CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

205382Orig1s000

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 205-382
Drug Name:	Umeclidinium
Indication:	Long-term, once-daily, maintenance bronchodilator treatment of airflow obstruction in patients with chronic obstructive pulmonary disease (COPD)
Applicant:	GlaxoSmithKline
Dates:	NDA submitted: April 30, 2013 PDUFA due date: April 30, 2014
Review Priority:	Standard
Biometrics Division:	Division of Biometrics II
Statistical Reviewer:	Gregory Levin, PhD
Concurring Reviewers:	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:	Angela Ramsey

Keywords: NDA review, clinical studies, missing data, surrogate outcomes, labeling

Table of Contents

LIST OF TABLES

LIST OF FIGURES

1 EXECUTIVE SUMMARY

This review considers the inhaled long-acting muscarinic antagonist umeclidinium (UMEC) for long-term, once-daily, maintenance bronchodilator treatment of airflow obstruction in patients with chronic obstructive pulmonary disease (COPD). We focus on three phase 3, multicenter, randomized, double-blind, parallel-group, placebo-controlled clinical trials designed to evaluate the efficacy of UMEC 62.5 and/or 125 mcg with respect to pulmonary function at 12 or 24 weeks. Patients in these studies had moderate-to-severe COPD, 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 salbutamol was provided for as-needed daily relief medication.

There was statistical evidence of benefit for UMEC 62.5 mcg with respect to the primary endpoint, change from baseline in trough FEV_1, in two independent phase 3 clinical trials. Treatment with UMEC 62.5 mcg provided 0.12 L (95% confidence interval [CI]: 0.08, 0.16) and 0.13 L (95% CI: 0.05, 0.20) mean improvements over placebo in trough FEV_1 in these trials, which were 24 weeks and 12 weeks in duration, respectively. Similar results were observed for the higher 125 mcg dose of UMEC in two studies. Estimated treatment effects for UMEC were largely consistent across subgroups of interest, including sex, age, race, and geographic region.

We consider FEV_1 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_1 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, and St. George's Respiratory Questionnaire (SGRQ) score. The observed trends toward benefit for these endpoints increase confidence that the treatment effect on the surrogate endpoint FEV_1 is likely to predict clinical benefit, i.e., improvements in how COPD patients function, feel, or survive.

There were substantial missing data in the phase 3 efficacy studies, with overall dropout rates ranging from 18–25%. 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 assumes that the treatment effect observed before dropout would have persisted even after patients stopped taking the therapy. This assumption is implausible because bronchodilators are generally considered symptomatic and not disease-modifying therapies, and any FEV_1 improvement attributable to a bronchodilator will likely go away within a few days of treatment discontinuation. Therefore, we gave importance to a supportive analysis that multiply imputed missing data under the assumption that dropouts on UMEC would have had outcomes similar to those that were observed among completers in the control group. Supportive analyses provided consistent evidence of superiority to placebo, but estimated treatment effect sizes were approximately 20–30% less than in the primary analyses. For example, in Study 373, the estimated mean improvement in FEV_1 on UMEC 62.5, relative to placebo, was 0.09 L (95% CI: 0.05, 0.13), as compared to 0.12 L (95% CI: 0.08, 0.16) in the primary analysis.

The complete safety evaluation was conducted by Dr. Jennifer Pippins, the Medical Reviewer, but we performed additional analyses to explore cardiovascular risk. Rates of major adverse cardiac events (MACE) were similar across the treatment arms, but an analysis of cardiovascular-related serious adverse events suggested a possible trend toward greater risk on UMEC as compared to placebo and tiotropium.

Small numbers of events led to considerable statistical uncertainty around the estimated differences in risks between the treatment arms. The interpretability of safety analyses is also clouded by the high rates of missing data in the phase 3 studies. The large amount of missing data was primarily due to the trial design, as patients who discontinued treatment early were not followed up for an evaluation of safety through the complete double-blind study duration.

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 steroids. Bronchodilators, usually administered through an inhaler, relax muscles around the airways in order to improve airflow and relieve symptoms. There are two major types of bronchodilators: β_2 agonists, which act on β_2 receptors, and muscarinic antagonists, which inhibit the action of cholinergic nerves. Bronchodilators may be either 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 β_2 agonist salbutamol, short-acting muscarinic antagonist ipratropium, long-acting β_2 agonists (LABAs) salmeterol and formoterol, and long-acting muscarinic antagonists (LAMAs) tiotropium and aclidinium. FDA has also approved inhalers that combine a LABA and inhaled corticosteroid (ICS), such as Advair (salmeterol and fluticasone propionate), Symbicort (formoterol and budesonide), and Breo (vilanterol and fluticasone furoate).

This review considers the inhaled long-acting muscarinic antagonist umeclidinium (UMEC) for long-term, once-daily, maintenance bronchodilator treatment of airflow obstruction in patients with COPD. Two doses of UMEC, 62.5 mcg once daily and 125 mcg once daily, were evaluated in the phase 3 clinical development program, but only the lower 62.5 mcg dose is proposed for approval. We often omit the mcg unit when referring to the dose of UMEC 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 for treatment of airflow obstruction in patients with COPD. The clinical development program for UMEC was introduced to the Division of Pulmonary, Allergy, and Rheumatology Products under IND 104,479. Anoro, a related GlaxoSmithKline combination product consisting of umeclidinium 62.5 and vilanterol (VI) 25, was approved in December 2013 for the long-term, once-daily, maintenance treatment of airflow obstruction in patients with COPD. Several of the phase 3 studies designed to evaluate the efficacy of the UMEC/VI combination product also included an evaluation of the efficacy of the UMEC

monotherapy. More details on the evaluation of the UMEC/VI combination product are available in the reviews of NDA 203-975.

We next summarize important meetings and correspondence with the applicant relevant to this review. An end-of-phase 2 meeting to discuss the development of UMEC/VI and the monotherapies was held on October 29, 2010. FDA generally agreed with the two proposed placebo-controlled phase 3 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_1) 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. It was also 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)

[redacted]

FDA submitted an information request to the applicant on February 24, 2013 (during the UMEC/VI NDA review) 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). Results for Study 408 based on these sensitivity analyses were later submitted, as well.

2.1.3 Specific Studies Reviewed

This review focuses on three placebo-controlled phase 3 clinical trials designed to evaluate the efficacy of UMEC 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. Study AC4115408 (408) was a 12-week, randomized, double-blind, parallel-group, placebo-controlled clinical

trial. Only Studies 408 and 373 included a UMEC 62.5 mcg treatment arm; the dose of UMEC was 125 mcg in Study 361.

We also discuss results from four additional phase 3 studies of umeclidinium. Study DB2113374 (374) was a 24-week, randomized, double-blind, parallel-group, active-controlled clinical trial in which the LAMA tiotropium was the active treatment for comparison. Studies DB2114417 (417) and DB2114418 (418) were 12-week, randomized, double-blind, placebo-controlled, incomplete block, cross-over clinical trials to evaluate efficacy 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 125 and UMEC/VI 125/25.

Finally, we briefly comment on several studies used to support the dose selection of umeclidinium. Studies AC4113589 (589), AC4113073 (73), and AC4115321 (321) were randomized, double-blind, placebo-controlled, dose-ranging studies.

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\NDA205382\205382.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 Studies 361, 373, and 374

Studies 361, 373, and 374 were designed to evaluate the 24-week efficacy of the once-daily bronchodilator UMEC/VI, as well as that of the UMEC and VI monotherapies, for treatment of airflow obstruction in COPD. The three 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 (≥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 ≥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 ≤1000mg/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 and 3.2.5.2 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 single 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_1 at earlier time points, mean Transition Dyspnea Index (TDI) focal score, St. George's Respiratory Questionnaire (SGRQ) total 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, or placebo. 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. Study 374 was a tiotropium-controlled trial with 1:1:1:1 randomization to UMEC/VI 125/25 mcg, UMEC 62.5/25 mcg, UMEC 125 mcg, or tiotropium. 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.

The sample size estimates in these studies account for the increased ***variability*** that would occur because of 30% expected missing data on the primary endpoint (assessed at 24 weeks). However, missing data can also induce ***bias***, which cannot be corrected for with an increase in the sample size. In addition, any attempts to correct for potential biases in analyses must be based on unverifiable assumptions about the nature of the missing data (see 3.2.5.2 for further discussion).

3.2.1.2 Study 408

Study 408 was a 12-week, phase 3, multicenter, randomized, double-blind, parallel-group, placebo-controlled clinical trial to evaluate the efficacy of UMEC 62.5 and 125 mcg with respect to airflow obstruction. COPD patients with an extensive smoking history, moderate-to-severe airflow obstruction,

and dyspnea were randomized 1:1:1 to UMEC 62.5, UMEC 125, or placebo. Concomitant use of systemic corticosteroids or additional long-acting bronchodilators was prohibited, but patients were permitted to use inhaled corticosteroids at a stable dose ≤1000mg/day and study-provided salbutamol for as-needed relief medication. There was a 5- to 9-day run-in period followed by a 12-week treatment period. Visits occurred at Days 1, 2, and 14, Weeks 4, 8, and 12, and 1 day after Week 12 (Day 85). Postdose 6-hour serial spirometry data was collected on Day 1, and Weeks 4 and 12. A total sample size of 198 patients was planned to provide 90% power to detect a 130 mL difference between two of the treatment groups (accounting for the increased variability induced with 15% missing data).

The primary endpoint was the change from baseline in predose trough FEV_1 on Day 85. The secondary endpoints were the weighted mean FEV_1 0-6 hours postdose on Day 85, and serial FEV_1 at 1, 3, 6, 23, and 24 hours after dosing at Day 1 and Week 12. Additional endpoints of interest included FEV_1 at earlier time points, TDI focal score, SGRQ score, rescue salbutamol use, and several other spirometric measures. 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).

3.2.1.3 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) ≥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 ≤1000mg/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.

The study was primarily designed to evaluate the efficacy of UMEC/VI, with co-primary endpoints of change from period baseline in exercise endurance time (EET) and trough FEV_1 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 to provide evidence of efficacy for UMEC/VI. Trough FEV_1 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_1 at Week 12.

Study 359 was a 52-week, randomized, double-blind, parallel-group, placebo-controlled clinical trial to evaluate the safety and tolerability of UMEC/VI and UMEC. 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. Study 359 was the only phase 3 trial in

which Holter monitoring was carried out in all patients. Monitoring occurred at the Month 3, 6, 9, and 12 visits. Possible reasons for withdrawal from the study included adverse event, lack of efficacy, and protocol-defined stopping criteria based on electrocardiogram (ECG), Holter, or other laboratory abnormalities.

Studies 589, 73, and 321 were randomized, double-blind, placebo-controlled, dose-ranging studies for 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, and Study 321 was an incomplete block, 3-period (7 days per period) cross-over 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.

3.2.2 Statistical Methodologies

3.2.2.1 Studies 361, 373, and 374

In Studies 361, 373, 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_1 at Day 169. The model used FEV_1 measurements at Days 2, 28, 56, 84, 112, 168, and 169, and included the following covariates: treatment group, baseline FEV_1, center group, smoking status, visit (categorical variable), visit-by-baseline FEV_1 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_1 measurements within patients over time. The MMRM analysis 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_1, 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 and similar confidence intervals (results not shown) to the primary analyses.

The analysis of the secondary endpoint, weighted mean FEV_1 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 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 were based on logistic regression models adjusting for baseline value, smoking status, and center group. Patients who withdrew from the study early were considered non-responders in analyses of the proportion achieving some threshold change in an outcome (e.g., a 4-unit or greater decrease in SGRQ total score).

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_1), and then for the secondary endpoint (weighted mean FEV_1). In Study 374, the following treatment comparisons were performed in order: (1) UMEC/VI versus tiotropium; (2) UMEC/VI versus UMEC. There were no planned direct comparisons between UMEC and the active comparator tiotropium. The applicant did not control for multiplicity across other efficacy

endpoints (e.g., SGRQ, SOBDA, rescue medication use, time to COPD exacerbation) in any of the studies.

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 analysis.

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, regardless of adherence, 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_1 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 UMEC were on control (rather than UMEC) 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 UMEC in all randomized participants (often called the intention-to-treat or de facto estimand).

The Jump to Reference multiple imputation approach was performed for the primary endpoint trough FEV_1, as well as for other supportive endpoints, including 0-6 hour weighted mean FEV_1 and SGRQ total score. 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.

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 Study 408

The statistical methods used for Study 408 were analogous to those described for Studies 361, 373, and 374. In particular, the primary efficacy analysis was based on an analogous mixed effects model for repeated measures to compare treatment groups with respect to the mean change from baseline in trough FEV_1 at 12 weeks. To account for the multiple statistical tests, the applicant sequentially performed the following comparisons in order: (1) UMEC 125 versus placebo; and (2) UMEC 62.5 versus placebo. There was no multiplicity control across analyses of secondary and other endpoints. The applicant carried out the same set of missing data sensitivity analyses as described previously.

3.2.2.3 Additional Studies

In Studies 417 and 418, the primary efficacy analyses were based on MMRMs to compare treatment groups with respect to the mean change from baseline in EET, and trough FEV_1, 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_1.

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_1) 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.

Study 359 was a safety trial and therefore did not have prespecified primary efficacy analyses. However, exploratory efficacy analyses were conducted for trough FEV_1, rescue puffs per day, and time to exacerbation using analogous methods to those described for the other phase 3 trials.

3.2.3 Dose Selection

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_1 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.

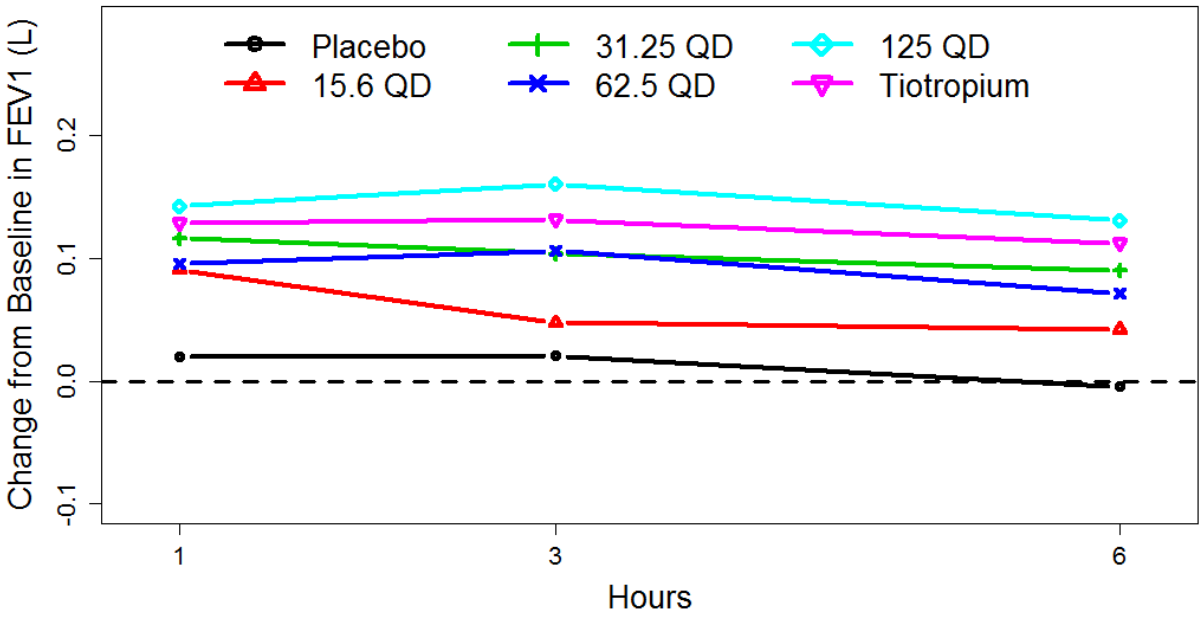
Study 321 was the only clinical trial that evaluated doses lower than 62.5 mcg. In the 6 hours postdose at Day 1, there was clear separation in efficacy between all once-daily UMEC doses and placebo (Figure 1). UMEC 15.6 mcg demonstrated the smallest improvement in FEV_1 and UMEC 125 mcg demonstrated 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 changes from baseline in trough FEV_1 at Day 8 (Figure 3). Figure 4 and Figure 5 present postdose time-response profiles at Days 1 and 7, respectively, for both once- and twice-daily doses of UMEC. Average FEV_1 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_1 at Day 8 (Figure 6).

Table 1. Mean Differences from Placebo in Change from Baseline in Trough FEV_1 for Different Once-Daily Doses of Umeclidinium in Studies 321, 73, 589, and 408

	Difference from Placebo for LS Mean Change from Baseline in Trough FEV_1 (L) (95% CI) by once daily UMEC dose (mcg) [a]						
Study	**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)			

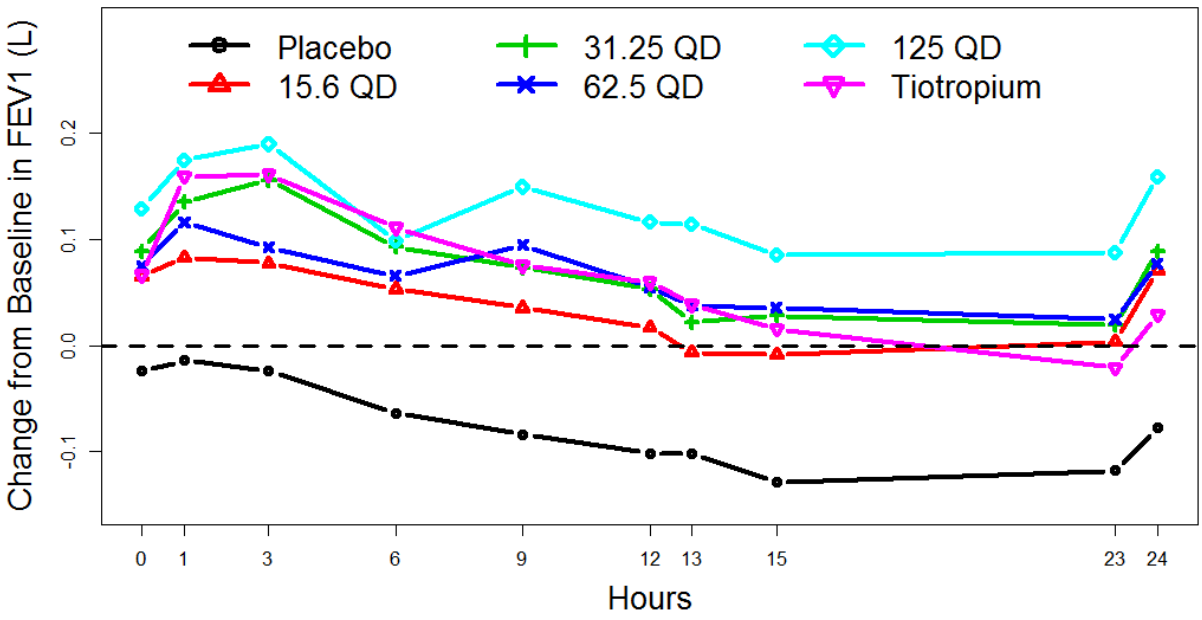
Source: Table 83, Applicant's Summary of Clinical Efficacy

Figure 1. Postdose 6-Hour Serial Mean Change from Baseline in FEV_1 at Day 1 for Placebo, Different Once-Daily Umeclidinium Doses, and Tiotropium, in Study 321



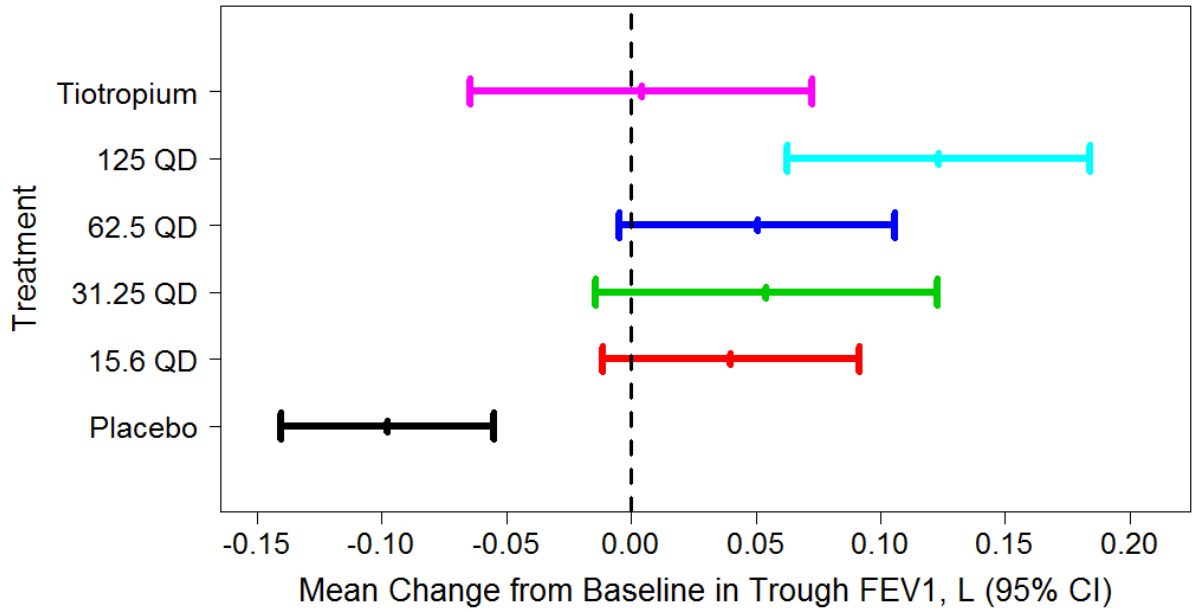
Abbreviations: QD = once-daily, BD = twice-daily

Figure 2. Postdose 24-Hour Serial Mean Change from Baseline in FEV_1 at Day 7 for Placebo, Different Once-Daily Umeclidinium Doses, and Tiotropium, in Study 321



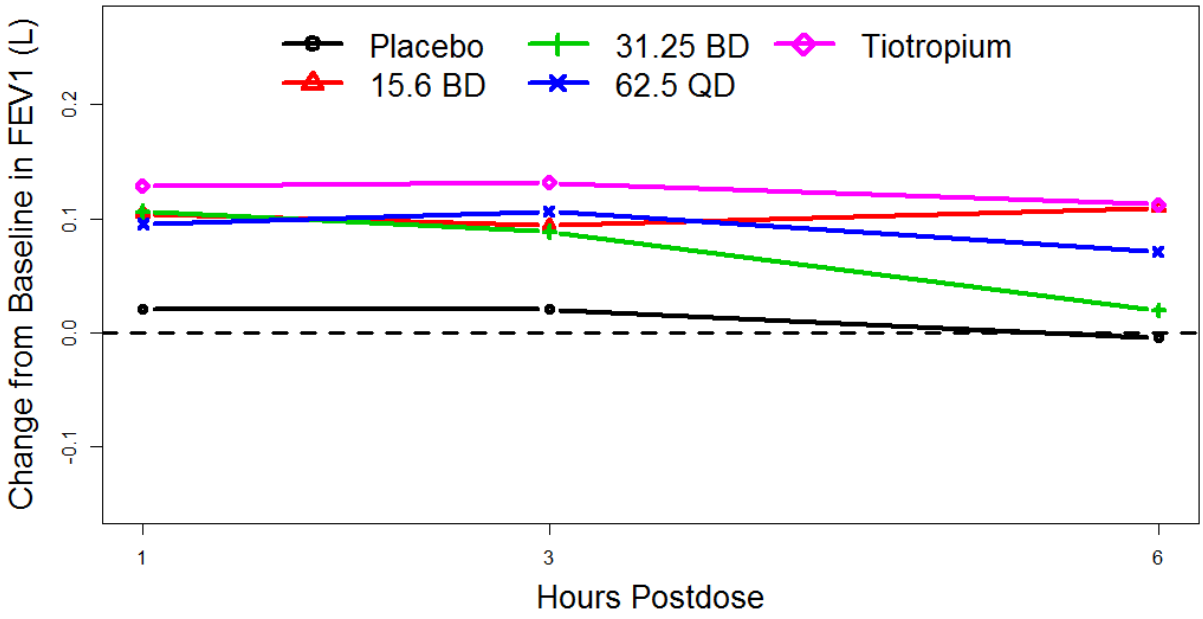
Abbreviations: QD = once-daily, BD = twice-daily

Figure 3. Mean Change from Baseline in Trough FEV_1 at Day 8 for Placebo, Different Once-Daily Umeclidinium Doses, and Tiotropium, in Study 321



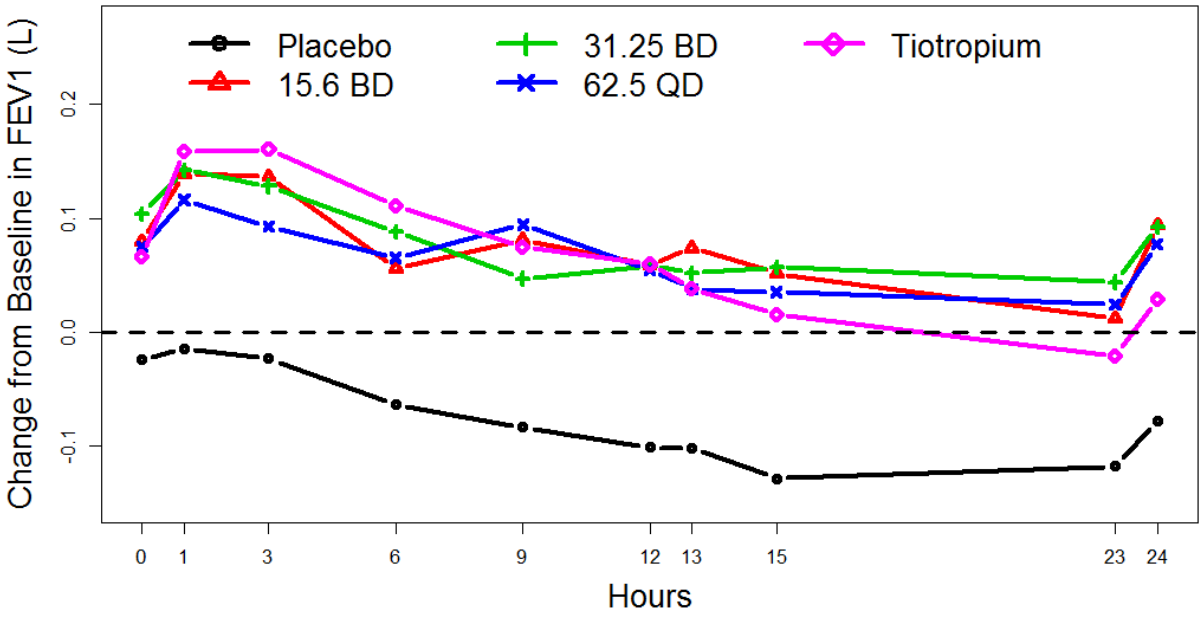
Abbreviations: QD = once-daily, BD = twice-daily

Figure 4. Postdose 6-Hour Serial Mean Change from Baseline in FEV_1 at Day 1 for Placebo, Different Once- and Twice-Daily Umeclidinium Doses, and Tiotropium, in Study 321



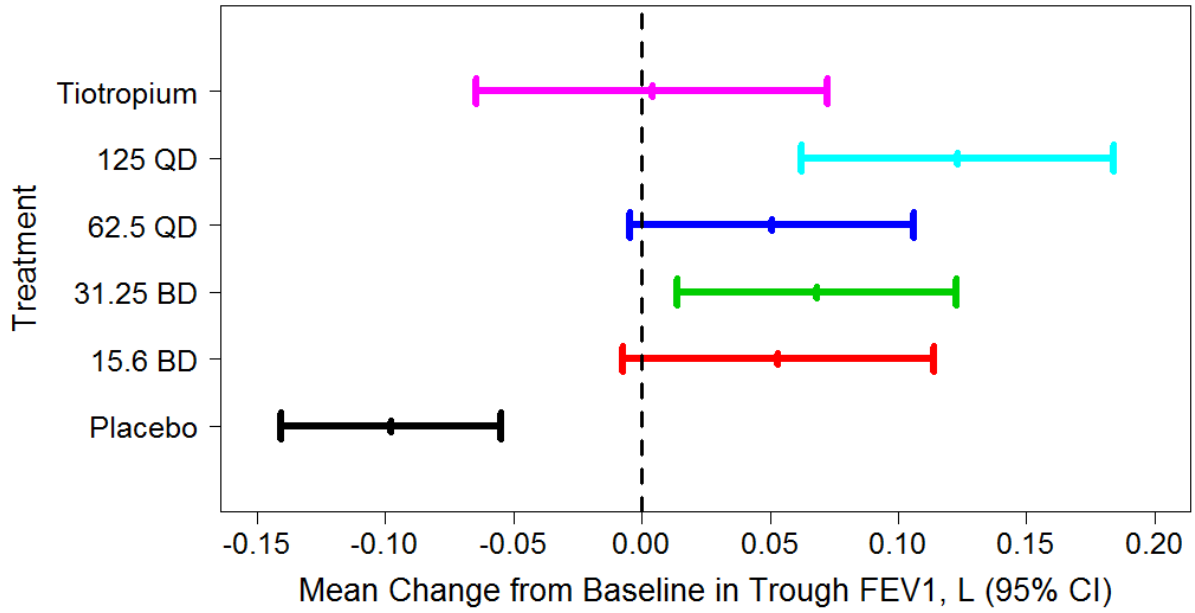
Abbreviations: QD = once-daily, BD = twice-daily

Figure 5. Postdose 24-Hour Serial Mean Change from Baseline in FEV_1 at Day 7 for Placebo, Different Once- and Twice-Daily Umeclidinium Doses, and Tiotropium, in Study 321



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

Figure 6. Mean Change from Baseline in Trough FEV_1 at Day 8 for Placebo, Different Once- and Twice-Daily Umeclidinium Doses, and Tiotropium, in Study 321



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

3.2.4 Patient Disposition, Demographic and Baseline Characteristics

Baseline characteristics were similar across Studies 361, 373, 374, and 408, which consisted of 1489, 1532, 869, and 206 patients, respectively (Appendix: Table 16, Table 17, Table 18, and Table 19). The combined population from Studies 361, 373, and 374 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_1 30–80%, 50% of subjects were current smokers, and 49% used inhaled corticosteroids. Patient characteristics in Study 408 were similar. Patients were enrolled at 153, 163, 95, and 27 different centers from several countries around the world in Studies 361, 373, 374, and 408, respectively. There were no large imbalances in baseline characteristics across the treatment arms in the four studies.

As described previously, the design of these phase 3 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, 374, and 408, 25%, 23%, 23%, and 18% of patients failed to complete the treatment period, respectively (Table 2, Table 3, Table 4, and Table 5). Dropout rates tended to be slightly higher on placebo than UMEC in Studies 361, 373, and 408, with the differences primarily attributable to greater placebo dropout because of lack of efficacy. Dropout rates, both overall and for reasons of lack of efficacy and adverse event, were slightly greater on UMEC 125 than tiotropium in Study 374. 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

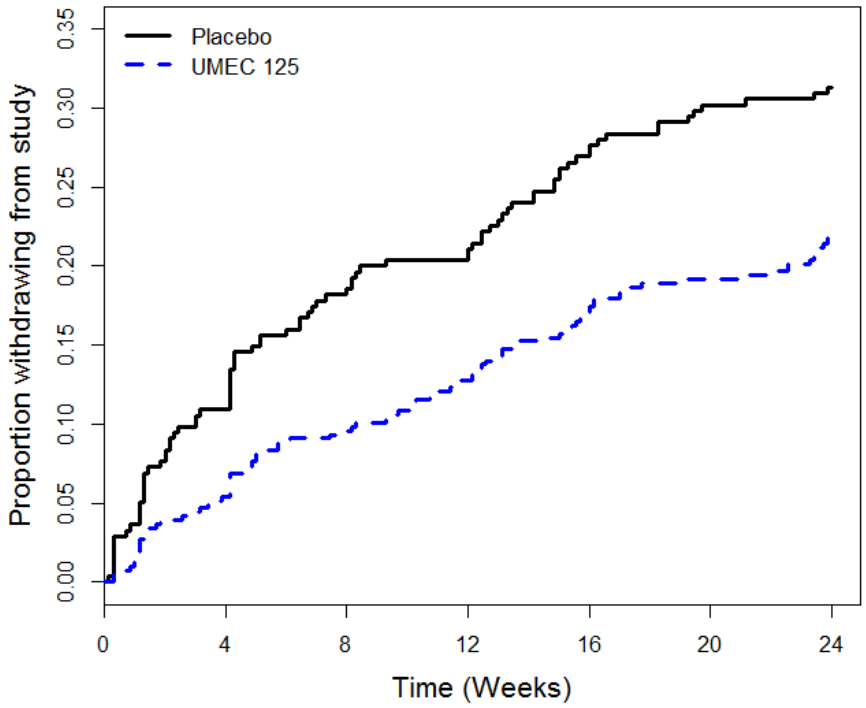


Figure 8. Proportion of Patients Withdrawing Early over Time in Study 373

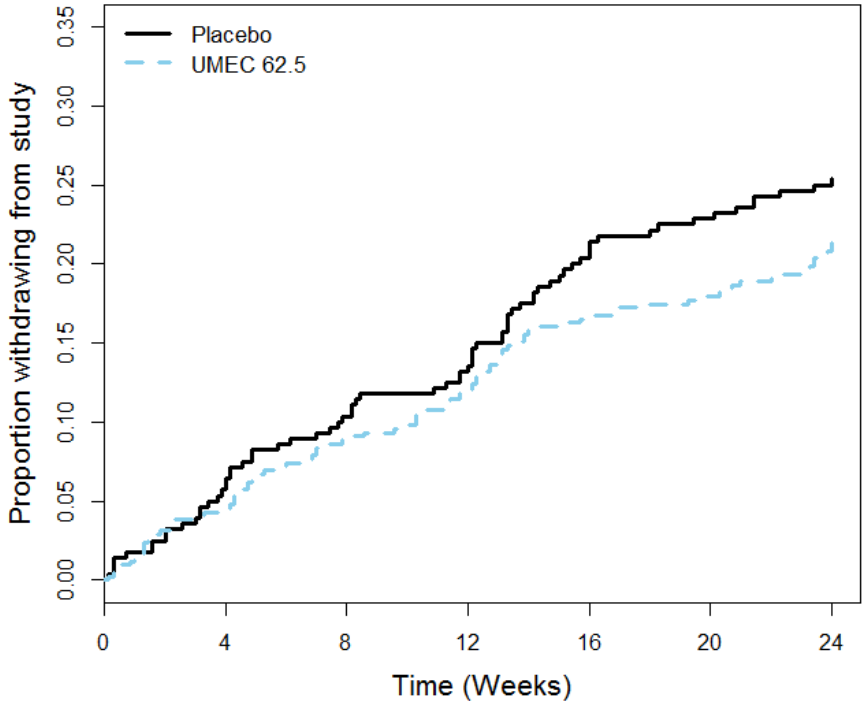


Figure 9. Proportion of Patients Withdrawing Early over Time in Study 374

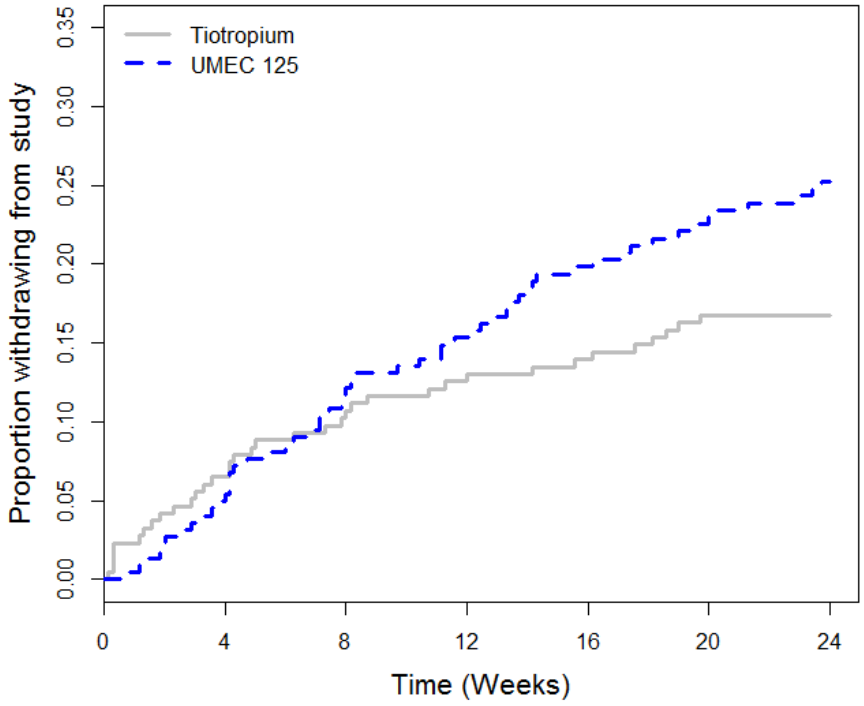


Figure 10. Proportion of Patients Withdrawing Early over Time in Study 408

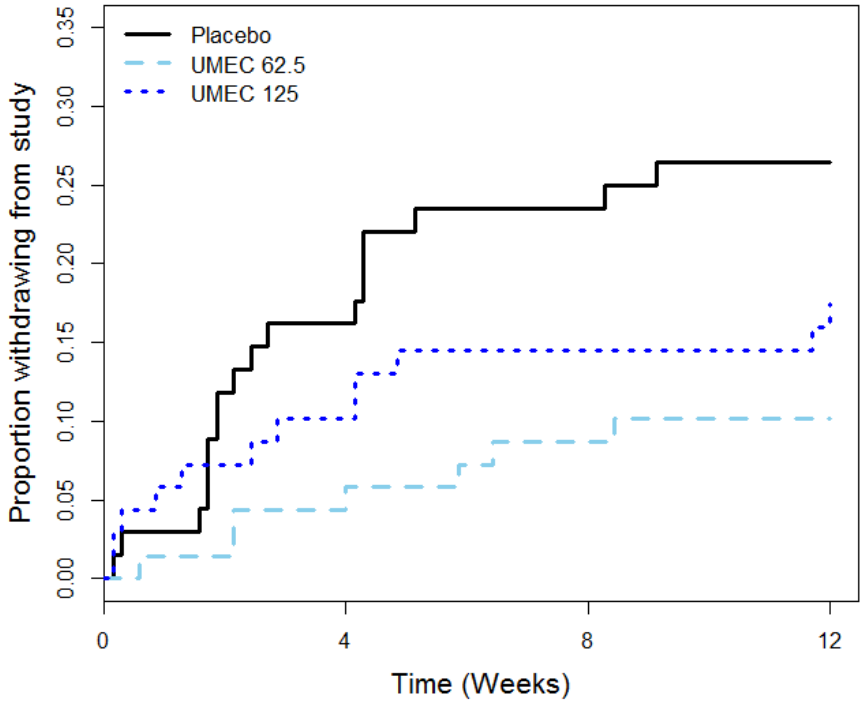


Table 2. Proportion of Patients Failing to Complete Study 361, by Reason for Withdrawal

	Placebo	UMEC 125	Overall[1]
Completed study	183 (67%)	312 (77%)	1118 (75%)
Did not complete study	92 (33%)	95 (23%)	371 (25%)
Adverse event	17 (6%)	24 (6%)	84 (6%)
Lack of efficacy	44 (16%)	38 (9%)	143 (10%)
Lost to follow-up	0 (0%)	2 (0%)	6 (0%)
Protocol deviation	4 (1%)	3 (1%)	23 (2%)
Protocol-defined stopping criteria	16 (6%)	15 (4%)	58 (4%)
Withdrew consent	11 (4%)	13 (3%)	57 (4%)

[1] Includes patients randomized to treatment with VI 25 and UMEC/VI 125/25 as well

Table 3. Proportion of Patients Failing to Complete Study 373, by Reason for Withdrawal

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

[1] Includes patients randomized to treatment with VI 25 and UMEC/VI 62.5/25 as well

Table 4. Proportion of Patients Failing to Complete Study 374, by Reason for Withdrawal

	Tiotropium	UMEC 125	Overall[1]
Completed study	176 (82%)	165 (74%)	670 (77%)
Did not complete study	39 (18%)	57 (26%)	199 (23%)
Adverse event	11 (5%)	17 (8%)	63 (7%)
Lack of efficacy	13 (6%)	22 (10%)	56 (6%)
Lost to follow-up	2 (1%)	0 (0%)	3 (0%)
Protocol deviation	1 (0%)	1 (0%)	10 (1%)
Protocol-defined stopping criteria	6 (3%)	7 (3%)	32 (4%)
Withdrew consent	6 (3%)	10 (5%)	35 (4%)

[1] Includes patients randomized to treatment with UMEC/VI 62.5/25 and UMEC/VI 125/25 as well

Table 5. Proportion of Patients Failing to Complete Study 408, by Reason for Withdrawal

	Placebo	UMEC 62.5	UMEC 125	Overall
Completed study	50 (74%)	62 (90%)	56 (81%)	168 (82%)
Did not complete study	18 (26%)	7 (10%)	13 (19%)	38 (18%)
Adverse event	0 (0%)	1 (1%)	3 (4%)	4 (2%)
Lack of efficacy	8 (12%)	5 (7%)	4 (6%)	17 (8%)
Lost to follow-up	0 (0%)	0 (0%)	1 (1%)	1 (0%)
Protocol deviation	6 (9%)	0 (0%)	5 (7%)	11 (5%)
Withdrew consent	4 (6%)	1 (1%)	0 (0%)	5 (2%)

Baseline characteristics in the cross-over Studies 417 and 418, and the long-term safety Study 359, were largely similar to those of Studies 361, 373, 374, and 408. 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 6). Dropout rates overall were similar between the placebo and UMEC 125 treatment arms. There was greater study withdrawal on placebo than UMEC for lack of efficacy and for adverse event, but greater withdrawal on UMEC because of protocol-defined stopping criteria. In particular, there was greater withdrawal on UMEC 125 (16%) than placebo (7%) because of either ECG or Holter abnormalities.

Table 6. Proportion of Patients Failing to Complete Study 359, by Reason for Withdrawal

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

[1] Includes patients randomized to treatment with UMEC/VI 125/25 as well
[2] Patients who dropped out because of protocol-defined stopping criteria could have had more than one abnormality

3.2.5 Results and Conclusions

3.2.5.1 Studies 361, 373, 374, and 408

Data are available for treatment comparisons of umeclidinium at the proposed 62.5 mcg dose against placebo from Studies 373 and 408. In each of these trials, treatment with UMEC resulted in a statistically significant, greater change from baseline in mean trough FEV_1, as compared to placebo (Table 7). In Study 373, the estimated difference in 24-week mean trough FEV_1 was 0.115 L (95% confidence interval [CI]: 0.076, 0.155; $p<0.0001$). In Study 408, the estimated difference in 12-week mean trough FEV_1 was 0.127 L (95% CI: 0.052, 0.202; $p=0.001$). There was also evidence of benefit over placebo with respect to trough FEV_1 for the higher 125 mcg dose of UMEC in Studies 361 and 408 (Table 7). All comparisons preceding the evaluation of UMEC in the sequential testing hierarchies used to account for multiplicity in Studies 361 and 373 were also statistically significant – see the statistical review of NDA 203-975 for more details.

Observed effects of UMEC on trough FEV_1 were evident as early as Day 2 and then remained relatively constant over the 24-week treatment periods in Studies 361 and 373 (Figure 11 and Figure 12), and the 12-week treatment period in Study 408 (Figure 13). There was also evidence of efficacy for UMEC with respect to the secondary endpoint 0–6 hour weighted mean FEV_1. Mean differences in weighted mean FEV_1 were slightly larger than the analogous trough FEV_1 comparisons, with strong statistical evidence against the null hypothesis of no treatment effect (Table 7). In addition, patients on UMEC demonstrated consistently higher mean FEV_1 levels than patients on placebo in the 24 hours postdose at 24 weeks in Studies 361 and 373 (Figure 14 and Figure 15), and at 12 weeks in Study 408 (Figure 16). There were similar trends in a subset of patients in Studies 361 and 373 who had more frequent 24-hour spirometry assessments. Finally, empirical distribution plots, in which dropouts were treated as the worst potential outcomes, suggested benefits of UMEC treatment with respect to summary measures of the FEV_1 distribution besides the mean, such as the median (Figure 17, Figure 18, and Figure 19). These figures

can also be used to descriptively compare treatment groups with respect to the proportion achieving certain threshold changes in FEV_1 at 24 weeks, such as improvements of at least 0.1 or 0.2 L.

Umeclidinium also showed trends toward benefit for additional non-spirometric endpoints of interest, including mean changes from baseline in SGRQ total score and average daily rescue mediation use (Table 8), and time to COPD exacerbation rate (Table 9). Estimated mean differences in SGRQ score between UMEC 62.5 and placebo in Studies 373 and 408 were -4.7 (95% CI: -7.1, -2.3) and -7.9 (95% CI: -12.2, -3.6), respectively. These estimated effects sizes are greater than the typically cited minimal clinically important difference of -4. However, there was evidence of benefit for UMEC 125 in only one of two studies, and the mean difference between UMEC 125 and placebo in Study 361 was only -0.3 (95% CI: -2.5, 1.8). Empirical distribution plots can be used to descriptively compare the proportions of patients achieving certain threshold changes in SGRQ score (Appendix: Figure 28, Figure 29, and Figure 31). In Study 373, 31% of patients on placebo, and 41% of patients on UMEC 62.5, remained in the study and had at least a 4-unit decrease in SGRQ score at Week 24 (odds ratio: 1.7; 95% CI: 1.2, 2.3).

In Study 373, 35 (13%) and 33 (8%) of patients on placebo and UMEC 62.5 suffered a COPD exacerbation, respectively, for an estimated reduction in exacerbation risk on UMEC 62.5 of 40% (95%CI: 4%, 63%). A similar trend was observed for UMEC 125 in Study 361 (Table 9). The separation between the treatment groups in the proportions suffering an exacerbation over time was also evident in Kaplan Meier plots (Appendix: Figure 32, Figure 33, and Figure 35). There were also trends toward benefit with respect to the mean 24-week change in SOBDA score for UMEC 62.5 in Study 373 (estimated difference: -0.10; 95% CI: -0.19, -0.00) and UMEC 125 in Study 361 (estimated difference: -0.08; 95% CI: -0.17, 0.02). SOBDA was not evaluated in Study 408.

The applicant did not control for multiplicity across these analyses of additional endpoints, and there may not have been substantial evidence of efficacy for each endpoint. However, the observed trends toward benefit provide support for the observed treatment effect on the primary endpoint.

In summary, there was strong statistical evidence of beneficial effects of UMEC 62.5, as compared to placebo, with respect to the primary and secondary FEV_1 endpoints, in addition to supportive trends across other spirometric and non-spirometric endpoints of interest. Findings were similar for the higher 125 mcg dose of umeclidinium.

Data are also available from Study 374 for comparisons of UMEC 125 against the active LAMA comparator tiotropium (Table 10, Figure 20, and Figure 21), although these were not prespecified analyses. There was a trend toward greater FEV_1 improvement on UMEC 125 than on tiotropium (mean difference: 0.04 L; 95% CI: -0.01, 0.09), but trends in the wrong direction with respect to SGRQ score (mean difference: 1.4; 95% CI: -1.3, 4.0; empirical distribution plot: Figure 30) and time to COPD exacerbation (hazard ratio: 1.8; 95% CI: 1.0, 3.5; Kaplan Meier plot: Figure 34).

Table 7. Comparisons of Umeclidinium against Placebo with Respect to the Primary and Secondary Endpoints in Studies 361, 373, and 408: Mean Differences in Changes from Baseline in Trough FEV_1 and 0-6 Hour Weighted Mean FEV_1

	Mean Change from Baseline in Trough FEV_1, L	Mean Difference versus Placebo in Trough FEV_1, L (95% CI) p-value	Mean Change from Baseline in 0-6 Hour Weighted Mean FEV_1, L	Mean Difference versus Placebo in 0-6 Hour Weighted Mean FEV_1, L (95% CI) p-value
Study 361: at 24 Weeks				
Placebo	-0.031		-0.018	
UMEC 125	0.129	0.160 (0.122, 0.198) <0.0001	0.16	0.178 (0.141, 0.216) <0.0001
Study 373: at 24 Weeks				
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
Study 408: at 12 Weeks				
Placebo	-0.007		-0.003	
UMEC 62.5	0.12	0.127 (0.052, 0.202) 0.001	0.163	0.166 (0.094, 0.239) <0.0001
UMEC 125	0.145	0.152 (0.076, 0.229) 0.0001	0.188	0.191 (0.117, 0.265) <0.0001

Figure 11. Mean Change from Baseline in Trough FEV_1 over Time in Study 361

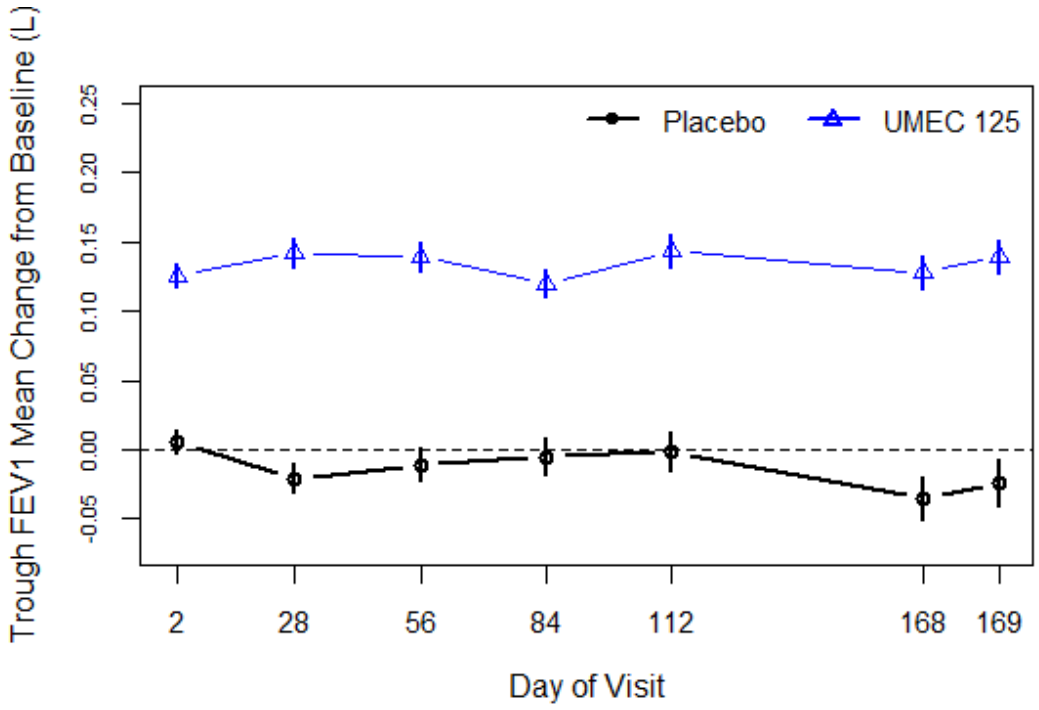


Figure 12. Mean Change from Baseline in Trough FEV_1 over Time in Study 373

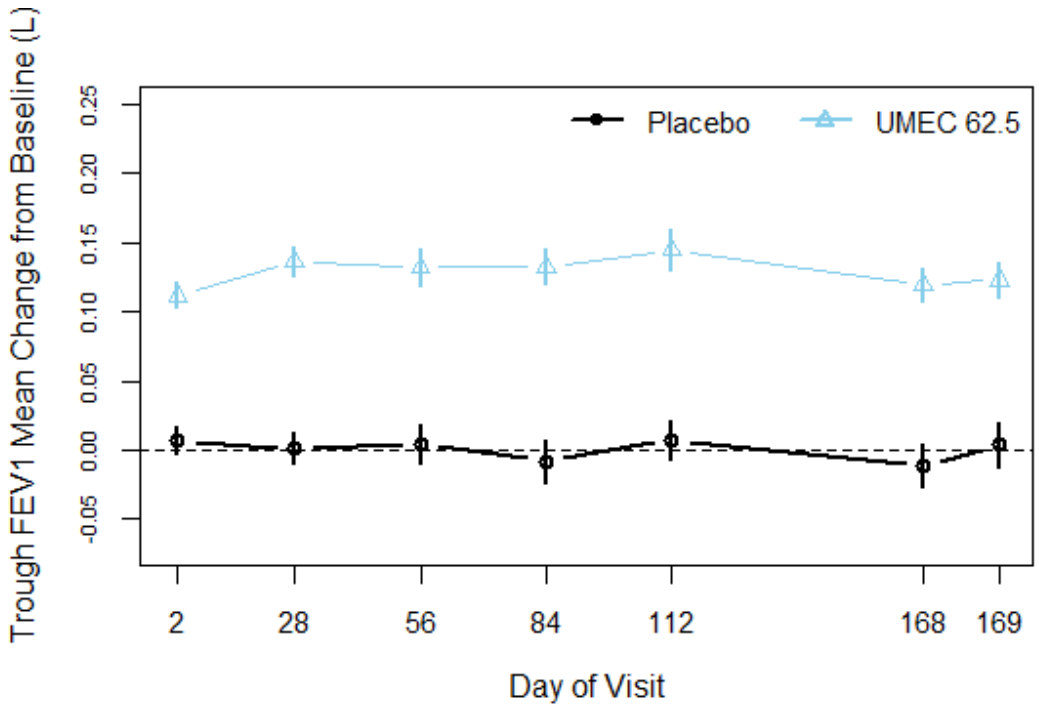


Figure 13. Mean Change from Baseline in Trough FEV_1 over Time in Study 408

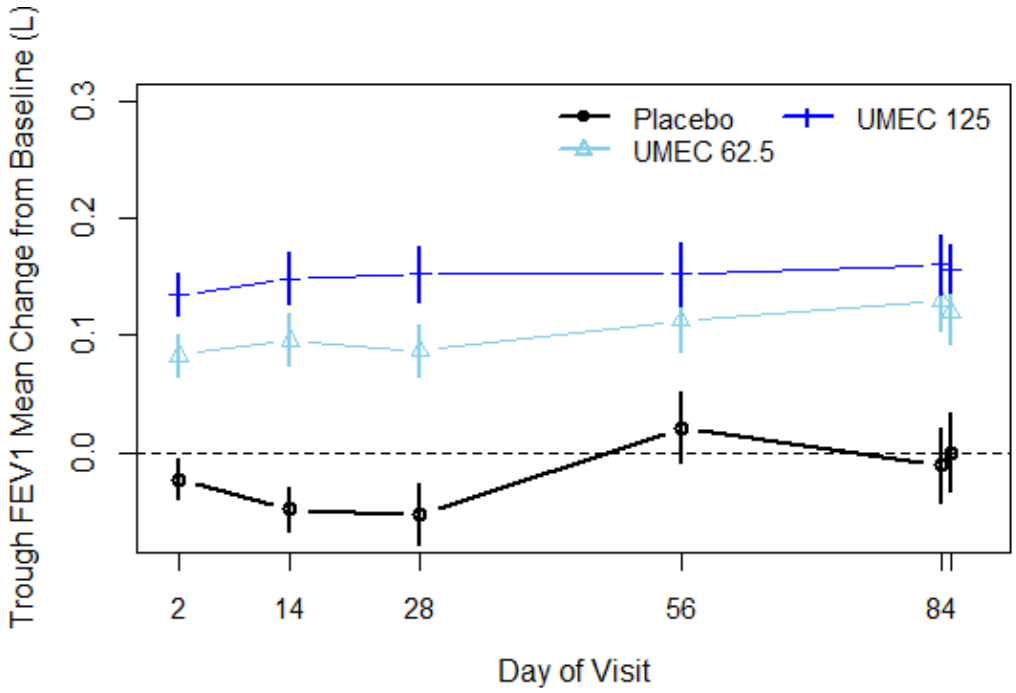


Figure 14. Postdose 24-Hour Mean Change from Baseline in FEV_1 at Day 168 in Study 361

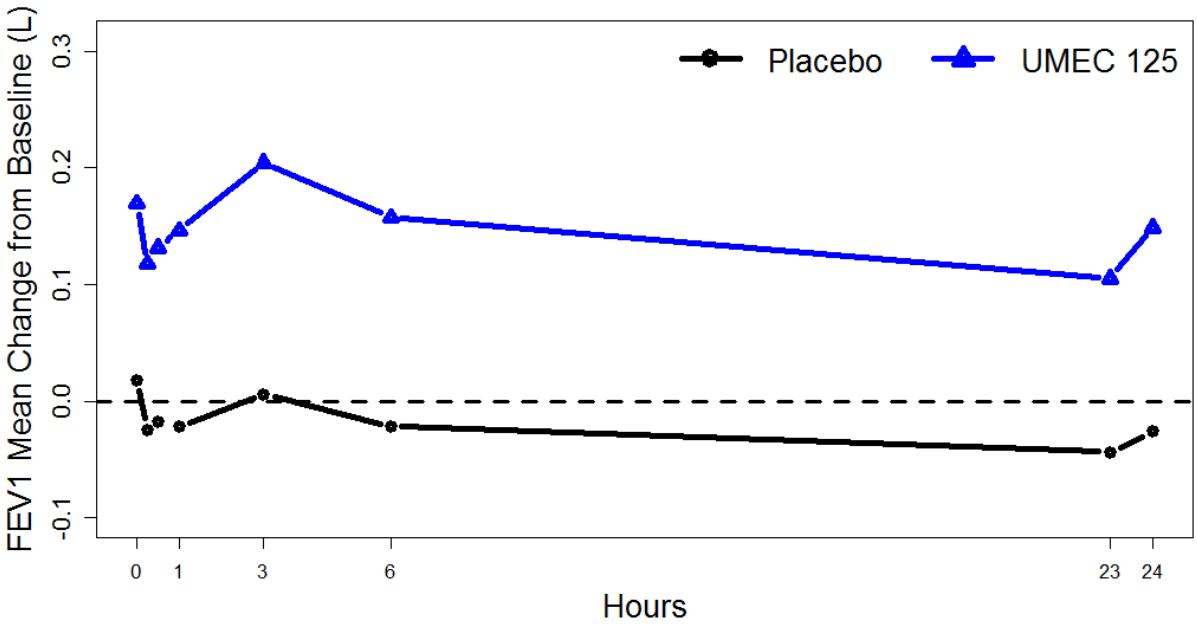


Figure 15. Postdose 24-Hour Mean Change from Baseline in FEV_1 at Day 168 in Study 373

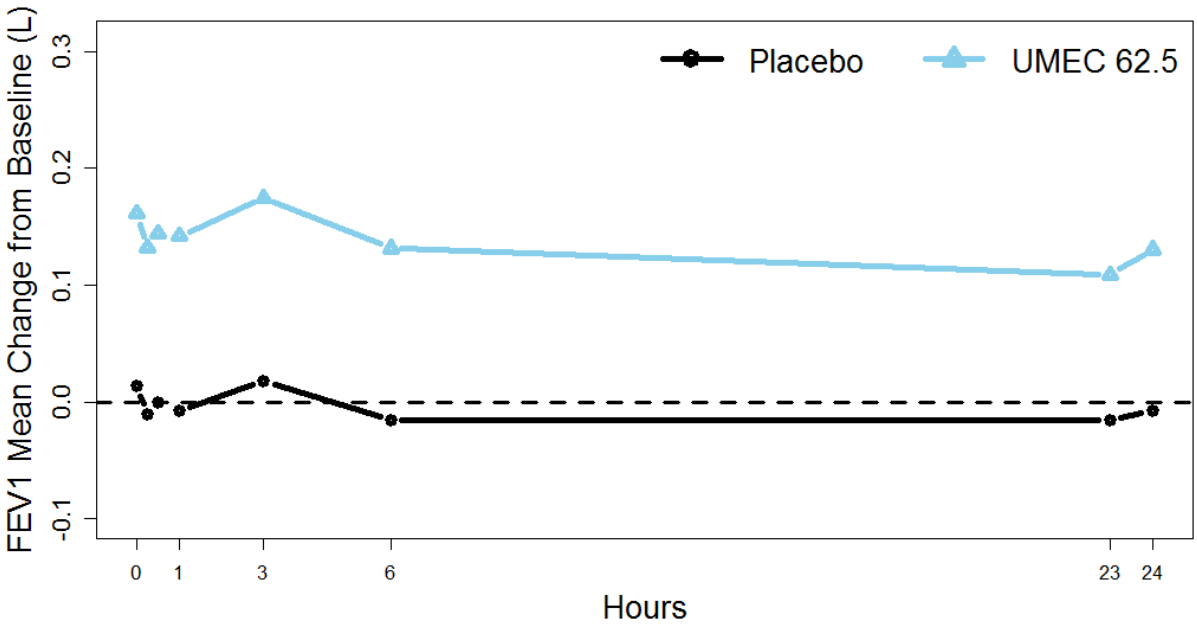


Figure 16. Postdose 24-Hour Mean Change from Baseline in FEV_1 at Day 84 in Study 408

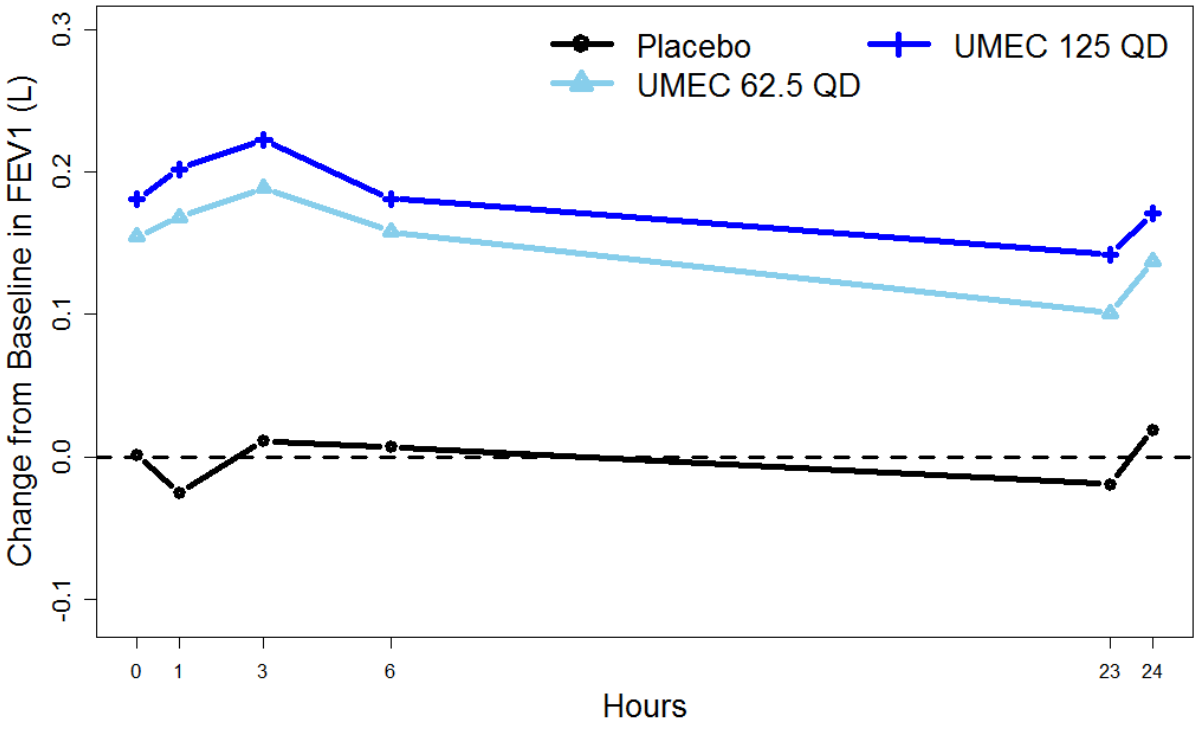


Figure 17. Empirical Distribution Function for Change from Baseline in Trough FEV_1 at 24 Weeks in Study 361

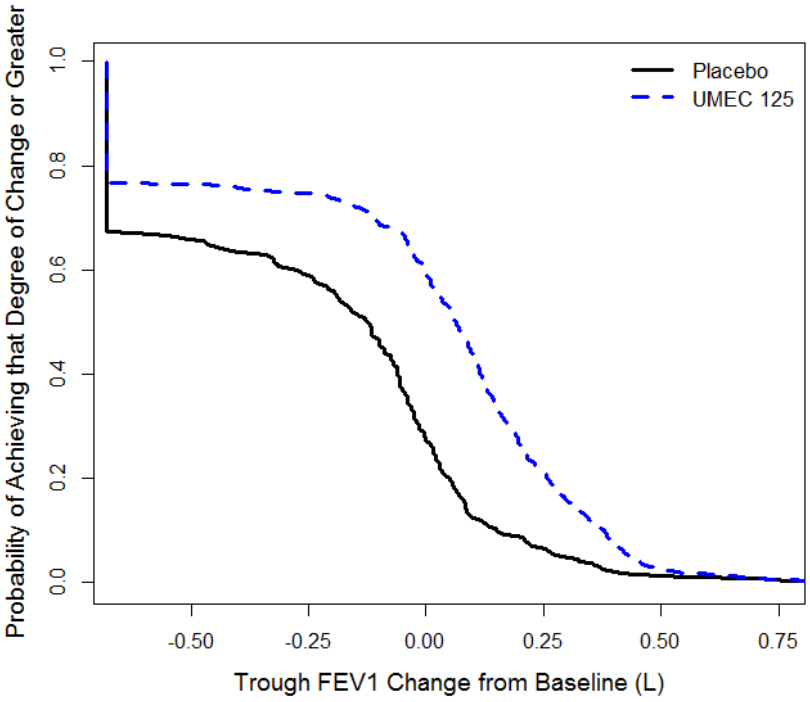


Figure 18. Empirical Distribution Function for Change from Baseline in Trough FEV_1 at 24 Weeks in Study 373

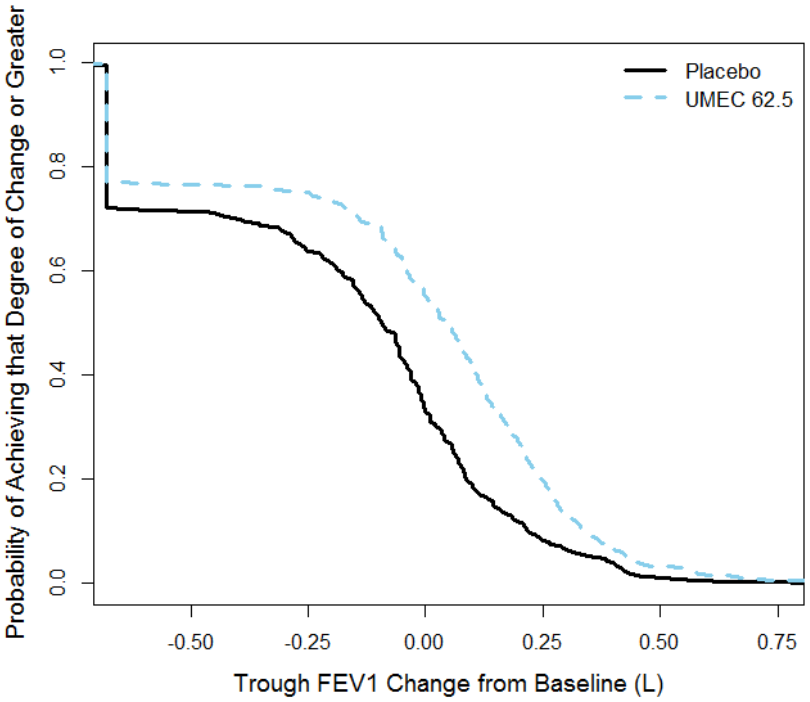


Figure 19. Empirical Distribution Function for Change from Baseline in Trough FEV_1 at 12 Weeks in Study 408

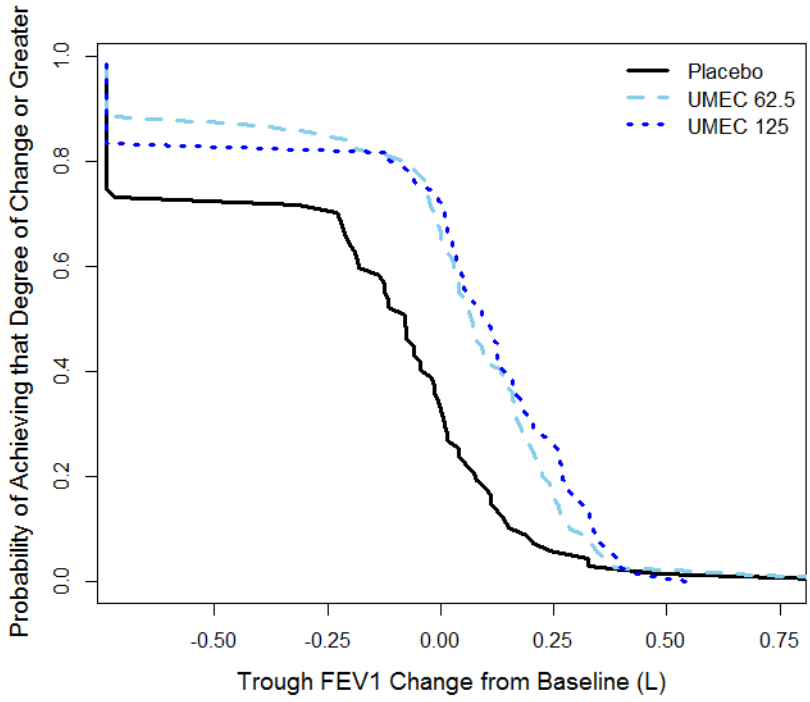


Table 8. Comparisons of Umeclidinium against Placebo with Respect to SGRQ Score and Daily Rescue Medication Use in Studies 361, 373, and 408

	Mean Change from Baseline in SGRQ Score	Mean Difference versus Placebo in SGRQ Score (95% CI) p-value	Mean Change from Baseline in Rescue Puffs per Day	Mean Difference versus Placebo in Rescue Puffs per Day (95% CI) p-value
Study 361: at 24 Weeks				
Placebo	-3.8		-0.7	
UMEC 125	-4.1	-0.3 (-2.5, 1.8) 0.78	-1.5	-0.8 (-1.3, -0.4) 0.0002
Study 373: at 24 Weeks				
Placebo	-2.6		-1.4	
UMEC 62.5	-7.3	-4.7 (-7.1, -2.3) 0.0001	-1.7	-0.3 (-0.8, 0.2) 0.28
Study 408: at 12 Weeks				
Placebo	4.8		-0.0	
UMEC 62.5	-3.1	-7.9 (-12.2, -3.6) 0.0004	-0.7	-0.7 (-1.3, -0.1) 0.03
UMEC 125	-6.1	-10.9 (-15.2, -6.5) <0.0001	-0.6	-0.6 (-1.2, 0.0) 0.07

Table 9. Comparisons of Umeclidinium against Placebo with Respect to Time to COPD Exacerbation in Studies 361 and 373

	Number (Percent) of Patients Suffering a COPD Exacerbation	Exacerbation Hazard Ratio versus Placebo (95% CI) p-value
Study 361		
Placebo	38 (14)	
UMEC 125	32 (8)	0.50 (0.31, 0.80) 0.0040
Study 373		
Placebo	35 (13)	
UMEC 62.5	33 (8)	0.60 (0.37, 0.96) 0.035

Table 10. Comparisons of Umeclidinium against Tiotropium with Respect to Trough FEV_1, SGRQ Score, and Time to COPD Exacerbation in Study 374

	Mean Change from Baseline in Trough FEV_1, L	Mean Difference vs Tiotropium in Trough FEV_1, L (95% CI)	Mean Change from Baseline in SGRQ Score	Mean Difference vs Tiotropium in SGRQ (95% CI)	Number (Percent) of Patients Suffering an Exacerbation	Exacerbation Hazard Ratio vs Tiotropium (95% CI)
Tiotropium	0.15		-9.8		14 (7)	
UMEC 125	0.19	0.04 (-0.01, 0.09)	-8.4	1.4 (-1.3, 4.0)	26 (12)	1.8 (1.0, 3.5)

Figure 20. Mean Change from Baseline in Trough FEV_1 over Time in Study 374

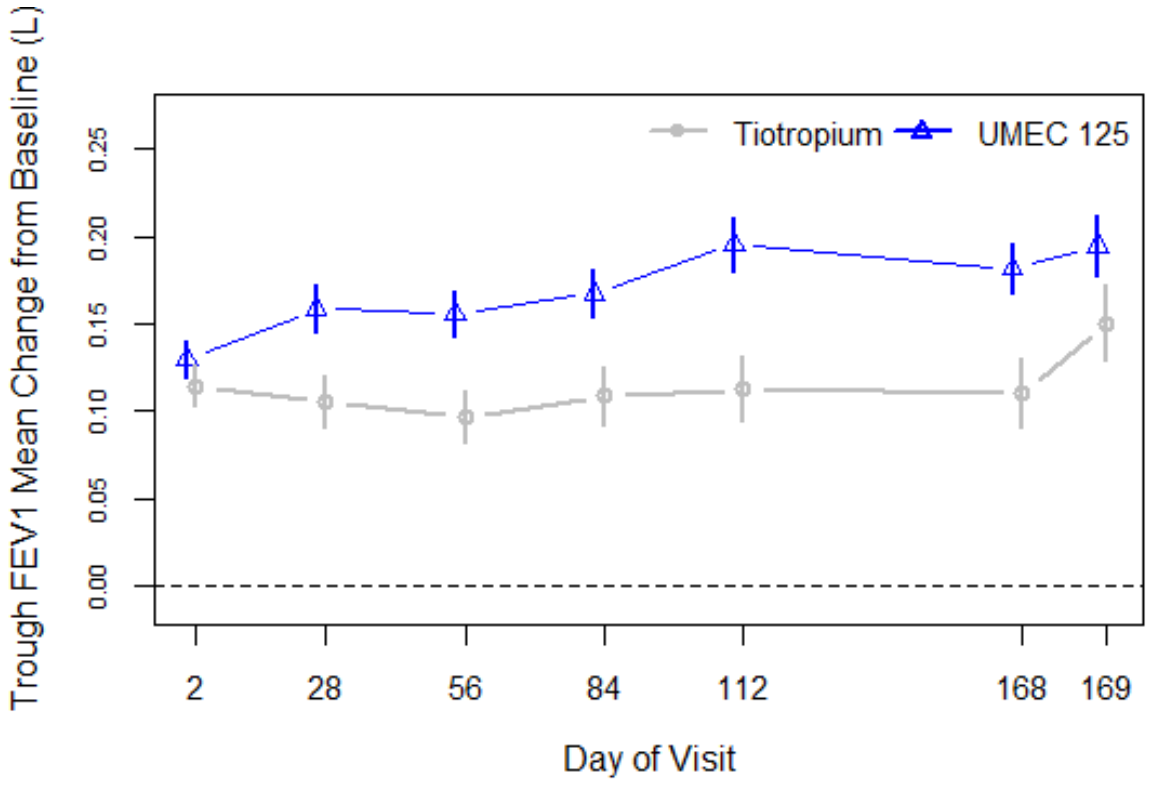
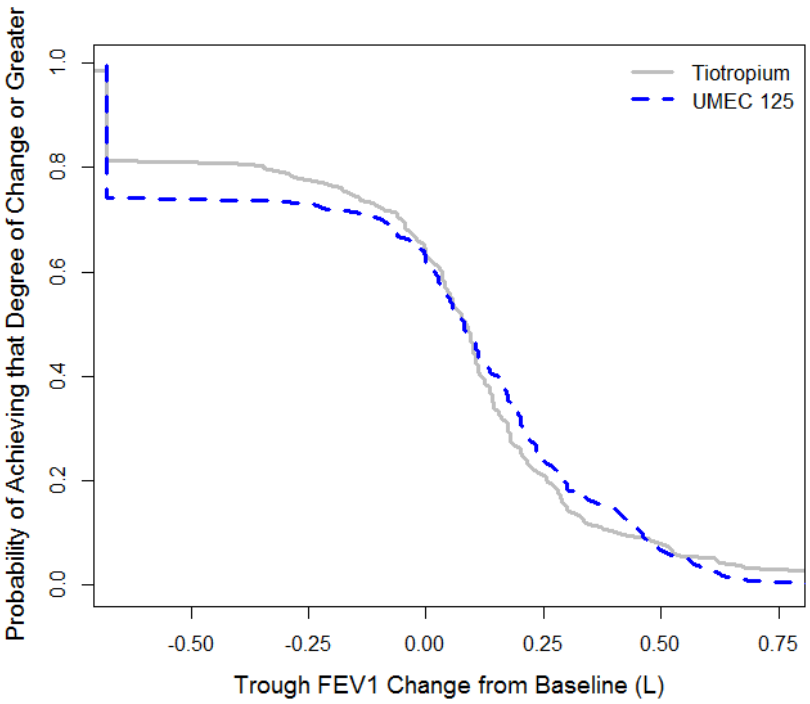


Figure 21. Empirical Distribution Function for Change from Baseline in Trough FEV_1 at 24 Weeks in Study 374



3.2.5.2 Potential Effect of Missing Data

As described in detail in 3.2.4, there were substantial missing data in Studies 361, 373, 374, and 408. Overall dropout rates ranged from 18% to 25%. 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 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 treatment period in Studies 361, 373, and 374 (Table 11). For example, 15% of dropouts had GOLD Stage IV COPD at baseline, as compared to 8% of completers. Demographic characteristics were largely similar between dropouts and completers. These trends were also evident when each study was evaluated separately, and similar patterns were observed in Study 408.

We also examined trends in trough FEV_1 before dropout within each treatment arm. Figure 22 displays average pulmonary function over time by dropout pattern, i.e., by the final visit at which FEV_1 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 UMEC 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 who withdrew from the study early went on to have substantially worse lung function at the end of the study 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 UMEC. However, these patterns also highlight important deficiencies in the primary MMRM analysis, as well as the majority of the sensitivity analyses proposed by the applicant.

If the estimand of interest is the hypothetical effectiveness of UMEC ***if all patients could tolerate and adhere to the therapy***, then the estimated treatment effect from the MMRM 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 (i.e., the difference in the mean change from baseline in trough FEV_1 at 12 or 24 weeks), ***at real world achievable adherence and tolerability***, the MMRM analysis likely does not produce a reliable estimate of the truth. The MMRM analysis, 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 (through 12 or 24 weeks), even after they stopped taking the therapy. Because bronchodilators are generally considered symptomatic and not disease-modifying therapies, and their effects on FEV_1 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 UMEC would have had outcomes similar to those that were observed among completers (with similar baseline characteristics) ***in the placebo group***. Under the Jump to Reference approach, statistical significance was maintained for all comparisons of UMEC against placebo in Studies 361, 373, and 408. However, estimated magnitudes of treatment effect were approximately 20–30% smaller than those based on the primary MMRM analysis (Table 12). For example, in Study 373, the estimated mean improvement in FEV_1 on UMEC 62.5, relative to placebo, was 0.09 L (95% CI: 0.05, 0.13), as compared to 0.12 L (95% CI: 0.08, 0.16) in the primary analysis. There were also 20–30% smaller estimates of the magnitude of the treatment effect on SGRQ score in missing data supportive analyses (Table 13).

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 UMEC could have experienced ***worse*** outcomes after discontinuation than dropouts on control. That being said, the observed trend toward greater FEV_1 on UMEC than placebo before dropout (Figure 22) somewhat mitigates this concern, at least with respect to pulmonary function. There remains the possibility that dropouts from UMEC could have gone on to experience ***worse*** outcomes with respect to important safety endpoints (see 3.3).

Table 11. Baseline Characteristics, Stratified According to Whether Patients Completed the Study, Based on Integrated Data from the Placebo and Umeclidinium treatment arms in Studies 361, 373, and 374

	Completer[1] (N=1,182)	Dropout (N=408)	Overall (N=1,602)
Female	374 (32%)	135 (33%)	516 (32%)
Age	62.9 (8.6)	64.1 (9.2)	63.2 (8.8)
Race			
White	1004 (85%)	345 (85%)	1361 (85%)
Black	28 (2%)	14 (3%)	42 (3%)
Asian	119 (10%)	42 (10%)	161 (10%)
Hispanic/Latino	80 (7%)	23 (6%)	105 (7%)
BMI (kg/m^2)	26.7 (5.9)	26.0 (5.4)	26.5 (5.8)
Current Smoker	618 (52%)	189 (46%)	814 (51%)
FEV_1 (L)	1.3 (0.5)	1.1 (0.5)	1.2 (0.5)
GOLD Stage ($ppFEV_1$)			
Stage II (50-80%)	557 (47%)	150 (37%)	711 (44%)
Stage III (30-50%)	524 (44%)	195 (48%)	723 (45%)
Stage IV (<30%)	97 (8%)	63 (15%)	164 (10%)
Chronic Bronchitis	773 (65%)	259 (63%)	1041 (65%)
Emphysema	719 (61%)	272 (67%)	997 (62%)
Duration of COPD, years			
<1	102 (9%)	24 (6%)	127 (8%)
1,5	455 (38%)	134 (33%)	594 (37%)
5,10	354 (30%)	133 (33%)	492 (31%)
10,15	169 (14%)	79 (19%)	249 (16%)
15-20	51 (4%)	19 (5%)	70 (4%)
20-25	28 (2%)	8 (2%)	36 (2%)
>25	23 (2%)	11 (3%)	34 (2%)
Inhaled Corticosteroid Use	589 (50%)	215 (53%)	811 (51%)
Reversible to Salbutamol	376 (32%)	116 (28%)	496 (31%)
Reversible to Salbutamol and Ipratropium	648 (55%)	219 (54%)	873 (54%)
At USA site	284 (24%)	111 (27%)	398 (25%)

Figure 22. Mean Change from Baseline in Trough FEV_1 by Treatment Group over Time, Stratified by Dropout Pattern, Based on Integrated Data from Studies 361 and 373

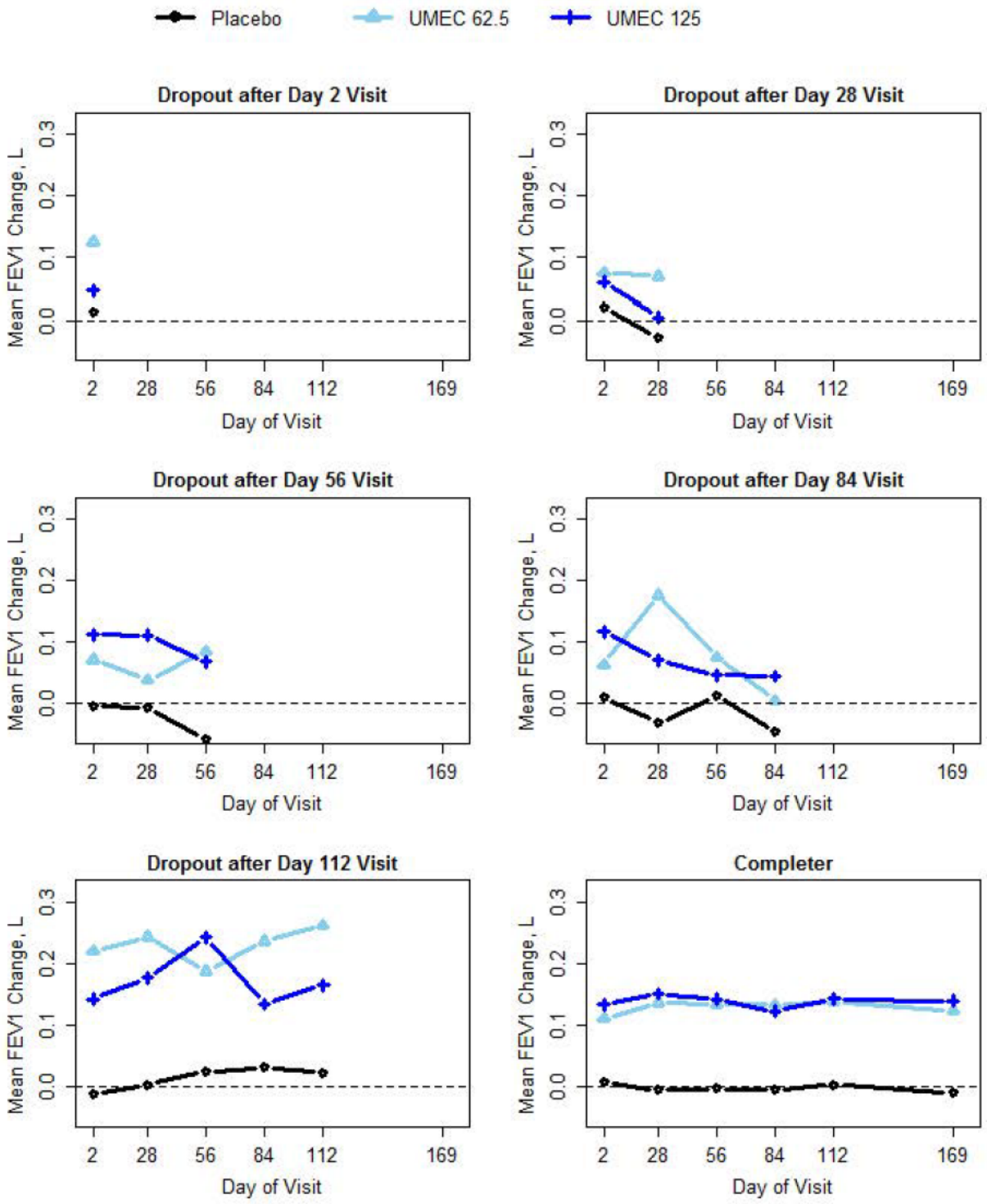


Table 12. Exploring the Potential Effect of Missing Data: Results for the Primary Endpoint Trough FEV_1 with the Primary Mixed Effects Analysis as Compared to a Multiple Imputation Supportive Analysis in Studies 361, 373, and 408

Study	Comparison	Mean Difference in Trough FEV_1 Change from Baseline[1], L (95% CI)	
		Primary	Supportive[2]
373	UMEC 62.5 vs. Placebo	0.12 (0.08, 0.16)	0.09 (0.05, 0.13)
408	UMEC 62.5 vs. Placebo	0.13 (0.05, 0.20)	0.11 (0.04, 0.19)
408	UMEC 125 vs. Placebo	0.15 (0.08, 0.23)	0.13 (0.05, 0.20)
361	UMEC 125 vs. Placebo	0.16 (0.12, 0.20)	0.12 (0.08, 0.16)

[1] Comparison is at 24 weeks in Studies 361 and 373, and at 12 weeks in Study 408
[2] Based on Jump to Reference Multiple Imputation Approach

Table 13. Exploring the Potential Effect of Missing Data: Results for SGRQ Score with the Prespecified Mixed Effects Analysis as Compared to a Multiple Imputation Supportive Analysis in Studies 361, 373, and 408

Study	Comparison	Mean Difference in SGRQ Score Change from Baseline[1] (95% CI)	
		Primary	Supportive[2]
373	UMEC 62.5 vs. Placebo	-4.7 (-7.1, 2.3)	-3.7 (-6.1, -1.4)
408	UMEC 62.5 vs. Placebo	-7.9 (-12.2, -3.6)	-7.2 (-11.5, -2.9)
408	UMEC 125 vs. Placebo	-10.9 (-15.2, 6.5)	-9.5 (-13.9, -5.1)
361	UMEC 125 vs. Placebo	-0.3 (-2.5, 1.9)	-0.3 (-2.4, 1.9)

[1] Comparison is at 24 weeks in Studies 361 and 373, and at 12 weeks in Study 408
[2] Based on Jump to Reference Multiple Imputation Approach

3.2.5.3 Additional Phase 3 Studies

In the cross-over Studies 417 and 418, comparisons between UMEC and placebo were among many secondary analyses, for which there was no multiplicity control. Study 417 also failed to achieve statistical significance in one of the two co-primary analyses (the comparison between UMEC/VI and placebo with respect to EET). Therefore, evaluations of effects of the UMEC monotherapy in these studies are considered exploratory. In addition, Studies 417 and 418 were primarily designed to compare UMEC/VI to placebo, so fewer patients were randomized to periods of treatment with the monotherapies than with the combination product. Nevertheless, patients treated with UMEC showed consistent trends

toward greater mean improvement in trough FEV_1 at 12 weeks, with effect sizes similar to those observed in the other phase 3 studies (Table 14). There were also some trends toward greater improvement in EET, although confidence intervals for the estimated effect sizes are wide. Note that the interpretation of results from these cross-over studies is clouded by the substantial patient dropout.

The 52-week Study 359 only included the 125 mcg dose of UMEC, and was designed to evaluate safety and tolerability, so no primary efficacy analyses were prespecified. However, exploratory efficacy results were generally supportive of findings in the phase 3 efficacy studies. Treatment with UMEC resulted in 0.16 L (95% CI: 0.08, 0.24) and 0.18 L (95% CI: 0.10, 0.26) greater mean trough FEV_1 changes at 6 and 12 months, respectively, as compared to placebo. In addition, there were trends toward a lower rate of first COPD exacerbation (hazard ratio: 0.6; 95% CI: 0.3, 1.0) and lesser daily rescue medication use (difference in mean puffs per day: -0.4; 95% CI: -0.9, 0.1) on UMEC than placebo.

Table 14. Comparisons of Umeclidinium against Placebo with Respect to Mean 12-Week Changes from Baseline in Exercise Endurance Time (EET) and Trough FEV_1 in the Cross-Over Studies 417 and 418

	Mean Difference in EET, s (95% CI)	Mean Difference in Trough FEV_1, L (95% CI)
Study 417		
UMEC 62.5	26.5 (-25.9, 78.9)	0.09 (0.03, 0.14)
UMEC 125	13.1 (-38.9, 65.1)	0.14 (0.08, 0.20)
Study 418		
UMEC 62.5	25.0 (-41.0, 91.1)	0.14 (0.09, 0.20)
UMEC 125	74.7 (6.0, 143.4)	0.26 (0.19, 0.32)

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. We also conducted additional analyses to further explore cardiovascular risk. The applicant prespecified a number of adverse events (AEs) of special interest based on potential pharmacologic class effects of LAMAs. 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 because of the small numbers of events within groups and the lack of a consistent dose-response. We report findings based on data from all phase 3 studies. 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 phase 3 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 (±2 days) following a patient's final 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; (3) our analyses 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; and (4) our analyses include results from the tiotropium-controlled Study 360, which did not include a UMEC monotherapy treatment arm but was part of the same UMEC/VI development program and was carried out in the same patient population as Studies 361, 373, and 374 (whereas the applicant's analyses do not include data from this study).

Incidence rates of MACE were largely similar across the treatment arms, with a slightly higher rate on placebo than UMEC and tiotropium (Table 15). There also was no evidence of a safety signal for MACE based on comparisons of the proportions of patients with events over time (Figure 23), nor was there evidence in regression analyses (hazard ratio for UMEC versus placebo: 0.8; 95% CI: 0.4, 1.4). 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 (hazard ratio for UMEC versus placebo: 0.8; 95% CI: 0.3, 2.2; incidence rates in Table 15; Kaplan Meier estimates in Figure 24).

Despite the lack of evidence for MACE, there was the suggestion of a possible trend toward greater cardiovascular risk on UMEC, as compared to both placebo and tiotropium, when evaluating cardiovascular-related serious adverse events. This imbalance was evident when examining incidence rates (Table 15) and proportions with events over times (Figure 25), as well as in regression analyses (hazard ratio versus placebo: 1.3, 95% CI: 0.6, 3.0; hazard ratio versus tiotropium: 2.8, 95% CI: 0.6, 11.9).

The most striking feature of these analyses of cardiovascular risk is the low overall numbers of events in the integrated phase 3 studies, which leads to considerable statistical uncertainty around estimated differences in risks between treatment arms. As an example, the 95% confidence interval for the hazard ratio for cardiovascular-related serious adverse events indicates that true differences ranging from a ***40% decreased risk*** on UMEC (as compared to placebo) to a ***3-fold increased risk*** on UMEC cannot be ruled out based on the observed data.

It is also important to note that missing data clouds the interpretability of safety analyses. It is reassuring that dropout rates because of adverse events on UMEC tended to be similar to the rate on placebo (Table 2, Table 3, Table 5, and Table 6). However, because patients were not followed up after treatment discontinuation for a safety evaluation through the complete double-blind period in each trial, we cannot rule out the possibility that: (1) differences in patient characteristics between dropouts on placebo and UMEC induce bias in safety comparisons; or (2) UMEC has residual effects that increase risk of adverse

events after patients stop taking the therapy. In particular, it is concerning that in the long-term, placebo-controlled safety Study 359, the only trial to include Holter monitoring in all randomized subjects, there was greater dropout on UMEC 125 (16%) than placebo (7%) because of ECG and/or Holter abnormalities. This trial contributed much of the statistical information to the analyses that evaluated cardiovascular risk.

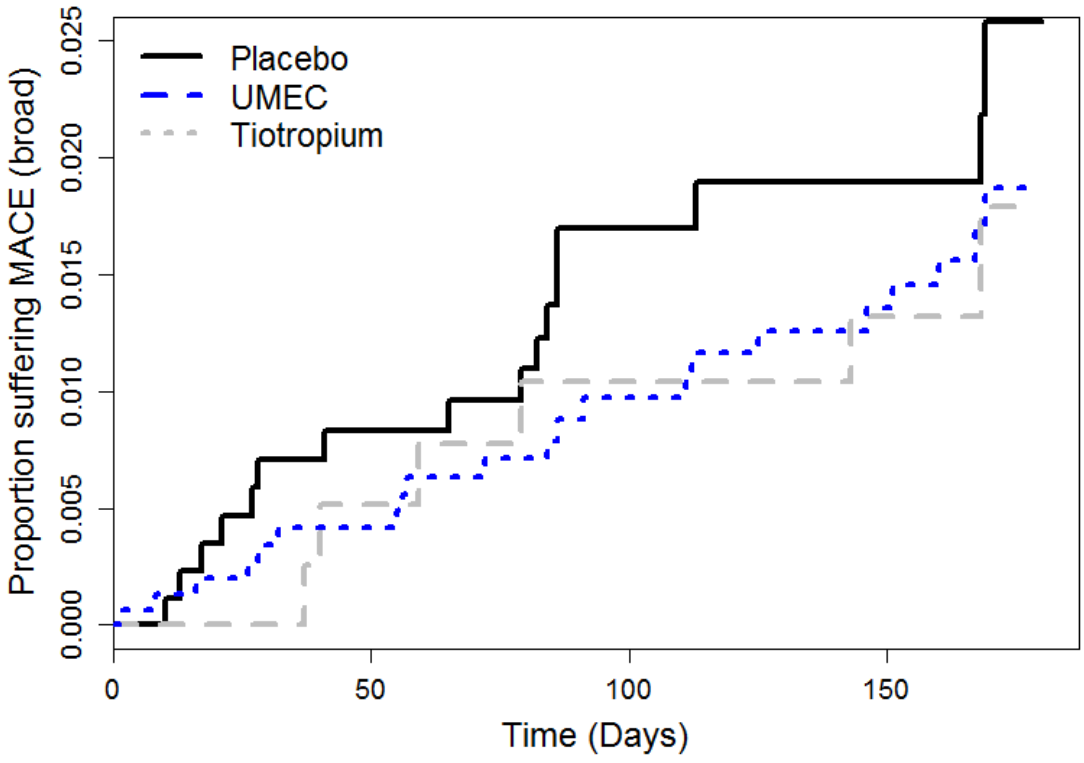
Table 15. Numbers of MACE and Adjudicated Cardiovascular Serious Adverse Events, and Unadjusted Pooled Incidence Rates, by Treatment Group, in all Phase 3 Studies[1]

Endpoint	Placebo (N=910)	UMEC[2] (N=1,512)	Tiotropium (N=423)
MACE (broad)[3]	17 (51)	23 (37)	6 (35)
MACE (narrow)[4]	6 (18)	9 (14)	1 (6)
Adjudicated Cardiovascular SAE	8 (24)	20 (32)	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

[1] All Phase 3 Studies = Studies 361, 373, 360, 374, 359, 408, and the first treatment periods of Studies 417 and 418.
[2] Combines the UMEC 62.5 and 125 mcg treatment groups
[3] MACE (broad) includes adjudicated cardiovascular death, non-fatal stroke adverse events of special interest, and non-fatal cardiac ischaemia adverse events of special interest
[4] 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 23. Proportion Suffering MACE (broad definition) over Time by Treatment Group Based on Data from all Phase 3 Studies



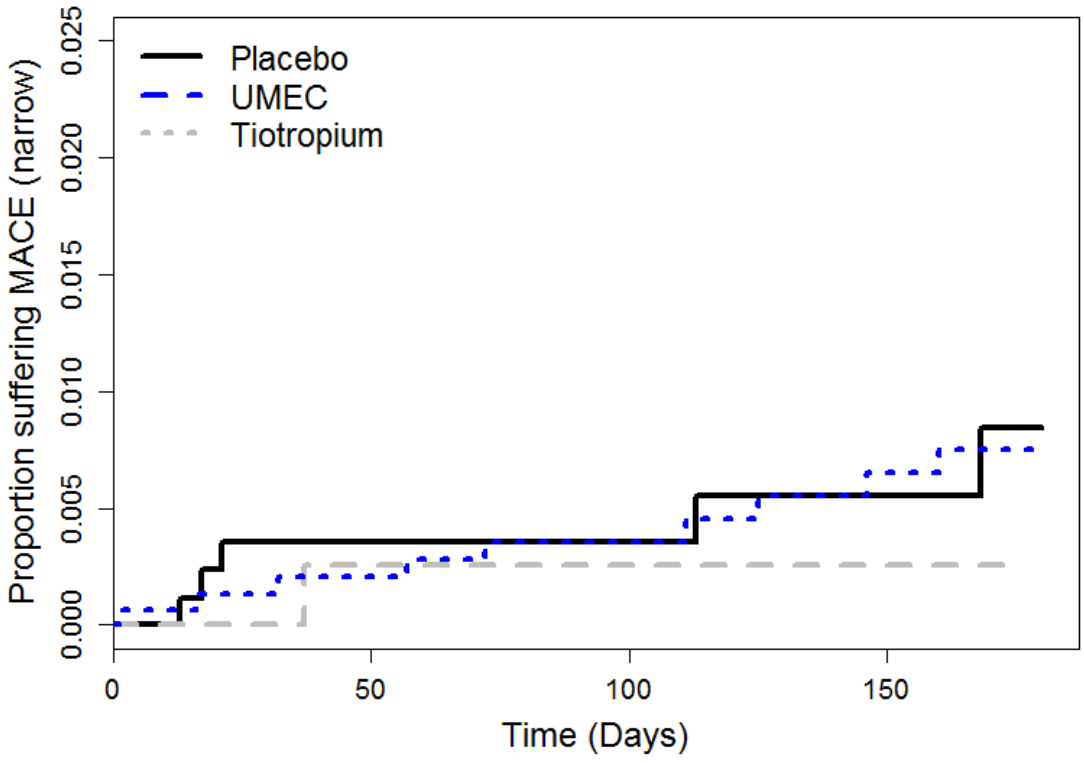
Number at risk

Placebo	910	780	512	479
Tiotropium	423	385	366	355
UMEC	1512	1357	1045	996

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 MACE (narrow definition) over Time by Treatment Group Based on Data from all Phase 3 Studies



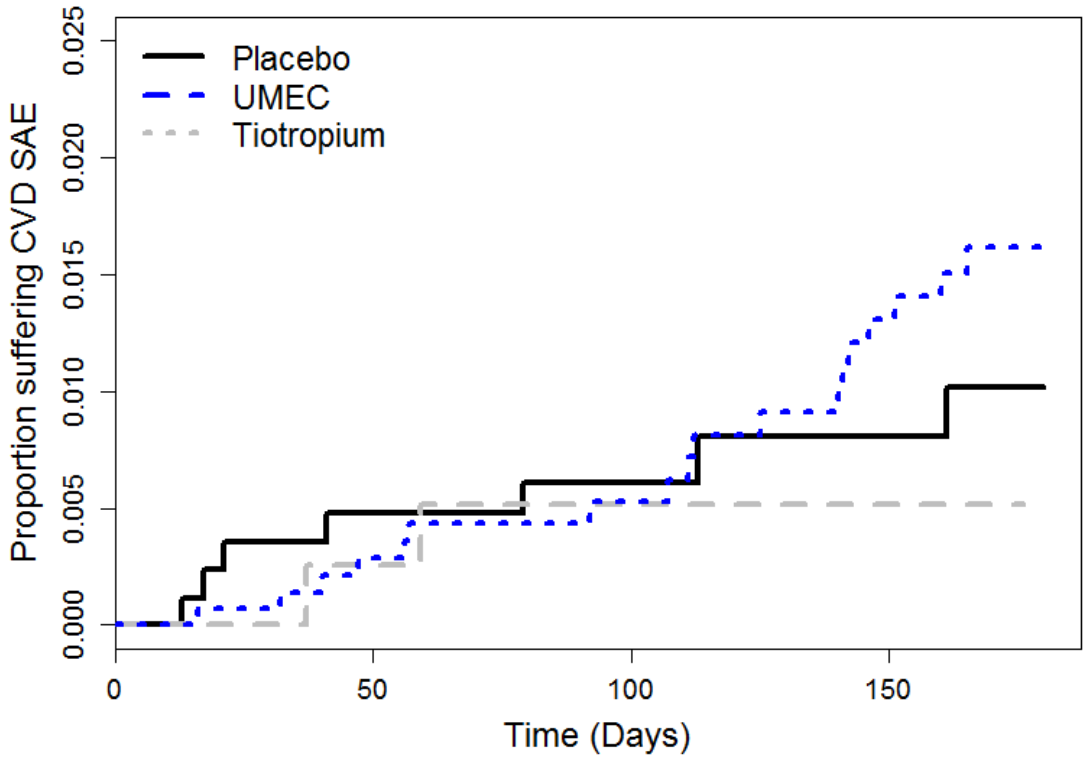
Number at risk

Placebo	910	784	517	483
Tiotropium	423	386	368	357
UMEC	1512	1359	1049	1001

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



Number at risk

Placebo	910	783	517	483
Tiotropium	423	386	368	357
UMEC	1512	1358	1048	996

Abbreviations: CVD SAE = cardiovascular serious adverse event

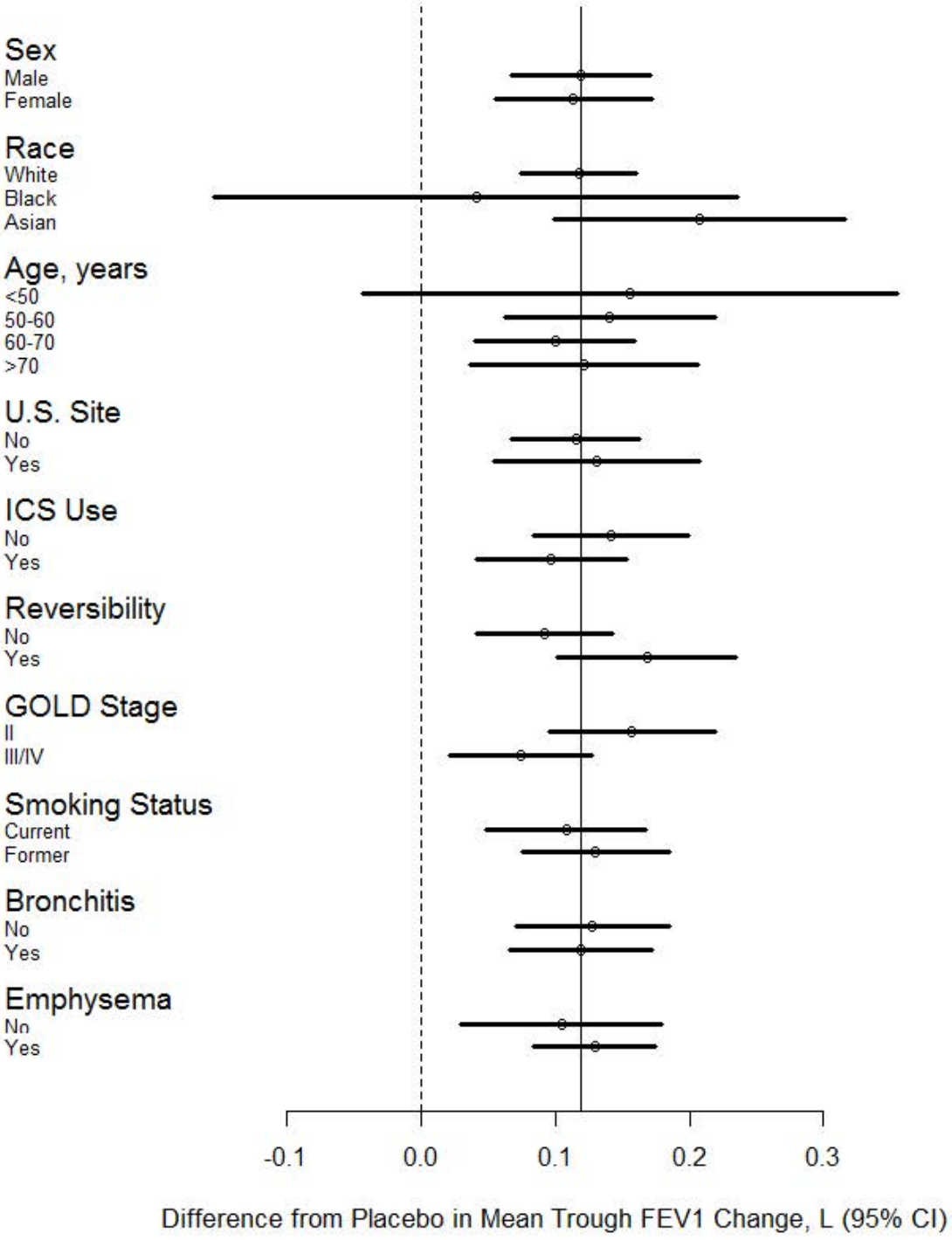
All Phase 3 Studies = Studies 361, 373, 360, 374, 359, 408, and the first treatment periods of Studies 417 and 418.

4 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

Figure 26 and Figure 27 present the results of subgroup analyses by a number of demographic and baseline characteristics based on data from Studies 373 and 408, respectively. We conducted subgroup analyses by sex, race (White, Black, or Asian), 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_1 at least 12% and 200 mL greater than pre-salbutamol FEV_1), COPD GOLD stage, smoking status, chronic bronchitis, and emphysema.

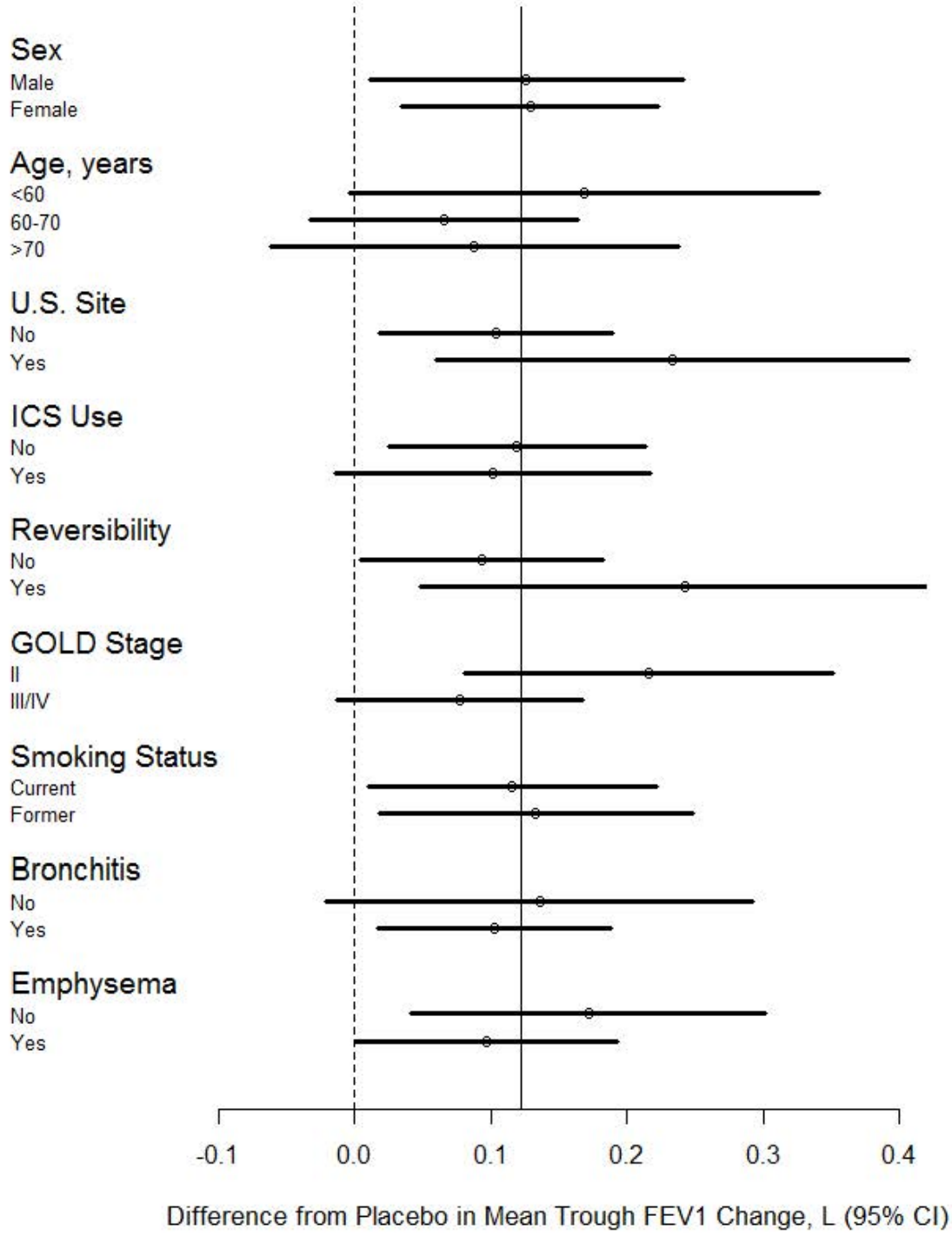
Estimated differences in mean trough FEV_1 comparing UMEC 62.5 with placebo were largely consistent across these subgroups. There was a trend toward a larger observed treatment effect within the subset of patients demonstrating reversibility to salbutamol at baseline. Importantly, the estimated treatment effect in patients who did not demonstrate reversibility, although smaller in magnitude, was still consistently in the direction of benefit (and was actually statistically significantly greater than zero in both studies). Also of note, the limited number of Black subjects led to large variability in the estimated treatment effect in this subgroup in Study 373 (see wide confidence interval in Figure 26). In Study 408, the numbers of non-White patients were too small to get reliable estimates of the treatment effect in these subgroups.

Figure 26. Estimated Mean Differences between UMEC 62.5 and Placebo with Respect to 24-Week Change from Baseline in Tough FEV_1 in Study 373, Stratified by Different Subgroups



Estimates based on linear regression models adjusting for baseline FEV_1, smoking status, and center grouping
Solid vertical line represents overall estimated effect size
Dashed vertical line represents no effect

Figure 27. Estimated Mean Differences between UMEC 62.5 and Placebo with Respect to 12-Week Change from Baseline in Tough FEV_1 in Study 408, Stratified by Different Subgroups



Estimates based on linear regression models adjusting for baseline FEV_1, smoking status, and center grouping
Solid vertical line represents overall estimated effect size
Dashed vertical line represents no effect

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 3.2.5.2. There were substantial missing data in the phase 3 efficacy studies, with overall dropout rates ranging from 18% to 25%. 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 analysis likely does not provide a reliable estimate of the truth. The MMRM analysis, 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_1 likely do not persist more than a few days after patients stop using them, this assumption is not scientifically plausible.

Therefore, we gave importance to a supportive analysis that multiply imputed missing data under the assumption that dropouts on UMEC 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 analysis. None of the sensitivity analyses proposed by the applicant allow for the possibility that dropouts on UMEC could have experienced ***worse*** outcomes after discontinuation than dropouts on control. However, the observed trend toward greater FEV_1 on UMEC 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 UMEC were similar to those on placebo. However, because patients were not followed up after treatment discontinuation for a safety evaluation through the complete double-blind period, we cannot rule out the possibility that: (1) differences in patient characteristics between dropouts on placebo and UMEC induce bias in safety comparisons; or (2) UMEC has residual effects that increase risk of adverse events after patients stop taking the therapy.

- Use of the surrogate marker FEV_1 as the primary efficacy endpoint

The primary endpoint in the phase 3 efficacy studies was the mean change from baseline in trough FEV_1 at 12 or 24 weeks. We consider FEV_1 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_1 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_1 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 FEV1, 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 additional endpoints ascertained in the phase 3 studies might be considered to provide some direct measure of how patients function or feel in daily life: COPD exacerbation, rescue medication use, and SGRQ score. In Study 373, UMEC 62.5 provided the following estimated benefits over placebo for these additional endpoints: mean difference in SGRQ of -4.7 (95% CI: -7.1, -2.3), mean difference in average daily rescue medication use of -0.3 (95% CI: -0.8, 0.2), and hazard ratio for incident COPD exacerbation of 0.60 (95% CI: 0.37, 0.96). There were also trends toward benefit for UMEC with respect to SGRQ and rescue medication use in Study 408. Therefore, results for these secondary assessments provide additional support for the effectiveness of UMEC in COPD. The observed trends toward benefit increase confidence that the treatment effect on the surrogate marker FEV_1 will reliably predict clinical benefit.

5.2 Collective Evidence

The collective evidence supports the effectiveness of umeclidinium 62.5 mcg for once-daily maintenance bronchodilator treatment of airflow obstruction in patients with chronic obstructive pulmonary disease. In Studies 373 and 408, treatment with UMEC 62.5 provided statistically significant 0.12 L (95% CI: 0.08, 0.16) and 0.13 L (95% CI: 0.05, 0.20) mean improvements over placebo, respectively, in the primary endpoint, change from baseline in trough FEV_1. Similar results were observed for the higher 125 mcg dose of UMEC. Missing data supportive 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 was also supported by trends toward benefit with respect to several additional endpoints, including SGRQ score, daily rescue medication use, and COPD exacerbation rate. These trends toward benefit increase confidence that the treatment effect on the surrogate endpoint FEV_1 is likely to predict clinical benefit, i.e., improvements in how COPD patients function, feel, or survive.

The complete safety evaluation was conducted by Dr. Jennifer Pippins, the Medical Reviewer, but we performed additional analyses to explore cardiovascular risk. Rates of MACE were similar across the treatment arms, but an analysis of cardiovascular-related serious adverse events suggested a possible trend toward greater risk on UMEC as compared to placebo and tiotropium. Small numbers of events led to considerable statistical uncertainty around the estimated differences in risks between the treatment arms. The interpretability of safety analyses is also clouded by the high rates of missing data in the phase 3 studies. The large amount of missing data is primarily due to the trial design, as patients who discontinued treatment early were not followed up for safety evaluation through the complete double-blind study duration.

5.3 Labeling Recommendations

(b) (4)

Data comparing UMEC 62.5 and placebo with respect to mean changes from baseline in daily puffs of rescue salbutamol, and SGRQ total score, are available from Studies 373 and 408. There was statistical evidence of a reduction in rescue medication use on UMEC 62.5 in Study 408 (estimated mean difference: -0.7; 95% CI: -1.3, -0.1), but not in Study 373 (estimate: -0.3; 95% CI: -0.8, 0.2). Results from Studies 361 and 408 for the higher 125 mcg dose of UMEC were similar (Table 8). In Studies 373 and 408, there was statistical evidence of an improvement in SGRQ score on UMEC 62.5, with estimated effect sizes of -4.7 (95% CI: -7.1, -2.3) and -7.9 (95% CI: -12.2, -3.6), respectively. There was evidence of an effect on SGRQ score for UMEC 125 in Study 408 but not Study 373 (Table 8; mean differences of -10.9 and -0.3, respectively).

The legal requirements for labeling (21 CFR 201.57) state that section 14 *Clinical Trials* "must discuss those clinical studies that facilitate an understanding of how to use the drug safely and effectively. Ordinarily, this section will describe the studies that support effectiveness for the labeled indication(s), including discussion of study design, population, endpoints, and results, but must not include an encyclopedic listing of all, or even most, studies performed as part of the product's clinical development program." Based on these requirements, it seems appropriate for section 14 *Clinical Trials* to include results on a selected set of safety and efficacy endpoints that a prescriber would likely consider important in his or her evaluation of how to use the drug safety and effectively. A prescriber might be interested in both: (1) whether or not there is evidence of a treatment effect on each important endpoint; and (2) what magnitudes of treatment effects are consistent with the observed data.

The regulations (21 CFR 201.56) also indicate that "labeling must be informative and accurate and neither promotional in tone nor false or misleading in any particular." One possible labeling approach would be to: (1) only report results on an efficacy endpoint if there is substantial evidence of a treatment effect for that endpoint; and/or (2) only report results on an efficacy endpoint if the estimated treatment effect exceeds some estimate of a *minimal clinically important difference* (e.g., 4 units for SGRQ total score). However, such an approach to only report estimates conditional on the results being favorable for the new drug will tend to result in those estimates being biased to the extreme (e.g., towards spuriously greater benefit for the drug with respect to an efficacy endpoint). Therefore, information might be included in labeling that is inaccurate, and potentially even misleading and promotional, thus violating the regulations. It is also possible that the absence of results in labeling for an endpoint that has been included in labeling for other drugs in the same class (and/or the same disease) might be misinterpreted to imply that the drug has no effect on that endpoint. Absence of evidence is not evidence of absence, and this is best conveyed by reporting a best estimate of the treatment effect, along with a range of effects that would be consistent with the data (i.e., a confidence interval).

Based on these arguments, it seems appropriate to include results in labeling on a selected set of important safety and efficacy endpoints whether or not there is substantial statistical evidence of a treatment effect, and whether or not the estimated effect size is of a certain magnitude. Results for an endpoint should include both the best point estimate of treatment effect, and an estimate of the statistical uncertainty around that point estimate (e.g., with a confidence interval). That being said, the requirement for "informative and accurate" labeling also suggests that treatment effects (or lack thereof) on those selected important endpoints should only be reported if they have been estimated with a reasonable degree of reliability and precision. If, for example, issues such as lack of blinding, measurement error, or missing data make the results unreliable and difficult to interpret, it may not be appropriate to include the results in labeling.

Based on this rationale, results on SGRQ and rescue salbutamol should be included in the labeling for umeclidinium if: (1) these endpoints convey important information to a prescriber to help ensure adequate use; and (2) the reported results for these endpoints have reasonable reliability and precision.

An evaluation of the first condition requires clinical expertise and will not be addressed in this review. It is potentially relevant to the evaluation that information on rescue short-acting bronchodilator use has been included in many of the labels for long-acting bronchodilators in COPD (including those for arformoterol, indacaterol, formoterol, formoterol/budesonide, tiotropium, and aclidinium), while information on SGRQ score has only been included in the indacaterol label. We next discuss the second condition regarding reliability and precision.

For mean daily puffs of rescue medication use, the applicant's proposed labeling includes (b) (4) Study 373, in which there was no evidence of a treatment effect for UMEC 62.5 (especially given that Study 373 was actually larger and had greater statistical precision than Study 408). If information regarding potential treatment effects on rescue medication use is deemed important knowledge for a prescriber, results from both Study 373 and Study 408 should be reported.

For mean SGRQ score, the applicant has proposed to (b) (4) in labeling. If it is determined that SGRQ score results provide important information for a prescriber, the applicant's approach appears reasonable (although we would recommend that the (b) (4) be removed). However, it might also be informative to include results from Studies 361 and 408 for the higher 125 mcg dose, because of the lack of consistency in estimates (Table 8; mean differences of -0.3 and -10.9, respectively).

An additional consideration is whether results are sufficiently reliable for labeling, given the substantial amount of missing data. Evidence of benefit for UMEC over placebo with respect to several important efficacy endpoints was consistent across many different analyses addressing the potential effect of missing data. In addition, results from phase 3 programs for other COPD drugs have been included in labeling in the presence of similar levels of missing data. Therefore, it seems that results from the phase 3 trials of UMEC should be reported in spite of the patient dropout. However, if the estimand of interest is the effectiveness of the assigned treatment in all randomized participants, at real world achievable adherence and tolerability, supportive multiple imputation analyses provide estimated treatment effect sizes of approximately 20–30% less than the primary analyses. Consideration should be given to the inclusion in labeling of a statement about the potential bias of estimates with respect to this particular estimand.

6 REFERENCES

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 16. Baseline Characteristics in Study 361

	Placebo (N = 275)	UMEC 125 (N = 407)	Overall[1] (N = 1489)
Female	100 (36%)	137 (34%)	515 (35%)
Age (years)	62.2 (8.5)	63.1 (8.5)	62.9 (8.5)
Race			
White	238 (87%)	363 (89%)	1314 (88%)
Black	9 (3%)	4 (1%)	24 (2%)
Asian	27 (10%)	40 (10%)	148 (10%)
Other	1 (0%)	0 (0%)	3 (0%)
Hispanic/Latino	1 (0%)	0 (0%)	3 (0%)
BMI (kg/m^2)	26.5 (6.1)	26.4 (5.8)	26.6 (5.8)
Current Smoker	143 (52%)	216 (53%)	769 (52%)
FEV_1 (L)	1.3 (0.5)	1.3 (0.5)	1.3 (0.5)
GOLD Stage ($ppFEV_1$)			
Stage II (50-80%)	121 (44%)	194 (48%)	699 (47%)
Stage III (30-50%)	132 (48%)	180 (44%)	660 (45%)
Stage IV (<30%)	21 (8%)	32 (8%)	124 (8%)
Chronic Bronchitis	199 (72%)	266 (65%)	1021 (69%)
Emphysema	160 (58%)	241 (59%)	858 (58%)
Duration of COPD, years			
<1	16 (6%)	39 (10%)	116 (8%)
1-5	94 (34%)	146 (36%)	542 (36%)
5-10	101 (37%)	117 (29%)	465 (31%)
10-15	48 (17%)	58 (14%)	249 (17%)
15-20	9 (3%)	25 (6%)	58 (4%)
20-25	3 (1%)	10 (2%)	34 (2%)
>25	4 (1%)	12 (3%)	25 (2%)
Inhaled Corticosteroid Use	138 (50%)	193 (47%)	698 (47%)
Reversible to Salbutamol	77 (28%)	132 (32%)	461 (31%)
Reversible to Salbutamol and Ipratropium	146 (53%)	228 (56%)	790 (53%)
At United States site	57 (21%)	87 (21%)	316 (21%)

Cell contents are mean (standard deviation) for continuous variables or frequency (percent) for categorical variables
Abbreviations: BMI = body mass index, $ppFEV_1$ = percent predicted forced expiratory volume in 1 second
[1] Includes patients randomized to treatment with VI 25 and UMEC/VI 125/25

Table 17. Baseline Characteristics in Study 373

	Placebo (N = 280)	UMEC 62.5 (N = 418)	Overall[1] (N = 1532)
Female	85 (30%)	120 (29%)	449 (29%)
Age (years)	62.2 (9.0)	64.0 (9.2)	63.1 (8.9)
Race			
White	237 (85%)	354 (85%)	1302 (85%)
Black	9 (3%)	14 (3%)	47 (3%)
Asian	22 (8%)	35 (8%)	126 (8%)
Other	12 (4%)	15 (4%)	57 (4%)
Hispanic/Latino	25 (9%)	37 (9%)	133 (9%)
BMI (kg/m^2)	26.9 (5.9)	26.5 (5.6)	26.8 (5.9)
Current Smoker	150 (54%)	207 (50%)	759 (50%)
FEV_1 (L)	1.2 (0.5)	1.2 (0.5)	1.2 (0.5)
GOLD Stage ($ppFEV_1$)			
Stage II (50-80%)	119 (42%)	191 (46%)	708 (46%)
Stage III (30-50%)	133 (48%)	172 (41%)	650 (43%)
Stage IV (<30%)	28 (10%)	54 (13%)	171 (11%)
Chronic Bronchitis	182 (65%)	274 (66%)	999 (65%)
Emphysema	173 (62%)	271 (65%)	953 (62%)
Duration of COPD, years			
<1	20 (7%)	36 (9%)	128 (8%)
1-5	107 (38%)	151 (36%)	575 (38%)
5-10	82 (29%)	127 (30%)	447 (29%)
10-15	51 (18%)	70 (17%)	257 (17%)
15-20	9 (3%)	15 (4%)	59 (4%)
20-25	6 (2%)	10 (2%)	37 (2%)
>25	5 (2%)	9 (2%)	29 (2%)
Inhaled Corticosteroid Use	137 (49%)	219 (52%)	780 (51%)
Reversible to Salbutamol	91 (32%)	121 (29%)	496 (32%)
Reversible to Salbutamol and Ipratropium	146 (52%)	223 (53%)	826 (54%)
At United States site	78 (28%)	118 (28%)	428 (28%)

Cell contents are mean (standard deviation) for continuous variables or frequency (percent) for categorical variables
Abbreviations: BMI = body mass index, $ppFEV_1$ = percent predicted forced expiratory volume in 1 second
[1] Includes patients randomized to treatment with VI 25 and UMEC/VI 62.5/25

Table 18. Baseline Characteristics in Study 374

	Tiotropium (N = 215)	UMEC 125 (N = 222)	Overall[1] (N = 869)
Female	62 (29%)	74 (33%)	280 (32%)
Age (years)	65.2 (8.3)	64.5 (8.3)	64.6 (8.4)
Race			
White	163 (76%)	169 (76%)	656 (75%)
Black	8 (4%)	6 (3%)	31 (4%)
Asian	36 (17%)	37 (17%)	145 (17%)
Other	8 (4%)	10 (5%)	37 (4%)
Hispanic/Latino	38 (18%)	42 (19%)	153 (18%)
BMI (kg/m^2)	26.4 (6.1)	26.4 (5.7)	26.5 (5.9)
Current Smoker	102 (47%)	98 (44%)	388 (45%)
FEV_1 (L)	1.2 (0.4)	1.1 (0.4)	1.1 (0.5)
GOLD Stage ($ppFEV_1$)			
Stage II (50-80%)	103 (48%)	86 (39%)	384 (44%)
Stage III (30-50%)	83 (39%)	106 (48%)	374 (43%)
Stage IV (<30%)	28 (13%)	29 (13%)	107 (12%)
Chronic Bronchitis	120 (56%)	120 (54%)	499 (57%)
Emphysema	136 (63%)	152 (68%)	556 (64%)
Duration of COPD, years			
<1	16 (7%)	16 (7%)	91 (10%)
1,5	83 (39%)	96 (43%)	333 (38%)
5,10	65 (30%)	65 (29%)	254 (29%)
10,15	34 (16%)	22 (10%)	114 (13%)
15-20	12 (6%)	12 (5%)	46 (5%)
20-25	3 (1%)	7 (3%)	16 (2%)
>25	2 (1%)	4 (2%)	15 (2%)
Inhaled Corticosteroid Use	115 (53%)	124 (56%)	455 (52%)
Reversible to Salbutamol	60 (28%)	75 (34%)	278 (32%)
Reversible to Salbutamol and Ipratropium	110 (51%)	130 (59%)	480 (55%)
At United States site	55 (26%)	58 (26%)	225 (26%)

Cell contents are mean (standard deviation) for continuous variables or frequency (percent) for categorical variables
Abbreviations: BMI = body mass index, $ppFEV_1$ = percent predicted forced expiratory volume in 1 second
[1] Includes patients randomized to treatment with UMEC/VI 62.5/25 and UMEC/VI 125/25

Table 19. Baseline Characteristics in Study 408

	Placebo (N=68)	UMEC 62.5 (N=69)	UMEC 125 (N=69)	Overall (N=206)
Female	26 (38%)	25 (36%)	27 (39%)	78 (38%)
Age	62.5 (8.7)	62.3 (9.5)	64.6 (8.0)	63.1 (8.8)
Race				
White	59 (87%)	61 (88%)	61 (88%)	181 (88%)
Black	8 (12%)	7 (10%)	6 (9%)	21 (10%)
Asian	1 (1%)	1 (1%)	2 (3%)	4 (2%)
Hispanic/Latino	0 (0%)	0 (0%)	0 (0%)	0 (0%)
BMI (kg/m^2)	28.0 (5.5)	27.6 (7.4)	25.4 (4.7)	27.0 (6.1)
Current Smoker	36 (53%)	37 (54%)	39 (57%)	112 (54%)
FEV_1 (L)	1.2 (0.4)	1.3 (0.6)	1.2 (0.4)	1.2 (0.5)
GOLD Stage ($ppFEV_1$)				
Stage II (50-80%)	33 (49%)	25 (36%)	34 (49%)	92 (45%)
Stage III (30-50%)	26 (38%)	30 (43%)	25 (36%)	81 (39%)
Stage IV (<30%)	9 (13%)	14 (20%)	10 (14%)	33 (16%)
Chronic Bronchitis	48 (71%)	50 (72%)	52 (75%)	150 (73%)
Emphysema	46 (68%)	48 (70%)	55 (80%)	149 (72%)
Duration of COPD, years				
<1	3 (4%)	6 (9%)	6 (9%)	15 (7%)
1,5	33 (49%)	25 (36%)	18 (26%)	76 (37%)
5,10	20 (29%)	18 (26%)	25 (36%)	63 (31%)
10,15	7 (10%)	11 (16%)	13 (19%)	31 (15%)
15-20	3 (4%)	6 (9%)	4 (6%)	13 (6%)
20-25	1 (1%)	2 (3%)	2 (3%)	5 (2%)
>25	1 (1%)	1 (1%)	1 (1%)	3 (1%)
Inhaled Corticosteroid Use	18 (26%)	15 (22%)	16 (23%)	49 (24%)
Reversible to Salbutamol	22 (32%)	13 (19%)	14 (20%)	49 (24%)
Reversible to Salbutamol and Ipratropium	31 (46%)	28 (41%)	33 (48%)	92 (45%)
At USA site	15 (22%)	17 (25%)	16 (23%)	48 (23%)

Figure 28. Empirical Distribution Function for Change from Baseline in SGRQ Score at 24 Weeks in Study 361

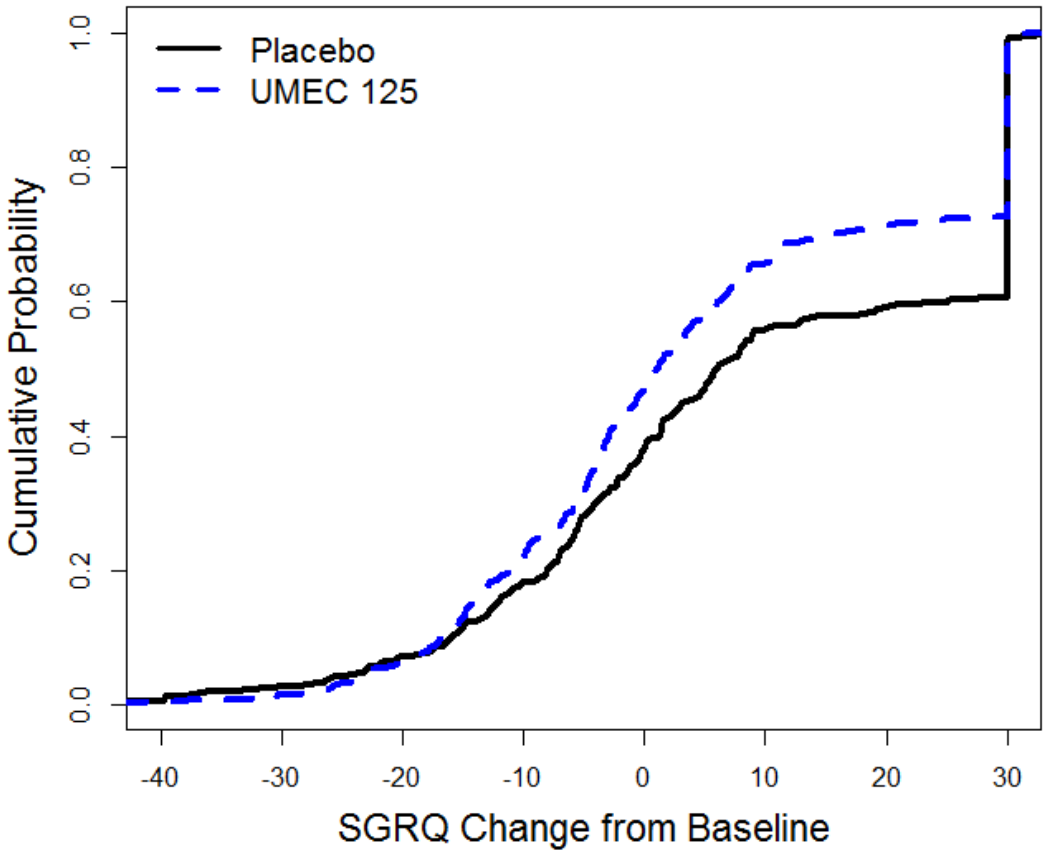


Figure 29. Empirical Distribution Function for Change from Baseline in SGRQ Score at 24 Weeks in Study 373

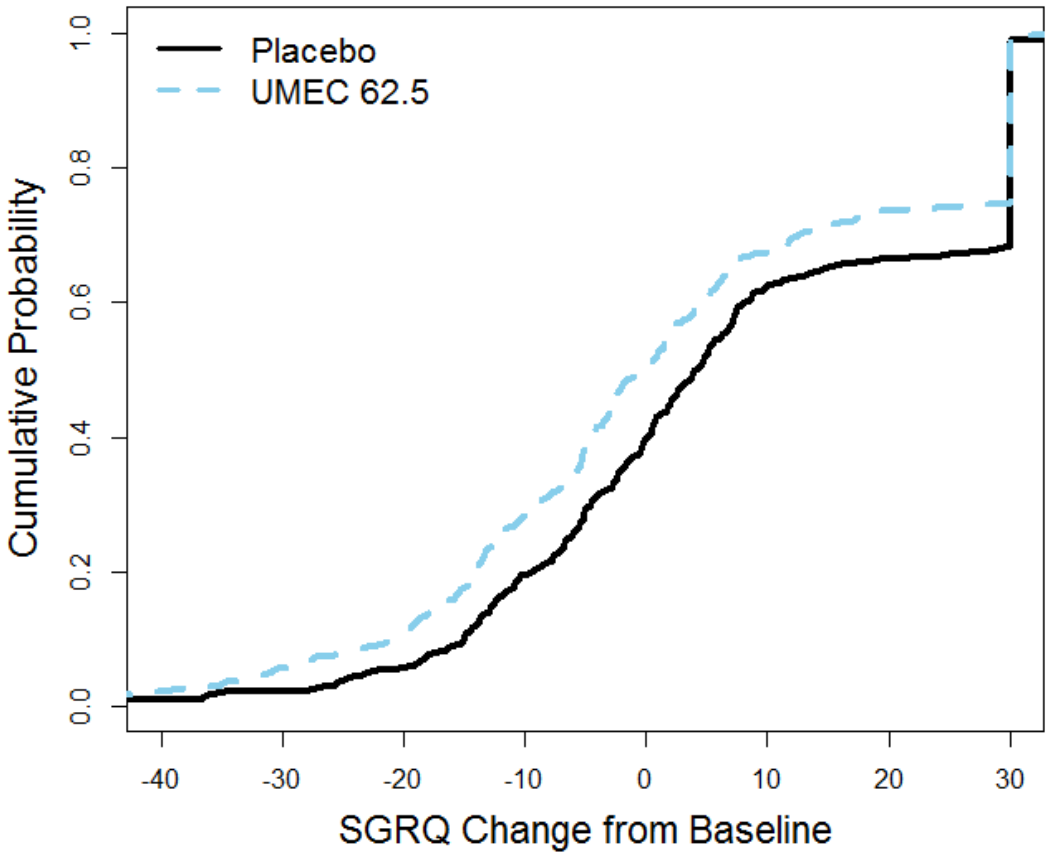


Figure 30. Empirical Distribution Function for Change from Baseline in SGRQ Score at 24 Weeks in Study 374

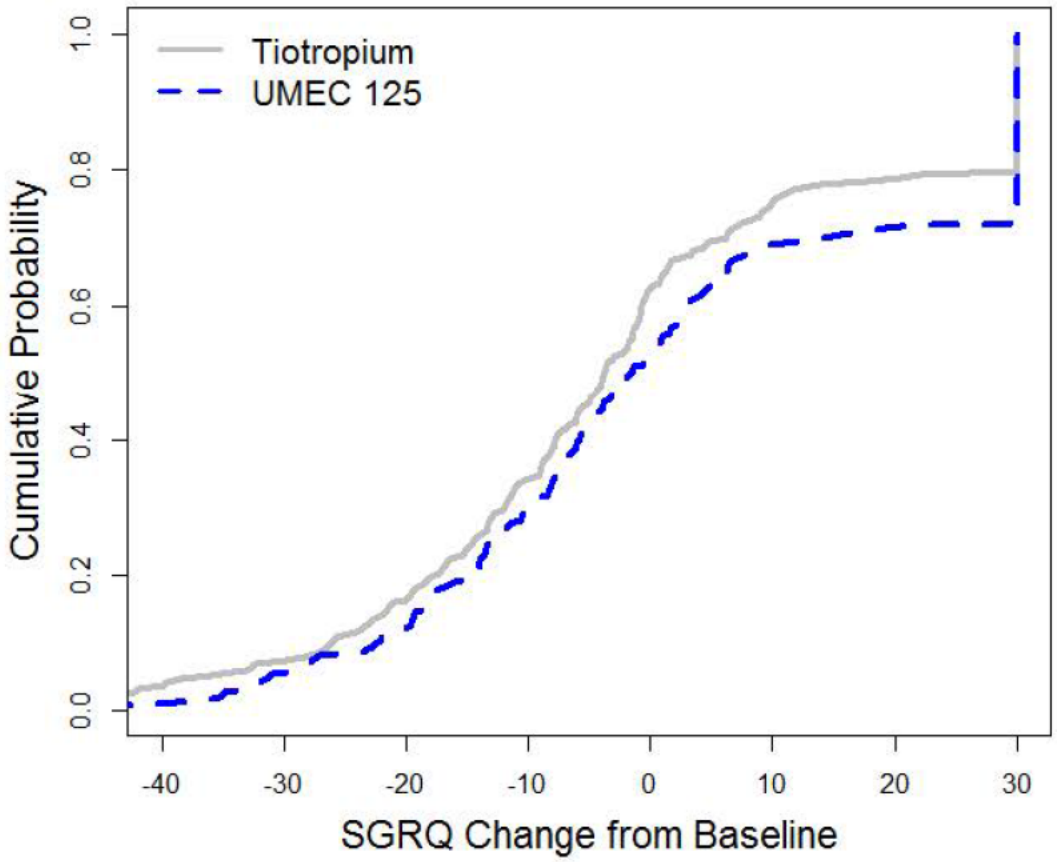


Figure 31. Empirical Distribution Function for Change from Baseline in SGRQ Score at 12 Weeks in Study 408

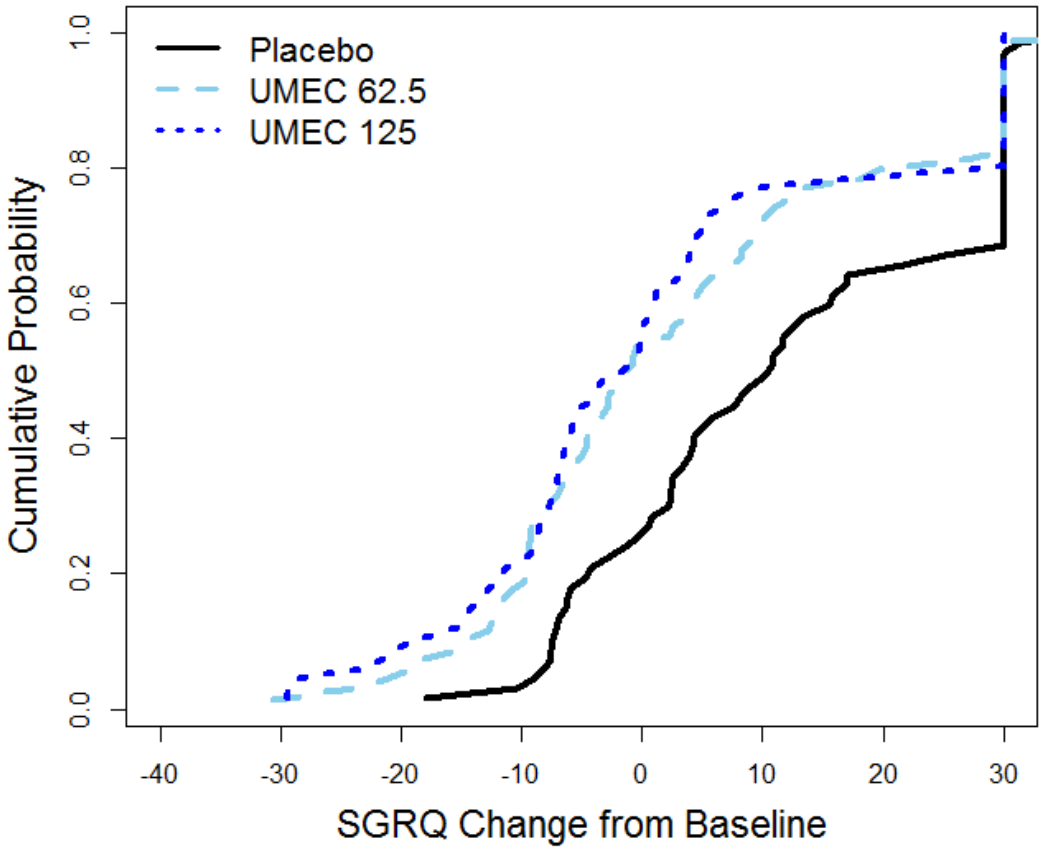
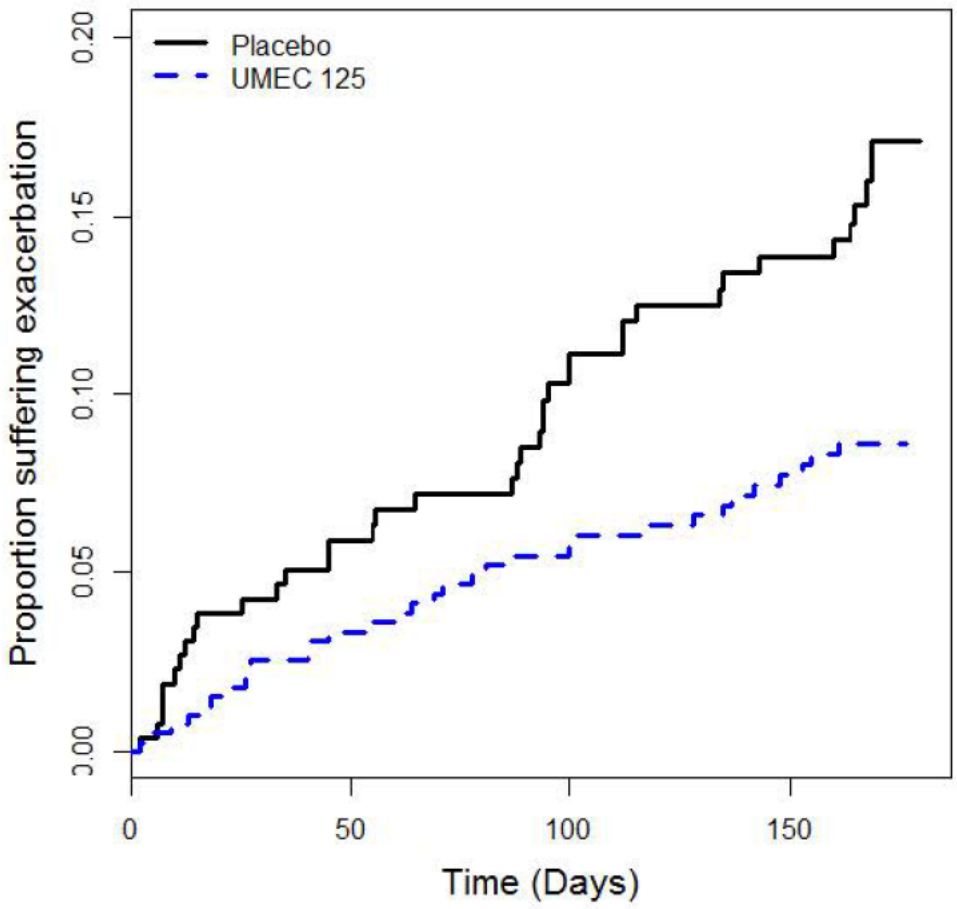


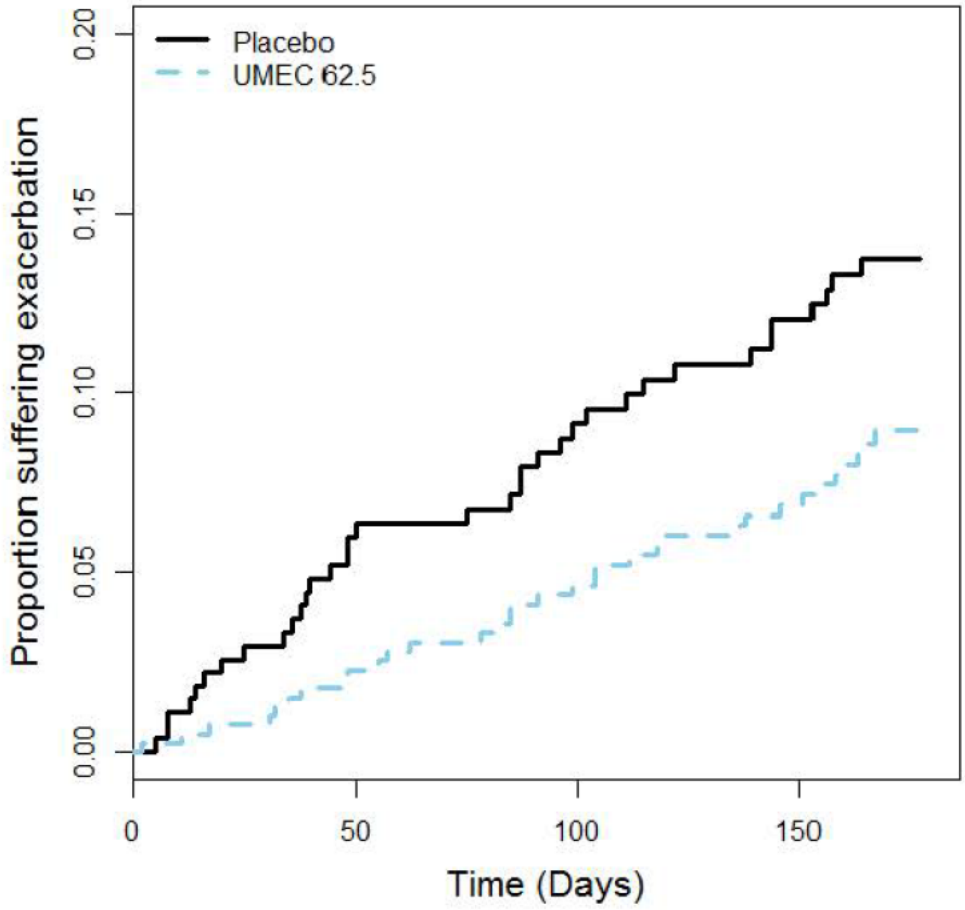
Figure 32. Proportion Suffering a COPD Exacerbation over Time in Study 361



Number at risk

Placebo	275	223	199	189
UMEC 125	407	368	337	324

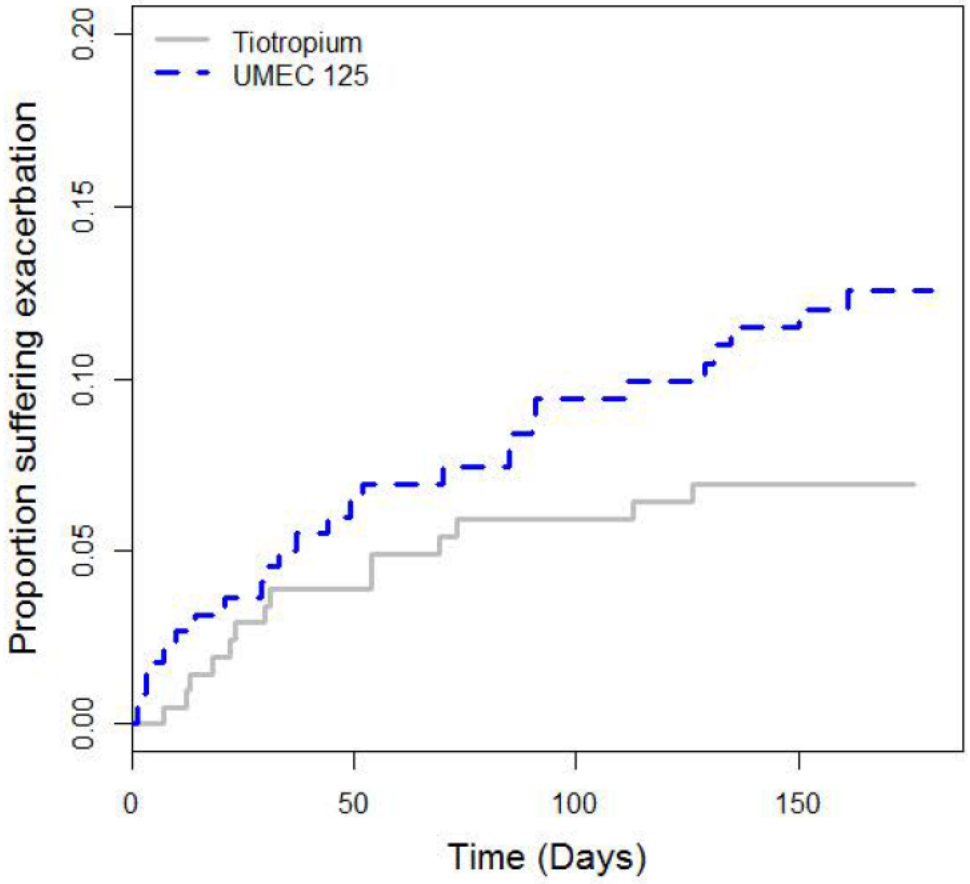
Figure 33. Proportion Suffering a COPD Exacerbation over Time in Study 373



Number at risk

Placebo	280	247	224	209
UMEC 62.5	418	378	349	334

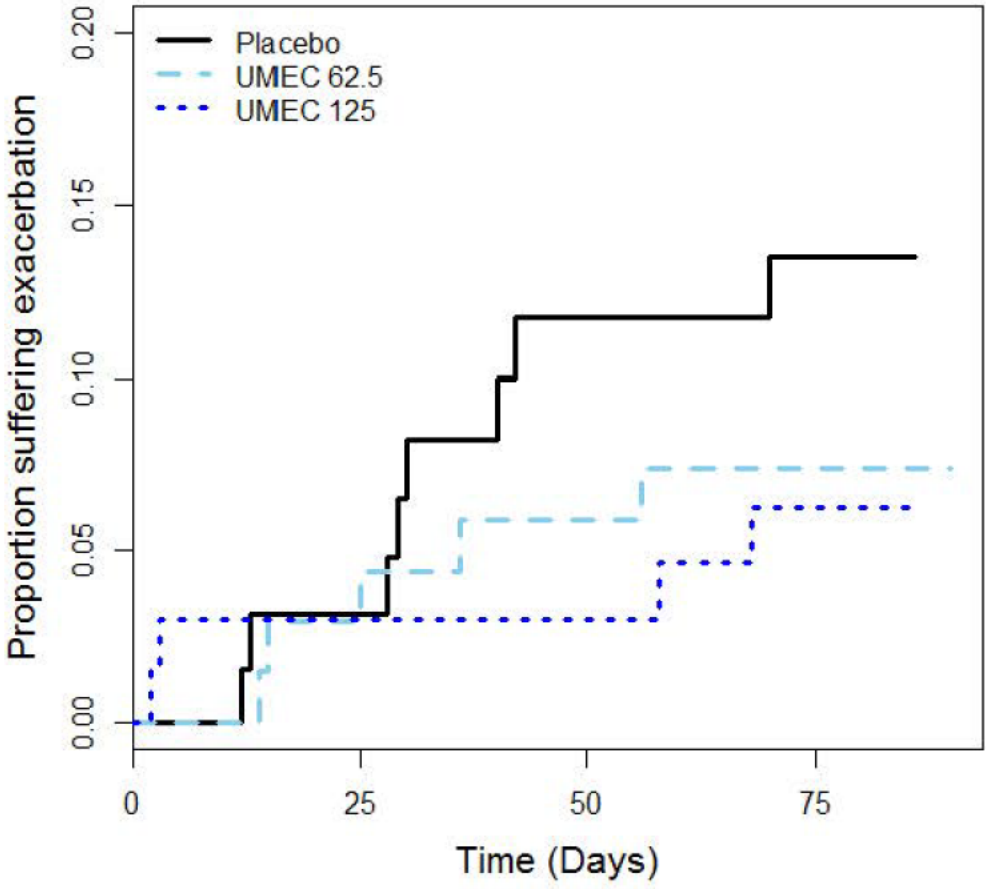
Figure 34. Proportion Suffering a COPD Exacerbation over Time in Study 374



Number at risk

Tiotropium	215	194	185	179
UMEC 125	222	196	178	167

Figure 35. Proportion Suffering a COPD Exacerbation over Time in Study 408



Number at risk

Placebo	68	57	50	49
UMEC 62.5	69	65	63	60
UMEC 125	69	62	59	57

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
02/04/2014

THOMAS J PERMUTT
02/04/2014
I concur

STATISTICAL REVIEW AND EVALUATION

Biometrics Division: VI

NDA	205382
CONSULT REQUESTED DATE	Oct 23, 2013
TRADE NAME	(b) (4) tm) ELLIPTA(tm)
ESTABLISHED NAME	umeclidinium
DOSAGE FORM	Inhaled Powder
STRENGTHS	62.5mcg
ROUTE OF ADMINISTRATION	Oral Inhalation
INDICATION	maintenance bronchodilator treatment of airflow obstruction in COPD, including chronic bronchitis and emphysema
SPONSOR	Glaxo Group Limited, England d/b/a GlaxoSmithKline
REVIEW FINISHED	Dec 2, 2013
STATISTICAL REVIEWER	Xiaoyu Dong, Ph.D.
CMC REVIEWER	CDER/OPS/ONDQA/DNDQAIII/ Arthur Shaw
PROJECT MANAGER	CDER/OPS/ONDQA/Youbang Liu

_____

Reviewer: Xiaoyu Dong, CDER/OTS/OB/DB VI

Concur: _____

Meiyu Shen, Ph.D, CDER/OTS/OB/DB VI

Distribution: NDA 205382
CDER/OTS/OB/DB VI/ Yi Tsong
CDER/OTS/OB/DB VI / Meiyu Shen
CDER/OTS/OB/ Lillian Patrician
CDER/OPS/ONDQA/DNDQAIII / Arthur Shaw
CDER/OPS/ONDQA/Youbang Liu, PM

The sponsor mentioned that they would like to have a discussion with FDA on "how sample size adjustments to the two 1-sided PTIT procedure should be implemented" (see Appendix). Biostatistics reviewer's recommendation on a generic approach on sample size adjustment for content uniformity test is as follows. For large sample content uniformity test, we recommend a one-tier parametric tolerance interval test for normally distributed data. To extend the PTI test from 30 inhalers to a larger sample size, we can find the corresponding k value and coverage by matching its OC curve with the OC curve of PTI test with 30 inhalers at the acceptance probability of (b) (4)%.

APPENDIX (FDA's Question 2 and GSK's Response)

CONFIDENTIAL

m1.11.1. Quality Information Amendment_Response to FDA Question dated Sep 26, 2013

Question 2

Regarding the Dose Content Uniformity and Dose Content Uniformity through Life test:

a. Prespecify the alternative sample sizes (b) (4)
(b) (4)

b. Your sampling approach should be such that the probability of a given batch passing will not change with a change in sample size. Your choice of the sample sizes for (b) (4)

c.

Response

Based on the Agency questions and GSK responses for NDA 204275 for Breo Ellipta Inhalation Powder and for NDA 203975 for Anoro Ellipta Inhalation Powder, GSK did not include a proposal for alternative sample sizes for Umeclidinium Inhalation Powder.

For clarity, the footnote to the specification table has been updated to remove (b) (4) when describing the number of determinations to be taken.

GSK remains concerned that there is a substantial gap between the Agency's and GSK's thinking on how sample size adjustments to the two 1-sided PTIT procedure should be implemented. GSK would like the opportunity for a non-product specific discussion on this matter with the aim of finding a generic approach to such adjustments for GSK products.

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

/s/

XIAOYU DONG
12/02/2013

MEIYU SHEN
12/02/2013

STATISTICAL REVIEW AND EVALUATION

Biometrics Division: VI

NDA	205382
CONSULT REQUESTED DATE	Oct 23, 2013
TRADE NAME	(b) (4) tm) ELLIPTA(tm)
ESTABLISHED NAME	umeclidinium
DOSAGE FORM	Inhaled Powder
STRENGTHS	62.5mcg
ROUTE OF ADMINISTRATION	Oral Inhalation
INDICATION	maintenance bronchodilator treatment of airflow obstruction in COPD, including chronic bronchitis and emphysema
SPONSOR	Glaxo Group Limited, England d/b/a GlaxoSmithKline
REVIEW FINISHED	Dec 02, 2013
STATISTICAL REVIEWER	Xiaoyu Dong, Ph.D.
CMC REVIEWER	CDER/OPS/ONDQA/DNDQAIII/ Arthur Shaw
PROJECT MANAGER	CDER/OPS/ONDQA/Youbang Liu

_____

Reviewer: Xiaoyu Dong, CDER/OTS/OB/DB VI

Concur: _____

Meiyu Shen, Ph.D, CDER/OTS/OB/DB VI

Distribution: NDA 205382
CDER/OTS/OB/DB VI/ Yi Tsong
CDER/OTS/OB/DB VI / Meiyu Shen
CDER/OTS/OB/ Lillian Patrician
CDER/OPS/ONDQA/DNDQAIII / Arthur Shaw
CDER/OPS/ONDQA/Youbang Liu, PM

TABLE OF CONTENTS

I. INTRODUCTION

The sponsor proposed parametric tolerance interval (PTI) test for content uniformity of the emitted dose for release. This two-tier test is outlined in the FDA's October 25, 2005 Advisory Committee of Pharmaceutical Science.

Tier 1: sample 10 inhalers with 20 results. Pass the batch if

- Mean content of the results is within (b) (4)% of target dose (55 mcg);
- The tolerance interval (Mean ± k_1 × SD) is within ((b) (4))% of target dose (55 mcg),

If the batch fails tier 1, then continue to tier 2.

Tier 2: sample additional 20 inhalers with 40 results. With a total of 60 samples, pass the batch if

- Mean content of the results is within (b) (4)% of target dose (55 mcg);
- The tolerance interval (Mean ± k_2 × SD) is within ((b) (4))% of target dose (55 mcg).

k_1 and k_2 are the tolerance factors with central (b) (4)% coverage, (b) (4)% confidence level and a sample size of 20 and 60, respectively.

A consultation request was sent to DB VI in the Office of Biostatistics regarding the sponsor's response to the Question 2 posed by the FDA (see Appendix). BD VI reviewer reviewed P.5.6 (Justification of the Specification for Umeclidimium Inhalation Powder) and evaluated the proposed PTI test. The comments are summarized in the next section. The detailed PTI test is described in Section III.

II. SUMMARY OF STATISTICAL COMMENTS

The statistical reviewer evaluated the proposed PTI test for content uniformity and the sponsor's response to Question 2 posed by the FDA. The summarized comments and statistical issues are listed below.

- If the data fit the normal distribution, in general, the sponsor's two-tier parametric tolerance interval (PTI) test for content uniformity is acceptable. As demonstrated by the sponsor's simulation study, PTI test is more stringent than the Zero Tolerance criteria in terms of passing rate. In addition, unlike PTI test, adding the second tier of the Zero Tolerance test provides little increase in the passing rate.
- The OC curves from PTI test should be simulated using batch mean and overall standard deviation, instead of using within-batch standard deviation as indicated in Page 12 of P.5.6.
- The sponsor did not include a proposal for an alternative sample size for this product.

III. SPONSOR'S ANALYSIS

The sponsor proposed parametric tolerance interval (PTI) test for content uniformity of the emitted dose for release. This two-tier test is outlined in the FDA's October 25, 2005 Advisory Committee of Pharmaceutical Science.

PTI Test

Tier 1: sample 10 inhalers with 20 results. Pass the batch if

- Mean content of the results is within (b) (4)% of target dose (55 mg);
- The tolerance interval (Mean ± k_1 × SD) is within ((b) (4))% of target dose (55 mg),

If the batch fails tier 1, then continue to tier 2.

Tier 2: sample additional 20 inhalers with 40 results. With a total of 60 samples, pass the batch if

- Mean content of the results is within (b) (4)% of target dose (55 mg);
- The tolerance interval (Mean ± k_2 × SD) is within (b) (4))% of target dose (55 mg).

k_1 and k_2 are the tolerance factors with central (b) (4)% coverage, (b) (4) confidence level and a sample size of 20 and 60, respectively. Their values (adjusted for the alpha-spending function) are listed in Table 1 below.

Table 1 – Sponsor's Table 9 in P.5.6.

Table 9 Example Coefficient Values

Coverage	K	
	First Tier Testing	Second Tier Testing
	n = 20	n = 60
(b) (4)	(b) (4)	

The OC curves of above proposed PTI test are compared to the zero tolerance test described in FDA Draft Guidance on Metered Dose Inhaler (MDI) and Dry Powder Inhaler (DPI) Drug Products - Chemistry, Manufacturing, and Controls.

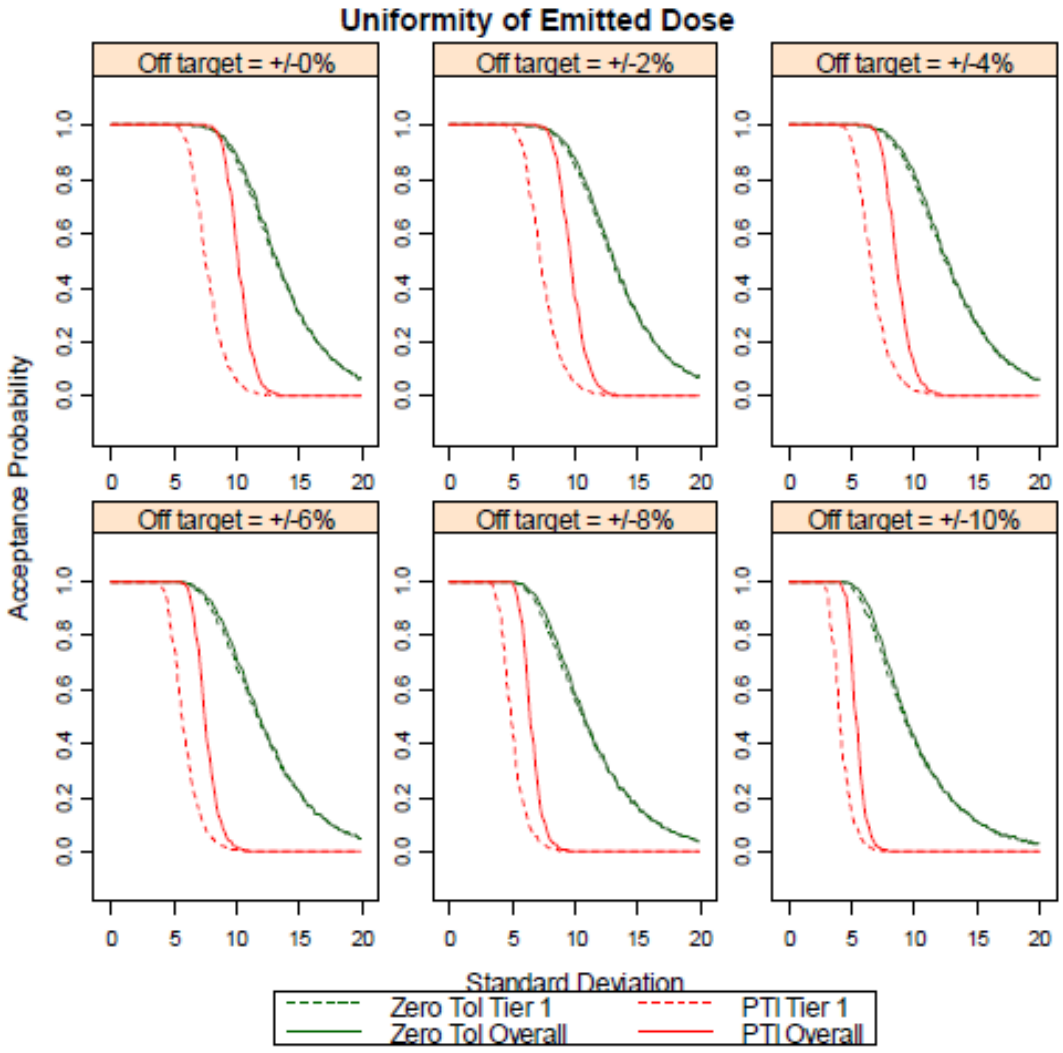
Zero Tolerance Test

Sample 10 doses (Tier 1) and 20 doses (Tier 2). Batch passes if the mean is within (b) (4) (b) (4)% of label claim (Target Dose) and (b) (4) of individual doses are within (b) (4)% of label claim (Target Dose) and (b) (4) of individual doses are within (b) (4)% of label claim (Target Dose).

The OC curves of PTI and zero tolerance tests are provided below.

Figure 1 – Sponsor's Figure 3 in P.5.6.

Figure 3 Operating Characteristic Curves for PTI Test and Zero Tolerance Criteria



As demonstrated by the sponsor's simulation study, PTI test is more stringent than the Zero Tolerance criteria in terms of passing rate. In addition, unlike PTI test, adding the second tier of the Zero Tolerance test provides little increase in the pasing rate.

IV. APPENDICES

IV.1 FDA's Question 2 and GSK's Response

CONFIDENTIAL

m1.11.1. Quality Information Amendment_Response to FDA Question dated Sep 26, 2013

Question 2

Regarding the Dose Content Uniformity and Dose Content Uniformity through Life test:

a. Prespecify the alternative sample sizes [redacted] (b) (4)

b. Your sampling approach should be such that the probability of a given batch passing will not change with a change in sample size. Your choice of the sample sizes for [redacted] (b) (4)

c. [redacted]

Response

Based on the Agency questions and GSK responses for NDA 204275 for Breo Ellipta Inhalation Powder and for NDA 203975 for Anoro Ellipta Inhalation Powder, GSK did not include a proposal for alternative sample sizes for Umeclidinium Inhalation Powder.

For clarity, the footnote to the specification table has been updated to remove [redacted] (b) (4) when describing the number of determinations to be taken.

GSK remains concerned that there is a substantial gap between the Agency's and GSK's thinking on how sample size adjustments to the two 1-sided PTIT procedure should be implemented. GSK would like the opportunity for a non-product specific discussion on this matter with the aim of finding a generic approach to such adjustments for GSK products.

IV.2 Specification Table and the Footnote

Table 1 Specification for Umeclidinium Inhalation Powder 62.5 microgram

Test	Acceptance Criteria	Method Number	
		(b) (4)	(b) (4)
Description [a]	A plastic inhaler with a light grey body, a light green mouthpiece cover and a dose counter, packed in a foil tray which contains a desiccant packet. The tray is sealed with a peelable lid. The inhaler contains one strip of either 30 or 7 regularly distributed blisters, each containing a white powder.	LAM 1918	PRS02166
Identification of Umeclidinium [b] by UV Spectrophotometry	The spectrum of the sample is concordant with that of the umeclidinium bromide reference material.	LAM 1918	ATM02319
Identification of Umeclidinium [c] by HPLC	The retention time of the principal peak in the HPLC chromatogram of the sample corresponds with the principal peak in the chromatogram for the umeclidinium bromide reference material.	LAM 1918	ATM02323/ ATM02324
Umeclidinium Content per Blister by HPLC (mcg/blister) [b,d]	(b) (4) (b) (4) of nominal blister content)	LAM 1918	ATM02320
Drug-related Impurities Content by HPLC (% w/w) [e] Any Unspecified Impurity [f] Total Impurities [g]	 Not greater than (b) (4) Not greater than (b) (4)	LAM 1918	ATM02325
Content Uniformity of Emitted Dose by HPLC (mcg/inhalation) Individual Emitted Dose Mean Emitted Dose Content L [j]	 (b) (4) coverage with (b) (4) confidence with goal posts of (b) (4) target [h,i] (b) (4) (b)	LAM 1918	ATM02323/ ATM02324

Cont'd

Notes:

a. Test performed on the packed inhaler.
b. For batch release, test may be performed on the blister strip prior to the (b) (4) of manufacture.
c. For batch release, test may be performed on the inhaler prior to the (b) (4) of manufacture.
d. For batch release the mean result from the (b) (4) IPC test is applied.
e. Drug-related impurities content not tested at release.

Notes Cont'd

f. The following synthetic impurities are controlled in the specification for umeclidinium bromide drug substance and are therefore not monitored in the drug product:
(b) (4)
Therefore the acceptance criteria for any unspecified impurity will be applied to any synthetic or degradation impurity that is not specified in either the umeclidinium bromide drug substance or drug product drug-related impurity specification.

g. Total drug-related impurities content includes total synthetic and total degradation impurities.

h. The target emitted dose for drug product at release and during storage in the secondary pack is 55 micrograms. For product stored outside of the secondary pack the target is the mean emitted dose.

i. **Tier 1:** Collect 20 individual determinations (b) (4) from 10 inhalers (2 determinations from each inhaler) sampled throughout the manufacturing process according to the defined sampling plan. For stability testing an alternative sampling scheme may be applied provided that the number of individual determinations is maintained. The batch passes if (b) (4)% Target $\leq (X_1 - K_1 \bullet S_1) \leq (X_1 + K_1 \bullet S_1) \leq$ (b) (4)% Target. X_1, S_1 are the sample mean of 20 determinations and sample standard deviation of 20 determinations, respectively, where K_1 = (b) (4). The mean of the batch is within (b) (4) of Target.
If the test is not passed at Tier 1, go to the second tier.
Tier 2: Collect a further 40 individual determinations (b) (4) from 20 inhalers (2 determinations from each inhaler) sampled throughout the manufacturing process according to the defined sampling plan. For stability testing an alternative sampling scheme may be applied provided that the number of individual determinations is maintained. Combine the results of both tiers.
The batch passes if (b) (4)% Target $\leq (X_2 - K_2 \bullet S_2) \leq (X_2 + K_2 \bullet S_2) \leq$ (b) (4)% Target. X_2, S_2 are the sample mean of 60 determinations and sample standard deviation of 60 determinations, respectively, where K_2 = (b) (4). The mean of the batch is within (b) (4) of Target.

j. Using the overall mean and standard deviation expressed as a % of Target, calculate the L value:
L Value = (b)
Where k = constant defined above for k_1 and k_2 as appropriate.

k. Each specification parameter is applied to individual inhalers. For all the inhalers tested, the total amount of umeclidinium recovered from the (b) (4) must be within (b) of the target emitted dose.

l. In-house limits are applied at time of batch release to ensure that the product will meet the regulatory specification throughout its shelf-life.

m. Test performed as an (b) (4).

n. If Water Activity criteria are met then MLT (Tier 2) testing is not required. If the acceptance criterion for Water Activity is not achieved, then MLT (Tier 2) testing must be performed.

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

/s/

XIAOYU DONG
12/02/2013

MEIYU SHEN
12/02/2013

STATISTICAL REVIEW AND EVALUATION

Biometrics Division: VI

NDA	205382
CONSULT REQUESTED DATE	Oct 17, 2013
TRADE NAME	(b) (4) tm) ELLIPTA(tm)
ESTABLISHED NAME	umeclidinium
DOSAGE FORM	Inhaled Powder
STRENGTHS	62.5mcg
ROUTE OF ADMINISTRATION	Oral Inhalation
INDICATION	maintenance bronchodilator treatment of airflow obstruction in COPD, including chronic bronchitis and emphysema
SPONSOR	Glaxo Group Limited, England d/b/a GlaxoSmithKline
REVIEW FINISHED	Nov 13, 2013
STATISTICAL REVIEWER	Xiaoyu Dong, Ph.D.
CMC REVIEWER	CDER/OPS/ONDQA/DNDQAIII/ Arthur Shaw
PROJECT MANAGER	CDER/OPS/ONDQA/Youbang Liu

_____

Reviewer: Xiaoyu Dong, CDER/OTS/OB/DB VI

Concur: _____

Yi Tsong, Acting Division Director, Ph.D, CDER/OTS/OB/DB VI

Distribution: NDA 205382
CDER/OTS/OB/DB VI/ Yi Tsong
CDER/OTS/OB/ Lillian Patrician
CDER/OPS/ONDQA/DNDQAIII / Arthur Shaw
CDER/OPS/ONDQA/Youbang Liu, PM

TABLE OF CONTENTS

I. INTRODUCTION

The sponsor proposed to apply the tolerance interval method to establish the acceptance criteria (AC) of the Aerodynamic Particle Size Distribution (APSD) by (b) (4) for Coarse Particle Mass (CPMass), Fine Particle Mass (FPMass), and Very Fine Particle Mass (vFPMass). In the sponsor's analysis, a mixed effect model was applied to compute the tolerance interval using pooled data from clinical batches, primary stability batches, batch characterisation, and developement batches. The original and revised proposed ACs are listed below.

Table 1 – Sponsor's Table 2 in Quality Information Amendment_Response to FDA Question dated Sept 26, 2013

Table 2 Umeclidinium Inhalation Powder 62.5 microgram APSD Acceptance Criteria

Test Parameter	GSK Original Proposed Acceptance Criteria (micrograms)[a]	GSK Revised Proposed Acceptance Criteria (micrograms)
Coarse Particle Mass	(b) (4)	(b) (4)
Fine Particle Mass		
Very Fine Particle Mass		

Note:
a. With (b) (4)

A consultation request was sent to DB VI in the Office of Biostatics regarding the validity of the proposed AC. BD VI reviewer performed an independent analysis of the submitted data. Please refer to Section II for the summarized results and Section VI for the detailed analysis.

II. SUMMARY OF STATISTICAL COMMENTS

The statistical reviewer performed an independent evaluation of the APSD data from from clinical batches, primary stability batches, batch characterization, and developement batches. The summarized comments and statistical issues are listed below.

- In general, acceptance criteria obtained by tolerance interval method is wider than the underlying unknown acceptance criteria interval. Considering this, we recommend using confidence limtis of percentiles to compute the AC interval. Such obtaied AC interval is tigher than the underlying AC interval, thus can better assure the product quality.
- The accpetance criteria for CPMass, FPMass, and vFPMass using our proposed method along with GSK proposed AC are listed in Table 2. As can be seen, the AC

proposed by GSK are all wider than our proposed values. With larger sample size, our proposed AC will be closer to the AC interval estimated by tolerance interval.

Table 2 – FDA Reviewer Proposed Acceptance Criteria using Confidence Limits of Percentiles and GSK Proposed Acceptance Criteria using Tolerance Interval Method

Test	FDA Proposed (All Data[1])	FDA Proposed (Clinical Batches)	GSK Original Proposed	GSK Revised Proposed
CPMass	(b) (4)		(b) (4)	(b) (4)
FPMass			(b) (4)	
vFPMass			(b) (4)	

[1]All data contains data from Clinical Batches, Batch Characterization, Development Batches, and Primary Stability Batches.

- In addition, the model GSK used is the mixed effect model with unbalanced data. However, the formula of tolerance factor they applied does not correspond to the above model. The correct tolerance factor should be derived based on the mixed effect with unbalanced data.

III. SPONSOR'S ANALYSIS

The summarized statistics for APSD data from Clinical Batches, Batch Characterization, Development Batches and Primary Stability Batches are listed in Table 3 below (Table 1 in sponsor's Quality Information Amendment_Response to FDA Question dated Sept 26, 2013). To compute the acceptance criteria for each APSD group, GSK applied the two one-sided tolerance interval method based on the formula below

$$AC = Mean \pm k \times SD$$

where mean is the average of the data, SD is the square root of (between batch variability + between time point variability + within batch and time point variability), and *k* is the one-sided tolerance factor with (b) (4)% coverage and (b) (4)% confidence level. The formula for k is given by

(b) (4)

(b) (4) quantile of the t-distribution with (n-1) degree of freedom and non-centrality parameter of $Z_{0.995}\sqrt{n}$. GSK proposed AC are listed in Table

4 (Table 2 in Quality Information Amendment_Response to FDA Question dated Sept 26, 2013). For Vary Fine Particle Mass, one-sided tolerance interval with coverage (b) (4)% and confidence level of (b) (4)% was applied.

Table 3 – Sponsor's Table 1 in Quality Information Amendment_Response to FDA Question dated Sept 26, 2013

Table 1 Summary of APSD data

		Umeclidinium (micrograms per inhalation)		
		CPMass	FPMass	vFPMass
Clinical Batches (n) = (b) (4) (d) = 40	Mean Range	(b) (4)		
Primary Stability Batches (n) = (b) (4) (d) = 119	Mean Range			
Batch Characterisation[a] (n) = (b) (4) (d) = 99	Mean Range			
Development Batches (n) = (b) (4) (d) = 369	Mean Range			
Overall (n) = (b) (4) (d) = 627	Mean Range			

Notes:
(n) = Number of Batches
(d) = Number of Determinations
[a] Data originally generated to 3 decimal places, reported here to 1 decimal place
[b] The 3 batch characterisation and primary stability batches are the same batches and have the same batch numbers
[1 & 4] Primary Stability Batches [5] Batch Characterisation on Primary Stability Batches
[2, 3 & 5] Development Batches

Table 4 – Sponsor's Table 2 in Quality Information Amendment_Response to FDA Question dated Sept 26, 2013

Table 2 Umeclidinium Inhalation Powder 62.5 microgram APSD Acceptance Criteria

Test Parameter	GSK Original Proposed Acceptance Criteria (micrograms)[a]	GSK Revised Proposed Acceptance Criteria (micrograms)
Coarse Particle Mass	(b) (4)	(b) (4)
Fine Particle Mass		
Very Fine Particle Mass		

Note:
a. With (b) (4)

IV. FDA STATISTICS REVIEWER'S ANALYSIS

IV.1 Comparison of Various Campaign Data

We analyzed the submitted data from Batch Characterization, Clinical Batches, Development Batches, and Primary Stability Batches for each APSD group. The summarized results are shown in Table 5.

Table 5 – FDA Reviewer's Analysis for Each APSD Group and Campaign

Analysis Variable : Mass							
Group	**Campaign**	**N Obs**	**Mean**	**Std Dev**	**Minimum**	**Maximum**	**Range**
CPMass	Batch Characterization	(b) (4)					
	Clinical						
	Development						
	Primary Stability						
FPMass	Batch Characterization						
	Clinical						
	Development						
	Primary Stability						
Total	Batch Characterization						
	Clinical						
	Development						
	Primary Stability						
vFPMass	Batch Characterization						
	Clinical						
	Development						
	Primary Stability						

We also displayed the data using histogram. As Figures 1a to 1d show, the data are roughly bell shaped and show similar distribution profiles for different campaigns. The bimodal pattern observed in certain histograms can be caused by the batch-to-batch variability. The histograms here were plotted by the pooled data from all batches. Thus each batch mean may locate differently.

Figure 1a – Histogram of CPMass data

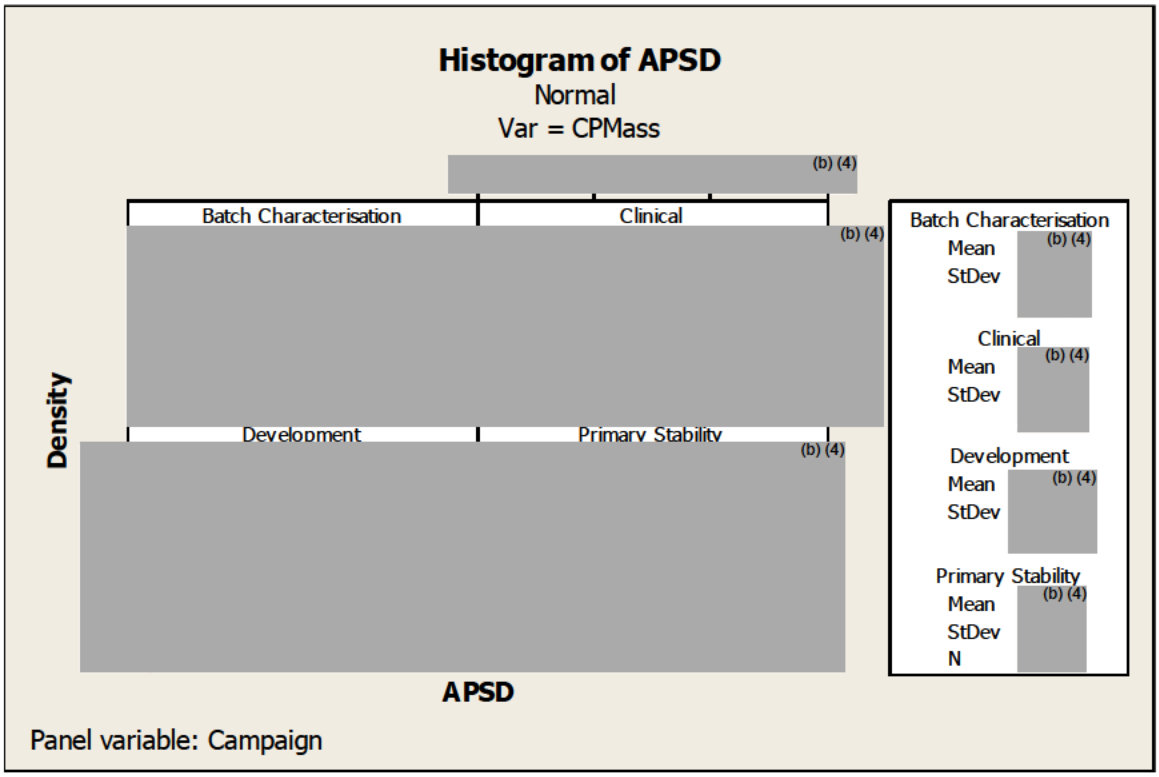


Figure 1b – Histogram of FPMass data

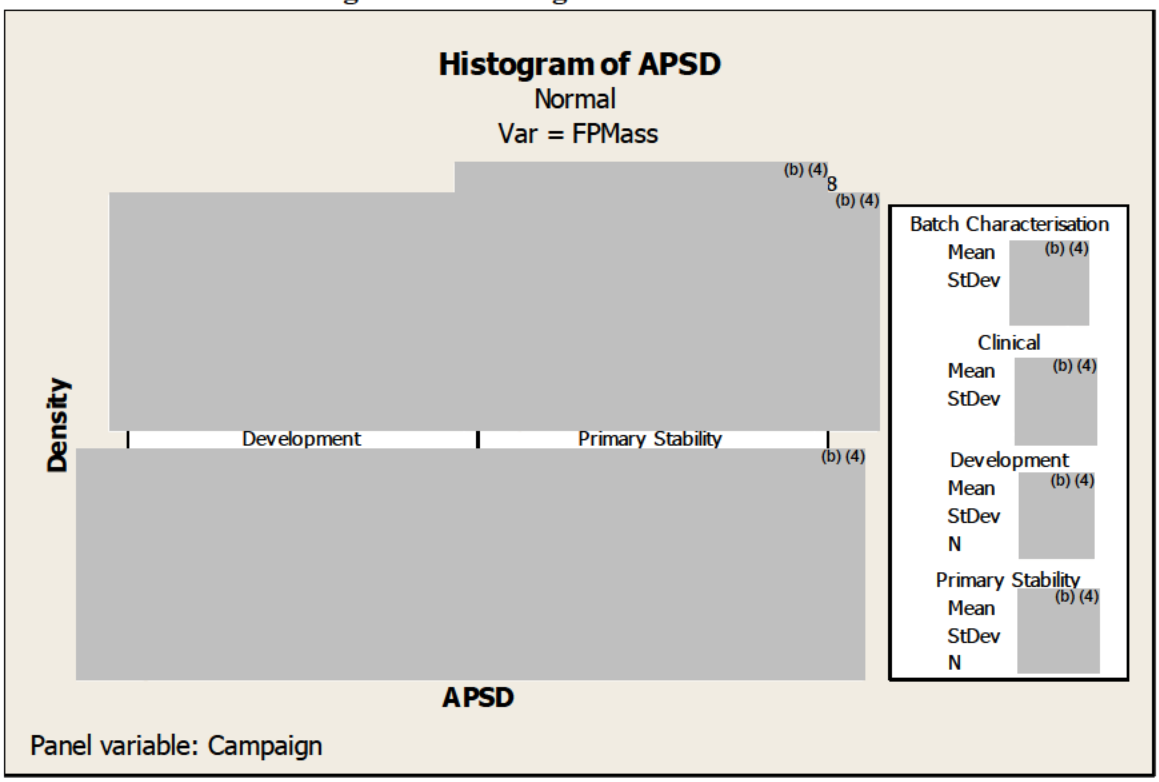


Figure 1c – Histogram of vFPMass data

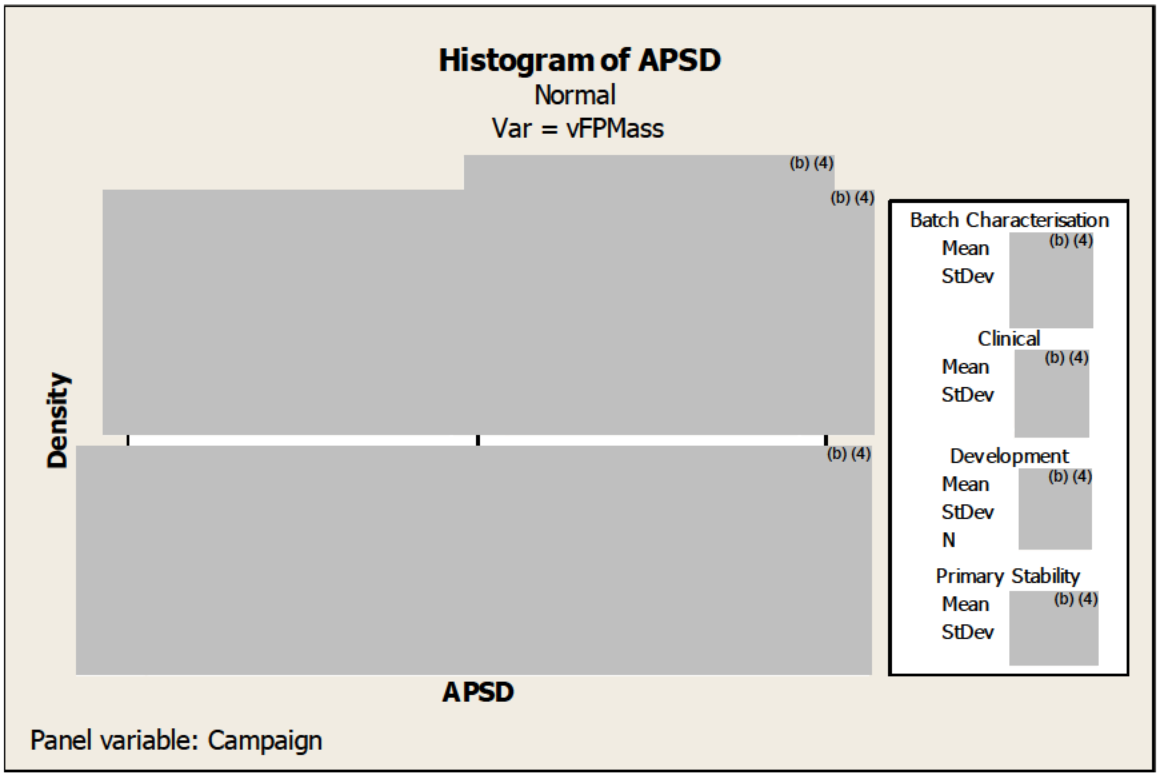
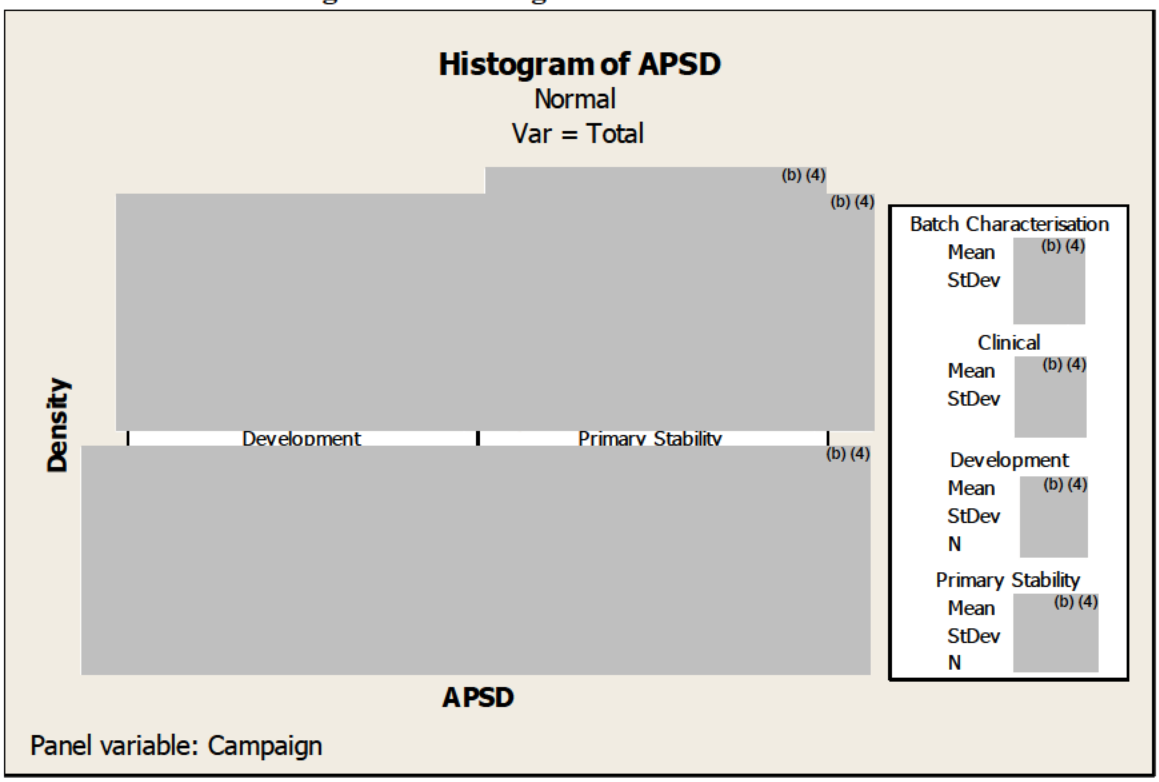


Figure 1d – Histogram of Total Mass data



IV.2 Batch-to-Batch Variability

Considering that there are multiple batches from each campaign, we compute the batch-to-batch variability in Table 6. As can be seen, such variability is noticeable. For FPMass, batch-to-batch variability for all campaigns is larger than the within batch variability with (b) (4) of the total variability for Batch Characterization, Clinical Batches and Development Batches, respectively.

Table 6 – Estimated Batch-to-batch Variability and Within Batch Variablity. Batch% is the Percentage of the Batch-to-batch Variability Relative to the Total Variability.

Group	Campaign	Batch-to-Batch	Within Batch	Batch%
CPMass	Batch Characterization	(b) (4)		
	Clinical			
	Development			
FPMass	Batch Characterization			
	Clinical			
	Development			
vFPMass	Batch Characterization			
	Clinical			
	Development			
Total	Batch Characterization			
	Clinical			
	Development			

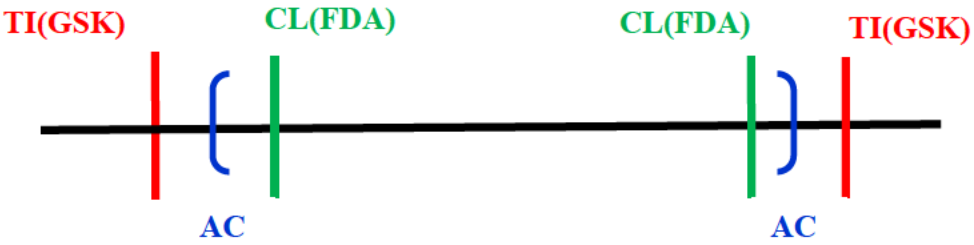
IV.3 FDA Proposed Acceptance Criteria

The AC proposed by the sponsor aim to cover central (b) (4) of the APSD distribution. In other words, the proportion of mass less than the lower limit of the AC is (b) (4); and the proportion of mass more than the upper limit of the AC is (b) (4) as well. With this requirement, we can estimate the AC limits using the confidence interval of the percentiles. Let (L, U) denote the target AC interval, then we can estimate L by $\hat{L}$ which is the (1-α)% upper confidence limit of the (b) (4)%-th percentile and estimte U by $\hat{U}$ which is the (1-α)% lower confidence limit of the (b) (4)%-th percentile.

The plot below displays the difference between the AC estimated by the tolerance interval (GSK proposed) and estimated by the confidence limit of percentiles (FDA proposed). As Figure 2 shows, tolerance interval will be wider than the target AC interval with high probability (say (b) (4)%). On the other hand, confidence limits for percentiles are

tighter than the target AC interval with high probability. Thus, using confidence limit of percentiles can better assure the product quality by applying a tighter AC. Please note, the target AC (in blue in Figure 2) is usually unknown due to limited data.

Figure 2 – Graphic Display of the underlying true AC (in blue), Tolerance Interval (in red), and Confidence Limits of Percentiles (in green)



We apply non-parametric method, which does not need any distribution assumption, to obtain the confidence limits $\hat{L}$ and $\hat{U}$ by the following simulation scheme.

(b) (4)

Apply the above simulation scheme, our proposed AC for each APSD group using all data (Clinical Batches, Batch Characterization, Development Batches and Primary Stability Batches) and using Clinical Batches only are listed in the table below.

Table 7 – FDA Reviewer Proposed Acceptance Criteria using Confidence Limits of Percentiles and GSK Proposed Acceptance Criteria using Tolerance Interval

Test	FDA Proposed (All Data)	FDA Proposed (Clinical Batches)	GSK Original Proposed	GSK Revised Proposed
CPMass	(b) (4)		(b) (4)	(b) (4)
FPMass			(b) (4)	
vFPMass			(b) (4)	

As expected, our proposed AC using confidence limits are tighter than GSK proposed AC using tolerance interval methods.

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

/s/

XIAOYU DONG
11/13/2013

YI TSONG
11/14/2013

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

NDA FILING REVIEW

NDA #:	205,382
Drug Name:	Umeclidinium inhalation powder
Indication(s):	Long-term, once-daily, maintenance bronchodilator treatment of airflow obstruction in patients with chronic obstructive pulmonary disease (COPD), including chronic bronchitis and emphysema
Applicant:	GlaxoSmithKline
Date(s):	June 13, 2013

Biometrics Division:	Division of Biometrics II
Statistical Reviewer:	Gregory Levin, PhD
Concurring Reviewers:	Joan Buenconsejo, 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:	Angela Ramsey

Keywords: NDA filing review

INTRODUCTION

The applicant has submitted the results of several studies to support the efficacy and safety of umeclidinium (UMEC) inhalation powder for once-daily maintenance bronchodilator treatment of airflow obstruction in patients with chronic obstructive pulmonary disorder (COPD). UMEC is a long-acting muscarinic antagonist that is also being developed in combination with the long-acting beta$_2$-antagonist vilanterol (VI). The UMEC/VI combination product is currently being reviewed under NDA 203,975.

The clinical development program for UMEC and UMEC/VI included two doses of UMEC (62.5 and 125 mcg once daily). The applicant has submitted results from the following seven phase 3 clinical trials in support of the efficacy and safety of the UMEC monotherapy: Studies AC4115408, DB2113361, DB2113373, DB2113374, DB2114417, DB2114418, and DB2113359, which we will refer to as Studies 408, 361, 373, 374, 417, 418, and 359, respectively. These were randomized, double-blind clinical trials in COPD patients with moderate to very severe airflow obstruction and an extensive cigarette smoking history. Concurrent use of systemic corticosteroids or long-acting bronchodilators was prohibited, but use of inhaled corticosteroids at a stable dose and rescue salbutamol was allowed. Study 408 was a 12-week, placebo-controlled, parallel-group trial, Studies 361 and 373 were 24-week, placebo-controlled, parallel-group trials, Study 360 was a 24-week, tiotropium-controlled, parallel-group trial, Studies 417 and 418 were 2-period (12 weeks per period), placebo-controlled, cross-over trials, and Study 359 was a 52-week, placebo-controlled, parallel-group, safety trial.

The primary placebo-controlled efficacy results for the proposed 62.5 mcg once daily dose of the UMEC monotherapy come from Studies 408 and 373. The prespecified primary efficacy endpoint was the mean change from baseline in trough FEV_1 at either Week 12 (Study 408) or Week 24 (Study 373). A prespecified secondary efficacy endpoint was the change from baseline in weighted mean FEV_1 0 to 6 hours post-dose. FEV_1 at 1, 3, 6, 23, and 24 hours post-dose on Days 1 and 84 were also secondary endpoints in Study 408. A number of other efficacy endpoints were prespecified in both studies, including additional spirometry measurements, daily rescue salbutamol use, and St. George's Respiratory Questionnaire (SQRQ) score. Time to first COPD exacerbation and the Shortness of Breath with Daily Activities (SOBDA) questionnaire score were additional efficacy endpoints only in Study 373.

FILING SUMMARY

There are no filing issues from a statistical perspective. We are able to locate necessary data files, summaries, and reports, and data sets are accessible and appropriately documented. Safety and efficacy were investigated by gender, racial, and age subgroups.

POTENTIAL REVIEW ISSUES

We have identified the following topics to be further assessed as part of the statistical review of this application: (1) the potential impact of missing data on the reliability of efficacy and safety results, and (2) evidence in support of the (b) (4)

With respect to the impact of missing data, we do not find the sensitivity analyses provided by the applicant for Study 408 to be sufficient. All four multiple imputation approaches (missing at random, copy differences from control, last mean carried forward, and last mean -25 mL/year carried forward) more or less impute post-dropout data by preserving the mean treatment effect that was observed prior to discontinuation. This may not be appropriate, since any positive effects of the bronchodilator on FEV_1 prior to dropout likely declined or went completely away once the patient stopped taking the therapy. We request that the applicant provides results based on additional sensitivity model(s) that do not preserve the pre-dropout treatment effect after patients stop taking the therapy. For example, the "copy reference" and "jump to reference" approaches that the applicant implemented under NDA 203,975 are additional models of interest.

FILING CHECKLIST

On **initial** overview of the NDA/BLA application for refuse-to-file:

	Content Parameter	Yes	No	NA	Comments
1	Index is sufficient to locate necessary reports, tables, data, etc.	X			
2	ISS, ISE, and complete study reports are available (including original protocols, subsequent amendments, etc.)	X			
3	Safety and efficacy were investigated for gender, racial, and geriatric subgroups (if applicable).	X			
4	Data sets in EDR are accessible and conform to applicable guidances (e.g., existence of define.pdf file for data sets).	X			

IS THE STATISTICAL SECTION OF THE APPLICATION FILEABLE? __YES_____

If the NDA/BLA is not fileable from the statistical perspective, state the reasons and provide comments to be sent to the Applicant.

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
06/13/2013

JOAN K BUENCONSEJO
06/13/2013
I concur