# CHIESI

# GLIMMER

A 6-WEEK, RANDOMIZED, DOUBLE-BLIND, PLACEBO AND ACTIVE-CONTROLLED, PARALLEL GROUP, DOSE RANGING STUDY TO EVALUATE THE EFFICACY AND SAFETY OF 4 DOSES OF CHF 5259 PMDI (GLYCOPYRRONIUM BROMIDE) IN SUBJECTS WITH CHRONIC OBSTRUCTIVE PULMONARY DISEASE (COPD)

Protocol: CCD-05993AA3-02

IND No: 133681

ClinicalTrials.gov ID: NCT03084796

Statistical Analysis Plan

Version 1.0

Final

20 SEP 2018

Author:

[redacted]

Table of Contents

# 1 SOPs to be followed

The statistical analysis will be carried out according to the following [redacted] SOPs:

| SOP Number | SOP Title | Effective Date & Version Number | [redacted] or Sponsor |
|---|---|---|---|
| SOP_ST_03 | Statistical Analysis Plan | 3.0 01AUG2015 | [redacted] |
| SOP_ST_04 | SAS Programming, QC and Validation | 3.0 21DEC2015 | [redacted] |
| SOP_ST_06 | Study Unblinding for Analysis | 3.0 12JUN017 | [redacted] |
| SOP_ST_07 | Statistical Report | 3.0 26JUN2017 | [redacted] |
| SOP_ST_08 | Trial Statistics File* | 2.0 15APR2014 | [redacted] |

* TSF SOP in combination with the ToC of Chiesi.

# 2 Abbreviations

| Abbreviation | Description |
|---|---|
| ADAM | Analysis Data Model |
| ADR | Adverse Drug Reaction |
| AE | Adverse Event |
| AUC | Area Under the Curve |
| BDI | Baseline Dyspnea Index |
| BID | Bis in die (twice a day) |
| BMI | Body Mass Index |
| CAT | COPD Assessment Test |
| CDISC | Clinical Data Interchange Standards Consortium |
| Cm | Centimeters |
| COPD | Chronic Obstructive Pulmonary Disease |
| CSR | Clinical Study Report |
| DBP | Diastolic Blood Pressure |
| ECG | ElectroCardioGram |
| eCRF | Electronic Case Report Form |
| E-RS | EXACT-Respiratory Symptom |
| $FEV_1$ | Forced Expiratory Volume in the 1$^{st}$ second |
| FVC | Forced Vital Capacity |
| GB | Glycopyrronium Bromide |
| H | Hours |
| HR | Heart Rate |
| IC | Inspiratory Capacity |
| ICS | Inhaled Corticosteroid |
| IRT | Interactive Response Technology |
| ITT | Intention-to-Treat (analysis population) |
| Kg | Kilograms |
| L | Liters |
| LABA | Long-Acting $\beta_2$-adrenergic receptor Agonist |
| LAMA | Long Acting Muscarinic Antagonist |
| µg | Micrograms |
| M | Meters |
| ms | Milliseconds |
| MedDRA | Medical Dictionary for Regulatory Activities |
| min | Minute |
| mL | Milliliter |
| PEF | Peak Expiratory Flow |
| pMDI | Pressurized Metered Dose Inhaler |
| PP | Per-Protocol (analysis population) |
| PR | PR Interval width (ECG) |
| PT | Preferred Term (MedDRA) |
| QC | Quality Control |
| Qd | Once a day |
| **Abbreviation** | **Description** |

| | |
|---|---|
| QRS | QRS interval width (ECG) |
| QTcF | QT interval width corrected according to Fridericia (ECG) |
| SAE | Serious Adverse Event |
| SAP | Statistical Analysis Plan |
| SBP | Systolic Blood Pressure |
| SDTM | Study Data Tabulation Model |
| SOC | System Organ Class (MedDRA) |
| SOP | Standard Operating Procedure |
| TDI | Transitional Dyspnea Index |
| TEAE | Treatment Emergent Adverse Event |
| WHO | World Health Organization |
| WHO-DD | World Health Organization – Drug Dictionary |

# 3 Protocol / Clinical Investigation Plan

This document presents the statistical analysis plan (SAP) for Chiesi Farmaceutici S.p.A., Protocol No. CCD-05993AA3-02: A 6-week, randomized, double-blind, placebo and active-controlled, parallel group, dose ranging study to evaluate the efficacy and safety of 4 doses of CHF 5259 pMDI (glycopyrronium bromide) in subjects with Chronic Obstructive Pulmonary Disease (COPD). This analysis plan is based on the final protocol (version 2.0) dated 19 MAY 2017 and the final electronic case report form (eCRF) (version 3.0) dated 28 JUL 2017. Text copied from Protocol is reported in this document in Italics to avoid unnecessary alterations to text approved in the Protocol.

The SAP provides the description of the final analyses. In case of deviations from the SAP, explanations will be provided in the Clinical Study Report (CSR).

## 3.1 Study Objectives

The primary objective of this clinical trial is

- *To evaluate the efficacy of CHF 5259 pMDI by comparison with placebo in terms of change from baseline in $FEV_1$ $AUC_{0-12h}$ normalized by time at Week 6.*

There are 2 secondary objectives:

- *To evaluate the effect of CHF 5259 pMDI on other lung function parameters and clinical outcome measures.*
- *To assess the safety and the tolerability of the study treatments.*

## 3.2 Study Design

The study is a Phase II multi-center, randomized, double-blind, placebo and active-controlled, parallel group, dose ranging study. Subjects will be randomized into 1 of 6 treatment groups presented below.

- Treatment A*: GB 12.5µg Total Daily Dose (GB 6.25µg per inhalation, 1 inhalation bid)
- Treatment B*: GB 25µg Total Daily Dose (GB 12.5µg per inhalation, 1 inhalation bid)
- Treatment C: GB 50µg Total Daily Dose (GB 12.5µg per inhalation, 2 inhalations bid)
- Treatment D: GB 100µg Total Daily Dose (GB 25µg per inhalation, 2 inhalations bid)
- Treatment E: Matched placebo (2 inhalations bid)
- Treatment F (open-label): SPIRIVA® HANDIHALER® (Tiotropium bromide inhalation powder) 18µg capsule (2 inhalations qd)

*An adequate number of inhalations from Placebo inhalers will be performed to maintain a double blind design.

**Background Medication for the Run-in and Treatment Periods**

*At the screening visit (Visit 1), all eligible subjects who have been receiving an ICS in combination with a LABA or LAMA will be switched to an equipotent dose of QVAR® as run-in / background therapy: QVAR® 80µg (HFA beclomethasone dipropionate - Teva Respiratory, LLC). The dose will be 1-2 inhalations twice per day (Total daily dose: beclomethasone dipropionate 160-320µg).*

*This treatment will be maintained at a stable dose and regimen (160-320µg daily) throughout the study from screening until the end of the study.*

## 3.3 Study Schedule

| | Pre-Screening | Screening | Treatment period | | | Follow-Up | ET |
|---|---|---|---|---|---|---|---|
| **Visits** | **V0** | **V1** | **V2** | **V3** | **V4** | **V5** | |
| **Time (Weeks)** | | **-2** | **0** | **3** | **6** | **7** | **Early Termination** |
| **Window (days)** | | | ±2 | ±2 | ±2 | ±2 | |
| | | | | | | | |
| Informed consent form | ✓ | | | | | | |
| IRT – visit confirmation call | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Demographic data | ✓ | | | | | | |
| BDI questionnaire | | ✓ | ✓ | | | | |
| TDI questionnaire | | | | ✓ | ✓ | | ✓ |
| COPD Assessment Test (CAT) | | ✓ | ✓ | | | | |
| Vital signs (SBP/DBP) [a] | | ✓ | ✓ | ✓ | ✓ | | ✓ |
| Weight and Height | | ✓ | | | | | |
| Physical examination | | ✓ | | | ✓ | | ✓ |
| Medical history/Previous medication | | ✓ | | | | | |
| Concomitant medications | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Smoking Status | | ✓ | ✓ | ✓ | ✓ | | ✓ |
| 12-lead ECG [b] | | ✓ | ✓ | ✓ | ✓ | | ✓ |
| Spirometry (pre & post BD) [c] | | ✓ | | | | | |
| Pre-dose spirometry [d] | | | ✓ | ✓ | ✓ | | ✓ |
| Post-dose serial spirometry 12h [e] | | | ✓ | | ✓ | | |
| Hematology and Blood Chemistry | | ✓ | | | ✓ | | ✓ |
| Serum pregnancy test [f] | | ✓ | | | ✓ | | ✓ |
| Urinary pregnancy test [f] | | ✓ | ✓ | ✓ | | | |
| Inclusion / exclusion criteria | | ✓ | ✓ | | | | |
| Eligibility recheck [g] | | | ✓ | | | | |
| Training for use of pMDI Inhalers and e-diary | | ✓ | ✓ | | | | |
| Dispensation of rescue albuterol [h] | | ✓ | | | | | |
| 24-hour holter recording [i] | | | ✓ | | ✓ | | |
| Schedule next visit | ✓ | ✓ | ✓ | ✓ | ✓ | | ✓ |
| Randomization | | | ✓ | | | | |
| e-diary completion | | ✓ **(daily)** | | | | | |
| Study drug dispensation (D) / Return (R) and Accountability | | | D | D/R | R | | R |
| Subject diary dispensation (D) /return (R) | | D | D/R | D/R | R | | R |
| Dispensation (D) of background ICS (QVAR®) [j] / return (R) | | D | D/R | D/R | R | | R |
| Adverse Events assessment | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Check for COPD Exacerbations | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

[a] At V2 and V4: SBP/DBP will be measured within 1h pre-dose and at 30 mins, 1.5h, 3.5h, 7h and 11h post-dose. At V1 and Early Termination: SBP/DBP will be measured within 1h before the expected time of study drug administration. At V3: SBP/DBP will be measured within 1h pre-dose.
[b] At V2 and V4: Local 12-lead ECG will be done within 1h prior to study drug intake and at 1.5h post-dose. At V1 and Early Termination: Local 12-lead ECG will be done within 1h before the expected time of study drug administration. At V3: Local 12-lead ECG will be done within 1h prior to study drug intake.
[c] Spirometry (FVC maneuver) will be carried out 15 min before bronchodilator and 30 to 45 minutes after the inhalation of 84µg of ipratropium pMDI at V1.
[d] Pre-dose $FEV_1$, FVC: T -45'and T -15' before the expected time of study drug administration. IC will be measured only once at T-45' using SVC maneuver, before the FVC maneuver.
[e] Post-dose serial spirometry ($FEV_1$, FVC): T15', T30', T45', T1h, T2h, T3h, T4h, T6h, T8h, T10h, T11.5h, T12h
[f] In women of childbearing potential only.
[g] Eligibility recheck only for exclusion criteria #1, 3, 4, 5, 10, 11 and inclusion criteria # 6-8.
[h] One commercial albuterol HFA MDI (200 actuations) will be prescribed and supplied by the investigator to each subject at V1, and resupplied as needed at V2-V3 based on assessment of doses used between visits.
[i] At V2: the 24h digital device will be placed the day before the visit and will be removed the day after the visit. At V4: the 24-hour digital device will be placed the day before the visit and will be removed upon arrival to the site (before any assessment).
[j] all subjects who have been receiving an ICS in combination with a LABA or LAMA at V1 will be switched to an equipotent dose of QVAR® and continued on same dose until V4 or Early Termination.

## 3.4 Randomization

*A balanced block randomization scheme stratified by US Region (based on US Census Bureau Regions: West, Midwest, South, Northeast) will be prepared via a computerized system. Subjects will be centrally assigned to one of the six treatment arms at the end of the run-in period through an IRT system (Interactive Response Technology) with a 1:1:1:1:1:1 ratio.*

*The IRT will allocate the subject to a certain treatment group using a list-based randomization algorithm and will assign the study drug kit number corresponding to the treatment group assigned to the subject. The IRT will also generate a confirmation after every IRT transaction is performed.*

*The Investigator will call the IRT at each visit (from pre-screening to follow-up call) to record the subject number at pre-screening, to enroll and randomize the subject, to obtain the medication kit numbers and to register the subject status in the system. Detailed instructions for use of IRT will be provided to the site.*

## 3.5 Sample Size Calculation

*The sample size has been calculated to evaluate the superiority of CHF 5259 pMDI at different doses over placebo in terms of change from baseline in $FEV_1$ $AUC_{0\text{-}12h}$ normalized by time at Week 6.*

*A total of 594 evaluable subjects (99 per group) will provide 80% power to detect a mean difference of 120 mL between each dose of CHF 5259 pMDI and placebo at a two-sided significance level of 0.0125 (since 4 dose levels will be tested, the Bonferroni adjustment has been taken into account: 0.0125 = 0.05/4), assuming a standard deviation of 250 mL.*

*Since four dose levels will be tested, the Edwards and Berry method will be used to control the family-wise Type I error rate at the 0.05 (two-sided) level. In the sample size calculation, the Bonferroni adjustment of the significance level has been taken into account (0.0125 = 0.05/4). This will ensure the required power for each test, since the Edwards and Berry method is uniformly more powerful than the Bonferroni procedure.*

*Considering a non-evaluable rate of 15%, a total of approximately 702 subjects (117 per group) will be randomized.*

## 3.6 Efficacy and Safety Variables

### *Primary Efficacy Variable*

*Change from baseline in $FEV_1$ $AUC_{0\text{-}12h}$ normalized by time at Week 6.*

### *Secondary Efficacy Variables*

The protocol defines 14 unique secondary efficacy variables but groups them into 8 bullet points in Section 12.3.5 of the protocol and 11 bullet points in Section 8. The variable numbering below is not meant to reflect a hierarchical ordering. Rather, the numbering is meant to allow the listing of all 14 unique secondary efficacy variables but provide the traceability back to the 8 bullet points as outlined in the protocol:

- 1: *Change from baseline in $FEV_1$ $AUC_{0\text{-}12h}$ normalized by time at Day 1*
- 2a: *Change from baseline in $FEV_1$ $AUC_{0\text{-}4h}$ normalized by time at Day 1 and Week 6*

- 2b: *Change from baseline in $FEV_1$ $peak_{0-4h}$ at Day 1 and Week 6*
- 2c: *Change from baseline in pre-dose morning $FEV_1$ at Week 3 and Week 6*
- 3a: *Change from baseline in FVC $AUC_{0-12h}$ normalized by time at Day 1 and Week 6*
- 3b: *Change from baseline in FVC $AUC_{0-4h}$ normalized by time at Day 1 and Week 6*
- 3c: *Change from baseline in FVC $peak_{0-4h}$ at Day 1 and Week 6*
- 4: *Time to onset of action (change from baseline in post-dose $FEV_1 \geq 100$ mL) at Day 1*
- 5: *Change from baseline in pre-dose morning IC at Week 3 and Week 6*
- 6: *TDI focal score at Week 3 and Week 6*
- 7: *TDI response (TDI focal score ≥ 1) at Week 3 and Week 6*
- 8a: *Change from baseline in percentage of rescue medication-free days during inter-visit periods and entire treatment period*
- 8b: *Change from baseline in average use of rescue medication (number of puffs/day) during inter-visit periods and entire treatment period*
- 8c: *Change from baseline in average E-RS total score and domain scores during inter-visit periods and entire treatment period*

**Safety Assessments**

The list of safety assessments presented below is from Section 8 of the protocol. Additional details regarding the safety variables collected during this clinical trial are presented in Section 12.3.6 of the protocol.

- *Adverse Events (AEs) and Adverse Drug Reactions (ADRs)*
- *Vital signs (systolic and diastolic blood pressure)*
- *24-hour digital Holter ECG parameters (HR, QTcF, QRS, PR)*
- *24-hour HR average, minimum and maximum and hourly average HR*
- *24-hour digital Holter ECG abnormal findings*
- *Standard blood chemistry and hematology*

## 3.7 Interim Analyses

No interim analyses are planned for this clinical investigation.

## 3.8 Changes in the Conduct of Study or Planned Analysis compared to the Study Protocol

No changes in conduct between the Statistical Analysis Plan (SAP) and the study protocol have been noted.

# 4 General Definitions

## 4.1 Report Language

The output of the analyses will be prepared in (USA) English.

## 4.2 Analysis Software

The statistical analysis will be performed using the SAS® statistical software package (Version 9.3).

# 5 Data Preparation

## 5.1 Data Handling and Medical Coding

### 5.1.1 Data Handling

For data quality control, please refer to the Data Management Plan (Version 1.0 from 22 JUN 2017) including the Data Validation Plan (Version 3.0 from 28 JUL 2017).

### 5.1.2 Coding

The following dictionaries will be used for coding in the analysis:

**Medical History, Concomitant Diseases, Surgeries and Procedures**

Medical Dictionary for Regulatory Activities (MedDRA) coding system, version 20.0.

**Prior and Concomitant Medications**

World Health Organization Drug Dictionary (WHO-DD version January 2017.

**Adverse Events**

Medical Dictionary for Regulatory Activities (MedDRA) coding system, version 20.0.

## 5.2 CDISC

All output as defined in the SAP will be generated based on CDISC ADaM datasets. Specifications for the ADaM datasets (as well as SDTM datasets) are described in a separate document.

## 5.3 SAS-Programming Quality Level

The following quality level of programming deliverables will be applied:

All statistical output will receive a tailored Quality Control (QC) approach by:

- Full independently double programmed reproduction (QC) of results of
  - CDISC ADaM datasets
  - Unique Tables and Graphs
  - All tables and figures reporting inferential analyses results
- Listings will not be double programmed. The programming to generate the listings will be reviewed in accordance with the [redacted] procedure for all SAS programs.
- All tables, listings, and graphs will undergo comparison with specifications (i.e. SAP and templates), cross checking with other tables, listings and graphs, the individual logs from the SAS programs will be reviewed to ensure all errors, warnings, and uninitialized variable messages have been rectified.
- A Senior Review will also be performed by a reviewer independent of the study team. The reviewer studies all tables, listings and graphs for consistency and correctness, and pre-empts customer comments. This allows points of interest to be highlighted and discussed at customer hand-over.

All SDTM datasets, ADaM datasets, and tables, listings, and figures will be QC'ed by independent programmers, the study biostatistician, and a senior review as per the [redacted] SOP ([redacted]) SAS Programming, Validation, and QC version 3.0.

## 5.4 Data from third parties

Data provided by third parties, not contained in the clinical database, will be included in the SAS data repository. These third parties are:

- [redacted]: central laboratory data.
- [redacted]: Holter/ECG data, drug intake details, PEF values, compliance values, spirometry data, rescue medication and EXACT-Respiratory Symptom (RS) scores.
- [redacted]: supplies the randomization numbers and treatment codes to [redacted]
- [redacted]: randomization data from the IRT system.

# 6 Analysis Populations and Subgroups

## 6.1 Analysis Populations

- ***All Enrolled Subjects***
- ***All Randomized subjects***
- ***Safety population:*** *all randomized subjects who receive at least one dose of study drug.*
- ***Intention-to-Treat population (ITT):*** *all randomized subjects who receive at least one dose of the study drug and with at least one available evaluation of efficacy (primary or secondary efficacy variables) after the baseline.*
- ***Per-protocol population (PP):*** *all subjects from the ITT population without any major protocol deviation (i.e., wrong inclusions, poor compliance and non-permitted medications). Exact definition of major protocol deviations will be discussed by the study team during the blind review of the data and described in the Data Review Report.*

*Since the superiority of CHF 5259 pMDI at different doses over placebo will be tested, the primary efficacy analysis will be based on the ITT population. The primary efficacy analysis will be also performed on the PP population for sensitivity purposes.*

All 3 populations (Safety, ITT, and PP) will be used for the presentation of demographics, COPD history and smoking, and spirometry and reversibility.

The medical history/concomitant diseases and exposure/compliance to treatment results will be presented using the ITT and Safety populations. Exposure/compliance to background medication will be presented using the ITT population.

Medications, CAT score, BDI, compliance with the use of e-diary, and protocol deviations will be reported using the ITT population.

The safety data will be summarized and reported using the Safety population.

## 6.2 Treatment Misallocation

In case of deviation between the as-randomized treatment and the treatment actually received, the treatment actually received will be used in the safety analyses (i.e. an as-treated analysis will be performed). The following rules will be applied in the construct of the populations:

- Subjects randomized but not treated will be excluded from the Safety, ITT and PP populations.
- Subjects treated but not randomized will be excluded from the ITT and PP population, given there will be no randomized treatment. These subjects will be summarized based on the treatment they received and will be counted in the Safety population.
- Subjects who were randomized, but took the incorrect study drug for the duration of the study, will be summarized based on their randomized treatment for the ITT population

and excluded from the PP population. These subjects will be summarized based on the treatment that they received for the safety analyses.

- Subjects who were randomized, but took the incorrect study drug for part of the study (e.g. the subject began the study by taking the incorrect study drug at Visit 2 but switched to randomized study drug at Visit 3) will be discussed during the Blind Data Review Meeting, and the decisions will be agreed prior to unblinding, and documented in the Data Review Report prior to unblinding.

### 6.3 Subgroup Definitions

There are no specific subgroups that are planned for analysis.

# 7 Definition of Time Points and Analysis Variables

## 7.1 Definition of Time Points

For each visit, the date recorded by the Investigator in the eCRF (variable SVSTDTC in the SDTM SV dataset) will be considered as the visit date in all the algorithms and the listings.

**Start/End of the Randomized Treatment Period**

Since many algorithms used in the statistical analyses require the start and/or the end of the randomized treatment period to be identified, ad-hoc variables specifying these dates will be defined. The date of start/end of randomized treatment period will be set according to the following rule:

- The date of **Start of the Randomized Treatment Period** should coincide with the date of Visit 2, the randomization date and the date of first randomized study drug intake. However, discrepancies between these dates may arise and the most appropriate date to be used in such situations requires a case-by-case evaluation. The date of the start of the randomized treatment period will be initially set equal to the date of first randomized study drug intake for all subjects. The need for deviations from this rule in single cases will be evaluated during the data review and documented in the Data Review Report. As a consequence, the distinction between diary data from the run-in and the treatment period will not be based on the EPOCH variable included in the SDTM datasets, but on the algorithms defined in the SAP.
- If Visit 4 or Early Termination visit was completed, then he date of **End of the Randomized Treatment Period** will be defined as Date of Visit 4/ Early Termination Visit. Otherwise, the date of End of Randomized Treatment Period will be defined as the maximum of [Date of Last Study Drug Intake, date of last clinic visit during treatment period (excluding Visit 5 and any unscheduled visits after last dose)].

**Period durations (days)**

Day 1 is defined as the date of first dose of randomized study drug. The day prior to Day 1 is Day -1. There is no Day 0.

Periods used in the analysis of this study are:

- Run-in Period: Morning of Visit 1 through the evening of the day before the date of Start of the Randomized Treatment Period.
  Run-in duration = date of the day before the date of the Start of the Randomized Treatment Period – date of Visit 1 + 1 day

- Treatment Period: (End of the Randomized Treatment Period minus Start of the Randomized Treatment Period) + 1 day

Inter-visit periods will be used in the analysis of the percentage of rescue medication-free days, the average number of rescue medication-puffs per day, and the average E-RS total score and domain scores. The periods used in these analyses are:

- Run-in Period: The Run-In Period starts at Visit 1 and runs through the day before the date of Start of the Randomized Treatment Period

- Inter-visit Period 1: Inter-visit Period 1 starts on the date of the Start of the Randomized Treatment Period and runs through the day before the subject returns to the clinic at Visit 3.

  If the subject withdraws prematurely after Visit 2 and before Visit 3 any value collected until the End of Randomized Treatment Period will be considered for the inter-visit period 1. By default inter-visit period 2 is missing.

- Inter-visit Period 2: Inter-visit Period 2 starts on the day the subject returns to the clinic at Visit 3 and runs through End of the Randomized Treatment Period.

  If the subject withdraws prematurely after Visit 3,any value collected until the End of the Randomized Treatment Period will be considered for the inter-visit period 2.

**Date of first/last study drug intake**

Date of first randomized study drug intake is derived as the minimum of 'Date/Time of administration' between the 2 inhalations (puffs) of the Morning Dose collected on the 'Study drug administration at the clinic' eCRF page. While subjects randomized to open-label label SPIRIVA® will only receive 1 inhaler (and only 1 date/time of inhalation is recorded), patients are required to take 2 inhalations of the content of 1 inhaler. For these patients, no minimum is required, since only 1 time is recorded.

Note: The date and time of drug administration at each visit are mandatory variables to be reported in the eCRF; no missing or partial data can be accepted. Only eCRF data will be considered for the date of first randomized study drug intake.

Date of last randomized study drug intake is derived as 'Date/Time of last intake of study drug' variable collected on the 'Study Termination form' eCRF page. If the date of the last intake of study drug is missing or partially missing, but the date of at least one dose of the study drug is recorded in the diaries or at the clinic, the date of the last randomized study drug intake will be imputed using the following rule:

max [date of last study drug intake in the diaries, date of last study drug intake at the clinic visits]

The need for deviations from these rules in single cases will be evaluated during the data review and documented in the Data Review Report.

**Time to discontinuation from the study**

Time to discontinuation will be calculated for all randomized subjects, including subjects who complete the study. For subjects randomized but not treated, the time to discontinuation from the study will be imputed as 0. Subject disposition data will be collected on the 'Study Termination form' eCRF page, which provides information about study completion status of a subject. Study discontinuation is recorded along with the main reason for withdrawal. The primary reason for discontinuation could be early withdrawal or lost to follow up. Patients completing the study will be censored at the Date of Completion. Patients lost to follow up will be considered as having an event at the date of Discontinuation recorded on the 'Study Termination" eCRF page.

The time to completion/discontinuation from the study (weeks) is defined as:

(date of completion/discontinuation - date of start of randomized treatment period ) / 7

**Baseline definitions**

For $FEV_1$ and FVC, the baseline value is defined as the average of the pre-dose measurements on Visit 2 (45 min and 15 min pre-dose). For IC, the measurement recorded at 45 minutes pre-dose at Visit 2 will be used as the baseline.

From Section 12.3.2 of the protocol*: If one of the spirometry measurements at 45 min and 15 min pre-dose is not available, then the non-missing measurement will be taken as the pre-dose value. If no measurement is* available, then the pre-dose value will be considered as missing. This rule applies to $FEV_1$ and FVC.

The analysis models of TDI will use the BDI focal score at Visit 2 as the baseline value.

The baseline used in the analysis of rescue medication use and E-RS score is based on the data recorded during the Run-in Period. Depending on the structure of the specific variable, this may be an average value or a percentage.

For vital signs, baseline values are those recorded pre-dose at Visit 2.

For laboratory the baseline values are the ones collected at Visit 1.

Subjects will have ambulatory 24-hour digital Holter recording for 24 hours before and 24 *hours after the 1st dose of study drug (Visit 2) and for 24 hours before the last dose of study drug (Visit 4).* The time-matched values recorded during the 24 hours before the 1st dose of study drug on Visit 2 will serve as the baseline for the Holter-extracted ECG parameters (HR, QTcF, QRS and PR). The time-averaged baseline score is the average of the Day -1 scores at +5m, +55m and +2.5h.

*To facilitate data quality, the subject will be asked to assume a supine position and remain quiet, but awake, beginning at 10 minutes before the nominal extraction time and then for 5 minutes afterward. For each subject, there will be:*

- *3 nominal extraction times in the day before Visit 2 (time-matched with the post-study drug extraction times on Visit 2)*
- *3 extraction times at +5 min, +55 min, and +2.5h after 1st dose of study drug on Visit 2; and*
- *3 extraction times at +5 min, +55 min, and +2.5h after the morning dose of study drug on the day before Visit 4.*

For 24-hour average, minimum and maximum HR the baseline values are those derived from the Holter recording during the 24 hours before the 1st dose of study drug on Visit 2.

## 7.2 Baseline and Derived Analysis Variables

The purpose of this section is to describe the calculation of all derived variables. All other variables that are obtained directly from the eCRF system with no derivation are not described in this section.

### 7.2.1 Demographic Characteristics

- Age of the subject will be calculated by the IVRS system based on the date of birth and date of Pre-Screening (Visit 0) entered into the system.

### 7.2.2 COPD History and Pre-Study Smoking Habits

- The time since first COPD diagnosis (months) will be calculated as the (Visit 1date - date of first diagnosis)/30.4375.
- Age at First COPD Diagnosis (years) will be calculated in SAS using the following formula: floor(yrdif(date of birth, date of first diagnosis, 'AGE')).
  The Duration of Smoking (years) will be calculated using the start/stop date of smoking and is calculated differently for current and ex-smokers:
    - For ex-smokers: (smoking stop date - smoking start date + 1)/365.25
    - For current smokers: (Visit 1 date - smoking start date + 1)/365.25

Only the month and the year (not the day) is recorded on the eCRF for Date of first diagnosis and smoking start/stop date. The first day of the month will be assumed for these dates in order to calculate time duration variables.

In order to calculate the duration, the following rules will be applied for the partial dates: if the month is missing, January 1st will be assumed. When the start date is completely missing, the time duration variable will not be calculated.

### 7.2.3 Baseline Subject Characteristics

- The Baseline Dyspnea Index (BDI) Focal Score will be calculated within the eCRF from the three individual domain scores (Functional Impairment, Magnitude of Task, and Magnitude of Effort).
- The COPD Assessment Test (CAT) is an assessment containing 8 elements, each scored on a scale from 0 to 5. A score of 0 describes the best possible status and a score of 5 describes the worst possible status. These scores will be summed in the eCRF to calculate the CAT Total Score.
- Medical/Surgical History and Concomitant Diseases will be collected during Visit 1. All conditions that are not indicated as ongoing will be considered as medical/surgical history, while conditions indicated as ongoing will be considered as concomitant diseases.
- Body Mass Index ($kg/m^2$) is calculated as Weight (kg) / [Height (m)]²

### 7.2.4 Spirometry Performed at Screening

Spirometry (FVC maneuver) will be carried out 15 minutes before bronchodilator and 30 to 45 minutes after the inhalation of 84µg of ipratropium pMDI at Visit 1 (screening). The reversibility parameters ΔFEV1 (mL) and % ΔFEV1 will be calculated in the eCRF.

If the reversibility test of Visit 1 is repeated in a (pre-randomization) rescheduled visit (called Visit 1.1 in the eCRF), then all spirometry values of the rescheduled visit will be considered as the Visit 1 assessment in the analysis.

### 7.2.5 Medications

Medications will be split into four categories:

- **Previous medications** are those medications started and stopped prior to the initial exposure to the study drug (medication start date < date of first study drug intake and medication stop date ≤ date of first study drug intake).
- **Maintained medications** are all medications started before initial exposure to the study drug and ongoing at initial exposure to the study drug (medication start date < date of first study drug intake and medication stop date > date of date of first study drug intake).
- **Concomitant medications** are all medications started during the treatment period (date of first study drug intake <= medication start date < date of last study drug intake).
- **Post-treatment medications** are all medications started on or after the last dose of study drug (medication start date >= date of last study drug intake)

In case of missing or incomplete dates not directly allowing allocation to any of the four categories of medications, a worst-case allocation will be performed according to the available parts of the start and the stop dates. The medications will be allocated to the first category allowed by the available data, according to the following order:

1. concomitant medication;
2. maintained medication;
3. post-treatment medication;
4. previous medication.

Information on concomitant medications is retrieved from the 'Prior and Concomitant medications' eCRF page.

### 7.2.6 Procedures

Medications will be split into two categories:

- **Previous procedures** are all procedures started and stopped prior to the initial exposure to the study drug (procedures start date < date of first study drug intake and procedures stop date ≤ date of first study drug intake).
- **Maintained procedures** are all procedures started before initial exposure to the study drug and ongoing at initial exposure to the study drug (procedures start date < date of first study drug intake and procedures stop date > date of date of first study drug intake).
- **Concomitant procedures** are all procedures started during the treatment period (date of first study drug intake <= procedures start date < date of last study drug intake).
- **Post-treatment procedures** are all procedures started on or after the last dose of study drug (procedures start date >= date of last study drug intake)

In case of missing or incomplete dates not directly allowing allocation to the two categories of procedures, a worst-case allocation will be performed according to the available parts of the start and the stop dates. The procedures will be allocated to the first category allowed by the available data, according to the following order:

1. concomitant procedures;
2. maintained procedures;
3. post-treatment procedures;

4. previous procedures.

Information on concomitant procedures is retrieved from the 'Concomitant procedures' eCRF page. Procedures will be listed only.

### 7.2.7 Treatment Exposure and Compliance – Background Medication

**Run-In Period / Exposure and Compliance**

Since only subjects receiving prior ICS treatment will be provided with QVAR at V1, treatment exposure and compliance with background ICS medication (QVAR) can only be performed on subjects provided with QVAR at V1. The Run-In Period starts at Visit 1 and runs through the day before Day 1 (Day -1). Exposure during the Run-in Period is based on actual dosing date/times and not visit dates.

Exposure will be calculated as the date of last dose of the Run-in Period – date of first dose of the Run-in Period + 1 day.

Compliance to background medication during the run-in period is based on eCRF and diary data. If, on a visit day, information is available from both eCRF and diaries, data entered by the investigator in the eCRF will be taken into account.

The evaluation of compliance will be based on the number of puffs following the formula presented below:

$$\frac{\text{Total number of administered puffs (as recorded in the Subject Diary and / or Background Medication Administration eCRF page)}}{\text{Total number of scheduled puffs}} x\ 100\%$$

The total number of **scheduled puffs per day** is the sum of the 'Number of puffs the subject has been instructed to take: Morning puffs' plus 'Evening puffs' as recorded on the Screening 'COPD Background Medication Administration' eCRF page.

The total scheduled number of puffs during the Run-in Period is defined as:

[Run-in period duration (days)] x [scheduled number of puffs per day]

The approach above defined for the calculation of compliance assumes no intake of the background medication in case of missing data.

**Treatment Period / Exposure and Compliance**

Compliance to background medication during the treatment period should be calculated considering data recorded from date of start of randomized treatment period to date of end randomized treatment period.

Exposure to background medication during treatment period (days) will be calculated as date of last background medication intake - date of first background medication intake + 1 day.

Compliance to background medication during the treatment period is based on eCRF and diary data. If on a visit day information is available from both eCRF and diaries, data entered by the

investigator in the eCRF will be taken into account. In the case that eCRF data is available for unscheduled visits within the Treatment period, it will also be considered.

Date of last background medication will be collected in the 'COPD Background medication administration' eCRF page of Visit 4. If the subject discontinued the study before Visit 4 or the date of the last background medication is missing or partially missing, but the date of at least one dose of the study drug is recorded in the diaries or at the clinic, the date of the last background medication will be imputed using the following rule:

max [date of last background drug intake in the diaries, date of last background drug intake at the clinic visits].

The evaluation of compliance will be based on the number of puffs following the formula presented for Run-in period.

The total scheduled number of puffs during the Treatment Period is defined as:

- For subjects who had their last background medication dose on the last scheduled dosing day in the period **[i.e. the "Week 6 visit"]** the scheduled number of puffs = Treatment exposure (days) x [scheduled puffs per day] minus [scheduled evening puffs on the last scheduled dosing day]

- For subjects who had their last background medication dose on any day before the last scheduled dosing day in the period **[i.e. an early discontinuation**] the scheduled number of puffs = Treatment exposure (days) x [scheduled puffs per day]

If the date of the last dose of Background treatment = date of first dose of Background treatment, the scheduled number of puffs will be 1 day * [scheduled puffs per day].

The approach above defined for the calculation of compliance assumes no intake of the background medication in case of missing data.

Any level of compliance in the interval [65%; 135%] is considered as satisfactory after randomization. Therefore, the following categories will be presented:

- Treatment compliance < 65%
- 65% ≤ Treatment compliance ≤ 135%
- Treatment compliance > 135%.

The compliance will also be summarized using the following categories for the Run-in Period and also for the Treatment period:

- [0%-10%]
- (10%-20%]
- (20%-30%]
- (30%-40%]
- (40%-50%]
- (50%-60%]
- (60%-70%]
- (70%-80%]
- (80%-90%]
- (90%-100%]
- (100%-110%]

- (110%-120%]
- (120%-130%]
- (130%-140%]
- >140%

### 7.2.8 Treatment Exposure and Compliance - Study Drug Administration

The actual date and time of study drug administration will be used to calculate exposure. For Treatment Arms A, B, C, D and E, the following dosing schedule will be followed for the 6-week treatment period; study drug will be administered twice-a-day (in the morning and in the evening), except for the last day of treatment (Visit 4) when only the morning dose is administered:

Morning administration (between 8-10 am):

One inhalation from the inhaler numbered 1
One inhalation from the inhaler numbered 2

Evening administration (between 8-10 pm):

One inhalation from the inhaler numbered 1
One inhalation from the inhaler numbered 2

If the settings in the electronic diary are incorrect at randomization for a subject (e.g., Run-in period settings are applied during the treatment period), such that data is not collected for study drug administration during some days during the treatment period, then compliance will not be calculated for that subject for days on which the settings were incorrect. Instead, compliance will be calculated based on days during which the settings were correct. The information recorded at the study clinic will always be considered in the calculation (i.e. day will be considered in calculation if exposure is recorded at clinic in a day where diary has an incorrect setting). For these subjects, the duration of the treatment period (as used for calculation of scheduled puffs) will be reduced by the number of days during which the settings were incorrect. In other words, days with incorrect diary settings will be excluded from the numerator and denominator of the compliance calculation for these subjects. If the diary settings were incorrect during the entire treatment period, then compliance will not be calculated for the subject. To be considered as incorrect, diary data should be nonmissing and questions answered consistent with a different period. For example (I=incorrect questions, C=correct questions, M=missing data):

- Day 1: I
- Day 2: I
- Day 3: M
- Day 4: M
- Day 5: C.

In this case, the last "incorrect day" is Day 2.

For Treatment Arm F (*Open Label arm*), there will be 1 morning administration (between 8-10 am): two inhalations of the powder contents of a single SPIRIVA® capsule (18µg) once daily.

The first and last dosing occasions for study drug are expected to be taken during the visit dates, however exposure and compliance are based on actual dosing date/times and not visit dates. See Section 7.1 for the definition of the study periods and period durations (days).

For the calculation of the number of administered and scheduled puffs, the dosing occasions from the first to the last study drug intake will be taken into account (i.e. any study drug recorded in the Subject Diary after the Date of last intake of study drug (as recorded on the 'Study Termination' eCRF page) will be ignored for study drug compliance calculations).

If Date of last intake of study drug (as recorded on the 'Study Termination' eCRF page) is missing or partially missing, dosing occasions of the study drug recorded in the diaries or at the clinic will be taken into account until the 'last randomized study drug intake' as defined in Section 7.1.

Exposure will be calculated as the Date of last intake of study drug – date of first intake of study drug + 1 day.

The evaluation of compliance is done using the following formula:

$$\frac{\text{Total number of administered puffs (as recorded in the Subject Diary and / or Study Drug Administrationat the Clinic eCRF page)}}{\text{Total number of scheduled puffs}} \; x \; 100\%$$

The **total number of administered puffs** will be calculated by adding up the number of puffs taken during the Treatment period, as recorded in the Subject Diary or 'Study Drug Administration at the Clinic' eCRF page.
It should be noted that 2 puffs should be taken twice a day (for treatments A-E) and 2 puffs should be taken once daily (for treatment F) during the Treatment period (from Day 1 (V2) until the day before V4) and 2 puffs should be taken once at V4 for all treatment arms [On the Week 6 visit day, the morning dose occurs in the clinic and there is no evening dose].

The total number of scheduled puffs is defined during the treatment period as:

- For subjects who had their last study drug dose on the last scheduled dosing day **[i.e. the "Week 6 visit"],** the scheduled number of puffs = Treatment exposure (days) x [scheduled puffs per day] minus [scheduled evening puffs on the last scheduled dosing day]

- For subjects who had their last study drug dose on any day before the last scheduled dosing day **[i.e. an early discontinuation**], the scheduled number of puffs = Treatment exposure (days) x [scheduled puffs per day].

- If a subject is randomized into the study and discontinues the same day, then the subject was included in the study for 1 day. Hence, the expected number of puffs would be 1 day * [scheduled puffs per day]

Information on study drug intake is retrieved from the eCRF (on visit days) and subject diaries. If on a visit day information is available from both eCRF and diaries, data entered by the investigator in the eCRF will be taken into account.

The approach above defined for the calculation of compliance assumes no intake of the study drug in case of missing data.

Any level of compliance in the interval [65%; 135%] is considered as satisfactory after randomization. Therefore, the following categories will be discerned for Treatment Period:

- Treatment compliance < 65%
- 65% ≤ Treatment compliance ≤ 135%
- Treatment compliance > 135%.

The compliance will also be summarized as per following categories:

- [0%-10%]
- (10%-20%]
- (20%-30%]
- (30%-40%]
- (40%-50%]
- (50%-60%]
- (60%-70%]
- (70%-80%]
- (80%-90%]
- (90%-100%]
- (100%-110%]
- (110%-120%]
- (120%-130%]
- (130%-140%]
- >140%

### 7.2.9 Compliance with the use of e-Diary

If the setting in the electronic diary are incorrect for a subject (e.g., Run-in settings applied during treatment period or vice-versa), then compliance with use of e-Diary will still be considered.

In these cases, if any questions were answered, the day will count as "data recorded in the [redacted] device".

See Section 7.1 for the definition of the study periods and period durations (days). Period duration will be calculated differently for completed subjects and discontinued subjects as defined in Section 7.1. Since diary data is not expected on the date of Visit 4, Period durations for periods including that date will be shortened by 1 day. If diary data is received for date of Visit 4, it will also be excluded from the numerator of the compliance calculation.

The evaluation of compliance per period (Run-in/Treatment Period) is done using the following formula:

$$\frac{\text{Total number of days in the treatment period with data recorded in the [redacted] device}}{\text{Period duration}} \; x \; 100\%$$

The following categories will be discerned for compliance:

- [0%-10%]
- (10%-20%]
- (20%-30%]
- (30%-40%]
- (40%-50%]
- (50%-60%]
- (60%-70%]
- (70%-80%]
- (80%-90%]
- (90%-100%]

### 7.2.10 Efficacy Variables

● Overview

All spirometry results will be graded by an independent reviewer at [redacted]. The grades after the Best Test Review are: U = Unacceptable, A = Acceptable, and B = Borderline Acceptable. All spirometry values, including the values scored with a grade "Unacceptable", will be considered in the statistical analyses. This follows the approach recommended by the paper by Hankinson et al., where it was concluded that quality assessment regarding the acceptability of individual blows should be primarily used as an aid to good quality during testing rather than a reason to subsequently disregard data.

Spirometry data together with the time of scheduled time points are imported from external [redacted] data.

In addition, the spirometry values excluded from the analysis based on the decisions taken during the Blind Data Review Meeting (e.g., due to technical issues) will be considered to be missing prior to the calculation of the derived variables. In the listings, these assessments will be shown and flagged, to highlight their exclusion from the analyses.

● Primary Efficacy Variable

The primary efficacy variable is the change from baseline in $FEV_1$ $AUC_{0\text{-}12h}$ normalized by time at Visit 4. The baseline value is the average of the pre-dose $FEV_1$ measurements on Day 1. $FEV_1$ is a continuous variable derived from spirometry data ([redacted]). $FEV_1$ $AUC_{0\text{-}12h}$ will be calculated at Week 6. AUC normalized by time will be calculated based on the actual times using the linear trapezoidal rule:

$$\left\{\sum_{\mathbf{i}}[(\mathbf{t_i}-\mathbf{t_{i-1}})(\mathbf{FEV1_i}+\mathbf{FEV1_{i-1}})/\mathbf{2}]\right\}/\mathbf{time}$$

where:

for $t_0$, $FEV1_0$ is the pre-dose average of T -45' and T -15' before study drug administration at that visit

$t_0$ is the actual time of the administration of the first morning dose of study drug,

for i = 15 min, 30 min, 45 min, 1h, 2h, 3h, 4h, 6h, 8h, 10h, 11.5h and 12h, $FEV1_i$ is the actual FEV1 value at each time point and $t_i$ is the actual time of sample i.

*time* is the actual elapsed time from $t_0$ until $t_{12}$.

For the calculation, all valid measurements i are taken and, in addition, the following imputation rules apply:

- If one of the spirometry measurements at 45 min and 15 min pre-dose is not available, then the non-missing measurement will be taken as the pre-dose value (corresponding to time 0 for the calculation of AUC). If no measurement is available, then the pre-dose value will be considered as missing.
- If the pre-dose value is not available, the entire curve will be considered as missing
- The linear trapezoidal rule will allow $AUC_{0\text{-}12h}$ to be calculated if a single, isolated missing value (not pre-dose or last value) exists. This is akin to replacing the missing value by linear interpolation using the adjacent values
- If the value at 12h post-dose is missing, it will be replaced by the value at 11h30; the time will be replaced with the planned 12h post-dose time .
- If two or more consecutive post-dose time points have a missing observation, the entire $FEV_1$ $AUC_{0\text{-}12h}$ for that day will be missing.
- If in total more than three of the post-dose time points have missing values, the $FEV_1$ $AUC_{0\text{-}12h}$ for that day will be missing.
- In case of missing actual time, the theoretical or planned time is used.

If a subject uses rescue medications within the 6-hour wash-out period prior to dosing or during the 12-hour period during which the $FEV_1$ measurements are obtained at Week 6, then all $FEV_1$ measurements will be used in the calculation of the $FEV_1$ $AUC_{0\text{-}12h}$ and analyzed as part of the ITT Population with no exception. In case of rescue medication intake during the serial spirometry (from pre-dose to 12 h post-dose) or during the 6 h wash-out pre-dose, cases will be reviewed during the DRM and documented in the DRR. As a general rule, all spirometry data recorded from the time of rescue medication intake onwards for 6 hours will be excluded from the analysis of primary efficacy endpoint for the PP population.

### *Secondary Efficacy Variables*

If a subject uses rescue medications within the 12-hour period/4-hour period during which spirometry measurements are obtained or during the 6 hour wash out period prior to spirometry then all spirometry measurements will be used and analyzed as part of the ITT Population with no exception.

- 1: *Change from baseline in $FEV_1$ $AUC_{0\text{-}12h}$ normalized by time at Day 1*

$FEV_1$ is a continuous variable derived from spirometry data ([redacted]). $FEV_1$ $AUC_{0\text{-}12h}$ will be calculated on Day 1. The baseline value is the average of the T-45 and T-15 min pre-dose $FEV_1$ measurements on Day 1. The calculation of the $FEV_1$ $AUC_{0\text{-}12h}$ on Day 1 will follow in the same manner as the calculation outlined above for the primary efficacy variable.

- 2a: *Change from baseline in $FEV_1$ $AUC_{0\text{-}4h}$ normalized by time at Day 1 and Week 6*

The baseline value is the average of the pre-dose $FEV_1$ measurements on Day 1. The change from baseline in $FEV_1$ $AUC_{0\text{-}4h}$ normalized by time at Day 1 and Week 6 will be calculated using the following equation:

$$\left\{\sum_{\mathbf{i}}[(\mathbf{t_i}-\mathbf{t_{i-1}})(\mathbf{FEV1_i}+\mathbf{FEV1_{i-1}})/\mathbf{2}]\right\}/\mathbf{time}$$

Where:

for $t_0$, $FEV1_0$ is the pre-dose average of T -45' and T -15' before the time of study drug administration at that visit

$t_0$ is the actual time of the administration of the first morning dose of study drug,

for i = 15 min, 30 min, 45 min, 1h, 2h, 3h and 4h, $FEV1_i$ is the actual FEV1 value at each time point and $t_i$ is the actual time of sample i.

*time* is the actual elapsed time from $t_0$ until *t4*.

For the calculation, all valid measurements i are taken and, in addition, the following imputation rules apply:

- If one of the spirometry measurements at 45 min and 15 min pre-dose is not available, then the non-missing measurement will be taken as the pre-dose value (corresponding to time 0 for the calculation of AUC). If no measurement is available, then the pre-dose value will be considered as missing.
- If the pre-dose value is not available, the entire curve will be considered as missing
- The linear trapezoidal rule will allow $AUC_{0\text{-}4h}$ to be calculated if a single, isolated missing value (not pre-dose or last value) exists. This is akin to replacing the missing value by linear interpolation using the adjacent values
- If the value at 4h post-dose is missing, it will be replaced by the value at 3h.
- If two or more consecutive post-dose time points until 4h have a missing observation, the entire $FEV_1$ $AUC_{0\text{-}4h}$ for that day will be missing.
- If in total three or more post-dose time points until 4h have missing values, the $FEV_1$ $AUC_{0\text{-}4h}$ for that day will be missing.
- In case of missing actual time, the theoretical or planned time is used.

- 2b: *Change from baseline in $FEV_1$ $peak_{0\text{-}4h}$ at Day 1 and Week 6*

$FEV_1$ $peak_{0\text{-}4h}$ at Day 1 and Week 6 will be calculated using the following formula:

Maximum value of all $FEV_1$ values at the time points 15 min, 30 min, 45 min, 1h, 2h, 3h, and 4h of the respective day. If 3 or more missing values are reported among the post-dose spirometry measurements at 45 min, 1h, 2h, 3h, 4h, , the $FEV_1$ $peak_{0\text{-}4h}$ for that day will be missing.

- 2c: *Change from baseline in pre-dose morning $FEV_1$ (average of pre-dose $FEV_1$ measurements) at Week 3 and Week 6*

If one of the spirometry measurements at 45 min and 15 min pre-dose is not available, then the non-missing measurement will be taken as the pre-dose value. If no measurement is available,

then the pre-dose value will be considered as missing. The baseline value is the average of the pre-dose $FEV_1$ measurements on Day 1.

In case of rescue medication intake during the 6 hours wash-out pre-dose, all spirometry data recorded will be considered.

- 3a: *Change from baseline in FVC $AUC_{0-12h}$ normalized by time at Day 1 and Week 6*

FVC is a continuous variable derived from spirometry data ([redacted]). The baseline value is the average of the pre-dose FVC measurements on Day 1. FVC $AUC_{0-12h}$ will be calculated on Day 1 and Week 6. The calculation of the FVC $AUC_{0-12h}$ will follow in the same manner as the calculation outlined above for the primary efficacy variable.

- 3b: *Change from baseline in FVC $AUC_{0-4h}$ normalized by time at Day 1 and Week 6*

FVC $AUC_{0-4h}$ will be calculated on Day 1 and Week 6. The baseline value is the average of the pre-dose FVC measurements on Day 1. The calculation of the FVC $AUC_{0-4h}$ will follow in the same manner as the calculation outlined above for $FEV_1$ $AUC_{0-4h}$.

- 3c: *Change from baseline FVC $peak_{0-4h}$ at Day 1 and Week 6*

FVC $peak_{0-4h}$ at Day 1 and Week 6 will be calculated using the following formula:

Maximum value of all FVC values at the time points 15 min, 30 min, 45 min, 1h, 2h, 3h, and 4h of the respective day. If 3 or more missing values are reported among the post-dose spirometry measurements at 45 min, 1h, 2h, 3h, 4h, the FVC $peak_{0-4h}$ for that day will be missing.

- 4: *Time to onset of action (change from baseline in post-dose $FEV_1 \geq 100$ mL) at Day 1*

In case of at least one post-dose timepoint with change from baseline in FEV1 ≥ 100 mL, then Time to onset of action will be equal to the first post-dose timepoint with change from baseline in FEV1 ≥ 100 mL in the time interval considered (12 h). Time to onset of action will be calculated as the difference in minutes from the actual spirometry time and the time of first drug intake. In case of no timepoint with change from baseline in FEV1 ≥ 100 mL, and less than four missing values among post-dose spirometry measurements at 15 min, 30 min, 45 min, 1h, 2h, 3h and 4h, then the subject will be considered as censored at the last available post-dose timepoint. Subjects with 4 or more of these data points missing will be excluded from the analysis of time to onset of action.

- 5: *Change from baseline in pre-dose morning IC at Week 3 and Week 6*

IC will be measured at V2, V3 and V4 only once at T-45' pre-dose using the SVC maneuver, before the FVC maneuver. If either a baseline (V2) or a post-baseline value is missing, then the change from baseline in IC value will be considered as missing at that visit.

- 6: *TDI focal score at Week 3 and Week 6*

The TDI Focal Score will be calculated within the eCRF from the three individual domain scores (Functional Impairment, Magnitude of Task, and Magnitude of Effort). The BDI and the TDI focal scores will be considered as missing if at least one response is included among the following: "W = Amount Uncertain", "X = Unknown", "Y = Impaired for Reasons other than Shortness of Breath" or "Z = Further Impairment for Reasons other than Shortness of Breath".

- 7: *TDI response (TDI focal score ≥ 1) at Week 3 and Week 6*

A TDI response occurs if the TDI focal score is ≥ 1. Subjects with missing data at the relevant visit will be considered as non-responders.

- 8a: *Change from baseline in percentage of rescue medication-free days during inter-visit periods and entire treatment period*

Information on rescue medication use is collected in the subject-completed e-diary. A patient is assumed to have used rescue medication if the answer to either of the following questions is >0:

- How many inhalations of rescue medication did you take since last evening?
- How many times did you take rescue medication since last evening?

In case of duplicates information reported at clinic and in diary, the greatest information will be used. The change from baseline in the percentage of rescue medication-free days is a continuous variable. The percentage of rescue medication-free days is calculated in each period as

$$\frac{\text{Number of rescue medication} - \text{free days}}{\text{Number of days with available data}} x100\%$$

A minimum of 7 days with available measurements will be required for each inter-visit period (including the run-in period) to consider the percentage of rescue medication-free days as non-missing. The percentage of rescue medication-free days of the entire treatment period will be derived only if both inter-visit period are non-missing. See Section 7.1 for the definition of the inter-visit periods.

- 8b: *Change from baseline in average use of rescue medication (number of puffs/day) during inter-visit periods and entire treatment period*

The change from baseline in the average use of rescue medication is a continuous variable. The average use of rescue medication is calculated in each period as

$$\frac{\sum \text{Number of puffs of rescue medication taken}}{\text{Number of days with available data}}$$

In case of duplicates information reported at clinic and in diary, the greatest information will be used. A minimum of 7 days with available measurements will be required for each inter-visit period (including run-in period) to consider the average use of rescue medication as non-missing. The percentage of rescue medication-free days of the entire treatment period will be

derived only if both inter-visit period are non-missing. See Section 7.1 for the definition of the inter-visit periods.

8c: *Change from baseline in average E-RS total score and domain scores during inter-visit periods and entire treatment period.*

Formerly known as the EXACT-Respiratory Symptom Scale (E-RS), the Evaluating Respiratory Symptoms of COPD measure (E-RS™:COPD) uses 11 respiratory symptom items from the 14-item EXACT. The E-RS total score quantifies respiratory symptom severity and 3 domains: breathlessness, cough and sputum, and chest symptoms. [redacted] The E-RS Total score will be calculated for each day of diary collection. [redacted]:

- [redacted]
- [redacted].

E-RS domain scores will be calculated for each day of diary collection.

[redacted]

- [redacted]
- [redacted]
- [redacted]

[redacted]

- [redacted]
- [redacted]

Higher E-RS scores indicate more severe symptoms and a declining total score indicates health improvement. [redacted] The average score will be calculated as the sum of the scores for the interval being summarized, divided by the number of scores in the interval.

The average E-RS total and domain scores will be calculated as the average of the total/domain scores for subjects with ≥7 days of measurements. A minimum of 7 days with available measurements will be required for each inter-visit period (including run-in period) to consider the average E-RS total/domain score as non-missing. The percentage of rescue medication-free days of the entire treatment period will be derived only if both inter-visit period are non-missing. Section 7.1 for the definition of the inter-visit periods.

### 7.2.11 Safety Variables

There are 6 sets of safety variables identified in the protocol. A description of each variable, and the timing of the recording of the information is presented below.

- 1: Treatment Emergent Adverse Events (TEAEs)

Three categories of AEs will be presented: Pre-Treatment, Treatment Emergent, and Post-Study. An AE will be classified as:

- pre-treatment AE: if AE starts before the first randomized study medication intake (AE onset date < date of first randomized study medication intake).
- Treatment emergent AEs (TEAEs): all adverse events starting on or after the first study drug intake, up to study completion/discontinuation. TEAEs are defined as AEs with date of first randomized study drug intake ≤ AE onset date ≤ date of study completion/discontinuation.
- post-study AE: if AE starts after the date of completion/discontinuation (AE onset date > date of completion/discontinuation).

Additional variables to define an adverse event include the following:

- Serious AE (SAE) is defined as any AE that has the question "Is the AE serious?" marked as "Yes" on the eCRF.
- ADR is defined as any AE that has the question "Is the AE Related to Study Drug?" marked as "Yes" on the eCRF.
- Serious ADR is defined as any ADR that are also SAE
- *Severe AE*, defined as any AE that has intensity marked as "Severe" on the eCRF.
- *AE leading to study drug discontinuation*, defined as a TEAE where the action taken on the eCRF is marked as "Drug withdrawn".
- *AE leading to death*, defined as a TEAE where the outcome is marked as "Fatal" on the eCRF.
- *Relative day of AE onset*, defined as the AE onset date – Date of first study drug intake + 1 if the AE onset date is greater or equal to date of first study drug intake. For pre-treatment AEs, the relative day is defined as the AE onset date – Date of first study drug intake. Relative day will not be calculated in case of partial/missing date of AE onset.
- *Duration of AE*, defined as the AE end date – AE onset date + 1 when the AE is resolved. If the AE is not resolved, the duration is defined as study completion/discontinuation date – AE onset date + 1 and the duration will be reported as ">x" in the listings. AE duration will not be calculated in the case of a partial/missing date of AE onset.

In case of missing or incomplete AE date not directly allowing allocation to any of the categories of AEs, a worst-case allocation will be done according to the available parts of the start and the stop date. The AE will be allocated to the first category allowed by the available data, according to the following order:

1. treatment emergent;

2. post-study;
3. pre-treatment.

- 2: Vital signs (systolic and diastolic blood pressure)

On Day 1 and at Week 6, vital signs (systolic and diastolic blood pressure) will be measured within 1hr pre-dose and at 30 mins, 1.5, 3.5, 7 and 11h post-dose. Blood pressure measurements are continuous variables and recorded in mmHg. At Screening and Early Termination, blood pressure will be measured within 1h before the expected time of study drug administration. At Week 3, blood pressure will be measured within 1h pre-dose. Change from baseline will be calculated using the pre-dose assessment on Day 1 as the baseline value. Change from pre-dose at Week 6 will be calculated using the pre-dose assessment at Week 6.

- 3: 24-hour digital Holter ECG parameters (HR, QTcF, QRS, PR)

Subjects will have an ambulatory 24-hour digital Holter recording for 24 hours before and 24 hours after the 1st dose of study drug (Day 1) and for 24 hours before the last dose of study drug (Week 6).

The changes from baseline will be based on the pre-dose assessment on Day 1, and the results will also be matched by time for 24 hours before and 24 hours after the 1st dose of study drug (Day 1) and for 24 hours before the last dose of study drug (Week 6). For each visit involving Holter, sites will begin the Holter at approximately the same time of day to allow for consistency of data collected at various time points across visits. Results will be presented for each time point. Change from baseline for a given time point will use the corresponding time point at baseline as reference. (e.g., If Visit 2 dosing occurs at 9:30 AM, and the "+55min" ECG on V2 occurs at 10:25 AM, then the +55 minute ECG on V2 will be "matched" to the corresponding pre-dose assessment which occurred on or around 10:25 AM on the day prior to Visit 2. Therefore, change from baseline to "Visit 2 +55 minutes" will be calculated as value at "Visit 2 +55 minutes" – value at "Day prior to V2 10:25 AM").
The following ECG parameters will be recorded: HR, QTcF, QRS and PR.

*For each subject, there will be:*

- *3 nominal extraction times in the day before Visit 2 (time-matched with the post-study drug extraction times on Visit 2)*
- *3 extraction times at +5 min, +55 min, and +2.5h after 1st dose of study drug on Visit 2; and*
- *3 extraction times at +5 min, +55 min, and +2.5h after the morning dose of study drug on the day before Visit 4.*

For each extraction time, ECG parameters will be obtained in triplicate, separate by 30-second intervals. The average of the triplicate will be used in the descriptive statistics and analysis of the Holter ECG results for each extraction time.

- 4: 24-hour HR average, minimum and maximum and hourly average HR using the 24-hour digital Holter ECG monitoring

Three distinct 24-hour monitoring periods will be collected. The first 24-hour monitoring period will start on the day prior to Visit 2. The Holter recordings will be started approximately 30 minutes prior to expected dosing time on the day before Visit 2 and continue until approximately 30 minutes prior to actual dosing time on Visit 2 (Day 1). The second 24-hour monitoring period will start approximately 30 minutes prior to dosing on Visit 2 and continue for 23 hours and 30 minutes after dosing on Visit 2. The third and final 24-hour monitoring period will start on the day prior to Visit 4. The Holter recordings will be started approximately 30 minutes prior to dosing on the day before Visit 4 and continue until approximately 30 minutes prior to dosing on Visit 4.

For 24 hour Holter assessments, records covering <18 hours will not be considered in the analysis.
Hourly heart rate: First and last hour are often <60 minutes. In these cases, if <30 mins are available, the record will not be used. If >=30 minutes are available, then it will be used.

For each of the distinct 24-hour digital Holter monitoring periods, the recordings for HR will be used to derive the 24-hour HR average, minimum and maximum HR (bpm). In addition, the average HR will be calculated on an hourly basis during each 24-hour monitoring period.

The following parameters will also be obtained from the 24 hour Holter monitor results:

- Longest Tachycardia Duration (min);
- Longest Tachycardia Maximum HR (bpm);
- Fastest Tachycardia Duration (min);
- Fastest Tachycardia Maximum HR (bpm);
- Longest Bradycardia Duration (min);
- Longest Bradycardia Minimum HR (bpm);
- Slowest Bradycardia Duration (min);
- Slowest Bradycardia Minimum HR (bpm);
- Time in Atrial Fibrillation Percentage;
- Atrial Fibrillation Peak Average Rate (bpm);
- Supraventricular Total;
- Ventricular Total;
- RR > 2 sec (yes/no: yes if number of RR > 2 sec > 0, no if number of RR > 2 sec = 0).

- 5: 24-hour digital Holter ECG abnormal findings

The number and percentage of subjects exhibiting arrhythmia will be summarized. The Holter ECG recordings will be scanned for the presence or absence of arrhythmia including, but not limited to:

- Atrial fibrillation,
- Sinus Pauses,
- PACs,
- SVT,
- PVCs,

- Ventricular tachycardia,
- AV Blocks.

For 24 hour Holter assessments, records covering <18 hours will not be considered in the analysis. All abnormal findings will be listed.

- 6: Standard blood chemistry and hematology

Blood samples will be collected for hematology and serum chemistry at Screening, Week 6 and Early Termination in the morning.

The following evaluations will be performed using a central laboratory:

- *Hematology test: red blood cells count (RBC), white blood cells count (WBC) and differential, total hemoglobin (Hb), hematocrit (Hct), platelets count (PLT).*
- *Serum chemistry test: blood urea nitrogen (BUN), cholesterol, triglycerides, creatinine, uric acid, aspartate aminotransferase (AST), alanine aminotransferase (ALT), Gamma-glutamyl transferase (γ-GT), total bilirubin, alkaline phosphatases, albumin, total proteins, and electrolytes (sodium, potassium, calcium, and chloride).*
- *Fasting blood glucose*
- *Serum pregnancy test (serum β-hCG) in women of child-bearing potential;* results will be presented as positive or negative. Urine β-hCG from visits 1, 2, and 3 will also be presented as positive or negative.

**Complete Blood Count** (abbreviation followed by the full name and units)

| | |
|---|---|
| HCT: | Hematocrit (L/L) |
| HGB: | Hemoglobin (g/L) |
| PLATE: | Platelets (10^9/L) |
| RBC: | Erythrocytes (10^12/L) |
| WBC: | Leucocytes (10^9/L) |

**Differential** (abbreviation followed by the full name and units)

| | |
|---|---|
| BASO: | Basophil (10^9/L) and (%) |
| EOSIN: | Eosinophil (10^9/L) and (%) |
| LYMPH: | Lymphocyte (10^9/L) and (%) |
| MONO: | Monocyte (10^9/L) and (%) |
| NEUT: | Neutrophil (10^9/L) and (%) |

**Blood Chemistry** (abbreviation followed by the full name and units)

| | |
|---|---|
| ALB: | Albumin (g/L) |
| ALKPH: | Alkaline Phosphatase (U/L) |
| ALT: | ALT (U/L) |
| AST: | AST (U/L) |
| BUN: | Blood Urea Nitrogen (mmol/L) |
| BILDIR: | Direct Bilirubin (µmol/L) |
| CA: | Calcium (mmol/L) |
| CHOL: | Cholesterol (mmol/L) |
| CL: | Chloride (mmol/L) |
| CREAT: | Creatinine (µmol/L) |
| GGT: | Gamma-GT (U/L) |

GLUC: Fasting Serum Glucose (mmol/L)
K: Potassium (mmol/L)
NA: Sodium (mmol/L)
PROT: Total Protein (g/L)
TRIG: Triglycerides (mmol/L)
URATE: Urate (μmol/L)

Results from the central laboratory will include the out of range flag based on the lower and upper limits of normal range. Categories will be

- Low CS
- Low NCS
- Normal
- High NCS
- High CS, and
- Missing.

# 8 Analysis Methods

## 8.1 General Methods

Hypothesis testing will be carried out at the alpha = 0.05 level (two-sided). Statistical significance will be declared if the p-value is less than 0.05, with p-values for the primary efficacy analysis adjusted for multiplicity as described in Section 8.2.1.4.

For **continuous variables**, summary statistics will include the number of non-missing observations, mean, standard deviation, 95% (or otherwise) confidence interval of the mean (only in the efficacy and safety analyses, unless otherwise specified), minimum, median and maximum.

For **qualitative variables**, the number (n) and percentage (%) of subjects with non-missing data per category will be the default summary presentation, and where appropriate and present, the number of missing values as a "Missing" category. Unless otherwise specified, the denominator for each percentage will be the number of non-missing observations within the analysis set and treatment group.

General rules to be used for reporting the number of decimal places for derived variables in the listings (in the analyses rounding will not be performed):

- BMI, treatment exposure (days): whole numbers;
- Time to discontinuation (weeks), duration of smoking (years), time since first COPD diagnosis (years), compliance, average use of rescue medication (number of puffs/day), percentage of rescue medication-free days, average E-RS scores: 1 decimal place;
- Change from baseline/pre-dose: same as the variable considered.

The following rules on decimal places will be considered for the results of the analyses (if the analyses are performed on derived variables, the level of precision of the actual data is derived from the previous list):

- min, max: same as actual data;
- mean and its confidence limits (unadjusted and adjusted), adjusted difference between means and its confidence limits, SD, median: actual data + 1 decimal place;
- percentage: 1 decimal place;
- Kaplan-Meier percentiles estimates and confidence limits: actual data + 1 decimal place (3 decimal places for probabilities).
- hazard ratio and its confidence limits, odds ratio and its confidence limits: 2 decimal places;
- P-values: 3 decimal places; if the p-value is less than 0.001, it will be presented as <0.001.

In general, additional to the analysis model results presented in TLFs, the full SAS output generated for all the analyses will be presented to the sponsor as a standalone document. SAS output results will be generated only after DB lock.

### 8.1.1 Rules for Handling Missing Data

Handling of missing data for the statistical analysis is described below:

*For the primary efficacy analysis, a linear mixed model for repeated measures will be used to handle missing data. Under the Missing At Random (MAR) assumption, this model provides an unbiased estimate of the treatment effect that would have been observed if all subjects had continued on treatment for the full study duration. This approach addresses efficacy (or de jure) hypotheses, estimating the causal effects of the initially randomized drug if taken as directed (in contrast with effectiveness, or de facto, hypotheses, evaluating the effect of the drug as actually taken). The efficacy estimand is considered appropriate in the context of a phase II study aiming at characterizing dose-response.*

*In the TDI responder analysis, subjects with missing data at the relevant visit will be considered as non-responders*. In order to avoid any exclusion from the responder analysis, for subjects with missing baseline value (i.e., with baseline value still missing after having applied ad-hoc rules potentially defined in the Data Review Report), the baseline value will be imputed as the overall (i.e., calculated jointly considering all treatment groups) mean baseline BDI focal score value calculated on subjects with available data on ITT population.

The specific methods to deal with missing data, in particular with partial or missing dates (e.g., of medication use), missing spirometry data and missing questionnaire data, are laid out in detail in the description of the variable derivation in the subsections under Section 7.2.

Other critical missing data, if any, will be discussed during the blind review of the data. Decisions will be fully documented in the Data Review Report.

**Calculation of adjusted means (least squares means):**

The approach described below will ensure that the least squares means calculated by SAS will be based on:

- coefficients for classification effects (i.e., the effects of categorical covariates) proportional to the margins observed in the group of subjects analysed;
- effects of quantitative covariates set equal to their mean values in the group of subjects analysed.

The analysis is based on the following steps:

1. generate a dataset by selecting:
   - in case of repeated post-randomisation measurements (e.g., $FEV_1$ at each visit, analysed using a linear mixed model for repeated measures): all the post-randomization records for subjects with at least one available and valid post-randomisation measurement and no missing covariates;
2. in case of repeated post-randomisation measurements, add to the dataset generated in step 1 the records for missing post-randomisation visits of the subjects included in the dataset. In the added records the value of the response variable will be missing, but the full information on covariates has to be included;
3. use the dataset obtained as the input dataset for the MIXED or the GENMOD procedure, specifying the following options in the LSMEANS statement:
   - in case of repeated post-randomisation measurements: OM AT MEANS;

– in case of single post-randomisation measurement: OM.

Example: analysis of change from baseline (Visit 2) at all visits (Visits 3 to 5) based on a mixed model for repeated measures including the effects of treatment, visit (categorical variable), treatment by visit interaction, baseline and another covariate. Visit 1 represents the screening visit.

Original dataset (X = available value, . = missing or invalid value):

| Patient | Treatment | Covariate | Baseline | Visit | Change from baseline |
|---|---|---|---|---|---|
| 1 | A | X | X | 1 | . |
| 1 | A | X | X | 2 | . |
| 1 | A | X | X | 3 | X |
| 1 | A | X | X | 4 | X |
| 1 | A | X | X | 5 | X |
| 2 | B | X | X | 1 | . |
| 2 | B | X | X | 2 | . |
| 2 | B | X | X | 3 | X |
| 3 | A | X | X | 1 | . |
| 3 | A | X | X | 2 | . |
| 3 | A | X | X | 3 | X |
| 3 | A | X | X | 4 | . |
| 3 | A | X | X | 5 | X |
| 4 | B | X | X | 1 | . |
| 4 | B | X | X | 2 | . |
| 5 | A | . | X | 1 | . |
| 5 | A | . | X | 2 | . |
| 5 | A | . | X | 3 | X |

Step 1 (visits 1 and 2 not selected since pre-randomisation, subject 4 not selected due to missing post-randomisation measurements, subject 5 not selected due to missing covariate):

| Subject | Treatment | Covariate | Baseline | Visit | Change from baseline |
|---|---|---|---|---|---|
| 1 | A | X | X | 3 | X |
| 1 | A | X | X | 4 | X |
| 1 | A | X | X | 5 | X |
| 2 | B | X | X | 3 | X |
| 3 | A | X | X | 3 | X |
| 3 | A | X | X | 4 | . |
| 3 | A | X | X | 5 | X |

Step 2 (added records in *italic*):

| Subject | Treatment | Covariate | Baseline | Visit | Change from baseline |
|---|---|---|---|---|---|
| 1 | A | X | X | 3 | X |
| 1 | A | X | X | 4 | X |
| 1 | A | X | X | 5 | X |
| 2 | B | X | X | 3 | X |
| *2* | *B* | *X* | *X* | *4* | . |

| | | | | | |
|---|---|---|---|---|---|
| *2* | *B* | *X* | *X* | *5* | . |
| 3 | A | X | X | 3 | X |
| 3 | A | X | X | 4 | . |
| 3 | A | X | X | 5 | X |

## 8.2 Specific Methods for Efficacy Analyses

**Analyses Based on Mixed Models**

### Primary efficacy endpoint

The change from baseline in $FEV_1$ $AUC_{0\text{-}12h}$ normalized by time at week 6 will be analyzed using a linear mixed model for repeated measurements including treatment, visit, treatment by visit interaction, US regions and smoking status at screening as fixed effects, and the baseline value (average of the pre-dose $FEV_1$ measurements on Day 1) and baseline by visit interaction as covariates. An unstructured covariance matrix will be assumed; the generic code for the model is presented below.

```
proc mixed data = <xxx>;
 class subject treatment visit usregion smoking;
 model <response> = treatment visit treatment*visit usregion smoking fev1base fev1base*visit /
  DDFM=KR;
 repeated visit / subject=subject type=UN;
 lsmeans treatment*visit / cl OM AT MEANS;
 lsmestimate treatment*visit 'Trt A vs Placebo week 6' 0 1 0 0 0 0 0 0 0 -1 0 0,
                'Trt B vs Placebo week 6' 0 0 0 1 0 0 0 0 0 -1 0 0,
                'Trt C vs Placebo week 6' 0 0 0 0 0 1 0 0 0 -1 0 0,
                'Trt D vs Placebo week 6' 0 0 0 0 0 0 0 1 0 -1 0 0 / cl adjust = simulate (seed = 12311980
acc = 0.0001);
 lsmestimate treatment*visit 'Trt D vs Trt A week 6' 0 -1 0 0 0 0 0 1 0 0 0 0/cl;
 lsmestimate treatment*visit 'Trt D vs Trt B week 6' 0 0 0 -1 0 0 0 1 0 0 0 0/cl;
 lsmestimate treatment*visit 'Trt D vs Trt C week 6' 0 0 0 0 0 -1 0 1 0 0 0 0/cl;
 lsmestimate treatment*visit 'Trt C vs Trt A week 6' 0 -1 0 0 0 1 0 0 0 0 0 0/cl;
 lsmestimate treatment*visit 'Trt C vs Trt B week 6' 0 0 0 -1 0 1 0 0 0 0 0 0/cl;
 lsmestimate treatment*visit 'Trt B vs Trt A week 6' 0 -1 0 1 0 0 0 0 0 0 0 0/cl;
 lsmestimate treatment*visit 'Trt F vs Placebo week 6' 0 0 0 0 0 0 0 0 0 0 -1 0 1/cl;
run;
```

where:

- *treatment* is the randomized treatment variable, codified and sorted as follow:
    1. Treatment A (GB 12.5 µg)
    2. Treatment B (GB 25 µg)
    3. Treatment C (GB 50 µg)
    4. Treatment D (GB 100 µg)
    5. Placebo
    6. Treatment F (Spiriva 18 µg)
- *visit* is the visit variable (Visit 2 [Day 1], and Visit 4 [week 6])
- *treatment*visit* is the treatment by visit interaction
- *usregion* is the region in the United States from which the subject entered the study (Midwest, Northeast, South and West Regions).
- *smoking* is the smoking status at baseline (ex-smoker, current smoker).
- *fev1base* is the baseline $FEV_1$ result

- *fev1base*visit* is the baseline $FEV_1$ result by visit interaction.

The adjusted means in each treatment group, the adjusted mean difference between treatments and their 95% Confidence Intervals (CIs) will be estimated by the model. In addition, the p-values for each of the factors in the model will be provided.

The CIs and the p-values of the comparisons between each dose of CHF 5259 pMDI and Placebo at Week 6 will be adjusted for multiplicity. For the primary endpoint, the adjustment will be based on the parametric simulation method by Edwards and Berry (adjust=simulate in the MODEL statement). The random number seed will be set equal to 12311980, and 0.0001 will be used for the target accuracy radius gamma (i.e. ACC=0.0001). At each dose level, the superiority of CHF 5259 pMDI will be demonstrated by a **statistically significant** difference (adjusted p-value < 0.05) favoring CHF 5259 pMDI.

The model will include also pairwise comparisons of treatment A through treatment D, comparing each dose with a lower dose, and treatment F versus Placebo. The CIs and the p-values of the comparisons between those pairwise comparisons at Week 6 will not be adjusted for multiplicity.

Kenward Rogers method will be used to approximate denominator degrees of freedom (/DDFM=KR in the MODEL statement); OM and AT MEANS in LSMEANS statement will be used to account for unbalanced population (see details in 8.1).

Sensitivity analysis:

Primary analysis will be repeated for sensitivity analysis excluding the subject that has been randomized few seconds after another subjects receiving the same random number.

On [redacted], two subjects [redacted] and [redacted], were assigned the same Randomization Number (Rand ID), [redacted]. Both subjects were randomized between [redacted] and [redacted] on [redacted].

The Subject report shows Rand ID [redacted] assigned to both subjects. However, the Subject Randomization report shows ID [redacted] being assigned to subject [redacted] and not to subject [redacted]. Subject [redacted] does not appear on the Subject Randomization report.

The subject [redacted] will be excluded from the sensitivity analysis for ITT population.

**Secondary efficacy endpoints**

For additional secondary efficacy analyses ($FEV_1$ $AUC_{0-12}$ at Day 1, $FEV_1$ $AUC_{0-4}$, $FEV_1$ peak 0-4, Pre-dose morning $FEV_1$, FVC $AUC_{0-12}$, FVC $AUC_{0-4}$, FVC peak 0-4, Pre-dose morning IC at week 3 and week 6, TDI focal score, Percentage of rescue medication free-days, Average use of rescue medication and E-RS total score and domain score) all the comparisons between treatments will be performed with no multiplicity adjustment.

The adjusted means in each treatment group, the adjusted mean difference between treatments and their 95% Confidence Intervals (CIs) and p-values will be estimated by the model. In addition, the p-values for each of the factors in the model will be provided.

Kenward Rogers method will be used to approximate denominator degrees of freedom (/DDFM=KR in the MODEL statement); OM and AT MEANS in LSMEANS statement will be used to account for unbalanced population.

The SAS code for the secondary efficacy endpoints using the same model as the primary efficacy analysis will be as below.

```
proc mixed data = <xxx>;
 class subject treatment visit usregion smoking;
 model <response> = treatment visit treatment*visit usregion smoking base base*visit / DDFM=KR;
 repeated visit / subject=subject type=UN;
 lsmeans treatment*visit / cl OM AT MEANS;
 lsmestimate treatment*visit 'Trt A vs Placebo day 1' 1 0 0 0 0 0 0 0 -1 0 0 0 /cl;
 lsmestimate treatment*visit 'Trt B vs Placebo day 1' 0 0 1 0 0 0 0 0 -1 0 0 0 /cl;
 lsmestimate treatment*visit 'Trt C vs Placebo day 1' 0 0 0 0 1 0 0 0 -1 0 0 0 /cl;
 lsmestimate treatment*visit 'Trt D vs Placebo day 1' 0 0 0 0 0 0 1 0 -1 0 0 0 /cl;
 lsmestimate treatment*visit 'Trt D vs Trt A day 1'   -1 0 0 0 0 0 1 0 0 0 0 0 /cl;
 lsmestimate treatment*visit 'Trt D vs Trt B day 1'   0 0 -1 0 0 0 1 0 0 0 0 0 /cl;
 lsmestimate treatment*visit 'Trt D vs Trt C day 1'   0 0 0 0 -1 0 1 0 0 0 0 0 /cl;
 lsmestimate treatment*visit 'Trt C vs Trt A day 1'   -1 0 0 0 1 0 0 0 0 0 0 0 /cl;
 lsmestimate treatment*visit 'Trt C vs Trt B day 1'   0 0 -1 0 1 0 0 0 0 0 0 0 /cl;
 lsmestimate treatment*visit 'Trt B vs Trt A day 1'   -1 0 1 0 0 0 0 0 0 0 0 0 /cl;
 lsmestimate treatment*visit 'Trt F vs Placebo day 1' 0 0 0 0 0 0 0 0 -1 0 1 0 /cl;
run;
```

where:

- *treatment* is the randomized treatment variable, codified and sorted as follow:
    1. Treatment A (GB 12.5 µg)
    2. Treatment B (GB 25 µg)
    3. Treatment C (GB 50 µg)
    4. Treatment D (GB 100 µg)
    5. Placebo
    6. Treatment F (Spiriva 18 µg)
- *visit* is the visit variable (Visit 2 [Day 1] <if applicable>, Visit 3 [week 3] <if applicable>, and Visit 4 [week 6])
- *treatment*visit* is the treatment by visit interaction
- *usregion* is the region in the United States from which the subject entered the study (Midwest, Northeast, South and West Regions).
- *smoking* is the smoking status at baseline (ex-smoker, current smoker).
- *base* is the baseline result, as defined in Section 7 per each variable.
- *base*visit* is the baseline result by visit interaction.

In the sas code model reported above, the LSMESTIMATE statement shows comparisons of the treatments at Day 1 for a variable assessed at Day 1 and Week 6. This statement will be adapted for week 3 or week 6 as appropriate for each secondary variables.

For the change from baseline in $FEV_1$ $AUC_{0\text{-}12h}$ normalized by time at Day 1, the adjusted means in each treatment group and the adjusted mean differences between treatments at Day 1 will be estimated with their 95% CIs and p-values by the model outlined above for the secondary efficacy endpoints.

Change from baseline in $FEV_1$ $AUC_{0\text{-}4h}$ normalized by time and in $FEV_1$ $peak_{0\text{-}4h}$ at Day 1 and Week 6 and change from baseline in pre-dose morning $FEV_1$ at Week 3 and Week 6 will be analyzed using the model outlined above for the secondary efficacy endpoints.

Change from baseline in FVC $AUC_{0\text{-}12h}$ normalized by time, in FVC $AUC_{0\text{-}4h}$ normalized by time and in FVC $peak_{0\text{-}4h}$ at Day 1 and Week 6 will be analyzed using the model outlined above for the secondary efficacy endpoints.

The change from baseline in pre-dose morning IC at Week 3 and Week 6 will be analyzed using the model outlined above for the secondary efficacy endpoints.

TDI focal score at Week 3 and Week 6 will be analyzed using the model outlined above for the secondary efficacy endpoints.

The change from baseline to each inter-visit period in percentage of rescue medication-free days, in average use of rescue medication and in average E-RS total score and domain scores will be analyzed using a similar model outlined above for the secondary efficacy endpoints. The inter-visit period will be considered instead of visit in the model. Comparison between treatments over the entire treatment period will also be derived from this model. In this comparison, equal weights will be assigned to the two inter-visit periods. The SAS code will be as below.

```
proc mixed data = <xxx>;
 class subject treatment inter_visit usregion smoking;
 model <response> = treatment  inter_visit treatment*inter_visit usregion  smoking  base  base*inter_visit /
  DDFM=KR;
 repeated  inter_visit / subject=subject  type=UN;
 lsmeans treatment  treatment*inter_visit / cl OM AT MEANS;
 lsmestimate treatment  "Trt A vs Placebo entire tmt period"  1 0 0 0 -1 0  /cl;
 lsmestimate treatment  "Trt B vs Placebo entire tmt period"  0 1 0 0 -1 0  /cl;
 lsmestimate treatment  "Trt C vs Placebo entire tmt period"  0 0 1 0 -1 0  /cl;
 lsmestimate treatment  "Trt D vs Placebo entire tmt period"  0 0 0 1 -1 0  /cl;
 lsmestimate treatment  "Trt B vs Trt A entire tmt period"    -1 1 0 0 0 0  /cl;
 lsmestimate treatment  "Trt C vs Trt A entire tmt period"    -1 0 1 0 0 0  /cl;
 lsmestimate treatment  "Trt D vs Trt A entire tmt period"    -1 0 0 1 0 0  /cl;
 lsmestimate treatment  "Trt C vs Trt B entire tmt period"    0 -1 1 0 0 0  /cl;
 lsmestimate treatment  "Trt D vs Trt B entire tmt period"    0 -1 0 1 0 0  /cl;
 lsmestimate treatment  "Trt D vs Trt C entire tmt period"    0 0 -1 1 0 0  /cl;
 lsmestimate treatment  "Trt F vs Placebo entire tmt period"  0 0 0 0 -1 1  /cl;
 lsmestimate treatment*inter_visit 'Trt A vs Placebo period 1' 1  0  0  0  0  0  0  0  -1  0  0  0 /cl;
 lsmestimate treatment*inter_visit 'Trt B vs Placebo period 1' 0  0  1  0  0  0  0  0  -1  0   0  0 /cl;
 lsmestimate treatment* inter_visit 'Trt C vs Placebo period 1' 0  0  0  0  1  0  0  0  -1  0  0  0 /cl;
 lsmestimate treatment* inter_visit 'Trt D vs Placebo period 1' 0  0  0  0  0  0  1  0  -1  0  0  0 /cl;
 lsmestimate treatment* inter_visit 'Trt D vs Trt A period 1'     -1  0  0  0  0  0  1  0  0  0  0  0 /cl;
 lsmestimate treatment* inter_visit 'Trt D vs Trt B period 1'     0  0  -1  0  0  0  1  0  0  0  0  0 /cl;
 lsmestimate treatment* inter_visit 'Trt D vs Trt C period 1'     0  0  0  0  -1  0  1  0  0  0  0  0 /cl;
 lsmestimate treatment* inter_visit 'Trt C vs Trt A period 1'     -1  0  0  0  1  0  0  0  0  0  0  0 /cl;
 lsmestimate treatment* inter_visit 'Trt C vs Trt B period 1'     0 0  -1  0  1  0  0  0  0  0  0  0 /cl;
 lsmestimate treatment* inter_visit 'Trt B vs Trt A period 1'     -1  0  1  0  0  0  0  0  0  0  0  0 /cl;
 lsmestimate treatment* inter_visit 'Trt F vs Placebo period 1' 0  0  0  0  0  0  0  0  -1  0  1  0 /cl;
lsmestimate  treatment* inter_visit "Trt A vs Placebo period 2" 0 1    0 0    0 0    0 0   0 -1   0 0 /cl;
…
 lsmestimate treatment* inter_visit  "Trt F vs Placebo period 2" 0 0    0  0    0  0    0  0   0  -1   0 1 /cl;
run;
```

## Analyses Based on the Time to Onset

A Cox proportional hazards model will be utilized to analyze the time to onset of action in $FEV_1$. The Cox proportional hazards model will include treatment, US Region and smoking status at screening as fixed effects, and baseline (average of pre-dose $FEV_1$ measurements on Day 1) as a covariate. The generic code for the Cox model is:

```
proc phreg data=<xxx>;
  class treatment usregion smoking;
  model time*censor(0) = treatment usregion smoking fev1base / risklimits ties=exact;
  contrast "Treatment A vs Placebo" treatment      1 0 0 0 -1  /estimate=exp;
  contrast "Treatment B vs Placebo" treatment      0 1 0 0 -1  /estimate=exp;
  contrast "Treatment C vs Placebo" treatment      0 0 1 0 -1  /estimate=exp;
  contrast "Treatment D vs Placebo" treatment      0 0 0 1 -1  /estimate=exp;
  contrast "Treatment D vs Treatment A" treatment -1 0 0 1 0  /estimate=exp;
  contrast "Treatment D vs Treatment B" treatment 0 -1 0 1 0  /estimate=exp;
  contrast "Treatment D vs Treatment C" treatment 0 0 -1 1 0  /estimate=exp;
  contrast "Treatment C vs Treatment A" treatment -1 0 1 0 0  /estimate=exp;
  contrast "Treatment C vs Treatment B"  treatment 0 -1 1 0 0  /estimate=exp;
  contrast "Treatment B vs Treatment A" treatment -1 1 0 0 0  /estimate=exp;
  contrast "Treatment F vs Placebo "         treatment 0 0 0 0 -1 /estimate=exp;
run;
```

The generic code for the Kaplan-Meier plot is:

```
proc lifetest data=<xxx>;
    time time*censor(0);
    strata treatment;
run;
```

where:

- *treatment* is the randomized treatment variable, codified and sorted as follow:
    1. Treatment A (GB 12.5 µg)
    2. Treatment B (GB 25 µg)
    3. Treatment C (GB 50 µg)
    4. Treatment D (GB 100 µg)
    5. Placebo
    6. Treatment F (Spiriva 18 µg)
- *usregion* is the region in the United States from which the subject entered the study (Midwest, Northeast, South, and West Regions).
- *smoking* is the smoking status at baseline (ex-smoker, current smoker).
- *fev1base* is the baseline $FEV_1$ result

A censor variable value of 0 will be used for subjects who do not experience the onset of action in $FEV_1$. If the option ties=exact requires a considerable amount of computer resources, the Efron approximation will be used (ties=efron).

The 1st quartile, median, and 3rd quartile for the time to onset of action in $FEV_1$ will be provided along with the number and percentage of subjects who were censored. The hazard ratios, 95% confidence intervals for the hazard-ratios, and p-values comparing treatment groups will be provided. P-values will not be adjusted for multiplicity. In addition, the p-values for each of the factors in the model will be provided.

The Kaplan-Meier plot will be repeated also for time to discontinuation from study.

**Analyses Based on logistic regression model**

A multiple logistic regression model will be utilized to analyze TDI Response (TDI Focal Score ≥ 1). The logistic regression model will model the probability of a subject successfully achieving the TDI Response. The model will include treatment, US Region, and smoking status as fixed effects and the Baseline Dyspnea Index (BDI) as a covariate. The logistic regression model will be based on the results from the Week 3 and Week 6 time points separately. The generic code for the model is:

```
proc genmod data = <xxx>;
 class treatment usregion smoking;
 model TDI_focal_response = treatment usregion smoking BDI / dist=binomial wald type3;
  lsmestimate treatment  "Treatment A vs. Placebo" 1 0 0 0 -1 0/ exp cl;
 lsmestimate treatment  "Treatment B vs. Placebo" 0 1 0 0 -1 0/ exp cl;
 lsmestimate treatment  "Treatment C vs. Placebo" 0 0 1 0 -1 0/ exp cl;
 lsmestimate treatment  "Treatment D vs. Placebo" 0 0 0 1 -1 0/ exp cl;
 lsmestimate treatment "Treatment D vs. Treatment A" -1 0 0 1 0 0/ exp cl;
 lsmestimate treatment "Treatment D vs. Treatment B " 0 -1 0 1 0 0/ exp cl ;
 lsmestimate treatment "Treatment D vs. Treatment C "0 0 -1 1 0 0/ exp cl;
 lsmestimate treatment "Treatment C vs. Treatment A" -1 0 1 0 0 0/ exp cl;
 lsmestimate treatment "Treatment C vs. Treatment B" 0 -1 1 0 0 0/ exp cl;
 lsmestimate treatment "Treatment B vs. Treatment A" -1 1 0 0 0 0/ exp cl;
 lsmestimate treatment "Treatment F vs. Placebo"0 0 0 0 -1 1/ exp cl;
run;
```

where:

- *treatment* is the randomized treatment variable, codified and sorted as follow:
    1. Treatment A (GB 12.5 µg)
    2. Treatment B (GB 25 µg)
    3. Treatment C (GB 50 µg)
    4. Treatment D (GB 100 µg)
    5. Placebo
    6. Treatment F (Spiriva 18 µg)
- *usregion* is the region in the United States from which the subject entered the study (Midwest, Northeast, South, and West Regions).
- *smoking* is the smoking status at baseline (ex-smoker, current smoker).
- *BDI* is the baseline result

The odds ratios, 95% confidence intervals for the odd-ratios, and p-values comparing treatment groups will be provided. P-values will not be adjusted for multiplicity. In addition, the odds ratios, 95% confidence intervals, and p-values for each of the factors in the model will be provided.

### 8.2.1 Statistical/Analytical Issues

#### 8.2.1.1 Handling of Dropouts or Missing Data

The rules to address missing data rules for specific variables are described in Section 7.2 and 7.1.1.

#### 8.2.1.2 Blind Review

A Blind Data Review Meeting is organized just before data base lock to finalize the list of protocol deviations which will exclude subjects from the Per Protocol analysis and also to address handling of any outstanding data issues. The Data Review Report will describe all agreed handling and will be finalized prior to data base lock and unblinding. The Data Review Plan will describe the tables, listings and figures used during this meeting, and falls outside the scope of the SAP.

#### 8.2.1.3 Multi-center Studies

No adjustment is done for the multiple centers, and no analysis per center is carried out. The adjustment in the model will be US region.

#### 8.2.1.4 Multiple Comparisons/Multiplicity

Adjustment for multiplicity is applied only to the primary efficacy analysis. The CIs and the p-values of the comparisons between each dose of CHF 5259 pMDI and placebo for the primary efficacy endpoint at Week 6 will be adjusted for multiplicity. The adjustment will be based on the parametric simulation method by Edwards and Berry. At each dose level, the superiority of CHF 5259 pMDI will be demonstrated by a **statistically significant** difference (adjusted p-value < 0.05) favoring CHF 5259 pMDI.

### 8.3 Specific Methods for Safety Analyses

**Analyses Based on Mixed Models**

The SAS code for the safety endpoints change from baseline (day before V2) in 24-hour average, minimum and maximum HR extracted from the Holter recordings will be:

```
proc mixed data = <xxx>;
  class subject treatment visit usregion smoking;
  model <response> = treatment  visit treatment*visit usregion  smoking  base  base*visit / DDFM=KR;
  repeated  visit / subject=subject  type=UN;
  lsmeans treatment*visit / cl OM AT MEANS alpha=0.05
lsmestimate treatment*visit "Trt A vs Placebo Visit 2"  1 0 0 0 0 0 0 0 -1 0 0 0  /cl alpha=0.05;
lsmestimate treatment*visit "Trt A vs Placebo Visit 4"  0 1 0 0 0 0 0 0 0 -1 0 0  /cl alpha=0.05;
…
lsmestimate treatment *visit "Trt D vs Placebo Visit 2"  0 0 0 0 0 0 1 0 -1 0 0 0  /cl alpha=0.05;
lsmestimate treatment *visit "Trt D vs Placebo Visit 4"  0 0 0 0 0 0 0 1 0 -1 0 0  /cl alpha=0.05;

lsmestimate treatment *visit "Trt B vs Trt A Visit 2"    -1 0 1 0 0 0 0 0 0 0 0 0  /cl alpha=0.05;
lsmestimate treatment *visit "Trt B vs Trt A Visit 4"    0 -1 0 1 0 0 0 0 0 0 0 0  /cl alpha=0.05;
…
```

```
lsmestimate treatment *visit "Trt F vs Placebo Visit 2"  0 0 0 0 0 0 0 0 1 0 -1 0   /cl alpha=0.05;
lsmestimate treatment *visit "Trt F vs Placebo Visit 4"  0 0 0 0 0 0 0 0 0 1 0 -1   /cl alpha=0.05;
run;
```

where:

- *treatment* is the randomized treatment variable, codified and sorted as follow:
    1. Treatment A (GB 12.5 µg)
    2. Treatment B (GB 25 µg)
    3. Treatment C (GB 50 µg)
    4. Treatment D (GB 100 µg)
    5. Placebo
    6. Treatment F (Spiriva 18 µg)
- *visit* is the visit variable (Visit 2 [Day 1] and Visit 4 [week 6])
- *treatment*visit* is the treatment by visit interaction
- *usregion* is the region in the United States from which the subject entered the study (Midwest, Northeast, South and West Regions).
- *smoking* is the smoking status at baseline (ex-smoker, current smoker).
- *base* is the baseline time-matched baseline as defined in Section 7.2.12
- *base*visit* is the baseline result by visit interaction.
- Treatment comparisons will be extended to comparisons between each dose of CHF 5259 pMDI and Placebo at Visit 2 and Visit 4 and all pairwise comparisons of treatment A through treatment D at Visit 2 and Visit 4.

The SAS code for the safety endpoints change from baseline (time-matched on day before V2) in ECG parameters extracted from the 24hr Holter (HR, QTcF, QRS and PR) and change from baseline (day before V2) in 24-hour average, minimum and maximum HR extracted from the Holter recordings will be:

```
proc mixed data = <xxx>;
 class subject treatment timePoint usregion smoking;
 model <response> = treatment  timePoint  treatment*timePoint usregion  smoking  base  base*timePoint
  timeavgBase / DDFM=KR;
 repeated  timePoint / subject=subject  type=UN;
 lsmeans treatment*timePoint   / cl OM AT MEANS alpha=0.1;
 lsmestimate treatment*timePoint "Trt A vs Placebo 1"   -1 0 0 0 0 0  0 0 0 0 0 0  0 0 0 0 0 0   0 0 0 0 0 0  1 0 0 0
0 0   0 0 0 0 0 0 /cl alpha=0.1;
 lsmestimate treatment*timePoint "Trt A vs Placebo 2"   0 -1 0 0 0 0  0 0 0 0 0 0  0 0 0 0 0 0   0 0 0 0 0 0  0 1 0 0
0 0   0 0 0 0 0 0 /cl alpha=0.1;
 lsmestimate treatment*timePoint "Trt A vs Placebo 3"   0 0 -1 0 0 0  0 0 0 0 0 0  0 0 0 0 0 0   0 0 0 0 0 0  0 0 1 0
0 0   0 0 0 0 0 0 /cl alpha=0.1;
 lsmestimate treatment*timePoint "Trt A vs Placebo 4"   0 0 0 -1 0 0  0 0 0 0 0 0  0 0 0 0 0 0   0 0 0 0 0 0  0 0 0 1
0 0   0 0 0 0 0 0 /cl alpha=0.1;
 lsmestimate treatment*timePoint "Trt A vs Placebo 5"   0 0 0 0 -1 0  0 0 0 0 0 0  0 0 0 0 0 0   0 0 0 0 0 0  0 0 0 0
1 0   0 0 0 0 0 0 /cl alpha=0.1;
 lsmestimate treatment*timePoint "Trt A vs Placebo 6"   0 0 0 0 0 -1  0 0 0 0 0 0  0 0 0 0 0 0   0 0 0 0 0 0  0 0 0 0
0 1   0 0 0 0 0 0 /cl alpha=0.1;

....

run;
```

where:

- *treatment* is the randomized treatment variable, codified and sorted as follow:
    1. Treatment A (GB 12.5 µg)
    2. Treatment B (GB 25 µg)
    3. Treatment C (GB 50 µg)
    4. Treatment D (GB 100 µg)
    5. Placebo
    6. Treatment F (Spiriva 18 µg)
- *timePoint* is the time point variable: 6 time points scheduled at Visit 2 +5m, +55m and +2.5h and Visit 4 +5m, +55m and +2.5h.
- *treatment*timePoint* is the treatment by time point interaction
- *usregion* is the region in the United States from which the subject entered the study (Midwest, Northeast, South and West Regions).
- *smoking* is the smoking status at baseline (ex-smoker, current smoker).
- *base* is the baseline time-matched baseline as defined in Section 7.2.12
- *base*timePoint* is the baseline result by visit interaction
- timeavgBase is the time-averaged baseline score be the average of the Day -1 scores at +5m, +55m and +2.5h.

Treatment comparisons will be extended for all 6 time points.

# 9 Interim analyses

Not applicable.

# 10 Overview of Tables, Listings and Figures

Demographics, COPD and smoking history, spirometry and reversibility will be presented using all three populations (ITT, Safety, PP). Medical history/concomitant diseases and compliance with study treatment will be presented using the ITT and Safety populations. Medications, CAT score, BDI, compliance use of e-diary, and protocol deviations will be presented using the ITT population. Vital signs, ECG, AEs and laboratory data will be presented using the Safety population.

## 10.1 Disposition of Subjects

A summary of the number of screened subjects, the number of screening failures and reasons for screening failure will be produced for all enrolled subjects.

A disposition summary of subjects by treatment and overall will include the number (N) of subjects randomized with the number (n) and percentage (%) of subjects who completed or discontinued the study. A subject is considered as completed if the Study Termination Form in the eCRF has “Completed” checked for the question “Specify the subject’s status”. All percentages will be based on the number of subjects randomized.

The reasons for study discontinuation will be also summarized by treatment using absolute counts and percentages, based on the number of subjects randomized. . A Kaplan Meier plot will be reported by treatment .

## 10.2 Protocol Deviations

The number of subjects who had at least one Major Protocol Deviation and all Major Protocol Deviations will be summarized by absolute counts (n) and percentages (%) by treatment for the Safety population. Percentages will be based on the number of subjects in the ITT population. The summary will be repeated for Minor Protocol Deviations.

## 10.3 Analysis Sets

The number of subjects in each population will be summarized by treatment and overall, by absolute counts (n) and percentages (%) for each treatment group and by US region. Percentages will be based on the number of randomized subjects.

## 10.4 Demographic and Other Baseline Characteristics

### 10.4.1 Demographic Characteristics, COPD History and Smoking Habits

Demographic data, COPD history and smoking habits will be summarized by summary statistics and absolute counts (n) and percentages (%) by treatment and overall for the Safety population, the ITT population and PP population.

### 10.4.2 Medical History

Medical and Surgical History and Concomitant Diseases will be summarized on a per-subject basis (i.e. if a subject reported the same event repeatedly the event will be counted only once at the specific level of display). Absolute counts (n) and percentages (%) by treatment and overall will be presented for the number of subjects with at least one Medical or Surgical History or Concomitant Disease, and per SOC and per PT within SOC, for the ITT and Safety population. Percentages will be based on the number of subjects in the population.

### 10.4.3 Baseline Subject Characteristics

Central spirometry and reversibility test data, Physical examination results, ECG results at Visit 1 and Visit 2 pre-dose, Baseline Dyspnea Index Questionnaire at Visit 1, and the COPD Assessment Test at Screening and Visit 2 will be summarized with summary statistics and absolute counts (n) and percentages (%) by treatment and overall for the populations described in Section 10, above.

### 10.4.4 Previous, Maintained and Concomitant Medications

Previous and maintained medications will be summarized by treatment and overall, while concomitant medications will be summarized by treatment (for the assignment of medications to a treatment, see section 7.2.7).

All previous, maintained, and concomitant medications will be presented in the tables and the listings.

Post-treatment medications will be only presented in a listing.

Previous, maintained and concomitant medications will be summarized on the ITT population according to Anatomical Main Group (1st level of ATC), Therapeutic Subgroup (2nd level of ATC), Chemical Subgroup (4th level of ATC) and preferred name.

Absolute counts (n) and percentages (%) will be presented for the number of subjects taking at least one medication, and per Anatomical Main Group and per Therapeutic Subgroup within Anatomical Main Group and per Chemical Subgroup within Therapeutic Subgroup, for the ITT populations. Percentages will be based on the number of subjects in the population.

## 10.5 Study drug / Investigational Medicinal Product

### 10.5.1 Background medication

Daily dose of previous ICS treatment for COPD will be summarized by descriptive statistics using absolute counts (n) and percentages (%) of subjects according to the dose category (Low Daily Dose and Medium Daily Dose), for the ITT population.

Similarly, descriptive statistics using only the subjects placed on QVAR® 80 µg will be used to calculate the QVAR Estimated Clinical Comparable Dose (µg/day), and the number of morning and evening puffs that the subject was instructed to take will be summarized using descriptive statistics.

Exposure and compliance using only the subjects placed on QVAR® 80 µg will be summarized by summary statistics and absolute counts (n) and percentages (%) by the treatment that the subject will be randomized to and overall during the Run-in Period, and by treatment during the Treatment Period, for ITT populations. In addition, the number and percentage of subjects with either satisfactory (65% to 135%) and unsatisfactory levels (<65% or > 135%) of compliance to the background medication will be summarized by the treatment the subject was randomized to and overall during the Run-in Period and by treatment and overall during Treatment Period for the ITT population. Compliance to background medication will also be categorized in the same manner in the following categories: [0%-10%], (10%-20%], (20%-30%], (30%-40%], (40%-50%], (50%-60%], (60%-70%], (70%-80%], (80%-90%], (90%-100%], (100%-110%], (110%-120%], (120%-130%], (130%-140%], >140%.

### 10.5.2 Study drug

Treatment exposure and compliance during the Treatment Period will be summarized by summary statistics and absolute counts (n) and percentages (%) by treatment using the ITT and Safety populations. In addition the number and percentage of subjects with satisfactory (65% - 135%) and unsatisfactory (<65%; >135%) levels of compliance to the study drug will be summarized by treatment using the ITT and Safety populations. Compliance to study medication will also be categorized in the same manner in the following categories: [0%-10%], (10%-20%], (20%-30%], (30%-40%], (40%-50%], (50%-60%], (60%-70%], (70%-80%], (80%-90%], (90%-100%], (100%-110%], (110%-120%], (120%-130%], (130%-140%], >140%.

## 10.6 Efficacy Results

### 10.6.1 Primary efficacy variable

The primary endpoint is the change from baseline in $FEV_1$ $AUC_{0\text{-}12h}$ at Week 6. The change from baseline in $FEV_1$ $AUC_{0\text{-}12h}$ normalized by time will be summarized using descriptive statistics and analyzed using a linear mixed model for repeated measurements as outlined in Section 8.2. The adjusted means in each treatment group and their corresponding 95% confidence intervals (CI), the adjusted mean difference between treatments (active – placebo) and their 95% CIs and p-values will be estimated by the model. The CIs and the p-values of the comparisons between each dose of CHF 5259 pMDI and placebo at Week 6 will be adjusted for multiplicity. The adjustment will be based on the parametric simulation method by Edwards and Berry. At each dose level, the superiority of CHF 5259 pMDI versus placebo will be demonstrated by a statistically significant difference (adjusted p-value < 0.05) favoring CHF 5259 pMDI. In addition, the p-values for each of the factors in the model will be displayed. Comparisons between the various active doses will be presented based on the model, but not adjusted for multiplicity.

The analysis will be performed using the ITT population (primary analysis) and will be repeated using the PP population.

The mean absolute $FEV_1$ values and the change form baseline will be summarized using descriptive statistics with the 95% CI of the mean and will be presented by treatment for the ITT and PP population.

The mean absolute $FEV_1$ values recorded on Day 1 and Week 6 will be plotted over time for each of the treatment groups. These figures will be based on data from subjects in both ITT and PP populations (separate graphs for ITT and PP).

The mean change from baseline $FEV_1$ values on Day 1 and Week 6 will be plotted over time for each of the treatment groups. These figures will be based on data from subjects in both ITT and PP populations (separate graphs for ITT and PP).

### 10.6.2 Secondary efficacy variables

The following secondary efficacy endpoints will be summarized using descriptive statistics with the 95% CI of the mean and will be presented by treatment for the ITT population; in addition the following analyses will be performed. There will be no adjustment for multiplicity to the CIs or the p-values for any of the secondary efficacy endpoints. The mean change from baseline of secondary efficacy endpoints values will be plotted over time for each of the treatment groups.

**Analysis of Secondary Efficacy Variable 1**: *Change from baseline in $FEV_1$ $AUC_{0\text{-}12h}$ normalized by time at Day 1*

The change from baseline in $FEV_1$ $AUC_{0\text{-}12h}$ normalized by time on Day 1 will be analyzed using a linear mixed model for repeated measurements as outlined in Section 8.2. The adjusted means in each treatment group with their corresponding 95% CIs, and the adjusted mean difference between treatments with their corresponding 95% CIs will be estimated by the model. In addition, the p-values for each of the factors in the model will be displayed.

**Analysis of Secondary Efficacy Variable 2a**: *Change from baseline in $FEV_1$ $AUC_{0-4h}$ normalized by time at Day 1 and Week 6*

The change from baseline in $FEV_1$ $AUC_{0-4h}$ normalized by time on Day 1 and Week 6 will be analyzed using a linear mixed model for repeated measurements as outlined in Section 8.2. The adjusted means in each treatment group with their corresponding 95% CIs, and the adjusted mean difference between treatments with their corresponding 95% CIs will be estimated by the model. In addition, the p-values for each of the factors in the model will be displayed.

**Analysis of Secondary Efficacy Variable 2b**: *Change from baseline in $FEV_1$ $peak_{0-4h}$ normalized by time at Day 1 and Week 6*

The change from baseline in $FEV_1$ $peak_{0-4h}$ normalized by time on Day 1 and Week 6 will be analyzed using a linear mixed model for repeated measurements as outlined in Section 8.2. The adjusted means in each treatment group with their corresponding 95% CIs, and the adjusted mean difference between treatments with their corresponding 95% CIs will be estimated by the model. In addition, the p-values for each of the factors in the model will be displayed.

**Analysis of Secondary Efficacy Variable 2c**: *Change from baseline in pre-dose morning $FEV_1$ (average of pre-dose $FEV_1$ measurements) at Week 3 and Week 6*

The change from baseline in pre-dose morning $FEV_1$ normalized by time at Week 3 and Week 6 will be analyzed using a linear mixed model for repeated measurements as outlined in Section 8.2. The adjusted means in each treatment group with their corresponding 95% CIs, and the adjusted mean difference between treatments with their corresponding 95% CIs will be estimated by the model. In addition, the p-values for each of the factors in the model will be displayed.

**Analysis of Secondary Efficacy Variable 3a**: *Change from baseline in FVC $AUC_{0-12h}$ normalized by time at Day 1 and Week 6*

The change from baseline in FVC $AUC_{0-12h}$ normalized by time on Day 1 and Week 6 will be analyzed using a linear mixed model for repeated measurements as outlined in Section 8.2. The adjusted means in each treatment group with their corresponding 95% CIs, and the adjusted mean difference between treatments with their corresponding 95% CIs will be estimated by the model. In addition, the p-values for each of the factors in the model will be displayed.

The mean absolute FVC values and the change form baseline will be summarized using descriptive statistics with the 95% CI of the mean and will be presented by treatment for the ITT population.

**Analysis of Secondary Efficacy Variable 3b**: *Change from baseline in FVC $AUC_{0-4h}$ normalized by time at Day 1 and Week 6*

The change from baseline in FVC $AUC_{0-4h}$ normalized by time on Day 1 and Week 6 will be analyzed using a linear mixed model for repeated measurements as outlined in Section 8.2. The adjusted means in each treatment group with their corresponding 95% CIs, and the adjusted mean difference between treatments with their corresponding 95% CIs will be estimated by the model. In addition, the p-values for each of the factors in the model will be displayed.

**Analysis of Secondary Efficacy Variable 3c**: *Change from baseline in FVC peak$_{0\text{-}4h}$ at Day 1 and Week 6*

The change from baseline in FVC peak$_{0\text{-}4h}$ normalized by time on Day 1 and Week 6 will be analyzed using a linear mixed model for repeated measurements as outlined in Section 8.2. The adjusted means in each treatment group with their corresponding 95% CIs, and the adjusted mean difference between treatments with their corresponding 95% CIs will be estimated by the model. In addition, the p-values for each of the factors in the model will be displayed.

**Analysis of Secondary Efficacy Variable 4**: *Time to onset of action (change from baseline in post-dose FEV$_1$ ≥ 100 mL) at Day 1 (min)*

Time to onset of action (i.e., change from baseline in post-dose $FEV_1 \geq 100$ mL at Day 1) will be analyzed using a Cox proportional hazard model as outlined in Section 8.2. The 1st quartile, median, and 3rd quartile for the time to onset of action in $FEV_1$ will be provided along with the number and percentage of subjects who were censored. The hazard ratios, 95% confidence intervals for the hazard-ratios, and p-values comparing each active treatment to placebo will be provided. In addition, the p-values for each of the factors in the model will be provided. Comparisons between the various active doses will be presented based on the model.

A Kaplan-Meier plot of the time to onset of action will also be presented.

**Analysis of Secondary Efficacy Variable 5**: *Change from baseline in pre-dose morning IC at Week 3 and Week 6*

The change from baseline in pre-dose morning IC at Week 3 and Week 6 will be analyzed using a linear mixed model for repeated measurements as outlined in Section 8.2. The adjusted means in each treatment group with their corresponding 95% CIs, and the adjusted mean difference between treatments with their corresponding 95% CIs will be estimated by the model. In addition, the p-values for each of the factors in the model will be displayed.

**Analysis of Secondary Efficacy Variable 6**: *TDI focal score at Week 3 and Week 6*

The TDI focal score at Week 3 and Week 6 will be analyzed using a linear mixed model for repeated measurements as outlined in Section 8.2. The adjusted means in each treatment group with their corresponding 95% CIs, and the adjusted mean difference between treatments with their corresponding 95% CIs will be estimated by the model. In addition, the p-values for each of the factors in the model will be displayed.

**Analysis of Secondary Efficacy Variable 7**: *TDI response (TDI focal score ≥ 1) at Week 3 and Week 6*

The TDI response at Week 3 and Week 6 will be analyzed using a multiple logistic regression model as outlined in Section 8.2. The odds ratios, 95% confidence intervals for the odd-ratios, and p-values comparing each active treatment to placebo will be provided. In addition, the odds ratios, 95% confidence intervals, and p-values for each of the factors in the model will be provided. The model will be run for Week 3 and Week 6 separately. Comparisons between the various active doses will be presented based on the model.

**Analysis of Secondary Efficacy Variable 8a**: *Change from baseline in percentage of rescue medication-free days during inter-visit periods and entire treatment period*

The change from baseline to each inter-visit period and to the entire treatment period in the percentage of rescue medication-free days will be analyzed using a linear mixed model for repeated measurements as outlined in Section 8.2. See Section 7.1 for the definition of the inter-visit periods. The adjusted means in each treatment group with their corresponding 95% CIs, and the adjusted mean difference between treatments with their corresponding 95% CIs will be estimated by the model. In addition, the p-values for each of the factors in the model will be displayed.

**Analysis of Secondary Efficacy Variable 8b**: *Change from baseline in average use of rescue medication (number of puffs/day) during inter-visit periods and entire treatment period*

The change from baseline to each inter-visit period and to the entire treatment period in the average use of rescue medication will be analyzed using a linear mixed model for repeated measurements as outlined in Section 8.2. See Section 7.1 for the definition of the inter-visit periods. The adjusted means in each treatment group with their corresponding 95% CIs, and the adjusted mean difference between treatments with their corresponding 95% CIs will be estimated by the model. In addition, the p-values for each of the factors in the model will be displayed.

**Analysis of Secondary Efficacy Variable 8c**: *Change from baseline in average E-RS total score and domain scores during inter-visit periods and entire treatment period*

The change from baseline to each inter-visit period and to the entire treatment period in the average E-RS total score and the domain scores will be analyzed using a linear mixed model for repeated measurements as outlined in Section 8.2. See Section 7.1 for the definition of the inter-visit periods. The adjusted means in each treatment group with their corresponding 95% CIs, and the adjusted mean difference between treatments with their corresponding 95% CIs will be estimated by the model. In addition, the p-values for each of the factors in the model will be displayed.

### 10.6.3 Other efficacy variables

PEF and $FEF_{25\%-75\%}$ are collected together with the spirometry assessments and will be presented in data listings only.

## 10.7 Safety Analyses

All safety analyses described below will be performed on the Safety population.

### 10.7.1 Adverse Events

As defined in the protocol, all AEs starting on or after the date of first study drug intake will be classified as TEAEs. Only TEAEs will be included in the tables. All AEs will be listed. AE onset date will be taken into account when assigning AEs as *treatment emergent*. Pre-treatment AEs (AE onset date < date of first randomized study drug intake) and post-study AEs (AE onset

date > date of completion/discontinuation) will be included in the subject listings, and flagged, but will be excluded from other summaries.

The number of TEAEs, ADRs, serious TEAEs, serious ADRs, severe TEAEs, TEAEs leading to study drug discontinuation and TEAEs leading to death, and the number and the percentage of subjects experiencing TEAEs, ADRs, serious TEAEs, serious ADRs, severe TEAEs, TEAEs leading to study drug discontinuation and TEAEs leading to death will be summarized by treatment and overall.

The number and percentage of subjects with at least one AE and the number of AEs will be presented by SOC and PT by treatment and overall for treatment-emergent AEs, ADRs, serious TEAEs, serious ADRs, severe TEAEs, TEAEs leading to study drug discontinuation and TEAEs leading to death.

Subjects who have multiple events in the same system organ class (SOC) and preferred term (PT) will be counted only once in the subject counts.

### 10.7.2 Vital Signs

Vital signs (systolic and diastolic blood pressure) and their changes from baseline (pre-dose on Day 1) and from pre-dose on Week 6 (for Week 6 post-dose only) will be summarized by treatment using descriptive statistics and the 95% CI of the mean.

### 10.7.3 ECGs

10.7.3.1 12-Lead ECGs

Recorded data from the local 12-Lead ECG will be presented in a listing.

10.7.3.2 Holter ECGs

ECG parameters extracted from the 24hr Holter (HR, QTcF, QRS, and PR) will be summarized for the +5 min, +55 min, and +2.5 hour nominal extraction times on Day -1, Day 1 and Week 6 by treatment using descriptive statistics and the 95% CI of the mean. The nominal recorded value will be used in the calculation. In addition, their changes from baseline (time-matched on the day before V2) will be summarized using descriptive statistics and 90% CI. The summaries will use the triplicate averages obtained at each nominal extraction time.

Change from baseline (time-matched on day before V2) in ECG parameters extracted from the 24hr Holter (HR, QTcF, QRS and PR) will be analyzed using a linear mixed model for repeated measurements including treatment, time point, treatment by time point interaction, US regions, and smoking status at screening as fixed effects, and the baseline value (time-matched on day before V2), baseline by time point interaction and time-averaged baseline as covariates. An unstructured covariance matrix will be assumed. The adjusted means in each treatment group with their corresponding 90% CIs, and the adjusted mean difference between all treatments and placebo with their corresponding 90% CIs will be estimated by the model. In addition, the p-values for each of the factors in the model will be displayed.

In addition, the number and percentage of subjects with prolonged QTcF will be summarized using counts and percentages for each nominal extraction and at any post-dose time point using the following thresholds:

- QTcF >450 ms for males and >470 ms for females
- QTcF >480 ms for males only
- QTcF >500 ms
- Change from baseline in QTcF >30 ms
- Change from baseline in QTcF >60 ms

For each of the three distinct 24-hour monitoring periods, the HR average, minimum, and maximum will be extracted from the Holter recordings and summarized on Day -1, Day 1, and Week 6 by treatment using descriptive statistics and the 95% CI of the mean. In addition, their changes from baseline will be summarized using descriptive statistics and 95% CI.

Change from baseline (day before V2) in 24-hour average, minimum and maximum HR extracted from the Holter recordings will be analyzed using a linear mixed model for repeated measurements including treatment, visit, treatment by visit interaction, US regions and smoking status at screening as fixed effects, and the baseline value (day before V2) and baseline by visit interaction as covariates. An unstructured covariance matrix will be assumed. The adjusted means in each treatment group with their corresponding 95% CIs, and the adjusted mean difference between treatments with their corresponding 95% CIs will be estimated by the model. In addition, the p-values for each of the factors in the model will be displayed.

The average HR will be calculated on an hourly basis during each 24-hour monitoring period and summarized by treatment using descriptive statistics and the 95% CI of the mean.

The Holter ECG recordings will be scanned for the presence or absence of arrhythmia including, but not limited to:

- Atrial fibrillation,
- Sinus Pauses,
- PACs,
- SVT,
- PVCs,
- Ventricular tachycardia,
- AV Blocks.

In addition, the following will be summarized by treatment using descriptive statistics and the 95% CI of the mean:

- Longest Tachycardia Duration (min);
- Longest Tachycardia Maximum HR (bpm);
- Fastest Tachycardia Duration (min);
- Fastest Tachycardia Maximum HR (bpm);
- Longest Bradycardia Duration (min);
- Longest Bradycardia Minimum HR (bpm);
- Slowest Bradycardia Duration (min);
- Slowest Bradycardia Minimum HR (bpm);

- Time in Atrial Fibrillation Percentage;
- Atrial Fibrillation Peak Average Rate (bpm);
- Supraventricular Total;
- Ventricular Total;

RR > 2 sec (yes/no: yes if number of RR > 2 sec > 0, no if number of RR > 2 sec = 0) will be summarized using counts and percentages.

For ECG analyses, numerical parameters (HR, QTcF, PR and QRS) will not be included in the statistical analysis in the following cases:

- patients with a pacemaker already in place at study entry, identified by the presence of at least one of the following Preferred Terms in the medical/surgical history or concomitant diseases: “Cardiac pacemaker battery replacement”, “Cardiac pacemaker evaluation”, “Cardiac pacemaker insertion”, “Cardiac pacemaker replacement”, “Electrocardiogram pacemaker spike”, “Pacemaker generated arrhythmia”, “Pacemaker generated rhythm”, “Pacemaker syndrome”, “Cardiac assistance device user”;
- patients with a pacemaker implanted during the study, identified by a procedure coded with the Preferred Term “Cardiac pacemaker insertion” (other relevant cases may be identified in the Data Review Report). In this case, only the parameters assessed in a date ≥ start date of the procedure will be excluded from the statistical analysis;
- patients with atrial fibrillation as concomitant disease, identified by the presence of at least one of the following Preferred Term: “Atrial fibrillation”, “Cardiac fibrillation”.
- ECGs with PR=0, since this is indicative of poor quality or inevaluble ECG result. Note that this does not apply to all ECGs for the patient, but only the ECGs with PR=0.

### 10.7.4 Clinical Laboratory Evaluation

Shift tables will be presented for hematology and biochemistry parameters by treatment group at Day 1 and Week 6.

The categories for the shift tables are as follows:

- Low clinically significant
- Low not clinically significant
- Normal
- High not clinically significant
- High clinically significant

In addition, all laboratory data will be listed, with abnormal values flagged.

## 10.8 Early Termination and Unscheduled Visits

The handling of the early termination efficacy and safety assessments in the statistical analyses will be discussed during the Data Review Meeting and the decisions will be fully documented in the Data Review Report.

With regards to unscheduled assessments,these measurements will be evaluated case by case during the Data Review Meeting, with handling described in the Data Review Report.

The following rules on data re-allocation will be considered:

- Data collected at multiple visits (spirometry, TDI, vital signs, laboratory data, and Holter ECG) recorded at the study termination visit for discontinued patients will be re-allocated by selecting the next visit at which the assessment was planned. For each assessment, only the visits at which the assessment was scheduled will be considered for re-allocation. This means that Holter ECG and laboratory data can be re-allocated only to Visit 4. For example, if the last visit performed before the study termination visit was Visit 2, the data recorded at the study termination visit will be re-allocated to Visit 3 for spirometry, TDI, and vital signs, and to Visit 4 for Holter ECG and laboratory data. If the study termination visit was performed less than 7 days after the preceding visit, then the data recorded at the study termination visit will not be re-allocated and they will excluded from the statistical analysis.
- For discontinued patients, efficacy data recorded in the diaries from the last visit performed before the study termination visit or the date of discontinuation onwards will be reallocated to the next expected inter-visit period;

- in case of missing intermediate visit not due to the re-allocation of data collected at the study termination visit (e.g., Visit 3 missing, but Visits 2 and 4 performed), an expected date for the missing Visit 3 will be imputed in order to define the inter-visit periods for diary data. The expected date for the missing Visit 3 will be imputed as follows: Visit 3 expected date = Visit 2 date + 28 days.

- Additional unscheduled/optional spirometries will be considered on a case-by-case basis and their inclusion in the analysis will be discussed at the DRM and confirmed in the DRR prior to data base lock and unblinding. As a general rule, unscheduled or optional FEV1 and FVC results which are recorded during the 12 hours post-dose at Visit 2 and Visit 4 will be considered in the AUC0-12h calculations, while results which are recorded during the 4 hours post-dose at Visit 2 and Visit 4 will be considered in the AUC0-4h and peak0-4h calculationsand as eligible for minimum/maximum assessments unless otherwise specified in the DRR.

- Holter ECG: In case of the memory card was not removed and replaced during Visit 2, such that 48 hours are recorded at the Visit 2, Day -1 time point, then the first 24 hours of data following the dosing time (first inhalation) will be re-allocated to the Visit 2 time point.

- Laboratories: the last assessment before the first randomised study medication intake of each parameter will be considered as from Visit 1 in the analysis. For WBC and the differential count parameters (lymphocytes, neutrophils, monocytes, eosinophils, basophils) the last complete assessment (i.e., with available measurements for all these parameters) before first study drug intake will be considered in the analysis. If no complete

assessment is available, the last assessment before first study drug intake with the highest number of available parameters will be considered in the analysis.

Potential issues of the approach above defined and other decisions regarding data re-allocation will be evaluated during the blind review of the data and documented in the Data Review Report.

# 11 Table of Contents for the Tables, Listings and Figures

Key Frist Results (KFR) are identified with an "*".

**Post-Text Tables**

## CSR Section 14.1: Demographic data



## CSR Section 14.2: Efficacy Data

Table 14.2.1.3: Summary of FEV1 AUC0-12h Normalized by Time and the Change from Baseline at Week 6 (PP Population)

* Table 14.2.1.4: Analysis of FEV1 AUC0-12h Normalized by Time at Week 6 - (PP Population)

Table 14.2.2.1: Summary of FEV1 AUC0-12h Normalized by Time and the Change from Baseline on Day 1 (ITT Population)

Table 14.2.2.2: Analysis of FEV1 AUC0-12h Normalized by Time on Day 1 (ITT Population)

* Table 14.2.3.1: Analysis of FEV1 AUC0-12h Normalized by Time on Day 1 (ITT Population)

* Table 14.2.4.1.1: Summary of FEV1 at Each Time Point on Day 1, Week 3, and Week 6 (ITT Population)

Table 14.2.4.1.2: Summary of FEV1 at Each Time Point on Day 1, Week 3, and Week 6 (PP Population)

Table 14.2.4.2: Summary of Change from Baseline in FVC at Each Time Point on Day 1 and Weeks 3 and 6 (ITT Population)

Table 14.2.5.1: Summary of FEV1 AUC0-4h Normalized by Time and the Change from Baseline on Day 1 and Week 6 (ITT Population)

* Table 14.2.5.2: Analysis of FEV1 AUC0-4h Normalized by Time on Day 1 and Week 6 (ITT Population)

Table 14.2.6.1: Summary of FEV1 Peak0-4h and the Change from Baseline on Day 1 and Week 6 (ITT Population)

* Table 14.2.6.2: Analysis of FEV1 Peak0-4h on Day 1 and Week 6 (ITT Population)

* Table 14.2.7.1: Analysis of Pre-Dose Morning FEV1 at Week 3 and Week 6 (ITT Population)

Table 14.2.8.1: Summary of Time to Onset of Action (Change from Baseline in Post-dose FEV1 ≥ 100 mL) on Day 1 (ITT Population)

Table 14.2.8.2: Analysis of Time to Onset of Action (Change from Baseline in Post-dose FEV1 ≥ 100 ml) on Day 1 (ITT Population)

Table 14.2.9.1: Summary of FVC at Each Time Point on Day 1, Week 3, and Week 6 (ITT Population)

Table 14.2.10.1: Summary of Change from Baseline in IC at Each Pre-Dose Time Point on Day 1, Week 3, and Week 6 (ITT Population)

Table 14.2.11.1: Summary of FVC AUC0-12h Normalized by Time and the Change from Baseline on Day 1 and Week 6 (ITT Population)

Table 14.2.11.2: Analysis of FVC AUC0-12h Normalized by Time on Day 1 and Week 6 (ITT Population)

Table 14.2.12.1: Summary of FVC AUC0-4h Normalized by Time and the Change from Baseline on Day 1 and Week 6 (ITT Population)

Table 14.2.12.2: Analysis of FVC AUC0-4h Normalized by Time on Day 1 and Week 6 (ITT Population)

Table 14.2.13.1: Summary of FVC Peak0-4h and the Change from Baseline on Day 1 and Week 6 (ITT Population)

Table 14.2.13.2: Analysis of FVC Peak0-4h on Day 1 and Week 6 (ITT Population)

Table 14.2.14.1: Summary of IC at Each Pre-dose Time Point on Day 1, Week 3, and Week 6 (ITT Population)

Table 14.2.14.2: Summary of Change from Baseline in IC at Each Time Point on Day 1 and Weeks 3 and Week 6 (ITT Population)

Table 14.2.14.3: Analysis of Pre-Dose Morning IC at Week 3 and Week 6 (ITT Population)

Table 14.2.15.1: Summary of the TDI Domains, Focal Score, and Response (TDI Focal Score ≥ 1 Unit) at Weeks 3 and 6 (ITT Population)

* Table 14.2.15.2: Analysis of the TDI Focal Score at Weeks 3 and 6 (ITT Population)

Table 14.2.16.1: Analysis of the TDI Response (TDI Focal Score ≥ 1) at Weeks 3 and 6 (ITT Population)

Table 14.2.17.1: Summary of the Percentage of Rescue Medication-free Days during Inter-Visit Periods and the Entire Treatment Period (ITT Population)

Table 14.2.17.2: Analysis of the Percentage of Rescue Medication-free Days during Inter-Visit Periods and the Entire Treatment Period (ITT Population)

Table 14.2.18.1: Summary of the Average Use of Medication during Inter-Visit Periods and the Entire Treatment Period (ITT Population)

Table 14.2.18.2: Analysis of the Average Use of Rescue Medication during Inter-Visit Periods and the Entire Treatment Period (ITT Population)

Table 14.2.19.1: Summary of the E-RS Total and Domain Scores and the Change from Baseline during the Inter-Visit Periods and the Entire Treatment Period (ITT Population)

Table 14.2.19.2: Analysis of the E-RS Total during the Inter-Visit Periods and the Entire Treatment Period (ITT Population)

## CSR Section 14.3: Safety

* Table 14.3.1.1.1: Summary of Treatment Emergent Adverse Events (Safety Population)
Table 14.3.1.2.1: Treatment Emergent Adverse Events by System Organ Class and Preferred Term (Safety Population)
Table 14.3.1.2.2: Serious Treatment Emergent Adverse Events by System Organ Class and Preferred Term (Safety Population)
* Table 14.3.1.3.1: Adverse Drug Reactions by System Organ Class and Preferred Term (Safety Population)
Table 14.3.1.3.2: Serious Adverse Drug Reactions by System Organ Class and Preferred Term (Safety Population)
Table 14.3.1.3.3: Severe Adverse Drug Reactions by System Organ Class and Preferred Term (Safety Population)
Table 14.3.1.4: Treatment Emergent Adverse Events Leading to Study Drug Discontinuation by System Organ Class and Preferred Term (Safety Population)
Table 14.3.2: Treatment Emergent Adverse Events Leading to Death by System Organ Class and Preferred Term (Safety Population)
Table 14.3.4.1: Vital Signs Results and Changes from Baseline (Safety Population)
Table 14.3.4.2: Vital Signs Changes from Pre dose (Safety Population)
Table 14.3.5.1.1: Summary of the 24-hour Holter ECG Results and Changes from Baseline (Safety Population)
Table 14.3.5.1.2: Analysis of the 24-hour Holter ECG Results (Safety Population)
Table 14.3.5.2.1: 24-hour Holter ECG – Prolonged QTcF (Safety Population - Males)
Table 14.3.5.2.2: 24-hour Holter ECG - Prolonged QTcF (Safety Population - Females)
Table 14.3.5.2.3: 12-Lead ECG - Prolonged QTcF – Changes from Baseline (Safety Population)
Table 14.3.5.3.1: Summary of the 24-hour Holter ECG Average, Minimum, and Maximum Heart Rate Results and Changes from Baseline (Safety Population)
Table 14.3.5.3.2: Analysis of the 24-hour Holter ECG Average, Minimum, and Maximum Heart Rate Results (Safety Population)
Table 14.3.5.4: Summary of the 24-hour Holter ECG Hourly Average Heart Rate Results (Safety Population)
Table 14.3.5.5: 24-hour Holter ECG Abnormalities (Safety Population)
Table 14.3.6: Shift Table of Hematology Results (Safety Population)
Table 14.3.7: Shift Table of Biochemistry Results (Safety Population)

## CSR Section Listings 16:

Listing 16.1.7: Randomization Schedule (All Randomized Subjects)
Listing 16.2.1.1: Subject Disposition (All Subjects)
Listing 16.2.1.2: Visit Dates (All Randomized Subjects)
Listing 16.2.2.1: Protocol Deviations (All Randomized Subjects)
Listing 16.2.2.2: Inclusion/Exclusion Criteria Not Met (All Randomized Subjects)
Listing 16.2.3: Analysis Sets (All Randomized Subjects)
Listing 16.2.4.1: Subject Demography (All Randomized Subjects)
Listing 16.2.4.2: COPD History and Smoking Habits (All Randomized Subjects)
Listing 16.2.4.3.1: Medical and Surgical History (All Randomized Subjects)
Listing 16.2.4.3.2: Concomitant Diseases (All Randomized Subjects)
Listing 16.2.4.4: Previous Medications (All Randomized Subjects)

Listing 16.2.4.5: Concomitant Medications (All Randomized Subjects)
Listing 16.2.4.6: Maintained Medications (All Randomized Subjects)
Listing 16.2.4.8: Concomitant Procedures (All Randomized Subjects)
Listing 16.2.4.9: Reversibility Test (All Randomized Subjects)
Listing 16.2.4.10: Baseline Dyspnea Index Questionnaire (All Randomized Subjects)
Listing 16.2.4.11: COPD Assessment Test (All Randomized Subjects)
Listing 16.2.5.1: Background Medication Administration (All Randomized Subjects)
Listing 16.2.5.2: Study Drug Administration (All Randomized Subjects)
Listing 16.2.5.3: Rescue Medication Administration (All Randomized Subjects)
Listing 16.2.5.4: Background Medication Exposure and Compliance (All Randomized Subjects)
Listing 16.2.5.5: Study Drug Exposure and Compliance (All Randomized Subjects)
Listing 16.2.6.2: Peak FEV1 and FVC (All Randomized Subjects)
Listing 16.2.6.3: Pre-Dose Morning FEV1 and IC Results (All Randomized Subjects)
Listing 16.2.6.4: FEV1 and FVC AUC0-12h and AUC0-4h Results (All Randomized Subjects)
Listing 16.2.6.5: Onset of Action (Change from Baseline in FEV1 ≥ 100 mL) on Day 1 (All Randomized Subjects)
Listing 16.2.6.6:Transition Dyspnea Index Questionnaire (All Randomized Subjects)
Listing 16.2.6.7: Rescue Medication Administration — Derived Variables (All Randomized Subjects)
Listing 16.2.6.8: E-RS Questionnaire (All Randomized Subjects)
Listing 16.2.7.1: Pre-Treatment Adverse Events (All Randomized Subjects)
Listing 16.2.7.2: Treatment Emergent Adverse Events (All Randomized Subjects)
Listing 16.2.7.3: Serious Treatment Emergent Adverse Events (All Randomized Subjects)
Listing 16.2.7.4: Adverse Drug Reactions (All Randomized Subjects)
Listing 16.2.7.5: Serious Adverse Drug Reactions (All Randomized Subjects)
Listing 16.2.7.6: Severe Treatment Emergent Adverse Events (All Randomized Subjects)
Listing 16.2.7.7: Treatment Emergent Adverse Events Leading to Study Drug Discontinuation (All Randomized Subjects)
Listing 16.2.7.8: Treatment Emergent Adverse Events Leading to Death (All Randomized Subjects)
Listing 16.2.7.9: Post-study Adverse Events (All Randomized Subjects)
Listing 16.2.8.1: Vital Signs (All Randomized Subjects)
Listing 16.2.8.2.1: 24-hour Holter ECG Results (All Randomized Subjects)
Listing 16.2.8.2.2: 24-hour Holter ECG – Prolonged QTcF (All Randomized Subjects)
Listing 16.2.8.2.3: 24-hour Holter ECG Abnormalities (All Randomized Subjects)
Listing 16.2.8.3.1: Laboratory Results – Hematology (All Randomized Subjects)
Listing 16.2.8.3.2: Laboratory Results – Biochemistry (All Randomized Subjects)
Listing 16.2.8.3.3: Laboratory Results – Pregnancy Test (All Randomized Females)
Listing 16.2.8.4: Physical Examination (All Randomized Subjects)
Listing 16.2.9: Comments (All Randomized Subjects)

## CSR Section Figure 14:

Figure 14.1.1: Kaplan-Meier Curves for Time until Discontinuation (All randomized subject)
Figure 14.2.1.1.1: Mean absolute FEV1 values on Day 1 with 95% Confidence Interval (ITT Population)
Figure 14.2.1.1.2: Mean absolute FEV1 values on Week 6 with 95% Confidence Interval (ITT Population)
Figure 14.2.1.1.3: Mean absolute FEV1 values on Day 1 with 95% Confidence Interval (PP Population)
Figure 14.2.1.1.4: Mean absolute FEV1 values on Week 6 with 95% Confidence Interval (PP Population)

## SAS Outputs Listings:

Listing 16.1.9.16: SAS Output of Analysis of the E-RS Total during the Inter-Visit Periods and the Entire Treatment Period (ITT Population)

# 12 References

1. Hankinson JL, Eschenbacher B, Townsend M, Stocks J, Quanjer PH. Use of forced vital capacity and forced expiratory volume in 1 second quality criteria for determining a valid test. European Respiratory Journal 2015, 45(5):1283-92.

2. Leidy NK, Murray LT. Patient-reported outcome (PRO) measures for clinical trials of COPD: the EXACT and E-RS. COPD. Jun 2013;10(3):393-398.

3. P.W. Jones, G. Harding, P. Berry, I. Wiklund, W-H. Chen and N. Kline Leidy. Development and first validation of the COPD Assessment Test. Eur Respir J 2009, 34: 648–654.

# APPENDIX I: EXACT-PRO QUESTIONNAIRE

## Scoring Algorithms and Analytical Definitions

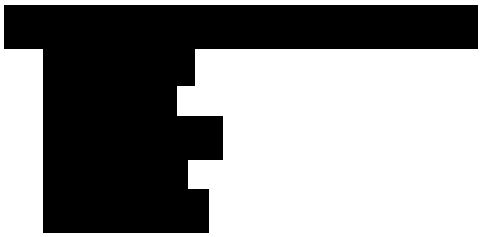
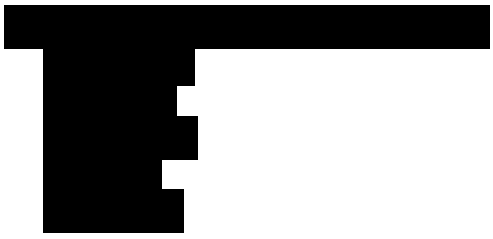
## COMPUTATIONAL INSTRUCTIONS

### Raw Scores

Raw scores are assigned to each item response option, as shown below.

### Annotated EXACT-PRO for Raw Score Assignment

The following annotates the raw score values associated with each response category for the EXACT-PRO items:

1. 
2. 
3. 
4. 
5. 

6. [redacted]

7. [redacted]

8. [redacted]

9. [redacted]

10. [redacted]

11. [redacted]