AstraZeneca

Statistical Analysis Plan

Study Code	D3250C00038
Edition Number	2.0
Date	13 Aug 2018

A Multicenter, Randomized, Double-blind, Parallel Group, Placebo-controlled, Phase 3b Study to Evaluate the Onset of Effect and Time Course of Change in Lung Function with Benralizumab in Severe, Uncontrolled Asthma Patients with Eosinophilic Inflammation

A Multicenter, Randomized, Double-blind, Parallel Group, Placebo-controlled, Phase 3b Study to Evaluate the Onset of Effect and Time Course of Change in Lung Function with Benralizumab in Severe, Uncontrolled Asthma Patients with Eosinophilic Inflammation

Study Statistician [redacted]

A Multicenter, Randomized, Double-blind, Parallel Group, Placebo-controlled, Phase 3b Study to Evaluate the Onset of Effect and Time Course of Change in Lung Function with Benralizumab in Severe, Uncontrolled Asthma Patients with Eosinophilic Inflammation

Global Product Statistician [redacted]

[redacted]
Date

TABLE OF CONTENTS PAGE

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbreviation or special term	Explanation
ACQ-6	Asthma Control Questionnaire 6
ADA	Anti-drug antibodies
AE	Adverse event
ALT	Alanine aminotransferase
AST	Aspartate aminotransferase
ATC	Anatomic Therapeutic Chemical
ATS/ERS	American Thoracic Society/European Respiratory Society
BD	Bronchodilator
BMI	Body Mass Index
CGI-C	Clinical global impression of change
CI	Confidence interval
CSP	Clinical study protocol
CSR	Clinical study report
ECG	Electrocardiogram
eCRF	Electronic case report form
EOT	End of treatment
ePRO	Electronic patient reported outcome
ERT	eResearch Technology, Inc.
FeNO	Fractional exhaled Nitric Oxide
FEV_1	Forced expiratory volume in 1 second
FRC	Functional residual capacity
FSH	Follicle-stimulating hormone
FVC	Forced vital capacity
Gamma-GT	Gamma-glutamyl transpeptidase
GCP	Good clinical practice
IC	Inspiratory capacity
ICF	Informed Consent Form
ICS	Inhaled corticosteroids
IP	Investigational product
IPD	Premature IP Discontinuation

Abbreviation or special term	Explanation
ITT	Intent-to-Treat
IVRS	Interactive Voice Response System
LABA	Long-acting β_2 agonists
LOCF	Last observation carried forward
LSMEANS	Least Square Means
MedDRA	Medical Dictionary for Regulatory Activities
nAb	Neutralizing antibodies
PK	Pharmacokinetic(s)
PGI-C	Patient Global Impression of Change
PGI-S	Patient Global Impression of Severity
Post-BD	Post-bronchodilator
Pre-BD	Pre-bronchodilator
PRO	Patient reported outcome
PT	Preferred term
Raw	Airway resistance
RBC	Red blood cell
RV	Residual volume
SABA	Short-acting β_2 agonists
SAE	Serious adverse event
SAP	Statistical Analysis Plan
SC	Subcutaneous
SGaw	Specific airway conductance
SGRQ	St. George's Respiratory Questionnaire
SOC	System Organ Class
TLC	Total lung capacity
ULN	Upper limit of normal
VC	Vital capacity
WBC	White blood cell
WBDC	Web-based Data Capture
WHODD	WHO drug dictionary

AMENDMENT HISTORY

Date	Brief description of change
27 November 2016 (v1.0)	Initial version
13 August 2018 (v 2.0)	SAP amendment for the following: • Clarification that no visit window will be applied for summary of ADA or safety labs other than WBC counts and differentials • Update that no upper bound will be applied for follow up visit in Table 1 • Update that no post-baseline mean score will be derived for ACQ6 • Added FEV1 responder definition and analysis as secondary endpoint • Updated visit schedule for ACQ6, ACQ5 responder analysis and ACQ6 control status summary • Clarification that on-study summary of AE will be provided, but not on-treatment or post-treatment summaries • Added supportive analysis of onset of effect modelling in FEV1 • Details of analysis method is added for overall standardized effect plot • Subgroup analysis by baseline FeNO was added, and cutoffs for subgroup analysis by baseline IgE updated • Added correlation plot of baseline EOS and FeNO • Clarification on how BLQ values to be included in PK summary, and subjects to be included in NCA • Simplified MI analysis approaches in Appendix 8.1 • Clarified imputation for missing dates in medical history in Appendix 8.2 • Specifications on IPD were added

1. STUDY DETAILS

This statistical analysis plan (SAP) outlines the analyses to be generated for the global clinical study report (CSR). Additional analyses required for regional submissions will be pre-specified in a separate analysis plan and will be submitted to the appropriate authorities.

1.1 Study objectives

1.1.1 Primary objective

Primary Objective	Outcome Measure
To determine the effect of benralizumab on the time course of change (onset and maintenance of effect) on lung function	The average over the mean differences between benralizumab and placebo for the change from baseline (Visit 4) at Day 28 (Visit 8), Day 56 (Visit 9) and Day 84 (Visit 10) in pre-BD FEV_1 will be used to determine if the study is positive and to determine maintenance of effect. The first post baseline timepoint where the p-value for the mean difference between benralizumab and placebo is less than or equal to 0.05 will be used to determine time to onset of effect.

1.1.2 Secondary objectives

Secondary Objectives	Outcome Measures
To determine time course of effect of benralizumab on blood eosinophils and correlate changes in eosinophil depletion with lung function.	Change from baseline in blood eosinophils (from hematology) to end of treatment together with the correlation with lung function
To determine the effect of benralizumab on the time course of change and maintenance on lung function.	• Change from baseline (Visit 4) in pre-BD FEV_1 at all clinic visits compared to placebo • Percent of responders in pre-BD FEV1 by pre-specified cutoffs compared to placebo at all visits. • Change from baseline (Visit 4) in pre-BD forced vital capacity (FVC) at all clinic visits compared to placebo
To determine the time course of the effect of benralizumab on asthma control metrics	Change from baseline (Visit 4) in Asthma Control Questionnaire-6 (ACQ-6) score at all post-baseline assessment timepoints
To determine the time course of effect of benralizumab on health-related quality-of-life	Change from baseline (Visit 4) in St. George's Respiratory Questionnaire (SGRQ) score at all post-baseline assessment timepoints

Secondary Objectives	Outcome Measures
To determine the effect of benralizumab on the time course of change in exhaled nitric oxide (FeNO)	Change from baseline in FeNO (ppb) compared to placebo to end of treatment (Visit 10)

1.1.3 Safety objective

Safety Objectives	Outcome Measures
To evaluate the pharmacokinetics and immunogenicity of benralizumab	• Serum PK • Anti-drug antibodies
To assess the safety and tolerability of benralizumab	• Adverse events (AEs) and serious adverse events (SAEs) • Laboratory variables • Physical Examination

1.1.4 Exploratory objectives

Exploratory Objective	Outcome Measures
To evaluate patient impression of overall asthma severity (PGI-S) and overall change from baseline as reported by the patient (PGI-C) and clinician (CGI-C)	• Change from baseline (Visit 4) in PGI-S at all post-baseline assessment timepoints • Clinician-reported global change in asthma from baseline (Visit 4) as measured by CGI-C • Patient-reported global change in asthma from baseline (Visit 4) as measured by PGI-C

1.1.5 Body plethysmography sub-study objectives

Sub-study Primary Objective	Outcome Measure
To determine the effect of benralizumab on the time course of change in lung function as assessed through body plethysmography	Change from baseline (Visit 4) to end of treatment (Visit 10) in the following measures compared to placebo: • Residual volume (RV)
Sub-study Secondary Objectives	**Outcome Measures**

To determine the effect of benralizumab on the time course of change in lung function as assessed through body plethysmography	Change from baseline (Visit 4) to end of treatment (Visit 10) in the following measures compared to placebo: • Total lung capacity (TLC) • RV/TLC ratio • Inspiratory capacity (IC) • Functional residual capacity (FRC) • Vital capacity (VC)
Sub-study Exploratory Objective	**Outcome Measures**
To determine the effect of benralizumab on the time course of change in lung function as assessed through body plethysmography	Change from baseline (Visit 4) to end of treatment (Visit 10) in the following measures compared to placebo: • Specific airway conductance (SGaw) • Airway resistance (Raw)

1.2 Study design

This is a randomized, double-blind, parallel group, placebo-controlled study designed to evaluate the onset of effect and safety of a fixed 30 mg dose of benralizumab administered subcutaneously (SC) to patients with severe, uncontrolled asthma. Approximately 230 patients with peripheral blood eosinophil counts ≥300 cells/μL will be randomized globally to receive SC benralizumab 30 mg or placebo.

After enrolment, eligible patients will enter a 5-week screening/run-in period. Patients who meet eligibility criteria will enter a 12-week treatment period and receive 30 mg benralizumab or placebo at Day 0, Day 28 (+/- 3 days), and Day 56 (+/- 3 days). An End-of-Treatment (EOT) visit will be conducted at Day 84 (Week 12) and a follow-up (FU) visit will be conducted at Day 112 (Week 16).

See the clinical study protocol (CSP) Section 4, Tables 2 and 3 for a detailed list of visits and assessments.

1.3 Number of subjects

The study will recruit approximately 230 patients in total and 115 patients in each treatment group based on the primary objective of the study to determine the effect of benralizumab on the time course of change (onset and maintenance of effect) of pre-bronchodilator (BD) FEV_1.

Patients will be randomized in a 1:1 ratio to benralizumab or placebo. In order to have a balanced number of patients for each region and the body plethysmography sub-study, the randomization will be stratified by these variables.

The study is powered for the primary objective of the study to determine the effect of benralizumab on the time course of change (onset and maintenance of effect) of pre-BD FEV_1.

The first post-baseline timepoint where the p-value for the mean difference between benralizumab and placebo is ≤0.05 will be used to determine time to onset of effect. The average over the mean differences between benralizumab and placebo based on the change from baseline (Visit 4) to Days 28 (Visit 8), 56 (Visit 9), and 84 (Visit 10) in pre-BD FEV_1 will be used to determine if the study is positive and to determine maintenance of effect.

The results of the Phase 2b study, MI-CP220 (Castro et al 2014) and the Phase 3 asthma exacerbation studies, SIROCCO (Bleeker et al 2016) and CALIMA (Fitzgerald et al 2016) were used as a basis for the sample sizing. From these studies, it is expected that around 115 patients per treatment arm (230 in total) will provide approximately 90% power for the primary endpoint pre-BD FEV_1 with type I error controlled at a two-sided alpha level of 0.05 if the true average over the mean treatment differences (benralizumab – placebo) across Days 28 (Visit 8), 56 (Visit 9), and 84 (Visit 10) is 138 mL. The sample size will also be sufficient to have approximately 80% power for an average over the mean treatment difference of 120 mL. The assumed variability has a within-group standard deviation of 375 mL and within-patient correlation of 0.6. If the within-patient variability is reduced to 350 mL, the sample size would be sufficient 80% power for an average of 110 mL. In addition, it is expected that around 115 patients per treatment arm will have 80% power to detect a true treatment difference (benralizumab – placebo) of 130 mL, assuming a within-group standard deviation of 350 mL or a true treatment difference of 140 mL, assuming a standard deviation of 375 mL. Based on these assumptions, the minimum difference that would be statistically significant at the 5% level ranges from approximately 85 to 100 mL. The sample size was estimated using nQuery + nTerim version 3.0 using methodology by Liu et al (Liu et al 2005).

2. ANALYSIS SETS

2.1 Definition of analysis sets

Five analysis sets are defined below:

- All patient analysis set
- Full analysis set (FAS)
- Sub-study analysis set
- Safety analysis set
- Pharmacokinetics (PK) analysis set

Patients must have provided their informed consent. If no signed informed consent is collected (major protocol deviation), the patient will be excluded from all analysis sets defined below.

All efficacy will be performed according to the Intent-to-Treat (ITT) principle based on the FAS. For consistency, demographic and baseline characteristics will be presented using the FAS. Safety objectives will be analyzed based on the safety analysis set.

2.1.1 All patients analysis set

This analysis set comprises all patients screened for the study and will be used for reporting of disposition and screening failures.

2.1.2 Full analysis set

All patients randomized and receiving at least 1 dose of investigational product (IP) will be included in the FAS, irrespective of their protocol adherence and continued participation in the study. Patients will be analyzed according to their randomized treatment, irrespective of whether or not they have prematurely discontinued, according to the ITT principle. Patients who withdraw consent, and assent when applicable, to participate in the study will be included up to the date of their study termination. All efficacy and patient-reported outcome (PRO) data will be based on this analysis set.

2.1.3 Body plethysmography sub-study analysis set

The subset of patients who are randomized as part of the body plethysmography sub-study and receive any IP will be analyzed as described for the FAS.

2.1.4 Safety analysis set

All patients who received at least 1 dose of IP will be included in the safety analysis set. Patients will be classified according to the treatment they actually received. A patient who has on 1 or several occasions received active treatment will be classified as active. All safety summaries and anti-drug antibodies (ADA) analyses will be based on this analysis set.

2.1.5 Pharmacokinetic analysis set

All patients who received benralizumab and from whom PK blood samples are assumed not to be affected by factors such as protocol violations and who had at least 1 evaluable serum PK observation post first dose will be included in the PK analysis set. All PK summaries will be based on this analysis set.

2.2 Violations and deviations

Patients who do not meet eligibility criteria but are still randomized will be analyzed according to the analysis sets described in Section 2.1. There is no intention to perform a per-protocol analysis in this study.

2.2.1 Important protocol deviations

Important deviations will be those which are considered to potentially impact upon the interpretation of the primary endpoint in the study. Only important protocol deviations will be listed and tabulated in the CSR.

The following categories of protocol deviations will be reviewed by medical advisors and statisticians prior to database lock to determine those which are considered important deviations as outlined below, with additional details provided in Appendix 8.3:

- Patients who do not meet key inclusion criteria (patients incorrectly randomised)
- Patients who meet any of the key exclusion criteria (patients incorrectly randomised)
- Patients who developed IP discontinuation criteria during the study but were not discontinued
- Patients who received prohibited/restricted concomitant medication(s)
- Deviations from study procedures

Patients for whom significant protocol deviations were recorded that impact the interpretation of the study safety outcomes will have a footnote added to applicable output to describe the deviation and its potential impact. Such patients will be identified as part of the protocol deviation review process prior to database lock.

Patients who received incorrect study treatment will be identified after database lock and unblinding. Patients will be classified according to the treatment they actually received for the safety analysis set according to the following rules, based on the unblinded data from the Interactive Voice Response System (IVRS) system:

- Patients assigned to benralizumab treatment regimen who receive only placebo will be assigned to the placebo group.
- Patients assigned to the placebo group who received 1 or more doses of benralizumab will be assigned to the benralizumab group.

2.2.2 Visit window definitions

No visit windows will be defined for screening visits. For local laboratory and all vital signs, safety labs other than white blood cell (WBC) cell count and differentials, and ADA/neutralizing antibodies (nAb), the visit recorded in web-based data capture (WBDC) will be used. For endpoints that present visit-based data including WBC cell count and differentials (including eosinophil count data), spirometry, FeNO, body plethysmography (sub-study), and patient reported outcomes (PROs), the variables will be summarized based on the scheduled days with adjusted analysis-defined visit windows. The adjusted analysis-defined visit windows will be based on the collection schedule listed in the protocol and variables will be windowed to the closest scheduled visit for that variable.

The adjusted analysis-defined visit windows for assessments conducted are summarized in Table 1.

Table 1 Visit windows for assessments

Adjusted defined windows visit	Scheduled study day	Maximum windows
Week 0 (Day 0)	1	Study Day=1
Day 3	4	2 ≤Study Days≤5
Day 7	8	6 ≤Study Days≤11
Day 14 (Week 2)	15	12 ≤Study Days≤22
Day 28 (Week 4)	29	23 ≤Study Days≤42
Day 56 (Week 8)	57	43 ≤Study Days≤70
Day 84 (Week 12/EOT)	85	71 ≤Study Days≤98
Day 112 (Week 16/Follow-up)	113	99 ≤Study Days

For assignment of data to adjusted analysis-defined visit windows, study day will be defined as follows: (Date of assessment – date of randomization) +1

For example, by this definition, the day of randomization will be Study Day 1 and the planned date of Week 4 will be Study Day 29 (=28+1).

If multiple assessments are recorded within a single visit window, please refer to the rules below:

- If there are 2 or more observations within the same visit window, then the non-missing one closest to the scheduled visit will be used in the analysis.
- If 2 observations are equidistant from the scheduled visit, then the non-missing observation with the earlier collection date will be used in the analysis.
- If 2 observations are collected on the same day, then the non-missing one with the earlier collection time will be included in the analysis.

If a visit window does not contain any observations, then the data will remain missing.

For pre-BD FEV_1 (L) and body plethysmography data, the non-missing value with acceptable quality (acceptable or borderline quality grade) which is closest to the scheduled visit will be included in the analysis. For post-BD FEV_1, the highest value with acceptable quality from the same date as the pre-BD FEV_1 result will be used.

2.2.3 The definition of baseline

In general, the last recorded value on or prior to the date of randomization will serve as the baseline measurement for efficacy endpoints, while the last recorded value prior to first dose of study treatment will serve as the baseline measurement for safety endpoints. If there is no value prior to randomization (or the first dose of study treatment, depending on the endpoint), then the baseline value will not be imputed and will be set to missing. No data known to be

collected post first dose will be used in determining the baseline value, unless otherwise specified.

For categorical (yes/no) baseline respiratory disease characteristics, if 'yes' is indicated in any visit during enrollment (scheduled or unscheduled), 'yes' will be used.

There are 2 pre-BD spirometry assessments – 30 and 60-mins prior to IP administration on Day 0 of randomization visit. Day 0 pre-BD FEV1 value will be set to:

- The average of the two if both have acceptable quality, or
- the value of the one with acceptable quality if only 1 has acceptable quality, or
- missing, if none of the two has acceptable quality

Day 0 pre-BD value of FVC, FEF 25-75% and ratio of FEV1/FVC will be calculated using the same logic as for Day 0 pre-BD FEV1.

For pre-BD FEV_1 (L) and sub-study assessments from whole body plethysmography, the last non-missing value with acceptable quality (acceptable or borderline quality grade) on or prior to the date of randomization will be used as baseline. The first dose of study treatment is scheduled to be administered on the date of randomization (Visit 4); however, if the first dose of study treatment is delayed until after the date of randomization, the last recorded value with acceptable quality prior to the first dose of study treatment will be used as baseline measurement for spirometry parameters. Baseline of pre-BD FVC will be assigned in the same way as for pre-BD FEV_1.

For FeNO, the mean of repeated assessments at Week 0 (Visit 4) prior to randomization will be used as baseline values. If the first dose of study treatment is delayed until after the date of randomization, the mean of last recorded values prior to the first dose of study treatment will be used as baseline. If the assessments prior to randomization/first dose are missing, the mean value from Day -21 (Visit 2) will be used for baseline.

ACQ-6, SGRQ, and PGI-S assessments at Week 0 (Visit 4) prior to randomization will be used for calculating the respective baselines. For ACQ-6, if the value at Week 0 (Visit 4) is missing, the non-missing value from Day -35 (Visit 1) will be used for baseline.

For summaries of vital signs and laboratory results by visit and in the calculation of change from baseline, the last non-missing assessment prior to first dose of study treatment will be used as baseline.

2.2.4 Prior/concomitant medications

All patients are required to be treated with stable dose inhaled corticosteroids (ICS)/ Long-acting β_2 agonists (LABA) for at least 30 days prior to Visit 1 and during the course of the study.

The background asthma controller medications should be maintained at a stable dose from Visit 1 until the end of the study. No changes are allowed to background medications throughout the duration of the study except during the treatment of an asthma exacerbation.

Short-acting β_2 agonists (SABAs) may be used as rescue medication during the study in the event of a worsening of asthma symptoms.

A medication will be regarded as 'prior' if it was stopped on or before the date of randomization (medication stop date ≤ date of randomization).

A medication will be regarded as 'concomitant' if the start date is after the date of randomization, or if it started on or prior to the date of randomization and was ongoing after the date of randomization.

3. EFFICACY AND SAFETY VARIABLES

3.1 Primary efficacy outcome variable

The primary efficacy variable is the change from baseline in pre-BD FEV_1. Details regarding the definition of baseline for pre-BD FEV_1 are provided in Section 2.2.3.

3.2 Secondary efficacy outcome variables

3.2.1 FEV1 responder

A FEV1 responder at each post-baseline visit will be defined as a patient who had improvement on FEV1 by the following cut-offs: 50, 100, 150, 200, 250 and 300 mL, ie, a subject with a change from baseline in FEV1 greater than or equal to the cut-off at a visit will be categorized as a responder at that visit. A patient with missing value of change from baseline at a visit will be excluded from the calculation of proportion of responders at that visit.

3.2.2 Eosinophil counts in the blood

Absolute and percentage changes from baseline Week 0 (Visit 4) in blood eosinophil counts (from hematology) at each post-baseline visit are the secondary efficacy variables.

These variables will also be used to assess the treatment effect over time of eosinophils in the blood.

3.2.3 Pre-bronchodilator forced vital capacity

The change from baseline (Visit 4) in FVC to each of the post-baseline visits will be a secondary efficacy variable for this study.

3.2.4 Asthma Control Questionnaire-6

The Asthma Control Questionnaire-6 (ACQ-6) is a shortened version of the ACQ that assesses asthma symptoms (nighttime waking, symptoms on waking, activity limitation, shortness of

breath, wheezing, and short-acting β_2 agonist use) omitting the FEV_1 measurement from the original ACQ-7 score.

Patients are asked to recall how their asthma has been during the previous week by responding to 1 bronchodilator use question and 5 symptom questions. Questions are weighted equally and scored from 0 (totally controlled) to 6 (severely uncontrolled). The mean ACQ-6 score is the mean of the responses. Mean scores of ≤0.75 indicate well-controlled asthma, scores between 0.75 and <1.5 indicate partly controlled asthma, and a score ≥1.5 indicates not well controlled asthma. Individual changes of at least 0.5 are considered to be clinically meaningful.

The questionnaire will be completed at the study center via provided site tablet (ePRO) in accordance with schedule provided in CSP Section 4, Tables 2 and 3.

The outcome variable for ACQ-6 will be the change in mean score from baseline to each of the post-randomization periods. There will be no imputation for missing values.

Patients will be categorized according to the following limits (Juniper et al 2005):

- ACQ-6 (post-baseline visit – baseline) ≤-0.5 → Improvement
- -0.5 < ACQ-6 (post-baseline visit – baseline) <0.5 → No change
- ACQ-6 (post-baseline visit – baseline) ≥0.5 → Deterioration

An ACQ-6 responder will be defined as a patient who had improvement on ACQ-6, ie, an ACQ-6 responder variable takes value 1 if change from baseline to post-baseline visit in ACQ-6 ≤-0.5 and 0 otherwise.

Furthermore, patients will be categorized according to their ACQ-6-defined asthma control status at the baseline and at each post-baseline visit using the following score thresholds (Juniper et al 2006):

- ACQ-6 ≤0.75 → Well-controlled
- 0.75< ACQ-6 <1.5 → Partly controlled
- ACQ-6 ≥1.5 → Not well-controlled

In addition to ACQ-6, the ACQ-5 score will be calculated by omitting the question related to rescue medication use. The change in mean score from baseline to each of the post-randomization periods will be calculated and the asthma control responder status will be derived as defined above for ACQ-6. Patients will not be categorized as well-controlled, partly controlled, and not well-controlled based on ACQ-5.

3.2.5 St. George's Respiratory Questionnaire

The St. George's Respiratory Questionnaire (SGRQ) is a 50-item PRO instrument developed to measure the health status of patients with airway obstruction diseases (Jones et al 1991). The questionnaire is divided into 2 parts:

- Part 1 consists of 8 items pertaining to the severity of respiratory symptoms in the preceding 4 weeks;
- Part 2 consists of 42 items related to the daily activity and psychosocial impacts of the individual's respiratory condition.

The SGRQ yields a total score and 3 domain scores (symptoms, activity, and impacts). The total score indicates the impact of disease on overall health status. This total score is expressed as a percentage of overall impairment, in which 100 represents the worst possible health status and 0 indicates the best possible health status. Likewise, the domain scores range from 0 to 100, with higher scores indicative of greater impairment. Specific details on the scoring algorithms are provided by the developer in a user manual (Jones et al 2009).

The change in mean total score in SGRQ from baseline (Visit 4) to each of the post-randomization periods in total score will be a secondary efficacy variable for this study.

For the responder analysis of SGRQ, a responder will be defined as an individual with a ≥4-point decrease (improvement) in SGRQ total score from baseline to each post-baseline visit.

3.2.6 Fractional exhaled nitric oxide measurement

The mean change from baseline of fractional exhaled nitric oxide (FeNO) to EOT visit will be a secondary efficacy variable.

Airway inflammation will be evaluated using a standardized single-breath FeNO (ATS 2005) test.

Since spirometry can potentially impact the nitric oxide measurement, the FeNO test needs be completed prior to spirometry. In addition, patients should not eat or drink 1 hour prior to having the FeNO, as this may affect the results.

While sitting, patients are to inhale to total lung capacity through the NIOX MINO® Airway Inflammation Monitor (Aerocrine, New Providence, New Jersey) and then exhale for 10 seconds at 50 mL/sec (assisted by visual and auditory cues). The value obtained will be recorded and the process will be repeated multiple times. Average of valid replicates provided by the ERT lab will be used in the analysis of the FeNO (ppb) data.

3.3 Exploratory efficacy outcome variables

3.3.1 Patient Global Impression of Severity

Patient Global Impression of Severity (PGI-S) is a single question asking the patient to rate the severity of their symptoms using a 6-point categorical response scale (0=no symptoms, 6=Very severe symptoms). Improvement from baseline using the PGI-S will include changes from moderate, severe, and very severe to no symptoms, very mild, and mild.

3.3.2 Clinician Global Impression of Change and Patient Global Impression of Change

Clinician Global Impression of Change (CGI-C) and Patient Global Impression of Change (PGI-C) instruments are used for an overall evaluation of response to treatment. The Investigator (clinician) and the patient will be asked to rate the degree of change in the overall asthma status compared to the start of treatment, ie, randomization visit. A 7-point rating scale will be used: 1=Very Much Improved, 2=Much Improved, 3=Minimally Improved, 4=No Changes, 5=Minimally Worse, 6=Much Worse, and 7=Very Much Worse.

The Investigator and the patient will be asked to rate the degree of change in the overall status compared with the randomization visit. There will be no imputation for missing values.

Patients will also be categorized according to the following responses post-baseline, separately for CGI-C and PGI-C:

- Very much improved, much improved, minimally improved → 'Improved'
- Very much improved, much improved → 'Much improved'
- Very much improved → 'Very much improved'

Agreement between CGI-C and PGI-C will be assessed at each visit, where agreement is achieved when both the Investigator and the patient provide the same response (eg, if both the Investigator and the patient indicate a response of 1 (very much improved) at a particular visit, agreement is achieved for that visit). Agreement will also be assessed for categorized responses at each visit.

A responder will be defined for CGI-C and PGI-C separately as anyone with a response of Improved, Much improved, or Very much improved at the EOT visit.

3.4 Sub-study efficacy outcome variables

Additional assessments will be performed for a subset of patients at designated sites, approximately 50 patients in the study, to determine the effect of benralizumab on the time course of change in lung function as assessed through body plethysmography.

The mean change from baseline (Visit 4) to each of the post-randomization visits up to and including the end of treatment visit (Visit 10) will be used as efficacy variables within the

body plethysmography sub-study analysis set. Baseline is defined as the last non-missing value prior to randomization.

3.4.1 Primary efficacy variable

The primary variable for the sub-study is the change from baseline in residual volume (RV).

3.4.2 Secondary efficacy variables

The secondary efficacy variables for the sub-study include the change from baseline in total lung capacity (TLC), ratio of RV/TLC, inspiratory capacity (IC), and functional residual capacity (FRC), vital capacity (VC).

3.4.3 Exploratory efficacy variables

The exploratory variables for the sub-study are change from baseline in Specific airway conductance (SGaw) and Airway resistance (Raw).

3.5 Safety outcome variables

The following safety data will be collected: reported AEs and SAEs, clinical chemistry, hematology, urinalysis, 12-lead electrocardiogram (ECG), physical examination, and vital signs.

All safety measurements will use all available data for analyses, including data from unscheduled visits and repeated measurements.

Change from baseline to each post-treatment timepoint where scheduled assessments were made will be calculated for relevant measurements. AEs will be summarized by means of using descriptive statistics and qualitative summaries.

No safety data will be imputed. The handling of partial/missing dates for AEs and prior/concomitant medications is detailed in Appendix 8.2.

3.5.1 Adverse events

Adverse events experienced by the patients will be collected throughout the entire study and will be coded using the latest version of the Medical Dictionary for Regulatory Activities (MedDRA) per the Data Management Plan.

Adverse event data will be categorized according to their onset date into the following study period:

- Adverse events in the on-study period are defined as those with onset date on or after the day of first dose of study treatment.

If an AE has a missing onset date, then unless the stop date of the AE indicates otherwise, this will be considered an ongoing AE. Similarly, if an AE has a partial onset date, then unless the partial onset date or the stop date indicates otherwise, this will be considered an ongoing AE.

3.5.2 Laboratory variables

Blood and urine samples for determination of clinical chemistry, hematology, and urinalysis parameters will be taken at the times detailed in the CSP and will be assessed in a central laboratory. The parameters outlined in Section 5.2.1, Table 4 of the CSP will be collected.

In summaries, figures, and listings, lab results and normal ranges will be presented in System International (SI) units. Eosinophil data will be presented in both SI and conventional units (cells/μL) in summaries.

For the purposes of clinical chemistry, hematology, and urinalysis shift tables, baseline will be defined as the last available non-missing assessment prior to first dose of randomized treatment, and maximum or minimum value post-baseline will be calculated over the entire post-baseline period, including the post-treatment period.

Changes in hematology and clinical chemistry variables between baseline and each post-baseline assessment will be calculated. The change from baseline is defined as the post-baseline visit value minus the baseline visit value. There will be no imputation for missing values. For values recorded with a leading greater than or less than ('>', '<') symbol, the reported numeric value will be used for analysis and the value with the symbol will be included in the listings, unless otherwise specified. For example, a value of <0.01 will be analyzed as 0.01 and listed as <0.01.

Absolute values will be compared to the relevant reference range and classified as low (below range), normal (within range or on limits), or high (above range). The central reference ranges will be used for laboratory variables. All values (absolute and change) falling outside the reference ranges will be flagged.

Urinalysis data will be categorized as negative (0), positive (+), or strongly positive (++, +++, or >+++) at each timepoint.

For the liver function tests: aspartate aminotransferase (AST), alanine aminotransferase (ALT), alkaline phosphatase (ALP), gamma-GT (GGT), and total bilirubin (TBL), the multiple of the central laboratory upper limit of the normal (ULN) range will be calculated for each data point.

$$\text{Multiple} = \text{Value}/\text{ULN}$$

That is, if the ALT value was 72 IU/L (ULN=36) then the multiple would be 2.

Patients who meet any of the following criteria at any point during the study will be flagged:

- AST ≥3x ULN
- ALT ≥3x ULN
- TBL ≥2xULN

3.5.3 Local electrocardiogram

Electrocardiogram (ECG) measurements will be performed at Visit 1 and the FU visit. The outcome of the overall evaluation is to be recorded as normal/abnormal in the electronic case report form (eCRF) by the Investigator/authorized delegate, with any abnormalities being recorded as not clinically significant or clinically significant.

3.5.4 Physical examination

Complete and brief physical examinations will be performed at timepoint specified in Tables 2 and 3 of the CSP.

Baseline data will be collected at Visit 1. Any new finding(s) or aggravated existing finding(s), judged as clinically significant by the Investigator, will be reported as an AE as described in Section 6.3 of the CSP.

3.5.5 Vital signs

Pre-dose vital signs (pulse, systolic blood pressure, diastolic blood pressure, respiration rate, and body temperature) will be obtained in accordance with the visit schedule provided in the CSP.

Changes in vital signs variables between baseline and each subsequent scheduled assessment will be calculated. Baseline is defined as the last value prior to the first dose of randomized treatment. The change from baseline is defined as the post-baseline visit value minus the baseline visit value. There will be no imputation for missing values.

Absolute values will be compared with the reference ranges in Table 2 and classified as low (below range), normal (within range or on limits), or high (above range). All values (absolute and change) falling outside the reference ranges will be flagged.

Table 2 Vital signs reference ranges

Parameter	Standard units	Lower limit	Upper limit
Diastolic Blood Pressure (DBP)	mmHg	60	120
Systolic Blood Pressure (SBP)	mmHg	100	160
Pulse Rate	Beats/min	40	120
Respiratory Rate	Breaths/min	8	28
Body Temperature	Celsius	36.5	38
Weight	kg	40	200

Body mass index (BMI) will be calculated from the height and weight as follows:

$$\text{BMI (kg/m}^2\text{)=weight (kg)/(height [m])}^2$$

3.6 Pharmacokinetic variables

Pharmacokinetic trough concentration will be assayed for each visit where serum samples are collected.

In addition, estimation of PK parameters, ie, C_{max}, t_{max}, and $AUC_{[0,28d]}$ will be computed for the first dosing cycle using a noncompartmental estimation method. C_{max} and T_{max} are direct measurements of the maximum observed concentration and the time to reach the C_{max}, respectively, during the first dosing cycle.

3.7 Immunogenicity variables

Anti-drug antibodies assessments will be conducted utilizing a tiered approach (screen, confirm, titer). The presence of nAb will be tested in all ADA-positive samples using a ligand binding assay.

For each patient, the following response variables will be evaluated. Samples taken post dose on the day of IP administrations are excluded from analysis.

- ADA positive at any visit (including baseline and unscheduled visits)
- ADA positive at both post-baseline and baseline
- ADA positive at post-baseline only
- ADA positive at baseline only
- nAb positive at any visit

Safety outcomes by ADA status (if the number of occurrences ADA responses is sufficient) will also be evaluated.

4. ANALYSIS METHODS

4.1 General principles

The analysis of the primary and secondary efficacy endpoints will include all data captured during the 12-week double-blind treatment period, defined as the period after administration of randomized IP at Visit 4 (Week 0) and the conclusion of Visit 10 (Week 12), inclusive. This includes data regardless of whether study treatment was prematurely discontinued or delayed, and/or irrespective of protocol adherence, unless the patient withdraws consent to study participation. The statistical analyses will compare benralizumab with placebo. The analysis of safety endpoints will include all data captured during the on-study period, defined as the period after first administration of randomized IP at Visit 4 (Week 0) and the conclusion of the scheduled post-treatment FU visit (Week 11), inclusive.

The data analyses will be conducted using the SAS® System 9.2 (SAS Institute Inc., Cary, NC) or higher version. Pharmacokinetic analyses will be performed using WinNonlin (version 6.3, or higher, Certara, L.P., St. Louis, MO).

Summary data will be presented in tabular format by treatment group. Descriptive statistics on continuous variables will be summarized by treatment group using mean, standard deviation, median, quartile, and range as applicable, while categorical data will be summarized using frequency counts and percentages. When data are summarized by time, the values recorded against the scheduled timepoints listed in the protocol will be used. When assessing minimum/maximum increases or decreases during the study, all assessments, including unscheduled assessments will be used. For analysis assessing change from baseline, only patients with both baseline and at least 1 evaluable post-baseline measure will be included.

Given there is a single primary endpoint and treatment comparison, no formal testing strategy will be used for the analysis in this study.

All hypothesis testing will be reported using 2-sided tests. All p-values apart from the primary analyses of the primary endpoint will be nominal. All p-values will be rounded to 4 decimal places.

The absolute change from baseline is computed as (*visit value – baseline value*). Percent change from baseline is computed as ([*visit value – baseline value*]/*baseline value*)×100%. If either a visit value or the baseline visit value is missing, the absolute change from baseline value and the percent change from baseline will also be set to missing.

The prior medications, categorized according to the WHO Drug Reference List dictionary which employs the Anatomical Therapeutic Chemical (ATC) classification system, will be summarized by treatment group as frequency and percentage of patients reporting usage.

The concomitant medication will be categorized according to the WHO Drug Reference List dictionary which employs the Anatomical Therapeutic Chemical (ATC) classification system. The frequency and percentage of patients taking concomitant medications and non-drug therapies during the treatment period will be summarized by drug class and drug name using the ATC codes.

4.2 Analysis methods

4.2.1 Patient disposition

Patient disposition will be summarized using the all patients analysis set. The total number of patients will be summarized for the following groups: those who enrolled, and those who were not randomized (and reason). The number and percentage of patients within each treatment group will be presented by the following categories: randomized, received treatment with study drug, did not receive treatment with study drug (and reason), completed treatment with study drug, discontinued treatment with study drug (and reason), discontinued treatment with

study drug but completed study follow-up, completed study, and withdrawn from study (and reason).

Screen failure information will be listed for the all patients analysis set.

The number of patients randomized by country and center will also be summarized by treatment group in the FAS.

4.2.2 Demography data and patient characteristics

Demography data such as age, gender, race, and ethnicity will be summarized by treatment group for all patients in the FAS. Age will be derived from the date of informed consent date-of-birth, rounded down to the nearest integer. For patients in country where date of birth is not recorded, the age as recorded in the eCRF will be used.

Various baseline characteristics will also be summarized by treatment for the FAS. These include maintenance asthma medications, maintenance ICS medications, medical and surgical histories, FEV_1 (pre- and post-BD) at baseline, and respiratory disease characteristics including age at onset of asthma, number of exacerbations in the previous 12 months, and number of exacerbations requiring hospitalizations in the previous 12 months.

The following will be summarized for patients in the FAS and repeated by region and for the body plethysmography sub-study set:

- Patient characteristics (weight, height, BMI, baseline eosinophil count)
- Baseline lung function data (FEV_1 [L], FEV_1 [% PN], FVC [L], FVC [% PN], FEV_1/FVC, FEF 25-75% [L/S], reversibility (%)
- ACQ-6 score
- Respiratory disease characteristics

Medical and surgical histories will be summarized by MedDRA Preferred Term (PT) within a MedDRA System Organ Class (SOC).

The number of patients remaining on treatment, patients discontinued IP but still in study follow-up, and patients who withdraw from the study will be presented by treatment group and scheduled visit.

4.2.3 Prior and concomitant medications

The number and percentage of patients taking maintenance asthma medications, including ICS/LABA fixed dose combinations, at baseline will be summarized.

The number and percentage of patients who take prior medications, those who take allowed concomitant medications, and those who take disallowed concomitant medications during the study will be presented by treatment group. Concomitant medications will be classified

according to the WHO Drug Dictionary (WHODD). The summary tables will present data by generic term within the ATC code.

4.2.4 Study Treatments

4.2.4.1 Exposure

Exposure to IP will be calculated in days as:

Last dose date of IP-first dose date of IP+1,

and will be summarized by treatment group for the safety analysis set.

4.2.4.2 Study treatment compliance

Study treatment compliance will be summarized by treatment group for the FAS and calculated as:

Study treatment compliance=(total doses administered/total doses expected)x100.

Patients who received no study treatment will have zero compliance. Total number of doses expected includes all visits with protocol scheduled IP administration on or before a subject's IP discontinuation or treatment complete date.

4.2.5 Primary outcome variable

4.2.5.1 Primary analysis

The change from baseline in pre-BD FEV_1 up until the EOT at Visit 10 will be compared between the 30 mg benralizumab group and placebo using a mixed-effect model for repeated measures (MMRM) analysis on patients with a baseline pre-BD FEV_1 and at least 1 post-randomization pre-BD FEV_1 in the FAS.

The dependent variable will be the change from baseline in pre-BD FEV_1 at post-baseline protocol specific visits (up to and including the EOT Visit). Treatment group will be fitted as the explanatory variable, region, visit and treatment*visit interaction as fixed effects, and baseline pre-BD FEV_1 will be fitted as a covariate. The variance-covariance matrix will be assumed to be unstructured. If the procedure does not converge, then a compound symmetric variance-covariance matrix will be used instead. The model is:

*Change in FEV_1=treatment+baseline pre-BD FEV_1+region+visit+treatment*visit*

The primary objective of the study is to determine the effect of benralizumab on the time course of change (onset and maintenance of effect) of pre-BD FEV_1 using the model described above. The first post-baseline timepoint where the p-value for the mean difference between benralizumab and placebo is ≤ 0.05 will be used to determine time to onset of effect. The average over the mean differences between benralizumab and placebo based on the change from baseline (Visit 4) to Days 28 (Visit 8), 56 (Visit 9), and 84 (Visit 10) in pre-BD FEV1 will be used to determine if the study is positive as well as to determine the

maintenance of effect. Contrasts will be used to obtain estimates of the treatment differences over Days 28, 56, and 84, as well as at each timepoint separately.

Results will be presented in terms of least square means (LSMEANS), treatment differences in LSMEANS, 95% confidence intervals (CI) of treatment differences, and p-values at each planned timepoint.

4.2.5.2 Sensitivity analysis

A sensitivity analysis of pre-BD FEV_1 may be performed as applicable by excluding patients who meet at least one of situations below:

- Who do not withhold their regular asthmas medication prior to spirometry assessment
- Who do not withhold their SABA for more than 6 hours prior to spirometry assessment

4.2.5.3 Subgroup analysis

To explore the uniformity of the detected overall treatment effect on the primary efficacy variable, subgroup analyses, including statistical modelling where the interaction between treatment and covariates to estimate the treatment effect within each subgroup, may be performed for the following factors for maintenance of effect: OCS use at baseline (yes/no), gender, race, age group categories (<65, ≥65 years), geographic region, BMI (≤35, >35 kg/m^2), the number of exacerbations during the previous year (2, ≥3 exacerbations), nasal polyps (yes/no), baseline immunoglobulin E (IgE) concentration (<150kU/L, ≥150kU/L), baseline EOS and FeNO. In addition, an overall standardized effects plot and forest plot across all subgroups will be produced for maintenance of effect. ANCOVA will be utilized in the overall standardized effects plot to assure convergence and avoid run-time issue. Average value from Day 28, 56 and 84 will be the response variable. Treatment group and baseline pre-BD FEV1 will be included in the model for overall effect model. Treatment group, subgroup being evaluated, interaction of treatment group and subgroup being evaluated, and baseline pre-BD FEV1 will be included in the subgroup effect model. Missing data will be imputed by last observation carried forward (LOCF) prior to the average of Day 28, 56 and 84 is derived.

Additionally, subgroup analysis of pre-BD FEV_1 will be performed to explore the relationship with baseline blood eosinophil counts using the categories ≥300-449/μL and ≥450/μL, and with baseline FeNO using the categories of ≤50 and >50 ppb. Blood eosinophil data collected at Screening visit will be used for this categorization to ensure that all subjects can be classified and included in the summary.

It is important to note that the study has not been designed or powered to assess efficacy within any of these pre-defined subgroups, and as such, these analyses are considered exploratory.

4.2.5.4 Supportive analyses

The modelled treatment response as well as differences between benralizumab and placebo using the mean change from baseline in pre-BD FEV_1 endpoint using an exponential model over time will be estimated. The model will be used to determine the modelled time to onset of effect.

Onset of effect will be defined as: (i) the time taken for benralizumab and placebo to reach a pre-defined change from baseline value, (ii) the time taken for the difference between benralizumab and placebo [change from baseline values to reach a pre-defined value.

As detailed in the study protocol, the main focus is to estimate onset of effect using the primary variable, pre-BD FEV1. Onset of effect will be estimated using change from baseline measurements which were scheduled on days 3, 7, 14, 28, 56, and 84, actual day that data is collected in study is utilized. An exponential relationship will be assumed over time (Gelman, et al 2013), using Bayesian methodology (Marostica, et al 2013) to fit a nonlinear mixed effects model to the observed data. The fitted model will then be used to estimate onset of effect for the set of change from baseline values 50, 100, 150, 200, 250, and 300 mL and the set of differences 50, 100, 150, 200, 250, and 300 mL. In addition, an average change from baseline will be estimated for days 1, 3, 7, 14, 21 and 28, 56, 84, and the proportion of subjects with a change from baseline value greater than or equal to {50, 100, 150, 200 and 250 and 300 mL} will be estimated for days 1, 3, 7, 14, 21, 28, 56, 84. Differences between proportions and odds ratios will also be estimated.

It is anticipated that an exponential model will adequately describe pre-BD FEV1 change from baseline over time, based on (Gelman, et al 2013) and modelling similar data taken from the Phase III SIROCCO and CALIMA trials. However, if the data show that an exponential model is a poor choice, an alternative model will be investigated following the same principles outlined below.

In addition to pre-BD FEV1, one or more secondary variables may also be investigated.

As described in the protocol, measurements for pre-BD FEV1 will be recorded on day 0, 3, 7, 14, 28, 56, 84. For every subject, pre-BD FEV1 change from baseline values will be calculated for day 3, 7, 14, 28, 56, 84. If a baseline value or post-baseline measurement is missing, the corresponding change from baseline value will be set to missing. Onset of effect will be estimated for subjects in the full analysis set (FAS) with a baseline value and at least one pre-BD FEV1 measurement recorded after the first dose of study drug.

Model

For each treatment separately, Bayesian methodology will be used to fit the following nonlinear mixed effects model to the observed pre-BD FEV1 change from baseline data

$$y_{ij} = C_i\left(1 - e^{-\theta_i^* t_{ij}}\right) + \varepsilon_{ij}$$

where y_{ij} denotes the change from baseline value for the i-th subject at the j-th time point $(i = 1, \cdots, n;\ j = 1, \cdots, 7)$

εij ~ t (0, σ^2, df)

$$C_i \sim N(C, \sigma_C^2)$$

$$\theta_i^* = e^{\log\theta_i}$$

$$\log\theta_i \sim N(\log\theta, \sigma_{\log\theta}^2)$$

Note that C is an asymptote, representing the maximum (or minimum) change from baseline over time and θ is a rate, constrained to be greater than zero, representing how quickly the change from baseline values increase or decrease over time. This model takes into account the inherent structure that is present in the observed data and gives an estimate of C and θ for a specific treatment group, and allows these parameters to vary between subjects.

Prior distributions for the model parameters will be:

$$C \sim N(0, 1e6)$$

$$\log\theta \sim N(0, 1e6)$$

σ^2 ~igamma (shape=0.01, scale=0.01)

σ_c^2 ~igamma (shape=0.01, scale=0.01)

$\sigma_{\log\theta}^2$ ~igamma (shape=0.01, scale=0.01)

df ~ uniform (2, 50)

Note that igamma refers to inverse gamma distribution.

For each parameter, a chain of 200,000 samples will be drawn from the posterior distribution, after discarding the first 100000 burn-in samples. The 200,000 samples will be thinned every 200th observation to ensure there is little correlation in the remaining 1,000 posterior samples. Following the same process, a further two chains will be run, whereby the seed and set of starting values will be changed for each of these additional chains. Seed numbers of 3598, 2917 and 6301 will be utilized for benralizumab, and seed numbers of 1256, 8734 and 7376 for placebo group. Starting values for both treatment groups will be 25, 50 and 75 mL for C, and -1 for $\log\theta$. Sensitivity analysis may be carried out for different starting values from those specified to assess the robustness of model fitting.

Note that when modelling the pre-BD FEV1 measurements, data collected beyond Day 98 (upper visit window of Day 84) will be excluded in the analysis.

Model checks

The convergence of the model will be investigated by assessing stationarity and mixing across chains. If convergence is not met, it may be deemed necessary to consider alternatives for (i) the set of priors, (ii) the distributions associated with the model parameters describing the data, (iii) the exponential model. Any deviation from the planned analysis will be documented as a change to the statistical analysis plan.

- Stationarity

  For each parameter, the 1000 posterior samples will be plotted against iteration number for each chain separately. Posterior summary will be provided for C and logθ from each chain, the efficiency associated with the effective sample size will also be checked for each parameter to ensure that this statistic is high (e.g. greater than 0.9) across chains. If an obvious pattern exists in the posterior samples or the efficiency < 0.9 for a specific chain, auto-correlation exists and stationarity has not been reached. Consequently, a larger number of samples will be discarded as burn-in, and the number of samples per chain will be increased with a greater amount of thinning.

- Mixing

  For each parameter, the 1000 posterior samples from each chain will be plotted together against iteration number, with a different coloured line used for each chain. If the chains overlap, they are deemed to mix well and they will be combined to give 3000 posterior samples for each model parameter. If the chains do not mix well, the proposed model will be reviewed and refined if deemed necessary.

Estimating the change from baseline at time points of interest in a treatment group, and difference in change from baseline at time points of interest between treatment groups

For a specific treatment group, let $\{C_{ik}, \log \theta_{ik}\}$ denote the set of estimated model parameters for the i-th subject from the k-th posterior sample, where $i = 1, \cdots, n$ and $k = 1, \cdots, 3000$ (assuming the 3 chains are stationary and mix well). Let t denote a time point of interest, where t=1, 3, 7, 14, 21, 28, 58, 84. The change from baseline at each time point of interest will be estimated for benralizumab and placebo by following the steps below:

i. For the k-th posterior sample $(k = 1, \cdots, 3000)$, the fitted change from baseline value at time point t for the i-th subject $(i = 1, \cdots, n)$ in a specific treatment group is

$$y_{ik}(t) = C_{ik}(1 - e^{-\theta^*_{ik} t})$$

where $\theta^*_{ik} = e^{log\theta_{ik}}$.

ii. The mean change from baseline value at time t for the k-th posterior sample will be calculated as

$$\bar{y}_k(t) = \frac{1}{n}\sum_{i=1}^{n} y_{ik}(t)$$

iii. The 50th and (2.5th, 97.5th) percentiles of $\{\bar{y}_1(t), \cdots, \bar{y}_{3000}(t)\}$ will be derived, giving the median and corresponding 95% credible interval for the average change from baseline value at time point t

iv. Steps (i) through to (iii) will be carried out for each treatment and all values of t.

Note that it is anticipated that some subjects may withdraw from the study early. For these subjects, their fitted response will be extrapolated beyond their last visit. Consequently, inferences will be based on all subjects in the analysis set.

Estimates and credible intervals for the change from baseline from the modelling above will be tabulated by treatment and time point. A figure will be provided for the change from baseline values over time.

To estimate the difference between treatment groups in change from baseline at time points of interest, assuming the 3 chains are stationary and mix well, by independence the k-th posterior sample for benralizumab will be associated with the k-th posterior sample for placebo $(k = 1, \cdots, 3000)$.

v. The difference between treatment groups in mean change from baseline value at time t for the k-th posterior sample will be calculated as (mean change from baseline value in benralizumab) – (mean change from baseline value in placebo), where mean within treatment group is derived as per step ii above.

vi. The 50th and (2.5th, 97.5th) percentiles of the differences {d1, …, d3000} at time t will be derived, giving the median and corresponding 95% credible interval for the difference between benralizumab and placebo in change from baseline at time t. If either the 50th, 2.5th or 97.5th percentile is greater than T=84, such values will be reported as ">84".

Estimating the time taken for a treatment to reach a pre-defined change from baseline value

For a specific treatment group, let $\{C_{ik}, \log \theta_{ik}\}$ denote the set of estimated model parameters for the i-th subject from the k-th posterior sample, where $i = 1, \cdots, n$ and $k = 1, \cdots, 3000$ (assuming the 3 chains are stationary and mix well). Let Δ denote a pre-defined change from baseline value, where Δ = 50, 100, 150. 200, 250, and 300 mL. Let T be the *planned* duration of the study in days, where T=84 is divided into 1000 time points to give $\{t_1, \cdots, t_{1000}\}$. The time taken to reach a pre-defined change from baseline value will be estimated for benralizumab and placebo by following the steps below:

i. For the k-th posterior sample $(k = 1, \cdots, 3000)$, the fitted change from baseline at time point t_j $(j = 1, \cdots, 1000)$ for the i-th subject $(i = 1, \cdots, n)$ in a specific treatment group is

$$y_{ik}(t) = C_{ik}(1 - e^{-\theta_{ik}^* t})$$

where $\theta_{ik}^* = e^{log\theta_{ik}}$.

ii. The mean change from baseline value at time t_j for the k-th posterior sample will be calculated as

$$\bar{y}_k(t_j) = \frac{1}{n}\sum_{i=1}^{n} y_{ik}(t_j)$$

iii. For pre-BD FEV1, an improvement is associated with $\Delta > 0$. Thus, from the set of values $\{\bar{y}_k(t_1), \cdots, \bar{y}_k(t_{1000})\}$, all values greater than or equal to a specific value of Δ will be retained. The earliest time point will be selected and labelled t_k^* (which gives the earliest time point at which $\bar{y}_k(t_j)$ is closest to Δ). If there are no values that are greater than or equal to Δ, the time t_k^* will be set to a large value greater than T (e.g. $t_k^* = 10000$). Consequently, all posterior samples will be included in the summary of $\{t_1^*, \cdots, t_{3000}^*\}$.

iv. The 50th and (2.5th, 97.5th) percentiles of $\{t_1^*, \cdots, t_{3000}^*\}$ will be derived, giving the median and corresponding 95% credible interval for the average time taken for the change from baseline to reach Δ. If either the 50th, 2.5th or 97.5th percentile is greater than T=84, such values will be reported as ">84".

v. Steps (i) through to (iv) will be carried out for each treatment and all values of Δ

Note that it is anticipated that some subjects may withdraw from the study early. For these subjects, their fitted response will be extrapolated beyond their last visit. Consequently, inferences will be based on all subjects in the analysis set.

Estimates and credible intervals for the time taken for a treatment to reach a pre-defined change from baseline value from the modelling above will be tabulated by treatment and pre-defined change from baseline value.

Estimating the time taken for the difference between benralizumab and placebo [change from baseline values] to reach a pre-defined value

Assuming the 3 chains are stationary and mix well, by independence the k-th posterior sample for benralizumab will be associated with the k-th posterior sample for placebo $(k = 1, \cdots, 3000)$. For the k-th posterior sample, let $\{C_{ik}, \log\theta_{ik}\}$ denote the set of estimated model parameters for the i-th subject in the benralizumab group ($i = 1, \cdots, n$) and let $\{D_{lk}, \log\beta_{lk}\}$ denote the set of estimated model parameters for the l-th subject in the placebo group ($l = 1, \cdots, m$). Let τ denote a pre-defined difference between benralizumab and

placebo change from baseline values, where τ = 50, 100, 150, 200, 250, and 300 mL. Let T be the *planned* duration of the study in days, where T=84 is divided into 1000 time points to give $\{t_1, \cdots, t_{1000}\}$. The time taken for the difference between benralizumab and placebo to reach a pre-define value will be calculated as follows:

i. For the k-th posterior sample $(k = 1, \cdots, 3000)$, the fitted change from baseline at time point t_j $(j = 1, \cdots, 1000)$ for the l-th subject ($l = 1, \cdots, m$) in the placebo group is

$$u_{lk}(t_j) = D_{lk}(1 - e^{-\beta^*_{lk} t_j})$$

where $\beta^*_{lk} = e^{log\beta_{lk}}$ and the mean change from baseline value at time t_j for the k-th posterior sample will be calculated as

$$\bar{u}_k(t_j) = \frac{1}{m}\sum_{l=1}^{m} u_{lk}(t_j)$$

ii. For the k-th posterior sample $(k = 1, \cdots, 3000)$, the fitted change from baseline at time point t_j $(j = 1, \cdots, 1000)$ for the i-th subject ($i = 1, \cdots, n$) in the benralizumab group is

$$v_{ik}(t_j) = C_{ik}(1 - e^{-\theta^*_{ik} t_j})$$

where $\theta^*_{ik} = e^{log\theta_{ik}}$ and the mean change from baseline value at time t_j for the k-th posterior sample will be calculated as

$$\bar{v}_k(t_j) = \frac{1}{n}\sum_{i=1}^{n} v_{ik}(t_j)$$

iii. For the k-th posterior sample $(k = 1, \cdots, 3000)$, the average difference between benralizumab and placebo at time point t_j will be calculated as

$$\bar{d}_k(t_j) = \bar{v}_k(t_j) - \bar{u}_k(t_j)$$

iv. For pre-BD FEV1, an improvement is associated with $\tau > 0$. Thus, from the set of values $\{\bar{d}_k(t_1), \cdots, \bar{d}_k(t_{1000})\}$, all values greater than or equal to a specific value of τ will be retained. The earliest time point will be selected and labelled t^*_k (which gives the earliest time point at which $\bar{d}_k(t_j)$ is closest to τ). If there are no values that are greater than or equal to τ, the time t^*_k will be set to a large value greater than T (e.g. $t^*_k = 10000$). Consequently, all posterior samples will be included in the summary of $\{t^*_1, \cdots, t^*_{3000}\}$.
Note that for a secondary variable an improvement could be associated with $\tau < 0$.

Thus, from the set of values $\{\bar{d}_k(t_1), \cdots, \bar{d}_k(t_{1000})\}$, *all values less than or equal to a specific value of* τ *will be retained. The earliest time point will be selected and labelled* t^*_k *(which gives the earliest time point at which* $\bar{d}_k(t_j)$ *is closest to* τ*). If there*

are no values that are less than or equal to τ, the time t_k^* will be set to a large value greater than T (e.g. $t_k^* = 10000$). Consequently, all posterior samples will be included in the summary of $\{t_1^*, \cdots, t_{3000}^*\}$.

v. The 50th and (2.5th, 97.5th) percentiles of $\{t_1^*, \cdots, t_{3000}^*\}$ will be derived, giving the median and corresponding 95% credible interval for the average time taken for the average difference between benralizumab and placebo to reach τ. If either the 50th, 2.5th or 97.5th percentile is greater than T=84, such values will be reported as ">84".

vi. Steps (i) through to (v) will be carried out for all values of τ

Estimates and credible intervals for the time taken for the difference between benralizumab and placebo [change from baseline values] from the modelling above will be tabulated by treatment and pre-defined difference value.

Estimating the proportion of subjects ≥ pre-defined thresholds at time points of interest

For a specific treatment group, let $\{C_{ik}, \log \theta_{ik}\}$ denote the set of estimated model parameters for the i-th subject from the k-th posterior sample, where $i = 1, \cdots, n$ and $k = 1, \cdots, 3000$ (assuming the 3 chains are stationary and mix well). Let t denote a time point of interest, where t=1, 3, 7, 14, 21, 28, 58, 84. Let Δ denote a pre-defined change from baseline value, where Δ= 50, 100, 150, 200, 250, and 300 mL. At a specific time point, the proportion of subjects $\geq \Delta$ will be estimated for benralizumab and placebo by following the steps below:

i. For the k-th posterior sample $(k = 1, \cdots, 3000)$, the fitted change from baseline at time point t for the i-th subject $(i = 1, \cdots, n)$ in a specific treatment group is

$$y_{ik}(t) = C_{ik}(1 - e^{-\theta_{ik}^* t})$$

where $\theta_{ik}^* = e^{log\theta_{ik}}$.

- If, at time point t:
  $y_{ik}(t) \geq \Delta$ set $I_{ik} = 1$
  otherwise set $I_{ik} = 0$

- The proportion of subjects will be calculated as

$$p_k = \frac{1}{n}\sum_{i=1}^{n} I_{ik}$$

iv. The 50th and (2.5th, 97.5th) percentiles of $\{p_1, \cdots, p_{3000}\}$ will be derived, giving the median and corresponding 95% credible interval for the proportion of subjects greater than or equal to a specific value of Δ at time point t.

v. Steps (i) through to (iv) will be carried out for each treatment and all values of t and Δ

Estimates and credible intervals for the proportion of responders ≥ pre-defined thresholds at time points of interest from the modelling above will be tabulated by treatment, pre-defined thresholds and timepoints of interest.

Estimating the difference between proportions

Following the steps outlined in Section 5.4, let $\{p_1, \cdots, p_{3000}\}$ denote the set of proportions for benralizumab and let $\{q_1, \cdots, q_{3000}\}$ denote the set of proportions for placebo for a specific value of t and Δ.

i. For the k-th posterior sample $(k = 1, \cdots, 3000)$, calculate the difference $d_k = p_k - q_k$

ii. The 50th and (2.5th, 97.5th) percentiles of $\{d_1, \cdots, d_{3000}\}$ will be derived, giving the median and corresponding 95% credible interval for the difference between proportions

iii. Steps (i) through to (ii) will be carried out for all values of t and Δ

Difference of proportions between treatment groups over time will be tabulated for each pre-defined threshold. Figures will be provided.

Estimating the odds ratio

Following the steps outlined in Section 5.4, let $\{p_1, \cdots, p_{3000}\}$ denote the set of proportions for benralizumab and let $\{q_1, \cdots, q_{3000}\}$ denote the set of proportions for placebo for a specific value of t and Δ.

i. For the k-th posterior sample $(k = 1, \cdots, 3000)$, calculate the odds ratio $o_k = \frac{p_k/(1-p_k)}{q_k/(1-q_k)}$

ii. The 50th and (2.5th, 97.5th) percentiles of $\{o_1, \cdots, o_{3000}\}$ will be derived, giving the median and corresponding 95% credible interval for the odds ratio

iii. Steps (i) through to (ii) will be carried out for all values of t and Δ

In the cases when proportions of responders or non-responders in either treatment group is zero, odds ratio would be estimated using the following approach:

i) calculation of odds ratio on iteration level: when proportion of responders or non-responders is 0 in a treatment group, make adjustment by adding 0.5 to all cell counts, eg out of N=100 in a treatment arm, if there is 0 responder, and 100 non-responders observed in an iteration, replace the count of responder with 0.5, and non-responder with 100.5 in that treatment arm. And add 0.5 to cell counts in the other treatment arm. Then calculate odds ratio for that

iteration. This applies to cases when 0 count occurs in either or both Benra and Placebo groups.

ii) summary statistics across 3000 iterations: if more than 5% odds ratios are estimated as outlined in step i) due to 0 count, then set the estimates (median, 95% credible intervals) to NC; otherwise, median, and 95% credible intervals are based on 3000 iteration results with estimated odds ratios (as outlined in step i) included.

Odds ratio over time will be tabulated for each pre-defined threshold. Figures will be provided.

Alternative models

Convergence issues may arise when carrying out the planned analysis and it may be deemed necessary to consider alternatives for (i) the set of priors, (ii) the distributions associated with the model parameters describing the data and (iii) the exponential model.
However, if such changes still lead to convergence issues, a simplified analysis will be carried out whereby an exponential model (or alternative) will be fitted through the mean change from baseline values for each treatment. Consequently, focus will be based on point estimates only, without any measures of precision, given that between-subject and within-subject components could not be incorporated in the model successfully. In this case, we will estimate

i. the change from baseline at t=3,7,14,28,56,84
ii. the time taken for benralizumab and placebo to reach Δ = 50, 100, 150, 200, 250, and 300 mL
iii. the time taken for the difference between benralizumab and placebo change from baseline values to reach τ = 50, 100, 150, 200, 250, and 300 mL

Note that all deviations from the planned analysis will be documented as a change in the statistical analysis plan

4.2.6 Secondary efficacy outcome variables

4.2.6.1 Secondary analyses

For each secondary efficacy variable described in Section 3.2, the objective is to determine the effect of benralizumab on the time course of change (onset and maintenance of effect).

FEV1 responder at each post-baseline visit will be defined as a subject with change from baseline ≥ prespecified cutoffs as in Section 3.1, and analysed using a logistic regression model with repeated measures using generalized equation estimate method. The model will include treatment, region, visit and treatment group by visit interaction as fixed factors, and baseline FEV1 scores as a covariate. The variance-covariance matrix will be assumed to be unstructured.

The percent change from baseline at each timepoint for the eosinophils counts, change from baseline at each timepoint for Pre-BD FVC, ACQ-6 score, SGRQ score, and FeNO will be compared between benralizumab and placebo using an MMRM analysis described in the primary analysis above. The least squares mean for the difference in treatment groups using the interaction between visit and treatment group, its 95% CI, and the 2-sided p-values will be reported for each post-baseline visit.

As for the primary endpoint, the first post-baseline timepoint where the p-value for the mean difference between benralizumab and placebo is less than or equal to 0.05 from the repeated measures analysis will be used to determine time to onset of effect for each secondary parameter. The average over the mean differences between benralizumab and placebo based on the change from baseline (Visit 4) to Days 28 (Visit 8), 56 (Visit 9), and 84 (Visit 10) will be used to determine maintenance of effect. Contrasts will be used to obtain estimates of the average treatment differences over Days 28, 56, and 84.

Results will be presented in terms of LSMEANS, treatment differences in LSMEANS, 95% CIs of treatment differences, and p-values at each planned timepoint. To explore correlations between changes in eosinophil depletion and lung function, and correlations between changes in FeNO and lung function, the following analyses will be conducted, blood eosinophil data collected at Screening visit will be used in the analyses below to be consistent with the subgroup analysis of pre-BD FEV1 by baseline EOS categories:

- Analyses will be conducted on change from baseline versus baseline eosinophil count at Day 28 and Day 84 in pre-BD FEV_1 using a moving average and MMRM approach. In this analysis, patients will be ordered by baseline eosinophil count and grouped by percentile into overlapping subgroups ($0 – 40^{th}$, $10^{th} – 50^{th}$, $20^{th} – 60^{th}$, $30^{th} – 70^{th}$, $40^{th} – 80^{th}$, 50^{th} - 90^{th}, $60^{th} –100^{th}$). The MMRM model will be then applied to each subgroup to estimate the treatment difference at Day 28 and Day 84 separately and associated 95% CI for each percentile grouping.

- A scatterplot of observed change in pre-BD FEV_1 at Day 28 and 84 versus baseline eosinophils at Day 28 and Day 84 will be prepared separately with locally weighted scatter plot smoothing curves (LOESS) and corresponding 95% CIs superimposed.

- Same kinds of MMRM and LOESS analyses will be performed using baseline FeNO data.

- If significant treatment effect is detected at timepoints earlier than Day 28, the same kinds of MMRM and LOESS analyses will also be performed at the earliest timepoint where significant treatment effect is detected.

A scatter plot with correlation coefficient will be provided for baseline FeNO and EOS count data to evaluate the association between baseline FeNO and EOS.

4.2.6.2 Proportion of responders by PRO measures

Responder at the post-baseline visit from the selected secondary endpoints below will be analyzed using a logistic regression model with repeated measures using generalized equation estimate method. The model will include treatment, region, visit and treatment group by visit interaction as fixed factors, and baseline scores as a covariate. The variance-covariance matrix will be assumed to be unstructured.

ACQ-6 responder at each post-baseline visit: Will be based on change from baseline to each post-baseline visit as defined in Section 3.3.2. Patients with missing or non-evaluable score at a post-baseline visit will be considered non-responders.

SGRQ responder at each post-baseline will be defined for each post-baseline visit as an individual with a ≥4-point decrease (improvement) from baseline in SGRQ total score. Patients with missing or non-evaluable score at any post-baseline visit will be considered non-responders.

4.2.7 Exploratory efficacy outcome variable

4.2.7.1 Exploratory analysis of PGI-S

The same analysis approach described in the primary analysis will be performed on the change from baseline in PGI-S at each post-baseline visit. A summary table will be produced showing the number and percentage of patients with improvement from baseline at each post-baseline visit.

Patients with a PGI-S improvement at the EOT visit will be analyzed using a logistic regression model with covariates of treatment and region.

4.2.7.2 Exploratory analysis of CGI-C and PGI-C

The CGI-C and PGI-C responses will be summarized by treatment group and visit. The number and percentage of patients will be presented for CGI-C, PGI-C, and for agreement in CGIC and PGIC responses as described in Section 3.3.2.

The number and percentage of patients defined as responders based on categorized responses for CGI-C and PGI-C (improved, much improved, very much improved) will also be presented by treatment group and visit.

Responders at the EOT visit will be analyzed using a logistic regression model with covariates of treatment and region.

4.2.8 Sub-study body plethysmography outcome variable

The objective of the sub-study is to determine the effect of benralizumab on the time course of change (onset and maintenance of effect) of body plethysmography outcome variables. Only the outcome values with acceptable quality (acceptable or borderline quality grade) will be used for analyses. All the observed and change from baseline of sub-study variables will be summarized by treatment and visit.

4.2.8.1 Sub-study primary analysis

The change from baseline in RV at each timepoint will be compared between benralizumab and placebo using an MMRM analysis in the sub-study population if sufficient acceptable quality data warrants.

The dependent variable will be the change from baseline in RV at post-baseline protocol specific visits (up to and including the EOT Visit). Treatment group will be fitted as the explanatory variable, visit and treatment*visit interaction as fixed effects, and baseline RV will be fitted as a covariate. Region may be fitted as a fixed effect. The variance-covariance matrix will be assumed to be unstructured. If the procedure does not converge, then a compound symmetric variance-covariance matrix will be used instead.

As for the primary efficacy analysis, contrasts will be used to obtain estimates of the treatment differences over Days 28, 56, and 84 as well as at each timepoint separately. Results will be presented in terms of LSMEANS, treatment differences in LSMEANS, 95% CIs of treatment differences, and p-values at each planned timepoint.

4.2.8.2 Sub-study secondary analysis

The change from baseline at each timepoint for TLC, ratio TLC/RV, VC, IC and FRC may be compared between benralizumab and placebo using an MMRM analysis in the sub-study population as for the primary efficacy analysis of sub-study as described above.

4.2.8.3 Sub-study explorative analysis

Only descriptive statistics for SGaw and Raw and change from baseline will be provided at each scheduled visit at which assessments are performed.

Subgroup analyses of RV and ratio TLC/RV may be performed for the age group categories (18-30 years, 31-65 years).

4.2.9 Safety outcome variables

All safety variables will be summarized using the safety analysis set and data presented according to actual treatment received.

4.2.9.1 Adverse events

Adverse events (AEs) will be summarized for the on-study period, as defined in Section 3.5.1. All summaries will be presented by treatment group.

An overall summary table will be produced showing the number and percentage of patients with at least 1 AE in any of the following categories: AEs, serious adverse events (SAEs), AEs with outcome of death, and AEs leading to discontinuation of IP (DAEs). The total number of AEs in the different AE categories in terms of AE counts will also be presented (ie, accounting for multiple occurrences of the same event in a patient).

Adverse events, AEs with outcome of death, and SAEs will be summarized by SOC and PT assigned to the event by MedDRA. For each PT, the number and percentage of patients reporting at least 1 occurrence will be presented (ie, multiple occurrences of an AE for a patient will only be counted once). Adverse events causing discontinuation of the study treatment and SAEs causing discontinuation from the study will also be summarized.

The rate of AEs per person-years at risk, calculated as (number of patients reporting AE)/(total period with patients at risk for AE), will also be reported. The total period at risk for each patient will be defined as the period from first dose of study treatment to the follow-up visit (Week 16). Rates will be expressed in terms of events per 100 patient-years.

A summary of the most common (frequency of ≥5%) AEs will be presented by PT. Adverse events will be summarized by PT and Investigator's causality assessment (related versus not related) and maximum intensity. If a patient reports multiple occurrences of the same AE within the same study period, the maximum intensity will be taken as the highest recorded maximum intensity (the order being mild, moderate, and severe). Adverse events of injection site reactions (high-level term of administration and injection site) and hypersensitivity (standardized MedDRA query of hypersensitivity) will be summarized by PT. The summary of injection site reactions will be summarized by injection site location and number of IP administrations.

Adverse events related to device malfunction will be listed. All adverse events will be listed.

4.2.9.2 Laboratory data

All continuous laboratory parameters will be summarized descriptively by absolute value at each visit by treatment group, together with the corresponding changes from baseline. All parameters will be summarized in SI units, with the exception of blood eosinophil counts which will be summarized in both SI and conventional units. Results which are reported from the central laboratory in conventional units will be converted to SI units for reporting.

Central laboratory reference ranges will be used for the identification of abnormalities, and a shift table will be produced for each laboratory parameter to display low, normal, high values, and missing values. The shift tables will present baseline and post-baseline maximum/minimum value, as applicable, for each parameter and will include patients with both baseline and post-baseline data.

Shift plots showing each individual patient's laboratory value at baseline and at maximum/minimum post-baseline will be produced for each continuous laboratory variable. If any laboratory variables show any unusual features (high or low values or a general shift in the data points) at other timepoints, then shift plots of these data may be produced.

Data for patients who have treatment-emergent changes outside central laboratory reference ranges will be presented. This data presentation will include all visits for this subset of patients.

Maximum post-baseline bilirubin elevations by maximum post-baseline ALT and AST will be presented, expressed as multiples of ULN. Bilirubin will be presented in multiples of the following ULN: ≤1.5, >1.5-2, and >2. Alanine aminotransferase and AST will be presented in multiples of the following ULN: ≤1, >1-3, >3-5, >5-10, and >10.

Maximum post-baseline total bilirubin (TBL) will be presented (<2 and ≥2xULN) and plotted against maximum post-baseline ALT (<3, ≥3-<5, ≥5-<10, and ≥10 x ULN), expressed as multiples of ULN. This will be repeated to show maximum post-baseline TBL against maximum post-baseline AST.

Data for patients with ALT or AST ≥3xULN, and TBL≥2xULN will be presented, which will include all visits for this subset of patients. A line plot of liver biochemistry test results (including ALP, ALT, AST, TBL, and GGT) over time will also be presented for this subset of patients.

For urinalysis data, a shift table will be generated to present changes from baseline to maximum post-baseline value for each parameter and will include patients with both baseline and post-baseline data.

Descriptive statistics will be presented for IgE baseline value by treatment group.

Any data outside the central laboratory reference ranges will be explicitly noted on the listings that are produced.

4.2.9.3 Electrocardiograms

A shift table will be produced for each ECG parameter to display normal, abnormal – not clinically significant, abnormal – clinically significant, and not done. The shift tables will present baseline and last observation post-baseline value, as applicable for each parameter.

4.2.9.4 Physical examination

Any new finding(s) or aggravated existing finding(s), judged as clinically significant by the Investigator will be reported as AEs.

4.2.9.5 Vital signs

Descriptive statistics and change from baseline for vital signs data will be presented for each treatment group by visit. Baseline to maximum post-baseline and baseline to minimum post-baseline value shift tables may be generated as applicable for each parameter and will include patients with both baseline and post-baseline data.

The normal reference ranges listed on Table 2 will be used for the identification of low, normal, and high values.

4.2.10 Pharmacokinetic analyses

The PK analyses will be performed at or under the guidance of AstraZeneca Research and Development.

Benralizumab serum concentrations and derived PK parameters will be summarized using descriptive statistics at each visit. In addition to the standard statistics, geometric mean and coefficient of variation (CV%) based on log-transformation will be presented for PK data. For the computation of geometric means, arithmetic means and medians, the BLQ data are set to LLOQ/2 before computation.

For non-compartment analysis (NCA), only patients with at least three evaluable serum PK observations post first dose (collected on day 3, day 7 and either day 14 or pre-dose on day 28) will be included.

When PK concentration data is missing on scheduled visit, data from unscheduled visits will be included based on the adjusted analysis defined visit window in Section 2.2.2; on dosing visits (ie, Baseline, Days 28 and 56), pre-dose values only are included. For subjects who prematurely discontinued study drug, PK data collected after the last dose date + 28 days (one dosing cycle) +3 days (visit window) will be excluded from PK analysis.

4.2.11 Immunogenicity analyses

Anti-drug antibody assessments will be conducted and analyzed as per the details in Section 3 of the "Statistical Analysis Plan for Benralizumab Anti-Drug Antibody data" with the following exceptions:

- Demographics/patient characteristics summary by ADA status and further by positive subgroup will not be produced.
- Summaries will not be produced by baseline blood eosinophils count cut off (≥300 or <300 cells/μL) due to all patients enrolled in this study having baseline blood eosinophils count ≥300 cells/μL.
- Summary of first positive post-baseline ADA response during the study will not be produced.
- Listing of patients who had ADA titres increased by >4-fold from onset to maximum titre will not be produced.
- Eosinophil levels will be summarized by visit for ADA positive and negative patients only.
- Summary of nAb will not be provided by subgroup of ADA positive.
- Relationship between ADA and efficacy will not be explored.
- Relative risk tables for select safety outcomes will be omitted due to low frequency of events.
- Listing of ADA titre by visit for ADA-positive patients will be produced.

5. INTERIM ANALYSES

No interim analysis is planned for this study.

6. CHANGES OF ANALYSIS FROM PROTOCOL

Protocol Section 8.5.2.3 mentions that ECG intervals will be summarized by presenting summary statistics of observed and change from baseline values. The (uncorrected) QT interval will be corrected according to the Fridericia's formula. However, this study collects only ECG outcome of the overall evaluation recorded as normal and abnormal; therefore, the summaries of observed and change from baseline values will not be performed.

7. REFERENCES

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

Bleeker et al 2016
Bleecker ER, FitzGerald JM, Chanez P, Papi A, Weinstein SF, Barker P, Sproule S, Gilmartin G, Aurivillius M, Werkström V, Goldman M; SIROCCO study investigators. Efficacy and safety of benralizumab for patients with severe asthma uncontrolled with high-dosage inhaled corticosteroids and long-acting β2-agonists (SIROCCO): a randomised, multicentre, placebo controlled phase 3 trial. Lancet. 2016;388:2115–27.

Castro et al 2014
Castro M, Wenzel SE, Bleecker ER, Pizzichini E, Kuna P, Busse WW, et al. Benralizumab, an anti-interleukin-5 receptor alpha monoclonal antibody, for uncontrolled eosinophilic asthma. Lancet Respir Med. 2014; 2:878-90.

Fitzgerald et al 2016
FitzGerald JM, Bleecker ER, Nair P, Korn S, Ohta K, Lommatzsch M, Ferguson GT, Busse WW, Barker P, Sproule S, Gilmartin G, Werkström V, Aurivillius M, Goldman M; CALIMA study investigators. Benralizumab, an anti–interleukin-5 receptor α monoclonal antibody, as add-on treatment for patients with severe, uncontrolled, eosinophilic asthma (CALIMA): a randomised, double-blind, placebo-controlled phase 3 trial. Lancet. 2016; 388:2128–41.

Jones et al 1991
Jones PW, Quirk FH, Baveystock CM. The St George's Respiratory Questionnaire. Respir Med. 1991; 85 Suppl B:25-31.

Jones et al 2009
Jones PW, Harding G, Berry P, Wiklund I, Chen W-H, Leidy NK. Development and firstvalidation of the COPD Assessment Test. Eur Respir J 2009, 34: 648–654.

Juniper et al 2005
Juniper EF, Svensson K, Mörk AC, Ståhl E. Measurement properties and interpretation of three shortened versions of the asthma control questionnaire. Respir Med. 2005 99:553–58.

Juniper et al 2006
Juniper EF, Bateman ED. The GOAL Committee. Identifying 'well-controlled' and 'not well-controlled' asthma using the Asthma Control Questionnaire. Respir Med. 2006 100:616–21.

Liu et al 2005
Liu, H, Wu T. Sample size calculation and power analysis of time-averaged difference. J Mod Appl Stat Methods. 2005 4(2): 434-45.

Gelman et al 2013
Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A., Rubin, D.B. (2013). Bayesian Data Analysis (Third Edition). Florida: CRC Press

Marostica et al 2013
Marostica, E., Yang, S.,Russu, A., De Nicolao, G., Zamuner, S., Beerahee, M. (2013). First-order longitudinal population model of FEV1 data: single-trial modeling and meta-analysis. Abstracts of the 22nd Annual Meeting of the Population Approach Group in Europe (PAGE)

8. APPENDIX

8.1 Accounting for missing data for pre-bronchodilator FEV_1

Patients dropping out of this study may lead to unobserved data. It is hoped the amount of missing data is minimized in this study as patients are encouraged to complete visits until they withdraw from the study.

This section summarizes how we will describe the pattern of and reasons for missing data from the study. It will also describe how we plan to account for missing data, including both the primary and sensitivity analyses to assess the robustness of the treatment effect under different underlying assumptions to account for missing data.

Missing data descriptions

Tabular summaries for the percentage of patients by the reason for discontinuation of randomized treatment as well as for withdrawal from the study will be presented by treatment to describe why patients discontinue from randomized treatment or withdraw from the study. The time to discontinuation of randomized treatment and withdrawal from the study will be presented using Kaplan Meier plots (overall and split by treatment related/not treatment

related reason for discontinuation, as defined in Tables 1 and 2). Dependent on these outputs additional exploratory analyses may be produced as deemed necessary to further understand the pattern of missing data.

Primary analysis under the Treatment Policy Estimand using the Missing at Random assumption

The primary analysis of the pre-BD FEV_1 is under the treatment policy estimand which allows for differences in outcomes over the entire study treatment period to reflect the effect of initially assigned randomized treatment as well as if subsequent treatments are taken. This primary analysis includes all data until patients withdraw from the study regardless of if they discontinue from randomised treatment. The Mixed Model Repeated Measures model (MMRM) used is a direct likelihood approach (DL) which is valid under the Missing at Random (MAR) assumption.

Sensitivity analysis under the Treatment Policy Estimand using both MAR and MNAR assumptions

To examine the sensitivity of the results of the primary analysis to departures from the underlying assumptions, sensitivity analyses of the repeated measures analyses will be performed for the pre-BD FEV_1 using controlled sequential multiple imputation methods based on pattern mixture models, as described in the European Medicines Agency Guideline on Missing Data in Confirmatory Clinical Trials (EMA 2010). As with the primary analysis, the sensitivity analyses include all data unitl patients withdrew from the study regardless of if patients discontinue from randomised treatment.

The imputation process consists of a sequence of multiple imputation (MI) steps, where each step is intended to impute missing values at 1 timepoint only. This model will assume that some pre-specified subset of patients who withdraw from the study have correlations with future (unobserved) visits similar to patients in the placebo arm. This allows us to assess various deviations from the MAR assumption.

The assumptions that will be used to impute the missing data who withdraw early are as follows:

(a) MAR: Assumes that the trajectory for patients who dropped out in each arm is similar to those observed in their own treatment arm

(b) DRMI: Assumes that the trajectory for patients in the benralizumab arms who dropped out for a treatment related reasons (according to the classification for DRMI) is similar to that of the placebo patients, whereas the remaining patients who has dropped out are imputed assuming MAR.

Approach b) can be considered more conservative than the approach for the primary analysis because the assumptions mean that as soon as patients withdraw for a treatment related reason, they begin to worsen immediately.

The MNAR imputation is achieved by using only appropriate data at each stage of the imputation. Imputation will be done in 2 steps, the non-monotone (intermediate) missing FEV_1 values will be imputed first (Markov chain Monte Carlo [MCMC] method is used to partially impute the data using SAS PROC MI) and then the missing value at each visit will be imputed using a sequential regression method (using MONOTONE REG option of SAS PROC MI).

For example, to impute missing values at time *t* for patients in the benralizumab arms who dropped out due to an AE, include only placebo observations up to and including time *t*, plus observations from patients in the benralizumab arms, that dropped out due to an AE, up to and including time *t*-1. This is carried out for each visit, 1 at a time using observed data, and missing values imputed up until that visit. Placebo missing observations and benralizumab observations that are not missing due to AEs are imputed assuming missing at random (MAR) and follow the pattern of observed placebo observations in each treatment arm, respectively. One-hundred imputations will be carried out. A seed of 784088 will be used for the monotone imputation step and a seed of 409345 will be used for the sequential regression imputation step. The analysis of each imputed dataset will be as described for the primary analysis in Section 4.2.5.1 and these will be combined using SAS procedure PROC MIANALYZE.

To avoid possible convergence issues when fitting MMRM models to 100 imputed datasets, the estimated unstructured covariance parameters from the first imputation where the model converges will be used as the starting values to fit the models for all imputed datasets.

A summary of reasons for patients withdrawing from the benralizumab treatment arm and the corresponding treatment arm used to calculate the imputation under MAR, and DRMI is given in Table 1.

Reason for withdrawal	**MAR**	**DRMI**
Adverse Event	Benralizumab	Placebo
Development of study-specific discontinuation criteria*	Benralizumab	Placebo
Death	Benralizumab	Placebo
Severe non-compliance to protocol	Benralizumab	Placebo
Eligibility criteria not fulfilled	Benralizumab	Benralizumab
Patient lost to follow up	Benralizumab	Benralizumab
Patient decision	Benralizumab	Based on review prior to study unblinding
Other	Benralizumab	Based on review prior to study unblinding

*Development of study-specific discontinuation criteria are based on the following: anaphylactic reaction to the IP requiring administration of epinephrine, development of helminth parasitic infestations requiring hospitalization, 2 consecutive doses of IP missed or more than 2 scheduled doses of IP are missed during course of the study, an asthma-related event requiring mechanical ventilation.

Together with the primary analysis, the sensitivity analyses are considered to cover the range from realistic to plausible worst case assumptions about missing data. The MAR multiple imputation approach is expected to correspond closely to the primary analysis, and is included to allow for comparisons with MNAR assumptions (specifically methods b) using the same multiple imputation methodology.

The dropout reason-based multiple imputation (DRMI) approach was selected as the most conservative approach based on the fact that placebo patients are receiving standard of care and are not expected to change to a substantially more effective treatment after withdrawing from study or study treatment. For patients receiving benralizumab who withdraw from the study due to treatment related reasons, it is assumed that, at worst, they would be on the standard of care treatment, ie, the placebo arm. For patients receiving benralizumab who withdraw from the study due to non-treatment related reasons, it seems reasonable to assume they would be similar to those patients who complete treatment.

Overall summary of analyses to account for missing data

A summary of the different analyses to be carried out under different estimands and assumptions are described in Table 3.

	Treatment Policy Estimand		
	DL	**MAR**	**DRMI**
Population	**On-treatment + post-discontinuation of randomized treatment**		
Estimand	**Treatment policy (MAR)**		**Treatment policy (MNAR)**
Imputation of pre-BD FEV1 in Benra arm	No explicit imputation*	Benra rate assumed for all reasons for withdrawal	Placebo rate assumed for AEs, Death, development of study specified reasons to stop active treatments and Severe non-compliance to protocol, otherwise Benra or based on review prior to study unblinding.

* Implicitly assumes unobserved data the same as observed

Forest plots will be used to show the primary and sensitivity analysis under the different estimands.

It is noted that if the primary analysis is statistically significant, it is not necessarily expected that all sensitivity analyses will also give statistically significant results. If the results of the sensitivity analyses provide reasonably similar estimates of the treatment effect to the primary analysis, this will be interpreted as providing assurance that neither the lost information nor the mechanisms which cause the data to be missing have an important effect on primary analysis conclusions. Based on these outputs and the drug's mechanism of action, the plausibility of the assumptions made about missing data in the different analyses will be considered and described in the clinical study report.

References

EMA 2010
European Medicines Agency. Committee for Proprietary Medicinal Products (CPMP). Guideline on Missing Data in Confirmatory Clinical Trials 2 July 2010
EMA/CPMP/EWP/1776/99 Rev. 1.

8.2 Partial dates for adverse events and prior/concomitant medication and medical history

Dates missing the day or both the day and month of the year will adhere to the following conventions in order to classify treatment-emergent AEs and to classify prior/concomitant medications:

Adverse events

- The missing day of onset of an AE will be set to:
  - First day of the month that the event occurred, if the onset YYYY-MM is after the YYYY-MM of first study treatment.
  - The day of the first study treatment, if the onset YYYY-MM is the same as YYYY-MM of the first study treatment.
  - The date of informed consent, if the onset YYYY-MM is before the YYYY-MM of the first treatment.
- The missing day of resolution of an AE will be set to:
  - The last day of the month of the occurrence. If the patient died in the same month, then set the imputed date as the death date.
- If the onset date of an AE is missing both the day and month, the onset date will be set to:
  - January 1 of the year of onset, if the onset year is after the year of the first study treatment.
  - The date of the first treatment, if the onset year is the same as the year of the first study treatment
  - The date of informed consent, if the onset year is before the year of the first treatment.
- If the resolution date of an AE or end date of a IP is missing both the day and month, the date will be set to:
  - December 31 of the year of occurrence. If the patient died in the same year, then set the imputed date as the death date.

Prior/concomitant medication, medical history

- The missing day of start date of a therapy (or symptom, diagnosis)will be set to the first day of the month that the event occurred.

- The missing day of end date of a therapy will be set to the last day of the month of the occurrence.

- If the start date of a therapy (or symptom, diagnosis) is missing both the day and month, the onset date will be set to January 1 of the year of onset.

- If the end date of a therapy is missing both the day and month, the date will be set to December 31 of the year of occurrence.

- If the start date of a therapy is null and the end date is not a complete date then the start date will be set to the date of the first study visit.

- If the start date of a therapy is null and the end date is a complete date

  – and the end date is after the date of the first study visit then the start date will be set to the date of the first study visit.

  – otherwise the start date will be set to the end date of the therapy.

- If the end date of a therapy is null and the start date is not a complete date then the end date will be set to the date of the study end date from the termination page If the end date of a therapy is null and the start date is a complete date

  – and the start date is prior to the date of the last study visit then the end date will be set to the date of the last study visit.

  – otherwise, the end date will be set to the start date of the therapy.

8.3 Important protocol deviations

As specified in Section 2.2.1, Important deviations will be those which are considered to potentially impact upon the interpretation of the primary endpoint in the study. The following IPDs will be summarized and listed in the CSR.

A. Eligibility deviations (patients incorrectly randomised and received IP)

1. No proof ICS/LABA used ≥ 30 days prior to Visit 1

or daily ICS dose < 500 µg powdered fluticasone propionate (or equivalent), or combined ICS/LABA use below locally approved maintenance dose.

2. Fewer than 2 documented asthma exacerbations in the 12 months prior to informed consent that required use of a systemic corticosteroid or temporary increase from the usual maintenance dose; re-screened patients based on physician's review

3. Pre-bronchodilator (pre-BD) FEV1 of ≥ 80% predicted at Visit 2 or Visit 3

4. Lack of evidence of asthma as documented by airway reversibility (FEV1 ≥12% and 200 mL) demonstrated at Visit 1, Visit 2 or Visit 3. For patients entering the body plethysmography sub-study, reversibility must be demonstrated at Visit 1 or at Visit 2 only.

5. Peripheral blood eosinophil count of < 300 cells/µL assessed by central lab at Visit 1

6. (substudy only) Residual volume < 125% of predicted at Visit 3

7. During the 7 days before randomization: < 2 days with a daytime or nighttime asthma symptoms; < 2 days of rescue SABA use; and no nights with awakenings due to asthma

8. Clinically important pulmonary disease (other than asthma), or diagnosed pulmonary or systemic disease (other than asthma) associated with elevated peripheral eosinophils.

9. Acute upper or lower respiratory infections requiring antibiotics or antiviral medication screening/run-in period, prior to randomization Visit 4

10. Upper respiratory tract infection or an asthma exacerbation that required treatment with systemic corticosteroids or an increase in regular maintenance dose of OCS during the screening/run-in period, prior to randomization Visit 4

11. Subject is a current smoker or former smoker with a smoking history of ≥ 10 pack-years

12. Subject previously randomised and dosed in the present study

13. Subject had greater than/equal to 20% change (or missing data for this assessment) in mean Pre BD FEV1 value (mean of the Pre-BD FEV1 taken 30 min (+/- 10 min) and 60 min

(+/- 10 min) prior to dosing) at randomization Visit 4 (Day 0) from the mean pre BD FEV1 calculated from the pre BD FEV1 recorded at Visit 2 and Visit 3.

B. Restricted and prohibited medications

1. Medium to high-dose systemic corticosteroids for non-asthma diagnosis as judged by Sponsor's medical review

C. Study treatment deviations

1 Patient received incorrect IP kit (active IP instead of placebo or placebo instead of active IP)

D. Study conduct

1. Changes in dose and regimen of asthma controller medications (ICS-LABA) or in dose and regimen of additional maintenance controllers done without medical need as judged by the Investigator
2. Regular LAMA- or LABA- containing maintenance therapy not withheld for 12-24h (twice daily scheme) or for >24h (once daily scheme) prior to scheduled lung function assessments based on physician's review
3. Administration of SABA within 6 hours prior to scheduled lung function assessments based on physician's review
4. Spirometry data missing for 2 or more visits between Visit 4 to Visit 7

E. Discontinuation and withdrawal

1. Subject met IP discontinuation criteria but was not discontinued from IP