[redacted]

Mylan Pharma UK Ltd.

REV-3001

A Randomized, Double Blind, Placebo-Controlled, Parallel Group Study of Nebulized Revefenacin Inhalation Solution in Chinese Subjects with Moderate to Very Severe Chronic Obstructive Pulmonary Disease (COPD)

Statistical Analysis Plan

Version: 2.0

[redacted] **Project Number:** [redacted]

NCT05046795
Document effective date: 16 Oct 2023

SPONSOR SIGNATURE PAGE

Approved by: Signed and dated electronically on last page _____

[redacted] Date

[redacted] Biostatistics

Mylan Pharma UK Ltd.

[redacted] SIGNATURE PAGE

Signature(s) below confirm that the Statistical Analysis Plan was developed in accordance with [redacted] and that it is approved for release.

This document has been approved and signed electronically on the final page by the following:

Signatory	
Author	[redacted] Project Role: Biostatistics[redacted]

TABLE OF CONTENTS

REVISION HISTORY

Version No.	Effective Date	Summary of Change(s)
1.0	14 Dec 21	New document
1.1	21 Jan 22	Corrected the items weight in Appendix 6.7.1 for question 14. Added description in Section 4.11.5 to describe algorithms of handling multiple ECG measurements performed at the same visit: the average of results used for continues variables, the worst result used for categorical variables.
	05 Feb 22	Corrected table number in Section 4.10
	15 Feb 22	Section 4.3.3, for the rule handling multiple records, Added "earliest" to illustrate that if there are multiple records are nearest, the earliest one will be used.
	27 Mar 22	Removed duplicate 'X' in Appendix 6.6
	10 May 22	Corrected the 'contrast' statement to 'estimate' statement for the model based average generation in Section 4.10.3.8 and 4.10.3.9 Corrected table number in Appendix 6.3
	13 Jul 22	Corrected spelling in Section 3.2.1
	10 Oct 23	Section 4.10.1.5, Examination of Subgroups Added subgroup analysis by ethnicity: North East Asian verses South East Asian
2.0	16 Oct 23	Finalized for final analysis

LIST OF ABBREVIATIONS

Abbreviation / Acronym	Definition / Expansion
AE	Adverse event
AECOPD	Acute exacerbation of COPD
ANCOVA	Analysis of covariance
ATC	Anatomical Therapeutic Chemical
AUC	Area Under the Curve
BDR	Blinded Data Review
BDI/TDI	Baseline Dyspnea Index/Transition Dyspnea Index
CBL	Change from Baseline
CI	Confidence Interval
COPD	Chronic Obstructive Pulmonary Disease
CRF	Case report form
CSR	Clinical Study Report
DBP	Diastolic Blood Pressure
ECG	Electrocardiogram
eCRF	electronic Case Report Form
EoT	End of Treatment
FAS	Full analysis set
FEV_1	Forced expiratory volume in 1 second
FVC	Forced vital capacity
GOLD	Global Initiative for Obstructive Lung Disease
HR	Heart rate
ICS	Inhaled Corticosteroid
IVRS/IWRS	Interactive Voice Response System/ Interactive Web Response System
LABA	Long Acting Beta2 Agonist
LAMA	Long Acting Muscarinic Antagonist
LOCF	Last observation carried forward
LS mean	Least squares mean
MedDRA	Medical Dictionary for Regulatory Activities
MAR	Missing at random
MI	Missing imputation
mMRC	modified Medical Research Council
MMRM	Mixed model repeated measures
MNAR	Missing not at random
OR	Odds Ratio
PP	Per protocol analysis set
PR	Pulse Rate
PRO	Patient Reported Outcome
PT	Preferred term
QTc	QT corrected
QTcB	QT corrected (Bazett's correction)
QTcF	QT corrected (Fridericia's correction)
PR	Pulse rate
SAE	Serious adverse event

Abbreviation / Acronym	Definition / Expansion
SAP	Statistical Analysis Plan
SBP	Systolic Blood Pressure
SGRQ	Saint George's Respiratory Questionnaire
SOC	System organ class
SS	Safety analysis set
TEAE	Treatment-emergent adverse events
WHODRUG	World Health Organization Drug Dictionary

1 INTRODUCTION

This initial statistical analysis plan (SAP) describes the planned summaries and statistical analyses of data for the clinical study REV-3001.

The overall objectives of this study are to confirm the efficacy and safety of revefenacin inhalation solution (175mcg once daily) for the treatment of Chinese subjects with moderate to very severe Chronic Obstructive Pulmonary Disease (COPD).

The analyses described in this SAP are based upon the following study documents:

- Study Protocol, Version 2.0 (July 15, 2020)
- electronic Case Report Form (eCRF), Version 4.0 (January 19, 2022)

2 STUDY OBJECTIVES

2.1 Primary Objective

The primary objective of the study is to confirm the efficacy of revefenacin inhalation solution 175 mcg administered once daily via nebulization for 12 weeks compared to placebo in a population of Chinese subjects with moderate to very severe COPD.

Note: "Revefenacin inhalation solution 175 mcg" will be referred to as "revefenacin 175 mcg" or "revefenacin" in subsequent sections of the SAP.

2.2 Secondary Objectives

The secondary objectives of the study are:

- To confirm the safety of revefenacin 175 mcg administered once daily via nebulization for 12 weeks compared to placebo in a population of Chinese subjects with moderate to very severe COPD.
- To characterize the efficacy of revefenacin 175 mcg on patient reported outcomes compared to placebo in a population of Chinese subjects with moderate to very severe COPD.
- To collect pharmacokinetics (PK) samples for population PK analysis of revefenacin 175 mcg in a population of Chinese subjects with moderate to very severe COPD.

3 INVESTIGATIONAL PLAN

3.1 Overall Study Design and Plan

3.1.1 Study Design

This is a randomized, double-blind, placebo-controlled, parallel-group, multi-center study in Chinese subjects with moderate to very severe COPD treated for 12 weeks.

The study population consists of

- Males and females of Chinese ethnicity, at least 40 years of age;
- A clinical diagnosis for at least 6 months prior to screening of COPD according to Global Initiative for Chronic Obstructive Lung Disease (GOLD) guidelines;
- Subjects capable of performing reproducible spirometry maneuvers and with a post-ipratropium Forced Expiratory Volume in 1 second/Forced Vital Capacity (FEV_1/FVC) ratio <0.7 at Screening Visit 2;
- Subjects with moderate to very severe COPD with a post-ipratropium FEV_1 <80% of predicted normal FEV_1 and an absolute FEV_1 > 700 mL at Screening Visit 2;
- Current smoker or ex-smoker, with a history of at least 10 pack-years of tobacco smoking. Ex-smokers must have stopped smoking >6 months prior to Visit 1.

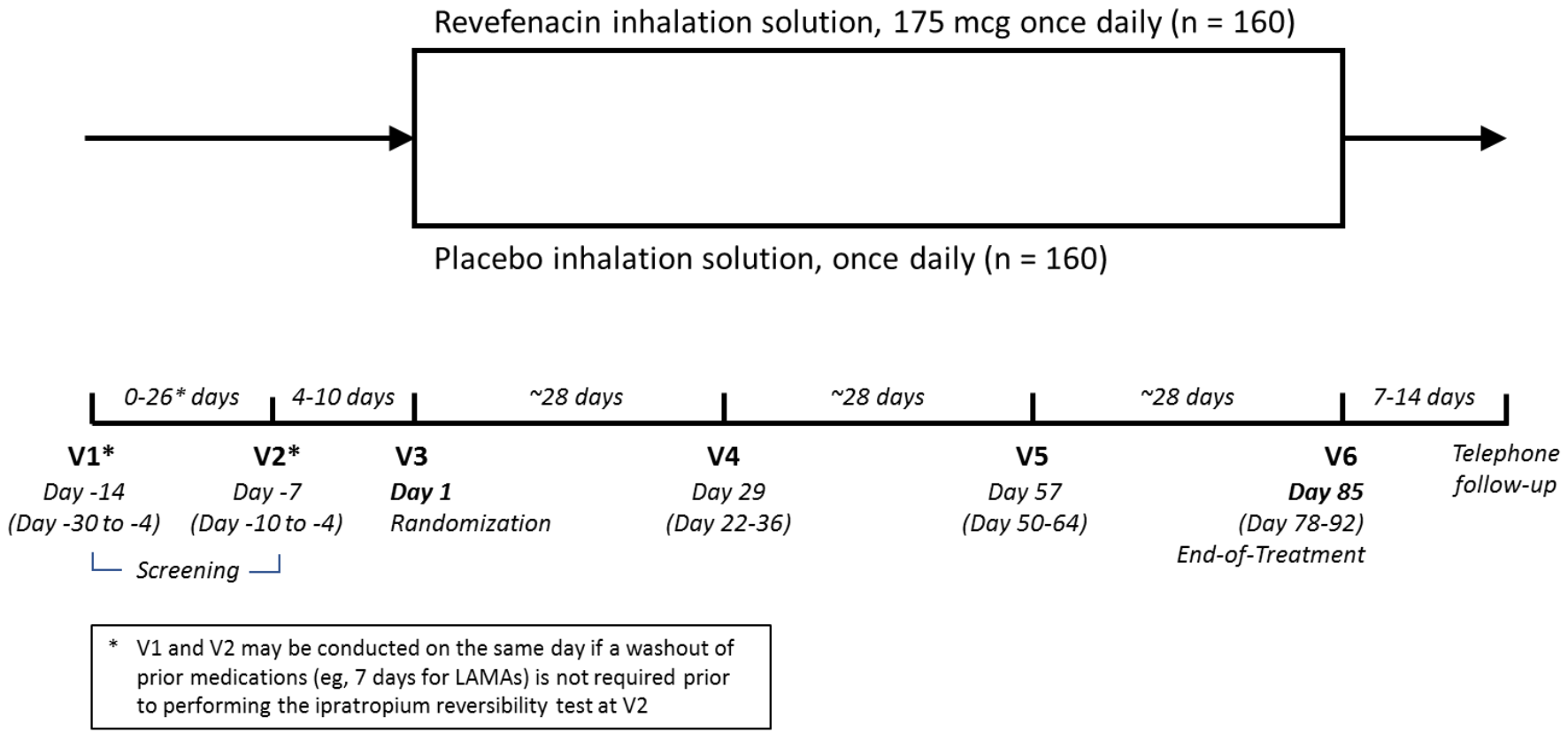
Subjects who are currently using stable doses of long-acting beta-adrenoceptor agonist (LABA)-containing products (LABA monotherapy or inhaled corticosteroid (ICS)/LABA fixed dose combination) will be permitted to enroll into the study and continue to take their LABA or ICS/LABA COPD medication. All other prohibited COPD medications will be required to be stopped. Salbutamol as rescue medication is permitted.

Eligible subjects will be centrally randomized to receive revefenacin 175 mcg or placebo in a 1:1 ratio, stratified by concomitant LABA use (Yes, No) and reversibility to ipratropium (Yes, No).

Each subject will receive treatment once daily in the morning for a total of 12 weeks. The two treatments will both be administered as a 3 mL solution by inhalation using a jet nebulizer. Subjects with a concomitant LABA will dose the LABA or ICS/LABA prior to nebulization.

The schematic of the study design is shown in Figure 1.

Figure 1: Study Diagram



3.1.2 Schedule of Activities

The schedule of study procedures is shown in Table 1.

Table 1: Schedule of Activities

VISIT	1 (Screen 1)	2 (Screen 2)[1]	3 Randomization	4	5	6 (EoT)	Early term (ET)	Follow up (Telephone)
NOMINAL DAY (ALLOWABLE WINDOW)	-14 (-30 to -4)	-7 (-10 to -4)	1	29 (22-36)	57 (50-64)	85 (78-92)		7-14 days after V6/ET
NOMINAL WEEK	-2	-1	0	4	8	12		
Written informed consent	x							
IVRS/IWRS transaction	x		x	x	x	x	x	x
Demography and medical history (incl. COPD, smoking, and exacerbation history)	x							
Review concomitant medications	x	x	x	x	x	x	x	x
Physical examination	x					x	x	
Chest X-ray (if not done in past 12 months)	x							
Height and weight	x							
Vital Signs (supine or semi-recumbent blood pressure, pulse rate)	x		x[2]			x[2]	x	
12-lead ECG (supine or semi-recumbent)	x		x[2]			x[2]	x	
Spirometry (FEV_1, FVC) pre-bronchodilator		x						
Spirometry (FEV_1, FVC) 45 min post-ipratropium (500 mcg nebulized)		x						
Laboratory Safety – Blood and Urinalysis	x					x	x	
Pregnancy Test	x		x	x	x	x	x	
Dispense Rescue Medication (salbutamol pMDI)	x		x	x	x			
Dispense Daily Diary	x		x	x	x			

VISIT	1 (Screen 1)	2 (Screen 2)[1]	3 Randomization	4	5	6 (EoT)	Early term (ET)	Follow up (Telephone)
NOMINAL DAY (ALLOWABLE WINDOW)	-14 (-30 to -4)	-7 (-10 to -4)	1	29 (22-36)	57 (50-64)	85 (78-92)		7-14 days after V6/ET
NOMINAL WEEK	-2	-1	0	4	8	12		
Record Adverse Events	x	x	x	x	x	x	x	x
Medication Training	x[3]		x					
Pre-dose Spirometry (FEV_1, FVC), -45 and -15 min pre-dose.			x	x	x	x	x	
Complete BDI/TDI[4]			x	x	x	x	x	
Complete SGRQ			x			x	x	
Complete mMRC			x					
Randomization			x					
Administer Study Medication in Clinic			x	x	x	x		
Dispense Outpatient Study Medication			x	x	x			
Post-dose Spirometry (FEV_1, FVC), 5, 15, and 30 minutes, and 1, 1.5, and 2h			x			x		
PK sampling			x[5]	x[5]	x[5]	x[5]		
Record Rescue Medication and Study Drug in Diary, Review at Clinic Visits		x	x	x	x	x	x	

1. Visit 2 (Screening Visit 2) can be conducted on the same day as Visit 1 (Screening Visit 1) if the subject does not require washout of other COPD medications (e.g., 14 days for LAMAs) prior to the ipratropium reversibility test at Visit 2.
2. Pre-dose and at approximately 1 hour post-dose.
3. Train subject on rescue medication use.
4. BDI is performed at V3 and TDI is performed at V4, V5, and V6.

5. PK sampling will be performed pre-dose (within 30 minutes of dosing) and at any time 1 – 30 minutes post-dose at Visit 3, 4, 5, and 6. At Visit 6, an additional PK sample will be collected during the sampling interval of 1 – 4 hours post-dose

3.1.2.1 Screening

Screening will involve one or two visits, depending on whether a washout period is required for prohibited COPD medications. The assessments for Visit 1 and Visit 2 may be conducted on the same day if a subject does not require washout of COPD medications prior to the spirometry assessments. Hence, initial screening can occur up to 30 days prior to randomization with key spirometry assessments occurring ~4-10 days prior to randomization.

Prior to the ipratropium reversibility testing during screening, short acting bronchodilators must be withheld for ≥6 hours.

A subject may be re-screened, however, the subject would need to be reconsented and would be assigned a new Subject ID via IWRS.

3.1.2.2 Study Treatment Period

Eligible subjects will be randomized on Day 1 (Visit 3) to double-blind treatment and enter a 12-week placebo-controlled study period where efficacy and safety are planned to be assessed in clinic at Visit 4 (Day 29), Visit 5 (Day 57) and Visit 6 (Day 85).

The first dose of study treatment will be administered on the morning of Day 1/Visit 3 at the study site. Both the start and end times of the administration of study treatment will be captured. Subsequent doses will be self-administered at home by the subject except on days with a scheduled visit to the study site. They will be advised to administer the study drug every morning at approximately the same time within the window of 6 am and 11 am. The time of administration will be recorded by the subjects in their diary.

Spirometry testing will be performed 45 minutes and 15 minutes prior to dosing of study drug at Visits 3-6, and post-dose at 5, 15, 30 minutes and 1, 1.5 and 2 hours after completion of nebulized dose of study drug at V3 and V6 only.

Salbutamol as a rescue medication is permitted but must be withheld for ≥6 hours before the first spirometry performed at each study visit and whilst serial spirometry is performed to 2 hours post-dose on Day 1 and Day 85. Use of salbutamol as a rescue medication will be documented in a medication diary and recorded in the eCRF.

Unless consent is withdrawn, subjects who prematurely terminate study drug will have an early termination (ET) visit scheduled as soon as possible after their last dose of study drug.

3.2 Endpoints

3.2.1 Efficacy Variables

In the following section, the efficacy endpoints have been defined more precisely than they were in the protocol. Therefore, definitions may differ slightly from the protocol. For example, "Change from baseline" is included in the primary endpoint definition.

- Primary endpoint
  - Change from baseline (CBL) Trough FEV_1 on Day 85
- Secondary endpoint

- CBL to Peak FEV_1 (0-2h) on Day 1
- Number (%) of St. George's Respiratory Questionnaire (SGRQ) responders (total score ≥4 units) on Day 85
- Transition Dyspnea Index (TDI) focal score on Day 85
- Average count of salbutamol puffs per day (Days 1-85)
- Percentage of salbutamol rescue free 24h periods (Days 1-85)
- CBL Trough FEV_1 on Days 29 and 57
- TDI focal score on Days 29 and 57
- CBL in Weighted Mean FEV_1 (0-2h) on Days 1 and 85
- CBL to Peak FEV_1 (0-2h) on Day 85
- CBL Trough FVC on Days 29, 57 and 85
- CBL SGRQ total score on Day 85
- Number (%) of TDI responders (focal score ≥1) on Days 29, 57 and 85

3.2.2 Safety Variables

- Adverse Events (AEs)
- Acute Exacerbation of COPD (AECOPD)
- Laboratory Tests
- Vital Signs (blood pressure and pulse rate)
- 12-lead ECG

3.2.3 Pharmacokinetic Variables

Pharmacokinetic (PK) data for Population PK analyses will be collected. The plans for the analysis and reporting of this data are outside the scope of this SAP and will be separately reported.

PK sampling will be performed pre-dose (within 30 minutes of dosing) and at any time 1 – 30 minutes post-dose at Visit 3, 4, 5, and 6. At Visit 6, an additional PK sample will be collected during the sampling interval of 1 – 4 hours post-dose.

4 STATISTICAL METHODS

4.1 Data Quality Assurance

All tables, figures and data listings to be included in the report will be independently checked for consistency, integrity and in accordance with standard [redacted] procedures.

4.2 Randomization and Unblinding Processes

Assignment of Subject Identification number (SID), randomization number and study drug, as well as site drug inventory control will be managed by an automated Interactive Web Response system (IWRS). A manual containing complete instructions for Web access and use will be provided to each site prior to study start. At their first clinic visit, the IWRS will assign a SID. Each SID will be unique and serve as the primary subject identifier throughout

all phases of the study. The SID must appear on all CRF pages, source documents, laboratory data, central spirometry, ECG, diary data, and PRO data. Subjects qualifying to enter the drug treatment phase, will be assigned an additional "randomization number" by the IWRS at randomization. Subjects will be allocated to treatment within blocks for the given allocation ratio.

At Visit 3 (Day 1 of dosing), eligible subjects will be randomized to receive revefenacin 175mg or matching placebo in a 1:1 ratio through the IWRS. The randomization will be stratified according to the subject's concomitant LABA use (Yes, No) and reversibility to ipratropium (Yes, No).

Emergency subject unblinding will be managed by an automated IWRS. A manual containing complete instructions for Web or telephone access and use will be provided to each site prior to study start.

Subjects may be allowed to continue on the study drug after emergency unblinding if the subject does not meet the criteria for study drug termination and/or withdrawal from the study or were accidentally unblinded not due to the reason of safety. These subjects will be considered as protocol deviations and excluded from the relevant analysis populations.

4.3 General Presentation Considerations

4.3.1 Treatment Arms and Labels

In tables, figures and listings, the following treatment labels or abbreviations will be used, depending on the column space available:

- Revefenacin 175 mcg; Placebo
- REV 175 mcg; Placebo

4.3.2 Definitions and Derivations

Baseline Assessments

Unless otherwise specified in Section 4.6, baseline is defined as the last non-missing observation before the first dose of study drug. Both date and time will be included in the comparison for the determination of baseline. For assessments that are required to be done on the date of first dose and pre-dose, but the time of the assessment is not collected, the last non-missing observation on the date of first dose will be considered as baseline assessment.

Reversibility

Reversibility to ipratropium is defined as post-bronchodilator FEV_1 change of ≥12% and ≥200 mL relative to the pre-bronchodilator FEV_1. Reversibility (Yes, No) is captured directly in the CRF as part of the randomization information.

GOLD Categories

GOLD 2019 Severity of Airflow Limitation Categories (GOLD 1-4) and GOLD 2019 categories (GOLD A-D) are defined in Table 2 and Table 3 below.

GOLD 1-4 will be derived from the percent predicted FEV_1 (%) collected at screening Visit 2. The modified Medical Research Council Dyspnea Scale (mMRC) captures the description of breathlessness on a 0-4 scale and will be collected in the CRF and reported. GOLD A-D will be as collected on the CRF at screening as part of COPD History.

Table 2: GOLD Severity of Airflow Limitation Categories

GOLD airflow category	Severity	FEV_1 threshold
GOLD 1	Mild	≥ 80% predicted
GOLD 2	Moderate	≥ 50%, <80% predicted
GOLD 3	Severe	≥ 30%, < 50% predicted
GOLD 4	Very severe	< 30% predicted

Table 3: GOLD 2019 Categories

Moderate or Severe Exacerbation History	Assessment of Symptoms / Risk of Exacerbations	
≥2 or ≥1 leading to hospital admission	**C**	**D**
0 or 1 (not leading to hospital admission)	**A**	**B**
	mMRC 0-1	mMRC ≥2
	Symptoms	

COPD Duration (years)

COPD duration is calculated as (date of informed consent – date of first diagnosis of COPD + 1) ÷ 365.25.
In case of missing or partial diagnosis date, the imputation rules below will be applied:

- If day is missing and the month and year are available, the diagnosis date is imputed as the mid day of the month (eg 15th).
- If day and month are missing and only a year is available, the diagnosis date is imputed as the mid day of the year (eg 1st July).
- If both month and year are missing, no imputation will be done.

The number of years smoked

As per the CRF completion guidelines, years of smoking is derived from only the YEAR component of a date. For ex-smoker, calculated as year stopped smoking – year started smoking; for current smoker, calculated as year of informed consent – year started smoking.

Study Day

When study day is used, the following algorithm will be used:

- study day = date of assessment – date of first dose of study medication + 1, if the date of the assessment is on or after the date of first dose;

- study day = date of assessment – date of first dose of study medication, if the date of the assessment is prior to the date of first dose.

4.3.3 Analysis Windows

The summary tables will only include the assessments at scheduled visits, except the spirometry measurements. Unscheduled assessments will be included in the listings.

Unscheduled spirometry measurements with study medication taken on the same day will be mapped to corresponding visit and timepoint according to Table 4 and Table 5. Unscheduled spirometry measurements with no study medication taken on the same day will not be mapped. Spirometry measurements at an early termination visit will be mapped to the schedule visit according to Table 4.

If multiple visits fall within a window, the earliest visit with acceptable spirometry data nearest to the scheduled day will be included in the analysis.

Table 4: Analysis Windows (Days)

Visit (Day) in Protocol	Analysis Visit	Target Study Day	Visit Window
Visit 3 (1)	Day 1 (Baseline)	1	1
Visit 4 (29±7)	Day 29	29	2≤ Study Day ≤ 43
Visit 5 (57±7)	Day 57	57	44 ≤ Study Day ≤ 71
Visit 6 (85±7)	Day 85	85	72 ≤ Study Day

Table 5: Analysis Windows (Timepoint)

Analysis Timepoint	Time Window
-45 Min	Timepoint < -30 Min
-15 Min	-30 Min ≤ Timepoint < 0 Min
5 Min	0 Min < Timepoint ≤ 10 Min
15 Min	10 Min < Timepoint ≤ 22 Min
30 Min	22 Min < Timepoint ≤ 45 Min
1 H	45 Min < Timepoint ≤ 75 Min
1 H 30 Min	75 Min < Timepoint ≤ 1H 45 Min
2 H	1H 45 Min < Timepoint

4.3.4 General Reporting Conventions

Continuous data will be summarized in terms of the mean, standard deviation (SD), median, minimum, maximum and number of observations (n), unless otherwise stated. The minimum and maximum will be reported to the same number of decimal places as the raw data recorded in the database. The mean and median will be reported to one more decimal place than the raw data recorded in the database. The standard deviation will be reported to two more decimal places than the raw data recorded in the database. In general, the maximum number of decimal places reported shall be four for any summary statistic.

Categorical data will be summarized in terms of the number of subjects providing data at the relevant time point (n), and percentages. Any planned collapsing of categories will be detailed in the SAP text and the data displays. Percentages will be presented to one decimal place. Percentages will not be presented for zero counts. Percentages will be calculated using the number of subjects in the analysis set (N) as the denominator.

Descriptive statistics for the 12-week study period will be presented by treatment group and visit, as applicable.

All statistical tests will be two-sided and will be conducted at the 5% significance level, unless otherwise specified.

P-values greater than or equal to 0.0001, in general, will be presented to four decimal places. P-values less than 0.0001 will be presented as "<0.0001".

4.3.5 Data Standards

The study data and analysis datasets will conform to the Clinical Data Interchange Standards Consortium (CDISC) standards including therapeutic area data standards for COPD. General standards include the Study Data Tabulation Model (SDTM) and Analysis Data Model (ADAM) and associated implementation guides.

4.3.6 Software

All report outputs will be produced using SAS® version 9.4 for UNIX in a secure and validated environment in accordance with [redacted] procedures.

4.4 Study Subjects

4.4.1 Disposition of Subjects

The total number of subjects screened and the reasons for failing screening will be reported.

Subject disposition information will be summarized for all randomized subjects by treatment group and overall. The summaries will include the following information:

- Subjects randomized
- Subjects randomized and Treated with Study Drug
- Subjects completed study treatment
- Subjects early terminated from study treatment (including by reason for early termination of study treatment, and if the reason for early termination of study treatment is related to public health emergency)
- Subjects completed study
- Subjects discontinuing from study (including by reason for discontinuation of study, and if the reason is related to public health emergency)

Data related to disposition (including the date of informed consent signed, the dates of first dose and last dose of study drug, primary reason for subject discontinuation of study treatment, the date of last visit, study completion status) will be listed.

Visit status (performed on time, out of window, missed), type of visit (on site, telephone), the impact of public health emergency to the visit will be summarized and listed.

4.4.2 Protocol Deviations

Major analysis protocol deviations that could potentially affect the efficacy conclusions of the study will be identified prior to database lock. Major analysis protocol deviations may include, but are not limited to:

- Randomized subjects who did not satisfy inclusion criteria
- Randomized subjects who did not satisfy exclusion criteria
- Subjects whose randomized and actual treatments differ at any time point
- Subjects who received an excluded concomitant treatment or medication
- Subjects who are not 80-120% compliant with study medication
- Subjects who received rescue medication within 6 hours prior to spirometry testing

Subjects with major analysis protocol deviations will be identified before the database lock and provided in a listing of subjects excluded from per-protocol (PP) efficacy analyses. In addition, a listing of all major deviations will be provided whether they impact the analysis or not. All decisions to exclude subjects from the PP analysis set will be made prior to unblinding in the final blinded data review (BDR) meeting and report.

The PD summaries will include the following information, using Randomized Set:

- the number and percentage of subjects with a major PD by treatment group and overall and by type of deviation
- the number and percentage of subjects with a major PD leading to exclusion from the PP by treatment group and overall, and by type of major deviation leading to exclusion
- the number and percentage of subjects with a major PD related to COVID-19 by treatment group and by type of deviation

4.5 Analysis Sets

The **Screened Set** consists of all subjects who sign the study informed consent. Screen failures are subjects who sign the study informed consent but fail to meet study requirements during screening or prior to randomization.

The **Randomized Set** consists of all subjects that are randomized to a treatment on Day 1. Subjects in this population will be reported according to the treatment they were randomized to, regardless of any dosing error. The Randomized Set will be predominantly used for listings of data.

The **Safety Analysis Set (SS)** is defined as all subjects who receive at least one dose of study drug. Subjects in this population will be summarized according to the actual treatment they received. The treatment received will only be different from the randomized treatment if the subject took the wrong dose or study drug for > 50% of the entire double-blind treatment period. All safety analyses will be performed using the SS.

The **Full Analysis Set (FAS)** consists of all randomized subjects who received at least one dose of study drug and have at least one recorded post-baseline efficacy assessment. Subjects in this set will be analyzed and summarized according to their assigned treatment and strata at randomization. The FAS will be used for all efficacy endpoints.

The **Per-Protocol (PP) Analysis Set** consists of subjects in the FAS who had no major protocol violation that would impact significantly on the primary endpoint. The list of major protocol deviations will be finalized prior to database lock. All decisions to exclude subjects from the PP analysis set will be made prior to unblinding in the final blinded data review (BDR) meeting and report. Data will be summarized and analyzed according to the treatment a subject actually received and strata they actually belong to. The PP analysis set will be used for the primary endpoint only.

Analysis Set information will be summarized for all randomized subjects by treatment group and overall. The table will include rows for the number (%) of subjects randomized, randomized but not treated, Safety Set (SS), Full Analysis Set (FAS), Per-Protocol (PP) Analysis Set.

4.6 Demographic and Other Baseline Characteristics

4.6.1 Demographics

Baseline demographics are collected at the Screening Visit 1. Descriptive statistics will be prepared for the demographics by treatment group and overall using the Full Analysis Set (FAS) population. If the Per-Protocol (PP) Analysis Set is different to FAS, a separate demographics table will be included for this population as well. If the Randomized Set is different to FAS, demographics will also be reported for Randomized Set. The following information will be included in the summary:

- Age in years
- Age group (<65 years, ≥65 years)
- Sex (Male, Female)
- Race (Asian - Chinese, Asian - Not Chinese)
- Height
- Weight
- Body mass index (kg/m^2)

Demographic data will be listed using the Randomized Set.

4.6.2 Baseline Smoking, Alcohol and Drug History

Baseline smoking, alcohol and drug history is collected at Screening Visit 1 and will be summarized by treatment group and overall (FAS population). The endpoints will include:

- Smoking status (Current, Former)
- The number of years smoked (derived)
- The number of pack-years smoked (as collected)
- Subject with history of alcohol abuse within the past 5 years (Yes, No)
- Subject with history of illegal or recreational drug within the past 5 years (Yes, No)

Smoking, alcohol and drug history will be listed using the Randomized Set.

4.6.3 Baseline Spirometry

Baseline spirometry data will be summarized by treatment group and overall (FAS population). Data will be collected at Screening Visit 2 for the following endpoints:

- pre- and post-bronchodilator FEV_1 (L);
- reversibility to ipratropium (Yes, No as collected);
- predicted normal FEV_1 (L) and percent predicted FEV_1 (%)
- pre- and post-bronchodilator FVC (L);
- pre- and post-bronchodilator FEV_1/FVC ratios.

All baseline spirometry data will be listed using the Randomized set. If the Randomized Set is different to FAS, demographics will also be reported for the FAS.

Data for baseline FEV_1 and FVC assessments used in the derivation of the primary and secondary endpoints will be collected at -45 and -15 minutes prior to dosing of study medication on Day 1 (Randomization Visit 3).

Baseline is defined as the average of the -45 and -15 minute measurements. If either of the two pre-dose measures is missing, the baseline will be set to the value of the remaining non-missing measure. If both pre-dose values are missing then the baseline will be set to missing.

This baseline spirometry data will be presented in tables with the primary and secondary spirometry endpoints.

4.6.4 Baseline COPD Disease-Related Characteristics

Data on COPD disease-related characteristics will be summarized (FAS population) by treatment group and overall. Presentations will include:

- COPD Duration in years (derived);
- Number of COPD Exacerbations in past 12 months;
- Number of COPD Exacerbations requiring hospitalization in previous 12 months
- Concomitant ICS use (Yes, No as collected);
- Concomitant LABA use (Yes, No as collected);
- Concomitant ICS/LABA use (Yes, No), derived from Concomitant ICS use and Concomitant LABA use;
- GOLD Airflow Categories 1-4 (derived);
- GOLD Categories A-D (as collected).

COPD Disease-Related Characteristics will be listed using the Randomized Set.

4.6.5 Patient Reported Outcomes (PROs)

Baseline data for PROs will be collected pre-dose on Day 1 (Randomization Visit 3) and summarized and listed to include:

- Proportion of subjects with baseline mMRC ≥2 (n, %)
- Baseline SGRQ component scores and total score
- BDI component scores and focal score

4.7 Medical History

The medical history events except for COPD history will be coded by System Organ Class (SOC) and Preferred Term (PT) using the Medical Dictionary for Regulatory Activities (MedDRA). The MedDRA version used for reporting the study will be described in the relevant table and listing footnotes. The number and percentage of subjects in each SOC and PT will be summarized based on FAS by treatment group and overall.

Each subject's medical history will be listed by SOC and PT for the Randomized Set.

4.8 Prior and Concomitant Medication

Data on prior and concomitant medication will be summarized based on the Safety Set by treatment group and overall.

Any medication taken prior to the first dose of double-blind study treatment will be documented as prior medication. Any medication taken after the first dose of study drug will be documented as a concomitant medication.

For partial/missing start/end date of the medication, the imputation rule below will be applied before the summary:

- Start date:
  - If Year and Month exists but Day is missing, impute the Day as 1;
  - If Year exists but Month and Day are missing, impute the Month and Day as January 1st;
  - If Year is missing or the entire date is missing, do not impute but assume the medication started after the first dose and to be considered as a concomitant medication.
- End date:
  - If Year and Month exists but Day is missing, impute the Day as the last day of the month;
  - If Year exists but Month and Day are missing, impute the Month and Day as December 31st;
  - If Year is missing or the entire date is missing, and the status is not ongoing, do not impute but assume the medication ended after the first dose and to be considered as a concomitant medication.

Both prior and concomitant medications will be coded using the World Health Organization Drug Dictionary (WHODRUG). The WHODRUG version used for reporting the study will be described in the relevant table and listing footnotes.

The Anatomical Therapeutic Chemical (ATC) level 1 term and PTs for prior medications will be summarized by treatment group and overall. Concomitant Medications for COPD (e.g.

concomitant bronchodilators, concomitant corticosteroids) will be identified and summarized in a similar manner in separate tables.

Each subject's prior and concomitant mediation will be listed by ATC level 1 and PT for the Randomized Set.

4.9 Treatment Compliance

Subjects will take nebulized study drug (3mL sealed plastic vial) once daily during the 12-week double-blind placebo-controlled treatment period.

The number of dispensed, used, unused and lost vials will be counted and recorded as part of the drug accountability. The primary assessment of compliance will be based on the vial counts.

Compliance is based on the count of number of study drug vials taken as recorded on the Drug Accountability eCRF page. Treatment compliance in percent will be calculated as:

Compliance (%) = (Number of vials taken / Number of expected vials) × 100

where Number of expected vials = 1 × study drug exposure

Details of study drug exposure are described in Section 4.11.1.

The compliance will be classified (e.g. <80%, 80%-120%, >120%). If the compliance is outside of 80%-120%, a subject is considered non-compliant. The number and percentage of subjects in the above categories will be summarized based on Safety Set by treatment group and overall.

4.10 Efficacy Evaluation

For all efficacy analyses the Full Analysis Set (FAS) will be used unless otherwise specified. Efficacy will be assessed based on spirometry, subject-reported outcomes (SGRQ, BDI/TDI) and rescue medication usage. A summary of the efficacy variables used in the statistical analyses is given in Table 6.

Table 6: Summary of Efficacy Endpoints

Efficacy Variable		**Analysis Set**	**Objective**	**Visit Day**	**Data Type**	**Analysis Model**
Spirometry	CBL Trough FEV_1	FAS	Primary	Day 85	Continuous	MMRM (MAR)
	CBL Trough FEV_1	PP	Supportive	Day 85	Continuous	MMRM (MAR)

	CBL Trough FEV_1	FAS	Sensitivity	Day 85	Continuous	ANCOVA (LOCF)
	CBL Trough FEV_1	FAS	Sensitivity	Day 85	Continuous	MMRM (Tipping Point MI)
	CBL Trough FEV_1	FAS	Sensitivity	Day 85	Continuous	MMRM (Subgroup Analyses)
	CBL Trough FEV_1	FAS	Secondary	Day 29, 57	Continuous	MMRM
	CBL Trough FVC	FAS	Secondary	Day 29, 57 and 85	Continuous	MMRM
	CBL Peak FEV_1 0-2h	FAS	Secondary	Day 1	Continuous	ANCOVA
	CBL Peak FEV_1 0-2h	FAS	Secondary	Day 85	Continuous	ANCOVA
	CBL Weighted Mean FEV_1 0-2h	FAS	Secondary	Day 1, 85	Continuous	ANCOVA
SGRQ	SGRQ responder	FAS	Secondary	Day 85	Binary	Logistic Regression
	CBL SGRQ total score	FAS	Secondary	Day 85	Continuous	ANCOVA
	CBL SGRQ component scores (3)	FAS	Other	Day 85	Continuous	ANCOVA
BDI/TDI	TDI focal score	FAS	Secondary	Day 85	Continuous	MMRM
	TDI focal score	FAS	Secondary	Day 29, 57	Continuous	MMRM
	TDI component scores (3)	FAS	Other	Day 29, 57, 85	Continuous	MMRM

	TDI responder	FAS	Secondary	Day 29, 57, 85	Binary	MMRM logistic model
Rescue Medication Usage	Average Count of Salbutamol Puffs Per Day	FAS	Secondary	Month 1,2,3 and Overall (Day1-85)	Continuous	MMRM
	Percentage of monthly salbutamol rescue free 24-hour periods	FAS	Secondary	Month 1,2,3 and Overall (Day1-85)	Continuous	MMRM

4.10.1 Analysis and Data Conventions

4.10.1.1 Multicenter Studies

This multicenter study utilizes a central randomization stratified by concomitant LABA use and reversibility to ipratropium. All centers will be in China. Center will not be included as a fixed effect covariate in the statistical analyses due to the small numbers at each site.

4.10.1.2 Adjustments for Covariates

As the randomization is stratified by concomitant LABA use (Yes, No) and reversibility to ipratropium (Yes, No), these will be included as covariates in all analysis models.

Other factors to be included in the statistical models for efficacy as fixed-effect class terms are smoking status (current smoker, ex-smoker); age category (<65, ≥65); Sex (Male, Female).

A covariate for baseline FEV_1 will be included in all FEV_1 analysis models. Full details of the covariates used in each statistical model are outlined by endpoint in Sections 4.10.2 and 4.10.3.

4.10.1.3 Handling Missing Data

Missing or unavailable data will be queried by Data Management. Data that remains missing will be identified in summaries and analyses of that data item.

For any subject who withdraws prematurely from study treatment, all available data up to the time of discontinuation of study treatment will be considered for inclusion in the analyses. An Early Termination visit will occur as soon as possible after their last dose of study drug.

The following graphical representations are examples of plots that may be used to explore the data and help to explain any dropout patterns. The actual plots used will be dependent on the final dropout rate/pattern for the study:

- Kaplan-Meier plots of time to discontinuation of study treatment split by treatment group.
- Kaplan-Meier plots of time to discontinuation of study treatment split by treatment group and reason for discontinuation (e.g. Lack of Efficacy, Adverse Events).
- Kaplan-Meier plots of time to discontinuation of study treatment split by treatment group and relationship to public health emergency (related, not related).

Imputing Missing Data

Missing data will only be imputed in the sensitivity analyses for the primary analysis of the primary endpoint (change from baseline trough FEV_1 at Day 85). Otherwise, missing data will not be imputed.

Missing data in the primary efficacy analysis is not directly imputed and is assumed to be missing at random (MAR). In sensitivity analyses the following approaches will be used for imputing missing data:

Non-monotone missing data (e.g. subjects have missing values for intermediate visits, but have available data at subsequent visits) will be assumed to be MAR. For monotone missing data the following missing data imputation methods will be used to assess the robustness of the conclusions to the MAR assumption since it may be biased and attribute treatment benefit to withdrawals even though they are no longer taking the study treatment:

- Last-Observation-Carried Forward (LOCF) single imputation method
- Two-dimensional Tipping-point multiple imputation method

LOCF Analysis

Missing trough FEV_1 data on Day 85 will be imputed using LOCF single imputation method. Changes from baseline on Day 85 that include the LOCF imputed values will then be analyzed using an Analysis of Covariance (ANCOVA) model.

Tipping-Point Analysis

A Tipping-point analysis will be performed. Monotone missing data at all missing scheduled visits will have the value of a shift parameter (e.g. 50 mL) subtracted from the revefenacin treatment arm (worsening results) and added to the placebo arm (improving results) for change from baseline trough FEV_1. The imputed datasets will then be analyzed using the MMRM primary analysis method and an estimate for the difference between the two treatments and associated 95% CI and p-value calculated. This method will be applied iteratively where at each iteration, monotone missing values are adjusted with a larger shift parameter than at the previous iteration (e.g. by 50 mL from 0 mL to -500 mL on revefenacin and 0 mL to +500 mL on placebo). This process continues until the difference between revefenacin and placebo is no longer statistically significant. This point is called the tipping point. The plausibility of the value of the shift parameters needed to change a statistically significant result ($p<=0.05$) to not significant ($p>0.05$) will be clinically assessed.

Details of tipping-point multiple imputation method are described can be found in Appendix.

4.10.1.4 Multiple Comparisons

No adjustment for multiple comparisons is required.

4.10.1.5 Examination of Subgroups

The intent of the subgroup analyses is to demonstrate the consistency of response of a treatment across a wide range of the intended treated population. The subgroups of interest are defined below:

- Baseline Smoking status [current smoker, former smoker]
- Age group [<65 years, ≥65 years]
- Sex [Male, Female]
- Reversibility to ipratropium [yes, no]
- Concomitant ICS use [yes, no]
- Concomitant LABA use [yes, no]
- Baseline post bronchodilator % predicted FEV_1[50% to <80%, 30% to <50%, <30%] (GOLD airflow categories [2, 3, 4])
- GOLD categories [A, B, C, D]
- Ethnicity [North East Asian, South East Asian], provided by external data vendor of spirometry assessments.
  North East Asian defined as China north of the Huaihe River and Qinling Mountains.
  South East Asian defined as China south of the Huaihe River and Qinling Mountains.

4.10.1.6 Interim Analyses

Not Applicable

4.10.2 Primary Efficacy Variable – CBL Trough FEV_1 on Day 85

The primary efficacy variable is the change from baseline trough FEV_1 on Day 85.

- Baseline FEV_1 is defined as the average of the -45 and -15 minute measurements prior to dosing of study medication on Day 1 (Visit 3).

- Trough FEV_1 measurements will be the pre-dose FEV_1 measures on Days 29, 57 and 85 calculated as the average of the -45 and -15 minute measurements prior to dosing of study medication.

- If a public health emergency prevents FEV_1 data from being measured at site then information on the reasons for the missing assessments will be recorded in the CRF

If either of the two pre-dose FEV_1 measures is missing, the baseline/trough will be set to the value of the remaining non-missing measure. If both pre-dose values are missing then the baseline/trough will be set to missing.

If no baseline can be calculated for the subject, then the change from baseline values will be recorded as missing. Change in FEV_1 is calculated as the Trough FEV_1 – Baseline FEV_1 difference such that a positive change indicates improvement.

The absolute values and the change from baseline trough FEV_1 will be summarized by treatment group and visit using summary statistics for continuous data.

4.10.2.1 Primary Analysis

The analysis of the primary endpoint will be based on a mixed model repeated measures (MMRM) analysis including changes from baseline trough FEV_1 measurements on Days 29, 57 and 85 as the dependent variable. The model will include independent fixed-effect terms for concomitant LABA use, reversibility to ipratropium, smoking status, age, sex, treatment and baseline FEV_1 as a covariate. A time effect (visit) and its interaction terms with treatment and baseline FEV_1 will be included. The variance estimation will be based on an unstructured covariance matrix. The Kenward-Roger method for approximating the denominator degrees of freedom will be used.

The primary analysis of the primary efficacy endpoint will be based on the Full Analysis Set (FAS) population and includes all randomized subjects who received at least one dose of study drug and have at least one recorded post-baseline FEV_1 assessment. Available data for a subject prior to discontinuing study treatment will be used in the MMRM analysis. Missing data is assumed as missing at random (MAR). The robustness of the conclusions to this assumption will be assessed with a missing data sensitivity analysis.

The difference between revefenacin and placebo on Day 85 will be estimated from the least squares means (LS means) along with the 95% confidence interval (CI) and associated 2-sided p-values. As a supportive analysis, the primary efficacy analysis will be performed using the Per-Protocol (PP) Analysis Set.

4.10.2.2 Primary Endpoint Sensitivity Analysis

Sensitivity analyses using missing data imputation methods are included to assess the robustness of the primary estimate with regard to missing data.

Multiple Imputation (MI) Tipping Point Analysis

A sensitivity analysis using a two-dimensional tipping point imputation method will be applied to the MMRM analysis (FAS population) to assess the robustness of the conclusions to the assumption that missing data is MAR. In the tipping point analysis a succession of delta adjustments via multiple imputation methods will be used to impute the missing revefenacin and placebo trough FEV_1 values at each timepoint. The tipping point is denoted as the mean trough FEV_1 value that would have to be imputed to overturn a statistically significant result. Clinical judgement will be applied to the plausibility of the tipping point. The detail of this imputation method and example SAS code is given in Appendix.

Single Imputation Last-Observation-Carried-Forward (LOCF) Approach

The LOCF method described in Section 4.10.1.3 will be used for subjects with missing trough FEV_1 data on Day 85. Change from baseline trough FEV_1 at Day 85 will be analyzed

using an analysis of covariance (ANCOVA) using the FAS population. The model will include treatment, and concomitant LABA use, bronchodilator reversibility to ipratropium, smoking status, age, sex, baseline FEV_1 as covariates.

Least squares means for Day 85 will be derived for both treatment groups. The difference between revefenacin and placebo on Day 85 will be estimated from the LS means along with the 95% CI and associated 2-sided p-values.

4.10.2.3 Primary Endpoint Subgroup Analysis

An assessment of whether the effect of treatment on the primary endpoint is modified by each Subgroup of Interest (see Section 4.10.1.5) will be made by fitting separate MMRM models, identical to the model described for the primary analysis, but including additional terms for the subgroup and treatment by subgroup interaction.

Based on this MMRM model, LS means and standard errors for treatment within subgroup and LS means and two-sided 95% CIs for treatment differences within subgroup will be provided.

LS mean differences between groups (calculated as Revefenacin – placebo) and associated CI will also be presented by Forest Plot for each subgroup and categories within the subgroup.

The number of subjects in each treatment group and subgroup will be reviewed. Subgroups may need to be regrouped or excluded from analyses where insufficient numbers exist. The reasons why subgroup analyses are not performed will be documented.

4.10.3 Secondary Efficacy Variables

In all the secondary efficacy analyses the FAS population will be used and missing data will not be imputed and will be assumed to be missing at random (MAR) unless otherwise stated. If a public health emergency prevents data from being measured or requires data to be assessed remotely via a virtual visit, then information on the reasons and the data impacted will be recorded in the CRF.

4.10.3.1 CBL Trough FEV_1 and FVC

The changes from baseline trough FEV_1 on Days 29 and 57 will be estimated from the MMRM analysis used for the primary analysis of the primary endpoint.

The changes from baseline trough FVC on Days 29, 57 and 85 will be analyzed using the same MMRM analysis described for the primary analysis of the primary endpoint. The baseline FEV_1 covariate will be replaced with the baseline trough FVC recorded on Day 1.

4.10.3.2 CBL to Peak FEV_1 (0-2h) on Days 1 and 85

Lung function assessments will be made after study treatments (5, 15, and 30 minutes, and 1, 1.5, and 2 hours) on Day 1 and Day 85. Peak FEV_1 is defined as the highest post-dose FEV_1 value within 2 hours after the dosing on Day 1.

Peak FEV_1 on Day 1 and Day 85 will be summarized by treatment group for the actual value and for the change from baseline on Day 1.

An ANCOVA model will be used to assess the difference between treatment groups for the change from baseline to Peak FEV_1 after the first dose on Day 1. The ANCOVA model will include treatment, and concomitant LABA use, reversibility to ipratropium, smoking status, age, sex, baseline FEV_1 as covariates.

The difference between revefenacin and placebo on Day 1 will be estimated from the LS means along with the 95% CI and associated 2-sided p-values.

A separate analysis will be performed for Peak FEV_1 on Day 85 using the same ANCOVA model.

4.10.3.3 CBL in Weighted Mean FEV_1 (0-2h) on Days 1 and 85

The area under the curve (AUC) [0 to 2] hours will be calculated using the trapezoidal rule, including observed data points from the pre-dose (0 hour) through to the 2 hour observation. For all observations the nominal time of the observation should be used. For the calculation of AUC, the first (0 hour), last (2 hour) and at least one time point between first and last must be available. Otherwise, the AUC is set to missing. The Baseline FEV_1 on Day 1 and Trough FEV_1 on Day 85 will be used as the 0 hour data point to calculate the AUC for Day 1 and Day 85, respectively.

The weighted mean (WM) FEV_1 is derived from dividing AUC by the total time (nominally 2 hours).

An ANCOVA model will be used to assess the difference between treatment groups for change from baseline in WM FEV_1 after the first dose on Day 1. The ANCOVA model will include treatment, and concomitant LABA use, reversibility to ipratropium, smoking status, age, sex, baseline FEV_1 as covariates.

The difference between revefenacin and placebo on Day 1 will be estimated from the LSmeans along with the 95% CI and associated 2-sided p-values.

A separate analysis will be performed for WM FEV_1 on Day 85.

4.10.3.4 CBL SGRQ Total Score on Day 85

The SGRQ total score and 3 component scores, including symptoms, activity, and impacts, will be derived. Baseline SGRQ endpoints will be those recorded on Day 1 (Visit 3). The baseline and Day 85 absolute values and the change from baseline in SGRQ total score and 3 component scores will be summarized by treatment group (continuous data).

An ANCOVA model will be used to assess the difference between treatment groups for change from baseline SGRQ total score on Day 85. The ANCOVA model will include treatment, and concomitant LABA use, reversibility to ipratropium, smoking status, age, sex, baseline SGRQ total score as covariates. For subjects who discontinue study treatment, the data from the SGRQ completed at the Early Termination (ET) visit will be carried forward in the analyses.

The difference between revefenacin and placebo on Day 85 will be estimated from the LSmeans along with the 95% CI and associated 2-sided p-values.

The three SGRQ components will be similarly analyzed.

4.10.3.5 Number (%) of SGRQ Responders (Total Score ≥4 Units) on Day 85

A decrease from baseline of ≥4 units in SGRQ total score is considered clinically meaningful and will be defined as a responder.

The number and percentage of responders on Day 85 in each treatment group will be summarized. A comparison between revefenacin and placebo will be evaluated with the odds ratio (OR) and corresponding 95% confidence interval from a logistic regression model. The model will include fixed-effect terms for concomitant LABA use, reversibility to ipratropium, smoking status, age, sex, treatment and baseline SGRQ total score as a covariate.

For subjects who discontinue study treatment, the SGRQ completed at the Early Termination (ET) visit will be carried forward and used to derive a responder.

4.10.3.6 TDI Focal Score on Days 29, 57 and 85

The BDI and TDI focal scores and associated components will be listed and summarized by treatment and visit using summary statistics. Baseline BDI scores will be recorded on Day 1 (Visit 3). The TDI questionnaire represents changes from baseline and will be recorded on Days 29, 57 and 85.

The BDI/TDI scores include three components: functional impairment, magnitude of task, and magnitude of effort. For BDI scores, each domain can be rated from Grade 0 (very severe) to 4 (no impairment). For TDI scores, each domain can be rated from Grade -3 (major deterioration) to + 3 (major improvement). The focal score is the sum of the three components (ie BDI ranges from 0 to 12 and TDI ranges from -9 to +9).

The TDI focal score will be analyzed using the same MMRM analysis described for the primary analysis of the primary endpoint. The baseline FEV_1 covariate will be replaced with the BDI focal score. Estimates for TDI on Days 29, 57 and 85 will be derived from the model.

The difference between revefenacin and placebo on Days 29, 57 and 85 will be estimated from the LSmeans along with the 95% CI and associated 2-sided p-values.

The three TDI components will be analyzed in a similar way.

Details of BDI/TDI are further described in Section 6.6.

4.10.3.7 Number (%) of TDI Responders (Focal Score ≥1) on Days 29, 57 and 85

A TDI focal score of ≥1 unit is considered clinically meaningful and will be defined as a responder.

The number and percentage of responders on Days 29, 57 and 85 in each treatment group will be summarized. A comparison between revefenacin and placebo will be evaluated with the odds ratio (OR) and corresponding 95% confidence interval from a logistic regression model

including the proportion of responders at each visit (Days 29, 57 and 85) as the dependent variable. The MMRM logistic (binomial distribution with logit link function) model will include the same independent fixed-effect terms as for the primary analysis of the primary endpoint. The baseline FEV1 covariate will be replaced with the BDI focal score as a covariate. Estimates for the proportion of TDI responders on Days 29, 57 and 85 will be derived from the model.

4.10.3.8 Average Count of Salbutamol Puffs Per Day (Days 1-85)

Use of salbutamol as a rescue medication will be documented in a medication diary and recorded in the eCRF. Rescue medication use should be recorded as the number of doses per day, where each puff of salbutamol is one dose.

The average count of rescue salbutamol puffs per day will be calculated for each monthly interval by dividing the total salbutamol puff counts for each of the 3 monthly intervals by the total number of days on study for that month.

> Average count of rescue salbutamol puffs per day = Total salbutamol puff counts / Number of non-missing days during each defined interval

where the "number of non-missing days" is the actual number of days in the interval whilst receiving study treatment.

If a subject withdraws from study treatment early, only the number of days on treatment in the interval will be included in the count of non-missing days and rescue medication information collected after withdrawal will not be included in the calculations.

The baseline average count of rescue salbutamol puffs per day will be calculated using the last 7 daily entries recorded within 14 days prior to randomization on Day 1.

The average count of salbutamol puffs per day will be analyzed using the same MMRM analysis described for the primary analysis of the primary endpoint with the time variable Month (1, 2, 3) replacing Visit.

The average count of salbutamol puffs per day for the study (Days 1-85) will be estimated from the model-derived average of each of the 3 months using an appropriate estimate statement as well as the LS mean differences, 95% CI and associated 2-sided p-values.

Estimates for each one month period (Months 1, 2 and 3) will also be reported.

4.10.3.9 Percentage of Salbutamol Rescue Free 24h Periods (Days 1-85)

A 24h rescue-free period is defined as a day without a "yes" for rescue use per the daily diary.

The percentage of salbutamol rescue free 24h periods will be calculated for each monthly interval by dividing the number of 24h rescue-free periods for each of the 3 monthly intervals by the total number of days on study for that month.

Percentage of salbutamol rescue free 24h periods = Number of 24h rescue-free periods / Number of non-missing days, and multiplying by 100 during each defined interval,

where the "number of non-missing days" is the actual number of days in the interval whilst receiving study treatment.

If a subject withdraws from study treatment early, only the number of days on treatment in the interval will be included in the count of non-missing days and rescue medication information collected after withdrawal will not be included in the calculations.

The baseline percentage of salbutamol rescue free 24h periods will be calculated using the last 7 daily entries recorded within 14 days prior to randomization on Day 1.

The percentage of salbutamol rescue free 24h periods will be analyzed using the same MMRM analysis described for the primary analysis of the primary endpoint with the time variable Month (1, 2, 3) replacing Visit.

The average percentage of salbutamol rescue free 24h periods for the study (Days 1-85) will be estimated from the model-derived average of each of the 3 months using an appropriate estimate statement as well as the LSmean differences, 95% CI and associated 2-sided p-values.

Estimates for each one month period (Months 1, 2 and 3) will also be reported.

4.11 Safety Evaluation

Safety analyses include extent of exposure, adverse event (AE), laboratory test, vital signs, electrocardiogram (ECG) and physical examination.

Safety endpoints will be summarized using the Safety Analysis Set (SS) population. All safety analyses will be presented by treatment group.

There are no planned statistical analyses for the comparison of treatments for the safety endpoints.

4.11.1 Extent of Exposure

Descriptive summaries (mean, SD, median, min, max) will be presented for the number of vials taken and the study drug exposure by treatment group (Safety Set).

Study drug exposure will be calculated as the number of days from first to last dose date of study drug for the treatment period.

Study drug exposure (Days) = Date of last dose – Date of first dose + 1

For subjects who early withdraw, and for whom the last dose date from the End of Treatment eCRF page is unknown, the date of last available in-clinic dosing, date of withdrawal, and/or dosing recorded via the diary will be used, whichever is later.

All data for study drug exposure will be listed.

4.11.2 Adverse Events

All AEs that occur after the first dose of double-blind study medication to 30 days after the last dose of study medication will be considered to be treatment-emergent adverse events (TEAEs).

Where dates are missing or partially missing, adverse events will be assumed to be treatment-emergent, unless there is clear evidence (through comparison of partial dates) to suggest that the adverse event started prior to the first dose of study medication or more than 30 days after the last dose of study medication.

The MedDRA will be used to code the primary SOC and the PT of the AEs. The MedDRA version used for reporting the study will be described in relevant table and listing footnotes.

There are three categories of AE severity: Mild, Moderate and Severe. For adverse events with missing intensity, they are counted in the most conservative way as Severe in the summary by intensity.

There are five categories of AE Relationship to study treatment: Definite, Probable, Possible, Unlikely, Unrelated/Not Related. An adverse event is considered related to the study treatment if the relationship to the study treatment is considered “Definite”, “Probably Related”, “Probable” or ‘Possible’.

All AEs and all TEAEs will be listed by subject.

All AE tables will be restricted to treatment-emergent AEs unless otherwise specified. Each subject will be counted only once within each PT or SOC. If a subject experiences more than one AE within a PT or SOC, only the AE with the strongest relationship or the worst severity, as appropriate, will be included in the summaries of relationship and severity. AEs with missing classifications regarding relationship to study treatment will be considered as related to study drug.

TEAEs will be sorted in descending order of frequency (total column) by SOC and by PT within each SOC, unless otherwise stated, alphabetical order will be used for same frequency of SOC or PT.

An overall summary of TEAEs for both the subjects with at least one AE and the total number of AEs in following categories are planned to be displayed: any TEAE, any serious TEAE, TEAE by relationship (related, not related), TEAE by severity, TEAE leading to treatment discontinuation, and any TEAEs related to a public health emergency.

The following summaries of the total number of TEAEs and the number and percentage of subjects with at least one TEAE will be provided. These will also be presented by primary SOC and PT for each treatment group:

- All TEAEs by SOC and PT
- All TEAEs by PT
- TEAEs occurring in more than 1% of study population by SOC and PT
- All TEAEs by SOC, PT and Severity
- Moderate to severe TEAEs by SOC and PT

- Drug-related TEAEs by SOC and PT
- Drug-related TEAEs by SOC, PT and Severity
- Drug-related moderate to severe TEAEs by SOC and PT
- Public health emergency related TEAEs by SOC and PT

AECOPD events as defined in the protocol will be captured as AEs with a Preferred Term of "COPD". AECOPD data will be summarized for All Exacerbations; Moderate to Severe Exacerbations; and Severe Exacerbations.

Antimuscarinic TEAEs are of interest as a common event in the LAMA drug class. The preferred terms for TEAEs considered antimuscarinic are constipation, dry mouth, dysuria, worsening of urinary retention, and worsening of narrow-angle glaucoma. The incidence of these events in this study will be presented.

4.11.3 Deaths, Serious Adverse Events, and Other Significant Adverse Events

The following summaries of the total number of TEAEs and the number and percentage of subjects with at least one TEAE will be provided. These will also be presented by primary SOC and PT for each treatment group:

- All serious TEAEs by SOC and PT
- Drug-related serious TEAEs by SOC and PT
- TEAEs leading to drug discontinuation
- Serious TEAEs leading to drug discontinuation
- TEAEs leading to drug interruption
- TEAEs leading to death

4.11.4 Clinical Laboratory Evaluation

Clinical laboratory tests will be performed at times defined in the study schedule. Any clinically significant changes in laboratory safety data should be recorded as an AE.

Hematology, chemistry and urinalysis will be collected. Laboratory data will be summarized in terms of observed values and changes from baseline and will be tabulated at each scheduled time point by treatment group and overall.

Hematology, chemistry and urinalysis data will be presented in tables as changes from baseline relative to normal ranges (e.g. shifts from normal to abnormal).

All hematology, serum chemistry and urinalysis data will be listed with values flagged if outside of the normal range.

4.11.5 Vital Signs, Physical Findings and Other Observations Related to Safety

<u>Vital Signs</u>

Supine or semi-recumbent systolic blood pressure (SBP), diastolic blood pressure (DBP) and pulse rate (PR) will be measured at screening and pre-dose and 1 hour post-dose on Day 1 and 85.

SBP, DBP and pulse rate will be listed and descriptively summarized by treatment group and visit. Change from baseline values will be similarly summarized.

All data related to vital signs will be listed.

Physical Findings

Physical Examination, including body system of respiratory, cardiovascular, gastrointestinal, musculoskeletal, neurological system, lymph node, skin, extremities, head, ears, eyes, nose, thyroid gland, will be measured at screening and Day 85. Clinically significant conditions at Screening will be reported in the Medical History. If conditions have worsened since randomization, changes should be captured as AE/SAEs as appropriate.

Screening measurements will be listed and summarized by treatment group and visit. Shift from baseline results will be similarly summarized.

All data related to physical examination will be listed.

Electrocardiogram

Twelve-lead ECG measurements will be measured from a single recording at screening and pre-dose and at 1 hour post-dose on Day 1 and 85.

Baseline and change from baseline for QT, QTcF, HR, QRS, RR and PR will be summarized using descriptive statistics by treatment group and visit.

For QTcF a classification of absolute values and increases from baseline will be performed. The number of subjects with maximum absolute QTcF <450 msec, 450 msec ≤QTcF <480 msec, 480 msec ≤QTcF <500 msec and QTcF values ≥500 msec will be tabulated by treatment and visit. The number of subjects with maximum increase from baseline QTcF <30 msec, 30 msec ≤QTcF <60 msec and QTcF ≥60 msec will be tabulated by treatment and visit.

when multiple assessments of ECG taken at a post-baseline timepoint, for the continues variables, the average of results at the same timepoint will be used, for the categorical variables, the worst ECG finding at that timepoint will be used.

All recorded values for the ECG parameters will be listed by subject. A separate listing of subjects with values of QTcF ≥500 msec or an increase ≥60 msec will be provided, as necessary.

4.12 Determination of Sample Size

The objective of this study is to confirm the efficacy of 12 weeks treatment with revefenacin 175 mcg compared with placebo administered once-daily in subjects with moderate to very severe COPD. The primary endpoint is change from baseline trough FEV_1 at Day 85. The comparison of interest for the primary analysis is revefenacin 175 mcg versus placebo.

For each test on each efficacy endpoint, the null hypothesis is that there is no difference between the treatment groups.

H_0: $\mu_{rev} - \mu_{pbo} = 0$

The alternative hypothesis is that there is a difference.

H_1: $\mu_{rev} - \mu_{pbo} \neq 0$

The estimate for the difference in mean change from baseline trough FEV_1 at Day 85 and the common standard deviation are based on the sample size assumptions for the revefenacin US pivotal studies (0126 and 0127) in a Western population and the GLOW7 study - a similarly designed study in a predominantly Chinese population (Wang et al., 2015).

A sample size of 256 evaluable subjects (128:128; a 1:1 ratio of revefenacin to placebo) is required to provide at least 90% power to detect a difference between revefenacin and placebo of 100 mL in the analysis of the change from baseline trough FEV_1 at Day 85 assuming a standard deviation of 245 mL and a two-sided 5% significance level.

To allow for a 20% dropout rate, approximately 320 subjects in total will be randomized. In the event of a public health emergency such as COVID-19, the Sponsor may consider increasing enrollment to overcome any additional loss of information for the primary endpoint due to the inability of subjects to attend for scheduled clinic visits. The maximum number of subjects to be randomized under such circumstances will be limited to 428 (an additional 20% increase in enrollment).

4.13 Changes in the Conduct of the Study or Planned Analysis

A final SAP (version 1.0) will be available after the clinical protocol is approved and a Case Report Form (CRF) is available, but before the first dry run or blinded data review is performed.

The finalization of the SAP (version 2.0) will be completed prior to the database lock. Any modifications to the planned analyses after that time will be documented within the Clinical Study Report (CSR).

The below clarifications to the wordings of the secondary efficacy variables used in protocol (version 2.0):

Wordings used in Protocol	Wordings used in SAP	Rational
Peak FEV_1 (0-2h) on Day 1	CBL to Peak FEV_1 (0-2h) on Day 85	to reflect the estimates for the analyses
SGRQ responders (total score ≥4 units) on Day 85	Number (%) of SGRQ responders (total score ≥4 units) on Day 85	to reflect the estimates for the analyses
BDI/TDI on Day 85	TDI on Day 85	BDI is measured only on Day 1
Trough FEV_1 on Days 29 and 57	CBL Trough FEV_1 on Days 29 and 57	to reflect the estimates for the analyses
BDI/TDI on Days 29 and 57	TDI on Days 29 and 57	BDI is measured only on Day 1

Weighted Mean FEV_1 (0-2h) on Day 1 and Day 85	CBL in Weighted Mean FEV_1 (0-2h) on Day 1 and Day 85	to reflect the estimates for the analyses
Trough FVC on Days 29, 57 and 85	CBL Trough FVC on Days 29, 57, and 85	to reflect the estimates for the analyses
TDI responders (focal score ≥1) on Days 29, 57 and 85	Number (%) of TDI responders (focal score ≥1) on Days 29, 57 and 85	to reflect the estimates for the analyses

5 REFERENCES

[1] Global initiative for chronic obstructive lung disease (GOLD). Global strategy for the diagnosis, management and prevention of chronic obstructive pulmonary disease. 2019. Available at: http://goldcopd.org

[2] Jones Paul W, Forde Y. St George's Repiratory Questionnaire Manual. 2009 June, Version 2.3.

6 Appendix

6.1 Sample SAS code for MMRM

[redacted]

6.2 Sample SAS code for ANCOVA

[redacted]

6.3 Tipping-Point Multiple Imputation Method

Multiple imputation can be performed in SAS using a general three-step approach.

1) Imputation Step

The efficacy data will need to be transformed so that for each subject there is a separate column variable for the efficacy value (change from baseline FEV_1) at each visit (e.g. SUBJID BASE CHG4 CHG8 CHG12).

PROC MI is applied to an input dataset containing some missing values, which results in creation of multiple copies of this dataset. All copies contain identical values of the non-missing data items, but different values imputed for missing items.

a. PROC MI distinguishes two types of missingness patterns: non-monotone and monotone.
b. In all analyses non-monotone missing data will be assumed to be missing at random

Non-monotone missing data will first be imputed using PROC MI code similar to that shown below:

[redacted]

Monotone missing data will then be imputed for a Tipping Point Analysis with incremental shifts (k) of +50 mL for the revefenacin arm and -50 mL for the placebo arm in 50 mL increments. This is illustrated by the example PROC MI code shown below:

2) Analysis Step

The data from the imputed datasets will then need to be transformed back prior to analysis so that for each subject at each visit a separate record is available (eg SUBJID VISIT CHG).

The following SAS code is an example of code used to summarize the primary analysis model on each of the imputed datasets:

3) Pooling Step

Results from all imputed datasets are then combined together for overall inference using PROC MIANALYZE.

The following SAS code is an example of code used to combine the individual model estimates for each imputed dataset by shift parameter:

[redacted]

P-values from the analyses will be presented in a two-dimensional table as illustrated in Table 7:

Table 7: Two-Dimensional Tipping Point Analysis Results (p-values)

		Shift Parameters for Missing Data on Revefenacin (mL)										
		0	-50	-100	-150	-200	-250	-300	-350	-400	-450	-500
Shift Parameters for Missing Data on Placebo (mL)	0											
	50											
	100											
	150											
	200											
	250											
	300											
	350											
	400											
	450											
	500											

6.4 Sample SAS code for Logistic Regression Model

[redacted]

6.5 Sample SAS code for MMRM logistic model:

[redacted]

6.6 BDI/TDI Component Score and Focal Score

Dyspnea Index (BDI) measures the severity of dyspnea at the beginning of the trial and the Transition Dyspnea Index (TDI) evaluates changes from this baseline (transition period).
The BDI as well as the TDI are composed of the three categories: Functional Impairment, Magnitude of Task, Magnitude of Effort.
The BDI includes five grades of severity from zero (very severe impairment) to four (no impairment) and the categories are summed to create the focal score (zero to twelve).

The TDI ranges from minus three (major deterioration) to plus three (major improvement) including a zero score to indicate "no change". Also for the TDI the three categories are added to obtain a focal score ranging from minus nine, including zero, to plus nine. Provision is made for circumstances when dyspnea cannot be rated: in the BDI, score "X" if no information on the severity can be obtained, "W" if there is generally insufficient information, or "Y" if the patient's capacity is compromised by factors other than respiratory. In the TDI, score "Z" if reduction of activities, effort or functional impairment is caused by reasons other than respiratory.
If a component score is missing or scored W, X , Y, Z then the focal score will be set to missing.

6.7 SGRQ Component Score and Total Score

The SGRQ is designed to measure health impairment in patients with asthma and COPD. It is in two parts. Part 1 (Question 1 to 8) produces the Symptoms score, and Part 2 (Question 9 to 16) the Activity and Impacts scores. A Total score is also calculated which summarises the impact of the disease on overall health status. Scores are expressed as a percentage of overall impairment where 100 represents worst possible health status and 0 indicates best possible health status.

6.7.1 Item Weights

Questions	Answers	Weights
Part 1		
Question 1	Most	80.6
	Several	63.2
	A few	29.3
	Only	28.1
	Not	0.0
Question 2	Most	76.8
	Several	60.0
	A few	34.0
	Only	30.2
	Not	0.0
Question 3	Most	87.2
	Several	71.4
	A few	43.7
	Only	35.7
	Not	0.0
Question 4	Most	86.2
	Several	71.0
	A few	45.6
	Only	36.4
	Not	0.0
Question 5	More than three	86.7
	3 attacks	60.3
	2 attacks	29.3
	1 attack	44.2

	None	0.0
Question 6	a week or more	89.7
	3 or more days	73.5
	1 or 2 days	58.8
	less than a day	41.9
Question 7	None	93.3
	1 or 2	76.6
	3 or 4	61.5
	nearly every day	15.4
	every day	0.0
Question 8	No	0.0
	Yes	62.0
Part 2		
Question 9	The most important problem I have	83.2
	Causes me quite a lot of problems	82.5
	Causes me a few problems	34.6
	Causes no problem	0.0
Question 10	My chest trouble made me stop work	88.9
	My chest trouble interferes with my work or made me change my work	77.6
	My chest trouble does not affect my work	0.0
Question 11	Sitting or lying still	90.6
	Getting washed or dressed	82.8
	Walking around the home	80.2
	Walking outside on the level	81.4
	Walking up a flight of stairs	76.1
	Walking up hills	75.1
	Playing sports or games	72.1
Question 12	My cough hurts	81.1
	My cough makes me tired	79.1
	I get breathless when I talk	84.5
	I get breathless when I bend over	76.8
	My cough or breathing disturbs my sleep	87.9
	I get exhausted easily	84.0
Question 13	My cough or breathing is embarrassing in public	74.1
	My chest trouble is a nuisance to my family, friends or neighbours	79.1
	I get afraid or panic when I cannot get my breath	87.7
	I feel that I am not in control of my chest problem	90.1
	I do not expect my chest to get any better	82.3
	I have become frail or an invalid because of my chest	89.9

	Exercise is not safe for me	75.7
	Everything seems too much of an effort	84.5
Question 14	My medication does not help me very much	88.2
	I get embarrassed using my medication in public	53.9
	I have unpleasant side effects from my medication	81.1
	My medication interferes with my life a lot	70.3
Question 15	I take a long time to get washed or dressed	74.2
	I cannot take a bath or shower, or I take a long time	81.0
	I walk more slowly than other people, or I stop for rests	71.7
	Jobs such as housework take a long time, or I have to stop for rests	70.6
	If I walk up one flight of stairs, I have to go slowly or stop	71.6
	If I hurry or walk fast, I have to stop or slow down	72.3
	My breathing makes it difficult to do things such as walk up hills, carry things up stairs, light gardening such as weeding, dance, play bowls or play golf	74.5
	My breathing makes it difficult to do things such as carry heavy loads, dig the garden or shovel snow, jog or walk at 5 miles per hour, play tennis or swim	71.4
	My breathing makes it difficult to do things such as very heavy manual work, run, cycle, swim fast or play competitive sports	63.5
Question 16	I cannot play sports or games	64.8
	I cannot go out for entertainment or recreation	79.8
	I cannot go out of the house to do the shopping	81.0
	I cannot do housework	79.1
	I cannot move far from my bed or chair	94.0
Question 17	It does not stop me doing anything I would like to do	0.0
	It stops me doing one or two things I would like to do	42.0
	It stops me doing most of the things I would like to do	84.2
	It stops me doing everything I would like to	96.7

6.7.2 Scoring Algorithm

Each component of the questionnaire is scored separately in three steps:

i. The weights for all items with a positive responses are summed.

ii The weights for missed items are deducted from the maximum possible weight for each component. The weights for all missed items are deducted from the maximum possible weight for the Total score.

iii. The score is calculated by dividing the summed weights by the adjusted maximum possible weight for that component and expressing the result as a percentage:

$$Score = 100 \times \frac{\text{Summed weights from positive items in that component}}{\text{Sum of weights for all items in that component}}$$

The Total score is calculated in similar way:

$$Score = 100 \times \frac{\text{Summed weights from positive items in the questionnaire}}{\text{Sum of weights for all items in the questionnaire}}$$

Sum of maximum possible weights for each component and Total: Symptoms 662.5, Activity 1209.1, Impacts 2117.8, Total 3989.4

It will be noted that the questionnaire requests a single response to questions 1-7, 9-10 and 17. If multiple responses are given to one of these questions then averaging the weights for the positive responses for that question will be applied.

Symptoms component is calculated from the summed weights for the positive responses to questions 1-8. Activity component is calculated from the summed weights for the positive responses to questions 11 and 15. Impacts component is calculated from the summed weights for the positive responses to questions 9-10, 12-14 and 16-17.

The Total score is calculated by summing all positive responses in the questionnaire and expressing the result as a percentage of the total weight for the questionnaire.

6.7.3 Handling Missed Items

The Symptoms component will tolerate a maximum of 2 missed items. The weight for the missed item is subtracted from the total possible weight for the Symptoms component (662.5) and from the Total weight (3989.4).

The Activity component will tolerate a maximum of 4 missed items. The weight for the missed item is subtracted from the total possible weight for the Activity component (1209.1) and from the Total weight (3989.4).

The Impacts component will tolerate a maximum of 6 missed items. The weight for the missed item is subtracted from the total possible weight for the Impacts component (2117.8) and from the Total weight (3989.4).

6.8 Best Test Result (BTR)

The BTR process is designed to deselect any unacceptable efforts that are being used to represent the Best value for a particular measurement or reselect good quality efforts that were deselected by the Principal Investigator. If the Principal Investigator did not disregard the unacceptable efforts, the Best value could be influenced due to the fact that the MasterScope will report the highest value for FEV_1 and FVC per measurement as the Best value regardless of the quality efforts.

The Best value definition in Forced Spirometry is the highest values of FEV_1 and FVC of the efforts submitted by the investigative site. The goal for BTR is to ensure that only acceptable efforts are used to determine these Best values. The data submitted to the sponsor after BTR will include the highest value (FEV_1 or FVC) reported from acceptable efforts. If there are no acceptable efforts available for a measurement then the data submitted will be unacceptable quality for that measurement.

[redacted]

[redacted]

[redacted]

Approval Signatures

Document Name: Statistical Analysis Plan 16 Oct 2023 REV 3001

Document Number: [redacted]

[redacted] **Version Number:**

System Version Number: 1 .0

Document Approvals	
Reason for signing: Approved	Name: [redacted] Role: Biostatistics Date of signature: [redacted]
Reason for signing: Approved	Name: [redacted] Role: Biostatistics Date of signature: [redacted]