Statistical Analysis Plan	
Study Code	D5980C00019
Edition Number	2.0
Date	26-Aug-2019

A Randomized, Double-blind, Two Treatment, Two Period, Chronic dosing (4 weeks), Cross-over, Multi-center Pilot study to evaluate the effects of Budesonide/Glycopyrronium/Formoterol Fumarate and Glycopyrronium/Formoterol Fumarate on Specific Image based Airway Volumes and Resistance in Subjects with Moderate to Severe Chronic Obstructive Pulmonary Disease

A Randomized, Double-blind, Two Treatment, Two Period, Chronic dosing (4 weeks), Cross-over, Multi-center Pilot study to evaluate the effects of Budesonide/Glycopyrronium/Formoterol Fumarate and Glycopyrronium/Formoterol Fumarate on Specific Image based Airway Volumes and Resistance in Subjects with Moderate to Severe Chronic Obstructive Pulmonary Disease

Study Statistician

[redacted]

27 August 2019 | 07:42 PDT

Date

A Randomized, Double-blind, Two Treatment, Two Period, Chronic dosing (4 weeks), Cross-over, Multi-center Pilot study to evaluate the effects of Budesonide/Glycopyrronium/Formoterol Fumarate and Glycopyrronium/Formoterol Fumarate on Specific Image based Airway Volumes and Resistance in Subjects with Moderate to Severe Chronic Obstructive Pulmonary Disease

Global Product Statistician

[redacted]

27 August 2019 | 13:23 PDT

Date

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

Abbreviation or special term	Explanation
AE	Adverse Event
AT	Air Trapping
ATC	Anatomic Therapeutic Class
ATS	American Thoracic Society
BDRM	Blinded Data Review Meeting
BID	Bis In Die, twice daily
CFD	Computational Fluid Dynamics
COPD	Chronic Obstructive Pulmonary Disease
CSR	Clinical Study Report
CT	Computed Tomography
eCRF	electronic Case Report Form
ERS	European Respiratory Society
FEF25-75	Forced Expiratory Flow between 25% to 75% of FVC
FEV_1	Forced Expiratory Volume in 1 Second
FRC	Functional Residual Capacity
FRI	Functional Respiratory Imaging
FVC	Forced Vital Capacity
GFF	Glycopyrronium and Formoterol Fumarate
HFA	Hydrofluoroalkane
HRCT	High Resolution Computed Tomography
IAD	Internal Airflow Distribution
IC	Inspiratory Capacity
IMP	Investigational Medicinal Product
iRaw	Image based Airway Resistance
ITT	Intent-To-Treat
iVaw	Image based Airway Volume
iVaww	Image based Airway Wall thickness
iVbv	Image based Blood vessel density or fibrosis score
iVlobe	Image based Lobar Volume
LAS	Low Attenuation or emphysema Score
LL	Lower Lobe

LLL	Left Lower Lobe
LUL	Left Upper Lobe
MDI	Metered Dose Inhaler
mITT	Modified Intent-To-Treat
MPD	Major Protocol Deviation
PFT	Pulmonary Function Test
Raw	Airway Resistance
RLL	Right Lower Lobe
RML	Right Middle Lobe
RUL	Right Upper Lobe
RV	Residual Volume
SAE	Serious Adverse Event
sGaw	Specific Airway Conductance
siRaw	Specific Image based Airway Resistance
siVaw	Specific Image based Airway Volume
sRaw	Specific Airway Resistance
TEAE	Treatment Emergent Adverse Event
TLC	Total Lung Capacity
TP	Treatment Period
UA	Upper Airway
UL	Upper Lobe
WHO-DD	World Health Organization Drug Dictionary

AMENDMENT HISTORY

Date	Brief description of change
26 August 2019	The ITT analysis set will be analysed according to treatment actually received rather than randomized treatment as assigned (Section 2.1.2), unless the same treatment is received in both periods. This ensures inferences are made relating to the correct treatments, given that one subject is known to have had the treatments for each period interchanged. This maximises the number of useable subjects that can be analysed using data from both treatments as opposed to the use of the mITT analysis set where such patients are removed. Sensitivity analyses using the mITT set will be conducted for the secondary endpoints as well as the primary endpoints in order to assess the robustness of findings in an analysis using treatment as randomised in patients without major protocol deviations.

1 STUDY DETAILS

1.1 Study objectives

Study objectives and endpoints are presented in Table 1. Detailed definitions of the endpoints are given in section 3.

Table 1 Study Objectives and Endpoints

Primary objective:	**Endpoint/variable:**
To assess the effects of BGF and GFF on specific image-based airway volumes and resistance in subjects with moderate to severe COPD following chronic twice-daily (BID) dosing after approximately four weeks of treatment	• Specific airway volume (siVaw) • Specific airway resistance (siRaw)
Secondary objective:	**Endpoint/variable:**
To assess the effects of BGF and GFF on various Functional Respiratory Imaging (FRI) parameters	• Airway volume (iVaw) • Airway resistance (iRaw)
To assess the effects of BGF and GFF on lung function parameters	Forced expiratory volume in one second (FEV_1)
To assess the effects of BGF and GFF on body plethysmography parameters	Functional Residual Capacity (FRC)
Safety objective:	**Endpoint/variable:**
To assess the safety of BGF and GFF	Adverse events (AEs), serious adverse events (SAEs), adverse events leading to treatment discontinuation (DAEs)
Exploratory	**Endpoint/variable:**
To assess the effects of BGF and GFF on other FRI parameters	• Lobe volumes (iVlobe) • Air Trapping (AT) • Internal lobar airflow distribution (IAD) • Low attenuation or emphysema score (LAS) • Blood Vessel density or fibrosis score (iVby) • Airway Wall Thickness (iVaww) • Mass of deposited particles per defined airway section

To assess the effects of BGF and GFF on other lung function parameters	• Forced vital capacity (FVC) • Tiffeneau index (FEV_1/FVC ratio) • Forced expiratory flow 25%-75% (FEF_{25-75}) • Inspiratory capacity (IC)
To assess the effects of BGF and GFF on other body plethysmography parameters	• Residual volume (RV) • Total Lung Capacity (TLC) • Airway resistance (Raw) • Specific airway resistance (sRaw) • Specific airway conductance (sGaw)

1.2 Study design

This is a phase IIIb randomised, controlled, two period cross-over, 4 weeks chronic dosing, study to evaluate the effects of BGF, 320/14.4/9.6 µg (160/7.2/4.8 µg per actuation), MDI 2 oral inhalations BID, morning and evening and GFF, 14.4/9.6 µg (7.2/4.8 µg per actuation), MDI 2 oral inhalations BID, morning and evening on specific image based airway volumes and resistance in subjects with moderate to severe COPD. In this study, airway dimension parameters will be calculated for each of the active compounds.

Study Periods:

There will be two treatment periods (TP) in this study. The first day on which a study treatment is received is Treatment Day 1 (Visit 3 for TP 1, Visit 5 for TP 2). Treatment day is numbered sequentially thereafter until the next scheduled visit (Visit 4 for TP 1, Visit 6 for TP 2), or until the subject withdraws from the study or is lost to follow-up. Each TP is intended to last until Day 29. Washout period between Visit 4 and Visits 5 is a minimum of 21 days to a maximum of 28 days. If any visits are rescheduled, the latest visit will be used.

It is expected that IMP dosing will not go beyond Visit 4 for TP 1, and beyond Visit 6 for TP 2. However, exposure to IMP will be calculated based on the day of actual last dose of each treatment.

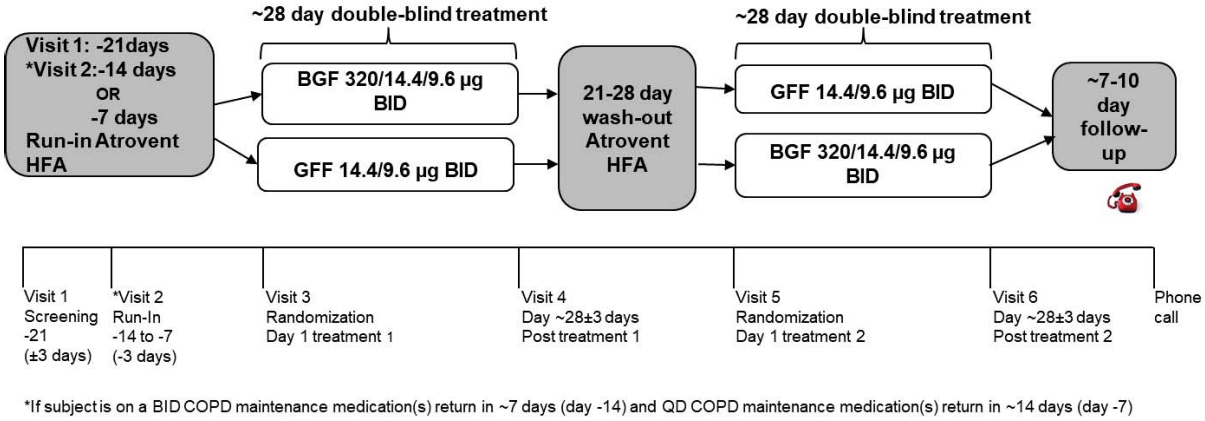
Treatments and treatment duration:
The 2 cross-over treatments will be:

BGF: Budesonide/Glycopyrronium/Formoterol Fumarate, 320/14.4/9.6 µg (160/7.2/4.8 µg per actuation), MDI (current phase III device) 2 oral inhalations in the morning and 2 oral inhalations in the evening.

GFF: Glycopyrronium/Formoterol Fumarate, 14.4/9.6 µg (7.2/4.8 µg per actuation), MDI 2 oral inhalations in the morning and 2 oral inhalations in the evening.

During the washout period, treatment with Atrovent HFA will be administered.

Study design chart



*If subject is on a BID COPD maintenance medication(s) return in ~7 days (day -14) and QD COPD maintenance medication(s) return in ~14 days (day -7)

Computed Tomography (CT)-scans during study:

On Day 1 of TP1 (Visit 3) pre-dose baseline measurement inspiratory scans (total lung capacity [$TLC_{(PRE)}$] scan) and expiratory scan (functional residual capacity [$FRC_{(PRE)}$] scan) will be conducted. During Visit 3 an additional scan of the upper airway (UA) will be taken. Post dose measurement inspiratory scan (total lung capacity [$TLC_{(POST)}$] scan) and expiratory scan (functional residual capacity [$FRC_{(POST)}$] scan) will be taken after approximately 4 weeks of each treatment with BGF MDI or GFF MDI on Day 29 ±3 days (Visit 4 and Visit 6). Post dose activities should be started 1 hour after dosing on Visit 4 and Visit 6 and should be concluded within 2.5 hours after dosing. Between the Treatment Periods there will be a washout period of approximately 21-28 days with treatment of Atrovent HFA.

1.3 Number of subjects

A total of approximately 20 patients with moderate to severe COPD will be randomized in a 1:1 scheme of BGF:GFF treatment sequences. No formal sample size calculation has been done. The proposed sample size for this study is based on previous experience, which included two previously conducted FRI studies with FLUIDDA (PT003018 and PT003019). Each of these studies included approximately 20 subjects.

PT003018 explored the effect of GFF on FRI parameters after 15 days of treatment. The observed mean change from baseline to Day 15 (on the log scale) and the effect size (expressed as ratios) for across-lobe average of siVaw and siRaw are presented in Table 2. It is expected that

similar effects will be seen in this study for the GFF treatment, while improved effects will be seen for the BGF treatment.

To estimate the treatment effect observed in PT003018, the definitions of co-primary endpoints of this study were applied to PT003018 data. For baseline, the scan at Day 1 of Period 1 was used; for post-baseline, the scan at Day 15 after GFF treatment was used. The averages across lobes were calculated for each subject, and the logarithm transformation was applied. The resulting mean differences and standard deviations of transformed data are shown in Table 2.

To estimate power, first, the power for the individual t-tests was calculated at half the nominal alpha (i.e. at alpha = 0.025), see Table 2. Under Hochberg procedure with familywise alpha=0.05, we are guaranteed to have statistical significance if $p<0.025$ for at least one of the variables, irrespective of what the p-value is for the other variable. Thus, a significance level of 0.025 was used in these calculations.

To attain 80% power after Hochberg procedure, we would need the true mean change from baseline in log-transformed siVaw to be 0.193 (siVaw ratio to baseline of 1.213), and -0.476 for log-transformed siRaw (ratio of 0.621), assuming the standard deviations from Table 2. These ratios to baseline are similar to those achieved by formoterol monotherapy in study PT003019 (1.23 on siVaw and 0.56 on siRaw), hence it is assumed that combination treatments each containing formoterol will achieve at least this effect.

Table 2 Power Estimates

Parameter	Mean Change (log scale) *	SD Change (log scale) *	Ratio **	Alpha (two-sided)	Power
siVaw	0.434	0.262	1.543	0.025	>99.9%
siRaw	-0.975	0.644	0.377	0.025	>99.9%

* Estimates are from PT003018, change from Day 1 period 1 scan to post-GFF scan.
** Ratio of geometric means on the original scale (same as exponent of mean change in log scale).

Given that a stronger BGF treatment effect is expected, the calculations presented above serve as conservative estimates for the BGF group as well.

2 ANALYSIS SETS

2.1 Definition of analysis sets

The following analysis populations are defined in this study:

2.1.1 Enrolled analysis set

Enrolled analysis set is defined as all subjects who sign the informed consent form (ICF).

2.1.2 Intent-to-Treat (ITT) analysis set

The Intent-to-Treat (ITT) analysis set is defined as all subjects who were randomized to study treatment. Patients will be analysed according to the treatment actually received. All available data will be utilized in the ITT analyses. However, for subjects who received the same treatment during both treatment periods, data from TP 1 only will be used for ITT analyses.

2.1.3 Modified ITT (mITT) analysis set

The modified ITT (mITT) analysis set is defined as all subjects in the ITT analysis set who completed both treatment periods and have FRI data at baseline and after approximately four weeks of treatment. Major protocol deviations (MPD) will be identified during the blinded data review meeting (BDRM) prior to database lock and unblinding. It will be determined in this meeting which data need to be excluded from the mITT analysis set due to the MPDs. See section 2.2 for examples of MPDs.

Subjects will be analysed according to the treatment they were assigned to at randomization.

2.1.4 Safety analysis set

The Safety analysis set is defined as all subjects who are randomized to study treatment and receive at least one dose of Investigational Medicinal Product (IMP) and for whom any post-dose data are available. Statistical analyses and tabulations will be by the treatment actually received.

2.1.5 Use of analysis sets

Analyses will be performed as follows:

Demographics and subject characteristics will be summarized for the ITT analysis set. Subjects who received same treatment in both treatment periods or who did not have any assessments in TP 2 will only be summarized for the treatment group of TP 1.

Extent of exposure and safety data will be summarized for the safety analysis set by actual treatment group. Subjects who did not receive any dose of IMP in TP 2 will only be summarized for the actual treatment group of TP 1.

The ITT analysis set will be considered the primary analysis population for efficacy. All primary and secondary efficacy endpoints will be analysed using the ITT analysis set. The mITT analysis set is a subset of the ITT analysis set. If the data contributing to the mITT analyses differs from the data contributing to the ITT analyses, then the mITT analysis set will be used to conduct sensitivity analyses for the primary and the secondary endpoints , see section 4.2.4.3.

2.2 Violations and deviations

Major Protocol Deviations (MPDs) will be listed and tabulated. These are defined as potential protocol deviations that may significantly affect the reliability of efficacy study data. These will be identified during the blinded data review meeting (BDRM) before database lock.

The following are a few examples that could be considered as major protocol deviations:

- Patient randomized despite not meeting key inclusion criteria, including:
  - Patients not meeting COPD diagnosis (as defined by ATS/ERS) and severity criteria as per inclusion criteria 3 and 5.
  - Patients meeting exclusion criteria for other respiratory disease as per exclusion criteria 2, including diagnosis of asthma.
- Patient received incorrect study drug
- Prohibited concomitant COPD medication taken during the study, defined as the classes of medications listed in table 6 of the protocol, received after Visit 3, other than those provided by the sponsor

Additional major protocol deviations may be identified during BDRM from a review of comments entered in electronic case report forms (eCRF) and of clinical protocol deviation report.

All data supporting identification of major deviations will be blinded.

Randomization errors

Randomization errors, such as a subject is given a treatment pack for a different patient or is randomized out of chronological order, will not necessarily be identified as MPDs. Only if the error caused the subject to receive a treatment other than randomized, the affected period will be excluded from mITT.

3 PRIMARY AND SECONDARY VARIABLES

3.1 Visit windows

No study time windows will be derived for the reporting of data. Data will be reported according to the protocol-scheduled day.

3.2 Efficacy Endpoints

3.2.1 Functional respiratory imaging (FRI)

Subjects will receive a total of seven high resolution computed tomography (HRCT) scans (Table 3). The first will be conducted at baseline, the beginning of TP 1 at Visit 3. The pre-dose CT-scan of the thorax will be taken on two breathing levels, $TLC_{(PRE)}$ and $FRC_{(PRE)}$. During Visit

3 an additional scan of the upper airway (UA) will be taken. The post-dose measurement will be performed on two breathing level, $TLC_{(POST)}$ and $FRC_{(POST)}$. The Post dose scans will be taken after approximately 4 weeks of each treatment (treatment 1 and treatment 2) of BGF MDI and GFF MDI on Day 29 (±3 days) (Visit 4 and Visit 6).

Table 3 Schedule of CT Scans

	Scan		
Visit	**FRC**	**TLC**	**UA**
Visit 3	X	X	X
Visit 4	X	X	
Visit 5			
Visit 6	X	X	

FRI parameters will be obtained from either TLC, FRC, or both scan types. The TLC scan will be used in the definition of primary and secondary endpoints. For the other variables, the TLC scan results are of primary interest, except for iVlobe, AT, and iVlobePP, which will use FRC scan data. For all variables, data obtained from the other scan type, where available, may be used for exploratory analyses.

Table 4 presents all collected or derived FRI parameters. Primary, secondary and exploratory parameters are identified in this table by footnotes 1, 2 and 3, respectively. Where trimming is applicable, they will be analysed using the untrimmed version (see section 3.2.1.1 for trimming discussion). All data will be listed, including both breathing levels, and both trimmed and untrimmed variable versions, where applicable.

The UA scan supports estimation of the mass of deposited particles, and otherwise does not provide variables for analysis. FRI parameters obtained from the FRC and TLC scans will be provided by the FRI data vendor (FLUIDDA), and are summarized in the table below.

Table 4 FRI Parameters

	Region											Scan	
Parameter	LLL	LUL	RML	RUL	RLL	Across-lobe Average	LL	UL	CENTRAL	DISTAL	TOTAL	FRC	TLC
iVaw†	Y	Y	Y	Y	Y	C[2]	Y	Y	Y	Y	Y	Y[3]	Y[2]
siVaw†	Y	Y	Y	Y	Y	C[1]	Y	Y	Y	Y	Y	Y[3]	Y[1]
iRaw†	Y	Y	Y	Y	Y	C[2]	Y	Y	Y	Y	Y	Y[3]	Y[2]
siRaw†	Y	Y	Y	Y	Y	C[1]	Y	Y	Y	Y	Y	Y[3]	Y[1]
iVlobe†	Y	Y	Y	Y	Y	C[3]	Y	Y	N/A	N/A	Y	Y[3]	Y[3]
LAS	Y	Y	Y	Y	Y	C[3]	Y	Y	N/A	N/A	Y	N/A	Y[3]
IAD	Y	Y	Y	Y	Y	C[3]	Y	Y	N/A	N/A	N/A	Y[3] **	
AT	Y	Y	Y	Y	Y	C[3]	Y	Y	N/A	N/A	Y	Y[3]	N/A
iVbv	Y	Y	Y	Y	Y	C[3]	Y	Y	N/A	N/A	Y	N/A	Y[3]
iVaww†	Y	Y	Y	Y	Y	C[3]	Y	Y	Y	Y	Y	N/A	Y[3]
siVaww†	Y	Y	Y	Y	Y	C[3]	Y	Y	Y	Y	Y	N/A	Y[3]

iVQ	Y	Y	Y	Y	Y	C[3]	Y	Y	N/A	N/A	Y	Y[3] **	
iVawPP	Y	Y	Y	Y	Y	C	Y	Y	Y	Y	Y	Y	Y
siVawPP	Y	Y	Y	Y	Y	C	Y	Y	Y	Y	Y	Y	Y
iRawPP	Y	Y	Y	Y	Y	C	Y	Y	Y	Y	Y	N/A	Y
siRawPP	Y	Y	Y	Y	Y	C	Y	Y	Y	Y	Y	N/A	Y
iVlobePP	Y	Y	Y	Y	Y	C	Y	Y	N/A	N/A	Y	Y	Y
IADPP	Y	Y	Y	Y	Y	C	Y	Y	N/A	N/A	N/A	Y**	
iVbvPP	Y	Y	Y	Y	Y	C	Y	Y	N/A	N/A	Y	N/A	Y
Deposition†*	Y	Y	Y	Y	Y	C[3]	Y	Y	Y	Y	Y	N/A	Y[3]

Lobe regions: L = left, R = right; L = lower, M = middle, U = upper.
Y = will be provided by FLUIDDA; C = calculated; N/A = not applicable.
[1] = defines a primary endpoint; [2] = defines a secondary endpoint; [3] = defines an exploratory endpoint.
† = data will be log-transformed.
* = sum of distal and peripheral results for lobes.
** = derived from both TLC and FRC scans.

For FRI parameters, data is generated within each of five lung lobes: right upper lobe (RUL), right middle lobe (RML), right lower lobe (RLL), left upper lobe (LUL) and left lower lobe (LLL). Each lobe comprises several segments (branches) that are further subdivided into typically 5-10 visible airway generations. Generation-level data (within segment within lobe) will not be provided by FLUIDDA.

For all FRI parameters, across-lobe averages will be derived as the simple average of available lobe-level data. The only exceptions are IAD and IADPP, where across-lobe averages will be derived as the simple average of available lobe-level data of the lower lobes.

LLL and RLL form the lower lobes region (LL), while LUL, RML and RUL form the upper lobes region (UL). The DISTAL region is formed from LL and UL (or all 5 lobes combined), CENTRAL region includes the airway segments before the 3rd bifurcation, and TOTAL region is CENTRAL and DISTAL combined. Further explanation of the parameters and derivation details are provided in subsequent sections.

3.2.1.1 Trimming

The visible airway generations that comprise the lobes can vary over time. The measurements of iVaw, iRaw, siVaw, siRaw, iVaww and siVaww can thus be performed in two ways: 1) for each segment use all the generations visible at the particular study visit scan (“untrimmed”), or 2) for each segment use only generations of airways that are visible in the baseline scan and in the Day 29 scan in each period (Visits 3-4, and Visits 3-6) (“trimmed”). The untrimmed version of iVaw, iRaw, siVaw, siRaw, iVaww and siVaww is of better interest, as it should allow to see better the effect of the ICS compound. These variables will thus be analysed using the untrimmed version. The trimmed values will be listed.

3.2.1.2 Airway volume (iVaw)

In the CT scans, the airways can be segmented up to the point where no distinction can be made between the intraluminal and alveolar air. This is where the airway diameter is around 1 – 2 mm, typically around the 5th to 10th bifurcation, depending mainly on the disease state of the individual patient. From the resulting model, the central and distal airway volumes can be assessed at individual airways or in different regions.

The distal airway volume is defined as the segmented airway volume starting from the third bifurcation that is from the 5 lung lobes without the central region. It is this distal region that will be used in the analyses in this study.

3.2.1.3 Lobe volume (iVlobe)

By identifying and grouping the voxels that represent the air in the lungs the lung volume can be determined from the scans. During segmentation, identifying the fissure planes on the CT images and using these surfaces as cutting objects can separate lung lobes. This means that not only the total lung volume is determined, but also the volume of each lobe individually.

Measurements of iVlobe will thus be available for individual lobes. iVlobe for the other regions will be obtained by adding the iVlobe for the lobes that comprise the regions:

- iVlobe LL = iVlobe RLL + iVlobe LLL
- iVlobe UL = iVlobe RUL + iVlobe RML + iVlobe LUL
- iVlobe TOTAL = iVlobe RLL + iVlobe LLL + iVlobe RUL + iVlobe RML + iVlobe LUL

3.2.1.4 Airway resistance (iRaw)

The airway resistance (iRaw) is determined using Computational Fluid Dynamics (CFD). During the CFD calculations, the outflow to each lobe is adjusted iteratively for each subject to match the internal flow rate distributions obtained from the segmentation of the CT scans. Hence, iRaw accounts for the patient-specific internal airflow distribution which might be greatly altered by the lung disease. Hence, the airflow distribution in the CFD calculation reflects the airflow distribution as derived from the expansion of the lung lobes from FRC to TLC. The iRaw is defined as the total pressure drop over an airway, divided by the flow rate through that airway.

iRaw for the lobes and the CENTRAL region will be provided by the FRI data vendor. iRaw for the other regions will be obtained by applying the following formulas to the available iRaw values:

- iRaw LL = 1 / (1/iRaw RLL + 1/iRaw LLL)
- iRaw UL = 1 / (1/iRaw RUL + 1/iRaw RML + 1/iRaw LUL)
- iRaw DISTAL = 1 / (1/iRaw RLL + 1/iRaw LLL + 1/iRaw RUL + 1/iRaw RML + 1/iRaw LUL)
- iRaw TOTAL = 1 / (1/iRaw CENTRAL + 1/iRaw RLL + 1/iRaw LLL + … + 1/iRaw LUL)

3.2.1.5 Specific image-based airway volume (siVaw)

The values of siVaw are derived from iVaw by dividing iVaw by the image based lobe volume (iVlobe). On the lobar scale, these calculations are:

- siVaw LLL = iVaw LLL / iVlobe LLL
- siVaw LUL = iVaw LUL / iVlobe LUL
- siVaw RML = iVaw RML / iVlobe RML
- siVaw RUL = iVaw RUL / iVlobe RUL
- siVaw RLL = iVaw RLL / iVlobe RLL

siVaw is also derived for the larger lung regions (LL, UL, DISTAL, CENTRAL, and TOTAL) by using the formulae below:

- siVaw LL = iVaw LL / iVlobe LL
- siVaw UL = iVaw UL / iVlobe UL
- siVaw CENTRAL= iVaw CENTRAL / iVlobe TOTAL
- siVaw DISTAL= iVaw DISTAL / iVlobe TOTAL
- siVaw TOTAL = iVaw TOTAL / iVlobe TOTAL

3.2.1.6 Specific image-based airway resistance (siRaw)

The values of siRaw are derived from iRaw by multiplying iRaw by the image based lobe volume (iVlobe). On the lobar scale, these calculations are:

- siVaw LLL = iVaw LLL * iVlobe LLL
- siVaw LUL = iVaw LUL * iVlobe LUL
- siVaw RML = iVaw RML * iVlobe RML
- siVaw RUL = iVaw RUL * iVlobe RUL
- siVaw RLL = iVaw RLL * iVlobe RLL

siRaw is also derived for the larger lung regions (LL, UL, DISTAL, CENTRAL, and TOTAL) by using the formulae below:

- siRaw LL = iRaw LL * iVlobe LL
- siRaw UL = iRaw UL * iVlobe UL
- siRaw CENTRAL = iRaw CENTRAL * iVlobe TOTAL
- siRaw DISTAL = iRaw DISTAL * iVlobe TOTAL
- siRaw TOTAL = iRaw TOTAL * iVlobe TOTAL

3.2.1.7 Percent predicted image-based lobe volume (iVlobePP)

Predicted volumes will be calculated based upon age, sex and height. Percent predicted iVlobe will be provided along with the absolute values of iVlobe.

3.2.1.8 Air trapping (AT)

Air trapping is defined as all the intrapulmonary voxels with Hounsfield Units between -1024 and -850 using the expiratory scans at FRC. This parameter is given as percentage of iVlobe.

3.2.1.9 Internal lobar airflow distribution (IAD)

By segmenting the lobes at FRC and TLC scans for each patient, the patient-specific airflow distribution can be established by assessing lobar volume expansion.

Internal Airflow Distribution (IAD) will be calculated according to the following formulas:

- IAD LLL = 100 * (iVlobe LLL TLC - iVlobe LLL FRC) / (iVlobe TOTAL TLC - iVlobe TOTAL FRC)
- IAD LUL = 100 * (iVlobe LUL TLC - iVlobe LUL FRC) / (iVlobe TOTAL TLC - iVlobe TOTAL FRC)
- IAD RML = 100 * (iVlobe RML TLC - iVlobe RML FRC) / (iVlobe TOTAL TLC - iVlobe TOTAL FRC)
- IAD RUL = 100 * (iVlobe RUL TLC - iVlobe RUL FRC) / (iVlobe TOTAL TLC - iVlobe TOTAL FRC)
- IAD RLL = 100 * (iVlobe RLL TLC - iVlobe RLL FRC) / (iVlobe TOTAL TLC - iVlobe TOTAL FRC)
- IAD LL = IAD RLL + IAD LLL
- IAD UL = IAD RUL + IAD RML + IAD LUL

3.2.1.10 Low attenuation or emphysema score (LAS)

Low attenuation score is defined as all the intrapulmonary voxels with Hounsfield Units between -1024 and -950 using the inspiratory scans at TLC. This parameter is given as percentage of iVlobe.

3.2.1.11 Blood vessel density or fibrosis score (iVbv)

Blood vessel density can be determined through segmentation and three-dimensional reconstruction of the blood vessels. The segmentation is based on a Hounsfield unit (HU) threshold between -600 and 600 and is performed on the TLC scan. The blood vessel density can be considered a surrogate for perfusion. This parameter is given as percentage of iVlobe.

3.2.1.12 Airway wall thickness (iVaww)

The airway wall volume (also called airway wall thickness) consists of all visible tissue in the CT scan that encompasses the airway wall.

Airway wall thickness (iVaww) will be calculated according to the following formulas:

- iVaww LL = iVaww RLL + iVaww LLL
- iVaww UL = iVaww RUL + iVaww RML + iVaww LUL
- iVaww DISTAL = iVaww RLL + iVaww LLL + iVaww RUL + iVaww RML + iVaww LUL
- iVaww TOTAL = iVaww CENTRAL + iVaww DISTAL

3.2.1.13 Mass of deposited particles per defined airway section

Regional aerosol deposition is determined by simulating the flow in the patient specific geometries using patient specific boundary conditions by means of CFD. While solving the flow equations, simultaneously particles are released in the flow and the force mass balance of the individual particles is determined through additional discrete phase computations. When a calculated particle trajectory intersects with the airway wall, the particle is trapped in that location. This allows determining the regional concentration of inhaled aerosols and consequently the effective lung dose of inhaled medication. These are measured in the full airway tree.

Mass of deposited particles is a useful way to see the particle deposition in each region of the lung. Distal and peripheral data will be summed. Upper lobe and lower lobe totals will also be derived.

3.2.2 Spirometry

Forced expiratory spirometry maneuvers will be assessed to derive the following parameters:

- Forced expiratory volume in 1 second (FEV_1), the volume of air exhaled under forced conditions in the first second.
- Forced vital capacity (FVC), the determination of the vital capacity from a maximally forced expiratory effort.
- Forced expiratory flow between 25% to 75% of FVC (FEF_{25-75}).
- Tiffeneau index, the FEV_1/FVC ratio.
- Inspiratory capacity (IC), the amount of air that can be inhaled after the end of a normal expiration.

At Visit 3 and Visit 5, spirometry will be obtained at the beginning of the visit. A second spirometry will be obtained before the study drug administration after the pre-dose CT-scans are taken (at Visit 3 only). The -60 minute FEV1 value obtained at Visit 5 prior to study drug administration will be used to check the stability criteria. At Visit 4 and Visit 6, spirometry will be

obtained before the study drug administration, and post-dosing of study drug. Post-dose spirometry assessments will not be conducted before the CT-scan is taken.

Morning pre-dose trough FEV_1 will be derived at Visits 4 and 6 as the average of the 60 and 30-minute pre-dose FEV_1 values. In patients missing either of the pre-dose assessments, the single value from that visit will be used as the morning pre-dose trough FEV_1. Similar definitions of morning pre-dose trough will be applied for FVC, FEF_{25-75}, and IC.

3.2.3 Body plethysmography

The following parameters will be measured using body plethysmography:

- Functional residual capacity (FRC), the volume remaining in the lungs at the end-expiratory position.
- Residual volume (RV), the volume of air remaining in the lungs after a maximal exhalation.
- Total lung capacity (TLC), the volume in the lungs at maximal inflation.
- Airway resistance (Raw), the resistance of the respiratory tract to airflow during inspiration and expiration.
- Specific airway resistance (sRaw), airway resistance corrected for differences in lung volume.
- Specific airway conductance (sGaw), derived as the mathematical inverse of airway resistance corrected for differences in lung volume.

On Day 1 of each treatment period (Visit 3 and Visit 5) the body plethysmography measurement will be performed after the CT-scan (at Visit 3 only) however, before the IP administration.

On Day 29 (±3 days) of each treatment period (Visit 4 and Visit 6) the body plethysmography measurement will be performed after the CT-scan and after the post-dose spirometry measurement.

3.2.4 Rescue Ventolin HFA use

The number of puffs of rescue Ventolin HFA will be collected in the subject's diary and then entered into eCRF. Entries will be made on a daily basis for the daytime and nighttime number of puffs.

Mean daily, daytime and nighttime number of puffs of Ventolin HFA will be calculated for each TP. The denominator for these means will be adjusted based on the number of days (including half days) with non-missing values. That is, the mean daily number of puffs of daytime rescue use will be set to the total number of daytime puffs divided by the number of half-days when daytime rescue use was recorded. The mean daily number of puffs of nighttime rescue use will be set to the total number of nighttime puffs divided by the number of half-days when the

nighttime rescue use was recorded. The mean daily rescue use is then calculated as the average of the daytime and nighttime means.

3.2.5 COPD exacerbations

COPD exacerbations are not collected in this study. Serious COPD exacerbations will be captured via the SAE eCRF.

3.2.6 Definition of baselines

Baseline for FRI endpoints

For FRI parameters, baseline will be the pre-dose value taken at Day 1 of Treatment Period 1 (Visit 3).

FRI parameters allow for estimates by lobe. Baseline for individual lobes is derived for each subject as the pre-dose lobe-specific baseline at Visit 3, and these will be used in the individual lobe descriptive statistics. Baseline "across-lobes" is derived for each subject as the mean of all available baselines across all lobes at Visit 3, and will be used in descriptive statistics, and statistical analyses that include a baseline.

Baseline for spirometry and plethysmography endpoints

The focus of this study is on estimation of the individual effects of treatment with BGF and GFF. For this purpose, period-dependent baselines will be used. For spirometry, period-dependent baselines will be defined as the mean of the pre-dose -60 and -30-minute values, at each of Visit 3 and Visit 5. In patients missing either of the pre-dose assessments, the single value from that visit will be used as the period-dependent baseline. For plethysmography, the pre-dose assessments at Visits 3 and 5 will be used as the period-dependent baselines. If Visit 5 was rescheduled due to FEV_1 stability criteria, only the assessments with the latest date will be used in the calculation of the baselines.

Additional, exploratory analyses will compare the distribution of spirometry and plethysmography endpoints between the two treatments, using a linear mixed model approach. For these analyses, subject-level baselines will be used. The subject-level baseline for a given endpoint will be defined as the average of the corresponding period-dependent baselines. If there is no period-dependent baseline for period 2, the subject-level baseline will be defined as the period-dependent baseline for period 1.

Baseline for rescue Ventolin HFA use

Mean daily, daytime and nighttime number of puffs will be calculated over the 7 days prior to Visit 3, using the same rules as described in section 3.2.4.

3.2.7 Primary efficacy endpoints

There will be two co-primary endpoints, defined and analysed similarly for the BGF and GFF treatment groups. The co-primary endpoints for each of the treatment groups (BGF and GFF) of this study are as follows:

- Change from baseline to Day 29 in across-lobe average siVaw (TLC scans, untrimmed values).
- Change from baseline to Day 29 in across-lobe average siRaw (TLC scans, untrimmed values).

There will be a total of four hypotheses associated with these endpoints (two variables evaluated for two treatments), described in section 4.1.2. The lobe-specific values of siVaw and siRaw will be used to perform exploratory analyses, presented in section 4.2.4.2.

3.2.8 Secondary efficacy endpoints

Secondary efficacy endpoints include two FRI endpoints, one spirometry endpoint, and one plethysmography endpoint, as follows:

- Change from baseline to Day 29 in across-lobe average iVaw (TLC scans, untrimmed values).
- Change from baseline to Day 29 in across-lobe average iRaw (TLC scans, untrimmed values).
- Change from period-dependent baseline to Day 29 in post-dose FEV_1
- Change from period-dependent baseline to Day 29 in post-dose plethysmography FRC

For the lung function objective, the post-dose FEV_1 value was chosen as a secondary endpoint. Post-dose spirometry aligns with the timing of scans, and thus allows comparability with the FRI results.

Each of the variables will be analysed for BGF and GFF. There will be no type I error control for these endpoints. Analysis will be similar to the primary endpoints analysis, but will be focused on estimation. Refer to sections 4.2.4 and 4.2.5 for details.

3.2.9 Exploratory efficacy endpoints

3.2.9.1 FRI parameters

siVaw, siRaw, iVaw and iRaw parameters evaluated at the individual lobes (TLC scans, untrimmed) will provide an alternative way to calculate the across-lobe average values, as explained in section 4.2.4.2. Analysis at the lobe level will also allow to investigate regional effects of the treatments.

The following FRI parameters listed in Table 4 will be considered exploratory endpoints:

- iVaw untrimmed at FRC,
- siVaw untrimmed at FRC,
- iRaw untrimmed at FRC,
- siRaw untrimmed at FRC,
- iVlobe at both TLC and FRC,
- LAS at TLC, IAD,
- AT at FRC,
- iVbv at TLC,
- iVaww untrimmed at TLC,
- siVaww untrimmed at TLC,
- iVQ,
- deposition untrimmed at TLC.

There is no planned analysis for the remaining FRI data. All available FRI data will be listed, including the TLC / FRC scan types, the trimmed / untrimmed parameter versions, and all regions.

3.2.9.2 Spirometry parameters

Change from period-dependent baseline to Day 29, and change from subject-level baseline to Day 29, in post-dose FVC, FEF_{25-75} and IC will be exploratory endpoints. Change from subject-level baseline to Day 29 in post-dose FEV_1 will also be an exploratory spirometry endpoint.

Change from period-dependent baseline to Day 29, and change from subject-level baseline to Day 29, in morning pre-dose FEV_1, FVC, FEF_{25-75} and IC, are exploratory endpoints.

Analysis methods will be similar to the primary and secondary endpoints, including the t-tests and mixed models, but will be considered exploratory and will focus on estimation. All spirometry data will be listed.

3.2.9.3 Body plethysmography parameters

Change from subject-level baseline to Day 29 in post-dose FRC will be an exploratory plethysmography endpoint.

Change from period-dependent baseline to Day 29, and change from subject-level baseline to Day 29, in post-dose RV, TLC, Raw, sRaw and sGaw, are exploratory endpoints. Analysis methods will be similar to the primary and secondary endpoints, but will be considered exploratory and focus on estimation.

3.2.9.4 Rescue Ventolin HFA use

Rescue medication use will be collected, but do not support any of the study objectives. Average number of puffs will be derived and listed for reference purposes.

3.3 Safety endpoints

Adverse events

Adverse Events (AEs), Serious Adverse Events (SAEs) and AEs leading to treatment discontinuation (DAEs) will be evaluated to meet the safety objectives of the study (Table 1). These endpoints will be collected in the adverse events eCRF.

AEs experienced by the subjects will be collected throughout the entire study and will be coded using the latest version of the Medical Dictionary for Regulatory Activities (MedDRA).

Only treatment emergent adverse events from the time of randomization onwards will be summarized. AEs will be assigned to the TP in which they first occurred, and summarized under the treatment group of that TP. AEs reported as starting during a Washout Period or follow-up period will be assigned to the last treatment received. If an AE has a missing or partial onset date, then the AE will be assumed treatment emergent, and assigned to the TP (or both TPs) where is could have possibly started.

Other safety assessments

The following assessments will be performed, but will not be used to define additional safety parameters for the study: clinical laboratory test, potential Hy's law case investigation results, physical examination, and 12-lead ECG. Data from these assessments will be captured in eCRF but will not be listed.

3.4 Handling of missing data

The lung comprises 5 lobes which are each divided into further segments. The airway continues to divide into smaller airways and each branching point defines a new generation of the airway. Data for FRI parameters will be provided for each subject for each lobe of the lung. Descriptive statistics and analyses may require that a subject's data are summarized by lobe and across lobes.

Where scans are not evaluable for a single assessment or period, this data is assumed to be missing at random. No imputation for entirely missing assessments is planned.

There may be missing data for certain lobes, if for example the subject is missing that lobe, or no airways lead to the lobe (giving a volume of zero and infinite airway resistance). If, for a subject, a particular lobe has missing data, then the following rules will be applied:

- **T-tests**: across-lobe averages will be computed using the non-missing data for matching lobes between the baseline and the post-baseline scan, provided that there are at least 4 matching lobes. For example, if data for LUL is present at visit 3, but missing at visit 4, the across-lobe averages at both visit 3 and visit 4 will be based on LLL, RUL, RML and RLL, using the denominator of 4.

- **Mixed models**: all data for individual lobes will be utilized.
- **Regional totals:** Summed totals of lobes will not be computed for that subject and visit for regions to which the missing lobe would belong.

If fissures between lobes cannot be distinguished, then only a combined estimate across multiple lobes may be produced. Such results from combined lobes will be handled as follows:

- **T-tests**: across-lobe averages will be computed, treating the parameter value for the combined lobes as the sum of individual lobe values in the numerator of the average. The number of lobes involved in the combination will count towards the denominator of the average. For example, if data are available for LUL, LLL, RUL+RML, and RLL, then the across-lobe average will be equal to the sum of these 4 values, divided by 5.
- **Mixed models**: the lobes involved in the combination will be regarded missing.
- **Regional totals:** the totals will be listed as provided by FLUIDDA. No calculations will be performed.

4 ANALYSIS METHODS

The primary objective of this study is to assess the effects of BGF and GFF on specific image-based airway volumes and resistance in subjects with moderate to severe COPD following chronic twice-daily (BID) dosing after approximately four weeks of treatment.

Additional exploratory analyses will compare the distribution of regional effects between the two treatments for the different FRI endpoints. The magnitude of differences between BFG and GFF effects overall and locally (by lobe) will be estimated. It will be investigated whether there is any evidence for differential patterns in the effects of each treatment.

4.1 General principles

4.1.1 Treatment effect assessment

For the efficacy analyses, BGF and GFF will be assessed for Day 29 effects. P-values for differences from baseline, and between treatments, as well as for tests of treatment by lobe interaction, will be reported as two-sided. For the co-primary endpoints, the other FRI parameters, as well as spirometry and plethysmography parameters, the Day 29 summary statistics for each treatment period will be reported and analysed.

The Intent-to-Treat (ITT) Population will be considered the primary analysis population for efficacy. Sensitivity analyses for the co-primary endpoints and secondary endpoints will be conducted using the modified Intent-to-Treat (mITT) Population, if the data included in the mITT analysis set is different from the data included in the ITT analysis set.

Continuous efficacy variables will be summarized with descriptive statistics (the number of non-missing values, mean, standard deviation, median, minimum and maximum). Where data have been logarithmically transformed for analysis, the summary statistics on the back-transformed data will include the geometric mean and the coefficient of variation (calculated as $100\sqrt{[\exp(s^2)-1]}$, where s is the standard deviation of the data on the log scale). Categorical variables will be summarized with frequency counts and percentages, by treatment.

4.1.2 Hypothesis testing

Formal hypothesis testing will only be performed for the two co-primary FRI endpoints (siVaw, siRaw) in each treatment group in the ITT analysis set.

Denote the log-transformed FRI value at baseline by $FRI_{(PRE)}$ and the post-treatment log-transformed FRI value at Day 29 by $FRI_{(POST)}$. The change from baseline for each subject will then be calculated as $FRI_{(CHG)} = FRI_{(POST)} - FRI_{(PRE)}$ and analysed for each treatment. The hypotheses for the respective "within-treatment" comparisons are:

H_0: $FRI_{(CHG)} = 0$, for BGF (no effect of treatment)
H_1: $FRI_{(CHG)} \neq 0$, for BGF

H_0: $FRI_{(CHG)} = 0$, for GFF (no effect of treatment)
H_1: $FRI_{(CHG)} \neq 0$, for GFF.

The hypothesis $FRI_{(CHG)} = 0$ will be tested using a one sample t-test within each treatment group and a two-sided p-value ≤ 0.05 will be regarded statistically significant.

Each t-test will be accompanied by the estimation of mean change from baseline and its 95% confidence interval, based on the t-distribution. This procedure also will be repeated for the endpoints other than primary, however, no formal statements will be made based on the resulting p-values. For those endpoints, the focus will be on the estimation of mean change from baseline, see section 4.2.4.1.

Additional exploratory analyses will compare the distribution of regional effects between the two treatments for the primary FRI endpoints. These treatment comparisons will not be considered to be formal hypothesis testing, but as estimation and investigation of regional treatment effects. Likewise, comparisons performed for the other efficacy endpoints will be considered descriptive.

Control of type I error

For the primary efficacy endpoints, Hochberg's step-up procedure will be used as a multiplicity adjustment to control the type I error within each treatment group using alpha level of 5%. Specifically, Hochberg's procedure will be applied once for the siVaw and siRaw endpoints for BGF, and then applied separately again for the same endpoints for GFF. No correction will be

performed for the secondary or exploratory endpoints. P-values for variables other than the primary endpoints will not be statistically interpreted.

4.1.3 Software

All statistical analysis will be conducted using SAS version 9.4, or other validated software as appropriate.

4.2 Analysis methods

4.2.1 Disposition, demographics and baseline characteristics

Descriptive summaries of disposition, demographics and baseline characteristics will be produced. Disposition tables will summarize the number of patients who received each treatment, the number who completed 1 or 2 treatment periods and the number of early discontinuations. The number of patients in each analysis set will be summarized along with any reasons for exclusion. Disposition tables will present data by the randomized treatment sequence.

Demographics and baseline characteristics at screening will be summarized descriptively, including COPD disease duration, disease severity, and symptom severity scores (COPD assessment test [CAT] and Modified Medical Research Council Dyspnea Scale Assessment [MMRC]).

Pre and post bronchodilator spirometry parameters (and their percent predicted values) and reversibility to Ventolin HFA will be summarized using descriptive statistics based upon the assessments at Screening. Subject-level baseline spirometry and body plethysmography parameters will be summarized. Predicted values for FEV_1 will be calculated using The Third National Health and Nutrition Examination Survey (NHANES III) reference equations. Predicted values for residual volume (RV) and functional residual capacity (FRC) will be calculated using Quanjer (1993).

Diffusion capacity parameters will be summarized at Screening.

All demographic and baseline characteristic summaries will be presented by treatment group, and for all subjects in the ITT analysis set. Subjects who had no assessments in TP 2 will not be summarized in the treatment group they were assigned to in TP 2. Thus, the treatment group summaries will potentially include fewer subjects than the ITT set.

4.2.2 Prior and concomitant medications

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

Concomitant medication is any medication reported as being taken after the start of the study medication in the study to the date prior to the last dose of study medication for the subject. A

medication with an onset date on or after the last dose of study medication for the subject will not be considered concomitant, but will be considered a **Post-Treatment medication**.

It will be identified for each concomitant medication whether it was taken during treatment with BGF, GFF, and during the washout, based on the definition of periods provided in section 1.2. Any medication/treatment which cannot be identified as Prior, Concomitant, or Post-Treatment will be considered as being in each of the categories that are possible from the available information.

All medications will be assigned a preferred term and an ATC (Anatomic Therapeutic Class) term using the latest version of the World Health Organization Drug Dictionary (WHO-DD) available. Prior medications will be summarized according to whether they are COPD related or non-COPD related. For concomitant medications, summaries will be presented for those medications started during a treatment period and those started during the washout period, separately for COPD related or non-COPD related. All medications will be listed.

4.2.3 Exposure and compliance

The number of days of exposure to each treatment (BGF, GFF) will be defined as ((End date of the last dose of IMP – Date of the first dose of IMP) + 1).

Exposure to IMP will be summarized by actual treatment using the safety analysis set.

Percent compliance in a given TP is defined as (total number of IMP puffs of study treatment taken on a study day/total expected IMP puffs taken on a study day) averaged across all days in the TP) x 100.

The expected number of IMP puffs is as follows:

- 2 puffs for a test day which is the last date of treatment.
- 4 puffs for the last date of treatment which is not a test day when an evening dose is taken.
- 2 puffs for the last date of treatment which is not a test day when an evening dose is not taken.
- 4 puffs on dates prior to the last date of treatment.

Treatment compliance will be categorized into 7 different groups depending on the degree of compliance: 0 – <20%, ≥20 – <40%, ≥40 – <60%, ≥60 – <80%, ≥80 – ≤100%, >100 – ≤120%, and >120%. Additionally, descriptive statistics (n, mean, standard deviation, median, minimum and maximum) will be provided for percent compliance by treatment group. Treatment compliance will also be listed.

4.2.4 Analysis of FRI data

Data for FRI endpoints will be logarithmically transformed prior to analysis, if so indicated in Table 4.

For FRI parameters the values for each lobe, the lower lobe total, upper lobe total, and overall total will be listed.

4.2.4.1 Across-lobes Analyses

The primary efficacy analyses will consist of within-treatment comparisons of Day 1 and Day 29, using a one-sample t-test applied to the change from baseline to Day 29 value. These comparisons will be done for each primary endpoint (siVaw, siRaw) within each treatment group (BGF, GFF), for a total of four comparisons. Point estimates and 95% confidence intervals based on the t-distribution will be reported for each of the parameter and treatment group, and displayed in bar charts.

Individual subject values of siVaw and siRaw at each visit will be presented graphically with line plots.

Where data are logarithmically transformed prior to analysis, across-lobe averaging will occur after applying the logarithm transformation. Estimates will be exponentiated, and ratios to baseline will be presented, along with 2-sided 95% confidence intervals and p-values.

The same methodology will be applied to the secondary endpoints (iVaw, iRaw) and the exploratory endpoints, listed in section 3.2.9.1. Note that for IAD and IADPP, across-lobe averaging is based on the lower lobes only.

4.2.4.2 Lobe-level Analyses

As exploratory analyses, the Day 29 value for each parameter in each period will be analysed using a linear mixed effect model. The data will be logarithmically transformed prior to analysis, where indicated in Table 4. A multi-level model will be used to incorporate the repeated measurements from the lobes for each subject.

The model will include fixed effects for period, treatment, lobe, and treatment-by-lobe interaction. The model will not include treatment sequence unless that term is determined to be important ($p<0.10$). Lobe will be included as a random effect, within each subject. Heterogeneity across lobes (within subject) will be modelled using an unstructured variance-covariance matrix, with independence assumed between subjects.

If this model fit fails to converge, a compound symmetry covariance structure will be considered to model correlation across lobes from the same patient.

The Kenward-Roger approximation for degrees of freedom and correction of downward bias in the standard error of fixed effect parameters will be used. The p-values for the treatment effect and the treatment-by-lobe interaction will be reported. Estimates will be produced of the within-treatment means and difference between treatments, for each lobe and across all lobes (across lower lobes for IAD). Where logarithmic transformations have been made, estimates will be exponentiated, and treatment effects presented as ratios. Estimates and standard errors will be reported along with 2-sided 95% confidence intervals and p-values.

Analyses similar to the primary endpoint analysis will be conducted for secondary and exploratory efficacy endpoints.

4.2.4.3 Sensitivity analysis

If the mITT population differs from the ITT population, the mITT population will be used to conduct sensitivity analyses for the co-primary endpoints and also for the secondary endpoints. The mITT population will only include data from subjects without major protocol deviations who complete both periods..

Additional sensitivity analyses may be performed for any efficacy endpoint on an ad hoc basis, for example if outliers are detected, or if the normality assumption cannot be justified.

4.2.5 Spirometry / body plethysmography data

The t-tests similar to those described in section 4.2.4.1 will be used for within-treatment comparisons of Day 1 and Day 29. Associated estimates of the change from period-dependent baseline to Day 29 will be reported for each treatment, with a 95% confidence interval and p-value.

For exploratory comparisons between treatment, the change from baseline to Day 29 for each endpoint will be analysed using a linear mixed effect model including subject-average baseline as a continuous covariate and treatment and period as fixed effects. The model will not include treatment sequence unless that term is determined to be important ($p<0.10$). Subject will be modelled as a random effect. The ITT population will be used. It is expected that plethysmography endpoints may require logarithmic transformation, but that spirometry endpoints will not. Estimates of the difference between treatments, with 95% confidence interval and p-value, will be reported.

4.2.6 Assumptions checks and removal of outliers in sensitivity analyses

The logarithmic transformations will be required for certain FRI endpoints (see Table 4) and expected for plethysmography endpoints. If data are transformed via the natural logarithm prior to analysis, the adjusted mean estimates and confidence intervals from the linear models will be exponentiated back for presentation. As such, the treatment effect for these endpoints will be presented as a ratio.

Hypotheses for within-treatment comparisons described in section 4.1.2 may be tested by means of the Wilcoxon signed rank test, in addition to the paired t-test, if the distribution of the data warrants it.

Under certain circumstances (e.g., during a COPD exacerbation unrelated to treatment), extreme and atypical values can arise. Such values may disproportionately affect model-based estimates of the fixed effect and variance parameters. Depending on the plots of the standardized residuals versus fitted values of the mixed-effect models, other data transformations may be considered. If necessary, this may include the fifth root, adapted for positive and negative values: fifth root (x) = sign (x) * (abs (x)) ^ (1/5).

For resistance measures, outliers can occur that are real values but affect the model estimates extremely. In these cases, values that are larger than the third quantile + 3 times the interquartile range may be removed. Infinite values for resistance are possible and these will be listed, but will be removed from the data before statistical analysis. If erroneous values are detected, every effort will be made to correct them prior to database lock. However, if these values cannot be corrected, they will be considered for removal from the analysis.

4.2.7 Exploratory analyses

Formal comparison of treatment groups was not planned. However, contrasts between treatment group means for FRI parameters will be obtained from the mixed models described in section 4.2.4.2. The focus will be on point estimates and confidence intervals.

The potential effect of blood eosinophils at screening on siVaw and siRaw (TLC, untrimmed) will be explored by additional modelling. For example, for the across-lobe average values siVaw and siRaw, simple linear regressions may be fit within each treatment group. The change from baseline in siVaw and siRaw will be the response, and the logarithm of blood eosinophils (log-EOS) at screening will be an independent variable. The effect of blood eosinophils at screening may be further explored by fitting the mixed model to the lobe-level values of siVaw and siRaw, with the additional terms of log-EOS and log-EOS by treatment interaction.

Scatter plots will be generated for each treatment group for FRI endpoints: siVaw, siRaw, iVaw, and iRaw, versus FRC by plethysmography, and for the same FRI endpoints versus post-dose FEV_1 (8 combinations in total). Pearson or Spearman's rank correlations between variables may be produced (depending upon behaviour of distributions). Additional scatter plots and correlations may be generated, if necessary.

4.2.8 Safety data

Adverse events will be summarized by the number and percentage of subjects experiencing an event. Tables will show the overall incidence of AEs, and the incidence for each treatment. Treatment emergent adverse events from the time of randomization onwards will be summarized.

AEs reported as starting during a Washout Period or follow-up period will be assigned to the last treatment received.

AEs will be presented according to MedDRA preferred term and system organ class. Summaries will be produced by intensity, seriousness, AEs leading to discontinuation, by causality assessment to study drug, and AEs leading to death. No hypothesis tests will be performed.

Other safety data obtained during the course of the study will be listed. Data obtained only at Visit 1 to determine subjects' eligibility will not be listed, but will be available in datasets. Refer to section 3.3 for the variables collected.

5 INTERIM ANALYSES (NOT APPLICABLE)

No interim analyses are planned in this study.

6 CHANGES OF ANALYSIS FROM PROTOCOL

COPD exacerbations were defined in section 8.1.11. Only serious COPD exacerbations will be captured as SAEs and will contribute to the SAE summaries. Severity of COPD exacerbations as defined in the protocol will not be determined.

The protocol was not specific about whether the secondary endpoint of change in FEV_1 was in regards to pre-dose or post-dose values. It has been clarified in the SAP that the post-dose FEV_1 will be the secondary endpoint.

7 REFERENCES

Quanjer, Ph., Tammeling, et al. (1993). Lung volumes and forced ventilatory flows. European Respiratory Journal - Eur Resp J. 6. Suppl. 16 5-40.

8 APPENDIX (NOT APPLICABLE)