Statistical Analysis Plan

Study Code	PT001102
NCT #	NCT03358147
Date	12 August 2019

A Randomized, Double-Blind, Parallel Group, Multi-Center 24 Week Study Comparing the Efficacy and Safety of Three Doses of PT001 to Placebo and Open-label Spiriva® Respimat® in Subjects With Persistent Asthma

STATISTICAL ANALYSIS PLAN FOR STUDY PT001102

Protocol Number:	PT001102-04 (Version 5.0)
Investigational Drug and Drug Number:	Glycopyrronium Inhalation Aerosol (GP MDI; PT001) Spiriva® Respimat®
Indication:	Asthma
Dosage Form/Dose:	• GP MDI 28.8 µg BID • GP MDI 14.4 µg BID • GP MDI 7.2 µg BID • Placebo MDI BID • Spiriva Respimat 2.5 µg QD

PT001102 Protocol Title: A Randomized, Double-Blind, Parallel Group, Multi-Center 24-Week Study Comparing the Efficacy and Safety of Three Doses of PT001 to Placebo and Open-label Spiriva Respimat in Subjects With Persistent Asthma

Date of Issue: 12 August 2019

Version: Version 1.0

Signed Agreement on Statistical Analysis Plan

FINAL SIGN-OFF SIGNATURES

Author: [redacted]

Peer Reviewer: [redacted]

Approved by: [redacted]

Approved by: [redacted]

Change Log			
Version No.	**Effective Date**	**Reason for the Change / Revision**	**Supersedes**

TABLE OF CONTENTS

List of Tables

List of Figures

LIST OF ABBREVIATIONS AND DEFINITIONS OF TERMS

ACQ	Asthma Control Questionnaire
ACQ-5	ACQ questions 1 through 5, which measure the frequency, intensity or limitations from asthma symptoms using 1-week recall
ACQ-6	ACQ questions 1 through 6, which is the ACQ-5 plus 1 item that scores the average number of daily puffs needed from rescue medication using 1-week recall
ACQ-7	ACQ questions 1 through 7, which is the ACQ-6 plus 1 item that scores lung function (FEV_1 percent predicted)
AE	Adverse event
AESI	Adverse event of special interest
ALT	Alanine aminotransferase
ANCOVA	Analysis of covariance
AQLQ+12	Asthma Quality of Life Questionnaire for 12 years and older
AR(1)	Autoregressive order 1
AST	Aspartate aminotransferase
ATS	American Thoracic Society
AUC_{0-4}	Area Under the Curve From 0 to 4 Hours
BID	Twice daily
BDRM	Blinded Data Review Meeting
BPM	Beats Per Minute
BMI	Body mass index
CI	Confidence Interval
CKD-EPI	Chronic Kidney Disease Epidemiology Collaboration
CompEx	A composite endpoint comprising clinically relevant deteriorations and severe exacerbations in asthma
CRF	Case Report Form
CTCAE	Common Terminology Criteria for Adverse Events
DMC	Data monitoring committee
ECG	Electrocardiogram

eCRF	Electronic case report form
eDiary	Electronic Diary
e.g.	*Exempli Gratia*, For Example
F_{ENO}	Fractional nitric oxide concentration in exhaled breath
FEV_1	Forced expiratory volume in 1 second
FVC	Forced vital capacity
H_0	Null Hypothesis
H_1	Alternative Hypothesis
GINA	Global Initiative for Asthma
GP MDI	Glycopyrronium Inhalation Aerosol
hCG	Human chorionic gonadotropin
HFA	Hydrofluoroalkane
HLGT	High Level Group Term
HLT	High Level Term
ICF	Informed consent form
ICS	Inhaled corticosteroid
i.e.	Id est; that is
ITT	Intent-to-treat
IWRS	Interactive web response system
L	Liter
LABA	Long-acting β_2 agonist
LAMA	Long-acting muscarinic antagonist
MDI	Metered dose inhaler
MedDRA	Medical Dictionary for Regulatory Activities
MI	Multiple imputation
µg	Microgram
mITT	Modified intent-to-treat

mL	Milliliter
mm	Millimeter
mmHg	Millimeter of mercury
msec (ms)	Millisecond
NHANES	National Health and Nutrition Examination Survey
OTC	Over-the-counter
PCS	Potentially clinically significant
PEFR	Peak expiratory flow rate
PFT	Pulmonary function test
ppb	Parts per billion
PT	Preferred Term
PVC	Premature ventricular contraction
QTcF	QT corrected using Fridericia's formula
SABA	Short-acting β_2-agonist
SAE	Serious adverse event
SAP	Statistical analysis plan
SD	Standard deviation
SMQ	Standard MedDRA Query
sPDP	Statistical Protocol Deviation Plan
TEAE	Treatment-emergent adverse event
ULN	Upper limit of normal
UN	Unstructured
US	United States

Trademark Information

Trademarks Not Owned By Pearl

KoKo Spirometer®

SAS® Software

Spiriva® Respimat®

1. INTRODUCTION

This Statistical Analysis Plan (SAP) outlines the statistical methods for the display, summary and analysis of data to be performed at the end of Pearl Therapeutics, Inc. (Pearl) Study PT001102. The SAP should be read in conjunction with the study protocol. This version of the SAP has been developed using the PT001102-04 Amended Protocol (Version 5.0 dated 08 March 2019) and the PT001102 case report form (CRF) (Version 07 dated 05 April 2018).

2. STUDY OBJECTIVES AND ENDPOINTS

2.1 Study Objectives

The overall objective is to assess the efficacy and safety of 3 doses of GP MDI (glycopyrronium inhalation aerosol meter dose inhaler) compared to Placebo MDI and Spiriva Respimat over 24 weeks in subjects with persistent asthma.

2.1.1 Primary Objective

- To assess the effect of 3 doses of GP MDI compared to Placebo MDI and Spiriva Respimat on lung function over 24 weeks in subjects with persistent asthma

2.1.2 Secondary Objective

- To assess the effect of 3 doses of GP MDI compared to Placebo MDI and Spiriva Respimat on exacerbations, quality of life, and symptoms over 24 weeks in subjects with persistent asthma

2.1.3 Safety Objective

- To assess the safety of 3 doses of GP MDI compared to Placebo MDI and Spiriva Respimat over 24 weeks in subjects with persistent asthma

2.1.4 24-Hour Holter Monitoring Sub-Study Objective

- To assess the cardiovascular safety of GP MDI compared to Placebo MDI and Spiriva Respimat as evaluated by 24-hr Holter monitoring

2.2 Study Endpoints

2.2.1 Primary Efficacy Endpoints

- Change from baseline in forced expiratory volume in 1 second (FEV_1) area under the curve from 0 to 4 hours (AUC_{0-4}) at Week 24

2.2.2 Secondary Efficacy Endpoints

- Change from baseline in morning pre-dose trough FEV_1 at Week 24
- Rate of moderate/severe asthma exacerbations

- Change from baseline in Asthma Control Questionnaire (ACQ)-7 at Week 24
- Change from baseline in ACQ-5 at Week 24
- Change from baseline in Asthma Quality of Life Questionnaire for 12 years and older (AQLQ+12) at Week 24

2.2.3 Other Efficacy Endpoints

With the exception of rate of moderate/severe asthma exacerbation, all of the primary and secondary endpoints shown above will also be assessed over the Treatment Period and at each clinic visit as other efficacy endpoints.

In addition, the following endpoints will be assessed over the Treatment Period, by visit for measures assessed at clinic visits (with the exception of responder endpoints, which will be assessed at Week 24 and over the Treatment Period), and over each 4-week interval for measures assessed by electronic diary (eDiary) unless specified otherwise below.

- Percentage of responders in ACQ-5 (≥0.5 decrease equals response)
- Change from baseline in ACQ-6
- Percentage of responders in ACQ-6 (≥0.5 decrease equals response)
- Percentage of responders in ACQ-7 (≥0.5 decrease equals response)
- Percentage of responders in AQLQ +12 (≥0.5 change equals response)
- Time to first moderate/severe asthma exacerbation (assessed over the Treatment Period only)
- Time to first severe asthma exacerbation (assessed over the Treatment Period only)
- Rate of severe asthma exacerbations (assessed over the Treatment Period only)
- Time to first asthma exacerbation of any severity (assessed over the Treatment Period only)
- Rate of asthma exacerbations of any severity (assessed over the Treatment Period only)
- Peak FEV_1
- Percentage of symptom-free days (24-hour period without symptoms)
- Percentage of rescue-free days (24-hour period without rescue medication use)
- Change from baseline in the mean number of puffs of rescue medication use (puffs/day)
- Percentage of asthma control days (24-hour period without rescue medication use or symptoms)
- Change from baseline in morning peak expiratory flow rate (PEFR)
- Change from baseline in evening PEFR
- Change from baseline in fractional nitric oxide concentration in exhaled breath (F_{ENO})
- Percentage of clinically meaningful changes from baseline in F_{ENO}
- Time to first CompEx event

2.2.4 Safety Endpoints

The safety endpoints for this study include:

- Adverse events (AEs)
- Clinical laboratory values
- Vital signs
- 12-lead electrocardiograms (ECGs)

2.2.5 24-Hour Holter Monitoring Sub-Study Endpoints

Primary Holter Monitoring Sub-Study Endpoint

- Change from baseline in mean heart rate (HR) over 24 hours

Secondary Holter Monitoring Sub-Study Endpoints

- Change from baseline in mean nighttime (22:00 to 06:00) and daytime (06:00 to 22:00) HR
- Change from baseline in the maximum 24-hour HR
- Change from baseline in the minimum 24-hour HR
- Change from baseline in the frequency of isolated ventricular events (premature ventricular contractions [PVCs])
- Change from baseline in the frequency of ventricular couplets (defined as 2 PVCs preceded or followed by regular beats)
- Change from baseline in the frequency of ventricular runs (defined as 3 or more PVCs preceded or followed by regular beats)
- Change from baseline in the frequency of supraventricular ectopic beats
- Change from baseline in the frequency of isolated supraventricular events
- Change from baseline in the frequency of supraventricular couplets
- Change from baseline in the frequency of supraventricular runs
- Incidence of atrial fibrillation with rapid ventricular response (>100 beats per minute (bpm))
- Proportion of subjects with maximum HR >180, >160 to 180, >140 to 160, >120 to 140, >100 to 120, and 100 bpm or less
- Proportion of subjects with minimum HR >60, >50 to 60, >40 to 50, and <40 bpm
- Proportion of subjects in each category of change from baseline in the number of PVCs per hour (no change; increase of >0 to <60, 60 to <120, and ≥120; and decrease of >0 to <60, 60 to <120, and ≥120)

3. STUDY DESIGN AND ANALYTICAL CONSIDERATIONS

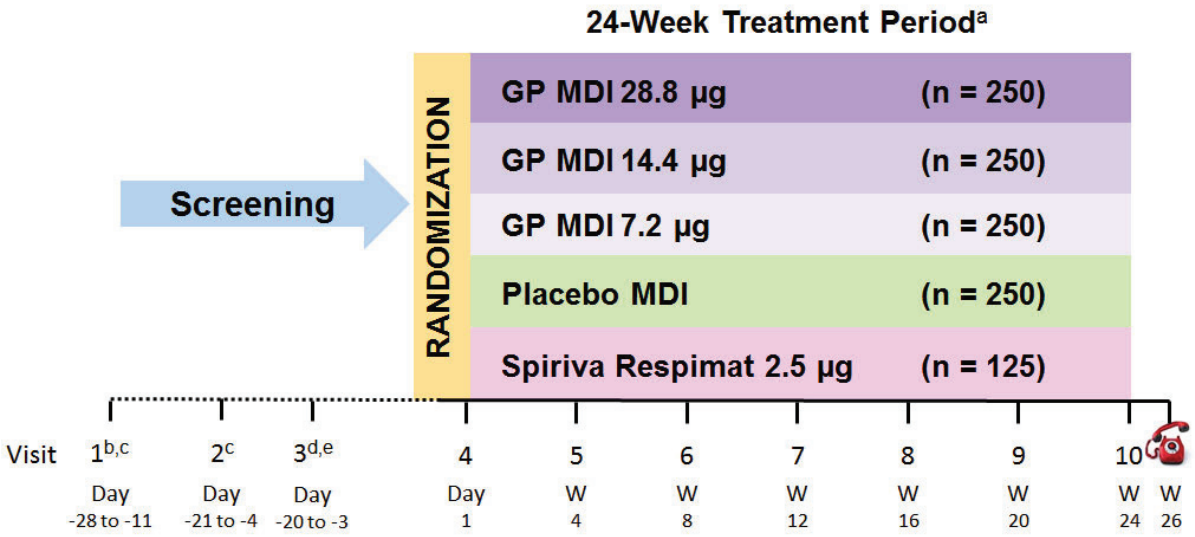
3.1 Study Design

3.1.1 Overall Study Design and Plan

This is a Phase II/III randomized, double-blind, parallel group, multi-center dose confirmation study comparing 3 doses of GP MDI (28.8, 14.4, and 7.2 µg) to Placebo MDI and open-label Spiriva Respimat (2.5 µg) in subjects with persistent asthma who are symptomatic despite treatment with an ICS/LABA or ICS/LABA + tiotropium. The 3 doses of GP MDI and the Placebo MDI will be administered twice daily (BID) and Spiriva Respimat will be administered once daily (QD), each for 24 weeks. The subjects will also continue on their own ICS/LABA regimen during the Screening Period and throughout the Treatment Period.

Figure 1 displays the overall study design.

Figure 1 Study Design Overview



Abbreviations: W=week.
[a] All GP MDI and Placebo MDI treatments are double-blind; Spiriva Respimat is open-label; all subjects will be on a fixed-dose combination ICS/LABA during screening and throughout the study
[b] For subjects on tiotropium, discontinue tiotropium and dispense ipratropium
[c] Subject must demonstrate reversibility to albuterol at either Visit 1 or Visit 2 for study eligibility
[d] Reversibility testing to ipratropium for characterization and stratification only
[e] Holter to be placed at Visits 3 and 7 and removed the following day

This study will be conducted at approximately 250 sites, contributing approximately 4 to 6 subjects per site. It is planned that approximately 2250 prospective subjects will be screened, with approximately 1125 subjects randomized to provide approximately 900 subjects who complete the study.

At Visit 1, subjects (or the parents or legal guardians of subjects <18 years of age) will sign an informed consent form (ICF); subjects <18 years of age will sign an assent form. All eligible subjects must be taking a stable regimen of an inhaled corticosteroid (ICS)/long-acting β2-agonist (LABA) for at least 4 weeks; this ICS/LABA regimen will be continued throughout the Screening and randomized Treatment Periods. Subjects who are also taking tiotropium will discontinue its use at Visit 1 and replace tiotropium with Sponsor-provided ipratropium bromide HFA MDI (hereafter referred to as ipratropium) taken as 2 inhalations 4 times daily (QID) during the Screening Period. Beginning with Visit 1, subjects will use an eDiary BID to record ICS/LABA use (Screening Period only), albuterol sulfate HFA MDI (hereafter referred to as albuterol) use, asthma symptom scores, PEFR, and study drug use (Treatment Period only).

Subjects must demonstrate an FEV_1 >40% and <85% of predicted (>40% and <90% of predicted for subjects 12 to <18 years of age) at Visit 1, Visit 2 (if applicable) and Visit 4. Reversibility to albuterol must be demonstrated at either Visit 1 or Visit 2. A subject will be considered reversible to albuterol if the improvement in FEV_1 from pre-albuterol to post-albuterol is ≥12% and ≥200 mL.

Subjects will return for Visit 2 after adequate washout of prohibited asthma medications. There will be a minimum of 7 days between Visit 1 and Visit 2. At Visit 2, the ACQ will be administered along with multiple other assessments including F_{ENO}. Spirometry will only be performed at Visit 2 if reversibility to albuterol was not demonstrated at Visit 1.

A subset of approximately 562 subjects (125 subjects from the GP MDI groups and Placebo MDI group and 62 subjects from the Spiriva Respimat group) will participate in the 24-hour Holter monitor sub-study. In these subjects, a 24-hour Holter monitor will be collected between Visit 3 and Visit 4 (Holter Monitor Baseline) and at Visit 7 (Week 12). Those subjects who have clinically significant abnormal findings defined as (but not limited to) criteria in protocol Section 5.5**Error! Reference source not found.** during the baseline 24-hour Holter monitor MUST be screen failed from the full study. Subjects unable to provide a minimum of 18 hours of acceptable quality recording in a 24-hour period after 2 attempts will not be eligible for the Holter monitoring sub-study, but DO NOT need to be screen failed from the full study.

Reversibility to ipratropium will be assessed at Visit 3 for characterization and randomization stratification purposes.

Subjects meeting randomization criteria at Visit 4 will discontinue ipratropium (if applicable), and be randomized in a 2:2:2:2:1 scheme to 1 of the following 5 treatment groups for 24 weeks:

- GP MDI 28.8 µg BID (n=250)
- GP MDI 14.4 µg BID (n=250)
- GP MDI 7.2 µg BID (n=250)
- Placebo MDI BID (n=250)
- Spiriva Respimat 2.5 µg QD (n=125)

Randomization will be stratified by baseline percent predicted FEV_1 (≤60% or >60%), reversibility to ipratropium (<12% or <200 mL vs ≥12% and ≥200 mL improvement in FEV_1), and by the ICS included in their fixed-dose ICS/LABA combination product (budesonide [e.g., Symbicort®, with formoterol fumarate dihydrate], mometasone furoate [e.g., Dulera®, with formoterol fumarate dihydrate], fluticasone furoate [e.g., Breo® Ellipta®, with vilanterol trifenatate], or fluticasone propionate [e.g., Advair® or AirDuo™ RespiClick®, with salmeterol]).

Following randomization, subjects will undergo 6 additional clinic visits over 24 weeks. For subjects participating in the Holter monitoring sub-study, a Holter monitor will also be placed at Visit 7 and removed the following day.

Note: The end of study is defined as the date on which data are collected 14 (+2) days after the last dose of study drug. Data collection can occur via a follow-up telephone call or a study visit, whichever provides 14 (+2) days of follow-up.

The Schedule of Events and Timed Assessments are available in the study protocol.

3.1.2 Prior, Concomitant, Post-Treatment, Prohibited Medications, and Other Restrictions

All prescription and over-the-counter (OTC) medications, as well as any herbal or vitamin supplements, taken by the subject within 30 days before Visit 1 (Screening) will be recorded on the prior/concomitant medications electronic CRF (eCRF). All concomitant medications taken during the study will be recorded on the Concomitant Medications eCRF page with indication, total daily dose, dose regimen, and route and dates of drug administration. Refer to the Protocol for information about prohibited medications.

3.2 Hypothesis Testing

For the primary comparisons, the general null hypothesis for each pair-wise comparison to a Placebo MDI dose will be that the mean treatment difference is zero (mean treatment effects are equal). The alternative two-sided hypothesis is that the mean treatment difference is not zero (mean treatment effects are not equal). All p-values will be reported as 2-sided.

The primary null (H_0) and alternative (H_1) hypotheses, with μ representing the mean, are presented below:

H0: $\mu_{GP\ xx} = \mu_{Placebo}$

H1: $\mu_{GP\ xx} \neq \mu_{Placebo}$

Secondary and other efficacy analyses will involve the above hypotheses applied to secondary and other efficacy endpoints. Other pairwise comparisons will be evaluated in a similar manner; comparison to Spiriva Respimat will focus on estimation.

3.3 Interim Analysis

No interim efficacy analyses are planned for this study.

The Data Monitoring Committee (DMC) will review safety data approximately every 6 months.

3.4 Sample Size

It is estimated that a sample size of 1125 subjects [redacted] The 2-sided alpha for each pairwise comparison is 0.05.

Comparisons of GP MDI to Spiriva Respimat will be focused on estimation as the study is not formally powered to demonstrate non-inferiority or superiority.

For the secondary endpoint of change from baseline in morning pre-dose trough FEV_1 at Week 24, it is assumed that the difference between GP MDI and Placebo MDI is 90 mL, with each treatment group having an SD of 317 mL at each visit. Under these assumptions, the study will have 80% power to declare GP MDI superior to Placebo MDI in the analysis of change from baseline in morning pre-dose trough FEV_1 at Week 24.

The Type I error rate will be strongly controlled for the primary analyses. The Type I error control strategy for the primary and secondary analyses are described in the Section 6.4.7. This sample size assumes that approximately 20% of randomized subjects will have discontinued study drug prior to Week 24.

4. DATA AND ANALYTICAL QUALITY ASSURANCE

The overall quality assurance procedures for the study data, statistical programming and analyses are described in Standard Operating Procedures (SOPs) of [redacted]. Detailed data management procedures are documented in the study Data Management Plan, Data Validation Check Specifications, and Integrated Safety Data Review Plan. Detailed statistical and programming quality control and quality assurance procedures are documented in the Statistical Analysis and Programming QC/QA Plan.

Transfer of PFT, eDiary, ECG and Holter monitor data from the central laboratory [redacted] to [redacted] will be defined in the [redacted] DMP (Data Management Plan), and data handling rules related to this data are included in Schwartz GJ, Brion LP, Spitzer A. The use of plasma creatinine concentration for estimating glomerular filtration rate in infants, children, and adolescents. Pediatr Clin North Am. 1987;34(3):571-590.

Appendix 1 of this SAP. The quality of all PFT's obtained at each time point will be graded independently at [redacted] by qualified personnel. Quality grading assessments will be based on American Thoracic Society (ATS)/ERS criteria and will be included in data transfers.

5. ANALYSIS SETS AND ESTIMANDS

5.1 Analysis Sets

5.1.1 Intent-to-Treat (ITT) Analysis Set

The Intent-To-Treat analysis set is defined as the data from all subjects who are randomized to treatment and use any amount of the study treatment. Subjects will be analyzed according to the treatment they were assigned at randomization. Data obtained after discontinuation of treatment, but prior to withdrawal from the study, will be included in the ITT analysis set.

5.1.2 Modified Intent-to-Treat (mITT) Analysis Set

The Modified Intent-to-Treat analysis set is a subset of the ITT analysis set, and is defined as the data from all subjects with post-randomization data obtained prior to discontinuation from treatment. Any data collected after completion of or discontinuation from randomized study medication will be excluded from the mITT analysis set but will still be included in the ITT analysis set. Subjects will be analyzed according to randomized treatment group. It is noted that a subject who used a study treatment, but took less than 1 full dose of treatment will qualify for this analysis set. The mITT analysis set will be the primary analysis set for all efficacy analyses. Note: The knowledge that a subject did not have an exacerbation constitutes an efficacy assessment.

5.1.3 Safety Analysis Set

The Safety analysis set is defined as data from all subjects who are randomized to treatment and receive any amount of the study treatment. However, subjects will be analyzed according to treatment received rather than randomized. If a subject receives more than one randomized treatment, they will be analyzed and included in summaries according to the treatment they received the most. Subjects receiving no study treatment will be excluded, as will subjects who have no post-dose safety assessments. Note: the statement that a subject has no AEs also constitutes a safety assessment.

5.1.4 Holter Monitoring Analysis Set

The Holter Monitoring analysis set is defined as all subjects in the Safety analysis set who had at least 18 hours of acceptable quality Holter monitoring data at both Visit 3 (Holter Baseline) and Visit 7 (Week 12). Data judged to be impacted by major protocol deviations will be determined prior to database lock and excluded per the Statistical Protocol Deviation Plan (sPDP).

5.2 Analysis Sets for Primary and Sensitivity Analyses

Demographics will be summarized for the mITT, Safety, and Holter Monitoring analysis sets.

Extent of exposure will be summarized for the Safety analysis set. The Safety analysis set will be used to summarize safety endpoints.

Holter monitoring data will be summarized for the Holter Monitoring analysis set.

Efficacy analyses will be performed for the ITT and mITT analysis sets. The mITT analysis set will be considered the primary analysis set for the efficacy analyses, with the ITT analysis set being considered supportive. Rescue medication endpoints will be analyzed with mITT and ITT analysis sets.

5.3 Estimands

5.3.1 Efficacy Estimand

The Efficacy Estimand, the primary estimand of interest, is defined as the effect of the randomized treatments in all subjects assuming continuation of randomized treatments for the duration of the study, regardless of actual compliance.

Analysis of the Efficacy Estimand will be conducted using the mITT analysis set where only data obtained prior to subjects discontinuing from randomized treatment will be utilized. This assumes that efficacy observed on treatment is reflective of what would have occurred after discontinuation of randomized treatment had they remained on treatment.

5.3.2 Attributable Estimand

The Attributable Estimand, the second estimand of interest, is defined as the effect of treatment in subjects attributable to the randomized treatment. For this estimand, discontinuation of randomized treatment for reasons of tolerability or lack of efficacy is considered a bad outcome.

Analysis of the Attributable Estimand will be conducted in the mITT analysis set, but data that are missing due to treatment discontinuation will be imputed based on the 95^{th} or the 5^{th} percentile of the reference arm's distribution if the reason is reasonably attributable to tolerability or lack of efficacy. The 95^{th} percentile would apply to an endpoint for which a higher value is a worse outcome, while the 5^{th} percentile would apply to an endpoint for which a higher value is a better outcome. For this purpose, the Placebo MDI treatment arm is considered the reference arm. Other missing data are to be imputed using the observed data model, i.e. assumed to be missing at random (MAR). The number of imputations used for the derivation of the attributable estimand will be between 100 and 1000. More detail about the computation of the attributable estimand will be provided in Appendix 6 of this SAP.

Treatment discontinuations reasonably attributable to tolerability or lack of efficacy will be identified during the blinded data review meeting (BDRM) and documented in the BDRM minutes prior to unblinding. Discontinuations will be attributed to tolerability if the subject had an adverse event determined by the investigator to be related to study drug, and for which study drug was permanently discontinued. Discontinuations will be attributed to lack of efficacy if 'lack of efficacy' is indicated to be the primary reason for discontinuation from study drug. For

the remaining discontinuation categories, where specific reasons or criteria frequently need to be considered, decisions will be made and documented at the BDRM.

5.3.3 Treatment Policy Estimand

The Treatment Policy Estimand, the third estimand of interest, is defined as the effect of randomized treatment over the study period regardless of whether randomized treatment is continued.

Analysis of the Treatment Policy Estimand will be conducted in the ITT analysis set, in which all observed data will be utilized regardless of whether subjects remain on randomized treatment.

6. STATISTICAL ANALYSIS

Analyses will be performed when the final database is available. All data collected contributing to the analysis will be provided in listings. Data for all subjects who are randomized will be included in the subject data listings. Data for non-randomized subjects will be listed where available.

All safety and efficacy parameters will be summarized by treatment unless specified otherwise.

Continuous variables will be summarized with descriptive statistics (the number of non-missing values, mean, standard deviation, median, minimum, and maximum). Additionally, the 25th and 75th percentiles will be presented when appropriate based on historical knowledge of the normality or non-normality distribution of underlying data.

Categorical variables will be summarized with frequency counts and percentages (where appropriate).

6.1 Data Handling Rules and Definitions, Including Handling of Missing Data

Missing data will be maintained as missing in the analysis datasets, unless specified otherwise. For variables where missing data are imputed, the analysis dataset will contain a new variable with the imputed value and the original variable value will be maintained as missing.

Data Imputation for Adverse Events Summaries by Severity and Relationship to Study Drug

For the AE summaries by severity (mild, moderate, or severe), an AE with missing severity will be deemed as severe. For the AE summaries by relationship to study drug, an AE with a missing relationship to study drug will be deemed as related. Imputed values will not be listed in data listings.

Data Imputation for Laboratory, Vital Sign, and ECG Summaries (Continuous Parameters)

Data from unscheduled visits will not be used for by-visit summaries. Data from both scheduled and unscheduled visits will be used for shift tables and for determining incidence of clinically significant values.

Data Imputation (All Laboratory Summaries)

Laboratory values of '>=x' or '<=x' will be taken as the value of x in the analyses. If a laboratory value is prefixed with '>': the available original value +0.001 will be used for table summaries; if a laboratory value is prefixed with '<', then the original value –0.001 will be used in table summaries.

Study Dates and Day of Assessment or Event

Study Day and Day of Assessment or Event definitions are provided in Appendix 1, Data Handling Rules.

On-treatment Asthma Exacerbations

An exacerbation will be considered "on-treatment" if its start date is before or on the last treatment date. For treatment discontinuations, this definition is extended to include exacerbations starting one day after the last treatment date. If it is decided during a clinic visit to discontinue study drug and to switch to a treatment for the ongoing exacerbation symptoms, the subject typically would not take the morning dose of study drug at that visit, and their exacerbation start date will be one day after the last treatment date. Such exacerbations will still be considered "on-treatment".

6.2 Subject Disposition and Analysis Sets

Disposition for all randomized subjects will be tabulated *(Table 1.1.1)* and listed *(Listing 1.2).* The tabulation will include the number of subjects in each randomized treatment group who received the study treatment, who were not treated, who completed treatment through Week 24, who discontinued treatment prematurely, who withdrew from the study prematurely, and who completed the study. The number and percentage of randomized subjects included in the mITT, ITT, Safety, and Holter Monitoring analysis sets will also be tabulated *(Table 1.1.1)*. Informed consent/assent (yes or no) and study day of informed consent/assent is listed in *Listing 2.1*.

The number of subjects randomized and in the analysis sets will be provided by center, and treatment in *Table 1.1.2*. The number of subjects randomized by stratification factor will be tabulated (*Table 1.1.4*), including baseline percent predicted FEV_1 (≤60% or >60%), reversibility to ipratropium (<12% or <200 mL vs ≥12% and ≥200 mL improvement in FEV_1), and by the ICS included in their fixed-dose ICS/LABA combination product (budesonide, mometasone furoate, fluticasone furoate, or fluticasone propionate). The stratification factors will be based on the actual data rather than on IWRS values. If there are any subjects who took study treatment

other than what was randomized during the study, both the treatment assigned at randomization and actual treatment(s) received during the Treatment Period will be listed (*Listing 1.3*). The duration of actual treatment will also be listed (*Listing 1.3*). A list of subjects with discrepant IWRS-based and actual stratification factors will also be provided (*Listing 1.6*).

A summary of reasons subjects were not randomized will be provided for all subjects not randomized (*Table 1.1.3*). A listing of reasons subjects were not randomized will also be provided *(Listing 1.4)*. Subjects excluded from the ITT, mITT, Safety, and Holter Monitoring analysis sets will be listed (*Listing 1.5)* for all subjects randomized. Reasons for premature discontinuation from study treatment will be summarized for the Safety Analysis Set *(Table 1.2.1).* Similarly, reasons for subjects' withdrawal from the study will be summarized for the ITT Analysis Set (*Table 1.2.2*).

Time to discontinuation of treatment and withdrawal from the study will be presented graphically by means of the Kaplan-Meier plots (*Tables and Figures 1.2.3 and 1.2.4*).

The attributable status for premature discontinuation from the study will be tabulated by treatment for the mITT Analysis Set (*Table 1.2.5*).

A Per Protocol population is not defined in this study. However, important protocol deviations will be identified according to the sPDP and blinded data review meeting (BDRM) process. These deviations will be summarized (*Table 1.3.1*) and listed (*Listing 1.7*).

A listing of subjects who did not comply with restrictions on illicit drugs or drugs of abuse, use of rescue medication, and xanthine-containing products will be provided in *Listing 6.1.1*. Use of rescue medication at pre-dose or during the post-dose assessments on each specific test day (yes/no), will be tabulated in *Listing 6.1.3*. In addition, the eligibility information (inclusion/exclusion criteria with any waivers granted) of all subjects who are randomized will be listed *(Listing 2.1).*

6.3 Demographic and Baseline Characteristics and Extent of Exposure

The definitions for the derived demographic or baseline characteristic variables can be found in Appendix 1.

6.3.1 Demography, Physical Characteristics

Subject demographics, use of inhaled corticosteroids (ICSs)/long-acting β2-agonists (LABAs) and the use of tiotropium (at least 4 weeks prior) at Screening, and smoking status/history will be summarized for the mITT, Safety and Holter Monitoring Analysis Sets and for Non-Randomized subjects *(Tables 1.4.1 through 1.4.4, respectively, and Listing 1.2)*. The ITT analysis set does not need to be tabulated because it is the same as the mITT analysis set for demographics and baseline characteristics. If the Safety Analysis Set has the same treatment assignment as the mITT, then these summaries will be identical as well and hence not produced.

Demographic and baseline characteristic variables summarized will include the following:

- Age
- Age Group (12 to <18, and 18 to 80 years)
- Age of onset of Asthma
- Gender
- Race
- Ethnicity (Hispanic or Non-Hispanic)
- PEFR stability limit at Visit 4
- ICS/LABA + tiotropium use (at least 4 weeks prior) at Screening (by the ICS type and use of tiotropium, Yes vs. No)
- GINA Classification (Steps 3 and 4/5, per GINA)
- ACQ-5 at Visit 2 (continuous and <1.5, ≥1.5, ≥2, ≥3, ≥4, ≥5)
- Baseline eosinophil count, calculated as the average of Visit 1 and Visit 4 values (<300 cells per mm^3 vs. ≥300 cells per mm^3)
- Number of severe asthma exacerbations in the 12 months prior to Visit 1 (Group as 0, 1 and ≥2)
- Baseline FE_{NO} (continuous and <40 ppb vs. ≥40 ppb)
- T-helper Cell Type 2 Status (Low: FE_{NO} < 40 ppb AND Baseline eosinophil count < 300 cells per mm^3; High (1): FE_{NO} ≥ 40 ppb OR Baseline eosinophil count ≥ 300 cells per mm^3; High (2): FE_{NO} ≥ 40 ppb AND Baseline eosinophil count ≥ 300 cells per mm^3)
- Smoking status (former smoker vs. non-smoker)
- Number of years smoked
- Average number of cigarettes smoked per day
- Number of pack years smoked, calculated as (number of cigarettes per day/20) x number of years smoked
- Weight
- Height
- Body mass index (BMI)

For this study, GINA classification will be determined by the total daily dose of ICS used at study entry and use of a concurrent LABA and/or tiotropium. Table 5-1 of the protocol provides the permissible ranges of ICS total daily doses. It is anticipated that all subjects will fall into GINA 3, 4, or 5. The table below presents how subjects will be categorized into GINA classifications. To account for subjects randomized in error, the table includes ICS doses that fall outside what is permissible in this study and definitions for GINA 1 and 2.

Table 1 GINA Classification of Asthma Treatment

Drug at Study Entry	GINA Classification			
	1	2	3	4 and 5
Controller Regimen	None	Low dose ICS	Low dose ICS/LABA	Med/high dose ICS/LABA Tiotropium
ICS (µg/day)				
Budesonide	0	200-400	200-400	>400
Fluticasone Furoate	0	100-199	100-199	≥200
Fluticasone Propionate	0	100-250	100-250	>250
Mometasone Furoate	0	110-220	110-220	>220
LABA	No	No	Yes	Yes
Tiotropium*	No	No	No	Yes or No

* Tiotropium is the only LAMA permitted in the study. Doses for MDIs are ex-valve doses.

Unusual combination therapies not included in the above table will be adjudicated on a case-by-case basis prior to database lock.

6.3.2 Asthma History, Screening/Baseline Spirometry, and Reversibility

Duration of asthma, the number of years prior to the start of study medication that asthma was first diagnosed (calculated as [Date of First Dose of Study treatment in the study – Date Asthma First Diagnosed] /365.25), will be summarized by treatment and all subjects for the mITT and Safety Analysis Sets and listed *(Tables 1.5.1, 1.5.2* and *Listing 4.1)*. A summary for the Safety Analysis Set will only be performed if the Safety Analysis Set is different from the mITT/ITT Analysis Set. Asthma control (GINA classifications) (GINA2018) will also be included in these summaries. History of severe asthma exacerbations will be listed (*Listing 4.2*).

Descriptive statistics will be provided for Screening period pre-bronchodilator and post-bronchodilator and baseline spirometry parameters *(Table 1.6.1 and Listing 2.2)* for the mITT analysis set. Spirometry data will be listed (*Listing 6.2).*

Characterization of Reversibility:

Reversibility to albuterol (Ventolin HFA) will be evaluated at either Visit 1 or Visit 2. Reversibility to ipratropium (Atrovent HFA) will be evaluated at Visit 3 and used as a stratification variable at randomization to ensure an even distribution of reversibility across the treatment arms.

Reversibility will be a comparison of the average of all best FEV_1 efforts obtained at -60 min and -30 min pre-bronchodilator to the best FEV_1 effort obtained at 30 minutes post-bronchodilator. However, if the subject fails the reversibility criteria at 30 minutes post-dose, any other available post-dose timepoint will be used instead to determine if reversibility was achieved. A subject is considered reversible to albuterol if the improvement in FEV_1 at 30 minutes (or up to 60 minutes) is ≥ 12% and ≥ 200 mL. A subject is considered reversible to ipratropium if the improvement in FEV_1 at 30 minutes is ≥ 12% and ≥ 200 mL.

Reversibility to albuterol at Screening Visits 1 or 2, with the latest value collected, will be summarized for the mITT Analysis Set and listed *(Table 1.7.1, Listing 2.2* and *Listing 5.2* for Ventolin HFA dispensing). Reversibility to ipratropium at Screening Visit 3 based on actual data will be summarized for the mITT Analysis Set and listed *(Table 1.7.2, Listing 2.2* and *Listing 5.2* for Atrovent HFA dispensing). The number and percentage of subjects reversible will be included in these summaries. Also included will be a summary of the change in FEV_1 from pre-dose FEV_1 to post-bronchodilator assessment. If multiple time points are available post-bronchodilator, then the one with the highest FEV_1 will be used.

Additionally, the number and percentage of subjects meeting each of the following response criteria will be summarized for Ventolin HFA and Atrovent HFA bronchodilators:

- ≥12% improvement post-bronchodilator in FEV_1 from pre-bronchodilator
- ≥15% improvement post-bronchodilator in FEV_1 from pre-bronchodilator
- ≥200 mL improvement post-bronchodilator in FEV_1 from pre-bronchodilator

6.3.3 Asthma Exacerbation, Medical and Surgical History at Screening, Reproductive Status and Pregnancy Testing

Severe asthma exacerbation history in the 12 months prior to Screening will be summarized for the Safety Analysis Set and listed for all randomization subjects *(Table 1.8.1 and Listing 4.2)*.

Medical and Surgical History at Screening will be summarized for the Safety Analysis Set and listed for all randomized subjects *(Table 1.8.2 and Listing 4.3)*.

Reproductive Status and Pregnancy Testing Results will be listed *(Listing 4.4)*.

6.3.4 Prior, Concomitant, and Post-Treatment Medications/Treatments

All prescription and OTC medications taken by the subject during 30 days before Screening and all concomitant therapy taken by the subject while on study will be recorded on the Prior and Concomitant Medications case report form (CRF) page.

Coding: Verbatim medication/treatment terms will be coded by [redacted] and will be assigned a preferred term and an ATC (anatomic therapeutic class) term using the latest version of the World Health Organization Drug Dictionary (WHO-DD) available (version: 1Q2018 or later).

Multiple ATC assignments: If there are multiple ATC codes assigned to the same concomitant medication, the "primary" one based on a Pearl medical evaluation will be used.

Prior medication/treatment is any medication/treatment taken prior to study treatment, even if this medication continued to be taken on the day of the start of study treatment in the study or afterward *(Appendix 1)*.

Concomitant medication/treatment is any medication/treatment reported as being taken any time between the date of first study treatment and the date that is one day before the date of discontinuation from or completion of study medication. Any medication that was used at any time on or after the day of treatment completion or treatment discontinuation will be considered a **Post-Treatment medication/treatment**.

Any medication/treatment which cannot be identified as Prior, Concomitant, or Post-Treatment will be considered as being in each of the categories that are possible from the available information.

Concomitant asthma, asthma-exacerbation, and non-asthma related medications/treatments will be summarized by preferred term and actual treatment received for the Safety Analysis Set (*Tables 1.10.1 to 1.10.3*). Asthma-related concomitant medication summaries will not include the asthma-exacerbation medications. Prior, concomitant/post-treatment asthma, asthma-exacerbation, and non-asthma medications will be displayed in separate listings (*Listings 4.5 to 4.7,* respectively).

Reported prior medications for asthma, asthma-exacerbation, and non-asthma-related medications will be tabulated for the Safety Analysis Set (*Tables 1.9.1, 1.9.3 to 1.9.4*) and listed separately *(Listings 4.5* to *4.7*, respectively*)*.

Prior asthma medications will be tabulated (for the Safety Analysis Set) for subjects having received any one, two, all three, or none of the following treatments (whether in fixed combination products or separately): (1) a muscarinic antagonist, (2) a β2 agonist, and (3) an inhaled corticosteroid *(Table 1.9.2)*. For this purpose, scheduled SAMA (Short-acting muscarinic antagonist) or SABA treatments are included. In addition, tabulations for long-acting muscarinic antagonists (LAMA) and long-acting β2 agonists (LABA) will also be included.

Reported post-treatment medications for asthma, asthma-exacerbation, and non-asthma-related medications will be tabulated for the Safety Analysis Set (*Tables 1.10.4, 1.10.6-1.10.7*) and listed separately (*Listings 4.5-4.7, respectively*). Post-treatment medications will be tabulated for subjects having received any one, two, all three, or none of the following treatments: (1) a muscarinic antagonist, (2) a β2 agonist, and (3) an ICS *(Table 1.10.5)*.

6.3.5 Extent of Exposure to Study Medication and Compliance

Subject's exposure to a study treatment will be determined by the duration of time (days) for which the doses were administered, defined as "([End date of treatment – Date of first dose of

treatment] + 1)". Percent compliance is defined as (total number of puffs of study treatment taken on a study day/total expected puffs taken on a study day) averaged across all days of a subject's dosing between start of study treatment and last day on study treatment) x 100.

For the blinded MDI treatments, the expected number of puffs for a test day which is the last date of treatment will be 2, and the expected number of puffs for the last date of treatment which is not a test day will be 4 when a PM dose is taken but will be 2 otherwise; the expected number of puffs on dates prior to the last date of treatment will be 4. For Spiriva Respimat, the expected number of puffs of study medication will be 2.

The number of days of exposure to study treatment will be summarized for each treatment for the Safety Analysis Set. The total person-years of exposure for a treatment group, defined as the total exposure in the study across all subjects in the treatment, will also be provided by treatment *(Table 1.11)*. In addition, treatment compliance will be provided in this summary. The treatment compliance will be categorized into 7 different groups depending on the degree of compliance: 0 – <20%, ≥20 – <40%, ≥40 – <60%, ≥60 – <80%, ≥80 – ≤100%, >100 – ≤120%, and >120%. Also provided in this summary will be descriptive statistics (n, mean, standard deviation, median, minimum and maximum) for percent compliance by treatment. Treatment compliance will be reported in *Listing 5.3*. A listing of treatment dosing and dispensing information will be provided in *Listing 5.1.* Any comments related to study medication or any other additional study comments will be listed *(Listing 9.4)*.

6.4 Efficacy Analyses

Three estimands of interest are defined for this study: efficacy estimand, attributable estimand, and treatment policy estimand. They have been introduced in Section 5.3.

There are three primary pairwise comparisons of treatments of interest, namely, GP MDI (3 doses: 28.8 µg, 14.4 µg and 7.2 µg) vs. Placebo MDI. Comparisons of GP MDI to Spiriva Respimat will be focused on estimation as the study is not formally powered to demonstrate non-inferiority or superiority. Estimation results will be provided by randomized treatment and for each treatment difference for all comparisons, in each estimand, and will be available in the post-text tables.

6.4.1 Baselines and Baseline Covariates for Analysis

The mean of all evaluable 60- and 30-minute pre-dose spirometry assessments conducted at Day 1 (Visit 4) will be used to establish baseline for FEV_1 $AUC_{0\text{-}4}$, trough and peak.

For the diary symptom score parameters and rescue medication usage, baseline will be the average of the non-missing values from the diary data collected in the last seven days of the Screening Period for morning and evening PEFR. The PEFR stability limit at Visit 4 will be calculated as the average of the available morning PEFR eDiary recordings during at least 4 of the last 7 days before Visit 4 (i.e., the baseline PEFR), multiplied by 0.8.

For severe asthma exacerbation history in the past 12 months before Visit 1, baseline is the number of severe asthma exacerbations reported on the History of Asthma Exacerbation CRF page. It will be grouped as 0 and ≥1 exacerbations for the purpose of modeling.

Identification of the fixed combinations of ICS/LABA taken within 4 weeks prior to Visit 1 will be based on a medical review of entries in the Prior and Concomitant Medications CRF page. The type of ICS in the fixed combination of ICS/LABA that the subject is taking prior to Visit 1 will be a covariate (budesonide, mometasone furoate, fluticasone furoate, or fluticasone propionate).

Baseline blood eosinophil count is the average of non-missing blood eosinophil count values at Visit 1 and Visit 4.

Baseline age is the age in years at the time of Informed Consent/Assent.

Baseline post-albuterol FEV_1 is the highest available value of FEV_1 obtained after dosing with albuterol (Ventolin HFA) at Visit 2 if reversibility to albuterol was performed at Visit 2; otherwise, it is the highest value of FEV_1 obtained after dosing with albuterol at Visit 1. Baseline percent reversibility to albuterol is 100 x (POST-PRE)/PRE, where PRE is the mean of the available 30 minute and 60 minute values of FEV_1 prior to dosing with albuterol at Visit 2 (or Visit 1), and POST is the post-albuterol FEV_1 value defined above.

Baseline post-ipratropium FEV_1 is the highest available value of FEV_1 obtained after dosing with ipratropium (Atrovent HFA) at Visit 3. Baseline percent reversibility to ipratropium is 100 x (POST-PRE)/PRE, where PRE is the mean of the available 30 minute and 60 minute values of FEV_1 prior to dosing with ipratropium at Visit 3, and POST is the post-ipratropium FEV_1 value defined above.

Baseline for the ACQ and AQLQ+12 questionnaire scores is the last available result at Visit 4. For ACQ, if there is no result at Visit 4, then the result at Visit 2 is used as the baseline.

Baseline F_{ENO} is defined taking into account the F_{ENO} testing procedures, which are outlined in protocol Appendix 11. Two valid, reproducible F_{ENO} measurements are required in order for the F_{ENO} test to be considered complete. The test is considered reproducible if ANY of these conditions is met:

- Either or both of 2 valid F_{ENO} measurements is < 30 ppb and the measurements are within 2 ppb of each other.
- Both measurements are > 30 ppb and within 10% of each other.
- Both values are below or above the measurement range (5-300 ppb).

If the reproducibility criteria are met within the first 2 exhalations, baseline F_{ENO} is defined the average of these 2 measurements. Otherwise, the participant has up to 2 additional exhalations to satisfy the reproducibility criteria (up to a total of 4 trials). Any pair of the completed attempts

can be assessed for reproducibility. If successful, the baseline is defined as the average of the last 2 reproducible measurements. For example, if attempts 1 and 2 do not meet reproducibility conditions, attempt 3 is performed. If pairs (1,3) and (2,3) both satisfy reproducibility, the pair (2,3) will be used to derive the baseline. If the reproducibility criteria are still not met for any pair of attempts after taking 4 measurements at Visit 4, then FE_{NO} measurements at Visit 2 will be considered, using the same rules. If no reproducible pair exists either at Visit 4 or Visit 2, baseline FE_{NO} will be missing.

6.4.2 Visits and Time Windows for Visit-Based Efficacy Assessments

Efficacy data obtained during unscheduled visits will not be used for any of the pre-defined efficacy analyses. Efficacy from scheduled and unscheduled visits will be listed.

As of protocol amendment v4.0, post-dose spirometry data are no longer collected at Week 4 (Visit 5). Any data available at Week 4 will be assigned to the corresponding analysis week (Week 4) in the analysis dataset.

Although no efficacy analysis based on time points were planned for spirometry, the change from baseline in PFT assessments will be allocated to derived nominal collection time windows using the time intervals specified below. This derivation is consistent with other Pearl studies. It will in particular ensure that no post-dose assessments are used in the calculation of trough, and no assessments later than 5 hours post-dose are used in the calculation of AUC_{0-4} and peak.

Table 2 Analysis Study Time Window for Spirometry Assessments

Calculated Study Time Window	Time Interval for the Study Time Window
Pre-dose 60 min.	≥45 minutes prior to dose
Pre-dose 30 min.	≥0 to <45 minutes prior to dose
Post-dose 15 min.	>0 to 22 min. post-dose
Post-dose 30 min.	23 to 44 min. post-dose
Post-dose 1 hr.	45 to 89 min. post-dose
Post-dose 2 hrs.	90 to 179 min. post-dose
Post-dose 4 hrs.	3 to <5 hrs. post-dose

Note: The minutes are rounded to the nearest whole number before applying time windows.

6.4.3 Asthma Exacerbations

The definitions of mild, moderate, and severe asthma exacerbations to be used in this study are shown below. These definitions are based on the ATS/ERS Statement "Asthma Control and Exacerbations, Standardizing Endpoints for Clinical Asthma Trials and Clinical Practice" (Reddel 2009).

Exacerbations will be considered **severe** if they result in:

- Use of systemic corticosteroids (tablets, suspension or injection) for at least 3 days or
- A hospitalization or ER visit because of asthma or
- Death due to asthma

Exacerbations will be considered **moderate** if they result in:

- A deterioration of asthma requiring a new or increased dose of ICS for at least 3 days.

Deterioration of asthma is defined by the occurrence of 1 or more of the following for ≥ 2 consecutive days:

  - PEFR: a decline of ≥ 20% from baseline
  - Rescue albuterol use: >4 puffs/day and ≥ 2 x baseline
  - Symptoms: nighttime score that is > baseline and ≥ 2 or a daytime score that is > baseline and ≥ 3

Exacerbations will be considered **mild** if they result in:

- A deterioration of asthma (as defined above)

  - Not requiring treatment or
  - Requiring treatment with a systemic corticosteroid or a new or increased dose of ICS for < 3 days

The conditions for moderate and severe asthma exacerbations will be captured in the eCRF. Mild exacerbations will be captured programmatically using data from the eDiary (*Listing 6.1.4*).

Additionally, the investigator may identify certain events (recorded on the same CRF page) which don't entirely meet the criteria above as exacerbations (for example, in the judgment of the investigator, severe symptoms lasting less than 2 days but requiring systemic steroids); the justifications supporting the investigator's judgment will be recorded on the same eCRF page.

For a **severe** asthma exacerbation, the start date is defined as the start date of prescribed treatment with a systemic corticosteroid or the date of ER visit or hospitalization for asthma or the date of death due to asthma. The stop date is defined as the last day on which systemic corticosteroids were used, or the discharge date from the ER or hospital for asthma, whichever is later.

For a **moderate** asthma exacerbation, the start date is defined as the first day of the increased dose of ICS or new ICS treatment (for example, adding inhaled budesonide). The end date is defined as the last day of this treatment. For moderate exacerbations that are longer than 14 days, a review will be done during the BDRM to confirm if this duration is justified. This review may result in a programmatic adjustment to the duration of the event.

For a **mild** asthma exacerbation, the start date is defined as the first day of 2 consecutive days of deterioration and the end date is defined as the first day the subject no longer meets the

deterioration criteria OR the last day of treatment, whichever is later. Treatment must be for less than 3 days to meet the criteria for a mild exacerbation.

There are three groups of asthma exacerbation severity analyzed in this study: severe, moderate/severe, and any severity.

For the severe and moderate/severe groups, in order to ensure that the same event is not counted twice, consecutive exacerbations with start and stop dates ≤7 days apart will be considered to be the same event of the highest severity. Any days of asthma deterioration where new or increased ICS or systemic CS were given, but which were not part of a moderate/severe event, will also be considered for this merger. **Note: This means that if a mild asthma exacerbation required treatment with a systemic corticosteroid or a new or increased dose of ICS for < 3 days, then those days under treatment will be merged into an exacerbation of worse severity if they occurred ≤7 days from another exacerbation. In the case where two mild events are within 7 days of each other, where ICS or systemic steroid treatment is required for each event, the merged event will be categorized as a moderate or severe event.** Asthma exacerbations will be considered separate events provided that there are more than 7 days between the recorded stop date of the earlier event and the start date of the later event.

Similarly, for asthma exacerbations of any severity, consecutive exacerbations with start and stop dates ≤7 days apart will be considered to be the same event.

For the efficacy estimand, the time at risk is defined for each severity group as time from the date of first dose until the last dosing date, minus the time during or within 7 days after all exacerbations from the given severity group. However, the start day of an asthma exacerbation will not be subtracted from the time at risk. Any days subsequent to the date of discontinuation from or completion of treatment will not be subtracted.

For the attributable estimand, the time at risk is defined as time from the date of first dose until the treatment completion date (for subjects who have completed treatment) or 24 weeks after the date of first dose (for subjects who have not completed treatment), minus the time during or within 7 days after all observed exacerbations from the given severity group. However, the start day of an asthma exacerbation will not be subtracted from the time at risk. For subjects who completed treatment, any days subsequent to the date of completion of treatment will not be subtracted. For subjects who discontinued treatment, all time between treatment discontinuation and 24 weeks after the date of first dose will be included into the time at risk.

For the treatment policy estimand, time at risk is defined as follow-up time up to the last recorded date (of any assessment or contact) for the subject (including telephone contact), minus the time during or within 7 days after all observed exacerbations from the given severity group. However, the start day of an asthma exacerbation will not be subtracted from the time at risk. Any days subsequent to the date of completion of study will not be subtracted.

It is noted that the time at risk will be different for each severity.

Analyses will be conducted on the efficacy estimand, on the attributable estimand, and on the treatment policy estimand. The primary analyses of the efficacy and treatment policy estimands will use only observed data. The attributable estimand will use the mITT Analysis Set but then impute missing data as described in Section 5.3.

6.4.4 Primary Efficacy Analysis

Analyses for the primary endpoint are presented in this section.

FEV_1 AUC_{0-4}

FEV_1 AUC_{0-4} at Week 24 is the primary efficacy endpoint. FEV_1 AUC_{0-4} at the other visits and over 24 weeks are "Other" endpoints.

FEV_1 AUC_{0-4} is the area under the curve for FEV_1 calculated using the trapezoidal rule, after the subtraction of the baseline FEV_1 value, and the AUC will be transformed into a weighted average by dividing by the time (in hours) from dosing to the last measurement included (typically 4 hours). For all estimands, only one non-missing post-dose value is required for the calculation of AUC. Actual time from dosing will be used if available; otherwise scheduled time will be used.

The computation of FEV_1 AUC_{0-4} after the discontinuation of study treatment will be based on the actual time of administration of asthma maintenance medication if known and applicable or otherwise using the nominal time points (e.g., 1 hr pre-dose, 30 min pre-dose, 15 min post-dose, 30 min post-dose, 1 hr post-dose, 2 hr post-dose, and 4 hr post-dose). The 1 hr pre-dose time point would be used as time zero for the computation of the AUC if the time of asthma maintenance medication is unknown.

The differences in change from baseline between treatment groups in FEV_1 AUC_{0-4} at Week 24, over 24 weeks, and at each post-randomization visit will be evaluated using a linear repeated measures analysis of covariance (ANCOVA) model. The model will include treatment, visit, background ICS/LABA, and treatment by visit interaction as categorical covariates and baseline FEV_1, percent reversibility to ipratropium, and logarithm of baseline blood eosinophil count as continuous covariates. An unstructured (UN) matrix will be used to model the variance-covariance structure. If the UN model fails to converge, then a first-order autoregressive (AR(1)) structure will be used instead. In the AR(1) model, subject will be included as a random effect.

Two-sided p-values and point estimates with two-sided 95% confidence intervals (CIs) will be produced for each treatment difference (*Tables and Figures 2.1.1 to 2.1.3* for the efficacy estimand, attributable estimand, and treatment policy estimand, respectively).

For the attributable estimand of FEV_1 AUC_{0-4}, data that are missing due to treatment discontinuation will be imputed as described in Section 5.3.2. Imputation for missing values classified as "attributable" will use mean changes from baseline based on the 5^{th} percentile of the

Placebo MDI arm's distribution. The analysis of this endpoint with the attributable estimand is considered secondary.

Individual data will be listed (*Listing 6.2).*

Assumptions Checks and Possible Removal of Outliers

In general, the distribution of spirometry measures is well-approximated by a normal distribution. Under some circumstances, atypical values can arise. Such values may disproportionately affect model-based estimates of the fixed effect and variance parameters. Prior to database lock and unblinding, the change from baseline values for efficacy endpoints will be examined as part of data quality management. This may include production of normal probability plots, kernel density estimates, and normal order outlier statistics. Based on this blinded evaluation, if atypical values are identified, nonparametric methods or data transformations (e.g. logarithmic or normal rank transformation) will be considered.

If erroneous values are detected, every effort will be made to correct them prior to database lock. If these values cannot be corrected, they will be considered for removal from analysis. These analyses will be conducted if warranted to demonstrate the robustness of the primary and secondary results and reported in the statistical methods appendix to the CSR.

The assumption of normality for FEV_1 $AUC_{0\text{-}4}$ data will be checked by visually inspecting the distribution of the residuals. Also, model fit and the assumption of homogeneity of variance among treatments will be verified by inspection of scatter plots of predicted vs. residuals, residuals vs. treatment, residuals vs. screening ICS/LABA use, and by box plots of residuals for model variables with a potential effect on variance (treatment, visit, and ICS/LABA use). Plots for scaled (marginal) residuals will be prepared (option=VCIRY on the model statement and ODS graphics option allows the production of plots using these residuals). If appropriate, the linear repeated measures model analysis may be conducted by allowing for heterogeneity of variance between treatments and ICS/LABA use categories. Note that the unstructured covariance structure allows for heterogeneity among the visits.

Further assumption checks are described in Appendix 6 to this SAP.

6.4.5 Secondary Efficacy Analysis

Similar to the primary efficacy analysis, the secondary efficacy analyses will be conducted for the efficacy estimand. Analyses will also be conducted for the attributable and treatment policy estimands.

Change from Baseline in Morning Pre-Dose Trough FEV_1

Change from baseline in morning pre-dose trough FEV_1 at Week 24 is a secondary endpoint. Change from baseline in morning pre-dose trough FEV_1 at the other visits and over 24 weeks are "Other" endpoints.

Change from baseline in morning pre-dose trough FEV_1 at each visit is defined as the average of the 60 and 30 minute pre-dose values minus baseline. Baseline is defined in Section 6.4.1. In subjects missing one of these pre-dose assessments, the value will be calculated from the single measurement. In subjects missing both pre-dose values, morning pre-dose trough FEV_1 at that visit will not be calculated. Spirometry data from unscheduled visits will not be used for this analysis.

The change from baseline in morning pre-dose trough FEV_1 will be analyzed using a linear repeated measures ANCOVA model. The model will include treatment, visit, background ICS/LABA, and treatment by visit interaction as categorical covariates and baseline FEV_1, percent reversibility to ipratropium, and logarithm of baseline blood eosinophil count as continuous covariates. An unstructured (UN) matrix will be used to model the variance-covariance structure. If the UN model fails to converge, then a first-order autoregressive (AR(1)) structure will be used instead. In the AR(1) model, subject will be included as a random effect. Contrasts will be used to obtain estimates of the treatment differences at Week 24, over 24 Weeks, and at each post-randomization visit.

Two-sided p-values and point estimates with two-sided 95% confidence intervals (CIs) will be produced for each treatment difference (*Tables and Figures 2.2.1 to 2.2.3* for the efficacy estimand, attributable estimand, and treatment policy estimand, respectively).

For the attributable estimand of change from baseline in morning pre-dose trough FEV_1, data that are missing due to treatment discontinuation will be imputed as described in Section 5.3.2. Imputation for missing values classified as "attributable" will use mean changes from baseline based on the 5^{th} percentile of the Placebo MDI arm's distribution.

Rate of Moderate/Severe Asthma Exacerbations

Only on-treatment exacerbations will be included for the analysis of rate of moderate/severe asthma exacerbations for the efficacy estimand (see Section 6.1). Exacerbations occurring after the premature discontinuation of treatment will be considered for the treatment policy estimand.

For the attributable estimand, missing data that have been reasonably attributed to tolerability or lack of efficacy will be imputed based on the 95^{th} percentile of the reference arm's (Placebo MDI) distribution. The imputed value will be drawn from a negative binomial distribution with mean exacerbation rate (and variance) based on the 95^{th} percentile of the reference arms' distribution, with estimates set to the estimate for Placebo MDI from the analysis of the exacerbation rates. Other missing data are to be imputed using the observed data model, i.e. assumed to be missing at random or missing completely at random. Further information about the computation of the attributable estimand is described in Appendix 6 to this SAP.

The rate of moderate/severe asthma exacerbations will be analyzed using negative binomial regression as implemented in SAS PROC GENMOD. Treatments will be compared adjusting for baseline post-albuterol FEV_1 percent predicted, percent reversibility to ipratropium and logarithm of baseline blood eosinophil count as continuous covariates and baseline severe asthma

exacerbation history, and background ICS/LABA as categorical covariates. Time at risk of experiencing an exacerbation will be used as an offset variable in the model.

Moderate/severe asthma exacerbation data will be listed (*Listing 6.1.3*). For exacerbations that were identified apart from an eDiary alert, the symptom information will be listed.

The number of moderate/severe exacerbations and the percentage of subjects who experience exacerbations, observed and adjusted exacerbation rates, and rate ratios comparing treatments will be summarized for the efficacy estimand (*Table 2.3.1*) and treatment policy estimand (*Table 2.3.3*); only adjusted rates and rate ratios will be presented for the attributable estimand (*Table 2.3.2).*

Follow-up time at risk for moderate/severe asthma exacerbations will be summarized for the efficacy estimand (*Table 2.3.1a*) and treatment policy estimand (*Table 2.3.3a*).

Change from Baseline in ACQ-7, ACQ-5, and AQLQ+12

Change from baseline in ACQ-7, ACQ-5 and AQLQ +12 scores at Week 24 are secondary endpoints. Change from baseline in ACQ-7, ACQ-5 and AQLQ +12 scores at the other visits and over 24 weeks are “Other” endpoints. ACQ-6 at all visits and over 24 weeks is an “Other” endpoint.

The Asthma Control Questionnaire (ACQ), comprising 7 questions, is completed in the clinic and requires subjects to recall their experiences during the previous week (7 days) prior to the study visit. The ACQ-5, -6, and -7 are each calculated from the ACQ with each using a different subset of questions (questions 1-5, 1-6, and 1-7, respectively). All three instruments have been validated. The ACQ-5 measures 5 symptoms (woken at night by symptoms, wake in the morning with symptoms, limitation of daily activities, shortness of breath, and wheeze); the ACQ-6 is these symptom measurements plus daily rescue medication use as recalled by the subject; and the ACQ-7 is the ACQ-6 plus airway caliber as measured by pre-bronchodilator FEV_1 percent predicted. Each question is scored on a 7-point scale (generally, 0=no impairment, 6=maximum impairment), the questions are equally weighted, and the ACQ score is the mean of the item responses and therefore ranges from 0 (totally controlled) to 6 (severely uncontrolled).

The Asthma Quality of Life Questionnaire (AQLQ +12) is a 32-item validated subject-administered questionnaire. 32 questions in the AQLQ +12 are in 4 domains (symptoms, activity limitation, emotional function, and environmental stimuli). All 4 domains contain standard specific questions relating to each domain. Subjects are asked to think about how they have been during the previous 2 weeks and to respond to each of the 32 questions on a 7-point scale (7=no impairment to 1=severe impairment). The overall AQLQ +12 score is the mean of all 32 responses, and the individual domain scores are the means of the items in those domains.

No imputation is planned for missing individual questionnaire items. Missing items will result in missing ACQ or AQLQ +12 score values for the visit.

Change from baseline in ACQ-7, ACQ-5, and AQLQ +12 will be analyzed using a linear repeated measures ANCOVA model. The model will include treatment, visit, background ICS/LABA, and treatment-by-visit interaction as categorical covariates and baseline post-albuterol FEV_1 percent predicted, baseline score for the patient-reported outcome instrument, percent reversibility to ipratropium, and logarithm of baseline blood eosinophil count as continuous covariates. A UN correlation matrix will be used to model additional autocorrelation within subject. If this model fails to converge, an AR(1) structure will be used instead. In the AR(1) model, subject will be included as a random effect.

Contrasts will be used to obtain estimates of the treatment differences at Week 24, over 24 Weeks, and at each post-randomization visit. Two-sided p-values and point estimates with 2-sided 95% confidence intervals will be produced for each treatment difference (*Tables and Figures 2.4.1-2.4.3 for ACQ-7, 2.5.1-2.5.3 for ACQ-5, 2.7.1-2.7.3 for AQLQ+12*) for the efficacy, attributable, and treatment policy estimands.

For the attributable estimand of change from baseline in ACQ and AQLQ +12 scores, data that are missing due to treatment discontinuation will be imputed as described in Section 5.3.2. Imputation for missing values classified as "attributable" will use mean changes from baseline based on the 95th percentile of the Placebo MDI arm's distribution for ACQ scores, and on the 5th percentile for AQLQ +12 score.

Individual responses will be listed (*Listing 6.1.5 for ACQ-5, ACQ-6 and ACQ-7, 6.1.6 for AQLQ+12, and 6.1.7 for AQLQ+12 Domain Scores*).

6.4.6 Other Efficacy Analysis

Other efficacy analyses will be conducted for the efficacy estimand only.

Percentage of Responders in ACQ-5, ACQ-6, ACQ-7, and AQLQ+12

The ACQ and AQLQ +12 instruments were introduced in Section 6.4.5.

Responder analyses will be performed for ACQ-5, ACQ-6, ACQ-7, and AQLQ +12 at Week 24, over 24 Weeks, and at each post-randomization visit. Responders are defined as subjects with an improvement of ≥0.5 points over baseline (decrease in ACQ scores, increase in AQLQ +12 score). Subjects who discontinue treatment for any reason will be classified as non-responders. Subjects who have a missing result at Week 24 will also be classified as non-responders for that endpoint at Week 24.

Logistic regression will be used to compare the treatment groups with treatment and background ICS/LABA as categorical covariates and baseline instrument score, logarithm of baseline blood eosinophil count, percent reversibility to ipratropium, and baseline post-albuterol FEV_1 percent predicted as continuous covariates. P-values and odds ratios with 95% confidence intervals will be produced for each treatment comparison (*Table 2.4.4 for ACQ-7, 2.5.4 for ACQ-5, 2.7.4 for AQLQ+12, and Table 2.6.2 for ACQ-6*).

Change from Baseline in ACQ-6

The ACQ instrument was introduced in Section 6.4.5, with ACQ-6 being one of the associated scores.

Change from baseline in ACQ-6 will be analyzed using a linear repeated measures ANCOVA model. The model will include treatment, visit, background ICS/LABA, and treatment-by-visit interaction as categorical covariates and baseline post-albuterol FEV_1 percent predicted, baseline score for the patient-reported outcome instrument, percent reversibility to ipratropium, and logarithm of baseline blood eosinophil count as continuous covariates. Contrasts will be used to obtain estimates of the treatment differences at Week 24, Over 24 weeks, and at each post-randomization visit. Two-sided p-values and point estimates with 2-sided 95% confidence intervals will be produced for each treatment difference (*Table and Figure 2.6.1*).

Rate of Asthma Exacerbations

The rates of severe asthma exacerbations and asthma exacerbations of any severity over the treatment period are "Other" endpoints. For details on the collection and operational definition of these events, and calculation of the time at risk, refer to Section 6.4.3.

The number and percentage of subjects with asthma exacerbations in each treatment group will be tabulated.

The rate of severe asthma exacerbations and asthma exacerbations of any severity will each be analyzed in the same way as the rate of moderate/severe asthma exacerbations. Treatments will be compared adjusting for baseline post-albuterol FEV_1 percent predicted, percent reversibility to ipratropium and logarithm of baseline blood eosinophil count as continuous covariates and baseline severe asthma exacerbation history, and background ICS/LABA as categorical covariates. Time at risk of experiencing an exacerbation will be used as an offset variable in the model.

The number of exacerbations and the percentage of subjects who experience exacerbations, exacerbation rates, and rate ratios comparing treatments will be summarized for the efficacy estimand for severe exacerbations (*Table 2.3.4*), and for exacerbations of any severity (*Table 2.3.5).*

Follow-up time at risk will be summarized and displayed graphically for the efficacy estimand for severe exacerbations (*Table 2.3.4a*), and for exacerbations of any severity (*Table 2.3.5a).*

Time to First Moderate/Severe Asthma Exacerbation, Severe Asthma Exacerbation, or Asthma Exacerbation of Any Severity

Time to first moderate/severe asthma exacerbation is the time from the first dose of study medication to the time of onset of the first moderate/severe asthma exacerbation. Time to first severe asthma exacerbation and time to first asthma exacerbation of any severity are defined similarly.

Only on-treatment exacerbations will be included for calculating the time to first asthma exacerbation for the efficacy estimand (see Section 6.1).

The time to first asthma exacerbation will be analyzed up through the Week 24 visit using a Cox regression model. Treatment comparisons will be performed using the model, adjusting for severe asthma exacerbation history and background ICS/LABA as categorical covariates, and percent reversibility to ipratropium, logarithm of baseline blood eosinophil count, and baseline post-albuterol FEV_1 percent predicted as continuous covariates. PROC PHREG will be used. Estimated adjusted hazard ratios relative to the comparator for each treatment comparison will be displayed along with the associated Wald two-sided 95% CIs and p-values (*Tables 2.3.6 to 2.3.8* for moderate/severe, severe, and any severity exacerbations).

Time to first asthma exacerbation will also be displayed graphically for each treatment using Kaplan-Meier curves (*Figures 2.3.6 to 2.3.8* for moderate/severe, severe, and any severity exacerbations). Subjects who complete the study (and the study treatment) and do not experience an asthma exacerbation over the treatment period will be censored at the Week 24 visit. Subjects who do not experience an asthma exacerbation and discontinue treatment early will be censored at the date of treatment discontinuation.

Peak Change from Baseline in FEV_1

The differences in peak change from baseline in FEV_1 within 4 hours post-dosing at Week 24, over 24 weeks, and at each post-randomization visit will be evaluated using a linear repeated measures analysis of covariance (ANCOVA) model. The model will include treatment, visit, background ICS/LABA, and treatment by visit interaction as categorical covariates and baseline FEV_1, percent reversibility to ipratropium, and logarithm of baseline blood eosinophil count as continuous covariates. An unstructured (UN) matrix will be used to model the variance-covariance structure. If the UN model fails to converge, then a first-order autoregressive (AR(1)) structure will be used instead. In the AR(1) model, subject will be included as a random effect.

Two-sided p-values and point estimates with two-sided 95% confidence intervals (CIs) will be produced for each treatment difference (*Table and Figure 2.8.1*) for the efficacy estimand.

Percentage of Symptom-Free Days, Rescue-Free Days, and Asthma Control Days

The percentages of symptom-free days, rescue-free days, and asthma control days over the treatment period are “Other” endpoints.

To calculate these percentages, the treatment period will be divided into 4-week intervals, as presented in Table 3. For each interval, the denominator will be the number of eDiary days in the interval, satisfying both conditions:

- The day is before or on the day of completion or discontinuation of study treatment
- eDiary data was collected on this day

The numerator will be the number of days contributing to the denominator, and also meeting the specific endpoint criteria: no symptoms (symptom-free days), no rescue medication use (rescue-free days), or no rescue medication use or symptoms (asthma control days).

If the treatment period extends beyond Day 168, the data from the days after Day 168 will not contribute to the calculation of these endpoints.

The percentages of symptom-free days, rescue-free days, and asthma control days will each be analyzed over the entire Treatment Period. Additionally, analyses will be conducted over each 4-week interval in the study. An ANCOVA model will be used with treatment, the number of the relevant 4-week interval (interval 1 to 6), baseline severe asthma exacerbation history, background ICS/LABA, and the treatment-by-4-week interval interaction as categorical covariates. Baseline percentage of endpoint days (where endpoint is symptom free, rescue free, or asthma control per model), baseline post-albuterol FEV_1 percent predicted, percent reversibility to ipratropium, and logarithm of baseline blood eosinophil count will be continuous covariates. An unstructured (UN) matrix will be used to model the variance-covariance structure. If the UN model fails to converge, then a first-order autoregressive (AR(1)) structure will be used instead. In the AR(1) model, subject will be included as a random effect. Contrasts will be used to obtain estimates of the treatment differences over 24 weeks.

Two-sided p-values and point estimates with two-sided 95% confidence intervals (CIs) will be produced for each treatment difference (*Table and Figure 2.9.1* for percentage of symptom-free days, *2.9.2* for number of rescue-free days, and *2.9.3* for number of asthma control days) for the efficacy estimand.

Table 3 4-Week Intervals

Interval	Time Period
1	Day 1 to Day 28
2	Day 29 to Day 56
3	Day 57 to Day 84
4	Day 85 to Day 112
5	Day 113 to Day 140
6	Day 141 to Day 168

Rescue Medication Usage

The mean daily number of puffs of rescue medication use will be calculated overall and for each of the 4-week intervals during the Treatment Period (see Table 3). For every period of time for which the mean number of puffs of rescue medication will be calculated, missing values will be ignored in both the numerator and denominator. As such, the denominator will be adjusted based on the number of days (including half days) with non-missing values. That is, the mean daily number of puffs of daytime rescue use (M_DT) will be set to the total number of daytime puffs

divided by the number of half-days when daytime rescue use was recorded. The mean daily number of puffs of nighttime rescue use (M_DN) will be set to the total number of nighttime puffs divided by the number of half-days when the nighttime rescue use was recorded. The mean daily rescue use (puffs) is then two multiplied by the mean of M_DT and M_DN.

A linear repeated measures ANCOVA model will be used to analyze change from baseline in average daily rescue albuterol use. The model will include treatment, the number of the relevant 4-week interval (interval 1 to 6, see Table 3), baseline severe asthma exacerbation history, background ICS/LABA, and the treatment-by-4-week interval interaction as categorical covariates and logarithm of baseline blood eosinophil count, baseline daily rescue albuterol use, percent reversibility to ipratropium, and baseline post-albuterol FEV_1 percent predicted as continuous covariates.

An unstructured variance-covariance matrix will be fit. If this model fails to converge, an AR(1) structure will be used instead. In the AR(1) model, subject will be included as a random effect. Contrasts will be used to obtain estimates of the treatment differences over 24 weeks.

Two-sided p-values and point estimates with two-sided 95% confidence intervals (CIs) will be produced for each treatment difference (*Table and Figure 2.10.1*) for the efficacy estimand.

Change from Baseline in Morning PEFR and Evening PEFR

Change from baseline in morning PEFR and evening PEFR over the treatment period are "Other" endpoints.

The PEFR is assessed using the Sponsor-provided PEFR meter, with the assessment data automatically transmitted from the PEFR to the eDiary. Subjects will complete the PEFR maneuver at home in the morning and in the evening before dosing with ICS/LABA (and ipratropium, if applicable) during the Screening Period, and in the morning and in the evening before dosing with study drug and ICS/LABA during the Treatment Period. In addition, PEFR should be measured before use of rescue albuterol.

The best (highest) of 3 PEFR efforts BID are to be captured in the eDiary. In case there is more than one assessment of PEFR available in the eDiary data, only the one with highest PEFR value will be retained for analysis (for the morning, and separately for the evening of a day). PEFR data with a "bad" vitalograph grade will not be used, where "Bad" is defined by the eDiary data vendor as unusable data. After this step, all morning and evening PEFR assessments will be allocated to 4-week intervals (see Table 3). Average morning and evening PEFR values will be calculated based on the data available within each interval.

Change from baseline in morning PEFR and evening PEFR will be analyzed using a linear repeated measures ANCOVA model. The model will include treatment, 4-week interval (interval 1 to 6), background ICS/LABA, and treatment by 4-week interval interaction as categorical covariates and baseline PEFR (morning or evening, depending on endpoint), percent reversibility to ipratropium, and logarithm of baseline blood eosinophil count as continuous covariates. An

unstructured (UN) matrix will be used to model the variance-covariance structure. If the UN model fails to converge, then a first-order autoregressive (AR(1)) structure will be used instead. In the AR(1) model, subject will be included as a random effect. Contrasts will be used to obtain estimates of the treatment differences over 24 weeks.

Two-sided p-values and point estimates with two-sided 95% confidence intervals (CIs) will be produced for each treatment difference (*Table and Figure 2.11.1* for morning PEFR, and *2.11.2* for evening PEFR) for the efficacy estimand.

PEFR data will be listed (*Listing 6.1.3).*

FE_{NO}

Change from baseline in FE_{NO} over the treatment period and each clinic visit, percentage of subjects with a clinically meaningful increase in FE_{NO} from baseline, and percentage of subjects with a clinically meaningful decrease from baseline in FE_{NO} are "Other" endpoints.

Measurement of FE_{NO} is a noninvasive and established clinical biomarker for airway inflammation. Two valid, reproducible FE_{NO} measurements are required, in accordance with testing procedures recommended by the manufacturer and similar to those published by the American Thoracic Society and European Respiratory Society (American Thoracic Society, European Respiratory Society 2005). The test is considered reproducible if ANY of these conditions is met:

- Either or both of 2 valid FE_{NO} measurements is < 30 ppb and the measurements are within 2 ppb of each other.
- Both measurements are > 30 ppb and within 10% of each other.
- Both values are below or above the measurement range (5-300 ppb).

If the reproducibility criteria are not met within the first 2 exhalations, a participant has up to 2 additional exhalations to satisfy the criteria (up to a total of 4 trials). The examinee's exhalations for which the device does not display a valid reading, i.e., exhalations that are either too strong or too weak, or any other reasons for failing to achieve a successful exhalation, will be recorded as attempts. Examinees may not always be able to meet the criteria for a reproducible measurement. The results of each attempt will be recorded in the appropriate eCRF. Any pair of the completed attempts can be assessed for reproducibility. If successful, the analysis value is defined as the average of the last 2 reproducible measurements. For example, if attempts 1 and 2 do not meet reproducibility conditions, attempt 3 is performed. If pairs (1,3) and (2,3) both satisfy reproducibility, the pair (2,3) will be used to derive the analysis value. If the reproducibility criteria are still not met, the value of Fe_{NO} will be missing.

Change from baseline in FE_{NO} will be analyzed using a linear repeated measures ANCOVA model, similar to the one used for FEV_1 $AUC_{0\text{-}4}$. The model will include treatment, visit, background ICS/LABA, and treatment by visit interaction as categorical covariates and baseline

FE_{NO}, percent reversibility to ipratropium, and logarithm of baseline blood eosinophil count as continuous covariates. An unstructured (UN) matrix will be used to model the variance-covariance structure. If the UN model fails to converge, then a first-order autoregressive (AR(1)) structure will be used instead. In the AR(1) model, subject will be included as a random effect.

Two-sided p-values and point estimates with two-sided 95% confidence intervals (CIs) will be produced for each treatment difference (*Table and Figure 2.12.1*) for the efficacy estimand.

The percentage of subjects with clinically meaningful increases or decreases from baseline in FE_{NO} will be summarized descriptively by visit (*Tables 2.12.2.1-2.12.2.2)*. The clinically meaningful changes are defined below:

- A clinically meaningful increase in FE_{NO} is defined as an increase of 20% or more from baseline if the baseline is 50 parts per billion (ppb) or more or an increase of more than 10 ppb from baseline if the baseline is less than 50 ppb.
- A clinically meaningful decrease in FE_{NO} is defined as a decrease of 20% or more from baseline if the baseline value is 50 ppb or more or a decrease of more than 10 ppb from baseline if the baseline value is less than 50 ppb.

At each visit, the percentage of subjects with FE_{NO} less than 25 ppb, between 25 and less than 50 ppb, and at least 50 ppb will be summarized descriptively by treatment (*Table 2.12.2.3)*.

Missing FE_{NO} data will not be imputed and will not contribute to the summary statistics. Percentages will be calculated out of the available FE_{NO} data at each visit. FE_{NO} data will be listed (*Listing 6.4).*

CompEx

A CompEx event is defined as the first occurrence of either a diary event (clinically relevant asthma deterioration) or a severe asthma exacerbation (Fuhlbrigge 2017).

The definition of a diary event will consider 6 endpoints: PEFR (morning and evening), rescue medication use (morning and evening), and asthma symptoms (morning and evening). For each endpoint, the fulfilment of deterioration criteria will be calculated as follows:

Thresholds: The change from baseline is calculated, where baseline for eDiary endpoints is defined in Section 6.4.1. If two consecutive days fulfil the chosen threshold limit (Table 4), the deterioration criterion is met.

Slopes: A slope is calculated via univariate linear regression over 5 days (described further in the below paragraph). If the slope fulfils the chosen limit (Table 4), the deterioration criterion is met.

Table 4 Asthma Deterioration Criteria

Variable	Threshold	Slope
PEFR (AM) PEFR (PM)	Decrease from BL ≥ 15% for two consecutive days	Decrease ≥ 3% per day
Number of puffs of rescue medication (AM) Number of puffs of rescue medication (PM)	Increase from BL ≥ 1.5 puffs for two consecutive days	Increase ≥ 0.3 puffs per day
Asthma symptoms score (Daytime)	Increase from BL ≥ 1 unit, or absolute score 5 for two consecutive days	Increase ≥ 0.2 units per day
Asthma symptoms score (Nighttime)	Increase from BL ≥ 1 unit, or absolute score 4 for two consecutive days	Increase ≥ 0.2 units per day

A diary event can occur when (i) the threshold deterioration criterion is met for at least two diary variables, or when (ii) the threshold deterioration criterion is met for one diary variable, and the slope criterion is fulfilled for all six variables. In case of (i), the diary event is defined to start on the first day of the two consecutive deterioration days (event days 0–1). Any missing data in this two-day window will make the event missing. In case of (ii), the event is defined to start on the first of the two days fulfilling the threshold criterion. This means that the slopes are calculated for days -4 to 0 of an event. At least two days with data are needed to calculate slopes, otherwise the event is considered missing. **The occurrence of a diary event within a subject is defined as the first occasion when either (i) or (ii) above occurs.**

Time to first CompEx event (i.e., a diary event or severe asthma exacerbation) will be analyzed with a Cox regression model to compare the treatment groups adjusted for baseline severe asthma exacerbation history, background ICS/LABA, percent reversibility to ipratropium, logarithm of baseline blood eosinophil count, and baseline post-bronchodilator FEV_1 percent predicted. Subjects who complete the study (and the study treatment) and do not experience a CompEx event will be censored at their Week 24 visit. Subjects who discontinue study treatment will be censored at the time of treatment discontinuation.

Two-sided p-values and point estimates with two-sided 95% confidence intervals (CIs) will be produced for each treatment difference (*Table and Figure 2.13.1*) for the efficacy estimand.

Differences from Fuhlbrigge:

- Fuhlbrigge defined baseline as “the mean over the 10 days ending just before the day of randomization for each of the diary variables. If no such data were available, the whole run-in period was used, and if missing, the first 3 days after randomization were used.” In this

study, baseline will be defined as the average of the non-missing values from the diary data collected in the last seven days of the Screening Period. This is consistent with Pearl analysis standards in which eDiary data are summarized as individual endpoints.

- Fuhlbrigge developed the CompEx endpoint using a paper diary that collected asthma symptoms on a 4-point scale. In this study, asthma symptoms are collected in a eDiary with different scales for AM and PM symptoms. Daytime symptom scores are collected in the evening on a 6-point scale. Nighttime symptom scores are collected in the morning on a 5-point scale. Although the scales are different, this study will retain the recommended thresholds and slopes for asthma symptoms score with a slight modification to account for the different possible maximum symptoms scores. Exploration of the impact of adjusting the thresholds and slopes to account for the different symptom scales may be pursued as post-hoc analyses.

6.4.7 Subgroup Analyses

Subgroup analyses will be performed for FEV_1 AUC_{0-4}, change from baseline in morning pre-dose trough FEV_1, and rate of moderate/severe asthma exacerbations (efficacy estimand only). The following subgroups will be considered:

- History of Severe Asthma Exacerbation in the last 12 Months:
  - None
  - One
  - At Least Two

If there are fewer than 20% of patients in the "At Least Two" subgroup, combine "One" and "At Least Two" into "At Least One".

- Baseline Eosinophil Count:
  - <300 cells per mm^3
  - ≥300 cells per mm^3

- T-helper cell type 2:
  - Low (Baseline eosinophils count < 300 cells per mm^3 AND baseline F_{ENO} < 40 ppb)
  - High (1) (Baseline eosinophils count ≥ 300 cells per mm^3 OR baseline F_{ENO} ≥ 40 ppb)
  - High (2) (Baseline eosinophils count ≥ 300 cells per mm^3 AND baseline F_{ENO} ≥ 40 ppb)

  Other cut-off values for FENO in this definition will be explored (e.g. 50).

- GINA Classification:
  - Step 3
  - Step 4 or 5

- Tiotropium Usage (Each subgroup must have ≥ 20% of study population):
  - Yes
  - No

The following subgroup analyses may also be performed in a post-hoc manner, where the cut point is to be determined based on the LOESS plots:

- Baseline Eosinophil count, with a cut point different from 300 cells per mm^3
- Baseline F_{ENO}
- T-helper cell type 2 with alternative cut points based on the LOESS plots of baseline eosinophils and F_{ENO}

Each subgroup will be analyzed separately using the same model that was used for the overall (combined subgroups) analysis. Estimates for the treatment effect and for the treatment differences will be displayed in the efficacy endpoint tables for each subgroup (*Tables 4.1.1* to *4.3.5*).

For each subgroup analysis, a test for the treatment-by-subgroup interaction will be performed using the same model that was used for the overall (combined subgroups) analysis but with the addition of terms for subgroup and the treatment-by-subgroup interaction. A table will be provided with the p-value for the test of the treatment-by subgroup interaction (*Table 4.4 and 4.4.1* for the efficacy estimand, for the planned and post hoc analyses, respectively).

Subgroup Cut Point Exploration

Subgroup analyses will be conducted in the baseline eosinophil count-high (≥300 cells per mm^3) and the baseline eosinophil count-low (<300 cells per mm^3) subgroups. It is acknowledged 300 cells per mm^3 may not ultimately be the appropriate threshold for evaluation of treatment benefit. Thus, additional analyses will evaluate alternative thresholds. This exploration will use Locally Weighted Scatter-plot Smoothing (LOESS) models to assess the impact of baseline eosinophil counts on AUC_{0-4} FEV_1 at Week 24 (*Figures 4.5.1 to 4.5.10*), on morning pre-dose trough FEV_1 at Week 24 (*Figures 4.6.1 to 4.6.10*) and on the rate of moderate/severe asthma exacerbations (*Figures 4.7.1 and 4.7.10*). Details of constructing these LOESS curves are described in the Appendix 6 to this SAP.

Similarly, LOESS curves will be created to explore the cut point for the baseline F_{ENO} (*Figures 4.5.11 to 4.5.20* for AUC, *4.6.11 to 4.6.20* for trough, and *4.7.11 to 4.7.20* for rate).

In addition, the cut point exploration for baseline eosinophils will be explored by general additive modeling (GAM) (*Figures 4.8.1 to 4.8.10* for AUC, *4.9.1 to 4.9.10* for trough, and *4.10.1 to 4.10.10* for rate). Details of performing GAM analysis are described in the Appendix 6 to this SAP.

6.4.8 Control of Type I Error

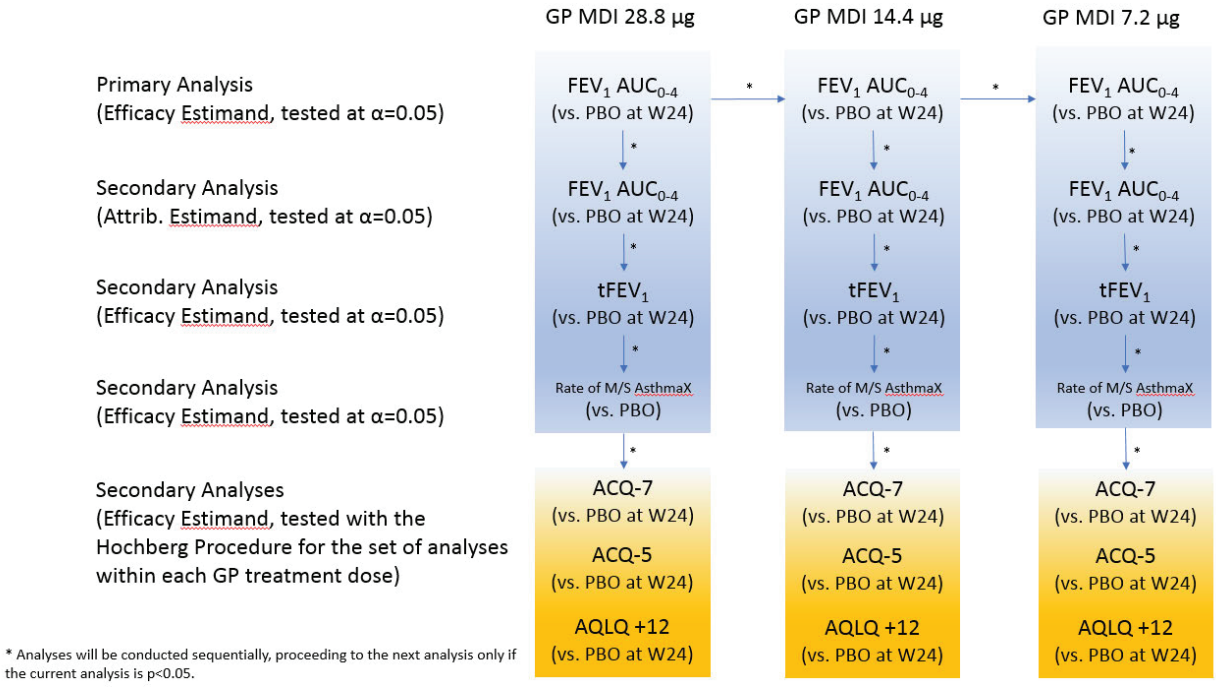
The Type I error rate will be controlled within the primary and secondary efficacy analyses. As a general strategy, the Type I error rate will be strongly controlled within the primary analysis. The Type I error rate will also be controlled within the family of secondary analyses within the GP MDI treatment groups.

Primary analysis testing will be conducted in descending dose order, with the highest dose of GP MDI tested against placebo at the nominal alpha. Testing will continue through the GP MDI doses as long as the results are successful. Testing will stop if any comparison of GP MDI versus placebo fails to meet statistical significance. Through this strategy, the Type I error rate of the primary analyses of all 3 GP MDI doses versus placebo will be strictly controlled to the nominal alpha.

The secondary analyses for each dose of GP MDI versus placebo will proceed only if the primary analysis for that GP MDI dose is successful. The Type I error rate within the secondary analyses of each GP MDI dose versus placebo will be controlled with a combination of sequential testing and the Hochberg procedure.

The Type I error control strategy is displayed in the figure below:

Figure 2 Order of Hypothesis Testing for Type I Error Control



6.4.9 Correlation Analysis

Pearson correlation coefficients will be generated between the primary and secondary continuous endpoints. The mITT Analysis Set will be used.

The correlations will be organized in a matrix, with its upper triangle filled with pairwise Pearson correlation coefficients. All treatment groups will be pooled (*Table 2.14.1*), and also analyzed individually (*Tables 2.14.2 to 2.14.6*).

6.5 Safety Analysis

All safety analyses are based on the Safety Analysis Set. Descriptive statistics will be used for safety analyses. Hypothesis testing will not be performed for any safety analyses. Comparisons performed for Holter Monitoring endpoints will be interpreted descriptively.

All AE data, clinically significant laboratory values, vital signs, and ECG values will be categorized according to their onset date into the following study periods:

- Events occurring during the treatment period ("On-Treatment") are events with an onset date on or after the first date of dose and up to and including the last day of randomized treatment (for study drug completers) or the last day of randomized treatment + 1 day (for premature treatment discontinuation). Events known to have occurred before the time of the first dose of study treatment are not included.

- Events occurring during the post-treatment-discontinuation follow-up are events with an onset date after the last day of randomized treatment (for study drug completers) or on or after the last day of randomized treatment + 2 days (for premature treatment discontinuation). The exception is that deaths are still considered to be during the Treatment Period if any adverse event that led to that death occurred during the Treatment Period.

Any AEs, clinically significant laboratory values, vital signs, and ECG values during the randomized-treatment period will be tabulated and listed. Beginning on the day after the date of discontinuation from or completion of study medication has passed, any new clinically significant ECGs, laboratory values, and vital signs will not be included in the tabulation or the computation of incidence rates, but will still be listed. Any new AEs, SAEs, and deaths during the post-randomized-treatment period will be tabulated and listed.

6.5.1 Adverse Events

The version of the Medical Dictionary for Regulatory Activities (MedDRA) that is current at the time of database lock will be used to code verbatim terms for AEs for final analysis of the data. A glossary of MedDRA preferred terms used for AEs reported in the study along with the associated Investigator's verbatim will be provided in *Listing 7.2*.

An AE is considered treatment-emergent if an event occurs after the first dose of randomized study medication in the study, or if the AE worsened during the study after the first dose of study

medication in the study (intensity and/or severity changed to a worsened grade), and the event onset is on or before the last day of randomized treatment (for study drug completers) or the last day of randomized treatment + 1 day (for premature treatment discontinuation). AEs with onset date after the last day of randomized treatment (for study drug completers) or the last day of randomized treatment + 1 day (for premature treatment discontinuation) will not be considered treatment-emergent, but will be listed in adverse event data listings, and will be tabulated separately. For the purposes of this SAP, the terms "treatment-emergent AE" and "On-Treatment AE" are synonymous. A more detailed definition may be found in Appendix 1 (Data Handling Rules, Category 16). AEs that occur between the time the subject signs the informed consent/assent form for the study and the time when that subject is randomized are to be recorded as medical history unless the event met the definition of an SAE.

The incidence of an AE will be defined as the number and percentage of subjects experiencing an event. Adverse events will be tabulated at the level of the Medical Dictionary for Regulatory Activities (MedDRA) preferred term and the MedDRA system organ class. No hypothesis tests will be performed.

An overview table will be prepared for the Safety Analysis Set with the incidences of subjects with at least one treatment-emergent adverse event (TEAE), at least one serious TEAE, at least one TEAE related to study treatment, at least one serious TEAE related to study treatment, at least one TEAE leading to premature treatment discontinuation, and a report of death (*Table 3.1*).

Events with Irregular Onset Dates

All treatment-emergent adverse events will be included in the data listings regardless of the completeness of the onset dates. Partial dates will be imputed in order to determine if an AE is treatment-emergent using the imputation rules in Appendix 1; however, imputed dates will not be provided in the data listings.

General Summarization and Listing Strategy for Adverse Events

All adverse events, whether treatment-emergent or not, will be included in the listings. Reported adverse events by system organ class, preferred term, treatment, country, center, subject and onset day will be provided *(Listing 7.1)*. Reported adverse events by treatment, country, center, subject, and onset date will be presented in *Listing 7.3*. SAE-specific report information will be listed in *Table 3.8.1* and *Table 3.8.2*.

The listing of adverse events will provide the severity, maximum severity, relationship to study drug, action taken and outcome for each adverse event. Any SAEs reported will be listed for all subjects screened (*Tables 3.8.1).* Adverse events leading to permanent discontinuation of study treatment will be listed for the Safety Analysis Set *(Table 3.6)*. A listing of any reported deaths during the study (prior to randomization, during the randomized-treatment period, or during the post- treatment discontinuation follow-up) will be provided (*Table 3.12*); study treatment taken

prior to the death and the number of days since the last dose of this study treatment at the time of the death will be included in the listing.

Summary tabulations of the following will be prepared for all subjects, for each treatment, for each primary system organ class, and for each preferred term within a system organ class:

- The incidence of all treatment-emergent adverse events *(Table 3.2.1.1)*
- The incidence of subjects with adverse events by SOC during the post- treatment discontinuation follow-up (*Table 3.2.1.2)*
- The incidence of treatment-emergent adverse events occurring in AEs of special interest *(Table 3.2.3)*
- The incidence of non-serious treatment-emergent adverse events occurring in ≥5% of subjects in a treatment *(Table 3.2.4)*
- The incidence of all treatment-related treatment-emergent adverse events *(Table 3.4)*
- The incidence of discontinuation from study treatment due to treatment-emergent adverse events *(Table 3.5)*
- The incidence of discontinuation from study treatment due to treatment-related treatment-emergent adverse events *(Table 3.5.1)*
- The incidence of treatment-emergent serious adverse events *(Tables 3.7.1)*
- The incidence of subjects with serious adverse events by SOC during the post-treatment discontinuation follow-up *(Tables 3.7.2)*
- The incidence of all treatment-related treatment-emergent serious adverse events *(Tables 3.9)*
- The incidence of all treatment-emergent adverse events by highest severity to treatment *(Tables 3.11.1* through *3.11.5* for the five treatments)
- The incidence of treatment-emergent adverse events occurring in at least 2% of subjects in any treatment *(Table 3.2.2)* sorted by descending frequency of events in a preferred term).
- In addition, to control for possible differences in exposure between the treatments, the following AE and SAE summaries will be presented with the frequency and rate of occurrence (total number of events per 1000 person-years of exposure) by treatment, primary system organ class, and preferred term:
    - Frequency and rate of AEs *(Table 3.3)*

- Frequency and rate of SAEs *(Table 3.8.3)*
- Frequency and rate of neoplasms (*Tables 3.10.1* and *3.10.3* – All Cancer, *3.10.2* and *3.10.4* - Excluding Non-Melanoma Skin Cancer).

Adverse Events of Special Interest

Adverse events of special interest (AESIs) have been defined based on known effects of LAMAs. These include anticholinergic effects, cardiovascular effects, cerebrovascular conditions, ocular disorders, paradoxical bronchospasm, urinary retention, and gastrointestinal disorders.

Standard MedDRA queries (SMQs) will be utilized when possible, and a selection of high-level terms (HLTs) and preferred terms (PTs) will be utilized to represent other situations. The definitions of AESIs associated with LAMAs are listed in Table 5.

Table 5 Adverse Events of Special Interest

Medical Concept	Selection of MedDRA Terms
Anticholinergic effects	Anticholinergic syndrome SMQ Dry mouth PT
Cardiovascular effects	Cardiac arrhythmias SMQ Cardiac failure SMQ Ischemic heart disease SMQ Torsades de Pointe/QT prolongation SMQ
Cerebrovascular conditions	CNS haemorrhages and cerebrovascular conditions SMQ
Gastrointestinal disorders	Gastrointestinal obstruction SMQ
Ocular (disorders) effects	Visual disorders HLT Glaucoma SMQ
Paradoxical bronchospasm	PT Bronchospasm paradoxical
Urinary retention	Collection of PTs: urinary retention, urinary incontinence, pollakiuria, dysuria

Appendix 5 (which will be based on the latest version of MedDRA available at the time of database lock) provides detail on selection of PTs.

AESIs will be tabulated by medical concept, preferred term and treatment group (*Table 3.2.3*).

Paradoxical Bronchospasm

During Visits 4, 5 (if available), 7 and 10, a paradoxical bronchospasm event is defined as a reduction in FEV_1 of >20% from baseline with associated symptoms of wheezing, shortness of breath, or cough that occur within 30 minutes post-dosing. Baseline is defined as the mean of the FEV_1 values obtained 60 and 30 minutes prior to study drug and ICS/LABA administration at that visit. Spontaneous reporting of paradoxical bronchospasm will occur at Visits 5 (for subjects without post-dose spirometry), 6, 8, and 9.

All paradoxical bronchospasm events will be captured on the AE CRF page. For those events that occurred during Visits 4, 5 (if available), 7 and 10, a programmatic check will be done to verify whether they satisfy the condition on the change in FEV_1, and will be queried as necessary. Paradoxical bronchospasms will be summarized by treatment during the randomized-treatment period *(Table 3.2.3)*.

6.5.2 Clinical Laboratory Measurements

Lab parameters collected include the following:

Table 6 Lab Parameters

Hematology	
Hemoglobin	Mean corpuscular hemoglobin
Hematocrit	Mean corpuscular hemoglobin concentration
White blood cell count with differential	Mean corpuscular volume
Red blood cell count	Eosinophil count
Platelet count	
Clinical Blood Chemistry	
Liver Enzyme and Other Liver Function Tests	**Other Clinical Blood Chemistry**
Alanine aminotransferase	Albumin
Aspartate aminotransferase	Blood Urea Nitrogen (BUN)
Alkaline phosphatase	Calcium
Bilirubin, total	Chloride
Gamma-glutamyl transferase	Cholesterol, total
	Bicarbonate
	Creatinine
	Glucose
	Magnesium
	Potassium
	Phosphate
	Protein, total
	Sodium
	Triglycerides
Urinalysis	
Macroscopic examination including specific gravity, pH, protein, glucose, ketones, blood, and urobilinogen.	
Other Tests:	
Pregnancy test (women of childbearing potential only): serum hCG at Visit 1 (Screening) and Final Visit (Visit 10) or Treatment Discontinuation/Study Withdrawal Visit, urine hCG at all other visits	
Creatinine clearance will be estimated by the CKD-EPI formula [Levey, 2009] for subjects >18 to 80 years of age or the Schwartz formula for subjects 12 to ≤18years of age [Schwartz, 1987].	
Abbreviations: CKD-EPI=Chronic Kidney Disease Epidemiology Collaboration; hCG=human chorionic gonadotropin	

A Clinically Significant Laboratory Abnormality as identified by the investigator after the start of study treatment will be recorded as an Adverse Event and tabulated as an AE in the AE analysis. Abnormalities occurring prior to the start of treatment will be noted in medical history and presented in a data listing. Per protocol, the criteria for a "clinically significant" laboratory abnormality are:

- A laboratory abnormality that leads to a dose-limiting toxicity (e.g., an abnormality that results in study drug dose reduction, suspension or discontinuation)
- A laboratory abnormality that results in any therapeutic intervention (i.e., concomitant medication or therapy)
- Other laboratory abnormality judged by the Investigator to be of any particular clinical concern (e.g., significant fall in hemoglobin not requiring transfusion)
- For laboratory abnormalities that do not meet the above criteria, but are outside of normal range (e.g., < or > normal reference range), the Investigator should indicate whether the value is clinically significant or not clinically significant for the subjects.

All laboratory data will be stored in the database with the units in which they were originally reported. Laboratory data not reported in International System of Units (SI units; Système International d'Unités) will be converted to SI units before data analysis.

Individual clinical laboratory variables for hematology and clinical chemistry and kidney function, including creatinine clearance, will be provided in listings *(Listing 8.1* for hematology*, Listing 8.2* for blood chemistry and kidney function, and *Listing 8.3* for urinalysis). Data will be listed in SI units where available. Comments for laboratory testing will be listed *(Listing 8.4)*. For listings, laboratory values will be flagged as Low or High based on the reference ranges provided by the central laboratory, Covance *(Appendix 4)*. These flags along with the reference ranges will be provided in the laboratory data listings.

The baseline measurement for a laboratory parameter will be the last available measurement prior to the start of dosing.

If there are multiple laboratory values for the same parameter at pre-dose of a visit, the last value will be chosen for analysis.

Summary statistics (n, mean, median, standard deviation, minimum, and maximum) for the baseline assessment and for the pre-dose value and change from baseline at each post-baseline visit and end of treatment for scheduled lab assessments of continuous laboratory variables including serum potassium and glucose will be tabulated. "End of Treatment" is defined as the last non-missing assessment during the treatment period. Data from unscheduled visits will not be used for the by-visit summaries but both scheduled and unscheduled-visit are candidates for the end-of-treatment summary. Data from both scheduled and unscheduled visits will be listed.

The summaries will be provided by treatment *(Tables 3.13.1* through *3.13.4,* for hematology, blood chemistry, kidney function, and urinalysis, respectively).

Data from unscheduled visits will not be used for the by-visit summaries but both scheduled-visit data and unscheduled-visit data are candidates for clinically significant values, for the end-of-treatment summary, and for shift tables. Shift tables will be produced using the categories defined by the Common Terminology Criteria for Adverse Events (CTCAE) Version 4.03 grades for the Safety Analysis Set *(Tables 3.13.5 to 3.13.7* for hematology, chemistry, and kidney function, respectively). For these shift tables, for each treatment, the subject's pre-dose grade will be cross-tabulated by the subject's maximum post-baseline grade during the treatment; also, the subject's maximum post-baseline grade during treatment will be tabulated for all baseline grades combined. Percentages of subjects in each maximum post-baseline grade for a treatment will be calculated for each pre-dose grade for the treatment and also for all baseline grades combined.

Potentially Clinically Significant (PCS) Laboratory Values Above/Below a Clinically Relevant Threshold on-treatment, based on CTCAE 4.03 and other criteria, will be identified based on the following thresholds:

Table 7 Potentially Clinically Significant Laboratory Parameter Criteria

Parameter	Post-Baseline Criteria
Hematology	
Hemoglobin	<8.0 g/dL (<80 g/L)
	Increase of >40 g/L to a value above the ULN (upper limit of normal)
White Blood Cell Count	<2000/µL
	>35,000/µL
Platelet Count	<50,000/µL
	>999,000/µL
Chemistry	
eGFR-EPI (where eGFR denotes estimated glomerular filtration rate)	Decrease from baseline in ≥ 1 CTCAE grade and ≥ 20% change from baseline (#)
AST (aspartate aminotransferase)	>3 x ULN
ALT (alanine aminotransferase)	>3 x ULN
Alkaline Phosphatase	>5 x ULN
Total Bilirubin	>2 x ULN
Blood Glucose* (random values)	<2.2 mmol/L (<39.6 mg/dL)
	>13.9 mmol/L (>250 mg/dL) if no history of diabetes, > 27.8 mmol/L (>500 mg/dL) regardless of baseline
Serum Potassium	<3.0 mmol/L
	>6.0 mmol/L

*CTCAE 4.03 criteria are based on fasting glucose values. However, subjects were not required to fast prior to obtaining blood glucose values.
Only a 20% decrease from baseline is considered here: a 20% increase would not be of concern.

Clinically significant laboratory values will be tabulated for the Safety Analysis Set (*Table 3.13.8*).

Since a reduction in potassium and an increase in blood glucose are known class effects of beta-agonists, all potassium or glucose assessments for subjects who experienced newly occurring or worsening potentially clinically significant values after start of the study treatment will be provided in separate listings *(Tables 3.13.9* and *3.13.10).* For all laboratory parameters other than glucose and potassium noted in Table 6, all laboratory data for the parameter identified as potentially clinically significant for a subject will be listed *(Table 3.13.11* - Safety Analysis Set).

6.5.3 Vital Signs

Changes from Baseline in on-treatment supine or seated systolic blood pressure, supine or seated diastolic blood pressure, and heart rate will be evaluated, where baseline is defined as the pre-dose measurement taken prior to the start of dosing at the Randomization Visit (Visit 4). If there are no Visit 4 pre-dose values, the baseline will be defined as the pre-bronchodilator values at Visit 3. No hypothesis testing will be performed.

A **Clinically Significant Abnormality** in vital signs identified by the investigator will be recorded as an Adverse Event if it occurs after the start of treatment. These adverse events will be included in the AE summaries; abnormalities prior to the start of treatment will be noted in medical history and listed.

Potentially clinically significant changes in systolic and diastolic blood pressure will be defined based on the following criteria provided by Pearl, Inc.:

Table 8 Potentially Clinically Significant Criteria for Systolic and Diastolic Blood Pressure Parameters

Parameter (mmHg)	Post-Baseline Criteria
Systolic Blood Pressure, increase	≥180 and increase from baseline ≥20
Systolic Blood Pressure, decrease	≤90 and decrease from baseline ≥20
Diastolic Blood Pressure, increase	≥105 and increase from baseline ≥15
Diastolic Blood Pressure, decrease	≤50 and decrease from baseline ≥15

Potentially clinically significant changes in heart rate will be assessed as follows:

Table 9 Potentially Clinically Significant Criteria for Heart Rate Parameters

Parameter	Post-Baseline Criteria
Tachycardia Event	≥110 bpm and increase ≥15% from baseline
Bradycardia Event	≤50 bpm and decrease ≥15% from baseline

bpm = beats per minute.

Vital sign measurements (Heart rate, systolic blood pressure, diastolic blood pressure, weight, height) during the study will be displayed in a vital signs listing *(Listing 9.1)*.

A summary of baseline weight, height, and BMI will be presented by treatment *(Tables 1.4.1-1.4.4)* for the mITT, ITT and Safety Analysis Sets, and all subjects not randomized, respectively. Adolescent subjects (12-18 years of age) will have height measurements collected at Visits 1, 4, 5, 7 and 10. A summary of height and change from baseline in height will be presented for the adolescent subjects in the mITT Analysis Set, by gender and treatment (*Table 1.4.5*).

Summary statistics (n, mean, median, standard deviation and range) of the absolute value and change from baseline for systolic blood pressure, diastolic blood pressure, and heart rate will be tabulated by treatment, visit, and time point. These summaries (*Table 3.14.1*) will be prepared for baseline and each scheduled post-baseline nominal time point at each scheduled post-baseline visit and end of treatment. End of Treatment will be summarized for each scheduled post-baseline time point (pre-dose, and post-dose 4 hours). “End of Treatment” for each of these assessment points is defined as the last non-missing on-treatment assessment available for the time point. Data from unscheduled visits will not be used for the by-visit summaries but both scheduled and unscheduled-visit data are candidates for clinically significant values and for the end-of-treatment summary. Data from both scheduled and unscheduled visits will be listed. Time windows will be derived for each post-baseline visit using the time intervals for the study time windows detailed in Table 10. No hypothesis tests will be performed.

Table 10 Analysis Study Time Windows for Vital Signs Assessments

Calculated Study Time Window	Time Interval for the Study Time Window
Pre-dose	≥0 min. prior to dose
Post-dose 4 hours	>0 to <6 hours post-dose

Note that minutes are rounded to the nearest whole number before applying time windows.

If there are multiple vital sign values for the same parameter at pre-dose assessments after Visit 4 or within the same post-dose study time window at a visit, the last value will be chosen for analysis.

The percentage of subjects with potentially clinically significant values for vital signs at any time post-baseline at a visit will be summarized by treatment based on the criteria in Table 8 and Table 9 (*Table 3.14.2*).

All vital sign assessments for subjects with potentially clinically significant values will be listed *(Tables 3.14.3 and 3.14.4)*.

6.5.4 12-Lead Electrocardiogram Measurements

Changes from baseline in Heart Rate, PR Interval, QRS Axis, QRS Interval, QT Interval and QTcF (Fridericia Corrected QT) interval will be calculated where baseline is defined as the pre-dose measurement taken prior to the start of treatment at the randomization visit (Visit 4). If there are no Visit 4 pre-dose values, the baseline will be defined as the value obtained at Visit 1. The QTcF is defined as $[QT/(RR^{1/3})]$. Heart rate (bpm) is estimated as 60,000/RR, where RR is in units of msec. These assessments will be tabulated for each treatment and assessment time.

A **Clinically Significant Abnormality** for a 12-Lead ECG measurement identified by the investigator as a clinically significant abnormality will be recorded as an Adverse Event if it occurred after the start of study treatment. These adverse events will be included in the AE summaries.

All 12-Lead ECG measurements for the Safety Analysis Set will be listed *(Listing 9.2)*. Summary statistics (mean, median, standard deviation and range) for raw values and change from baseline values in Heart Rate, PR Interval, QRS Axis, QRS Interval, QT Interval and QTcF interval will be calculated. These assessments will be tabulated for each treatment and each scheduled nominal time point at each visit and at end of treatment (*Table 3.15.1*). End of Treatment is defined as the last non-missing on-treatment assessment available. Data from unscheduled visits will not be used for the by-visit summaries but both scheduled and unscheduled-visit data are candidates for clinically significant values and for the end-of-treatment summary. Data from both scheduled and unscheduled visits will be listed. Mean pre-dose change from baseline for ECG parameters will be plotted across post-baseline visits by treatment *(Figures 3.15.1 – 3.15.5)*. ECG data from subjects with pacemakers will not be included in analyses, but will be listed.

If there are multiple ECG values for the same parameter at pre-dose of a visit date (other than for Visit 4), the last value will be chosen for analysis.

Other than for the change from baseline analyses mentioned above, all available data after randomization including data from unscheduled visits will be used for ECG parameter analyses.

Table 11 Criteria for PCS ECG Values

Parameter	Post-Baseline Criteria
QTcF Prolongation	(1) ≥ 500 msec if < 500 msec at study baseline and ≥ 15 msec change from study baseline (2) ≥ 530 msec if ≥ 500 msec at study baseline and ≥ 15 msec change from study baseline (3) ≥ 500 msec and ≥ 15 msec change from study baseline (4) Change of ≥60 msec from study baseline regardless of initial value (5) ≥ 500 msec and ≥60 msec change from study baseline

Potentially clinically significant ECG parameter values will be identified based on criteria listed in Table 11. The number and percentage of subjects who had such values observed any time post-baseline will be tabulated for each treatment (*Table 3.15.2*) and listed (*Tables 3.15.3* for QTcF prolongation). No hypothesis tests will be performed.

6.5.5 24-Hour Holter Monitoring Sub-Study

For subjects participating in the Holter monitoring sub-study, continuous Holter monitoring will be collected over 24 hours at Visit 3 and at Visit 7 (Week 12). If continuous Holter monitoring is unacceptable (less than 18 hours of acceptable quality), it may be repeated. If repeated, then the repeated values will be used in all analyses. However, any incidence of AEs indicated by the incomplete Holter findings will be captured. Holter monitoring data will be listed for subjects in the Holter Monitoring analysis set (*Listing 9.3.1, 9.3.2 and 9.3.3*). For subjects excluded from the Holter Monitoring analysis set, investigator's assessment of Holter data will be listed where available (*Listing 9.3.4*).

Baseline for each endpoint will be defined using the acceptable 24-hour Holter assessment at Visit 3 (Screening).

Note that [redacted] will provide mean 24-hr heart rate, mean daytime heart rate, and mean nighttime heart rate, which will be an average across hourly estimates collected during the specific Holter monitoring period (24-hour, daytime, nighttime).

Primary Analysis

The change from baseline in mean 24-hour heart rate (HR) obtained using Holter monitoring at Week 12 will be analyzed using an ANCOVA model to evaluate treatment differences with

baseline mean 24-hour HR (obtained during 24-hour Holter monitoring at Screening) as a covariate. LS means and estimated treatment differences with 95% CIs will be provided (*Table 3.16.1* and *Listing 9.3.1*). The raw mean values and change from baseline values will also be summarized by treatment. A schematic box-plot will display the distribution of change from baseline in mean heart rate by treatment, with extreme values that are 1.5*IQR above/below the upper/lower quartiles identifiable, where IQR is the interquartile range (*Figure 3.16.1*).

Secondary and Other Holter Monitoring Data Safety Analyses

The changes from baseline at Week 12 (Visit 7) for the mean daytime (06:00 to 22:00) HR, mean nighttime (22:00 to 06:00) HR, maximum 24-hour HR, and minimum 24-hour HR will be summarized and analyzed in a similar manner to the primary Holter endpoint (*Tables* and *Figures 3.16.2, 3.16.3, 3.16.4.1,* and *3.16.5.1* and *Listing 9.3.1*).

A frequency distribution of the following will be provided:

- Proportion of subjects with maximum heart rate during treatment of >180, >160-≤180, >140-≤160, >120-≤140, >100-≤120, and 100 or less (*Table 3.16.4.2*).
- The number and percentage of subjects who had sustained ventricular tachycardia (defined as PVCs lasting>30 seconds) will be tabulated by treatment (*Table 3.16.9*).
- Proportion of subjects with minimum heart rate during treatment of >60, >50-≤60, >40-≤50, and ≤40 (*Table 3.16.5.2*).
- Proportion of subjects with change from baseline in mean heart rate of >5, >10, >20 bpm (*Table 3.16.1.1*).
- Proportion of subjects with change from baseline in maximum heart rate of >5, >10, >20 bpm (*Table 3.16.4.3*).
- Proportion of subjects with change from baseline in minimum heart rate of >5, >10, >20 bpm (*Table 3.16.5.3*).

The change from baseline in the number of Holter ventricular and supraventricular events will be summarized descriptively (mean, median, range, etc.) and analyzed with nonparametric methods. The Wilcoxon Rank Sum test will be used to produce p-values for the pairwise comparison of treatments. The median treatment differences will be presented with 95% confidence intervals based on the Hodges-Lehmann approach for the location shift (median treatment difference). This analysis will be performed for change from baseline for the following parameters (calculated per hour): number of isolated ventricular events (PVCs), number of ventricular couplets, number of ventricular runs, number of isolated supraventricular events, number of supraventricular couplets, number of supraventricular runs, and number of supraventricular ectopic beats (*Tables* and *Figures 3.16.6.1, 3.16.7.1, 3.16.8.1, 3.16.10.1, 3.16.11.1, 3.16.12.1, and 3.16.13.1,* respectively, *Listing 9.3.2* for ventricular and supraventricular events).

The number of subjects experiencing an isolated ventricular event, ventricular couplet, ventricular run, isolated supraventricular event, supraventricular couplet, supraventricular run and supraventricular ectopic beat at Week 12 will be analyzed with a logistic regression model

comparing across the treatment groups with baseline value of the Holter variable as a continuous covariate, and treatment and background ICS/LABA as categorical covariates. P-values and odds ratios with 95% CIs will be produced for each treatment comparison (*Tables 3.16.6.2, 3.16.7.2, 3.16.8.2, 3.16.10.2, 3.16.11.2, 3.16.12.2, and 3.16.13.2, respectively*).

Additionally, the following will be tabulated:

- The proportion of subjects in each category of change from baseline in the number of isolated PVCs per hour (no change, increase of >0-<60, ≥60-<120, and ≥120, and decrease of >0-<60, ≥60-<120, and ≥120) (*Table 3.16.6.3*).
- The number of subjects who had atrial fibrillation with a rapid ventricular response (>100 bpm) will be tabulated by treatment (*Table 3.16.14*).
- Bradycardia and tachycardia episodes and Holter assessment will be listed (*Listing 9.3.3*).

6.5.6 Physical Examination

Any physical examination abnormality reported after the start of treatment for a subject is to be reported as an adverse event. Thus, these will be included in listings of adverse events and summarized in adverse event summaries. Abnormalities seen at the Screening physical examinations will be recorded as Medical History and listed.

7. CHANGES FROM METHODS PLANNED IN THE PROTOCOL

None.

8. STATISTICAL SOFTWARE

Data processing, statistical screening, descriptive reporting and analysis of the efficacy and safety data will be performed using SAS (Version 9.4 or higher). Graphs may also be produced using R (R Development Core Team, 2003).

9. REFERENCES

American Thoracic Society, European Respiratory Society. ATS/ERS recommendations for standardized procedures for the online and offline measurement of exhaled lower respiratory nitric oxide and nasal nitric oxide, 2005. Am J Respir Crit Care Med. 2005;171(8):912-930.

Fuhlbrigge AL, Bengtsson T, Peterson S, et al. A novel endpoint for exacerbations in asthma to accelerate clinical development: a post-hoc analysis of randomized controlled trials. Lancet Respiratory Medicine 2017. DOI: 10.1016/S2213-2600(17)20218-7.

Global Initiative for Asthma (GINA). Global Strategy for Asthma Management and Prevention, 2017. http://ginasthma.org/. Accessed 17 September 2017.

Levey AS, Stevens LA, Schmid CH, et al. A new equation to estimate glomerular filtration rate. Ann Intern Med. 2009;150(9):604-612.

Reddel HK, Taylor DR, Bateman ED, et al. An official American Thoracic Society/European Respiratory Society statement: asthma control and exacerbations: standardizing endpoints for clinical asthma trials and clinical practice. Am J Respir Crit Care Med. 2009;180(1):59-99.

SAS Institute Inc. 2008. SAS/STAT 9.2 User's Guide. Cary, NC: SAS Institute Inc.

Schwartz GJ, Brion LP, Spitzer A. The use of plasma creatinine concentration for estimating glomerular filtration rate in infants, children, and adolescents. Pediatr Clin North Am. 1987;34(3):571-590.