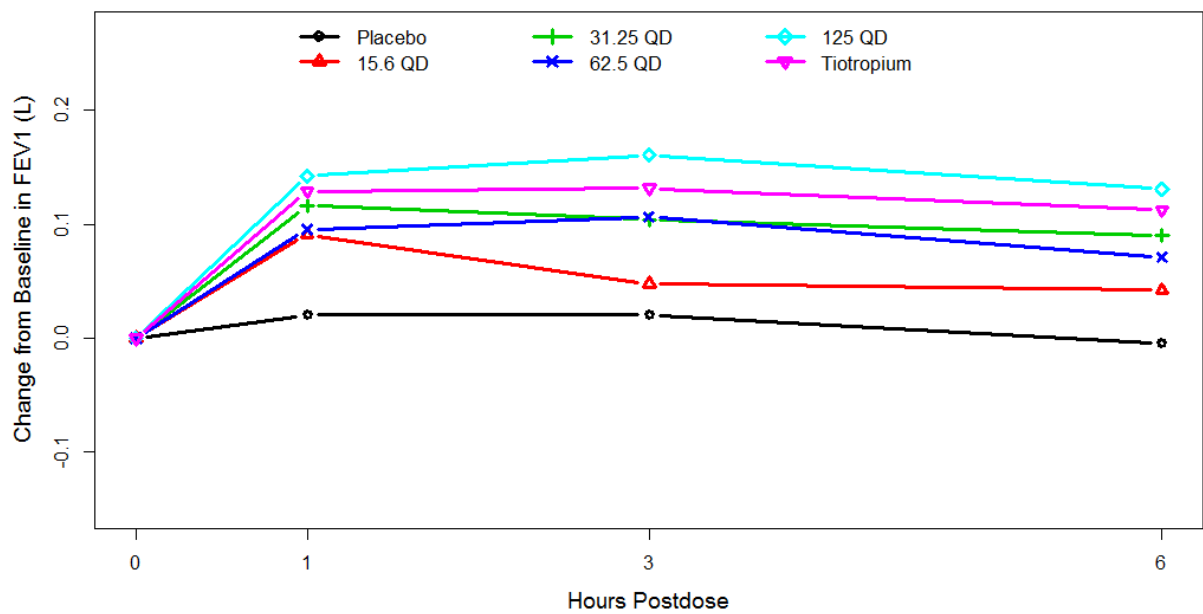
**Table 1. Differences from Placebo in Mean Change from Baseline in Trough FEV<sub>1</sub> for Different Once-Daily Doses of the Umeclidinium Monotherapy in Studies 321, 73, 589, and 408**

Study	Difference from Placebo for LS Mean Change from Baseline in Trough FEV <sub>1</sub> (L) (95% CI) by once daily UMEC dose (mcg) <sup>a</sup>						
	15.6	31.25	62.5	125	250	500	1000
AC4115321 at Day 8	0.113 (0.058, 0.168)	0.101 (0.045, 0.158)	0.124 (0.068, 0.179)	0.183 (0.127, 0.239)			
AC4113073 at Day 15			0.128 (0.060, 0.196)	0.147 (0.077, 0.216)	0.095 (0.027, 0.162)	0.140 (0.074, 0.205)	0.186 (0.113, 0.259)
AC4113589 at Day 29				0.159 (0.088, 0.229)	0.168 (0.099, 0.238)	0.150 (0.080, 0.220)	
AC4115408 at Day 85			0.127 (0.052, 0.202)	0.152 (0.076, 0.229)			

Source: Table 2, Clinical Overview

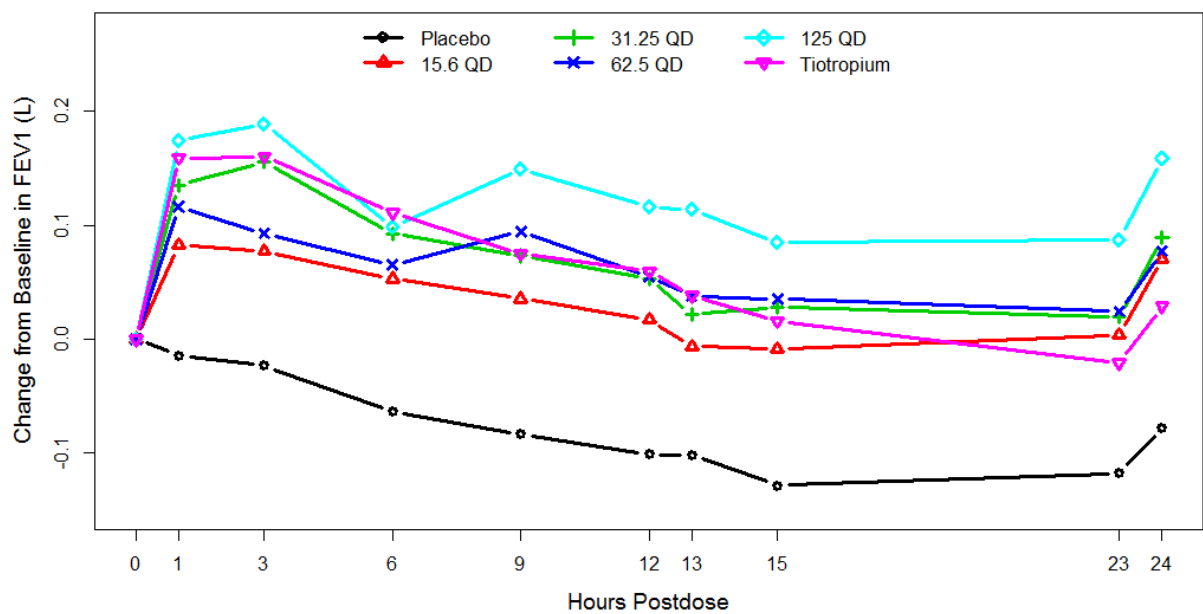


**Figure 1. Postdose 6-Hour Serial Mean Change from Baseline in FEV<sub>1</sub> at Day 1 for Different Once-Daily Umeclidinium Doses, and Tiotropium, in Study 321**



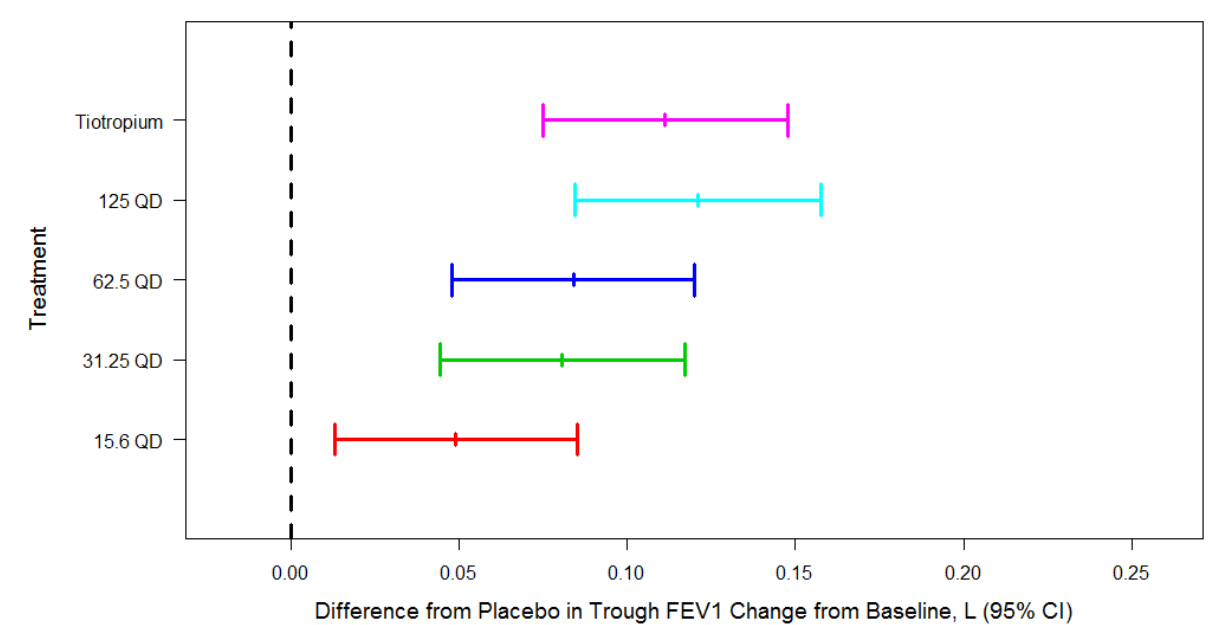
Abbreviations: QD = once-daily, BD = twice-daily

**Figure 2. Postdose 24-Hour Serial Mean Change from Baseline in FEV<sub>1</sub> at Day 7 for Different Once-Daily Umeclidinium Doses, and Tiotropium, in Study 321**



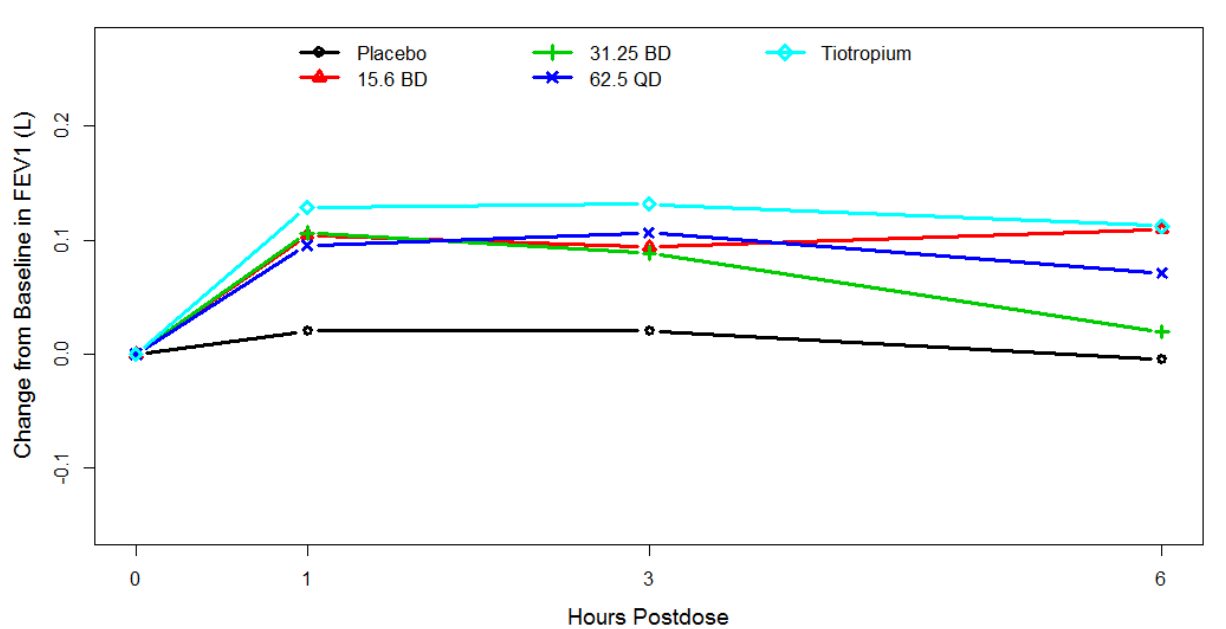
Abbreviations: QD = once-daily, BD = twice-daily

**Figure 3. Difference from Placebo in Mean Change from Baseline in Postdose 0-6 Hour Weighted Mean FEV<sub>1</sub> at Day 1 for Different Once-Daily Umeclidinium Doses, and Tiotropium, in Study 321**



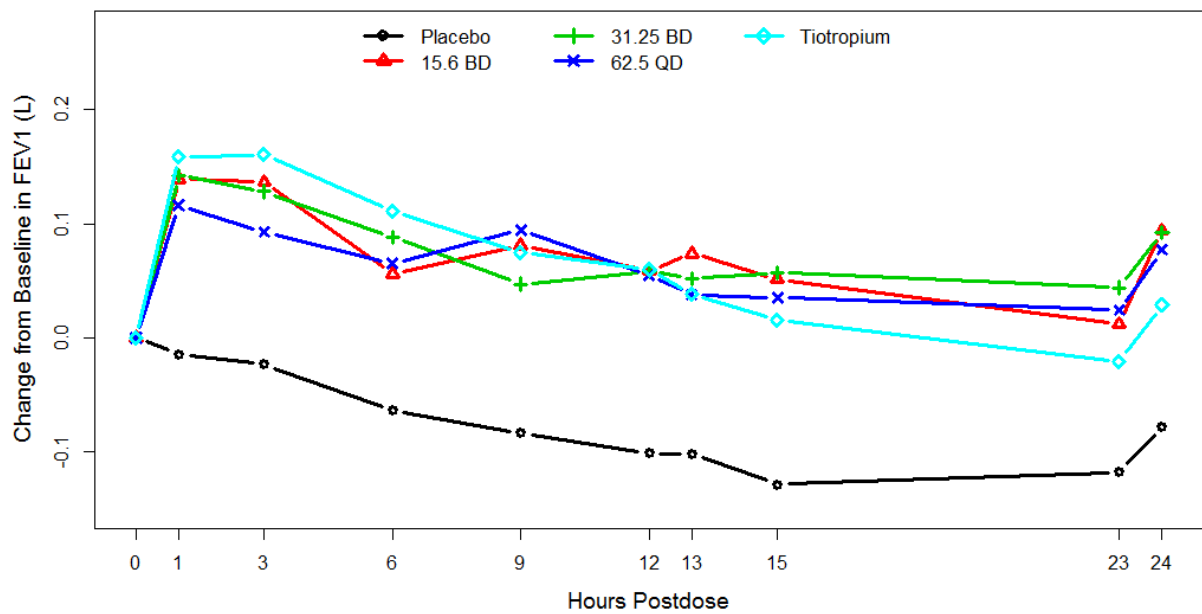
Estimates from mixed effects model adjusting for period baseline, mean period baseline, and period, with random subject effect  
 Abbreviations: QD = once-daily, BD = twice-daily

**Figure 4. Postdose 6-Hour Serial Mean Change from Baseline in FEV<sub>1</sub> at Day 1 for Different Once- and Twice-Daily Umeclidinium Doses, and Tiotropium, in Study 321**



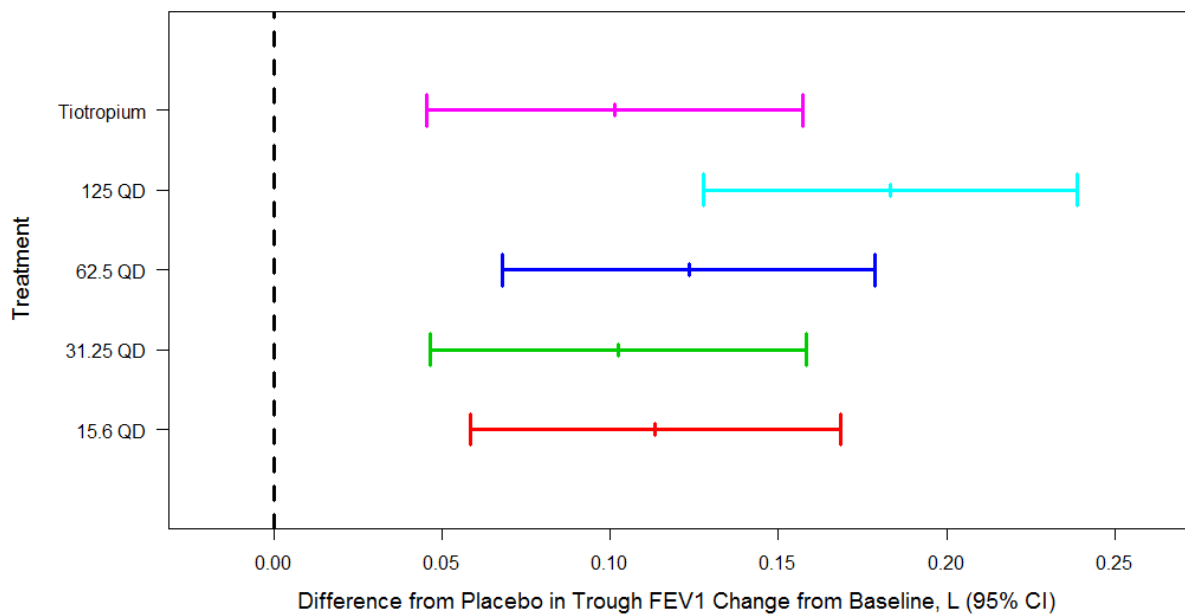
Abbreviations: QD = once-daily, BD = twice-daily

**Figure 5. Postdose 24-Hour Serial Mean Change from Baseline in FEV<sub>1</sub> at Day 7 for Different Once- and Twice-Daily Umeclidinium Doses, and Tiotropium, in Study 321**



Abbreviations: QD = once-daily, BD = twice-daily

**Figure 6. Difference from Placebo in Mean Change from Baseline in Trough FEV<sub>1</sub> at Day 8 for Different Once-Daily Umeclidinium Doses, and Tiotropium, in Study 321**



Estimates from mixed effects model adjusting for period baseline, mean period baseline, and period, with random subject effect  
Abbreviations: QD = once-daily, BD = twice-daily

**Table 2. Differences from Placebo in Mean Change from Baseline in Trough FEV<sub>1</sub> for Different Once-Daily Doses of the Vilanterol Monotherapy in Studies 45, 575, and 310**

	Difference from Placebo for LS Mean Change from Baseline in Trough FEV <sub>1</sub> (L) (95% CI) by once daily VI dose (mcg) <sup>a</sup>				
Study	3	6.25	12.5	25	50
<b>COPD</b>					
B2C111045 at Day 29	0.092 (0.039, 0.144)	0.098 (0.046, 0.150)	0.110 (0.057, 0.162)	0.137 (0.085, 0.190)	0.165 (0.112, 0.217)
<b>Asthma</b>					
B2C109575 at Day 28	0.064 (-0.036, 0.164)	0.069 (-0.029, 0.168)	0.130 (0.030, 0.230)	0.121 (0.023, 0.220)	0.162 (0.062, 0.261)
HZA113310 at Day 7		0.094 (0.049, 0.140)	0.102 (0.057, 0.147)	0.125 (0.080, 0.170)	

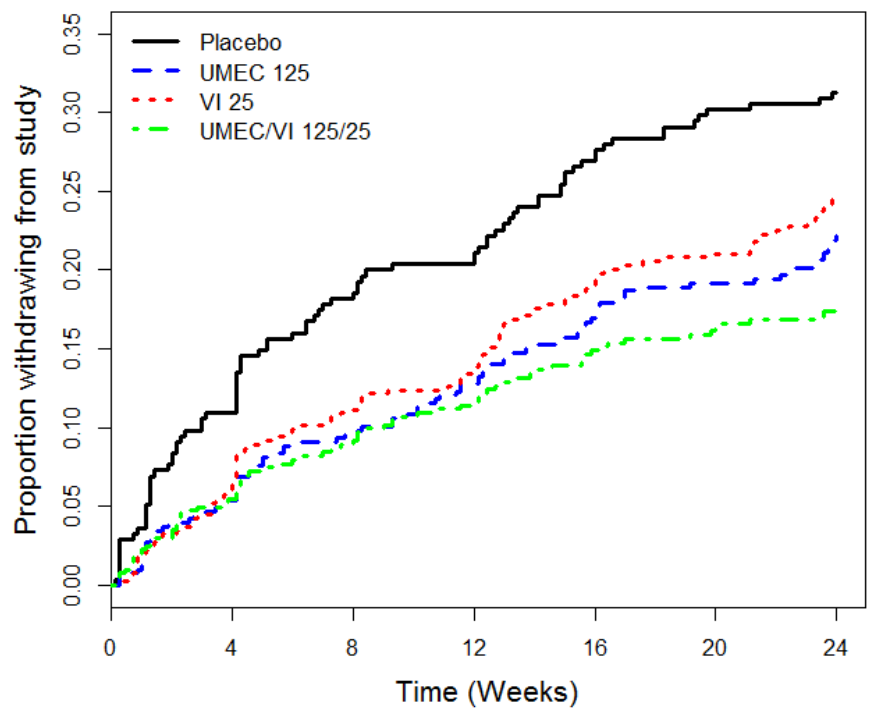
Source: Table 3, Clinical Overview

### 3.2.4 Patient Disposition, Demographic and Baseline Characteristics

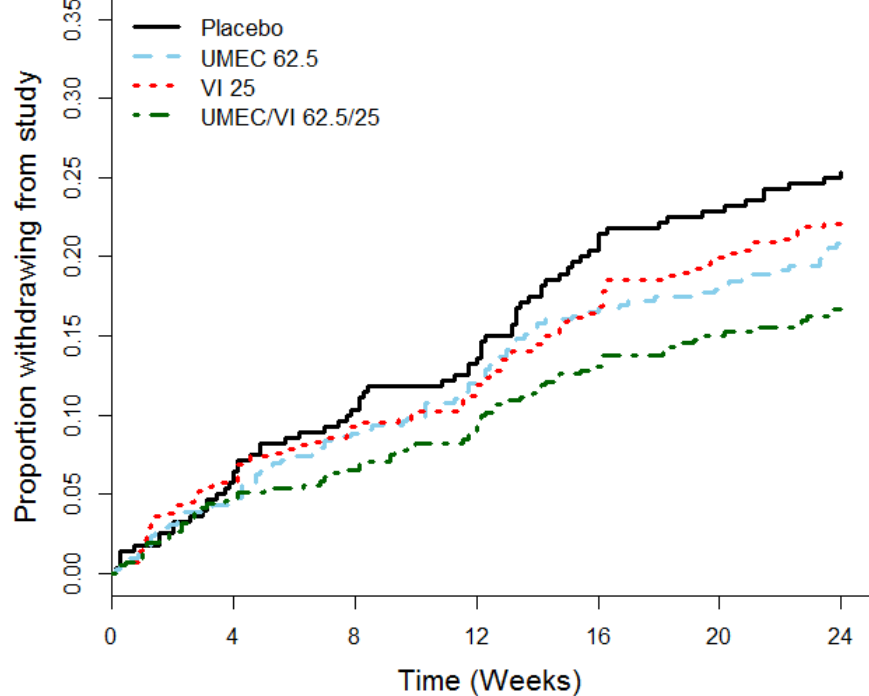
Baseline characteristics were similar across the primary efficacy Studies 361, 373, 360, and 374, which consisted of 1,489, 1,532, 843, and 869 patients, respectively (Appendix: Table 21, Table 22, Table 23, and Table 24). The combined study population was predominantly male (68%), White (84%), and older in age (mean 63 years). Twenty-five percent of patients were treated at U.S. sites. Only 3% and 9% of patients were Black and Asian, respectively. Within U.S. sites, 10% of patients were Black. Ninety percent of subjects had percent predicted FEV<sub>1</sub> 30–80%, 50% of subjects were current smokers, and 49% used inhaled corticosteroids. Patients were enrolled at 153, 163, 91, and 95 different centers from several countries around the world in Studies 361, 373, 360, and 374, respectively. There were no large imbalances in baseline characteristics across the treatment arms in the four studies.

As described previously, the design of the primary efficacy studies was such that subjects who stopped treatment early would also be withdrawn from the study. There were many prespecified reasons for withdrawal, such as adverse event, lack of efficacy (e.g., COPD exacerbation), and protocol deviation. As a result, there was substantial patient dropout. The proportions of subjects withdrawing from the four trials over time are displayed by treatment group in Figure 7, Figure 8, Figure 9, and Figure 10. In Studies 361, 373, 360, and 374, 25%, 23%, 17%, and 23% of patients failed to complete the 24-week treatment period, respectively. Dropout rates tended to be slightly higher on placebo than the other treatment arms in Studies 361 and 373, with the differences primarily attributable to greater placebo dropout because of lack of efficacy (Table 3 and Table 4). The most common reasons for study withdrawal across all four studies were adverse event, lack of efficacy, protocol-defined stopping criteria, and withdrawal of consent.

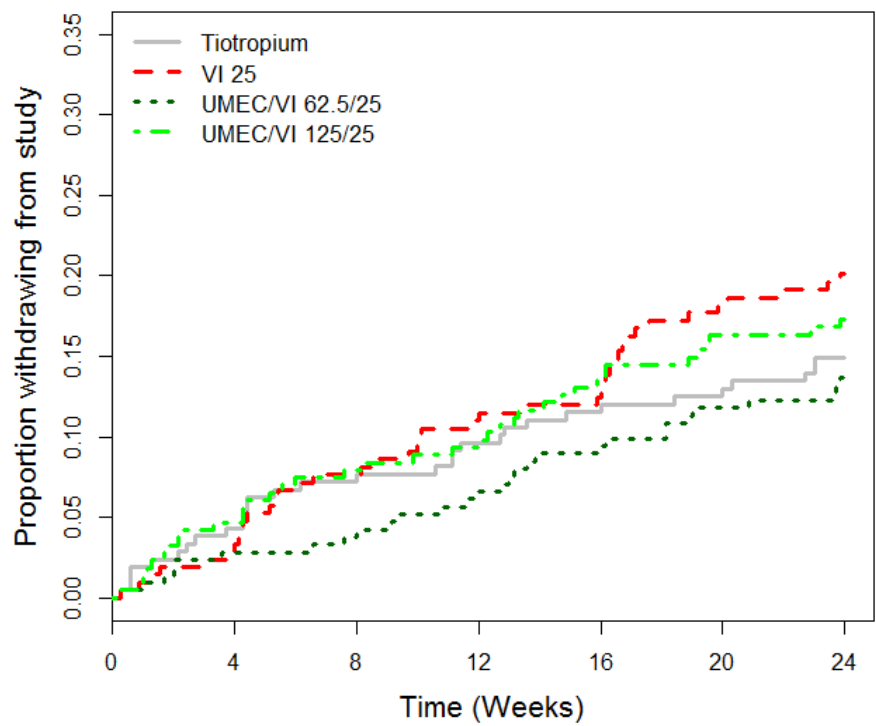
**Figure 7. Proportion of Patients Withdrawing Early over Time in Study 361**



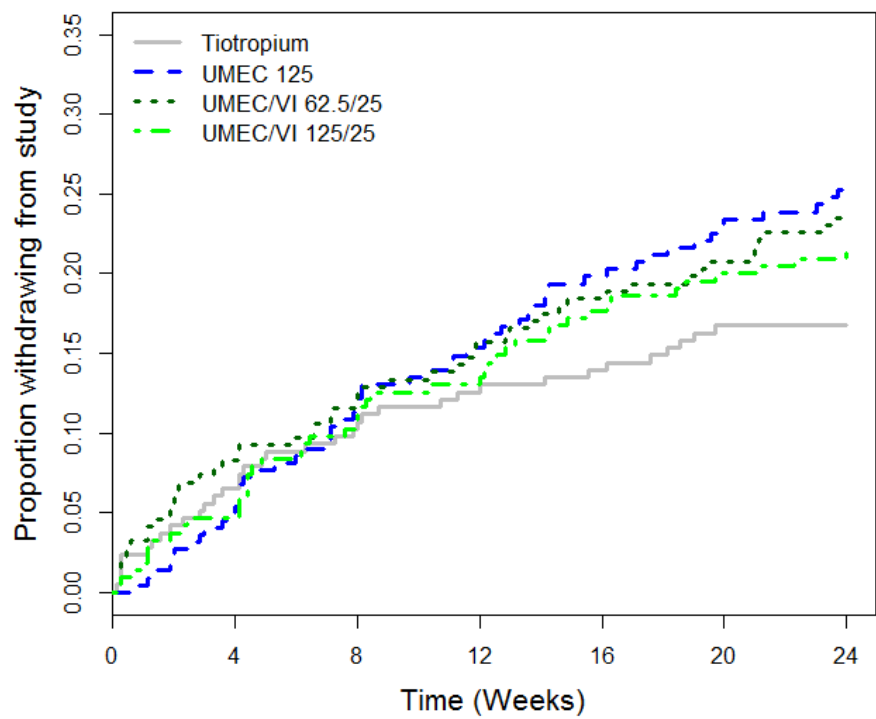
**Figure 8. Proportion of Patients Withdrawing Early over Time in Study 373**



**Figure 9. Proportion of Patients Withdrawing Early over Time in Study 360**



**Figure 10. Proportion of Patients Withdrawing Early over Time in Study 374**



**Table 3. Proportion of Patients Failing to Complete Study 361, by Reason for Withdrawal**

	Placebo	UMEC 125	VI 25	UMEC/VI 125/25	Overall
Completed study	183 (67%)	312 (77%)	298 (74%)	325 (81%)	1118 (75%)
Did not complete study	92 (33%)	95 (23%)	106 (26%)	78 (19%)	371 (25%)
Adverse event	17 (6%)	24 (6%)	25 (6%)	18 (4%)	84 (6%)
Lack of efficacy	44 (16%)	38 (9%)	37 (9%)	24 (6%)	143 (10%)
Lost to follow-up	0 (0%)	2 (0%)	1 (0%)	3 (1%)	6 (0%)
Protocol deviation	4 (1%)	3 (1%)	11 (3%)	5 (1%)	23 (2%)
Protocol-defined stopping criteria	16 (6%)	15 (4%)	14 (3%)	13 (3%)	58 (4%)
Withdrew consent	11 (4%)	13 (3%)	18 (4%)	15 (4%)	57 (4%)

**Table 4. Proportion of Patients Failing to Complete Study 373, by Reason for Withdrawal**

	Placebo	UMEC 62.5	VI 25	UMEC/VI 62.5/25	Overall
Completed study	204 (73%)	324 (78%)	318 (76%)	332 (80%)	1178 (77%)
Did not complete study	76 (27%)	94 (22%)	103 (24%)	81 (20%)	354 (23%)
Adverse event	9 (3%)	34 (8%)	24 (6%)	23 (6%)	90 (6%)
Lack of efficacy	37 (13%)	20 (5%)	32 (8%)	20 (5%)	109 (7%)
Lost to follow-up	1 (0%)	0 (0%)	3 (1%)	2 (0%)	6 (0%)
Protocol deviation	4 (1%)	7 (2%)	5 (1%)	6 (1%)	22 (1%)
Protocol-defined stopping criteria	9 (3%)	13 (3%)	24 (6%)	15 (4%)	61 (4%)
Withdrew consent	16 (6%)	20 (5%)	15 (4%)	15 (4%)	66 (4%)

**Table 5. Proportion of Patients Failing to Complete Study 360, by Reason for Withdrawal**

	Tiotropium	VI 25	UMEC/VI 62.5/25	UMEC/VI 125/25	Overall
Completed study	177 (85%)	165 (79%)	181 (85%)	173 (81%)	696 (83%)
Did not complete study	31 (15%)	44 (21%)	31 (15%)	41 (19%)	147 (17%)
Adverse event	9 (4%)	10 (5%)	10 (5%)	15 (7%)	44 (5%)
Lack of efficacy	7 (3%)	17 (8%)	9 (4%)	5 (2%)	38 (5%)
Lost to follow-up	1 (0%)	1 (0%)	0 (0%)	1 (0%)	3 (0%)
Protocol deviation	0 (0%)	7 (3%)	1 (0%)	4 (2%)	12 (1%)
Protocol-defined stopping criteria	5 (2%)	2 (1%)	3 (1%)	10 (5%)	20 (2%)
Withdrew consent	9 (4%)	7 (3%)	8 (4%)	6 (3%)	30 (4%)

**Table 6. Proportion of Patients Failing to Complete Study 374, by Reason for Withdrawal**

	<b>Tiotropium</b>	<b>UMEC 125</b>	<b>UMEC/VI 62.5/25</b>	<b>UMEC/VI 125/25</b>	<b>Overall</b>
Completed study	176 (82%)	165 (74%)	163 (75%)	166 (77%)	670 (77%)
Did not complete study	39 (18%)	57 (26%)	54 (25%)	49 (23%)	199 (23%)
Adverse event	11 (5%)	17 (8%)	20 (9%)	15 (7%)	63 (7%)
Lack of efficacy	13 (6%)	22 (10%)	12 (6%)	9 (4%)	56 (6%)
Lost to follow-up	2 (1%)	0 (0%)	1 (0%)	0 (0%)	3 (0%)
Protocol deviation	1 (0%)	1 (0%)	4 (2%)	4 (2%)	10 (1%)
Protocol-defined stopping criteria	6 (3%)	7 (3%)	8 (4%)	11 (5%)	32 (4%)
Withdrew consent	6 (3%)	10 (5%)	9 (4%)	10 (5%)	35 (4%)

Baseline characteristics in the cross-over Studies 417 and 418, and the long-term safety Study 359, were largely similar to those of the primary efficacy studies. One notable difference was that all patients in Studies 417 and 418 had lung hyperinflation, resulting in mean percent predicted normal FRC values of 153.6% and 151.6%, respectively. There were no noticeable imbalances in baseline characteristics across the randomized treatment arms in these three studies.

There was substantial patient dropout in Studies 417, 418, and 359. In Study 417, 95 (27%) of the 348 randomized subjects failed to remain in the study through both 12-week treatment periods. In Study 418, 96 (31%) of the 307 randomized subjects failed to do so. The most common reasons for dropout were adverse event and lack of efficacy. In Study 359, 220 (39%) of the 562 randomized subjects did not complete the 52-week study (Table 7). Dropout rates overall were similar across the placebo, UMEC 125, and UMEC/VI 125/25 treatment arms. There was greater study withdrawal on placebo than the active arms for lack of efficacy and for adverse event, but greater withdrawal on the active arms because of protocol-defined stopping criteria. In particular, there was greater withdrawal on both UMEC 125 (16%) and UMEC/VI 125/25 (16%) than placebo (7%) because of either ECG or Holter abnormalities.

**Table 7. Proportion of Patients Failing to Complete Study 359, by Reason for Withdrawal**

	<b>Placebo</b>	<b>UMEC 125</b>	<b>UMEC/VI 125/25</b>	<b>Overall</b>
Completed study	66 (61%)	133 (59%)	143 (63%)	342 (61%)
Did not complete study	43 (39%)	94 (41%)	83 (37%)	220 (39%)
Adverse event	13 (12%)	21 (9%)	17 (8%)	51 (9%)
Lack of efficacy	9 (8%)	3 (1%)	1 (1%)	13 (2%)
Protocol-defined stopping criteria <sup>1</sup>	8 (7%)	37 (16%)	36 (16%)	81 (14%)
ECG abnormality	0 (0%)	12 (5%)	13 (6%)	25 (4%)
Holter abnormality	8 (7%)	26 (11%)	26 (12%)	60 (11%)
Lab abnormality	0 (0%)	1 (0%)	0 (0%)	1 (0%)
Other	13 (12%)	33 (15%)	29 (13%)	75 (13%)

<sup>1</sup> Patients who dropped out because of protocol-defined stopping criteria could have had more than one abnormality



### 3.2.5 Results and Conclusions

We largely focus on findings from the four primary efficacy studies (Studies 361, 373, 360, and 374). We discuss comparisons of UMEC/VI and its components against placebo in 3.2.5.1, the contribution of each component to the efficacy of the combination in 3.2.5.2, and comparisons against the approved LAMA tiotropium in 3.2.5.3. We evaluate the results of sensitivity analyses to explore the potential impact of missing data in 3.2.5.4, and summarize findings from the cross-over Studies 417 and 418 and the long-term safety Study 359 in 3.2.5.5.

#### 3.2.5.1 Primary Efficacy Studies: Placebo Comparisons

Data are available for treatment comparisons of UMEC/VI against placebo only from Studies 361 (at the 125/25 mcg dose) and 373 (at the 62.5/25 mcg dose). In each of these trials, treatment with the combination product resulted in a statistically significant, greater change from baseline in the mean trough FEV<sub>1</sub> at 24 weeks, as compared to placebo (Table 8). In Study 361, the estimated difference in mean trough FEV<sub>1</sub> change between UMEC/VI 125/25 and placebo was 0.238 L (95% confidence interval [CI]: 0.200, 0.276;  $p < 0.0001$ ). In Study 373, the estimated difference in mean trough FEV<sub>1</sub> change between UMEC/VI 62.5/25 and placebo was 0.167 L (95% CI: 0.128, 0.207;  $p < 0.0001$ ).

Observed effects of the combination product on trough FEV<sub>1</sub> were evident as early as Day 2 and then remained relatively constant over the 24-week treatment period (Figure 11 and Figure 12). There was also evidence of efficacy for UMEC/VI with respect to the secondary endpoint, 0–6 hour weighted mean FEV<sub>1</sub>. Mean differences in weighted mean FEV<sub>1</sub> between the experimental treatments and placebo in Studies 361 and 373 were slightly larger than the analogous trough FEV<sub>1</sub> comparisons, with strong statistical evidence against the null hypothesis of no treatment effect (Table 8). In addition, data from the subset of patients with 24-hour serial FEV<sub>1</sub> assessments demonstrated consistently higher mean FEV<sub>1</sub> levels with the combination product than placebo throughout the 24 hours postdose (Figure 13 and Figure 14). Finally, empirical distribution plots, in which dropouts were treated as the worst potential outcomes, suggested benefits of UMEC/VI treatment with respect to summary measures of the FEV<sub>1</sub> distribution besides the mean, such as the median (Figure 15 and Figure 16). These figures can also be used to descriptively compare treatment groups with respect to the proportion achieving certain threshold changes in FEV<sub>1</sub> at 24 weeks, such as an improvements of at least 0.1 or 0.2 L.

UMEC/VI also showed trends toward benefit for additional non-spirometric endpoints of interest, including mean changes from baseline in the SOBDA and SGRQ scores at 24 weeks, mean puffs of rescue medication per day, percent of rescue-free days, and exacerbation rate over 24 weeks (Table 9). In Study 373, 35 (13%) and 27 (7%) patients suffered a COPD exacerbation on placebo and UMEC/VI 62.5/25, respectively. The separation between the treatment groups in the proportions suffering an exacerbation over time was also evident in Kaplan Meier plots (Appendix: Figure 27 and Figure 28). The applicant did not control for multiplicity across these additional comparisons, but the strength of statistical evidence without adjustment was high ( $p < 0.005$  for all comparisons). Although these trends toward benefit may not be sufficient to support labeling claims, they do provide support for the observed treatment effect on the primary endpoint.

In summary, there was strong statistical evidence of beneficial effects of both UMEC/VI 62.5/25 and UMEC/VI 125/25, as compared to placebo, with respect to the primary and secondary endpoints, in addition to supportive trends across a range of other spirometric and non-spirometric endpoints of interest. However, it is important to note that there was evidence of superiority against placebo for the

proposed 62.5/25 mcg dose of the combination product from only one of the primary efficacy studies (Study 373; see 5.1 for further discussion).

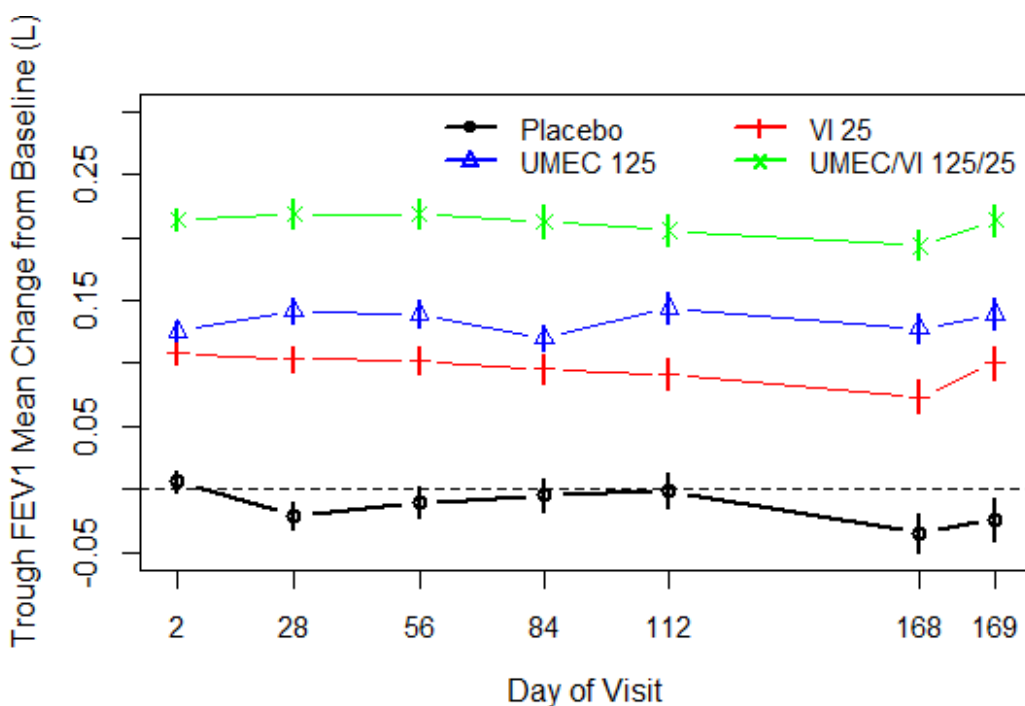
There was also statistical evidence of benefit for the umeclidinium and vilanterol monotherapies relative to placebo with respect to trough FEV<sub>1</sub> at 24 weeks (Table 8). The estimated difference in mean trough FEV<sub>1</sub> change between UMEC 125 and placebo was 0.160 L in Study 361, and the difference between UMEC 62.5 and placebo was 0.115 L in Study 373. There was replicate evidence of efficacy for UMEC 62.5 in phase 3 Study 408, where the estimated treatment effect on trough FEV<sub>1</sub> at 12 weeks was 0.127 L (95% CI: 0.052, 0.202; p<0.001). The estimated differences in mean trough FEV<sub>1</sub> change between VI 25 and placebo were 0.124 and 0.072 L in Studies 361 and 373, respectively. Efficacy of the monotherapies relative to placebo was also supported by comparisons of FEV<sub>1</sub> at earlier time points (Figure 11 and Figure 12), comparisons of 0–6 hour weighed mean FEV<sub>1</sub> (Table 8), and by comparisons of FEV<sub>1</sub> throughout the 24 hours postdose (Figure 13 and Figure 14). Evidence of benefits for the UMEC and VI monotherapies relative to placebo for additional non-spirometric endpoints of interest was generally not as strong as for the combination product, but estimates trended in the positive direction (Table 9).

**Table 8. Comparisons of UMEC, VI, and UMEC/VI against Placebo with Respect to the Mean Changes from Baseline in Trough FEV<sub>1</sub> (Primary Endpoint) and 0–6 Hour Weighted Mean FEV<sub>1</sub> (Secondary Endpoint) at 24 Weeks in Studies 361 and 373**

	Mean Change from Baseline in Trough FEV <sub>1</sub> (L)	Mean Difference <sup>1</sup> in Trough FEV <sub>1</sub> (L) (95% CI) p-value	Mean Change from Baseline in 0–6 Hour Weighted Mean FEV <sub>1</sub> (L)	Mean Difference <sup>1</sup> in 0–6 Hour Weighted Mean FEV <sub>1</sub> (L) (95% CI) p-value
<i>Study 361</i>				
Placebo	-0.031		-0.018	
UMEC 125	0.129	0.160 (0.122, 0.198) <0.0001	0.160	0.178 (0.141, 0.216) <0.0001
VI 25	0.093	0.124 (0.086, 0.162) <0.0001	0.127	0.145 (0.107, 0.182) <0.0001
UMEC/VI 125/25	0.207	0.238 (0.200, 0.276) <0.0001	0.269	0.287 (0.250, 0.324) <0.0001
<i>Study 373</i>				
Placebo	0.004		0.001	
UMEC 62.5	0.119	0.115 (0.076, 0.155) <0.0001	0.151	0.150 (0.110, 0.190) <0.0001
VI 25	0.076	0.072 (0.032, 0.112) 0.0004	0.123	0.122 (0.082, 0.162) <0.0001
UMEC/VI 62.5/25	0.171	0.167 (0.128, 0.207) <0.0001	0.243	0.242 (0.202, 0.282) <0.0001

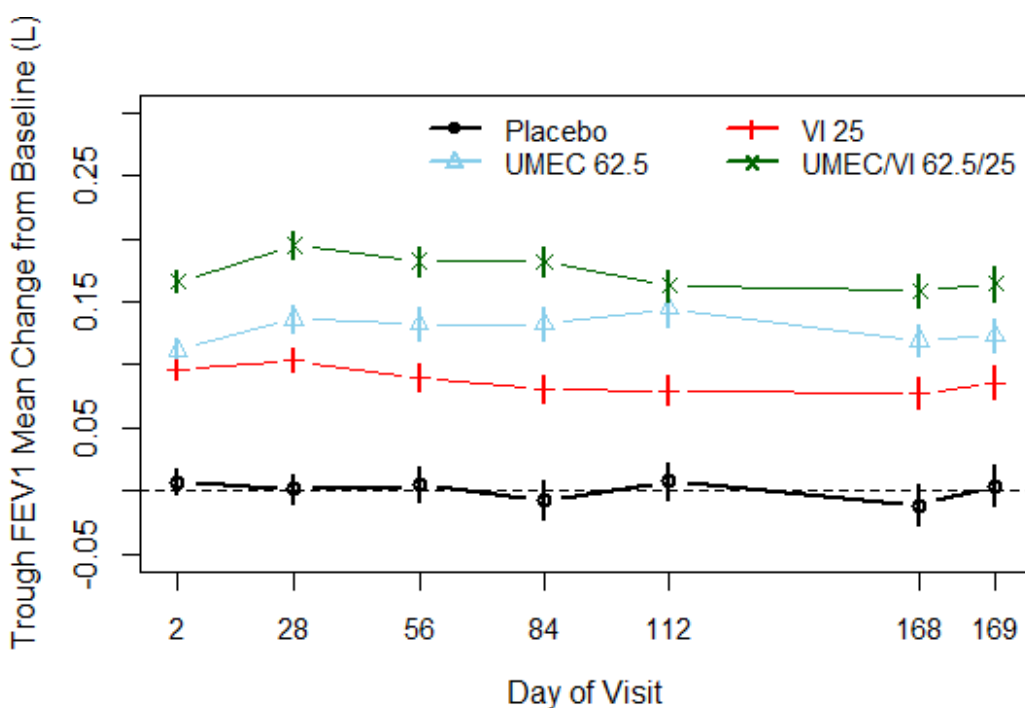
<sup>1</sup> Estimated differences, as compared to placebo, from linear mixed effects models with the following covariates: treatment group, baseline FEV<sub>1</sub>, center group, smoking status, visit, visit-by-baseline FEV<sub>1</sub> interaction, and visit-by-treatment group interaction

**Figure 11. Mean Change from Baseline in Trough FEV<sub>1</sub> over Time in Study 361**



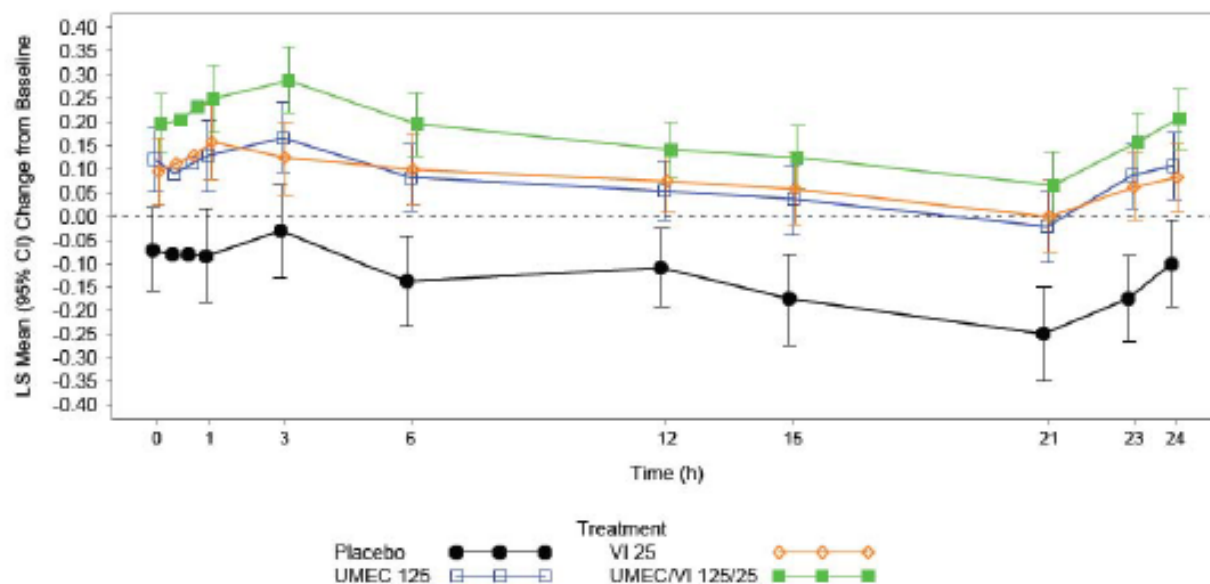
Unadjusted means based on observed data are displayed  
Error bars represent  $\pm 1$  standard error

**Figure 12. Mean Change from Baseline in Trough FEV<sub>1</sub> over Time in Study 373**



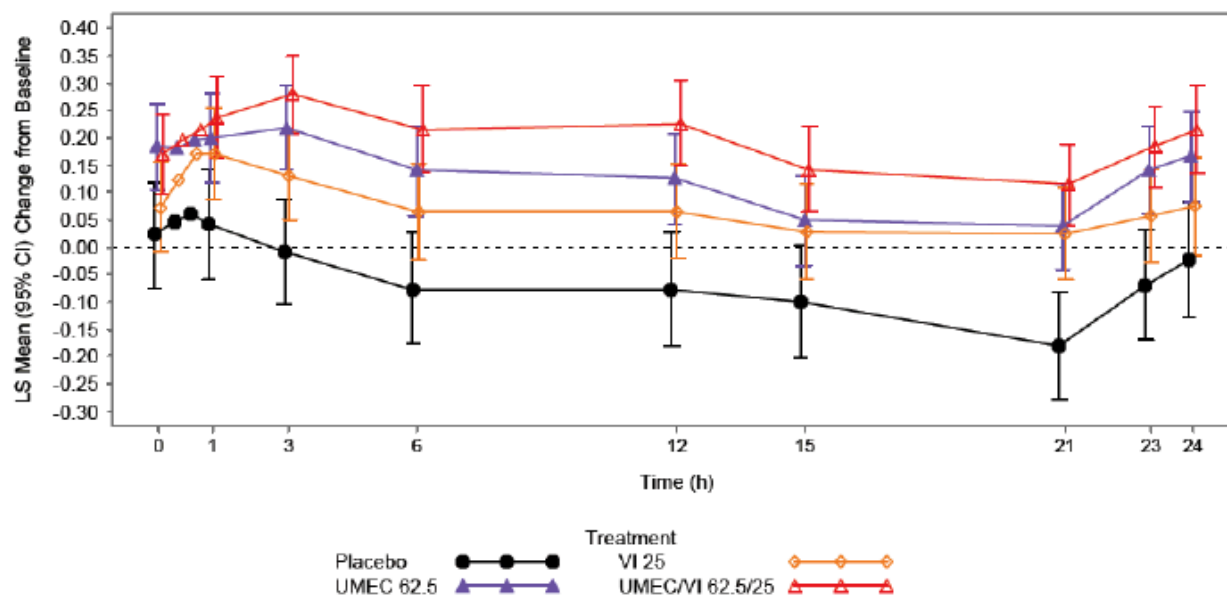
Unadjusted means based on observed data are displayed  
Error bars represent  $\pm 1$  standard error

**Figure 13. Mean Change from Baseline in Serial FEV<sub>1</sub> over 0 to 24 Hours Postdose at Day 168 in the Subset of Study 361 with 24-Hour measurements**



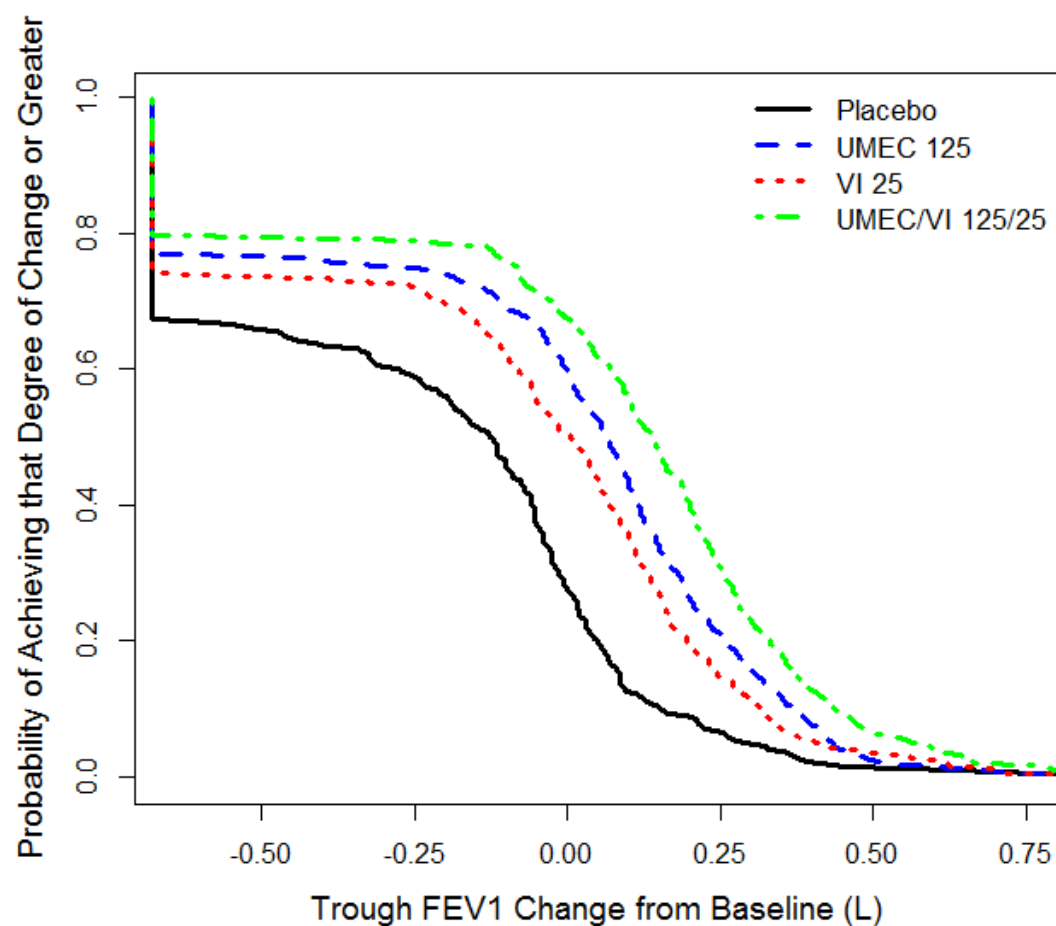
Source: Figure 15, Applicant's Clinical Study Report  
Error bars represent 95% confidence intervals for the mean

**Figure 14. Mean Change from Baseline in Serial FEV<sub>1</sub> over 0 to 24 Hours Postdose at Day 168 in the Subset of Study 373 with 24-Hour measurements**



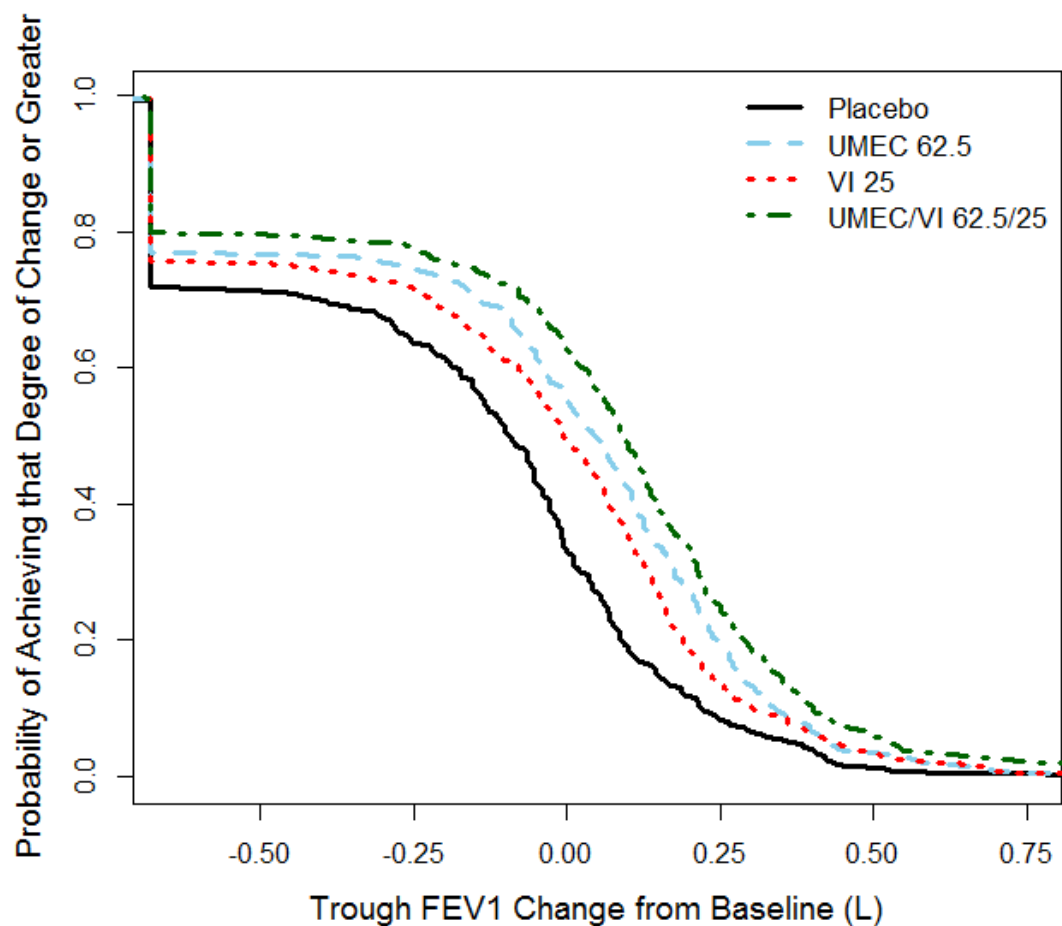
Source: Figure 15, Applicant's Clinical Study Report  
Error bars represent 95% confidence intervals for the mean

**Figure 15. Empirical Distribution Function for Change from Baseline in Trough FEV<sub>1</sub> at 24 Weeks in Study 361**



Plot displays one minus the empirical distribution function  
Early study withdrawal was considered the worst possible outcome

**Figure 16. Empirical Distribution Function for Change from Baseline in Trough FEV<sub>1</sub> at 24 Weeks in Study 373**



Plot displays one minus the empirical distribution function  
Early study withdrawal was considered the worst possible outcome

**Table 9. Comparisons against Placebo for Additional Supportive Endpoints in Studies 361 and 373**

	Mean SOBDA Score at 24 Weeks <sup>1</sup> (95% CI)	Mean SGRQ Score at 24 Weeks <sup>1</sup> (95% CI)	Mean Rescue Puffs per Day over 24 Weeks <sup>1</sup> (95% CI)	Mean Percent Rescue-Free Days over 24 Weeks <sup>1</sup> (95% CI)	Exacerbation Rate over 24 Weeks <sup>2</sup> (95% CI)
<i>Study 361</i>					
UMEC 125	-0.08 (-0.17, 0.02)	-0.31 (-2.47, 1.85)	-0.84 (-1.29, -0.39)	9.17 (3.51, 14.84)	0.50 (0.35, 0.93)
VI 25	-0.03 (-0.13, 0.06)	-0.87 (-3.05, 1.30)	-0.79 (-1.24, -0.34)	9.56 (3.85, 15.26)	0.52 (0.32, 0.83)
UMEC/VI 125/25	-0.15 (-0.24, -0.06)	-3.60 (-5.76, 1.44)	-1.49 (-1.94, -1.04)	16.74 (11.08, 22.41)	0.36 (0.21, 0.60)
<i>Study 373</i>					
UMEC 62.5	-0.10 (-0.19, -0.00)	-4.69 (-7.07, 2.31)	-0.27 (-0.77, 0.22)	8.40 (3.17, 13.62)	0.60 (0.37, 0.96)
VI 25	-0.14 (-0.24, -0.05)	-5.19 (-7.58, 2.80)	-0.92 (-1.41, -0.43)	13.55 (8.32, 18.78)	0.71 (0.45, 1.12)
UMEC/VI 62.5/25	-0.17 (-0.26, -0.08)	-5.51 (-7.89, -3.13)	-0.83 (-1.32, -0.34)	12.51 (7.31, 17.71)	0.48 (0.29, 0.79)

<sup>1</sup> Estimated differences, as compared to placebo, from linear mixed effects models with the following covariates: treatment group, baseline value, center group, smoking status, visit, visit-by-baseline interaction, and visit-by-treatment group interaction

<sup>2</sup> Estimated hazard ratios, as compared to placebo, from Cox proportional hazards models with the following covariates: treatment group, center group, smoking status

### 3.2.5.2 Primary Efficacy Studies: Contributions of Components to Efficacy of Combination

The contributions of the components to the efficacy of a combination product can be established by comparisons of the combination to each monotherapy. The contribution of UMEC was evaluated through a comparison between UMEC/VI and VI, and the contribution of VI was evaluated through a comparison between UMEC/VI and UMEC (Table 10). There was consistent evidence of the contribution of UMEC to the efficacy of the combination product with respect to the primary endpoint of mean change from baseline in trough FEV<sub>1</sub> at 24 weeks (evidence for the contribution of UMEC 62.5 to UMEC/VI 62.5/25 from Studies 373 and 360, evidence for the contribution of UMEC 125 to UMEC/VI 125/25 from Studies 361 and 374). However, there was statistical evidence of the contribution of VI to the efficacy of the combination product at the proposed 62.5/25 dose from only one of the primary efficacy studies (Study 373; difference in mean trough FEV<sub>1</sub>: 0.052 L, p=0.0021). There was evidence of the contribution of VI to the efficacy of the higher dose combination product in Study 361, but not in Study 374.

The contributions of both the UMEC and VI components were supported by analyses of the secondary endpoint 0–6 hour weighted mean FEV<sub>1</sub> (Table 10). There was also some separation between the monotherapy and combination treatment arms in descriptive figures presenting trough FEV<sub>1</sub> at earlier

time points (Figure 17 and Figure 18), in addition to empirical FEV<sub>1</sub> distribution functions (Figure 19 and Figure 20). There was not consistent evidence, however, for the contribution of the components with respect to other non-spirometric endpoints of interest (Table 11), although estimates generally trended in the direction of benefit (for the efficacy of the combination relative to the monotherapies).

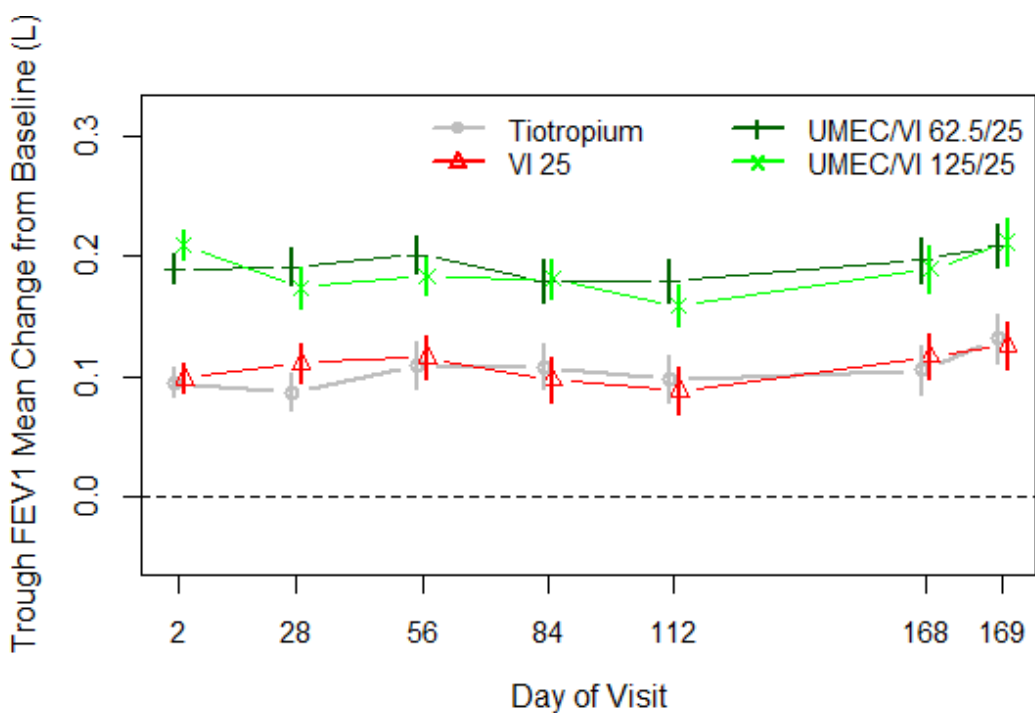
**Table 10. Contributions of the Umeclidinium and Vilanterol Components to the Efficacy of the Combination Product with Respect to the Mean Changes from Baseline in Trough FEV<sub>1</sub> (Primary Endpoint) and 0–6 Hour Weighted Mean FEV<sub>1</sub> (Secondary Endpoint) at 24 Weeks in Studies 361, 373, 360, and 374**

	Mean Difference <sup>1</sup> in Trough FEV <sub>1</sub> (L) (95% CI) p-value	Mean Difference <sup>1</sup> in 0–6 Hour Weighted Mean FEV <sub>1</sub> (L) (95% CI) p-value
<i>Study 361</i>		
UMEC 125 Contribution (UMEC/VI 125/25 vs. VI 25)	0.114 (0.081, 0.148) <0.0001	0.142 (0.109, 0.175) <0.0001
VI 25 Contribution (UMEC/VI 125/25 vs. UMEC 125)	0.079 (0.046, 0.112) <0.0001	0.109 (0.076, 0.141) <0.0001
<i>Study 373</i>		
UMEC 62.5 Contribution (UMEC/VI 62.5/25 vs. VI 25)	0.095 (0.060, 0.130) <0.0001	0.120 (0.084, 0.155) <0.0001
VI 25 Contribution (UMEC/VI 62.5/25 vs. UMEC 62.5)	0.052 (0.017, 0.087) 0.0021	0.092 (0.056, 0.127) <0.0001
<i>Study 360</i>		
UMEC 62.5 Contribution (UMEC/VI 62.5/25 vs. VI 25)	0.088 (0.037, 0.139) 0.0007	0.077 (0.025, 0.128) 0.0035
UMEC 125 Contribution (UMEC/VI 125/25 vs. VI 25)	0.093 (0.041, 0.144) 0.0004	0.088 (0.037, 0.140) 0.0008
<i>Study 374</i>		
VI 25 Contribution (UMEC/VI 125/25 vs. UMEC 125)	0.037 (-0.012, 0.087) 0.142	0.076 (0.029, 0.122) 0.0014

<sup>1</sup> Estimate differences, comparing the combination to monotherapy, from linear mixed effects models with the following covariates: treatment group, baseline FEV<sub>1</sub>, center group, smoking status, visit, visit-by-baseline FEV<sub>1</sub> interaction, and visit-by-treatment group interaction

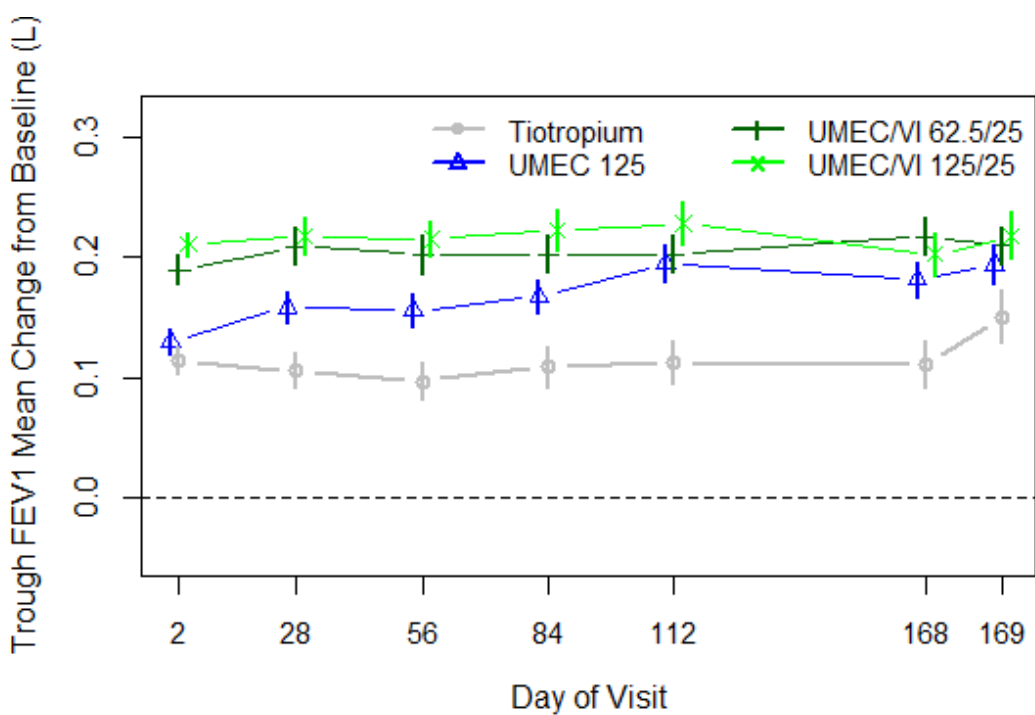


**Figure 17. Mean Change from Baseline in Trough FEV<sub>1</sub> over Time in Study 360**



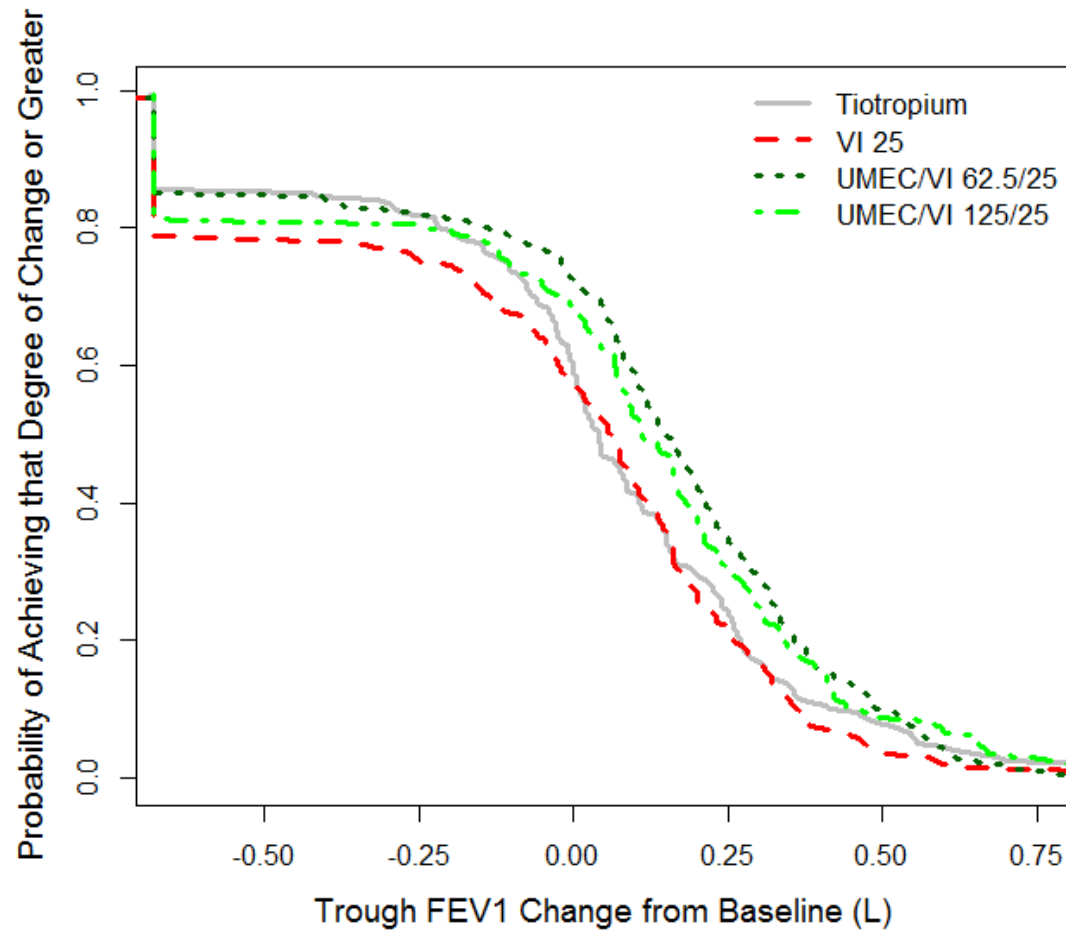
Unadjusted means based on observed data are displayed  
Error bars represent  $\pm 1$  standard error

**Figure 18. Mean Change from Baseline in Trough FEV<sub>1</sub> over Time in Study 374**



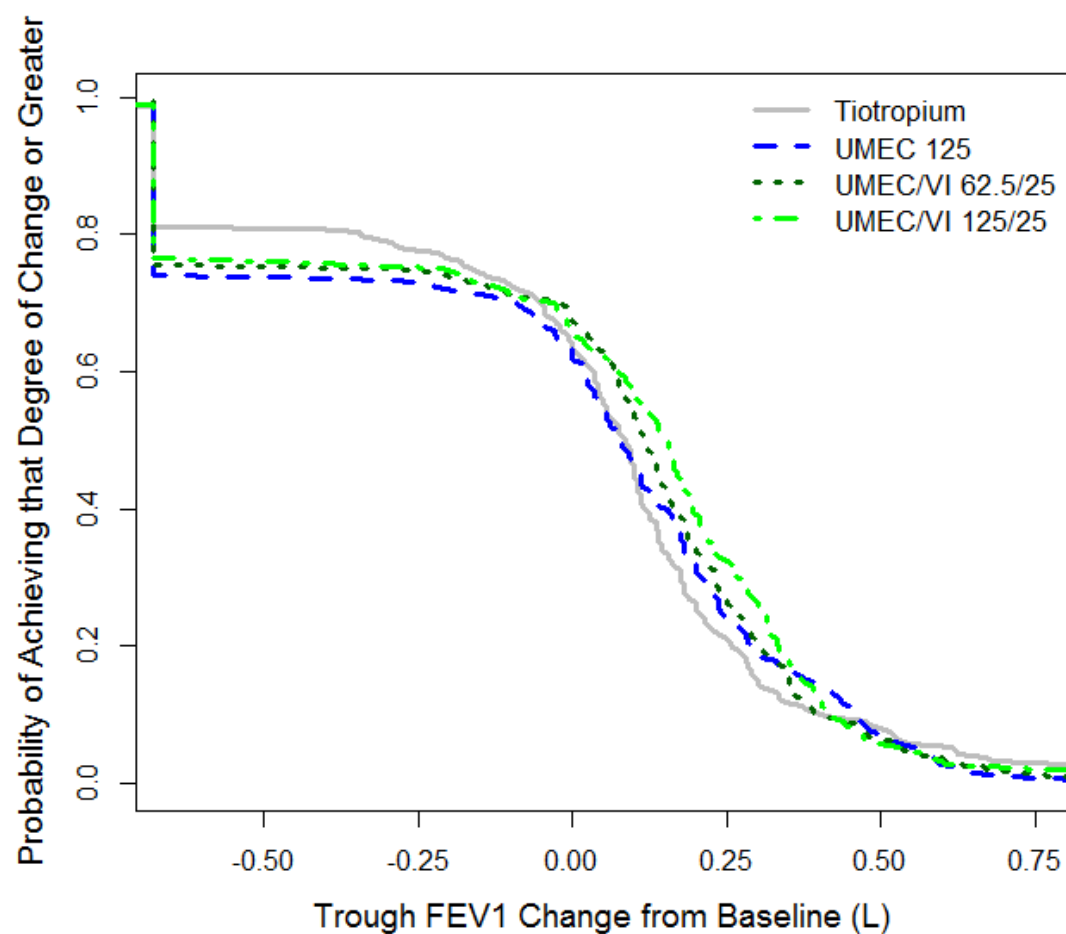
Unadjusted means based on observed data are displayed  
Error bars represent  $\pm 1$  standard error

**Figure 19. Empirical Distribution Function for Change from Baseline in Trough FEV<sub>1</sub> at 24 Weeks in Study 360**



Plot displays one minus the empirical distribution function  
Early study withdrawal was considered the worst possible outcome

**Figure 20. Empirical Distribution Function for Change from Baseline in Trough FEV<sub>1</sub> at 24 Weeks in Study 374**



Plot displays one minus the empirical distribution function  
Early study withdrawal was considered the worst possible outcome

**Table 11. Contributions of the Umeclidinium and Vilanterol Components to the Efficacy of the Combination Product with Respect to Additional Supportive Endpoints in Studies 361, 373, 360, and 374**

	Mean SOBDA Score at 24 Weeks <sup>1</sup> (95% CI)	Mean SGRQ Score at 24 Weeks <sup>1</sup> (95% CI)	Mean Rescue Puffs per Day over 24 Weeks <sup>1</sup> (95% CI)	Mean Percent Rescue-Free Days over 24 Weeks <sup>1</sup> (95% CI)	Exacerbation Rate over 24 Weeks <sup>2</sup> (95% CI)
<i>Study 361</i>					
UMEC 125 (UMEC/VI 125/25 vs. VI 25)	-0.12 (-0.20, -0.04)	-2.72 (-4.59, -0.86)	-0.70 (-1.09, -0.31)	7.19 (2.22, 12.16)	0.69 (0.40, 1.18)
VI 25 (UMEC/VI 125/25 vs. UMEC 125)	-0.07 (-0.15, 0.01)	-3.29 (-5.13, -1.44)	-0.65 (-1.04, -0.26)	7.57 (2.64, 12.50)	0.71 (0.42, 1.22)
<i>Study 373</i>					
UMEC 62.5 (UMEC/VI 62.5/25 vs. VI 25)	-0.03 (-0.11, 0.05)	-0.32 (-2.41, 1.78)	0.09 (-0.35, 0.53)	-1.04 (-5.68, 3.60)	0.68 (0.42, 1.11)
VI 25 (UMEC/VI 62.5/25 vs. UMEC 62.5)	-0.08 (-0.16, 0.01)	-0.82 (-2.90, 1.27)	-0.55 (-0.99, -0.11)	4.11 (-0.53, 8.76)	0.80 (0.48, 1.33)
<i>Study 360</i>					
UMEC 62.5 (UMEC/VI 62.5/25 vs. VI 25)	-0.02 (-0.14, 0.10)	1.42 (-1.46, 4.30)	-0.25 (-0.83, 0.32)	0.32 (-6.63, 7.28)	0.75 (0.37, 1.52)
UMEC 125 (UMEC/VI 125/25 vs. VI 25)	-0.02 (-0.14, 0.10)	-0.74 (-3.68, 2.20)	-0.22 (-0.80, 0.35)	1.61 (-5.41, 8.63)	0.64 (0.30, 1.37)
<i>Study 374</i>					
VI 25 (UMEC/VI 125/25 vs. UMEC 125)	-0.14 (-0.26, -0.03)	-2.12 (-4.80, 0.57)	-1.12 (-1.75, -0.49)	12.60 (5.35, 19.84)	0.63 (0.34, 1.17)

<sup>1</sup> Estimated differences, comparing the combination to monotherapy, from linear mixed effects models with the following covariates: treatment group, baseline value, center group, smoking status, visit, visit-by-baseline interaction, and visit-by-treatment group interaction

<sup>2</sup> Estimated hazard ratios, comparing the combination to monotherapy, from Cox proportional hazards models with the following covariates: treatment group, center group, smoking status

### 3.2.5.3 Primary Efficacy Studies: Tiotropium Comparisons

Data are available from Studies 360 and 374 for comparisons of UMEC/VI, and the monotherapy components, to the LAMA tiotropium. In Study 360, both UMEC/VI 62.5/25 and UMEC/VI 125/25 showed statistically significant, greater improvements in trough FEV<sub>1</sub> than tiotropium (Table 12; differences of 0.088 and 0.093 L, respectively). However, in Study 374, statistical significance was not achieved for the UMEC 125 versus UMEC/VI 125/25 comparison, which came before the tiotropium comparisons in the sequential multiple testing framework. Therefore, Study 374 *did not* provide replicate evidence for the superiority of the combination product (at either dose) to tiotropium. In addition, other non-spirometric endpoints of interest did not provide additional support for a benefit of the combination product relative to tiotropium (Table 13). In particular, the estimated effects of UMEC/VI 62.5/25 mcg on SOBDA and SGRQ in Studies 360 and 374 were similar to those of tiotropium, and actually trended in the wrong direction for exacerbation rate (hazard ratios in Studies 360 and 374 of 1.16 [95% CI: 0.53, 2.56] and 1.86 [95% CI: 0.97, 3.56], respectively).

**Table 12. Comparisons against Tiotropium with Respect to the Mean Changes from Baseline in Trough FEV<sub>1</sub> (Primary Endpoint) and 0–6 Hour Weighted Mean FEV<sub>1</sub> (Secondary Endpoint) at 24 Weeks in Studies 360 and 374**

	Mean Change from Baseline in Trough FEV <sub>1</sub> (L)	Mean Difference <sup>1</sup> in Trough FEV <sub>1</sub> (L) (95% CI) p-value	Mean Change from Baseline in 0–6 Hour Weighted Mean FEV <sub>1</sub> (L)	Mean Difference <sup>1</sup> in 0–6 Hour Weighted Mean FEV <sub>1</sub> (L) (95% CI) p-value
<i>Study 360</i>				
Tiotropium	0.121		0.181	
VI 25	0.121	0.000 (-0.052, 0.051) 0.995	0.178	-0.005 (-0.057, 0.047) 0.853
UMEC/VI 62.5/25	0.211	0.088 (0.038, 0.139) 0.0006	0.254	0.072 (0.021, 0.123) 0.0060
UMEC/VI 125/25	0.209	0.093 (0.042, 0.144) 0.0004	0.263	0.083 (0.032, 0.135) 0.0015
<i>Study 374</i>				
Tiotropium	0.149		0.18	
UMEC 125	0.186	0.037 (-0.012, 0.086) 0.138	0.206	0.026 (-0.020, 0.072) 0.273
UMEC/VI 62.5/25	0.208	0.060 (0.010, 0.109) 0.018	0.276	0.096 (0.050, 0.142) <0.0001
UMEC/VI 125/25	0.223	0.074 (0.025, 0.123) 0.0031	0.282	0.101 (0.055, 0.148) <0.0001

<sup>1</sup> Estimate differences, as compared to tiotropium, from linear mixed effects models with the following covariates: treatment group, baseline FEV<sub>1</sub>, center group, smoking status, visit, visit-by-baseline FEV<sub>1</sub> interaction, and visit-by-treatment group interaction

**Table 13. Comparisons against Tiotropium for Additional Supportive Endpoints in Studies 360 and 374**

	Mean SOBDA Score at 24 Weeks <sup>1</sup> (95% CI)	Mean SGRQ Score at 24 Weeks <sup>1</sup> (95% CI)	Mean Rescue Puffs per Day over 24 Weeks <sup>1</sup> (95% CI)	Mean Percent Rescue-Free Days over 24 Weeks <sup>1</sup> (95% CI)	Exacerbation Rate over 24 Weeks <sup>2</sup> (95% CI)
<i>Study 360</i>					
VI 25	0.02 (-0.10, 0.14)	-0.67 (-3.60, 2.26)	-0.41 (-0.99, 0.16)	5.66 (-1.36, 12.69)	1.56 (0.73, 3.33)
UMEC/VI 62.5/25	-0.00 (-0.12, 0.12)	0.75 (-2.12, 3.63)	-0.67 (-1.24, -0.10)	5.98 (-0.96, 12.93)	1.16 (0.53, 2.56)
UMEC/VI 125/25	-0.00 (-0.12, 0.12)	-1.41 (-4.34, 1.52)	-0.64 (-1.22, -0.06)	7.27 (0.23, 14.31)	1.00 (0.43, 2.31)
<i>Study 374</i>					
UMEC 125	0.02 (-0.10, 0.13)	1.38 (-1.28, 4.05)	0.03 (-0.59, 0.65)	2.49 (-4.63, 9.62)	1.82 (0.95, 3.49)
UMEC/VI 62.5/25	-0.08 (-0.20, 0.03)	-0.17 (-2.85, 2.52)	-0.58 (-1.21, 0.04)	5.16 (-2.08, 12.39)	1.86 (0.97, 3.56)
UMEC/VI 125/25	-0.13 (-0.24, -0.02)	-0.74 (-3.41, 1.93)	-1.09 (-1.72, -0.46)	15.09 (7.83, 22.35)	1.15 (0.56, 2.35)

<sup>1</sup> Estimated differences, as compared to tiotropium, from linear mixed effects models with the following covariates: treatment group, baseline value, center group, smoking status, visit, visit-by-baseline interaction, and visit-by-treatment group interaction

<sup>2</sup> Estimated hazard ratios, as compared to tiotropium, from Cox proportional hazards models with the following covariates: treatment group, center group, smoking status

### 3.2.5.4 Primary Efficacy Studies: Potential Effect of Missing Data

As described in detail in 3.2.3, there were substantial missing data in the primary efficacy studies. Dropout rates ranged from 15% to 33%, depending on the treatment arm and study. We used a number of approaches to investigate the potential effect of missing data on the reliability of efficacy results. First, we explored whether patients who dropped out were similar to patients who completed the 24-week studies. Patients who would go on to withdraw early tended to have slightly greater disease burden at baseline than patients who would go on to complete the 24-week primary efficacy studies (Table 14). For example, 14% of dropouts had GOLD Stage IV COPD at baseline, as compared to 9% of completers. Demographic characteristics were largely similar between dropouts and completers. These trends were also evident when each study was evaluated separately.

We also examined trends in trough FEV<sub>1</sub> before dropout within each treatment arm. Figure 21 displays average pulmonary function over time by dropout pattern, i.e., by the final visit at which FEV<sub>1</sub>

measurements were available, for the placebo-controlled Studies 361 and 373. Two general patterns were evident: (1) in all treatment arms, patients' pulmonary function tended to be relatively constant, or in slight decline, across the visits immediately preceding withdrawal; and (2) patients on the active treatment arms (both the monotherapies and combination product) tended to have better pulmonary function than placebo patients (both placebo completers and dropouts) before study withdrawal. These patterns were also generally observed within each study separately.

Based on these trends, it seems unlikely that patients treated with UMEC/VI who withdrew from the study early went on to have substantially worse lung function at 24 weeks than patients treated with placebo who dropped out. This is reassuring, especially in combination with the observation of greater dropout on placebo because of lack of efficacy (including COPD exacerbation) than on the experimental treatment arms. However, these patterns also highlight important deficiencies in the primary MMRM model, as well as the majority of the sensitivity analyses proposed by the applicant.

If the estimand of interest is the hypothetical effectiveness of the assigned treatment *if all patients could tolerate and adhere to the combination product*, then the estimated treatment effect from the MMRM model may provide a reliable estimate of the truth. However, if the estimand of interest is the effectiveness of the assigned treatment in all randomized participants, *at real world achievable adherence and tolerability*, the MMRM model likely does not produce a reliable estimate of the truth. The MMRM model, as well as the three missing data sensitivity analyses (MAR, CDC, LMCF) originally proposed by the applicant, essentially assumes that the observed treatment effect before dropout would have persisted in patients, even after they stopped taking the therapy. Because bronchodilators are generally considered symptomatic and not disease-modifying therapies, and their effects on FEV<sub>1</sub> likely do not persist more than a few days after patients stop using them, this assumption is not plausible scientifically.

Therefore, we focused on the Jump to Reference multiple imputation method, which essentially presumes that dropouts on all treatment arms would have had outcomes similar to those that were observed among completers (with similar baseline characteristics) *in the control group* (placebo in Studies 361 and 373, and tiotropium in Studies 360 and 374). Under the Jump to Reference approach, statistical significance was maintained for all relevant treatment comparisons (both against placebo and for contributions of the components). However, estimated magnitudes of treatment effect were approximately 20–30% smaller than those based on the primary MMRM model (Table 15 and Table 16). For example, in Study 373, the estimated mean improvement in FEV<sub>1</sub> on UMEC/VI 62.5/25, relative to placebo, was 0.132 L (95% CI: 0.092, 0.173), as compared to 0.167 L (95% CI: 0.128, 0.207) in the primary analysis. Estimated treatment effects of UMEC, VI, and UMEC/VI on weighted mean FEV<sub>1</sub> and SGRQ were also attenuated toward the null by about 20–30% in sensitivity analyses (Table 17). Estimated effects on SOBDA score were attenuated by up to 50% because of the substantial missing data (e.g., 50% missing Week 24 data in Study 373). The greater missing data on SOBDA as compared to other efficacy endpoints was in part attributable to the algorithm used by the applicant to compute the Week 24 mean score (see 3.2.2.1).

Although the scientific justification of the Jump to Reference assumptions seems reasonable, this and all other potential missing data sensitivity analyses rely on untestable assumptions about unobserved data. In addition, none of the sensitivity analyses conducted by the applicant allow for the possibility that dropouts on active treatment could have experienced *worse* outcomes after discontinuation than dropouts on control. That being said, the observed trend toward greater FEV<sub>1</sub> on active treatment than placebo before dropout (Figure 21) somewhat mitigates this concern, at least with respect to pulmonary function. There remains the possibility that dropouts from the active treatment arms could have gone on to experience *worse* outcomes with respect to important safety endpoints (see 3.3).

**Table 14. Baseline Characteristics, Stratified According to Whether Patients Completed the Study, Based on Integrated Data from Studies 361, 373, 360, and 374**

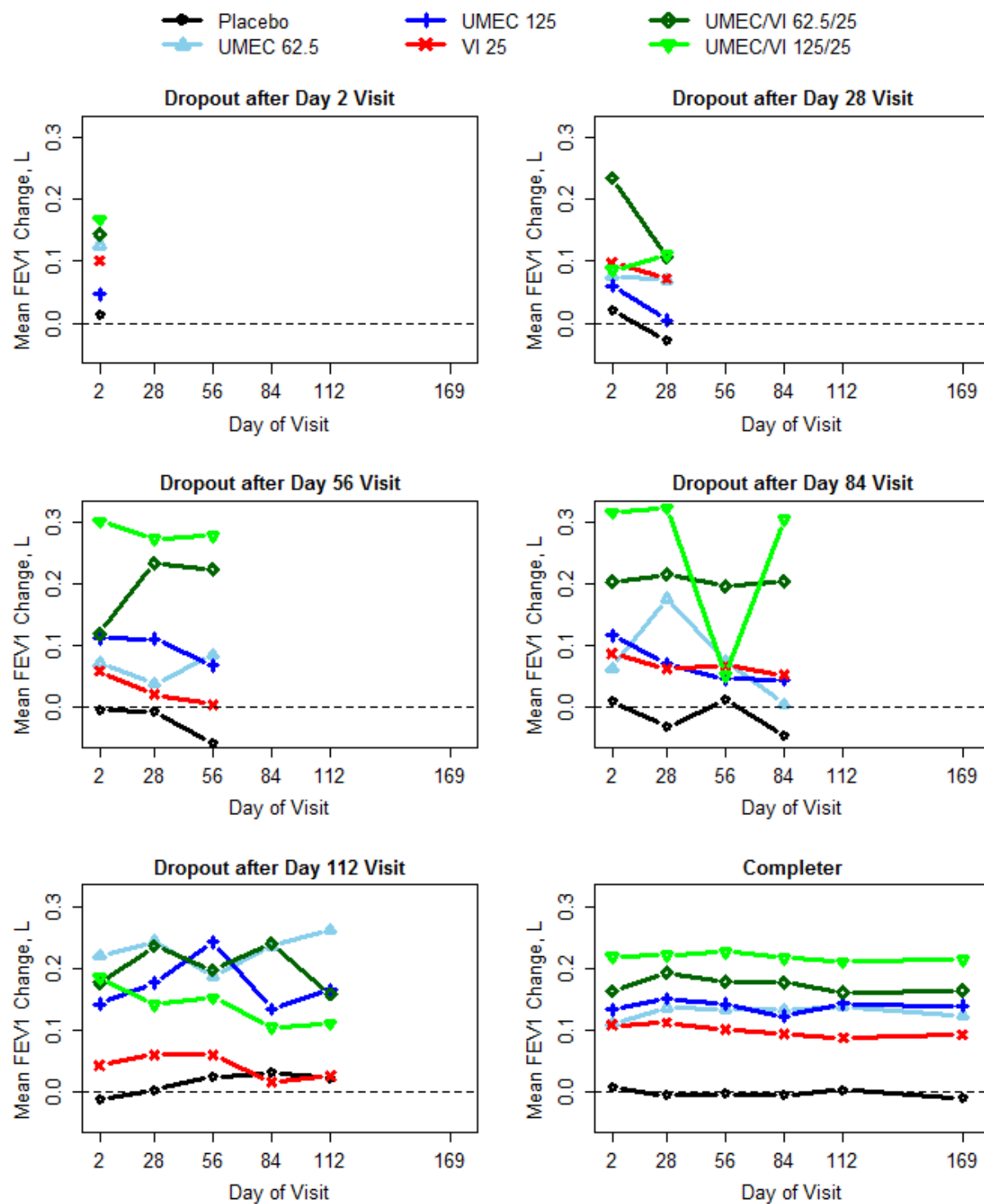
	<b>Completer<sup>1</sup> (N=3650)</b>	<b>Dropout (N=1051)</b>	<b>Overall (N=4733)</b>
Female	1141 (31%)	354 (34%)	1505 (32%)
Age	63.1 (8.6)	63.8 (9.1)	63.3 (8.7)
Race			
White	3076 (84%)	893 (85%)	3995 (84%)
Black	84 (2%)	40 (4%)	127 (3%)
Asian	328 (9%)	94 (9%)	425 (9%)
Other	162 (4%)	24 (2%)	186 (4%)
Hispanic/Latino	305 (8%)	75 (7%)	382 (8%)
BMI (kg/m <sup>2</sup> )	26.8 (5.8)	26.6 (5.8)	26.8 (5.8)
Current Smoker	1823 (50%)	497 (47%)	2343 (50%)
FEV <sub>1</sub>	1.3 (0.5)	1.1 (0.5)	1.2 (0.5)
GOLD Stage (ppFEV <sub>1</sub> )			
Stage II (50-80%)	1764 (49%)	408 (39%)	2184 (46%)
Stage III (30-50%)	1535 (42%)	486 (47%)	2034 (43%)
Stage IV (<30%)	338 (9%)	150 (14%)	494 (10%)
Chronic Bronchitis	2427 (66%)	658 (63%)	3106 (66%)
Emphysema	2153 (59%)	688 (65%)	2860 (60%)
Duration of COPD, years			
<1	331 (9%)	73 (7%)	407 (9%)
1,5	1379 (38%)	356 (34%)	1751 (37%)
5,10	1073 (29%)	322 (31%)	1405 (30%)
10,15	556 (15%)	204 (19%)	762 (16%)
15-20	163 (4%)	49 (5%)	212 (4%)
20-25	92 (3%)	25 (2%)	118 (2%)
>25	56 (2%)	22 (2%)	78 (2%)
Inhaled Corticosteroid Use	1775 (49%)	517 (49%)	2306 (49%)
Reversible to Salbutamol	1122 (31%)	317 (30%)	1452 (31%)
Reversible to Salbutamol and Ipratropium	1961 (54%)	534 (51%)	2512 (53%)
At United States site	854 (23%)	329 (31%)	1196 (25%)

Abbreviations: BMI = body mass index, ppFEV<sub>1</sub> = percent predicted forced expiratory volume in 1 second

<sup>1</sup> Patients classified as completers if they had a Day 169 visit. Numbers differ slightly from Tables 1-4, which were based on investigator reporting.



**Figure 21. Mean Change from Baseline in Trough FEV<sub>1</sub> by Treatment Group over Time, Stratified by Dropout Pattern, Based on Integrated Data from Studies 361 and 373**



**Table 15. Exploring the Potential Effect of Missing Data: Results for the Primary Endpoint Trough FEV<sub>1</sub> with the Primary Mixed Effects Analysis as Compared to a Multiple Imputation Sensitivity Analysis in Studies 361 and 373**

	Study 361: Mean Difference in Trough FEV <sub>1</sub> (L) (95% CI)		Study 373: Mean Difference in Trough FEV <sub>1</sub> (L) (95% CI)	
	Primary	Sensitivity <sup>1</sup>	Primary	Sensitivity <sup>1</sup>
<i>Placebo Comparisons</i>				
UMEC 62.5			0.115 (0.076, 0.155)	0.089 (0.049, 0.129)
UMEC 125	0.160 (0.122, 0.198)	0.122 (0.083, 0.162)		
VI 25	0.124 (0.086, 0.162)	0.092 (0.052, 0.131)	0.072 (0.032, 0.112)	0.053 (0.013, 0.093)
UMEC/VI 62.5/25			0.167 (0.128, 0.207)	0.132 (0.092, 0.173)
UMEC/VI 125/25	0.238 (0.200, 0.276)	0.192 (0.153, 0.231)		
<i>Contributions to Combination</i>				
UMEC 62.5 Contribution			0.095 (0.060, 0.130)	0.079 (0.044, 0.115)
UMEC 125 Contribution	0.114 (0.081, 0.148)	0.100 (0.065, 0.135)		
VI 25 Contribution	0.079 (0.046, 0.112)	0.070 (0.035, 0.104)	0.052 (0.017, 0.087)	0.043 (0.008, 0.079)

<sup>1</sup> Based on multiple imputation and the Jump to Reference model

**Table 16. Exploring the Potential Effect of Missing Data: Results for the Primary Endpoint Trough FEV<sub>1</sub> with the Primary Mixed Effects Analysis as Compared to a Multiple Imputation Sensitivity Analysis in Studies 360 and 374**

	Study 360: Mean Difference in Trough FEV <sub>1</sub> (L) (95% CI)		Study 374: Mean Difference in Trough FEV <sub>1</sub> (L) (95% CI)	
	Primary	Sensitivity <sup>1</sup>	Primary	Sensitivity <sup>1</sup>
<i>Tiotropium Comparisons</i>				
UMEC/VI 62.5/25	0.088 (0.038, 0.139)	0.078 (0.027, 0.129)	0.060 (0.010, 0.109)	0.046 (-0.003, 0.095)
UMEC/VI 125/25	0.093 (0.042, 0.144)	0.071 (0.019, 0.123)	0.074 (0.025, 0.123)	0.057 (0.008, 0.106)
<i>Contributions to Combination</i>				
UMEC 62.5 Contribution	0.088 (0.037, 0.139)	0.078 (0.026, 0.129)		
UMEC 125 Contribution	0.093 (0.041, 0.144)	0.071 (0.019, 0.123)		
VI 25 Contribution			0.037 (-0.012, 0.087)	0.029 (-0.020, 0.078)

<sup>1</sup> Based on multiple imputation and the Jump to Reference model

**Table 17. Exploring the Potential Effect of Missing Data: Results for Secondary and Additional Endpoints with the Primary Mixed Effects Analysis as Compared to a Multiple Imputation Sensitivity Analysis in Studies 361 and 373**

	Mean Difference in 0-6 Hour Weighted Mean FEV <sub>1</sub> at 24 Weeks (L) (95% CI)		Mean Difference in SOBDA Score at 24 Weeks (95% CI)		Mean Difference in SGRQ Score at 24 Weeks (95% CI)	
	Primary	Sensitivity <sup>1</sup>	Primary	Sensitivity <sup>1</sup>	Primary	Sensitivity <sup>1</sup>
<i>Study 361: Placebo Comparisons</i>						
UMEC 125	0.178 (0.141, 0.216)	0.136 (0.096, 0.175)	-0.08 (-0.17, 0.02)	-0.04 (-0.14, 0.05)	-0.31 (-2.47, 1.85)	-0.25 (-2.39, 1.88)
VI 25	0.145 (0.107, 0.182)	0.107 (0.067, 0.146)	-0.03 (-0.13, 0.06)	-0.02 (-0.12, 0.08)	-0.87 (-3.05, 1.30)	-0.68 (-2.83, 1.48)
UMEC/VI 125/25	0.287 (0.250, 0.324)	0.226 (0.186, 0.265)	-0.15 (-0.24, -0.06)	-0.10 (-0.19, 0.00)	-3.60 (-5.76, 1.44)	-3.02 (-5.17, -0.88)
<i>Study 373: Placebo Comparisons</i>						
UMEC 62.5	0.150 (0.110, 0.190)	0.116 (0.075, 0.157)	-0.10 (-0.19, -0.00)	-0.05 (-0.15, 0.04)	-4.69 (-7.07, 2.31)	-3.74 (-6.12, -1.37)
VI 25	0.122 (0.082, 0.162)	0.090 (0.049, 0.131)	-0.14 (-0.24, -0.05)	-0.07 (-0.17, 0.02)	-5.19 (-7.58, 2.80)	-4.12 (-6.50, -1.74)
UMEC/VI 62.5/25	0.242 (0.202, 0.282)	0.194 (0.153, 0.235)	-0.17 (-0.26, -0.08)	-0.10 (-0.19, -0.01)	-5.51 (-7.89, -3.13)	-4.54 (-6.92, -2.17)

<sup>1</sup> Based on multiple imputation and the Jump to Reference model

### 3.2.5.5 Additional Phase 3 Studies

In the cross-over Studies 417 and 418, treatment with UMEC/VI 125/25 resulted in statistically significantly greater mean changes from baseline in trough FEV<sub>1</sub> than placebo (Table 18 and Table 19). There was evidence of benefit on FEV<sub>1</sub> for UMEC/VI 62.5/25 over placebo in Study 418, but the same comparison failed to achieve statistical significance in Study 417 because it was positioned after the high dose EET comparison (which failed – see Table 19) in the statistical multiple testing hierarchy. Estimated magnitudes of treatment effect on FEV<sub>1</sub> for the two doses of combination product were similar to those observed in the primary efficacy studies.

There was not replicate evidence of benefit for UMEC/VI (at either dose) relative to placebo with respect to the co-primary endpoint exercise endurance time (Table 18 and Table 19). In Study 418, patients treated with UMEC/VI 62.5/25 had a statistically significant 69 second greater mean change from baseline in EET than those on placebo (95% CI: 25, 114 s; p=0.003). However, in Study 417, there was not statistical evidence of benefit for UMEC/VI 62.5/25 (estimated difference in means: 22 s; 95% CI: -14, 58 s; p=0.23). In addition, the estimated magnitude of benefit for the combination product (at either dose) in Studies 417 and 418 did not attain the minimal clinically important difference (70 seconds) approximated by the sponsor before the trial.

The interpretation of results from these cross-over studies is clouded by the substantial patient dropout, as the primary MMRM analyses rely on untestable and likely implausible assumptions about the missing data. Therefore, it is of interest to examine treatment effects within the stratum of patients who completed both 12-week treatment periods. Such an analysis maintains the integrity of randomization because of the cross-over design, although results should only be generalized to the subset of the patient population able to tolerate the therapy and willing to complete such a study. In Study 417, among patients who completed one period on UMEC/VI 62.5/25 and one period on placebo, there was a statistically significant treatment effect on FEV<sub>1</sub>, with an estimated difference in means at 12 weeks of 0.212 L (N=39; 95% CI: 0.131, 0.293; p<0.0001). There was no evidence of a treatment effect on EET for UMEC/VI 62.5/25, with an estimated difference in means at 12 weeks of 15.4 seconds (N=39; 95% CI: -41.4, 72.3 s; p=0.59). Similarly, in the subset of completers in Study 418, there was evidence of a treatment effect on FEV<sub>1</sub> (N=37; estimate=0.277; 95% CI: 0.208, 0.347; p<0.0001), but not EET (N=37; estimate=74.1 s; 95% CI: -3.3, 151.5 s; p=0.06). Similar results were observed for the higher dose of the combination product.

We also evaluated whether pulmonary function data from Studies 417 and 418 provided support for the contribution of vilanterol to the efficacy of the combination product. Such comparisons were not included in the multiple testing framework and therefore are only considered supportive. In both studies (Table 18 and Table 19), there was support for a contribution of VI 25 to the efficacy of the combination at the lower 62.5/25 dose (estimate mean differences in FEV<sub>1</sub> of 0.124 and 0.099 L), but not at the higher 125/25 dose.

The 52-week Study 359 only included placebo, UMEC 125, and UMEC/VI 125/25 treatment arms, and was designed to evaluate safety and tolerability, so no primary efficacy analyses were prespecified. Nevertheless, efficacy results were generally supportive of findings in the primary efficacy studies. Treatment with UMEC/VI 125/25 resulted in 0.197 L (95% CI: 0.121, 0.272) and 0.231 L (95% CI: 0.153, 0.310) greater mean trough FEV<sub>1</sub> changes at 6 and 12 months, respectively, as compared to placebo. In addition, there were lower rates of first COPD exacerbation (hazard ratio: 0.4; 95% CI: 0.3, 0.8) and daily rescue medication use (difference in mean puffs per day: -1.0; 95% CI: -1.4, -0.5) on the combination product, as compared to placebo.

**Table 18. Treatment Effects on the Co-Primary Endpoints Exercise Endurance Time (EET) and Trough FEV<sub>1</sub> at 12 Weeks in the Cross-Over Study 417**

	Mean Difference <sup>1</sup> in EET (s) (95% CI) p-value	Mean Difference <sup>1</sup> in Trough FEV <sub>1</sub> (L) (95% CI) p-value
<i>Placebo Comparisons</i>		
UMEC 62.5	26.5 (-25.9, 78.9) 0.32	0.087 (0.030, 0.143) 0.003
UMEC 125	13.1 (-38.9, 65.1) 0.62	0.140 (0.084, 0.197) 0.03
VI 25	-10.0 (-55.5, 35.4) 0.67	0.099 (0.050, 0.148) <0.0001
UMEC/VI 62.5/25	21.9 (-14.2, 58.0) 0.23	0.211 (0.172, 0.249) <0.0001
UMEC/VI 125/25	32.4 (-3.9, 68.8) 0.08	0.169 (0.129, 0.209) <0.0001
<i>Contributions to Combination</i>		
UMEC 62.5 Contribution (UMEC/VI 62.5/25 vs. VI 25)	31.9 (-14.1, 77.9) 0.17	0.111 (0.062, 0.161) <0.0001
UMEC 125 Contribution (UMEC/VI 125/25 vs. VI 25)	42.4 (-3.8, 88.7) 0.07	0.070 (0.019, 0.120) 0.007
VI 25 Contribution (UMEC/VI 62.5/25 vs. UMEC 62.5)	-4.6 (-57.6, 48.4) 0.86	0.124 (0.067, 0.182) <0.0001
VI 25 Contribution (UMEC/VI 125/25 vs. UMEC 125)	19.3 (-33.4, 71.9) 0.47	0.029 (-0.028, 0.086) 0.32

<sup>1</sup> Estimates from linear mixed effects models with the following covariates: treatment group, period baseline value, mean of period baseline values, period, center group, smoking status, visit, visit by mean baseline value interaction, and visit-by-treatment group interaction

**Table 19. Treatment Effects on the Co-Primary Endpoints Exercise Endurance Time (EET) and Trough FEV<sub>1</sub> at 12 Weeks in the Cross-Over Study 418**

	Mean Difference <sup>1</sup> in EET (s) (95% CI) p-value	Mean Difference <sup>1</sup> in Trough FEV <sub>1</sub> (L) (95% CI) p-value
<i>Placebo Comparisons</i>		
UMEC 62.5	25.0 (-41.0, 91.1) 0.46	0.144 (0.086, 0.203) <0.0001
UMEC 125	74.7 (6.0, 143.4) 0.03	0.256 (0.193, 0.318) <0.0001
VI 25	30.6 (26.8, 88.0) 0.30	0.112 (0.061, 0.163) <0.0001
UMEC/VI 62.5/25	69.4 (24.5, 114.4) 0.003	0.243 (0.202, 0.284) <0.0001
UMEC/VI 125/25	65.8 (20.3, 111.3) 0.005	0.261 (0.220, 0.303) <0.0001
<i>Contributions to Combination</i>		
UMEC 62.5 Contribution (VI 25 vs. UMEC/VI 62.5/25)	38.8 (-18.9, 96.5) 0.19	0.132 (0.081, 0.183) <0.0001
UMEC 125 Contribution (VI 25 vs. UMEC/VI 125/25)	35.2 (-22.7, 93.1) 0.23	0.150 (0.098, 0.201) <0.0001
VI 25 Contribution (UMEC 62.5 vs. UMEC/VI 62.5/25)	44.4 (-21.8, 110.6) 0.19	0.099 (0.041, 0.157) 0.0009
VI 25 Contribution (UMEC 125 vs. UMEC/VI 125/25)	-8.9 (-77.8, 60.1) 0.80	0.006 (-0.055, 0.067) 0.85

<sup>1</sup> Estimates from linear mixed effects models with the following covariates: treatment group, period baseline value, mean of period baseline values, period, center group, smoking status, visit, visit by mean baseline value interaction, and visit-by-treatment group interaction

### 3.3 Evaluation of Safety

Dr. Jennifer Pippins, the Medical Reviewer, conducted the safety evaluation, and the reader is referred to Dr. Pippins' review for detailed information on the safety profile of UMEC/VI. We also conducted some additional analyses to further explore a potential cardiovascular safety signal. The applicant prespecified a number of adverse events (AEs) of special interest based on potential pharmacologic class effects of LAMAs and LABAs. One group of special interest consisted of cardiovascular adverse events, including acquired long QT interval, cardiac arrhythmias, cardiac failure, cardiac ischemia, hypertension, sudden death, and stroke. All serious adverse event (SAE) narratives were adjudicated by an independent, blinded adjudication committee. The applicant also classified events according to the major adverse cardiac events (MACE) criteria. MACE included adjudicated cardiovascular death, non-fatal stroke AEs of special interest, and non-fatal cardiac ischaemia AEs of special interest.

We compared treatment groups with respect to adjudicated cardiovascular serious adverse events and MACE, using unadjusted incidence rates, Kaplan Meier plots, and Cox proportional hazards regression analyses. We combined UMEC 62.5 and 125 into one UMEC group, and combined UMEC/VI 62.5/25

and 125/25 into one UMEC/VI group, because of the small numbers of events within groups and the lack of a consistent dose-response. We report findings both for the primary efficacy studies, and based on data from all phase 3 studies (Studies 361, 373, 360, 374, 359, 408, and the first treatment periods of Studies 417 and 418). Analyses based on pooled data across randomized clinical trials can be influenced by confounding by study if randomization ratios and outcome risks differ across studies. This concern is somewhat mitigated here because the patient populations were very similar across the primary efficacy studies. In addition, we adjusted for study as a covariate in regression models.

Many of the numbers of events and event rates presented here differ from those in the Clinical Review and the applicant's summaries because: (1) our analyses include post-treatment events, which were generally captured if they occurred in the week ( $\pm 2$  days) following a patient's 24-week or early withdrawal visit, whereas the applicant's analyses only consider on-treatment data; (2) our analyses of CVD SAEs only include adjudicated events, whereas the applicant's analyses largely focus on event reports prior to adjudication; and (3) our analyses of data from "All phase 3 studies" include results from only the first treatment periods of cross-over Studies 417 and 418, whereas the applicant's results include data from the second treatment periods, as well.

Incidence rates of MACE were largely similar across the treatment arms (Table 20). There also was no evidence of a safety signal for MACE based on comparisons of the proportions of patients with events over time (Figure 22 and Figure 23), nor was there evidence in regression analyses (hazard ratio for UMEC/VI versus placebo in the primary efficacy studies: 1.2 [95% CI: 0.5, 2.7]; hazard ratio in all phase 3 studies: 0.8 [95% CI: 0.4, 1.5]). Similar results were observed when evaluating a narrow definition of MACE that only included cardiovascular death, non-fatal stroke AEs of special interest, and non-fatal myocardial infarction AEs of special interest (Table 20).

Despite the lack of evidence for MACE, there was the suggestion of a possible trend toward greater cardiovascular risk on UMEC, VI, and UMEC/VI, as compared to both placebo and tiotropium, when evaluating cardiovascular-related serious adverse events in the primary efficacy studies. This imbalance was evident when examining incidence rates (Table 20) and proportions with events over times (Figure 24), as well as in regression analyses (hazard ratio: 1.9; 95% CI: 0.5, 6.7). However, these trends largely went away when including data from all phase 3 studies (Table 20 and Figure 25; hazard ratio: 1.0 [95% CI: 0.4, 2.3]). Of note, the long-term, placebo-controlled safety Study 359 was the primary additional source of data in analyses that included all phase 3 studies. In Study 359, the only trial to include Holter monitoring in all randomized subjects, there was greater dropout on UMEC 125 (16%) and UMEC/VI 125/25 (16%) than placebo (7%) because of ECG and/or Holter abnormalities.

Finally, it is important to note that missing data clouds the interpretability of safety analyses. It is reassuring that dropout rates because of adverse events on the active arms (6-7% in the primary efficacy studies) were similar to the rate on placebo (5%). However, because patients were not followed up after treatment discontinuation for a complete 24-week safety evaluation, we cannot rule out the possibility that (1) differences in patient characteristics between dropouts on the placebo and active arms induce bias in safety comparisons, or (2) the active treatments have residual effects that increase risk of adverse events after patients stop taking them.

**Table 20. Numbers of MACE and Adjudicated Cardiovascular Serious Adverse Events, and Unadjusted Pooled Incidence Rates, by Treatment Group and Source of Data**

<i>All Phase 3 Studies<sup>1</sup> (N=6,156)</i>					
<b>Endpoint</b>	<b>Placebo (N=910)</b>	<b>UMEC<sup>2</sup> (N=1,512)</b>	<b>VI (N=1,111)</b>	<b>UMEC/VI<sup>3</sup> (N=2,200)</b>	<b>Tiotropium (N=423)</b>
MACE (broad) <sup>4</sup>	17 (51)	23 (37)	16 (38)	35 (38)	6 (35)
MACE (narrow) <sup>5</sup>	6 (18)	9 (14)	7 (16)	10 (11)	1 (6)
Adjudicated Cardiovascular SAE	8 (24)	20 (32)	15 (35)	22 (23)	2 (12)
<i>Primary Efficacy Studies (N=4,733)</i>					
<b>Endpoint</b>	<b>Placebo (N=555)</b>	<b>UMEC<sup>2</sup> (N=1,047)</b>	<b>VI (N=1,034)</b>	<b>UMEC/VI<sup>3</sup> (N=1,674)</b>	<b>Tiotropium (N=423)</b>
MACE (broad) <sup>4</sup>	8 (39)	16 (39)	16 (39)	27 (40)	6 (35)
MACE (narrow) <sup>5</sup>	3 (14)	5 (12)	7 (17)	8 (12)	1 (6)
Adjudicated Cardiovascular SAE	3 (14)	15 (36)	15 (37)	17 (25)	2 (12)

Cell contents are number of events (incidence rate, per 1,000 person-years)

Abbreviations: MACE = major adverse cardiac events; SAE = serious adverse event

<sup>1</sup> All Phase 3 Studies = Studies 361, 373, 360, 374, 359, 408, and the first treatment periods of Studies 417 and 418

<sup>2</sup> Combines the UMEC 62.5 and 125 mcg treatment groups

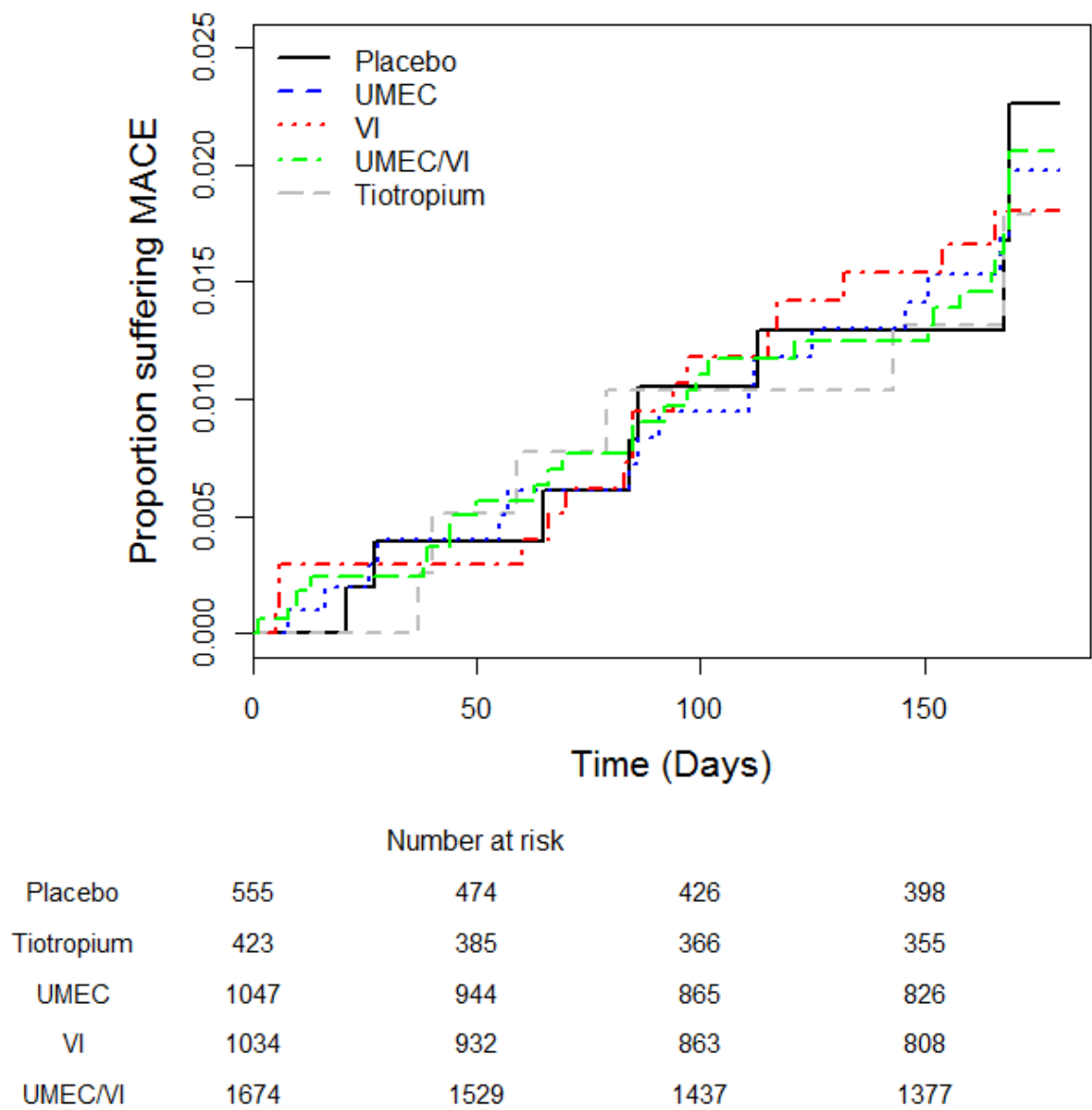
<sup>3</sup> Combines the UMEC/VI 62.5/25 and 125/25 mcg treatment groups

<sup>4</sup> MACE (broad) includes adjudicated cardiovascular death, non-fatal stroke adverse events of special interest, and non-fatal cardiac ischaemia adverse events of special interest

<sup>5</sup> MACE (narrow) includes adjudicated cardiovascular death, non-fatal stroke adverse events of special interest, and non-fatal myocardial infarction adverse events of special interest

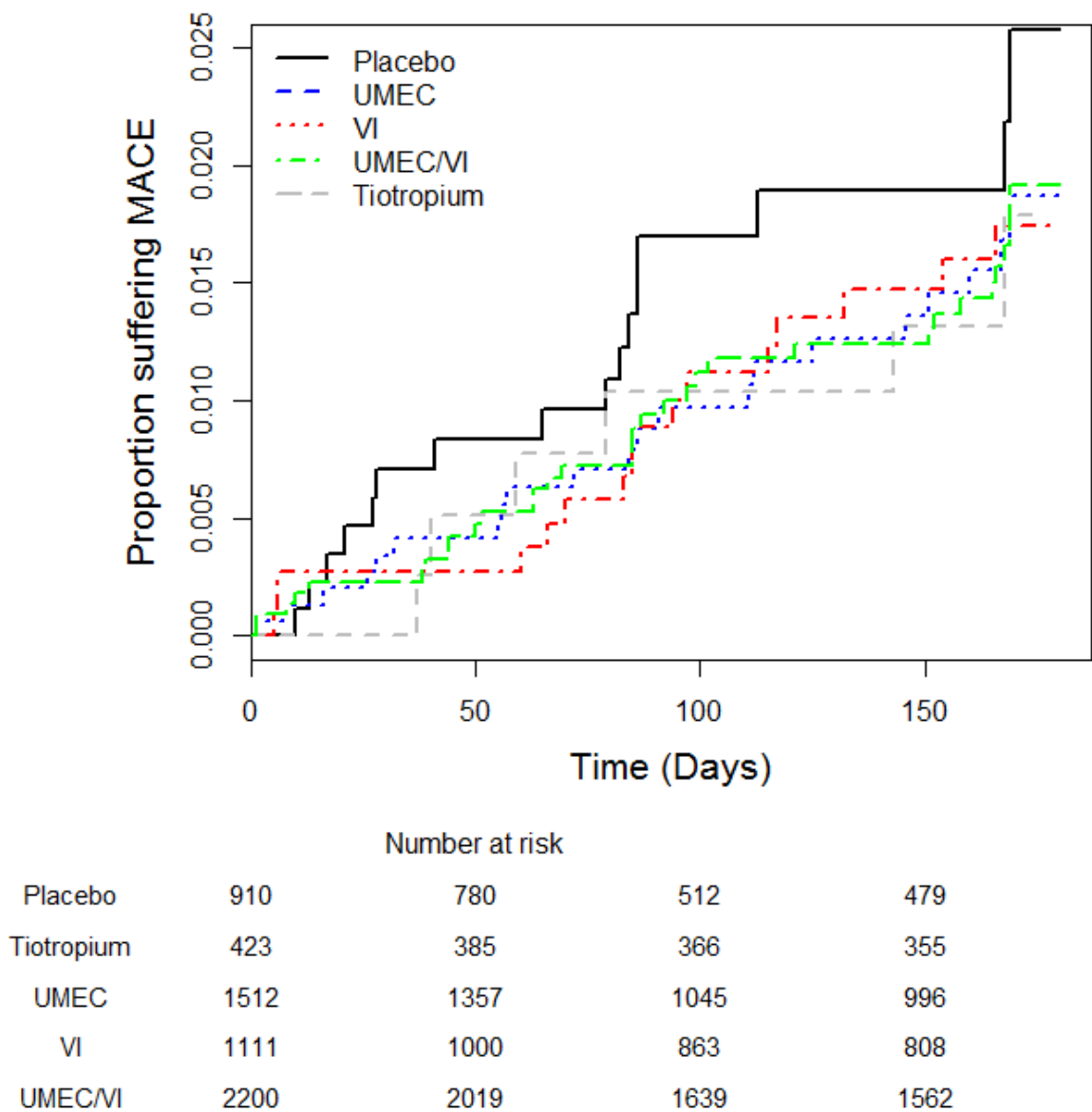


**Figure 22. Proportion Suffering MACE over Time by Treatment Group Based on Data from the Primary Efficacy Studies**



Abbreviations: MACE = major adverse cardiac events

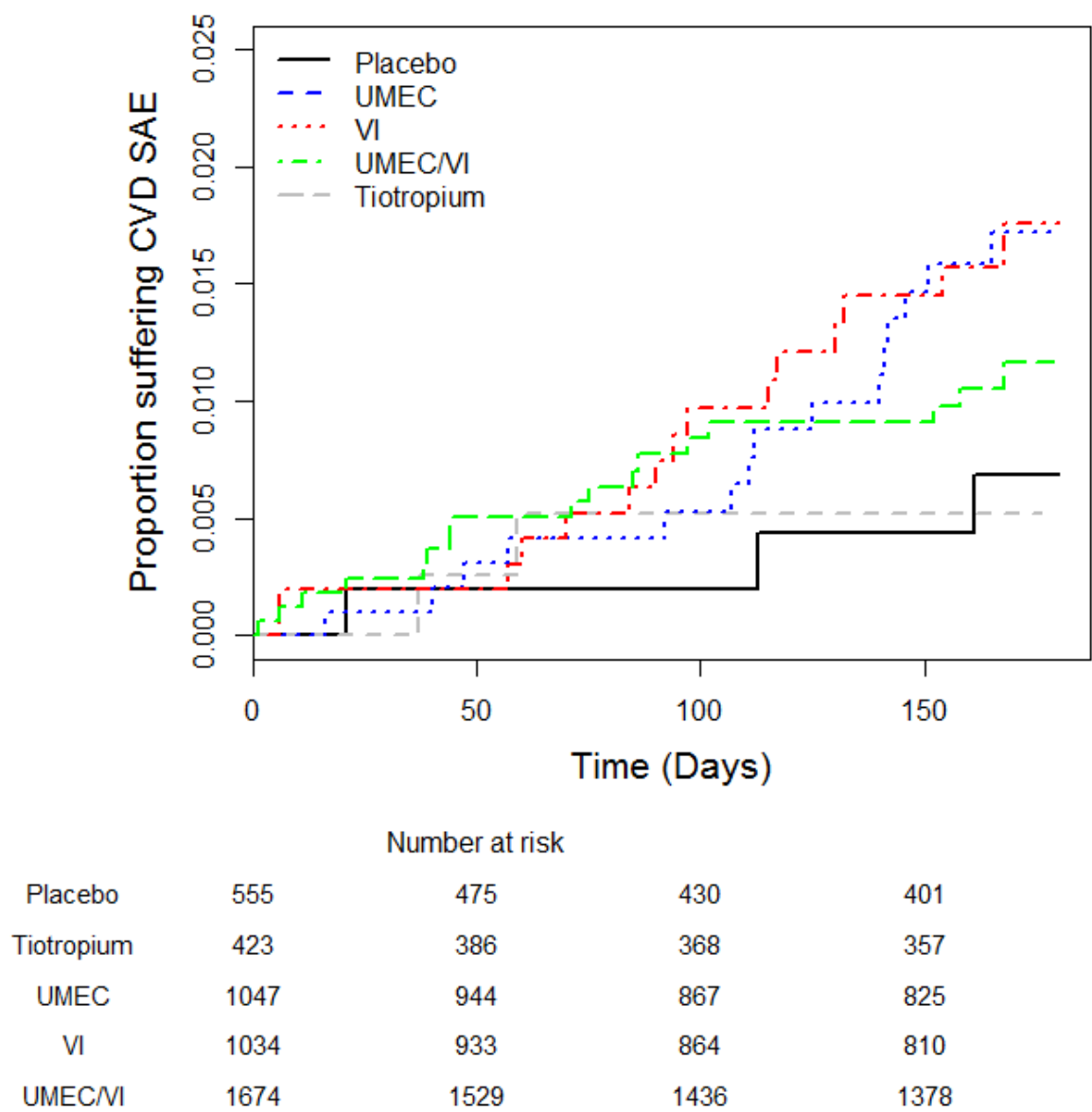
**Figure 23. Proportion Suffering MACE over Time by Treatment Group Based on Data from all Phase 3 Studies**



Abbreviations: MACE = major adverse cardiac events

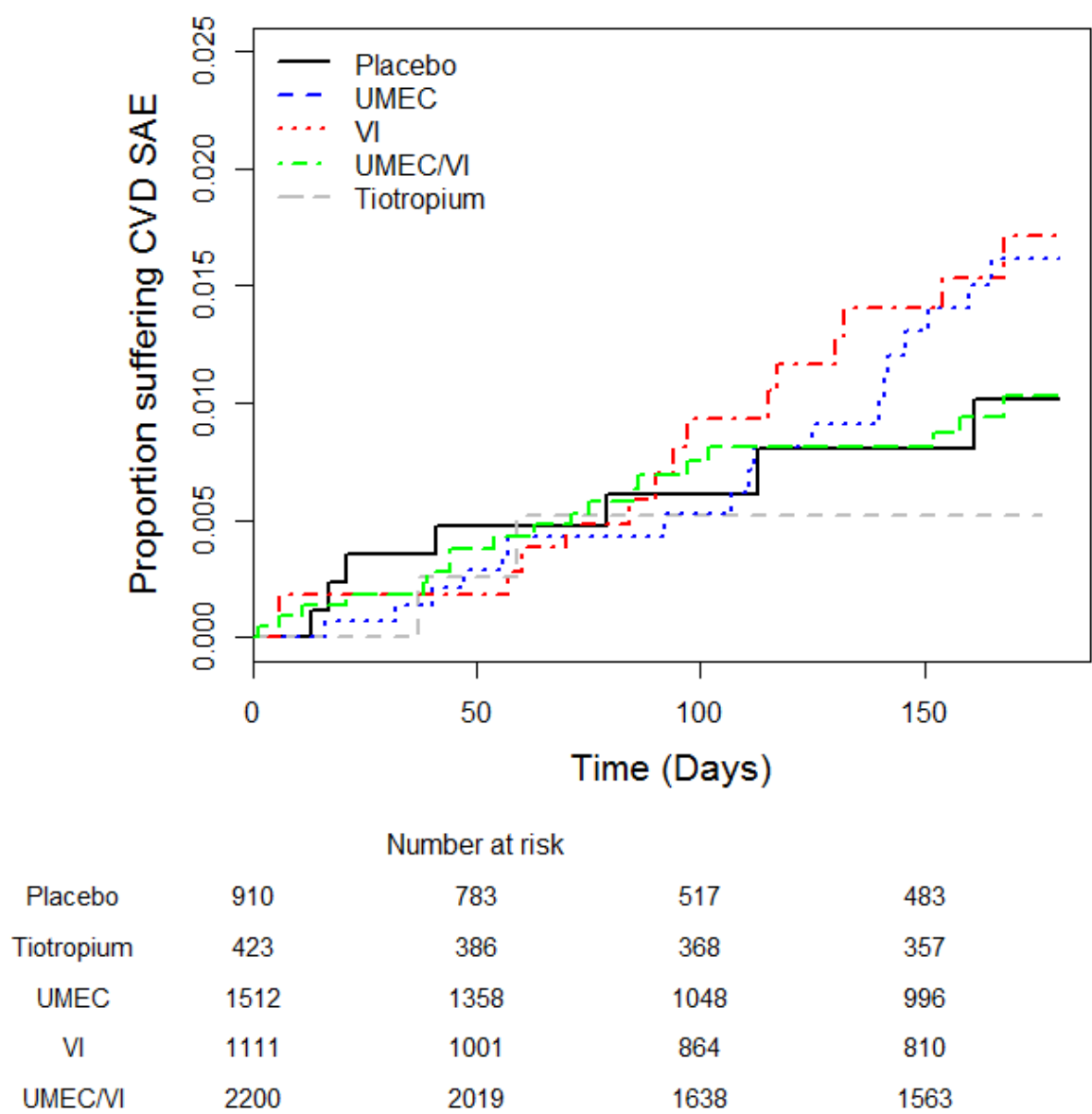
All Phase 3 Studies = Studies 361, 373, 360, 374, 359, 408, and the first treatment periods of Studies 417 and 418

**Figure 24. Proportion Suffering Adjudicated Cardiovascular Serious Adverse Events over Time by Treatment Group Based on Data from Primary Efficacy Studies**



Abbreviations: CVD SAE = cardiovascular serious adverse event

**Figure 25. Proportion Suffering Adjudicated Cardiovascular Serious Adverse Events over Time by Treatment Group Based on Data from all Phase 3 Studies**



Abbreviations: CVD SAE = cardiovascular serious adverse event

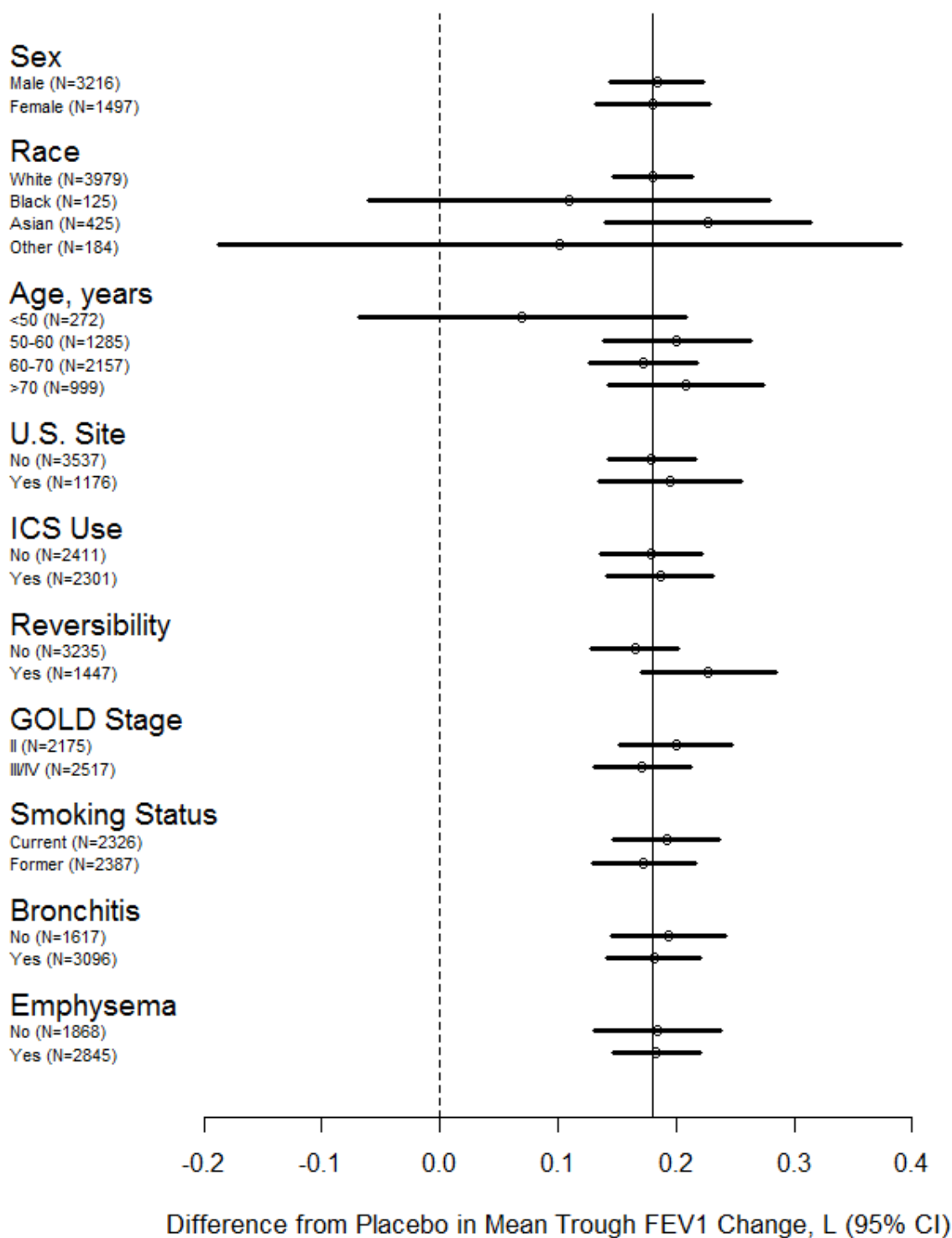
All Phase 3 Studies = Studies 361, 373, 360, 374, 359, 408, and the first treatment periods of Studies 417 and 418

## 4 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

Figure 26 presents the results of subgroup analyses by a number of demographic and baseline characteristics based on integrated data from the primary efficacy studies. We conducted subgroup analyses by sex, race (White, Black, Asian, or Other), age (<50, 50–60, 60–70, >70 years), geographic region (Non-U.S., U.S.), inhaled corticosteroid use (ICS) use, reversibility to salbutamol (defined by post-salbutamol FEV<sub>1</sub> at least 12% and 200 mL greater than pre-salbutamol FEV<sub>1</sub>), COPD GOLD stage, smoking status, chronic bronchitis, and emphysema. Estimated differences in mean trough FEV<sub>1</sub> comparing UMEC/VI 62.5/25 with placebo were largely consistent across these subgroups. Similar results were observed when restricting to Study 373, the only trial containing both placebo and UMEC 62.5/25 treatment arms. Of note, the limited number of Black subjects led to large variability in the estimated treatment effect in this subgroup (see wide confidence interval in Figure 26).

The suggestion in Figure 26 of possible differences in treatment effect by race and age was not replicated when we examined the higher 125/25 dose. The tendency for a larger observed treatment effect within the subset of patients demonstrating reversibility to salbutamol at baseline, however, was consistent across doses and studies. Importantly, the estimated treatment effect in patients who did not demonstrate reversibility, although smaller in magnitude, was still statistically significantly greater than zero.

**Figure 26. Estimated Treatment Effect of UMEC/VI 62.5/25 on Mean Trough FEV<sub>1</sub> at 24 Weeks, Stratified by Different Subgroups, Based on Integrated Data from Studies 360, 361, 373, and 374**



Estimates based on linear regression models adjusting for baseline FEV<sub>1</sub>, smoking status, center grouping, and study Subgroup sample sizes (N) based on complete integrated cohort, not subset receiving placebo or UMEC/VI 62.5/25

## 5 SUMMARY AND CONCLUSIONS

### 5.1 Statistical Issues

During this statistical review, we identified the following important issues:

- Potential effect of missing data on the reliability of efficacy results

This issue was discussed in detail in 3.2.4.4. There were substantial missing data in the primary efficacy studies, with dropout rates ranging from 15% to 33%, depending on the treatment arm and study. If the estimand of interest is the effectiveness of the assigned treatment in all randomized participants, at real world achievable adherence and tolerability, the MMRM model likely does not provide a reliable estimate of the truth. The MMRM model, as well as the three missing data sensitivity analyses (MAR, CDC, LMCF) originally proposed by the applicant, essentially assumes that the observed treatment effect before dropout would have persisted in patients, even after they stopped taking the therapy. Because bronchodilators are generally considered symptomatic and not disease-modifying therapies, and their effects on FEV<sub>1</sub> likely do not persist more than a few days after patients stop using them, this assumption is not plausible scientifically.

Therefore, we gave importance to a sensitivity analysis that multiply imputed missing data under the assumption that dropouts on all treatment arms would have had outcomes similar to those that were observed among completers (with similar baseline characteristics) *in the control group*. Statistical significance was maintained for all relevant treatment comparisons, but estimated magnitudes of treatment effect were approximately 20-30% smaller than those based on the primary MMRM model. None of the sensitivity analyses proposed by the applicant allow for the possibility that dropouts on active treatment could have experienced *worse* outcomes after discontinuation than dropouts on control. However, the observed trend toward greater FEV<sub>1</sub> on active treatment than placebo before dropout somewhat mitigates this concern, at least with respect to pulmonary function.

The presence of missing data also clouds the interpretation of safety comparisons. It is reassuring that dropout rates because of adverse events on the active arms (6-7% in the primary efficacy studies) were similar to the rate on placebo (5%). However, because patients were not followed up after treatment discontinuation for a complete 24-week safety evaluation, we cannot rule out the possibility that (1) differences in patient characteristics between dropouts on the placebo and active arms induce bias in safety comparisons, or (2) the active treatments have residual effects that increase risk of adverse events after patients stop taking them.

- Quantity of evidence of effectiveness for UMEC/VI at the proposed 62.5/25 mcg dose

From the four primary efficacy studies, a direct comparison between the proposed 62.5/25 mcg dose of UMEC/VI and placebo is possible only from Study 373. Therefore, we must evaluate whether results from this single study, combined with supportive data from additional studies, meet the standard for substantial evidence of effectiveness. The FDA guidance for industry *Providing Clinical Evidence of Effectiveness for Human Drug and Biological Products* indicates situations in which a single study of a new treatment may be combined with independent substantiation from related, supportive study data to provide evidence of effectiveness. In particular, the Guidance notes that supportive data may come from studies of a different dose, studies in a slightly different patient population, or from studies of the

monotherapies when evaluating a combination product, depending on the quality and outcomes of such related studies.

Therefore, we must evaluate the totality of the evidence in support of a UMEC/VI treatment effect on FEV<sub>1</sub>. First, we note that there was **strong** statistical evidence ( $p < 0.0001$ ) of a treatment effect in the single study evaluating UMEC/VI 62.5/25 versus placebo with respect to 24-week mean change in FEV<sub>1</sub>. Second, there was supportive evidence for FEV<sub>1</sub> improvement from a number of related studies. In the primary efficacy Study 360, there was evidence of superiority over the approved active comparator tiotropium (estimate=0.088 L;  $p = 0.0006$ ). In addition, in Study 418, which was a cross-over study designed to evaluate both FEV<sub>1</sub> and exercise endurance time (in a slightly different patient population than the primary efficacy studies), there was a statistically significant treatment effect on trough FEV<sub>1</sub> at 12 weeks (estimate=0.243 L;  $p < 0.0001$ ). Finally, there was strong statistical evidence ( $p < 0.0001$ ) of efficacy for the higher UMEC/VI 125/25 dose (which consistently showed similar benefit over placebo as the proposed 62.5/25 dose) in Study 361.

- Use of the surrogate marker FEV<sub>1</sub> as the primary efficacy endpoint

The primary endpoint in the primary efficacy studies was change from baseline in trough FEV<sub>1</sub> at 24 weeks. We consider FEV<sub>1</sub> to be a surrogate endpoint, because it does not directly measure how a patient functions or feels in daily life, or how long a patient survives (Fleming 2012). Spirometric assessments like FEV<sub>1</sub> provide standardized, easy to perform, and reproducible assessments of airflow obstruction and are commonly used and accepted by the Agency as primary efficacy endpoints in COPD clinical trials. However, because they do not directly measure the COPD symptoms (e.g., chronic cough, excess sputum production, dyspnea, exacerbation, and reduced exercise capacity) that are important to patients, the claim of effectiveness based on the primary analyses relies on the conclusion that the treatment effect on FEV<sub>1</sub> will reliably predict effects on a clinically meaningful endpoint. Therefore, we also considered the analyses of several secondary endpoints to be important in the overall evaluation of effectiveness. Such an approach is supported by the FDA draft guidance for industry *Chronic Obstructive Pulmonary Disease: Developing Drugs for Treatment*: “In studies where an objective measure is used as an endpoint, such as FEV<sub>1</sub>, use of subjective measures as important secondary assessments may be particularly useful in judging the value of mean changes in the primary endpoint.”

The following secondary endpoints ascertained in the UMEC/VI primary efficacy studies might be considered to provide some direct measure of how patients function or feel in daily life: COPD exacerbation, rescue medication use, SGRQ score, and SOBDA score. Although the level of evidence for these endpoints may not have been sufficient to support additional labeling claims, observed trends toward benefit can increase confidence that the treatment effect on FEV<sub>1</sub> will reliably predict clinical benefit. In Study 373, UMEC/VI 62.5/25 provided the following estimated benefits over placebo for these secondary endpoints: mean difference in SOBDA of -0.17 (95% CI: -0.26, -0.08); mean difference in SGRQ of -5.51 (95% CI: -7.89, -3.13), mean difference in mean daily rescue medication use of -0.83 (95% CI: -1.32, -0.34), and hazard ratio for incident COPD exacerbation of 0.48 (95% CI: 0.29, 0.79). Therefore, results for these secondary assessments provide additional support for the effectiveness of UMEC/VI in COPD, given the use of the surrogate marker FEV<sub>1</sub> as the primary endpoint.

- Quantity of evidence of the contribution of vilanterol to the effectiveness of UMEC/VI

From the four primary efficacy studies, a direct comparison between UMEC/VI 62.5/25 and UMEC 62.5 to evaluate the contribution of VI 25 at the proposed dose of the combination product is possible only from Study 373. Therefore, we must evaluate whether results from this single study, combined with supportive data from additional studies, meet the standard for substantial evidence of effectiveness.



First, it is important to note that there was **strong** statistical evidence in the single study evaluating the contribution of VI with respect to 24-week mean change in FEV<sub>1</sub> (p=0.0021). Second, there was some limited supportive evidence from related studies. There was statistical evidence of the contribution of VI to the higher 125/25 mcg dose of the UMEC/VI combination product in Study 361 (estimate=0.079 L; p<0.0001) but not Study 374 (estimate=0.037 L; p=0.142). There was also some support from comparisons of UMEC/VI 62.5/25 to UMEC 62.5 in the cross-over Studies 417 and 418 (estimated mean differences of 0.124 and 0.099 L, respectively), although these comparisons were not among those in the prespecified framework to account for multiple testing. Finally, there was evidence of a treatment effect for the VI 25 monotherapy, relative to placebo, in Studies 361 (estimate=0.124 L; p<0.0001) and 373 (estimate=0.072 L; p=0.0004).

- Dose selection and evidence of effectiveness for umeclidinium

Data to support the dose selection, safety, and effectiveness of vilanterol were reviewed as part of the Breo Ellipta program. However, umeclidinium is a new molecular entity, and therefore requires a more comprehensive evaluation. The findings of UMEC dose-ranging studies in COPD suggested that the 62.5 and 125 mcg doses selected for phase 3 study were reasonable, although there was little separation in efficacy between UMEC 31.25 and 62.5 mcg in Study 321. From the four primary efficacy studies, a direct comparison between UMEC 62.5 and placebo is possible only from Study 373. In Study 373, the estimated difference in mean trough FEV<sub>1</sub> change between UMEC 6.25 and placebo was 0.115 L (95% CI: 0.076, 0.155; p<0.0001). The efficacy of UMEC was also supported by trends toward benefit (relative to placebo) with respect to additional non-spirometric endpoints of interest, including mean changes from baseline in the SOBDA and SGRQ scores at 24 weeks, mean puffs of rescue medication per day, and exacerbation rate over 24 weeks. There was also evidence of efficacy in phase 3 Study 408, where the estimated treatment effect of UMEC 62.5 on trough FEV<sub>1</sub> at 12 weeks was 0.127 L (95% CI: 0.052, 0.202; p<0.001).

## 5.2 Collective Evidence

The collective evidence supports the effectiveness of UMEC/VI 62.5/25 mcg for once-daily maintenance bronchodilator treatment of airflow obstruction in patients with chronic obstructive pulmonary disease, including chronic bronchitis and/or emphysema. In the only one of the four primary efficacy studies (Study 373) that included both placebo and UMEC/VI 62.5/25 mcg treatment arms, the combination product provided a statistically significant 0.167 L (95% confidence interval: 0.128, 0.207) improvement over placebo in the primary endpoint, 24-week mean change from baseline in trough FEV<sub>1</sub>. There was also independent, supportive evidence of a treatment effect on FEV<sub>1</sub> from a 12-week phase 3 cross-over study, from comparisons against the active comparator tiotropium, and from results for the higher 125/25 mcg dose. Missing data sensitivity analyses demonstrated consistent evidence of superiority to placebo, but provided estimated treatment effect sizes of approximately 20–30% less than the primary analyses.

The effectiveness of UMEC/VI 62.5/25 was also supported by trends toward benefit with respect to several additional endpoints, including SGRQ score, SOBDA score, daily rescue medication use, and rate of COPD exacerbation. Although the level of evidence for these endpoints may not have been sufficient to support additional labeling claims (see 5.3), observed trends toward benefit increase confidence that the treatment effect on the surrogate endpoint FEV<sub>1</sub> is likely to predict clinical benefit, i.e., improvements in how COPD patients function, feel, or survive.

There was evidence of the contribution of UMEC 62.5 mcg to the efficacy of the UMEC/VI 62.5/25 mcg combination product in Studies 360 and 373. There was evidence of the contribution of VI 25 mcg to the efficacy of the combination in only a single study (Study 373), but there was supportive evidence from independent, related study data. Supportive data included results from Study 361 for the higher 125/25 mcg dose, exploratory analyses from two phase 3 cross-over studies, and findings of efficacy relative to placebo for the VI monotherapy. There was also replicate evidence of efficacy, relative to placebo, for the UMEC monotherapy, which is a new molecular entity.

The complete safety evaluation was conducted by Dr. Jennifer Pippins, the Medical Reviewer, but we performed additional analyses to explore potential cardiovascular safety signals. Rates of MACE were similar across the treatment arms, but an analysis of cardiovascular-related serious adverse events in the primary efficacy studies suggested a possible trend toward greater risk on the UMEC, VI, and UMEC/VI treatment arms, as compared to placebo and tiotropium. This imbalance in the rates of cardiovascular-related SAEs was not evident in analyses that included data from all of the phase 3 studies.

### 5.3 Labeling Recommendations



## 6 REFERENCES

FDA guidance for industry, 1998, Providing Clinical Evidence of Effectiveness for Human Drug and Biological Products.

Fleming, Thomas R and John H Powers, 2012, Biomarkers and Surrogate Endpoints in Clinical Trials, *Statistics in Medicine*, 31: 2973–2984.

FDA draft guidance for industry, 2007, Chronic Obstructive Pulmonary Disease: Developing Drugs for Treatment.

## 7 APPENDIX

**Table 21. Baseline Characteristics in Study 361**

	<b>Placebo</b> (N = 275)	<b>UMEC 125</b> (N = 407)	<b>VI 25</b> (N = 404)	<b>UMEC/VI 125/25</b> (N = 403)	<b>Overall</b> (N = 1489)
Female	100 (36%)	137 (34%)	139 (34%)	139 (34%)	515 (35%)
Age (years)	62.2 (8.5)	63.1 (8.5)	62.8 (8.8)	63.4 (8.1)	62.9 (8.5)
Race					
White	238 (87%)	363 (89%)	354 (88%)	359 (89%)	1314 (88%)
Black	9 (3%)	4 (1%)	7 (2%)	4 (1%)	24 (2%)
Asian	27 (10%)	40 (10%)	42 (10%)	39 (10%)	148 (10%)
Other	1 (0%)	0 (0%)	1 (0%)	1 (0%)	3 (0%)
Hispanic/Latino	1 (0%)	0 (0%)	0 (0%)	2 (0%)	3 (0%)
BMI (kg/m <sup>2</sup> )	26.5 (6.1)	26.4 (5.8)	27.2 (6.0)	26.5 (5.1)	26.6 (5.8)
Current Smoker	143 (52%)	216 (53%)	210 (52%)	200 (50%)	769 (52%)
FEV <sub>1</sub> (L)	1.3 (0.5)	1.3 (0.5)	1.3 (0.5)	1.3 (0.5)	1.3 (0.5)
GOLD Stage (ppFEV <sub>1</sub> )					
Stage II (50-80%)	121 (44%)	194 (48%)	207 (51%)	177 (44%)	699 (47%)
Stage III (30-50%)	132 (48%)	180 (44%)	159 (40%)	189 (47%)	660 (45%)
Stage IV (<30%)	21 (8%)	32 (8%)	36 (9%)	35 (9%)	124 (8%)
Chronic Bronchitis	199 (72%)	266 (65%)	273 (68%)	283 (70%)	1021 (69%)
Emphysema	160 (58%)	241 (59%)	230 (57%)	227 (56%)	858 (58%)
Duration of COPD, years					
<1	16 (6%)	39 (10%)	32 (8%)	29 (7%)	116 (8%)
1-5	94 (34%)	146 (36%)	167 (41%)	135 (33%)	542 (36%)
5-10	101 (37%)	117 (29%)	116 (29%)	131 (33%)	465 (31%)
10-15	48 (17%)	58 (14%)	70 (17%)	73 (18%)	249 (17%)
15-20	9 (3%)	25 (6%)	13 (3%)	11 (3%)	58 (4%)
20-25	3 (1%)	10 (2%)	3 (1%)	18 (4%)	34 (2%)
>25	4 (1%)	12 (3%)	3 (1%)	6 (1%)	25 (2%)
Inhaled Corticosteroid Use	138 (50%)	193 (47%)	191 (47%)	176 (44%)	698 (47%)
Reversible to Salbutamol	77 (28%)	132 (32%)	119 (29%)	133 (33%)	461 (31%)
Reversible to Salbutamol and Ipratropium	146 (53%)	228 (56%)	203 (50%)	213 (53%)	790 (53%)
At United States site	57 (21%)	87 (21%)	86 (21%)	86 (21%)	316 (21%)

Cell contents are mean (standard deviation) for continuous variables or frequency (percent) for categorical variables  
Abbreviations: BMI = body mass index, ppFEV<sub>1</sub> = percent predicted forced expiratory volume in 1 second

**Table 22. Baseline Characteristics in Study 373**

	<b>Placebo</b>	<b>UMEC 62.5</b>	<b>VI 25</b>	<b>UMEC/VI 62.5/25</b>	<b>Overall</b>
	<b>(N = 280)</b>	<b>(N = 418)</b>	<b>(N = 421)</b>	<b>(N = 413)</b>	<b>(N = 1532)</b>
Female	85 (30%)	120 (29%)	136 (32%)	108 (26%)	449 (29%)
Age (years)	62.2 (9.0)	64.0 (9.2)	62.7 (8.5)	63.1 (8.7)	63.1 (8.9)
Race					
White	237 (85%)	354 (85%)	363 (86%)	348 (84%)	1302 (85%)
Black	9 (3%)	14 (3%)	9 (2%)	15 (4%)	47 (3%)
Asian	22 (8%)	35 (8%)	34 (8%)	35 (8%)	126 (8%)
Other	12 (4%)	15 (4%)	15 (4%)	15 (4%)	57 (4%)
Hispanic/Latino	25 (9%)	37 (9%)	36 (9%)	35 (8%)	133 (9%)
BMI (kg/m <sup>2</sup> )	26.9 (5.9)	26.5 (5.6)	26.6 (5.9)	27.3 (6.0)	26.8 (5.9)
Current Smoker	150 (54%)	207 (50%)	199 (47%)	203 (49%)	759 (50%)
FEV <sub>1</sub> (L)	1.2 (0.5)	1.2 (0.5)	1.2 (0.5)	1.3 (0.6)	1.2 (0.5)
GOLD Stage (ppFEV <sub>1</sub> )					
Stage II (50-80%)	119 (42%)	191 (46%)	197 (47%)	201 (49%)	708 (46%)
Stage III (30-50%)	133 (48%)	172 (41%)	179 (43%)	166 (40%)	650 (43%)
Stage IV (<30%)	28 (10%)	54 (13%)	44 (10%)	45 (11%)	171 (11%)
Chronic Bronchitis	182 (65%)	274 (66%)	260 (62%)	283 (69%)	999 (65%)
Emphysema	173 (62%)	271 (65%)	273 (65%)	236 (57%)	953 (62%)
Duration of COPD, years					
<1	20 (7%)	36 (9%)	36 (9%)	36 (9%)	128 (8%)
1-5	107 (38%)	151 (36%)	157 (37%)	160 (39%)	575 (38%)
5-10	82 (29%)	127 (30%)	115 (27%)	123 (30%)	447 (29%)
10-15	51 (18%)	70 (17%)	73 (17%)	63 (15%)	257 (17%)
15-20	9 (3%)	15 (4%)	19 (5%)	16 (4%)	59 (4%)
20-25	6 (2%)	10 (2%)	12 (3%)	9 (2%)	37 (2%)
>25	5 (2%)	9 (2%)	9 (2%)	6 (1%)	29 (2%)
Inhaled Corticosteroid Use	137 (49%)	219 (52%)	212 (50%)	212 (51%)	780 (51%)
Reversible to Salbutamol	91 (32%)	121 (29%)	155 (37%)	129 (31%)	496 (32%)
Reversible to Salbutamol and Ipratropium	146 (52%)	223 (53%)	230 (55%)	227 (55%)	826 (54%)
At United States site	78 (28%)	118 (28%)	117 (28%)	115 (28%)	428 (28%)

Cell contents are mean (standard deviation) for continuous variables or frequency (percent) for categorical variables  
Abbreviations: BMI = body mass index, ppFEV<sub>1</sub> = percent predicted forced expiratory volume in 1 second

**Table 23. Baseline Characteristics in Study 360**

	<b>Tiotropium</b>	<b>VI 25</b>	<b>UMEC/VI 62.5/25</b>	<b>UMEC/VI 125/25</b>	<b>Overall</b>
	<b>(N = 208)</b>	<b>(N = 209)</b>	<b>(N = 212)</b>	<b>(N = 214)</b>	<b>(N = 843)</b>
Female	68 (33%)	66 (32%)	64 (30%)	63 (29%)	261 (31%)
Age (years)	62.6 (9.4)	63.2 (9.1)	63.0 (8.7)	62.9 (8.9)	62.9 (9.0)
Race					
White	177 (85%)	184 (88%)	182 (86%)	180 (84%)	723 (86%)
Black	6 (3%)	3 (1%)	7 (3%)	9 (4%)	25 (3%)
Asian	2 (1%)	0 (0%)	3 (1%)	1 (0%)	6 (1%)
Other	23 (11%)	22 (11%)	20 (9%)	24 (11%)	89 (11%)
Hispanic/Latino	23 (11%)	21 (10%)	24 (11%)	25 (12%)	93 (11%)
BMI (kg/m <sup>2</sup> )	27.6 (5.5)	27.3 (5.7)	27.4 (6.1)	26.5 (5.1)	27.2 (5.6)
Current Smoker	99 (48%)	106 (51%)	98 (46%)	124 (58%)	427 (51%)
FEV <sub>1</sub> (L)	1.3 (0.5)	1.4 (0.5)	1.3 (0.5)	1.3 (0.5)	1.3 (0.5)
GOLD Stage (ppFEV <sub>1</sub> )					
Stage II (50-80%)	96 (47%)	94 (46%)	104 (49%)	99 (47%)	393 (47%)
Stage III (30-50%)	87 (42%)	91 (44%)	85 (40%)	87 (41%)	350 (42%)
Stage IV (<30%)	23 (11%)	21 (10%)	22 (10%)	26 (12%)	92 (11%)
Chronic Bronchitis	149 (72%)	147 (70%)	147 (69%)	144 (67%)	587 (70%)
Emphysema	125 (60%)	116 (56%)	123 (58%)	129 (60%)	493 (58%)
Duration of COPD, years					
<1	20 (10%)	13 (6%)	20 (9%)	19 (9%)	72 (9%)
1,5	79 (38%)	73 (35%)	75 (35%)	74 (35%)	301 (36%)
5,10	54 (26%)	62 (30%)	63 (30%)	60 (28%)	239 (28%)
10,15	34 (16%)	40 (19%)	30 (14%)	38 (18%)	142 (17%)
15-20	14 (7%)	13 (6%)	11 (5%)	11 (5%)	49 (6%)
20-25	6 (3%)	7 (3%)	8 (4%)	10 (5%)	31 (4%)
>25	1 (0%)	1 (0%)	5 (2%)	2 (1%)	9 (1%)
Inhaled Corticosteroid Use	93 (45%)	84 (40%)	93 (44%)	103 (48%)	373 (44%)
Reversible to Salbutamol	47 (23%)	52 (25%)	57 (27%)	61 (29%)	217 (26%)
Reversible to Salbutamol and Ipratropium	99 (48%)	98 (47%)	113 (53%)	106 (50%)	416 (49%)
At United States site	53 (25%)	55 (26%)	59 (28%)	60 (28%)	227 (27%)

Cell contents are mean (standard deviation) for continuous variables or frequency (percent) for categorical variables  
Abbreviations: BMI = body mass index, ppFEV<sub>1</sub> = percent predicted forced expiratory volume in 1 second

**Table 24. Baseline Characteristics in Study 374**

	<b>Tiotropium</b>	<b>UMEC 125</b>	<b>UMEC/VI 62.5/25</b>	<b>UMEC/VI 125/25</b>	<b>Overall</b>
	<b>(N = 215)</b>	<b>(N = 222)</b>	<b>(N = 217)</b>	<b>(N = 215)</b>	<b>(N = 869)</b>
Female	62 (29%)	74 (33%)	77 (35%)	67 (31%)	280 (32%)
Age (years)	65.2 (8.3)	64.5 (8.3)	65.0 (8.6)	63.8 (8.5)	64.6 (8.4)
Race					
White	163 (76%)	169 (76%)	164 (76%)	160 (74%)	656 (75%)
Black	8 (4%)	6 (3%)	8 (4%)	9 (4%)	31 (4%)
Asian	36 (17%)	37 (17%)	35 (16%)	37 (17%)	145 (17%)
Other	8 (4%)	10 (5%)	10 (5%)	9 (4%)	37 (4%)
Hispanic/Latino	38 (18%)	42 (19%)	38 (18%)	35 (16%)	153 (18%)
BMI (kg/m <sup>2</sup> )	26.4 (6.1)	26.4 (5.7)	26.7 (6.1)	26.6 (5.8)	26.5 (5.9)
Current Smoker	102 (47%)	98 (44%)	92 (42%)	96 (45%)	388 (45%)
FEV <sub>1</sub> (L)	1.2 (0.4)	1.1 (0.4)	1.2 (0.5)	1.1 (0.5)	1.1 (0.5)
GOLD Stage (ppFEV <sub>1</sub> )					
Stage II (50-80%)	103 (48%)	86 (39%)	106 (49%)	89 (42%)	384 (44%)
Stage III (30-50%)	83 (39%)	106 (48%)	83 (38%)	102 (48%)	374 (43%)
Stage IV (<30%)	28 (13%)	29 (13%)	27 (12%)	23 (11%)	107 (12%)
Chronic Bronchitis	120 (56%)	120 (54%)	134 (62%)	125 (58%)	499 (57%)
Emphysema	136 (63%)	152 (68%)	132 (61%)	136 (63%)	556 (64%)
Duration of COPD, years					
<1	16 (7%)	16 (7%)	28 (13%)	31 (14%)	91 (10%)
1,5	83 (39%)	96 (43%)	80 (37%)	74 (34%)	333 (38%)
5,10	65 (30%)	65 (29%)	53 (24%)	71 (33%)	254 (29%)
10,15	34 (16%)	22 (10%)	37 (17%)	21 (10%)	114 (13%)
15-20	12 (6%)	12 (5%)	10 (5%)	12 (6%)	46 (5%)
20-25	3 (1%)	7 (3%)	3 (1%)	3 (1%)	16 (2%)
>25	2 (1%)	4 (2%)	6 (3%)	3 (1%)	15 (2%)
Inhaled Corticosteroid Use	115 (53%)	124 (56%)	103 (47%)	113 (53%)	455 (52%)
Reversible to Salbutamol	60 (28%)	75 (34%)	64 (29%)	79 (37%)	278 (32%)
Reversible to Salbutamol and Ipratropium	110 (51%)	130 (59%)	115 (53%)	125 (58%)	480 (55%)
At United States site	55 (26%)	58 (26%)	59 (27%)	53 (25%)	225 (26%)

Cell contents are mean (standard deviation) for continuous variables or frequency (percent) for categorical variables  
Abbreviations: BMI = body mass index, ppFEV<sub>1</sub> = percent predicted forced expiratory volume in 1 second

Figure 27. Proportion Suffering a COPD Exacerbation over Time in Study 361

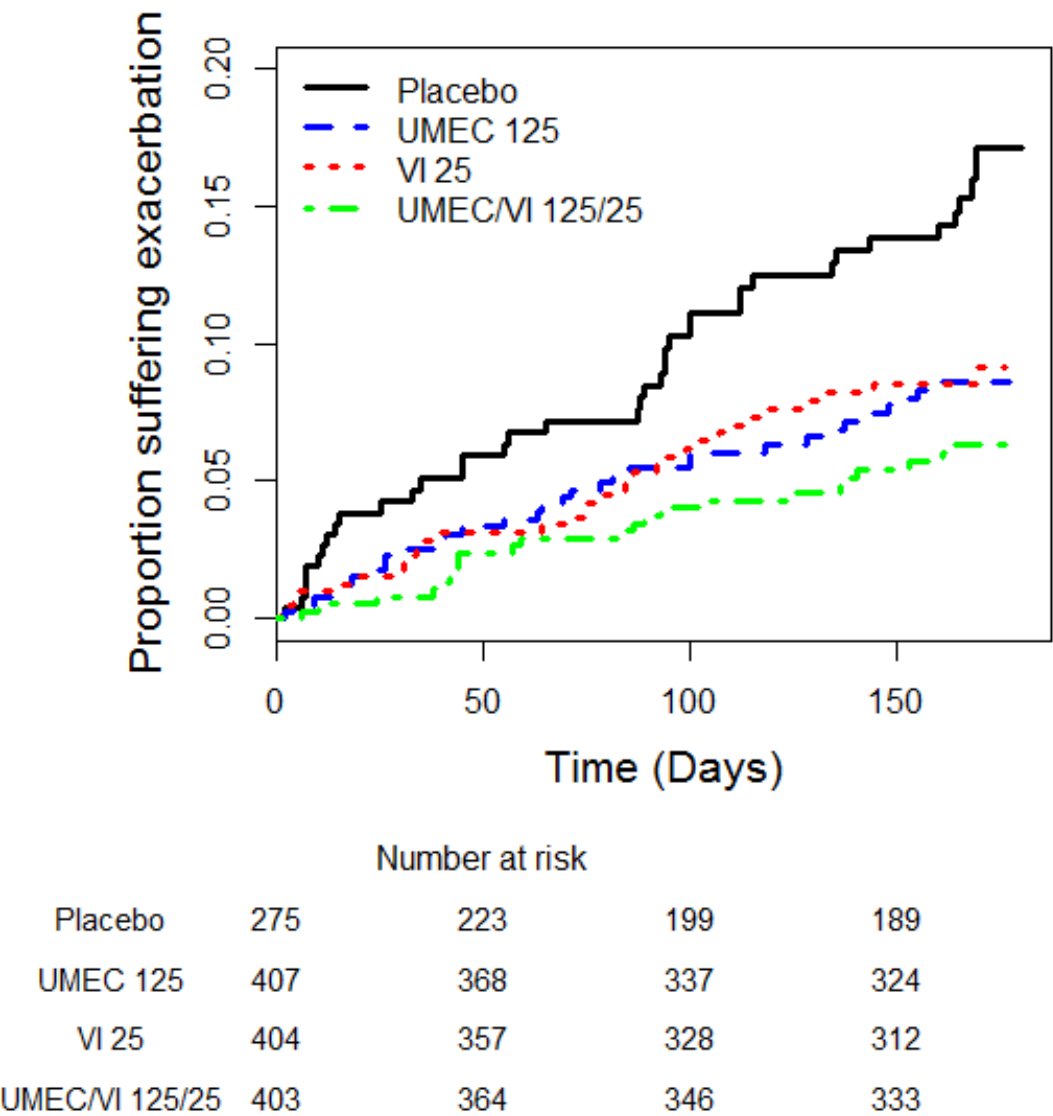


Figure 28. Proportion Suffering a COPD Exacerbation over Time in Study 373

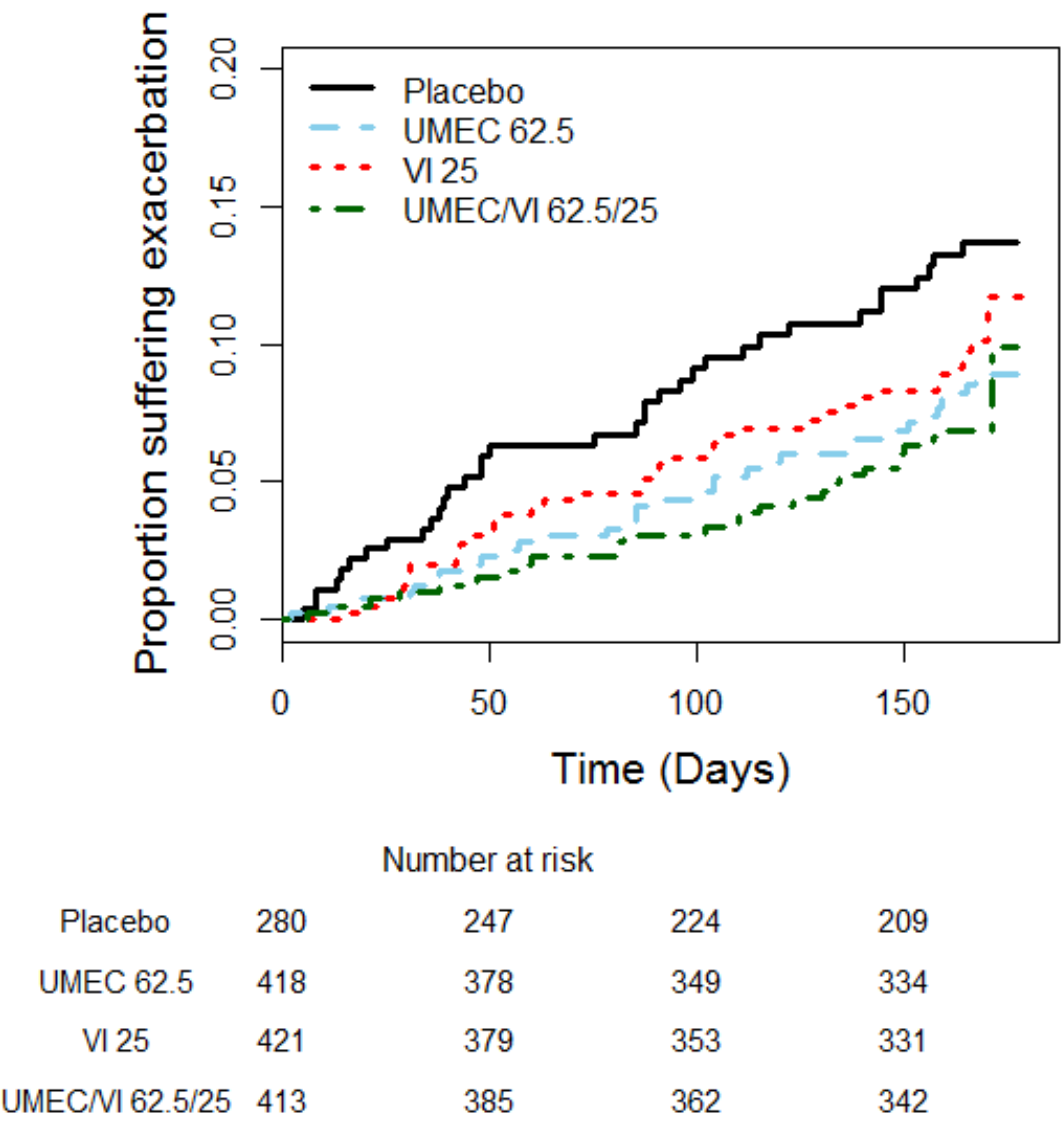
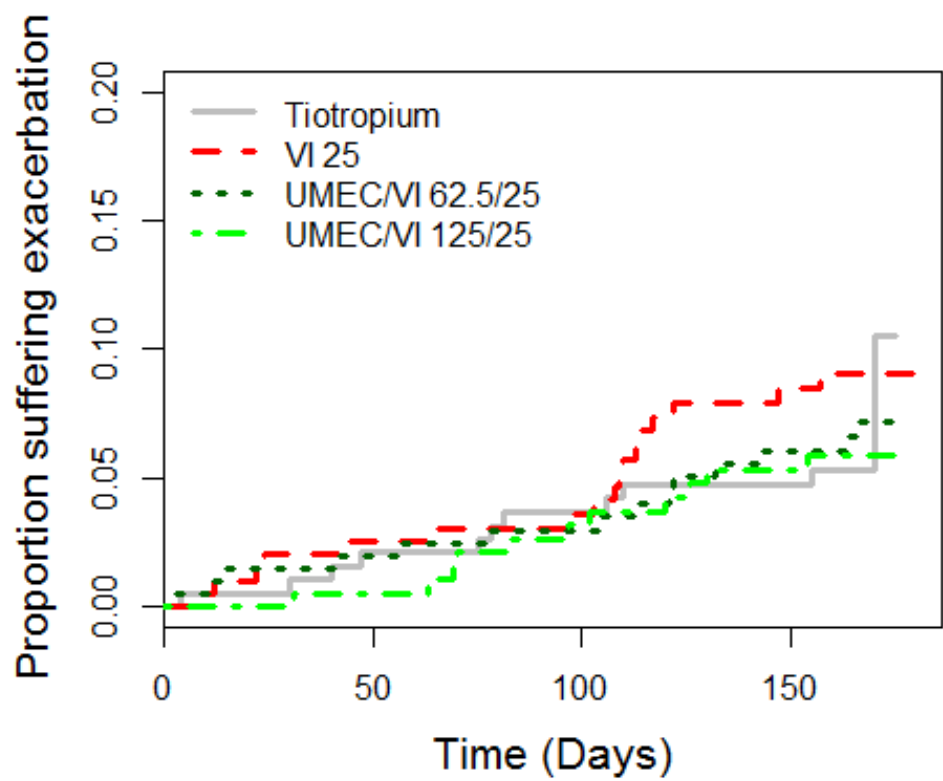


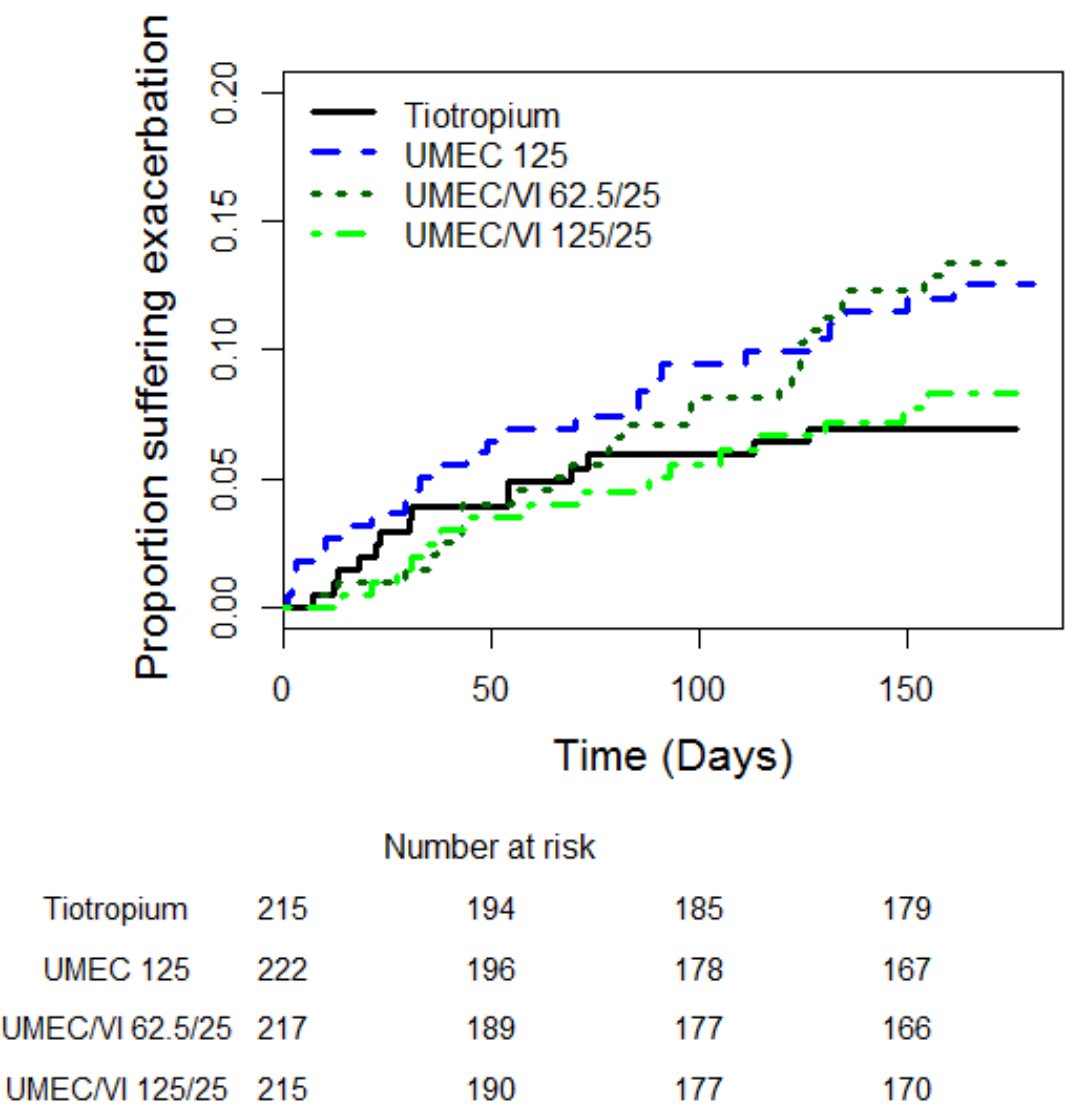


Figure 29. Proportion Suffering a COPD Exacerbation over Time in Study 360



Number at risk				
Tiotropium	203	187	179	175
VI 25	205	189	178	166
UMEC/VI 62.5/25	207	198	188	181
UMEC/VI 125/25	208	192	179	173

Figure 30. Proportion Suffering a COPD Exacerbation over Time in Study 374



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

/s/  
-----

GREGORY P LEVIN  
08/15/2013

JOAN K BUENCONSEJO  
08/15/2013  
I concur.

THOMAS J PERMUTT  
08/16/2013  
concur

## Addendum-1



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

### **Statistical Review and Evaluation**

#### **CARCINOGENICITY STUDIES**

**IND/NDA Number:** NDA 203-975  
**Drug Name:** GSK573719A  
**Indication(s):** 104 Week Carcinogenicity Studies in Rats and Mice  
**Applicant:** **Sponsor:** GlaxoSmithKline,  
Park Road, Ware, Hertfordshire, SG12 0DP, England

**Research Laboratory:** (b) (4)

**Documents Reviewed:** **Electronic submission:** Dated December 18, 2012  
**Electronic data:** Dated December 18, 2012

**Review Priority:** Standard

**Biometrics Division:** Division of Biometrics -6  
**Statistical Reviewer:** Mohammad Atiar Rahman, Ph.D.  
**Concurring Reviewer:** Karl Lin, Ph.D.

**Medical Division:** Division of Anesthesia Analgesia and Rheumatology Products  
**Reviewing Pharmacologist:** Jane Sohn, Ph.D.  
**Project Manager:** Leila Hann

**Keywords:** Carcinogenicity, Dose response

**Introduction:** A statistical review of this submission was issued on 4/2/2013. In that review there are two places in the interpretation of mortality data need to be clarified. This addendum contains the clarification of the interpretation.

**1. Rat study:**

**The original interpretation:** This reviewer's analysis showed statistically significant dose response relationship in mortality across treatment groups in female rats. The pairwise comparisons in female rats showed statistically significant increased mortality in high dose group compared to the control.

**The revised interpretation:** This reviewer's analysis showed statistically significant *negative* dose response relationship in mortality across treatment groups in female rats. The pairwise comparisons in female rats showed statistically significant *decreased* mortality in high dose group compared to the control.

**2. Mouse study:**

**The original interpretation:** This reviewer's analysis did not show statistically significant dose response relationship in the mortality across treatment groups in either sex. The pairwise comparisons showed statistically significant increased mortality in female mice low dose group compared to the control.

**The revised interpretation:** This reviewer's analysis did not show statistically significant dose response relationship in the mortality across treatment groups in either sex. The pairwise comparisons showed statistically significant *decreased* mortality in female mice low dose group compared to the control.

Mohammad Atiar Rahman, Ph.D.  
Mathematical Statistician

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

cc:  
Archival NDA 203-975  
Dr. Sohn  
Ms. Hann

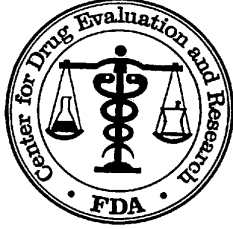
Dr. Machado  
Dr. Lin  
Dr. Rahman  
MS. Patrician

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

/s/  
-----

MOHAMMAD A RAHMAN  
04/09/2013

KARL K LIN  
04/09/2013  
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 and Evaluation**  
**CARCINOGENICITY STUDIES**

**IND/NDA Number:** NDA 203-975  
**Drug Name:** GSK573719A  
**Indication(s):** 104 Week Carcinogenicity Studies in Rats and Mice  
**Applicant:** **Sponsor:** GlaxoSmithKline,  
Park Road, Ware, Hertfordshire, SG12 0DP, England

**Research Laboratory:** (b) (4)

**Documents Reviewed:** **Electronic submission:** Dated December 18, 2012  
**Electronic data:** Dated December 18, 2012  
**Review Priority:** Standard

**Biometrics Division:** Division of Biometrics -6  
**Statistical Reviewer:** Mohammad Atiar Rahman, Ph.D.  
**Concurring Reviewer:** Karl Lin, Ph.D.

**Medical Division:** Division of Anesthesia Analgesia and Rheumatology Products  
**Reviewing Pharmacologist:** Jane Sohn, Ph.D.  
**Project Manager:** Leila Hann

**Keywords:** Carcinogenicity, Dose response

## Table of Contents

1.....	Background	3
2.....	Rat Study	3
2.1.	Sponsor's analyses.....	3
2.1.1.	Survival analysis.....	3
	<b>Sponsor's findings</b> .....	3
2.1.2.	Tumor data analysis.....	4
	<b>Sponsor's findings</b> .....	4
2.2.	Reviewer's analyses .....	4
2.2.1.	Survival analysis.....	4
	<b>Reviewer's findings</b> .....	4
2.2.2.	Tumor data analysis.....	5
	<b>Multiple testing adjustment</b> .....	5
	<b>Reviewer's findings</b> .....	6
3.....	Mouse Study	6
3.1.	Sponsor's analyses.....	7
3.1.1.	Survival analysis.....	7
	<b>Sponsor's findings</b> .....	7
3.1.2.	Tumor data analysis.....	7
	<b>Sponsor's findings</b> .....	7
3.2.	Reviewer's analyses .....	7
3.2.1.	Survival analysis.....	7
	<b>Reviewer's findings</b> .....	7
3.2.2.	Tumor data analysis.....	7
	<b>Reviewer's findings</b> .....	8
4.....	Evaluation of validity of the design	8
4.1.	Rat Study.....	9
4.2.	Mouse Study .....	10
5.....	Summary	11
6.....	Appendix	13
7.....	References	28



## 1. Background

In this submission the sponsor included reports of two animal carcinogenicity studies, one in rats and one in mice. These studies were intended to assess the carcinogenic potential of GSK573719A in rats and mice when administered by snout-only inhalation at appropriate drug levels for 104 weeks. Results of this review have been discussed with the reviewing pharmacologist Dr. Sohn.

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

## 2. Rat Study

Two separate experiments were conducted, one in males and one in females. In each of these two experiments there were three treated groups and one control group. Two hundred and sixty Crl:CD®(SD) rats of each sex were randomly assigned to the treated and control groups in equal size of 65 animals per group. The initially selected target doses were 0, 30, 100 and 300 µg/kg/day. These doses were given to males and females for Weeks 1-72. However, during Weeks 0-72, there was a statistically significant dose related decrement in the body weight gain in both sexes at all dose levels. This was accompanied by a minimal (statistically significant in females in all treated groups), reduction in food consumption. As a result of this decrement in body weight gain, the dose levels were reduced in all treated groups in both sexes from Week 73. From Week 73, the target doses were 0, 15, 50 and 150 µg/kg/day. In this review these dose groups will be referred to as the low, medium, and high dose group, respectively. All rats received daily snout-only inhaled doses of GSK573719 or vehicle for 60 minutes/day (Weeks 1 to 72) and for 30 minutes/day (Week 73 onwards). Control animals received lactose (vehicle) with 1% (w/w) magnesium stearate alone.

During the administration period all animals were checked daily for survival, general physical condition, and behavior. Detailed clinical examinations were performed prior to initiation of dosing, and weekly until necropsy. The animals were palpated weekly for the presence and growth of masses. New masses were recorded at detection in respect to mass number and location, date, size and description.

The animals were weighed twice during Week -1, on the day that treatment commenced, weekly for the first 16 weeks and once every 4 weeks thereafter throughout the remaining treatment period and on the day of necropsy. A complete histopathological examination was performed on all animals from all groups found dead, killed moribund, or sacrificed during or at the end of the experiment.

### 2.1. Sponsor's analyses

#### 2.1.1. Survival analysis

Survival function of each treatment group was estimated using the Kaplan-Meier product limit method and was presented graphically. Statistical analysis of the data was performed using the logrank tests for a dose response across the groups and pairwise comparisons of the treated groups with the control. Where the test for trend was statistically significant, the highest dose group was excluded and the dose response test repeated (using a one-tailed test), until the test was no longer statistically significant..

**Sponsor's findings:** Sponsor's analysis showed 57%, 65%, 63%, and 71% survival of male rats and 38%, 48%, 54%, and 62% survival of female rats in control, low, medium, and high dose groups, respectively. The sponsor analysis did not show statistically significant dose response relationship in mortality in male rats. For

female rats, the dose response relationship test, when all treated groups were included, was statistically significant ( $p = 0.037$ ). Upon exclusion of the high dose group, the dose response relationship test was no longer significant ( $p = 0.101$ ). The pairwise comparison of the high dose group with control was statistically significant ( $p = 0.035$ ).

### 2.1.2. Tumor data analysis

Tumor data were analyzed using the methods outlined in the paper of Peto et al. (1980) for dose response and pairwise comparisons of the treated group with the control. The time intervals for the sponsor's analysis were based on the suggestions made in the draft FDA guidance (2001) for carcinogenicity data analysis which are Weeks 1 - 52, 53 - 78, 79 - 92, 93 - 104 and terminal sacrifice. The score used were 0, 25, 85, and 254, which are the weighted average of the two sets of weights used before Week 73 and after Week 73. Tumors with a total tumor incidence of less than or equal to ten were analyzed using the exact permutation trend test.

Where the test for dose response was statistically significant, the highest dose group was excluded and the dose response test was repeated, using a one-tailed test until the test was no longer statistically significant. The significance levels were adjusted using a continuity correction where there was one degree of freedom.

**Adjustment for multiple testing:** For the adjustment for multiple testing for dose response relationship tests the sponsor used the method described in the book of Lin (2000), namely, the use of test levels of 0.005 for common tumors and rare tumors, respectively.

**Sponsor's findings:** Sponsor's analyses showed statistically significant dose response relationship among the treated groups in benign granular cell tumor in the brain in male rats and benign basal cell tumor in the skin in female rats. Although the dose responses in these tumors were statistically significant, the sponsor did not consider these incidences to be related to the treatment with GSK573719. The sponsor also mentioned that the ranges and distributions of the neoplastic lesions seen in this study were similar to those seen previously in this strain of rat in this laboratory. The sponsor concluded that there were no test article-related neoplastic changes.

## 2.2. Reviewer's analyses

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

### 2.2.1. Survival analysis

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

**Reviewer's findings:** This reviewer's analysis showed 58%, 65%, 63% and 72% survival of male rats, and 38%, 48%, 54% and 62% survival of female rats in control, low, medium, and high dose groups, respectively.

This reviewer's analysis showed statistically significant dose response relationship in mortality across treatment groups in female rats. The pairwise comparisons in female rats showed statistically significant increased mortality in high dose group compared to the control.

**Reviewer's comment:** *The sponsor's calculated percentage of survivals in male rat control and high dose groups were 57% and 71%, respectively, while those calculated by this reviewer were 58% and 72%, respectively. These differences were due to the fact that there was one animal in control group (animal #56) and one animal in high dose group (animal #252) that died due to natural causes during the terminal sacrifice weeks. In their calculations the sponsor did not count them with the terminally sacrificed animals, while this reviewer counted them with the terminally sacrificed animals.*

### 2.2.2. Tumor data analysis

The tumor data were analyzed for dose response relationships and pairwise comparisons of control group with each of the treated groups. Both the dose response relationship tests and pairwise comparisons were performed using the Poly-k method described in the paper of Bailer and Portier (1988) and Bieler and Williams (1993). In this method an animal that lives the full study period ( $w_{\max}$ ) or dies before the terminal sacrifice but develops the tumor type being tested gets a score of (b) (4). An animal that dies at week  $w_h$  without a tumor before the end of the study gets a score of (b) (4). The adjusted group size is defined as (b) (4). As an interpretation, an animal with score (b) (4) can be considered as a whole animal while an animal with score (b) (4) can be considered as a partial animal. The adjusted group size  $\Sigma s_h$  is equal to N (the original group size) if all animals live up to the end of the study or if each animal that dies before the terminal sacrifice develops at least one tumor, otherwise the adjusted group size is less than N. These adjusted group sizes are then used for the dose response relationship (or the pairwise) tests using the Cochran-Armitage test. One critical point for Poly-k test is the choice of the appropriate value of k, which depends on the tumor incidence pattern with the increased dose. For long term 104 week standard rat and mouse studies, a value of k=3 is suggested in the literature. Hence, this reviewer used k=3 for the analysis of this data. For the calculation of p-values the exact permutation method was used. The tumor rates and the p-values of the tested tumor types are listed in Tables 3A and 3B in the appendix for male and female rats, respectively. This reviewer used the same scores for his analysis as the sponsor used namely, 0, 25, 85, and 254 for both sexes of rats.

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

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

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

**Tumor Types with P-Values  $\leq 0.05$  for Dose Response Relationship or Pairwise Comparisons in Rats**

Sex	Organ Name	Tumor Name	Cont	Low	Med	High	P-Value			
			N=65	N=65	N=65	N=65	Dose Resp	C vs. L	C vs. M	C vs. H
=====										
Male	BRAIN	GRANULAR CELL TUMOUR	0	5	1	0	0.9287	0.0312*	0.5047	.
	H-POIETIC TUMOU	MALIGNANT LYMPHOMA	0	1	0	3	0.0395	0.5138	.	0.1391
	PITUITARY	ADENOMA, PARS DISTA	16	16	27	15	0.6510	0.5824	0.0315	0.5227
		ADN[PRS_DST+PRS_INT]	16	17	28	15	0.6766	0.5237	0.0205	0.5227
	WHOLE_BODY	LYMPHOMA	0	1	0	3	0.0395	0.5138	.	0.1391
Female	SKIN	BASAL CELL TUMOUR	0	0	0	3	0.0181*	.	.	0.1403

Based on the criteria of adjustment for multiple testing discussed above, the incidence of basal cell tumor on skin was considered to have statistically significant dose response relationship in female rats. The pairwise comparisons showed statistically significant increased incidence of granular cell tumor in brain in male rats low dose group compared to their control.

### 3. Mouse Study

Two separate experiments were conducted, one in males and one in females. In each of these two experiments there were three treated groups and one control group. Three hundred Crl:CD1(ICR) mice of each sex were randomly assigned to the treated and control groups in equal size of 75 animals per group. The initially selected target doses for male mice were 0 (vehicle), 58.6, 188 or 533  $\mu\text{g}/\text{kg}/\text{day}$ . These doses were given to males in Weeks 1-66. However, body weight gain of male mice given GSK573719 was lowered over Weeks 1-66, with slightly lowered food consumption noted in Week 1 for all male groups given the test article and throughout the study for male mice given 533  $\mu\text{g}/\text{kg}/\text{day}$ . As a result of this decrement in body weight gain, the dose levels were reduced in all treated groups in male mice from Week 67. From Week 67, the target doses for male mice were 0 (vehicle), 32.2, 102 or 295  $\mu\text{g}/\text{kg}/\text{day}$ . For female mice the targeted dose were 0 (vehicle), 20.8, 63.7 or 200  $\mu\text{g}/\text{kg}/\text{day}$  throughout the study. In this review these dose groups will be referred to as the low, medium, and high dose group, respectively. All rats received daily snout-only inhaled doses of GSK573719 or vehicle for 60 minutes/day (Weeks 1 to 72) and for 30 minutes/day (Week 73 onwards). Control animals received lactose (vehicle) with 1% (w/w) magnesium stearate alone.

During the administration period detailed observations in relation to dose administration were recorded daily during the first week of treatment, a minimum of twice weekly during Weeks 2 to 4, weekly during Weeks 5 to 13, once every two weeks during Weeks 14 to 52 and once every four weeks thereafter during the treatment period. Detailed clinical examinations were performed prior to initiation of dosing, and weekly until necropsy. The animals were palpated weekly for the presence and growth of masses. New masses were recorded at detection in respect to mass number, location, date, size and description.

The animals were weighed on the day following animal arrival, on the day that treatment commenced (Week 0), weekly for the first 16 weeks, and once every 4 weeks thereafter during the remaining treatment period, and on the day of necropsy. Male animals were also weighed at the beginning of Week 67 when the dose levels were reduced.

### 3.1. Sponsor's analyses

#### 3.1.1. Survival analysis

The sponsor analyzed the survival data of the mouse study using the same statistical methodologies as they used to analyze the survival data of rat study.

**Sponsor's findings:** Sponsor's analysis showed 53%, 40%, 44% and 53% survival of male mice and 40%, 59%, 39% and 45% survival of female mice in control, low, medium, and high dose groups, respectively. The sponsor analysis did not show statistically significant dose response relationship in mortality among the treated groups in either sex. None of the pairwise comparisons of treated group with the control was statistically significant in either sex.

#### 3.1.2. Tumor data analysis

The sponsor also analyzed the tumor data from the mouse study using the same statistical methodologies as they used to analyze the tumor data from rat study. The score used for male mice were 0, 49.0, 157 and 446, which are the weighted average of the two sets of weights used before Week 67 and after Week 67. The score used for female mice were 0, 20.8, 63.7 and 200, which were the targeted dose for female mice throughout the study.

**Sponsor's findings:** Sponsor's analyses did not show statistically significant dose response relationship among the treated groups, or higher tumor rates in the treated groups compared to the control in any of the tested tumor types in either sex.

### 3.2. Reviewer's analyses

This reviewer independently performed survival and tumor data analyses from the mouse study. For the mouse data analyses this reviewer used similar methodologies as he used to analyze the data from the rat study. Data used in this reviewer's analyses were provided by the sponsor electronically. For tumor data analysis, this reviewer used the same scores for his analysis as the sponsor used namely, 0, 49.0, 157 and 446 for male mice and 0, 20.8, 63.7 and 200 for female mice.

#### 3.2.1. Survival analysis

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

**Reviewer's findings:** This reviewer's analysis showed 53%, 40%, 44% and 53% survival of male mice, and 40%, 59%, 39% and 45% survival of female mice in Control, Low, Medium, and High dose groups, respectively. This reviewer's analysis did not show statistically significant dose response relationship in the mortality across treatment groups in either sex. The pairwise comparisons showed statistically significant increased mortality in female mice low dose group compared to the control.

#### 3.2.2. Tumor data analysis

The tumor rates and the p-values of the tumor types tested for dose response relationship and pairwise

comparisons of control and treated groups are given in Table 6A and 6B in the appendix for male and female mice, respectively.

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

**Tumor Types with P-Values  $\leq 0.05$  for Dose Response Relationship or Pairwise Comparisons in Mice**

Sex	Organ Name	Tumor Name	Cont	Low	Med	Hi gh	P-Val ue			
			N=75	N=75	N=75	N=75	Dose Resp	C vs. L	C vs. M	C vs. H
~~~~~										
Female	H-POI ETI C TUMOU	HI STI OCYTI C SARCOMA	1	1	5	5	0.0478	0.2708	0.0896	0.1206
	LUNGS + BRONCHI	BRONCHI OLOALVEOLAR AD O		2	5	3	0.1729	0.2708	0.0269*	0.1356
	WHOLE_BODY	HI STI OCYTI C_SARCOMA	1	1	5	5	0.0478	0.2708	0.0896	0.1206

Based on the multiple testing adjustment procedure discussed in the rat data analysis section, the incidence of none of the tested tumor types was considered to have statistically significant dose response relationship. The pairwise comparison showed statistically significant increased incidence of bronchiole alveolar adenoma collectively in lungs and bronchi in male mice medium dose group considered to be statistically significant compared to the control.

#### 4. Evaluation of validity of the design

As has been noted, none of the observed tumor types from rat or mouse study showed statistically significant dose response relationship. However, before drawing any conclusion regarding the carcinogenic or non-carcinogenic potential of a compound it is important to look into the following two issues, as have been pointed out in the paper by Haseman (1984).

- (i) Were enough animals exposed, for a sustained amount of time, to the risk of late developing tumors?
- (ii) Were dose levels high enough to pose a reasonable tumor challenge to the animals?

There is no consensus among experts regarding the number of animals and length of time at risk, although most carcinogenicity studies are designed to run for two years with about fifty to sixty animals per treatment group. The following are some rules of thumb regarding these two issues as suggested by experts in this field.

Haseman (1985) has done an investigation on the first issue. He gathered data from 21 studies using Fischer 344 rats and B6C3F1 mice conducted at the National Toxicology Program (NTP). It was found that, on the average, approximately 50% of the animals in the high dose group survived the two-year study period. Also, in a personal communication with Dr. Karl Lin of Division of Biometrics-6, Haseman suggested that, as a rule of thumb, a 50% survival of 50 initial animals or 20 to 30 animals still alive in the high dose group, between weeks 80-90, would be consider as a sufficient number and adequate exposure. In addition Chu, Cueto and Ward (1981), suggested that "to be considered adequate, an experiment that has not shown a chemical to be carcinogenic should have groups of animals with greater than 50% survival at one-year."

It appears, from these three sources that the proportions of survival at 52 weeks, 80-90 weeks, and two years are of interest in determining the adequacy of exposure and number of animals at risk.

Regarding the question of adequate dose levels, it is generally accepted that the high dose should be close to the maximum tolerated dose (MTD). In the paper of Chu, Cueto and Ward (1981), the following criteria are mentioned for dose adequacy. A high dose is considered as close to MTD if any of the criteria is met.

- (i) “A dose is considered adequate if there is a detectable loss in weight gain of up to 10% in a dosed group relative to the controls.”
- (ii) “The administered dose is also considered an MTD if dosed animals exhibit clinical signs or severe histopathologic toxic effects attributed to the chemical.”
- (iii) “In addition, doses are considered adequate if the dosed animals show a slight increased mortality compared to the controls.”

We will now investigate the validity of the GSK573719A rat and mouse carcinogenicity study, in the light of the above guidelines.

#### 4.1. Rat Study

The following is the summary of survival data of rats in the high dose groups:

##### Percentage of Survival in the High Dose Group at the End of Weeks 52, 78, and 91 in Rats

	Percentage of Survival		
	End of 52 weeks	End of 78 weeks	End of 91 weeks
Male	98%	87%	85%
Female	95%	83%	78%

Based on the survival criterion Haseman proposed, it may be concluded that enough rats were exposed to the high dose for a sufficient amount of time in both sexes.

The following table shows the percent difference in mean body weight gain in rats from the concurrent control, defined as

$$\text{Percent difference} = \frac{(\text{Final BW} - \text{Baseline BW})_{\text{Treated}} - (\text{Final BW} - \text{Baseline BW})_{\text{Control}}}{(\text{Final BW} - \text{Baseline BW})_{\text{Control}}} \times 100$$

##### Percent Difference in Mean body Weight Gain from Controls in Rats

Male			Female		
Low	Medium	High	Low	Medium	High
-3.76	-13.03	-16.29	-9.36	-8.51	-22.13

Source: Tables 7 of sponsor's submission

Therefore, relative to the control the male and female rat high dose group had about 16% and 22% decrement in their body weight gain, respectively.

The mortality rates at the end of the experiment were as follows:

#### **Mortality Rates at the End of the Experiment in Rats**

	Control	Low	Medium	High
Male	42%	35%	37%	28%
Female	62%	52%	46%	38%

This shows that the mortality rate of in the male high dose group is 14% lower than the control and that in female high dose group is 24% lower than the control.

Thus, even though the mortalities in the high dose group in both male and female rats are less than their respective controls, from the body weight gain data it can be concluded that the used high dose level might have reached the MTD in both sexes. For a final determination of the adequacy of the doses used, other clinical signs and histopathological toxic effects must be considered.

#### **4.2. Mouse Study**

The following is the summary of survival data of mice in the high dose groups:

#### **Percentage of Survival in the High Dose Group at the End of Weeks 52, 78, and 91 in Mice**

	Percentage of Survival		
	End of 52 weeks	End of 78 weeks	End of 91 weeks
Male	89%	76%	64%
Female	93%	80%	64%

Based on the survival criterion Haseman proposed, it may be concluded that enough mice were exposed to the high dose for a sufficient amount of time in both sexes.

The following table shows the percent difference in mean body weight gain in mice from the concurrent control,

#### **Percent Difference in Mean body Weight Gain from Controls in Mice**

Male			Female		
Low	Medium	High	Low	Medium	High
-4.63	-5.56	-8.33	8.16	1.02	4.08

Source: Tables 7 of sponsor's submission

Therefore, relative to control the male mice had less than 8% decrement in their body weight gain, while the female mice in high dose group had about 4% increment in their body weight gain.



The mortality rates at the end of the experiment were as follows:

#### Mortality Rates at the End of the Experiment in Mice

	Control	Low	Medium	High
Male	47%	60%	56%	47%
Female	60%	41%	61%	55%

This shows that the mortality rate of in the male mice high dose group is similar to the control, while in female mice the high dose group had 5% lower mortality than their control.

Thus, from the body weight gain data it can be concluded that the used high dose level might have reached the MTD in males. For female mice it might not have reached the MTD. For a final determination of the adequacy of the doses used for both male and female mice, other clinical signs and histopathological toxic effects must be considered.

### 5. Summary

In this submission the sponsor included reports of two animal carcinogenicity studies, one in rats and one in mice. These studies were intended to assess the carcinogenic potential of GSK573719A in rats and mice when administered by snout-only inhalation at appropriate drug levels for about 104 weeks. Results of this review have been discussed with the reviewing pharmacologist Dr. Sohn.

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

**Rat Study:** Two separate experiments were conducted, one in males and one in females. In each of these two experiments there were four treated groups and one control group. Two hundred and sixty Crl:CD®(SD) rats of each sex were randomly assigned to the treated and control groups in equal size of 65 animals per group. The initially selected target doses were 0, 30, 100 and 300 µg/kg/day. These doses were given to males and females for Weeks 1-72. However, during Weeks 0-72, there was a statistically significant dose-related decrease in the body weight gain in both sexes at all dose levels. This was accompanied by a minimal (statistically significant in females in all treated groups), reduction in food consumption. As a result of this decrement in body weight gain, the dose levels were reduced in all treated groups in both sexes from Week 73. From Week 73, the target doses were 0, 15, 50 and 150 µg/kg/day. All rats received daily snout-only inhaled doses of GSK573719 or vehicle for 60 minutes/day (Weeks 1 to 72) and for 30 minutes/day (Week 73 onwards). Control animals received lactose (vehicle) with 1% (w/w) magnesium stearate alone.

The tests showed statistically significant dose response relationship in mortality across treatment groups in female rats. The pairwise comparisons in female rats showed statistically significant increased mortality in high dose group compared to the control.

Tests showed statistically significant dose response relationship the incidence of basal cell tumor on skin in female rats. The pairwise comparisons showed statistically significant increased incidence of granular cell tumor in brain in male rats low dose group compared to the control.

**Mouse Study:** Two separate experiments were conducted, one in males and one in females. In each of these two experiments there were four treated groups and one control group. Three hundred Crl:CD1(ICR) mice of

each sex were randomly assigned to the treated and control groups in equal size of 75 animals per group. The initially selected target doses for male mice were 0 (vehicle), 58.6, 188 or 533 µg/kg/day. These doses were given to males Weeks 1-67. However, body weight gain of male mice given GSK573719 was lower over Weeks 1-66, with slightly lower food consumption noted in Week 1 for all male groups given the test article and throughout the study for male mice given 533 µg/kg/day. As a result of this decrement in body weight gain, the dose levels were reduced in all treated groups in male mice from Week 67. From Week 67, the target doses for male mice were 0 (vehicle), 32.2, 102 or 295 µg/kg/day. For female mice the targeted dose were 0 (vehicle), 20.8, 63.7 or 200 µg/kg/day throughout the study. All rats received daily snout-only inhaled doses of GSK573719 or vehicle for 60 minutes/day (Weeks 1 to 72) and for 30 minutes/day (Week 73 onwards). Control animals received lactose (vehicle) with 1% (w/w) magnesium stearate alone.

Tests did not show statistically significant dose response relationship in the mortality across treatment groups in either sex. The pairwise comparisons showed statistically significant increased mortality in female mice low dose group compared to the control. Tests did not show statistically significant dose response relationship among treatment groups in the incidence of any of the observed tumor types. The pairwise comparison showed statistically significant increased incidence of bronchiole alveolar adenoma collectively in lungs and bronchi in male mice medium dose group considered be statistically significant compared to the control.

**Evaluation of the study design:** The body weight gain data indicates that the used high dose level might have reached the MTD in both sexes of rats. The body weight gain data also indicates that the used high dose level might have reached the MTD in male mice. For female mice it might not have reached the MTD. For a final determination of the adequacy of the doses used for both rats and mice, other clinical signs and histopathological toxic effects must be considered.

Mohammad Atiar Rahman, Ph.D.  
Mathematical Statistician

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

cc:  
Archival NDA 203-975  
Dr. Sohn  
Ms. Hann

Dr. Machado  
Dr. Lin  
Dr. Rahman  
Ms. Patrician

## 6. Appendix

**Table 1A: Intercurrent Mortality Rate  
Male Rats**

Week	0 mg kg day		30/15mg kg day		100/50mg kg day		300/150mg kg day	
	No. of		No. of		No. of		No. of	
	Death	Cum. %	Death	Cum. %	Death	Cum. %	Death	Cum. %
~~~~~								
0 - 52	2	3.08	2	3.08	3	4.62	1	1.54
53 - 78	7	13.85	6	12.31	4	10.77	7	12.31
79 - 91	5	21.54	6	21.54	9	24.62	2	15.38
92 - 104	13	41.54	9	35.38	8	36.92	8	27.69
Ter. Sac.	38	58.46	42	64.62	41	63.08	47	72.31
-----								
Total	N=65		N=65		N=65		N=65	

**Table 1B: Intercurrent Mortality Rate  
Female Rats**

Week	0 mg kg day		30/15mg kg day		100/50mg kg day		300/150mg kg day	
	No. of		No. of		No. of		No. of	
	Death	Cum. %	Death	Cum. %	Death	Cum. %	Death	Cum. %
~~~~~								
0 - 52	3	4.62	5	7.69	2	3.08	3	4.62
53 - 78	4	10.77	8	20.00	9	16.92	8	16.92
79 - 91	15	33.85	11	36.92	7	27.69	3	21.54
92 - 104	18	61.54	10	52.31	12	46.15	11	38.46
Ter. Sac.	25	38.46	31	47.69	35	53.85	40	61.54
-----								
Total	N=65		N=65		N=65		N=65	

**Table 2A: Intercurrent Mortality Comparison  
Male Rats**

Test	Statistic	P_Value
~~~~~		
Dose-Response	Likelihood Ratio	0.1604
Homogeneity	Log-Rank	0.4580

**Table 2B: Intercurrent Mortality Comparison  
Female Rats**

Test	Statistic	P_Value
~~~~~		
Dose-Response	Likelihood Ratio	0.0161
Homogeneity	Log-Rank	0.0863

**Table 3A: Tumor Rates and P-Values for Dose Response Relationship and Pairwise Comparisons  
Male Rats**

Organ Name	Tumor Name	0mg	30/15mg	100/50mg	300/150mg	P-Value			
		Cont N=65	Low N=65	Med N=65	High N=65	Dose Resp	C vs. L	C vs. M	C vs. H
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%									

**Table 3A: Tumor Rates and P-Values for Dose Response Relationship and Pairwise Comparisons  
Male Rats**

Organ Name	Tumor Name	0mg	30/15mg	100/50mg	300/150mg	P-Value			
		Cont N=65	Low N=65	Med N=65	High N=65	Dose Resp	C vs. L	C vs. M	C vs. H
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%									

**Table 3A: Tumor Rates and P-Values for Dose Response Relationship and Pairwise Comparisons  
Male Rats**

		0mg	30/15mg	100/50mg	300/150mg	P-Value			
Organ Name	Tumor Name	Cont N=65	Low N=65	Med N=65	High N=65	Dose Resp	C vs. L	C vs. M	C vs. H
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%									



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

		0mg	30/15mg	100/50mg	300/150mg	P-Value			
Organ Name	Tumor Name	Cont N=65	Low N=65	Med N=65	High N=65	Dose Resp	C vs. L	C vs. M	C vs. H
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%									



**Table 4A: Intercurrent Mortality Rate in Male Mice**

Week	0 mg/kg/day		58.6/32.2mg/kg/day		188/102mg/kg/day		533/295 mg/kg/day	
	No. of		No. of		No. of		No. of	
	Death	Cum. %	Death	Cum. %	Death	Cum. %	Death	Cum. %
~~~~~								
0 - 52	6	8.00	10	13.33	6	8.00	8	10.67
53 - 78	9	20.00	14	32.00	14	26.67	10	24.00
79 - 91	8	30.67	9	44.00	10	40.00	9	36.00
92 - 104	12	46.67	12	60.00	12	56.00	8	46.67
Ter. Sac.	40	53.33	30	40.00	33	44.00	40	53.33
-----								
Total	N=75		N=75		N=75		N=75	

**Table 4B: Intercurrent Mortality Rate Female Mice**

Week	0 mg/kg/day		20.8 mg/kg/day		63.7 mg/kg/day		200 mg/kg/day	
	No. of		No. of		No. of		No. of	
	Death	Cum. %	Death	Cum. %	Death	Cum. %	Death	Cum. %
~~~~~								
0 - 52	6	8.00	4	5.33	6	8.00	5	6.67
53 - 78	13	25.33	12	21.33	15	28.00	10	20.00
79 - 91	9	37.33	8	32.00	12	44.00	12	36.00
92 - 104	17	60.00	7	41.33	13	61.33	14	54.67
Ter. Sac.	30	40.00	44	58.67	29	38.67	34	45.33
-----								
Total	N=75		N=75		N=75		N=75	

**Table 5A: Intercurrent Mortality Comparison Male Mice**

Test	Statistic	P_Value
~~~~~		
Dose-Response	Likelihood Ratio	0.7215
Homogeneity	Log-Rank	0.2576

**Table 5B: Intercurrent Mortality Comparison Female Mice**

Test	Statistic	P_Value
~~~~~		
Dose-Response	Likelihood Ratio	0.8530
Homogeneity	Log-Rank	0.0640

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

		0 mg	58.6/ 32.2mg	188/ 102mg	533/ 295mg	P-Val ue			
Organ Name	Tumor Name	Cont N=75	Low N=75	Med N=75	Hi gh N=75	Dose Resp	C vs. L	C vs. M	C vs. H
f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f									

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

		0 mg	58.6/ 32.2mg	188/ 102mg	533/ 295mg	P-Val ue			
Organ Name	Tumor Name	Cont N=75	Low N=75	Med N=75	Hi gh N=75	Dose Resp	C vs. L	C vs. M	C vs. H
f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f f									

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

Organ Name	Tumor Name	0 mg	20. 8mg	63. 7mg	200mg	P-Val ue			
		Cont N=75	Low N=75	Med N=75	Hi gh N=75	Dose Resp	C vs. L	C vs. M	C vs. H
ADIPOSE TISSUE	HIBERNOMA	1	0	0	0	0. 7546	0. 5225	0. 4854	0. 5093
ADRENALS	PHAEOCHROMOCYTOMA	1	0	0	0	0. 7546	0. 5225	0. 4854	0. 5093
	SUBCAPSULAR CELL ADE	1	1	1	1	0. 4172	0. 2708	0. 7377	0. 2570
BONE	CHONDROSARCOMA	0	0	1	1	0. 1824	.	0. 4854	0. 5093
	OSTEOSARCOMA	0	1	0	0	0. 4839	0. 5268	.	.
BRAIN	ASTROCYTOMA	0	0	0	1	0. 2546	.	.	0. 5093
COLON	ADENOCARCINOMA	0	0	1	0	0. 2546	.	0. 4854	.
GALL BLADDER	PAPILLOMA	0	0	1	0	0. 2546	.	0. 4854	.
H-POIETIC TUMOU	HISTIOCYTIC SARCOMA	1	1	5	5	0. 0478	0. 2708	0. 0896	0. 1206
	MALIGNANT LYMPHOMA	11	9	11	8	0. 7258	0. 6600	0. 5378	0. 7063
HARDERIAN GLAND	ADENOMA	5	6	3	7	0. 2454	0. 5742	0. 5981	0. 3936
ILEUM	HAEMANGIOMA	0	0	0	1	0. 2546	.	.	0. 5093
LIVER	HAEMANGIOMA	0	1	1	0	0. 4980	0. 5225	0. 4904	.
	HEPATOCELLULAR ADENO	1	0	0	0	0. 7546	0. 5225	0. 4854	0. 5093
LN MESENTERIC	HAEMANGIOMA	1	0	0	0	0. 7512	0. 5179	0. 4808	0. 5046
LUNGS + BRONCHI	BRONCHIOALVEOLAR A	0	2	5	3	0. 1729	0. 2708	0. 0269*	0. 1356
		7	7	5	8	0. 3790	0. 4732	0. 5778	0. 5610
MAMMARY	ADENOACANTHOMA	0	0	1	0	0. 2546	.	0. 4854	.
	MAMMARY ADENOCARCINO	2	3	0	0	0. 9726	0. 5511	0. 7377	0. 7615
	MAMMARY ADENOMA	1	0	0	0	0. 7546	0. 5225	0. 4854	0. 5093
	MAMMARY_ADE+ADENOCAR								
	+ ADENOACANTHOMA	3	3	1	0	0. 9734	0. 3919	0. 6688	0. 8853
	MAMMARY_ADENOMA								
	+ ADENOCARCINOMA	3	3	0	0	0. 9892	0. 3919	0. 8675	0. 8853
OESOPHAGUS	SQUAMOUS CELL PAPILL	0	0	0	1	0. 2581	.	.	0. 5138
OVARIES	CYSTADENOMA	0	0	0	1	0. 2546	.	.	0. 5093
	CYSTADENOMA+TUBELOST	0	1	0	2	0. 1127	0. 5225	.	0. 2570
	HAEMANGIOMA	1	0	1	0	0. 6230	0. 5225	0. 7377	0. 5093
	HAEMANGIOSARCOMA	0	0	0	1	0. 2581	.	.	0. 5138
	LUTEOMA	1	2	0	2	0. 3394	0. 5341	0. 4854	0. 5140
	MESOVARIAN LEIOMYOSA	2	0	0	0	0. 9407	0. 7743	0. 7377	0. 7615
	TUBULOSTOMAL ADENOMA	0	1	0	1	0. 3198	0. 5225	.	0. 5093
PANCREAS	ISLET CELL ADENOMA	1	0	0	0	0. 7546	0. 5225	0. 4854	0. 5093
PITUITARY	ADENOMA, PARS DISTA	0	2	2	1	0. 4389	0. 2753	0. 2381	0. 5093

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

		0 mg	20.8mg	63.7mg	200mg	P-Value			
Organ Name	Tumor Name	Cont N=75	Low N=75	Med N=75	Hi gh N=75	Dose Resp C	vs. L	C vs. M	C vs. H
PI TUI TARY	ADENOMA, PARS I NTERM	0	1	0	0	0.4861	0.5225	.	.
SALI VARY GLANDS	OSTEOSARCOMA	0	0	1	0	0.2535	.	0.4904	.
SKELETAL MUSCLE	FIBROSARCOMA	0	0	1	0	0.2546	.	0.4854	.
SKI N	FIBROSARCOMA	0	0	0	1	0.2546	.	.	0.5093
	FIBROUS HISTIOCYTOMA	1	0	1	0	0.6230	0.5225	0.7377	0.5093
	SQUAMOUS CELL CARCIN	1	0	0	0	0.7546	0.5225	0.4854	0.5093
	SQUAMOUS CELL PAPILL	0	2	0	2	0.1937	0.2708	.	0.2570
	SQUAMOUS_CELL_PAPILL	0	2	0	2	0.1937	0.2708	.	0.2570
STOMACH	ADENOMA	0	0	0	1	0.2546	.	.	0.5093
	OSTEOSARCOMA	0	0	1	0	0.2535	.	0.4904	.
	SQUAMOUS CELL PAPILL	1	0	0	0	0.7546	0.5225	0.4854	0.5093
URI NARY BLADDER	MESENCHYMAL TUMOUR	0	1	0	0	0.4861	0.5225	.	.
UTERI NE CERVI X	GRANULAR CELL TUMOUR	0	1	0	0	0.4861	0.5225	.	.
	LEI OMYOMA	0	1	0	1	0.3198	0.5225	.	0.5093
	SQUAMOUS CELL PAPILL	0	0	0	1	0.2546	.	.	0.5093
	STROMAL CELL SARCOMA	0	2	0	1	0.4384	0.2753	.	0.5093
UTERUS	DECI DUOMA	1	1	0	0	0.8053	0.2659	0.4808	0.5046
	ENDOMETRI AL ADENOCAR	0	1	0	0	0.4861	0.5225	.	.
	ENDOMETRI AL POLYP	3	3	6	2	0.6864	0.3741	0.2162	0.5088
	ENDOMETRI AL STROMAL	2	0	1	2	0.2861	0.7698	0.4709	0.3160
	ENDOMETRI AL_POLYP+ADE	3	4	6	2	0.7406	0.5401	0.2162	0.5088
	HAEMANGI OMA	2	1	0	3	0.1681	0.5335	0.7328	0.5088
	HAEMANGI OSARCOMA	0	0	0	1	0.2546	.	.	0.5093
	LEI OMYOMA	2	2	4	2	0.5024	0.3439	0.3209	0.3231
	LEI OMYOSARCOMA	1	0	0	0	0.7546	0.5225	0.4854	0.5093
	OSTEOSARCOMA	1	0	0	0	0.7546	0.5225	0.4854	0.5093
VAGI NA	POLYP	1	0	0	0	0.7546	0.5225	0.4854	0.5093
WHOLE_BODY	HAEMANGI OMA+HAEMANGI	4	2	2	5	0.1884	0.6934	0.6241	0.5115
	HISTIOCYTIC_SARCOMA	1	1	5	5	0.0478	0.2708	0.0896	0.1206
	LYMPHOMA	11	9	11	8	0.7258	0.6600	0.5378	0.7063

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

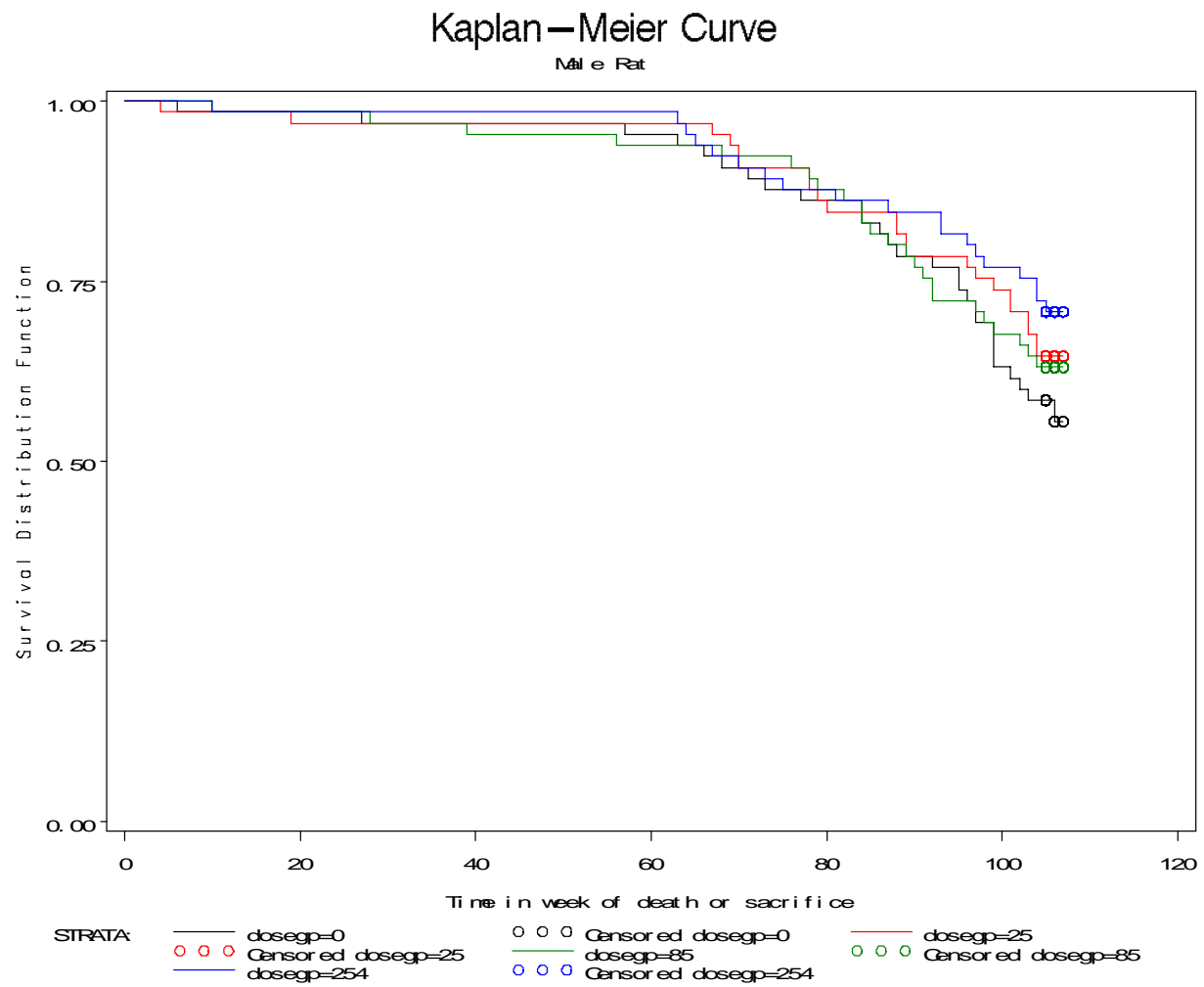


Figure 1B: Kaplan-Meier Survival Functions for Female Rats

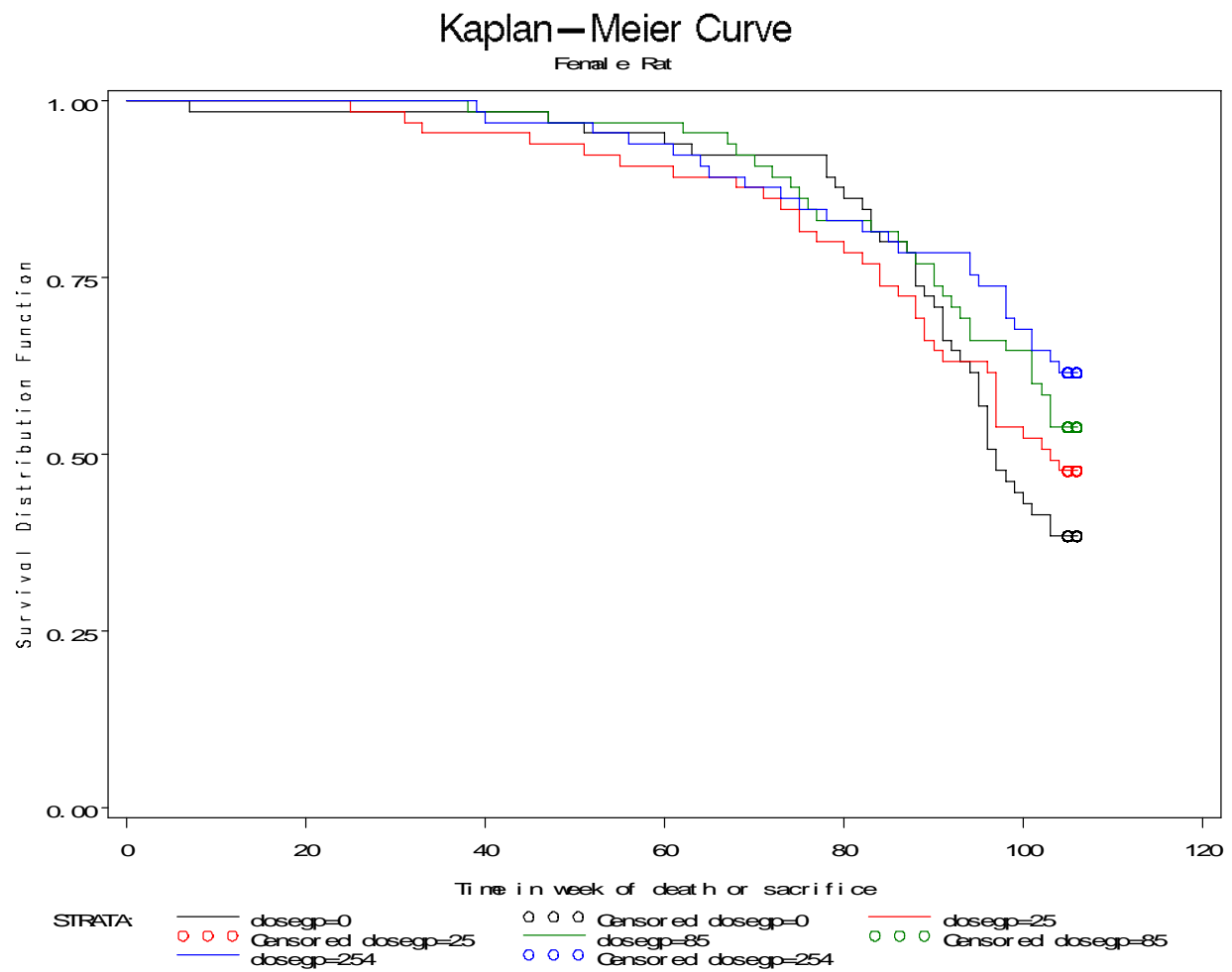


Figure 2A: Kaplan-Meier Survival Functions for Male Mice

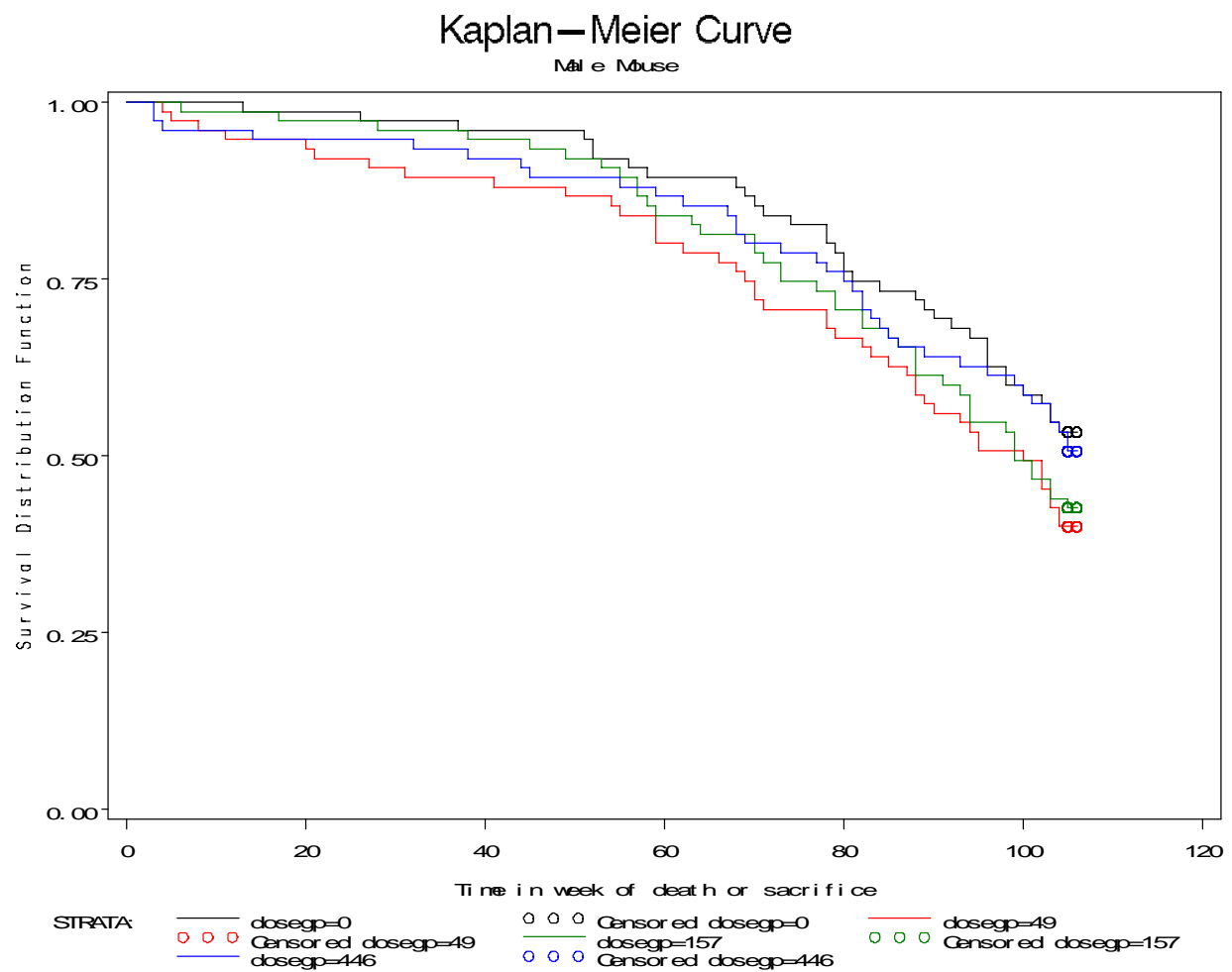
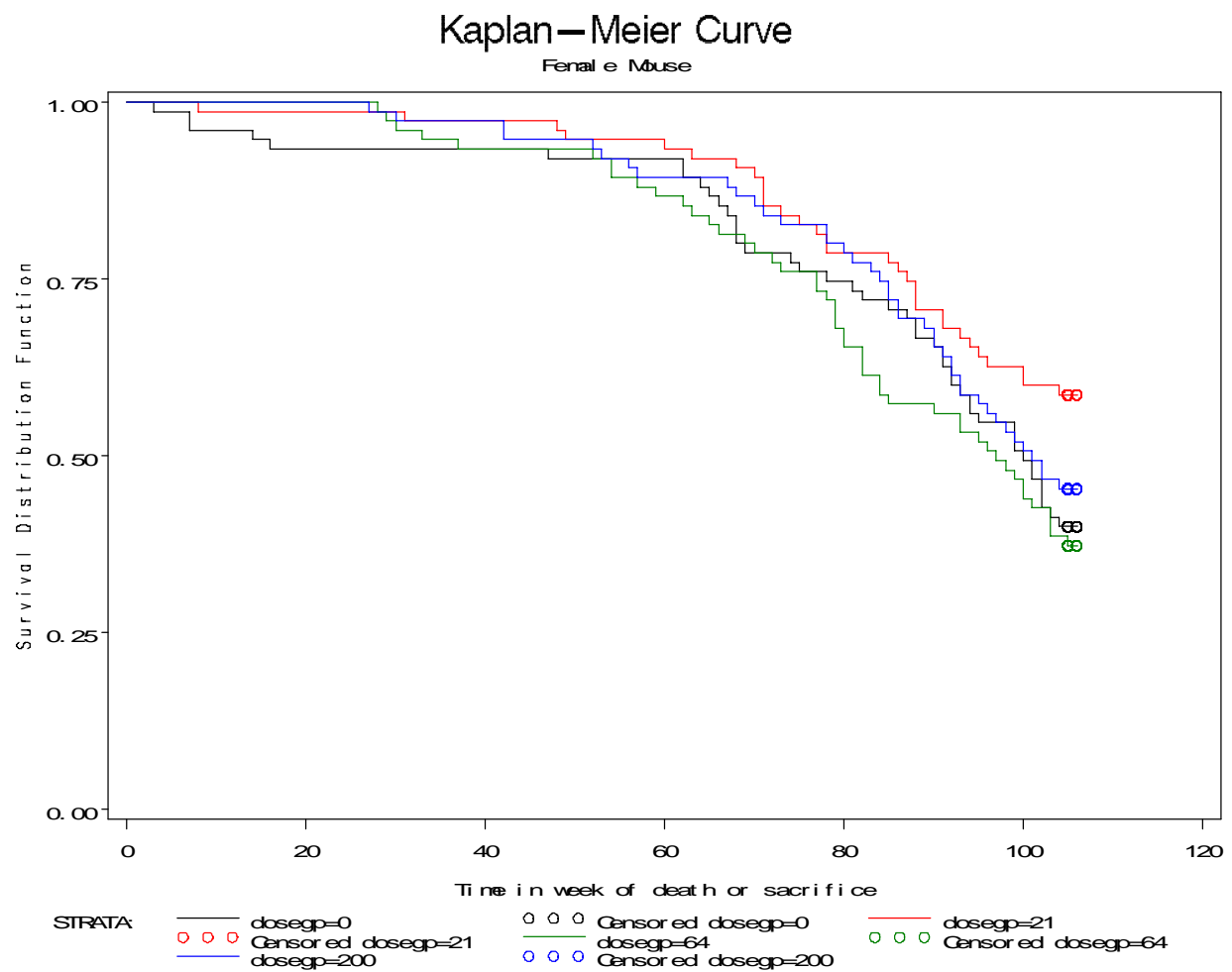




Figure 2B: Kaplan-Meier Survival Functions for Female Mice



## 7. References

1. Peto, R., M.C. Pike, N.E. Day, R.G. Gray, P.N. Lee, S. Parish, J. Peto, Richards, and J. Wahrendorf, "Guidelines for sample sensitive significance test for carcinogenic effects in long-term animal experiments", Long term and short term screening assays for carcinogens: A critical appraisal, International agency for research against cancer monographs, *Annex to supplement, World Health Organization, Geneva*, 311-426, 1980.
2. Bailer AJ, Portier CJ (1988). "Effects of treatment-induced mortality and tumor-induced mortality on tests for carcinogenicity in small samples." *Biometrics*, 44, 417-431.
3. Bieler, G. S. and Williams, R. L. (1993). "Ratio estimates, the delta method, and quantal response tests for increased carcinogenicity". *Biometrics* 49, 793-801.
4. LIN, K.K.(2000) Carcinogenicity Studies of Pharmaceuticals. In: *Encyclopedia of Biopharmaceutical Statistics*, ed. Shein-Chung Chow, Marcel Dekker, New York.
5. Lin K.K. and Rahman M.A., "Overall false positive rates in tests for linear trend in tumor incidence in animal carcinogenicity studies of new drugs", *Journal of Biopharmaceutical Statistics*, 8(1), 1-15, 1998.
7. Haseman, J, "A re-examination of false-positive rates for carcinogenesis studies", *Fundamental and Applied Toxicology*, 3: 334-339, 1983.
8. Guidance for Industry. Statistical Aspects of the Design, Analysis, and Interpretation of Chronic Rodent Carcinogenicity Studies of Pharmaceuticals (Draft Guidance). U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), May 2001.

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

/s/  
-----

MOHAMMAD A RAHMAN  
04/02/2013

KARL K LIN  
04/02/2013  
Concur with review