Statistical Analysis Plan Amendment 1

Study ID: 209692

Official Title of Study: A randomised, double-blind, placebo controlled, parallel group Phase III study to assess the clinical efficacy and safety of 100 mg SC Mepolizumab in adults with chronic rhinosinusitis with nasal polyps (CRSwNP) / eosinophilic chronic rhinosinusitis (ECRS)

Date of Document: 31-MAR-2023

Information Type	:	Statistical Analysis Plan (SAP)

Title Page

Protocol Title: A randomised, double-blind, placebo controlled, parallel group Phase III study to assess the clinical efficacy and safety of 100 mg SC Mepolizumab in adults with chronic rhinosinusitis with nasal polyps (CRSwNP) / eosinophilic chronic rhinosinusitis (ECRS)

Study Number: 209692

Compound Number: SB240563

Abbreviated Title: Efficacy and safety of mepolizumab in adults with CRSwNP / ECRS

Acronym: MERIT: Mepolizumab in Eosinophilic chronic RhinosinusITis study

Sponsor Name:

GlaxoSmithKline Research & Development Limited

Regulatory Agency Identifier Number(s)

Registry	ID
ClinicalTrials.gov	NCT04607005

SAP Amendment Number: 1

TABLE OF CONTENTS

PAGE

Version history

Table 1 SAP Version History Summary

SAP Version	Document Date	Protocol Version (Date) on which SAP is based	Change	Rationale
SAP	03-FEB-2021	Protocol Amendment 1 (08-OCT-2020)	Not Applicable	Original version
SAP Amendment 1	31 Mar 2023	Protocol Amendment 3 (21-JUL-2022)	SAP document updated to change the primary analysis method to be a difference in means and clarify further intercurrent events Analysis of Time to 'Need for Surgery' changed to 'Need for Surgery' at Week 52. Other administrative	Protocol Amendment 3 included a change in approach to analysis for primary and secondary endpoint requiring amendment to SAP.. Protocol Amendment 3 included newly defined Intercurrent event of 2 or more missed consecutive doses and clarified COVID-19 related intercurrent events requiring update to SAP. Participants enter the trial with severe NP symptoms.

SAP Version	Document Date	Protocol Version (Date) on which SAP is based	Change	Rationale
			updates have also been made within this Amendment, including: - Change of reference from PSAP - COVID-19 SMQ has been removed	- Detail of all Adverse Events of Special Interest (AESI) and lab ranges of Potential Clinical Importance (PCI) are now detailed in the OPS. - COVID-19 SMQ is not required within the reporting of the study.

1. INTRODUCTION

The aim of Study 209692 is to assess the efficacy and safety of mepolizumab on top of standard of care (SoC) therapy in the treatment of chronic rhinosinusitis with nasal polyps (CRSwNP) / eosinophilic chronic rhinosinusitis (ECRS) for the purpose of registration in Japan and China.

The purpose of this SAP is to describe the planned analyses to be included in the Clinical Study Report (CSR) for Study 209692. Details of the planned final analyses are provided.

Descriptive study population analyses such as summary of demography and baseline characteristics and additional detail with regards to data handling conventions and the specification of data displays will be provided in the Output and Programming Specification (OPS) document.

1.1. Objectives, Estimands and Endpoints

1.1.1. Objectives and Endpoints

Objectives	Endpoints
Primary	
• To evaluate the efficacy of mepolizumab 100mg SC compared to placebo at Week 52 in patients with a diagnosis of CRSwNP / ECRS	The primary estimands are defined as follows: **Treatment Comparison:** Mepolizumab 100 mg SC compared to placebo **Population:** entire trial population of patients with a diagnosis of CRSwNP / ECRS randomised and receiving study treatment **Co-primary variables**: a) Change from baseline in total endoscopic NP score at Week 52 b) Change from baseline in mean nasal obstruction VAS score during the 4 weeks prior to Week 52 **Summary measure**: Difference in mean scores between mepolizumab and placebo **Main Intercurrent events (ICE) anticipated**: a) Premature discontinuation of study treatment unrelated to the COVID-19 pandemic – to be handled using a treatment policy strategy b) Changes in background medication or start of a prohibited medication unrelated to the COVID-19 pandemic (e.g. start INCS therapy where absent at

Objectives	Endpoints
	baseline) – to be handled using a treatment policy strategy c) Premature discontinuation of study treatment, change in background medication or start of prohibited medication related to the COVID-19 pandemic – to be handled using a hypothetical strategy d) Surgery, which includes any procedure involving instruments resulting in incision and removal of tissue from the nasal cavity (e.g. polypectomy/functional endoscopic sinus surgery [FESS]) – to be handled using a composite strategy by incorporating occurrence of the event into the definition of the endpoint. Specifically, participants who undergo surgery will be assigned the worst possible score for the endpoint from the day after surgery onwards for inclusion in the analysis e) Course of systemic CS for CRSwNP / ECRS – to be handled using a treatment policy strategy f) Interruption to investigational product of 2 or more consecutive doses – to be handled using a treatment policy
Secondary	
• To evaluate the impact on quality of life of 100mg mepolizumab compared to placebo at Week 52 in patients with a diagnosis of CRSwNP / ECRS	• Change from baseline in Sino-Nasal Outcomes Test-22 (SNOT-22) total score at Week 52
• To evaluate the efficacy of 100mg mepolizumab compared to placebo at Week 52 in terms of mean overall VAS symptom score, mean composite VAS score, Lund Mackay CT score, mean individual VAS symptom score for loss of smell and impact on time to first nasal surgery or course of systemic CS in patients	• Change from baseline in mean overall VAS symptom score during the 4 weeks prior to Week 52 • Change from baseline in the mean composite VAS score (combining VAS scores for nasal obstruction, nasal discharge, mucus in the throat and loss of smell) during the 4 weeks prior to Week 52 • Change from baseline in Lund Mackay CT score at Week 52 • Change from baseline in mean individual VAS symptom score for loss of smell during the 4 weeks prior to Week 52 • Time to first nasal surgery or course of systemic CS for CRSwNP / ECRS up to Week 52

Objectives	Endpoints
with a diagnosis of CRSwNP / ECRS	
Other	

CCI

[redacted]

Objectives	Endpoints
CCI [redacted]	

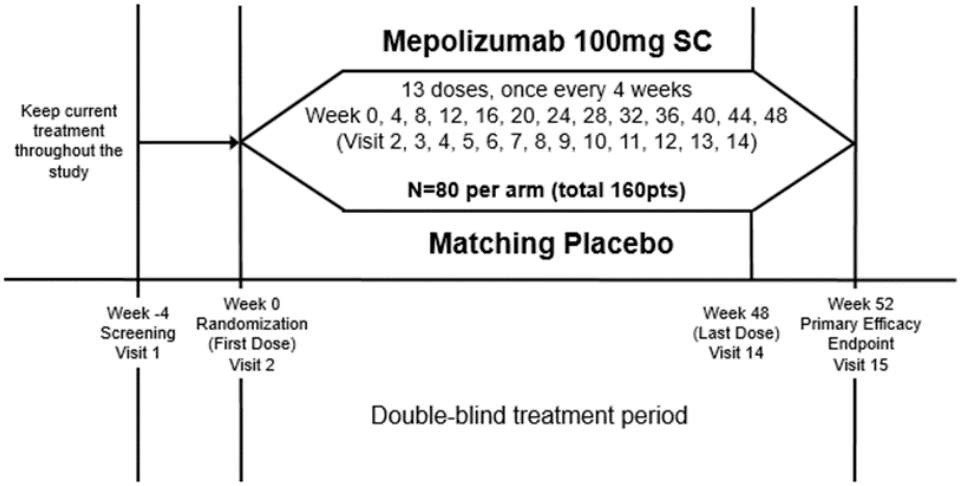
Objectives	Endpoints
CCI [redacted]	
Safety	
• To evaluate the safety and immunogenicity of 100mg mepolizumab compared placebo in patients with a diagnosis of CRSwNP / ECRS	• Frequency of Adverse events (AEs)/ Serious adverse events (SAEs) including systemic reactions and local injection site reactions reported • Vital signs (pulse rate, systolic and diastolic blood pressure) • Haematological and clinical chemistry parameters • 12 lead ECG derived endpoints • Presence of anti-mepolizumab antibodies and neutralising antibodies
Pharmacokinetics and pharmacodynamics	
• To evaluate the pharmacokinetics and pharmacodynamics of 100mg mepolizumab in a subgroup of participants from Japan and China with a diagnosis of CRSwNP / ECRS	• Plasma concentration of mepolizumab • PK/PD (blood eosinophil count) analysis

1.1.2. Estimands

The primary estimands are the difference between mepolizumab 100mg SC and placebo in a) mean change from baseline in total endoscopic NP score to Week 52 and b) mean change from baseline in mean nasal obstruction VAS score during the 4 weeks prior to Week 52 in participants with a diagnosis of CRSwNP / ECRS, regardless of IP discontinuation or changes in background medication/starting a prohibited medication unrelated to the COVID-19 pandemic, use of systemic CS for CRSwNP / ECRS or interruption of 2 or more consecutive doses of IP, with participants experiencing surgery being assigned the worst possible score from the day after surgery onwards.

Secondary estimands for SNOT-22, VAS scores and Lund Mackay CT score will use the same population, summary measure and strategies for intercurrent events as for the primary estimands. Time to first nasal surgery or course of systemic CS for CRSwNP / ECRS up to Week 52 will be summarised by the hazard ratio between mepolizumab and placebo. The same population and strategies for intercurrent events of treatment discontinuation, changes in background medication/starting a prohibited medication, interruptions of 2 or more consecutive doses of IP and COVID-19 pandemic related intercurrent events will be used as for the primary estimands. Surgery or a course of systemic CS for CRSwNP / ECRS will both be considered events within the analysis of this endpoint and therefore neither will be considered an intercurrent event. Supplementary estimands for the two co-primary endpoints will use the same population and strategies for intercurrent events of treatment discontinuation, changes in background medication/starting a prohibited medication, interruptions of 2 or more consecutive doses of IP and COVID-19 pandemic related intercurrent events as for the primary estimand. Within the first supplementary estimand participants experiencing surgery or receiving a course of systemic CS for CRSwNP / ECRS will be assigned the worst possible score. Within the second supplementary estimand participants experiencing surgery will be assigned the participant's own worst observed score prior to the surgery event, the intercurrent event of course of systemic CS for CRSwNP / ECRS will be handled using a treatment policy strategy as in the primary estimand.

1.2. Study Design

Overview of Study Design and Key Features	
	
Design Features	• This is a randomised, double-blind, placebo-controlled, parallel group, multicentre Phase III study designed to assess the clinical efficacy and safety of 100 mg SC Mepolizumab treatment in adults with CRSwNP / ECRS. • The study will include a 4-week run-in period followed by randomisation to a 52-week treatment period as a double-blind, placebo-controlled phase. Throughout the entire study period (run-in + treatment period), participants will be on the SoC for CRSwNP / ECRS. • A participant is considered to have completed study treatment if he/she receives study treatment at Week 48 (Visit 14). • A participant is considered to have completed the study if he/she has completed all phases of the study including the last visit (Visit 15 / Week 52). • Initiation or changes in the background SoC dosing regimen are not permitted from screening to end of the study. Participants are expected, if possible, to continue with these treatments with no interruption nor alteration throughout the study duration. • A participant may discontinue from study treatment at any time at his/her own request, or at the discretion of the investigator. Participants who discontinue from study treatment prematurely (for any reason) should, where possible, continue to be followed-up as per protocol until the completion of the Week 52 Visit assessments. • Participants that discontinue from study treatment should return to clinic 4 weeks after the last dose for an Investigational Product (IP) Discontinuation Visit. Participants that withdraw from the study should return to clinic 4 weeks after the previous visit for an Early Withdrawal Visit. • Participants are permitted nasal surgery during the study without discontinuation from study treatment. The participant's nasal polyps surgical status will be tracked for the duration of the study. Following discontinuation of study treatment, as a minimum, participants will be contacted by telephone to enquire regarding any safety assessments and any nasal polyps surgery events (received a nasal polyps surgical procedure or entry into a waiting list for nasal polyps surgery). • For the purpose of this study, nasal polyps surgery is defined as any procedure involving instruments with resulting incision (cutting open) and removal of the polyp tissue from the nasal cavity (polypectomy / functional endoscopic sinus surgery [FESS]). Any procedure involving instrumentation in the nasal cavity resulting in dilatation of the nasal passage such as balloon sinuplasty, insertion

Overview of Study Design and Key Features	
	of coated stents or direct injection of steroids or other medication without any removal of nasal polyps tissue does not fulfil this criterion.
Study Treatment	• The treatment period will consist of thirteen doses of mepolizumab or matched placebo, delivered by a pre-filled safety syringe device (SSD) subcutaneous injection every 4 weeks for 52 weeks in addition to SoC therapy. The last dose of interventional product is to be administered at Week 48 (Visit 14). • For non-Japanese participants, study treatment must be administered by a healthcare professional (HCP) until week 52 (last dose at Week 48). • Japanese participants who are willing can self-administer study treatment under observation of the principal investigator or delegate from Week 32 onwards. Before allowing the self-administration, the principal investigator or delegate must assess the participant's knowledge/technique and document that they are competent to undertake self-administration. All self-administered injections will be assessed for success based on direct observation. • No dose modifications or dose reductions are permitted within the study protocol. • Participants meeting a treatment discontinuation criteria (which include pregnancy, liver stopping criteria or ECG stopping criteria as defined in Protocol Section 7.1.2) must have study treatment discontinued. Restart, re-challenge or resumption of study treatment following a stopping criteria being met is not permitted.
Study Treatment Assignment	• Approximately 160 participants to be randomised in a 1:1 ratio to mepolizumab or placebo treatment in accordance with the randomisation schedule (approximately 80 participants per arm). • The randomisation schedule will be generated using the GSK validated randomisation software RandAll NG. Eligible participants will be assigned to study treatment randomly using RAMOS NG, an Interactive Web Response System (IWRS). • The study will be randomised separately for each country and the randomisation will be stratified by background intranasal corticosteroids (INCS) use. Equal numbers of participants will be allocated to each treatment.
Interim Analysis	• No interim analysis of data is planned for this study.

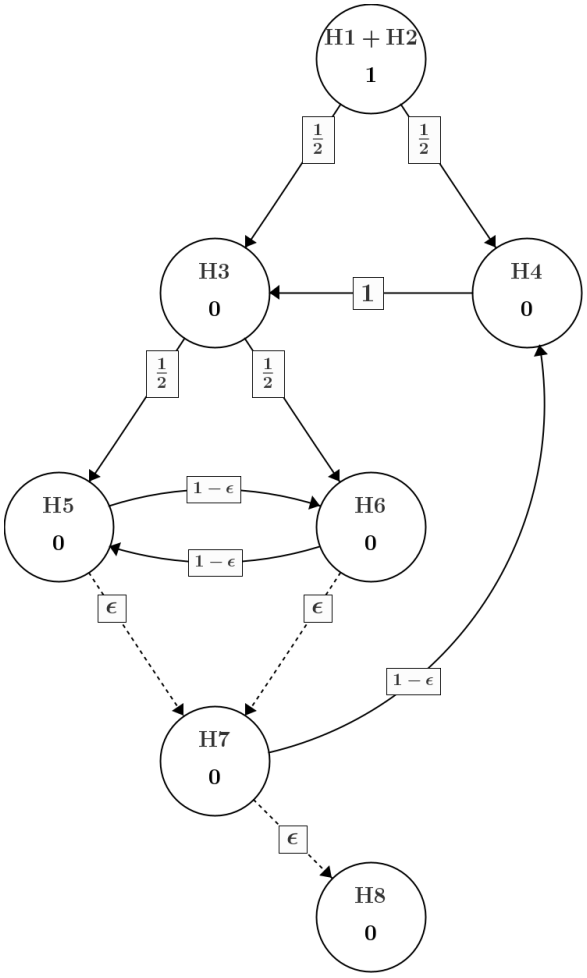
2. STATISTICAL HYPOTHESES

The study is designed to test the superiority of mepolizumab 100mg SC vs. placebo in both co-primary efficacy endpoints of total endoscopic nasal polyp score at Week 52 and nasal obstruction VAS score during the 4 weeks prior to Week 52. Each co-primary endpoint will be tested at the two-sided 5% alpha level (one-sided 2.5%) and both tests are required to be significant to achieve the primary objective of the study.

2.1. Multiplicity Adjustment

Multiplicity arising from the multiple secondary endpoints will be controlled using the closed testing procedure specified below for strong control of type I error. Hypotheses associated with the two primary (H_1 and H_2) and six secondary endpoints (H_3 to H_8) will be tested using a gatekeeping procedure based on the graphical approach to sequentially rejective multiple test procedures [Bretz, 2009]. The two primary hypotheses H_1 and H_2 will each be tested first and will be allocated the level α, where α = 0.05 (two-sided). If these are both rejected at level α, the procedure then is as follows: the hypotheses H_i , i = 3, ..., 8 are tested each at its local significance level α_i . If a hypothesis H_i can be rejected, its level is reallocated to one of the other hypotheses according to the pre-specified rules represented by Figure 1. The reallocation weights are updated in the reduced graph and the testing step repeated for the remaining, non-rejected hypotheses with the updated local significance levels. This possibly leads to further rejected null hypotheses with associated reallocation of the local significance levels. The procedure is repeated until no further hypothesis can be rejected. The reallocation of the local alpha levels is fully determined by the initial graph (Figure 1) and the update algorithm described by [Bretz, 2009].

Figure 1 Multiplicity Testing Strategy Across Primary and Secondary Endpoints



Primary:
H_1: Total Endoscopic Nasal Polyps score at Week 52
H_2: Nasal Obstruction VAS score during the 4 weeks prior to Week 52
Secondary:
H_3: Overall VAS score during the 4 weeks prior to Week 52
H_4: Lund Mackay CT score at Week 52
H_5: Composite VAS score of Nasal Symptoms during the 4 weeks prior to Week 52
H_6: SNOT-22 Total score at Week 52
H_7: Loss of smell VAS score during the 4 weeks prior to Week 52
H_8: Time to First Nasal Surgery or course of systemic CS for CRSwNP / ECRS up to Week 52
where ε reflects an infinitesimally small value, indicating the potential for alpha to be reallocated dependent on the rejection of all previous tests

For the initial graph (Figure 1) the resulting test procedure can be summarized as follows: If both co-primary endpoints (H_1 and H_2) are rejected, the first two secondary endpoints (Overall VAS Symptom Score [H_3] and Lund Mackay CT score [H_4]) will each be tested at a significance level of α/2 (0.025). If either endpoint is rejected at the α/2 (0.025) level, the α/2 (0.025) from that given endpoint will be reallocated and re-used within the subsequent test of the next secondary endpoint according to the pre-specified rules represented by Figure 1.

- If the Overall VAS score endpoint (H_3) endpoint is rejected at the α/2 (0.025) level, the Composite VAS score of Nasal Symptoms (H_5) and SNOT-22 total score (H_6) will then also be tested. Only following rejection of both endpoints (H_5 and H_6) will the Loss of Smell VAS symptom endpoint (H_7) be tested.
- If able to reject the loss of smell VAS endpoint (H_7), the α/2 level from this endpoint will be reallocated to the testing of the Lund Mackay CT score (H_4), to allow a test at a full level of α, if not already rejected at the α/2 significance level.
- Only following rejection of all other secondary endpoints (H_3 to H_7), will the final secondary endpoint of Time to First Nasal Surgery or course of systemic CS for CRSwNP / ECRS (H_8) be tested.

Lund Mackay CT score (H_4) is an important endpoint in the clinical practice of Japan and China, however, has not been studied previously within the mepolizumab clinical development program. The above testing strategy demonstrates the importance of this endpoint, whilst permitting testing of other secondary endpoints if not initially rejected at a significance level of α/2. The Composite VAS of Nasal Symptoms, SNOT-22 and Loss of Smell VAS secondary endpoints are all correlated patient reported endpoints measuring improvements in symptoms; these endpoints are grouped within the above strategy for testing in sequence following the rejection of the of Overall VAS Score (H_3) to increase power across these symptom endpoints. If all symptom endpoints (H_3, H_5, H_6, H_7) are rejected, the above strategy permits a test of the Lund Mackay CT score (H_4) a full level of α, if not already rejected at the initial α/2 significance level.

For secondary endpoints unadjusted p-values and p-values adjusted for multiplicity, based on the above hierarchy, will be presented. Adjusted p-values will be calculated using the gMCP package within R ™.

3. ANALYSIS SETS

Analysis Set	Definition / Criteria	Analyses Evaluated
All Participants Enrolled	• All participants enrolled and for whom a record exists on the study database	• Study Population
Randomised	• All participants who were randomly assigned to study treatment in the study • Any participant who receives a treatment randomisation number will be considered to have been randomised	• Study Population
Intent-to-Treat (ITT)	• All randomised participants who receive at least 1 dose of study treatment • Participants will be analysed according to the treatment they are allocated at randomisation	• Efficacy
Safety	• All randomised participants who receive at least 1 dose of study treatment • Participants will be analysed according to the treatment they actually received for more than 50% of treatment administrations	• Safety
Pharmacokinetic (PK)	• All participants in the Safety analysis set who have at least 1 non-missing PK assessment (Non-quantifiable [NQ] values will be considered as non-missing values). • Participants will be analysed according to the treatment they actually received for more than 50% of treatment administrations	• PK

Note: "Enrolled" means a participant's, or their legally acceptable representative's (LAR's), agreement to participate in a clinical study following completion of the informed consent process.

4. STATISTICAL ANALYSES

4.1. General Considerations

4.1.1. General Methodology

The Intent-to-Treat (ITT) analysis set will be used for all Study Population analyses and Efficacy analyses, and the Safety analysis set will be used for all Safety analyses.

In the case of wrong stratification assigned at the time of randomisation, the analyses will be performed based on the data collected in the eCRF, not the assigned stratum at randomisation.

Confidence intervals will use 95% confidence levels unless otherwise specified.

Unless otherwise specified, continuous data will be summarized using descriptive statistics: n, mean, standard deviation (SD), median, minimum and maximum. Categorical data will be summarized as the number and percentage of participants in each category.

4.1.2. Baseline Definition

4.1.2.1. Visit-Based Endpoints

For all visit-based endpoints, including the co-primary endpoint of endoscopic NP score, Baseline will be defined as the latest non-missing measurement collected prior to the first dose of study treatment. This will generally be from the Day 1 (Visit 2) assessment but may be from an unscheduled assessment. Where time is collected, an assessment is valid for use as Baseline if the date and time is prior to or the same as the date and time of first dose. Where time is not collected, an assessment is valid for use as Baseline if the date of assessment is on or before the date of first dose.

Unless otherwise stated, if baseline data is missing no derivation will be performed and baseline will be set to missing.

4.1.2.2. eDiary Data

Participants will report daily on the severity of individual and overall NP symptoms using an eDiary; this includes information for the co-primary endpoint of mean nasal obstruction VAS score. The baseline value for endpoints reported in the eDiary will be determined as the average score from the 7 days of diary data collected prior to Day 1 (using non-missing data recorded between Day -7 and Day -1 inclusive). Participants are required to have at least 4 days of non-missing diary data available within the baseline period (7 days prior to Day 1) to meet the necessary study randomisation criteria.

4.1.3. Multicentre Studies

In this multicentre global study, the number of enrolled participants and the number randomised to each treatment arm will be presented by centre and country. Separate subgroup analyses will be performed by country (as detailed in Section 4.11.1).

The study will be randomised separately for each country; therefore country will be included as a fixed effect covariate in the analysis models for all efficacy endpoints. Small numbers of participants in a treatment group within a country may result in model

convergence issues, in these cases further combining of countries may be considered or exclusion of the country covariate if required.

4.2. Co-primary Endpoints Analyses

4.2.1. Definition of Endpoint(s)

The co-primary endpoints for analysis are:

- Change from baseline in total endoscopic nasal polyp score at Week 52 (based on centrally read data)
- Change from baseline in mean nasal obstruction VAS score during the 4 weeks prior to Week 52

4.2.1.1. Total Endoscopic NP Score

Total endoscopic nasal polyp score is collected at clinical visits. Independent reviewers, blinded to treatment, grade the total endoscopic nasal polyp score from image recording of endoscopies. The total score is reported as the sum of the right and left nostril scores and ranges from 0 to 8, with higher scores indicating greater disease severity.

The primary assessment will be at Week 52. Baseline will be defined as the latest non-missing measurement collected prior to the first dose of study treatment (Section 4.1.2.1).

4.2.1.2. Mean Individual VAS Symptom Scores

Every day each participant will be asked to indicate on a visual analogue scale (VAS) the severity of six nasal polyposis symptoms (one VAS for each symptom and one for symptoms overall) at its worst over the previous 24 hours. These VAS scores will include: (1) CCI [redacted]; (2) CCI [redacted]; (3) CCI [redacted]; (4) CCI [redacted] (5) CCI [redacted] and (6) CCI [redacted] VAS score.

Participants will rate symptoms daily throughout the study on a VAS using an eDiary on a line with 101 individually selectable points ranging from 0 CCI [redacted] to 100 CCI [redacted]). The final VAS scores for inclusion in summary and analysis tables will be derived from the electronically captured score by dividing each score by 10, and therefore will range from 0 to 10 (reported to 1 decimal place).

Symptoms will be calculated as the mean of all relevant VAS score measurements during the 4 weeks prior to Week 52. Within the 4-week period each participant must have at least 4 days of non-missing diary data available within each given week (7 day) period for a week to be evaluable and must have at least 3 weeks of evaluable data for the 4-week period average to be derived. The mean VAS score over the last 7 days before the first dose of study treatment will be used to determine the baseline value of each VAS score (Section 4.1.2.2).

The individual VAS symptom score of nasal obstruction will be analysed as the co-primary endpoint in this study.

4.2.2. Main Analytical Approach

The treatment effect to be estimated will be the comparison of mepolizumab 100 mg SC + SoC to placebo + SoC. The population of interest will be the entire trial population (patients meeting the study inclusion/exclusion criteria with a diagnosis of CRSwNP / ECRS) who are randomised and receiving study treatment (ITT analysis set).

The summary measure of treatment effect for the primary estimand of both co-primary endpoints will be the difference in mean scores between mepolizumab and placebo.

The main intercurrent events (ICEs) anticipated which may affect subsequent scores for the co-primary endpoints are:

a) Premature discontinuation of study treatment unrelated to COVID-19 pandemic– to be handled using a treatment policy strategy. The study will continue collecting data for participants who prematurely discontinue from randomised treatment. Off-treatment data collected for these participants following treatment discontinuation will be included in the analysis.
b) Changes in background medication or to start a prohibited medication (e.g. start INCS therapy where absent at baseline) unrelated to the COVID-19 pandemic – to be handled using a treatment policy strategy. These events will be captured as a protocol deviations and the study will continue treatment and to collect data for participants following this change in background medication / starting of a prohibited medication. Participants will not be expected to discontinue treatment or withdraw from the study.
c) Premature discontinuation of study treatment or change in background medications or start of prohibited medication related to the COVID-19 pandemic (due to quarantines, site closures or other related issues) – to be handled using a hypothetical strategy. Data impacted by COVID-19 pandemic related events will not be included in the statistical analyses and will be treated as missing data.
d) Surgery, which includes any procedure involving instruments resulting in incision and removal of tissue from the nasal cavity (e.g. polypectomy/FESS) – to be handled using a composite strategy by incorporating occurrence of the event into the definition of the endpoint. Specifically, participants who undergo surgery will be assigned the worst possible score for inclusion in the analysis of all subsequent visits (or 4-week periods) following surgery. Assessments taking place on the same day as surgery are assumed to have been carried out before the surgery.
e) Course of systemic CS for CRSwNP / ECRS – to be handled using a treatment policy strategy. Data collected following a course of systemic CS for CRSwNP / ECRS will be included in the analysis where systemic CS will be considered as SoC treatment.
f) Interruption to investigational product of 2 or more consecutive doses – to be handled using a treatment policy. These events will be recorded as protocol deviations and participants are expected to continue investigational product per-protocol after interruption. Data collected will be included in the analysis regardless of IP interruption.

This analysis will utilise the change from baseline in each of the co-primary endpoints where data for participants with surgery (prior to visit) are based on the worst possible score from the day after surgery onwards.

There are two expected sources of missing data namely, participants who withdrew early from study and participants who experienced an intercurrent event due to COVID-19 pandemic related restrictions. For the primary analysis, all missing data is assumed to be missing at random (MAR).

Analysis will be carried out using a mixed model repeated measures analysis with covariates of treatment group, baseline value, log of baseline blood eosinophil count, background INCS use (actual strata based on information related to INCS use recorded on eCRF), country and time point, plus interaction terms for time point by baseline value and time point by treatment group. Models will be checked for issues with convergence.

An unstructured covariance structure for the R matrix will be used by specifying 'type=UN' on the REPEATED line. The Kenward and Roger method for approximating the denominator degrees of freedom and correcting for bias in the estimated variance-covariance of the fixed effects will be used. The estimated means for each treatment group will be calculated from the model using the observed marginal distributions of the sample covariates (OM option in LSMEANS statement).Model-estimates of the absolute mean and mean change from baseline will be presented with corresponding standard errors for each treatment group by visit. Estimated treatment differences (Mepolizumab - Placebo) with corresponding 95% CIs and p-values will also be presented.

A cumulative distribution function (CDF) plot will also be provided for the change from baseline in both co-primary endpoints by treatment group: (1) endoscopic nasal polyps score at Week 52 and (2) mean nasal obstruction VAS score during the 4 weeks prior to Week 52.

When providing descriptive summaries of the co-primary endpoint data:

1) The diary scores for each 4-week period will be calculated as the mean of the scores reported between timepoints (e.g. Weeks 1-4, Weeks 5-8, Weeks 9-12, etc.). The 4-week period average will be derived as defined within Section 4.2.1.2. Further details of these 4-weekly periods will be defined within the OPS.
2) Participants with the intercurrent event of nasal surgery (as discussed above) will be assigned the worst possible score from the day after surgery onwards.

4.2.3. Sensitivity Analyses

For the primary estimand missing data for any reason is considered to be missing at random. A sensitivity analysis will be carried out whereby missing data due to premature withdrawal from the study will be imputed using off-treatment data and missing data due to occurrence of intercurrent events related to COVID-19 will be imputed as missing at random. An additional strategy will be implemented for missing data from participants who prematurely withdrew from the study. Off-treatment imputation assumes that future outcomes for those who withdraw can be predicted from a combination of participant characteristics, the participant's past observations and the patterns of response of participants who withdrew from investigational product but continued in the trial off-treatment. Off-treatment is defined as a visit occurring over 28 days after the last dose of interventional product.

For intercurrent events related to the COVID-19 pandemic, missing data is assumed MAR and will be imputed MAR. The imputation assumes that future missing outcomes for those who withdraw can be predicted from a combination of participant characteristics, the participant's past observations and the patterns of response of participants who remain in the trial without any intercurrent events related to COVID-19 (i.e. conditional on the data observed for each participant, their unavailable data are randomly missing). This corresponds to a standard likelihood analysis using MMRM.

Both imputation strategies will be implemented using multiple imputation.

Stepwise imputation methods will be implemented where imputations for a visit are conditioned on data from previous visits. Missing values will be imputed sequentially at post-baseline visits *1,…,T*. At a given visit *t* the imputation regression model will include both observed and imputed outcome values from previous visits *1,…,t-1*, baseline score, log of baseline blood eosinophil count, background INCS use, country and treatment at visit *t*. Both observed and imputed outcome values from previous visits will be used in the regression model.

The treatment covariate at visit *t* will be equal to the randomised treatment if the participant was still on-treatment or a generic 'off-treatment' if the participant had already discontinued from treatment. For participants who had already discontinued from the study at visit *t*, missing values will be imputed based on the regression model and assuming those observations are off-treatment.

Analysis will be performed using mixed model repeated measures as described for the primary analysis. For off-treatment imputation, the results will be combined across at least 2000 imputations using Rubin's rules [Rubin, 1987].

4.2.4. Supplementary Estimands

4.2.4.1. Course of Systemic Corticosteroids for CRSwNP / ECRS Handled as Failure

Supplementary estimands for the two co-primary endpoints will be performed and analysed as detailed within Section 4.2.2, with the summary measure of treatment effect as the difference in mean scores between mepolizumab and placebo.

The main intercurrent events (ICEs) anticipated and how they are to be handled are identical to the co-primary estimands defined within Section 4.2.2, with the exception of:

- Course of systemic CS for CRSwNP / ECRS – to be handled using a composite strategy by incorporating occurrence of the event into the definition of the endpoint. Specifically, participants who receive a course of systemic CS for CRSwNP / ECRS will be handled as a failure and assigned the worst possible score for inclusion in the analysis of all subsequent visits (or 4-week periods) following the initiation of a course of systemic CS. Assessments taking place on the same day as the initiation of a course of systemic CS are assumed to have been carried out before the CS medication.

4.2.4.2. Worst Observed Score Prior to Surgery

A second supplementary estimand for the two co-primary endpoints will be performed and analysed as detailed within Section 4.2.2, with the summary measure of treatment effect as the difference in mean scores between mepolizumab and placebo.

The main intercurrent events (ICEs) anticipated and how they are handled are identical to the co-primary estimands defined within Section 4.2.2, with the exception of:

- Surgery, which includes any procedure involving instruments resulting in incision and removal of tissue from the nasal cavity (e.g. polypectomy/FESS) – to be handled using a composite strategy by incorporating occurrence of the event into the definition of the endpoint. Specifically, participants who undergo surgery will be assigned the participant's worst observed score prior to the surgery event for inclusion in the analysis of all subsequent visits (or 4-week periods) following surgery. Assessments taking place on the same day as surgery are assumed to have been carried out before the surgery.

4.3. Secondary Endpoint(s) Analyses

4.3.1. Definition of Endpoint(s)

The secondary endpoints for analysis are:

- Change from baseline in SNOT-22 total score at Week 52
- Change from baseline in mean overall VAS symptom score during the 4 weeks prior to Week 52
- Change from baseline in mean composite VAS score (combining VAS scores for nasal obstruction, nasal discharge, mucus in the throat and loss of smell) during the 4 weeks prior to Week 52
- Change from baseline in mean VAS for loss of smell score during the 4 weeks prior to Week 52
- Change from baseline in Lund Mackay CT score at Week 52
- Time to first nasal surgery or course of systemic CS for CRSwNP / ECRS up to Week 52

4.3.1.1. SNOT-22 Total Score

The 22 questions of the Sino-Nasal Outcomes Test-22 (SNOT-22) questionnaire are each graded on a 6-point scale ranging from 0 = '[redacted]' to 5 = '[redacted]'. The scores for each of the questions are summed to derive the total score for that participant at that visit, which ranges from 0 to 110 with higher scores representing worse quality of life [Hopkins, 2009].

The primary assessment will be at Week 52. Baseline will be defined as the latest non-missing measurement collected prior to the first dose of study treatment (Section 4.1.2.1).

4.3.1.2. Mean Individual, Overall and Composite VAS Symptom Scores

The collection of six nasal polyposis symptoms VAS scores (one VAS for each symptom and for symptoms overall) are described within Section 4.2.1.2.

Two separate composite VAS scores will be derived using the individual symptom VAS scores. One composite VAS score will be derived using the individual VAS scores of symptoms of (1) [redacted], (2) [redacted], (3) [redacted] and (4) [redacted]:

Composite VAS score of Nasal Symptoms =
Sum of the four individual symptoms VAS scores / 4

A second composite VAS score will be derived using the individual VAS scores of symptoms of (1) [redacted], (2) [redacted], (3) [redacted], (4) [redacted] and (5) [redacted]:

Composite VAS score of Nasal Symptoms and Facial Pain =
Sum of the five individual symptoms VAS scores / 5

The composite VAS score of Nasal Symptoms, overall VAS symptom score and individual VAS symptom score of loss of smell will be analysed as secondary endpoints in this study.

4.3.1.3. Lund Mackay CT Score

The Lund Mackay CT scoring system is based on localization with points given for degree of opacification: 0 = [redacted], 1 = [redacted], 2 = [redacted]. These points are then applied to the maxillary, anterior ethmoid, posterior ethmoid, sphenoid, frontal sinus on each side. The osteomeatal complex (OC) is graded as 0 = [redacted], or 2 = [redacted]. This scoring system has been validated in several studies [Bhattacharyya, 1999; Lund, 1993]. For patients in whom the OC is missing (because of a previous surgery) the reader should consider the location of the former OC and provide a scoring (as if the OC was there).

A CT scan should be performed anytime during the run-in, and at Week 52 and IP discontinuation/Early Withdrawal Visit. CT scans central reading for Lund Mackay CT score will be used in the statistical analysis. The primary assessment will be at Week 52. Baseline will be defined as the latest non-missing measurement collected prior to the first dose of study treatment (Section 4.1.2.1).

4.3.1.4. Time to First Nasal Surgery or Course of Systemic CS for CRSwNP / ECRS

Nasal polyps surgery is defined as any procedure involving instruments with resulting incision (cutting open) and removal of the polyp tissue from the nasal cavity (polypectomy/FESS). Any procedure involving instrumentation in the nasal cavity resulting in dilatation of the nasal passage such as balloon sinuplasty, insertion of coated stents or direct injection of steroids or other medication without any removal of nasal polyps tissue does not fulfil this surgery criterion. A clinical review will be carried out prior to the unblinding of treatment codes to identify all events to be considered as part of this evaluation.

Courses of systemic CS for CRSwNP / ECRS will be collected as a concomitant medication of systemic corticosteroids (taken via the following routes: oral, intravenous or intramuscular) with a collected reason for concomitant medication of ‘Nasal Polyps’.

The time to first nasal surgery or course of systemic CS for CRSwNP / ECRS will be determined as the number of days from the date of first dose of study medication to the date of first event (Nasal Surgery or Course of Systemic CS for CRSwNP / ECRS, whichever is sooner). For participants who complete to Week 52 without experiencing nasal surgery or a course of systemic CS for CRSwNP / ECRS, event times will be censored at the time of the Week 52 visit. If a participant withdraws from the study before Week 52 and before experiencing nasal surgery or having a course of systemic CS for CRSwNP / ECRS, the event time will be censored at the time of study withdrawal.

4.3.2. Main Analytical Approach

The following secondary endpoints will be analysed as detailed for the two co-primary endpoints within Section 4.2.2:

- Change from baseline in SNOT-22 total score at Week 52
- Change from baseline in mean overall VAS symptom score during the 4 weeks prior to Week 52

- Change from baseline in mean composite VAS score (combining VAS scores for nasal obstruction, nasal discharge, mucus in the throat and loss of smell) during the 4 weeks prior to Week 52
- Change from baseline in mean VAS for loss of smell score during the 4 weeks prior to Week 52
- Change from baseline in Lund Mackay CT score at Week 52 (based on ANCOVA analysis)

4.3.2.1. Time to First Nasal Surgery or Course of Systemic CS for CRSwNP / ECRS

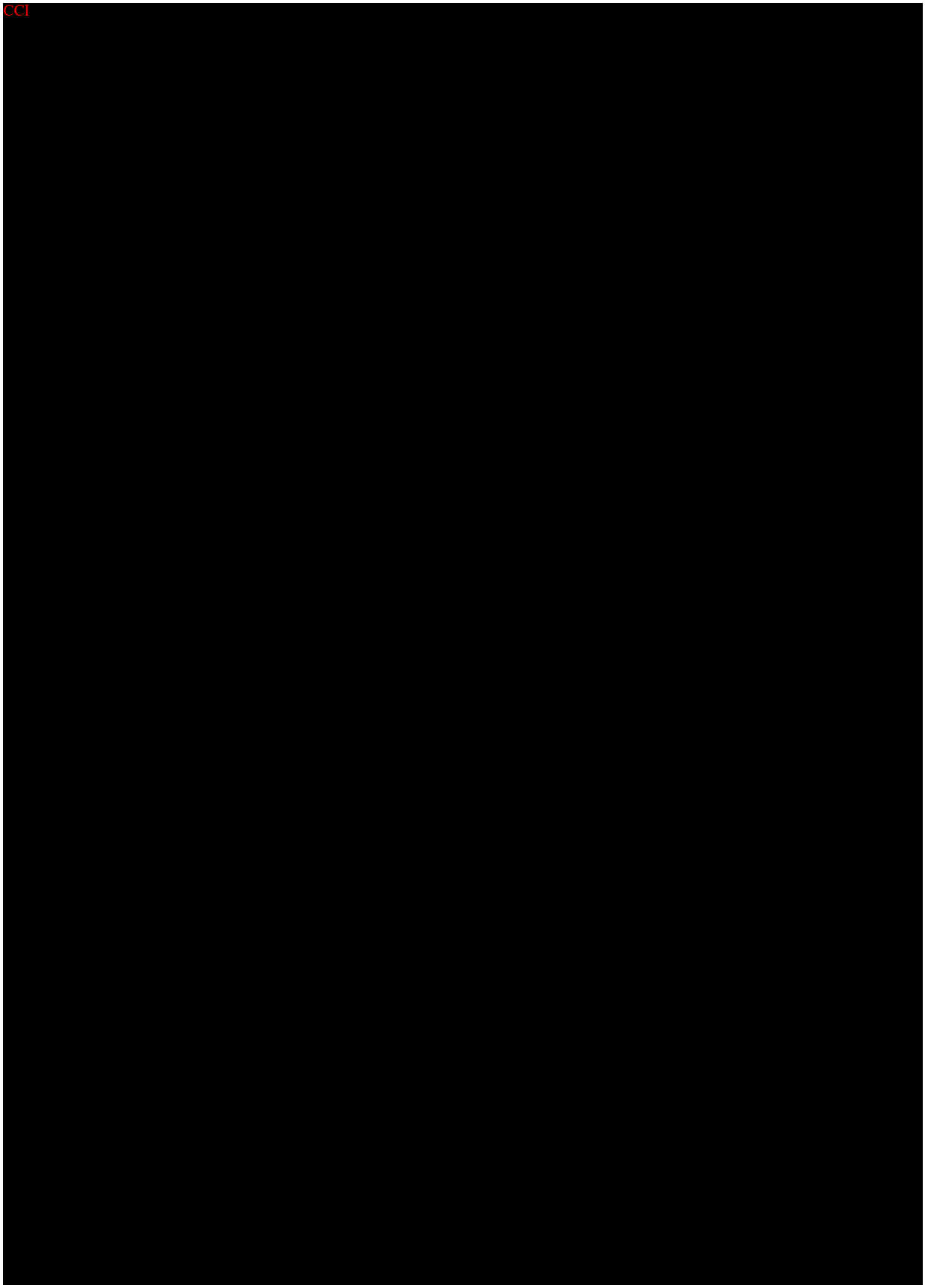
For the endpoint of time to first nasal surgery or course of systemic CS for CRSwNP / ECRS up to Week 52, the treatment effect to be estimated will be the comparison of mepolizumab 100 mg SC + SoC to placebo + SoC. The population of interest will be the entire trial population (patients meeting the study inclusion/exclusion criteria with a diagnosis of CRSwNP / ECRS) who are randomised and receiving study treatment (ITT analysis set) and the summary measure of treatment effect will be the hazard ratio (mepolizumab / placebo). The hazard ratio represents the relative probability of nasal surgery or course of systemic steroids for CRSwNP / ECRS at any time during the study.

The anticipated intercurrent events of premature discontinuation of study treatment unrelated to the COVID-19 pandemic, changes in background medication/starting of a prohibited medication unrelated to the COVID-19 pandemic or interruptions of 2 or more consecutive doses to investigational product will be handled using a treatment policy strategy, such that available event times will be included in the analysis regardless of whether the event (surgery or course of systemic CS for CRSwNP / ECRS) occurred before or after discontinuation of study treatment or changes in background medication/starting of a prohibited medication. Data occurring after intercurrent events of premature discontinuation of study treatment or changes in background medication/starting of a prohibited medication related to the COVID-19 pandemic will not be included in the statistical analyses and each participant will be censored at the time of the COVID-19 pandemic related intercurrent event. If a participant withdraws from the study before experiencing nasal surgery or having a course of systemic CS for CRSwNP / ECRS, the event time will be censored at the time of study withdrawal.

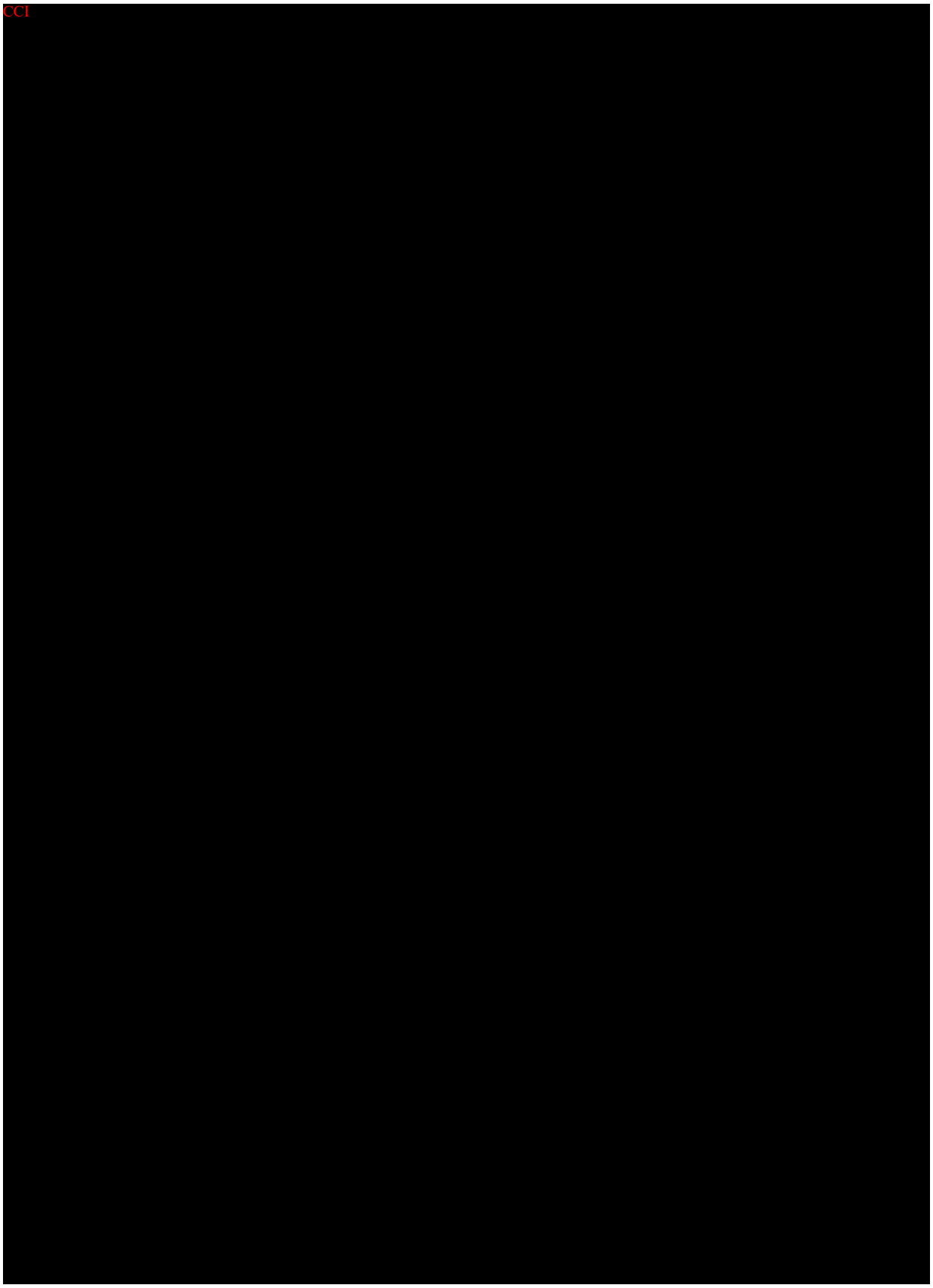
Statistical analysis will use the Cox proportional hazards model with covariates of treatment group, log of baseline blood eosinophil count, number of previous surgeries (0, 1, 2, >2; ordinal), background INCS use and country. A log(-log(survival)) plot by treatment group will be used to assess the assumption of proportional hazards.

The number of participants with an event, the number censored due to study withdrawal , due to occurrence of COVID-19 related intercurrent event, and at study completion will be presented by treatment arm. The hazard ratio with corresponding 95% CI and p-value from the Cox proportional hazards model will be presented. Additionally, a summary and graph of the Kaplan-Meier estimates of the cumulative proportion of participants with nasal surgery or a course of systemic CS for CRSwNP / ECRS within each treatment arm over time will be produced.

CCI



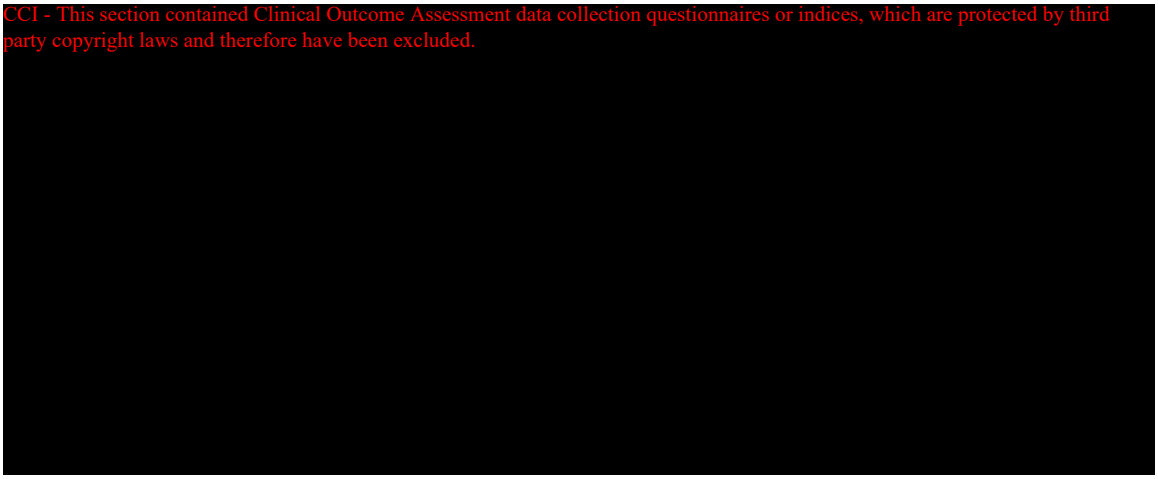
CCI



CCI

CCI

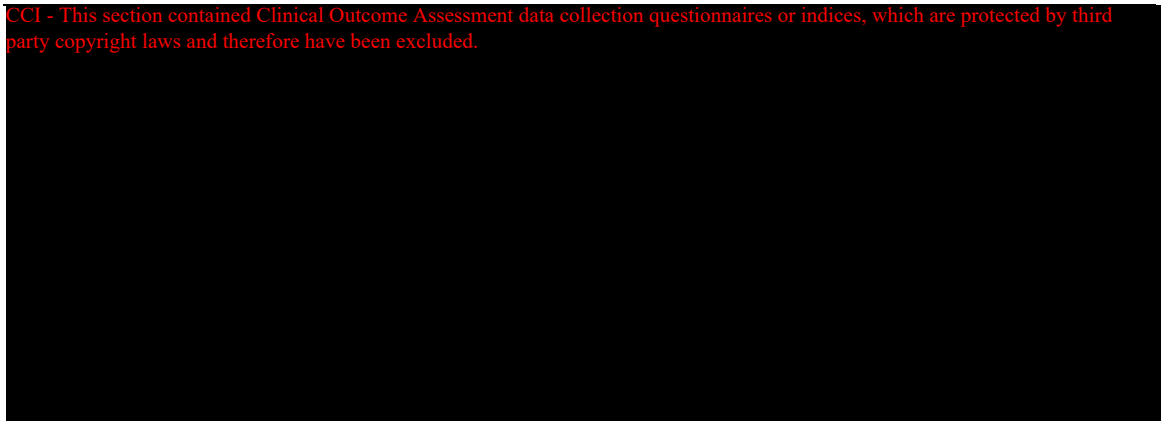
CCI - This section contained Clinical Outcome Assessment data collection questionnaires or indices, which are protected by third party copyright laws and therefore have been excluded.



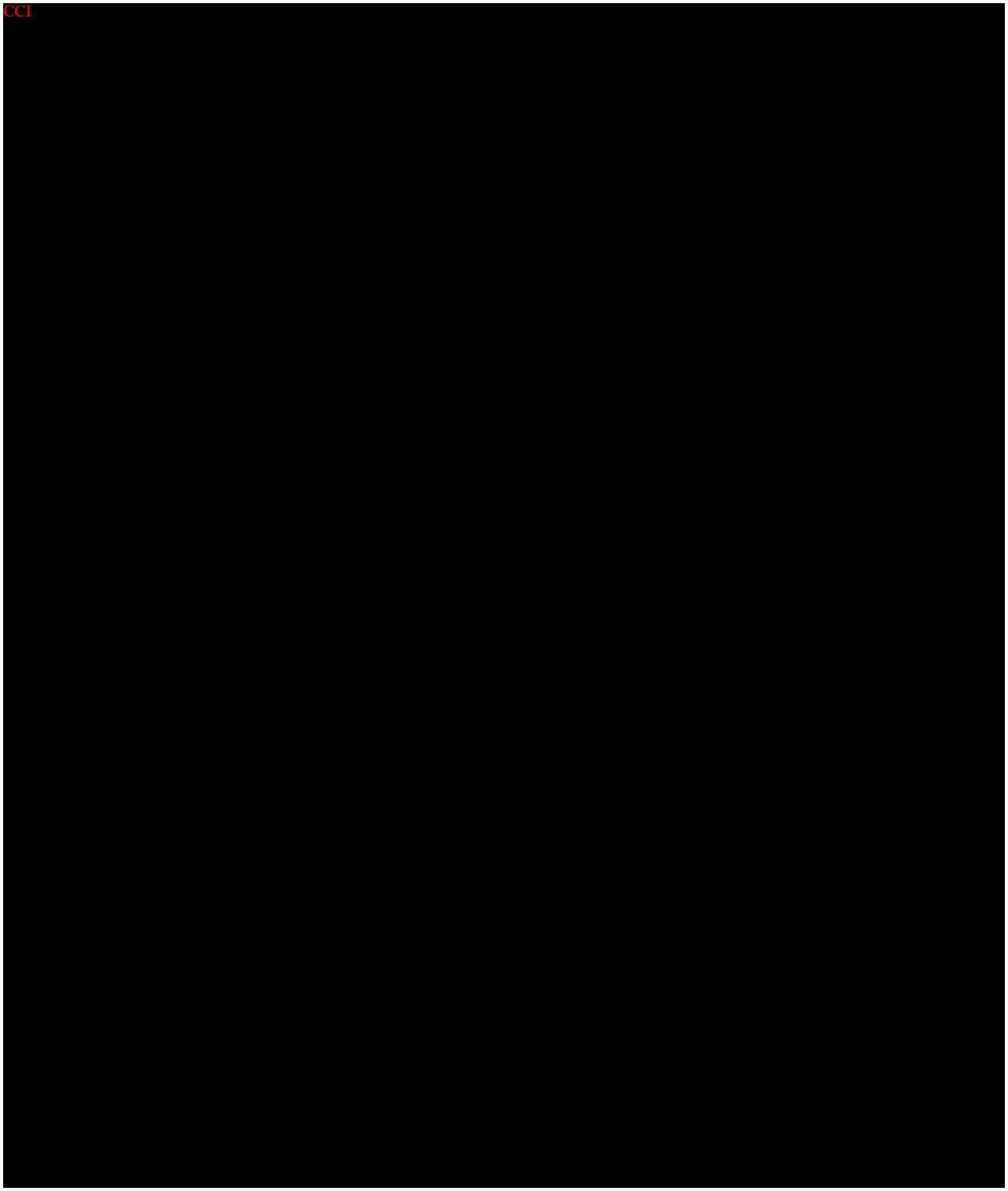
CCI

CCI

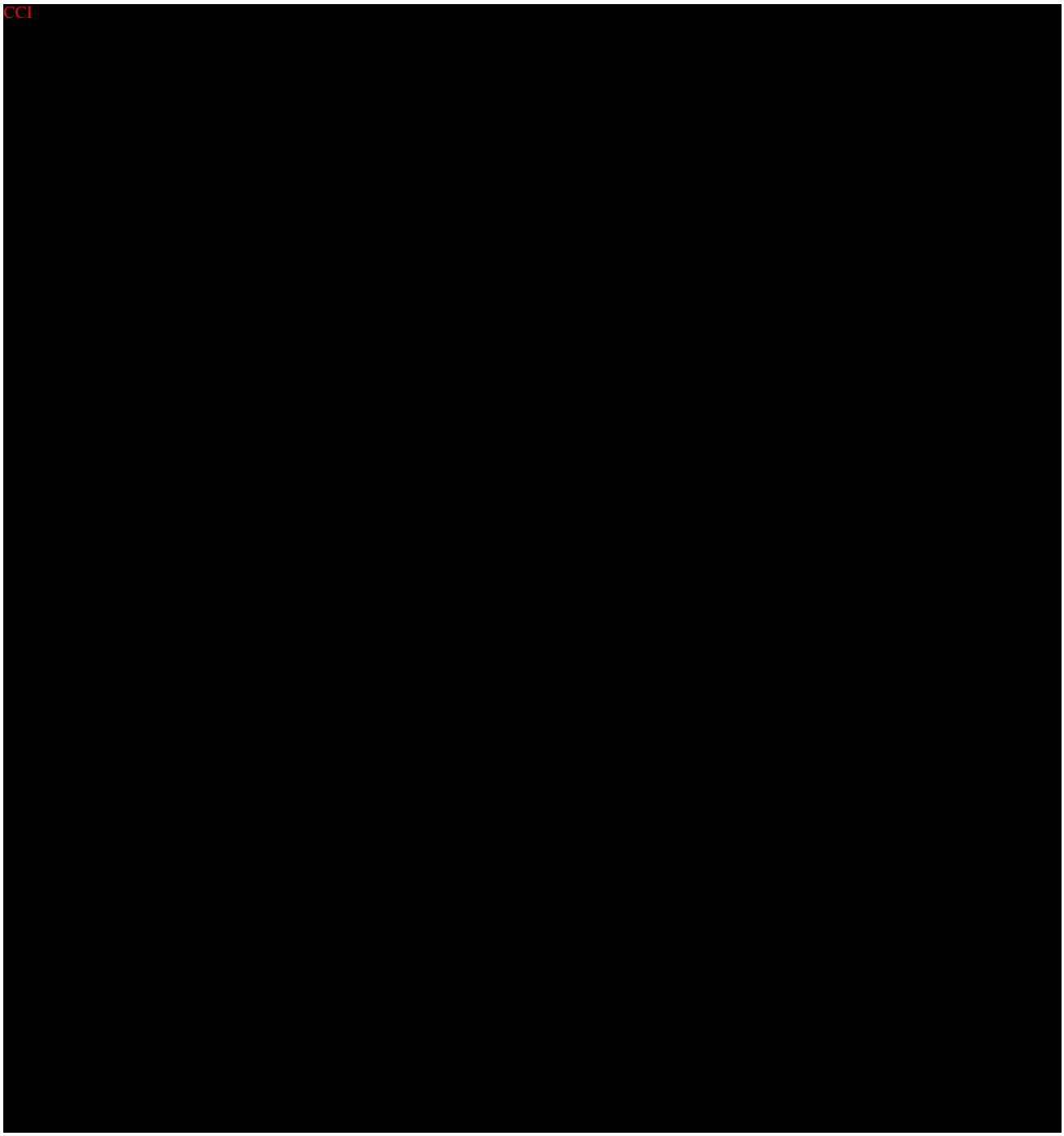
CCI - This section contained Clinical Outcome Assessment data collection questionnaires or indices, which are protected by third party copyright laws and therefore have been excluded.



CCI



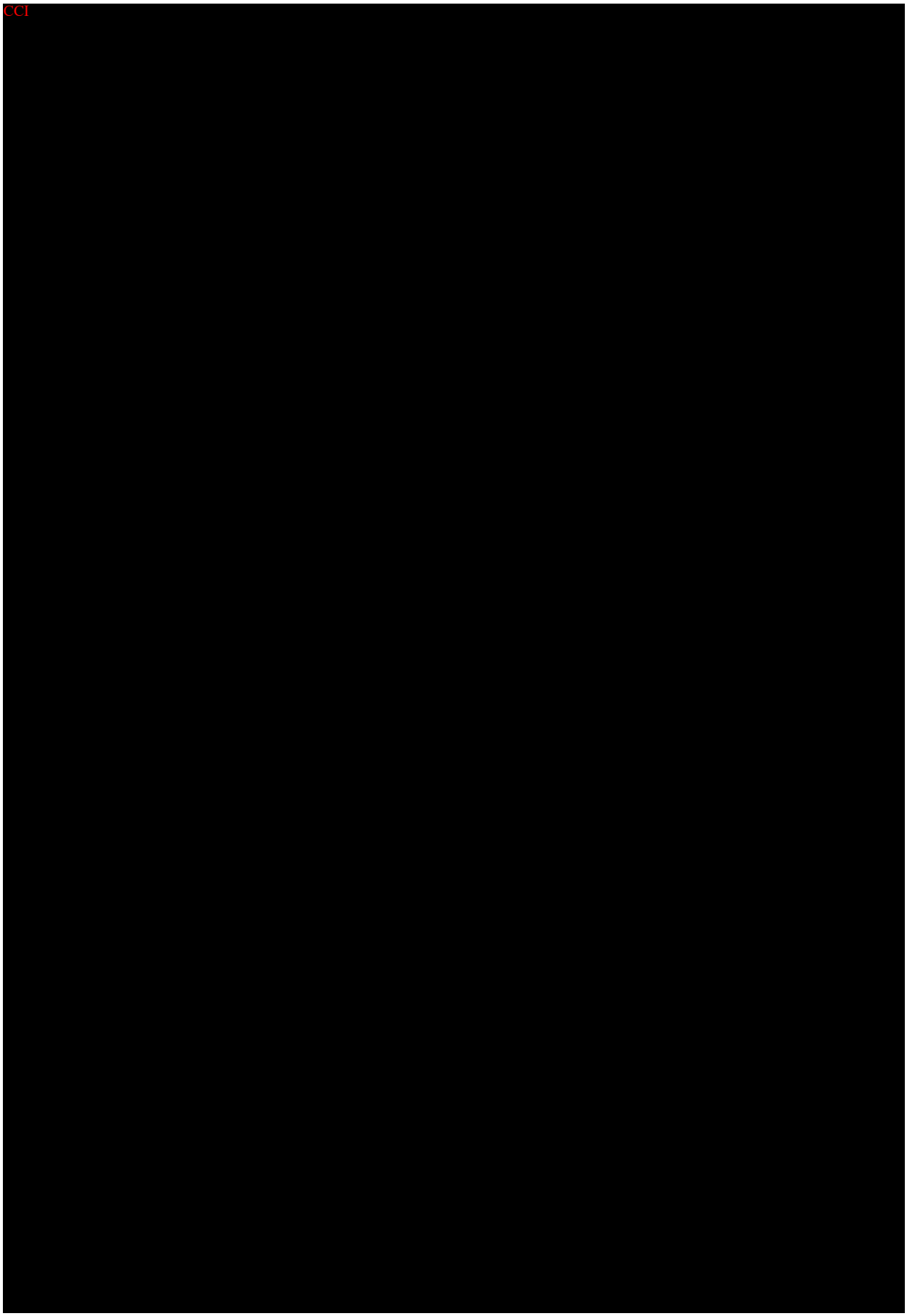
CCI



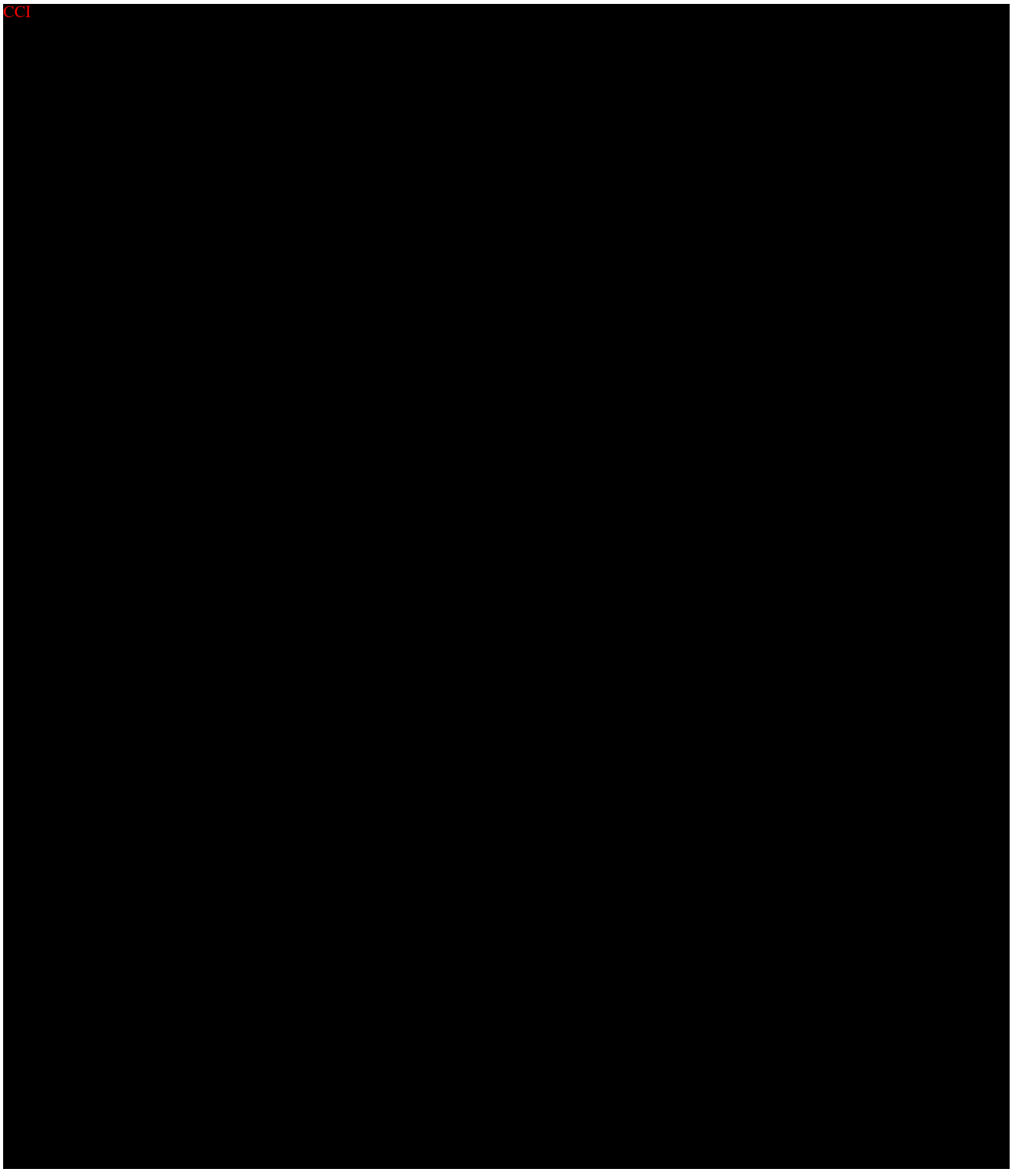
CCI

[redacted]

CCI



CCI



4.5. Safety Analyses

The safety analyses will be based on the Safety analysis set, unless otherwise specified.

4.5.1. Extent of Exposure

Study treatment is administered (subcutaneously injected) approximately every 4 weeks and each dose viewed as providing therapeutic coverage for 4 weeks (28 days).

Duration of exposure is calculated based on the therapeutic coverage as follows:

- Duration (days) =
  (Last IP Administration Date – First IP Administration Date + 29)
- Duration (months) =
  (Last IP Administration Date – First IP Administration Date + 29)*12/365.25
- Participant-years exposure =
  (Last IP Administration Date – First IP Administration Date + 29)/365.25

The extent of exposure to study treatment (mepolizumab/placebo) will be reported, including descriptive statistics of duration of exposure in addition to the number and percentage of participants exposed for ≥1-<3, ≥3-<6, ≥6-<9, ≥9-<12 and ≥12 months, based on each participant's duration of exposure rounded to the nearest whole month. The number of treatments administered and total participant-years exposure will also be summarised.

4.5.1.1. Self-administration of Study Treatment

A subset of Japanese participants who are willing can self-administer study treatment under observation of the principal investigator or delegate from Week 32 onwards.

A self-administered injection is considered successful if the following criteria are met:

- Use of a correct injection site (abdomen or thigh)
- Full dose administered: participant fully inserts the needle, slowly presses the plunger until the stopper reaches the bottom of the syringe and activates the needle guard by moving the thumb up

If the above criteria are met, the investigator will confirm the injection was successfully delivered.

The number and proportion of participants with at least one successful self-administered injection and the number and proportion of participants with all self-administered injections successfully delivered will be summarised. Participants with at least one unsuccessful self-administered injection and reasons for the unsuccessful delivery of study treatment will be listed.

4.5.2. Adverse Events

Adverse events analyses including the analysis of adverse events (AEs), Serious AEs (SAEs) and other significant AEs will be based on GSK Core Data Standards.

Adverse events will be coded using the standard Medical Dictionary for Regulatory Affairs (MedDRA dictionary), with the maximum severity of each AE determined by the investigator (as mild, moderate or severe).

An overview summary of AEs, including counts and percentages of participants with any AE, AEs related to study treatment, AEs leading to permanent discontinuation of study treatment, AEs leading to withdrawal from the study, AEs leading to dose delays, SAEs, SAEs related to study treatment, fatal SAEs, and fatal SAEs related to study treatment will be produced.

AEs will be considered on-treatment for at least 4 weeks (28 days) following the last dose of study treatment. AEs which occur more than 28 days following the last dose of study treatment will be considered as post-treatment. The number and percentage of participants experiencing at least one AE during the on-treatment and post-treatment periods will be summarised separately. Additionally, summaries of the number and percentage of participants with any on-treatment AEs by maximum severity will be produced. If an adverse event severity is missing, the severity is to be populated as 'UNKNOWN'.

The frequency and percentage of AEs will be summarized and displayed in two ways: 1) in descending order by SOC and PT and 2) in descending order by PT only.

Common AEs, defined as ≥3% (prior to rounding) in any treatment group, will be reported. The summary table will be displayed by PT only.

A separate summary will be provided for study treatment-related AEs. A study treatment-related AE is defined as an AE for which the investigator classifies the possible relationship to study treatment as "Yes". A worst-case scenario approach will be taken to handle missing relatedness data, i.e. the summary table will include events with the relationship to study treatment as 'Yes' or missing. The summary table will be displayed by SOC and PT.

All SAEs during the on-treatment and post-treatment periods will be tabulated based on the number and percentage of participants who experienced an event within each period. In addition, fatal and non-fatal SAEs will be reported separately.

Separate summaries will also be provided for study treatment-related SAEs. A study treatment-related SAE is defined as an SAE for which the investigator classifies the relationship to study treatment as "Yes". A worst-case scenario approach will be taken to handle missing data, i.e. the summary table will include events with the relationship to study treatment as 'Yes' or missing. The summary tables will be displayed by SOC and PT.

AEs leading to discontinuation of study treatment, AEs leading to withdrawal from the study and AEs reported on the day of dosing with study treatment will be both summarised. The summary tables will be displayed by SOC and PT.

Additional summaries of on-treatment AEs by time to onset (0-<12 weeks, 12-<24 weeks, 24-<36 weeks, 36-<48 weeks and ≥48 weeks) and on-treatment AEs in participants who have at least one anti-drug antibody (ADA) positive result and those with all negative ADA results post-baseline will be reported.

AEs and SAE summaries will also be presented by Background INCS use subgroup (defined within Section 4.11.1).

The incidence of AEs and SAEs (Fatal and Non-Fatal) of COVID-19, COVID-19 AEs leading to study drug discontinuation, COVID-19 AEs leading to study withdrawal, and COVID-19 AEs by maximum severity, will be obtained from the standard AE and SAE summaries.

4.5.2.1. Adverse Events of Special Interest

Adverse events of special interest (AESIs) are adverse events which are to be closely monitored as the development of mepolizumab progresses.

Systemic reactions are AESIs which are collected using targeted eCRF pages and further categorised by the investigator as either an allergic (type I hypersensitivity) or other systemic reaction. These events are required to be assessed against Sampson's diagnostic criteria for anaphylaxis [Sampson, 2006]. Local injection site reactions are also collected using targeted eCRF pages within the study.

AESIs of potential opportunistic infections, malignancies, serious cardiac, vascular and thromboembolic (CVT) events and serious ischemic events will be identified from a list of relevant preferred terms maintained within a project level reference dataset created based on the latest version of the MedDRA dictionary available at the time of source data lock for this study. Further details of how relevant preferred terms are identified for the AESIs are given in the OPS.

Separate summary tables showing the number and percent of participants with each type of AESI, broken down by preferred term will be created. These summaries will be reported as part of the standard AE/SAE tables for the AESIs of infections, serious infections, neoplasms, cardiac disorders and serious cardiac disorders. The relative risk and risk difference of each AESI between mepolizumab and placebo will be presented with 95% confidence intervals.

For each AESI a profile summary table will be produced containing information on event characteristics including, but not be limited to, the number of participants with the AESI, the number of occurrences of the AESI, the number of participants with a serious event or a drug related event, maximum severity, outcome and action taken.

Separate listings will be provided of participants with systemic reactions identified by the investigator as meeting the criteria for anaphylaxis, systemic reactions categorised as

allergic (type I hypersensitivity), systemic reactions categorised as other, and all other AESIs.

Summaries of AESIs will also be presented by Background INCS use subgroup (defined within Section 4.11.1).

4.5.2.2. Impact of COVID-19 Pandemic on Adverse Event Reporting

The number of participants with suspected, probable or confirmed for COVID-19 infection will be reported. This display will also summarise the number of participants with a COVID-19 diagnosis test performed and the number of participants with positive, negative, or indeterminate results.

Additionally, a listing of COVID-19 assessments and symptoms for participants with COVID-19 AEs will be generated.

4.5.2.3. Cardiovascular Events and Deaths (All Causes)

Cardiovascular events and deaths (all causes) will be captured on targeted CV event eCRF pages for the following AEs and SAEs:

- Arrhythmias
- Congestive heart failure
- Cerebrovascular events/stroke and transient ischemic attack
- Deep venous thrombosis/pulmonary embolism
- Myocardial infarction/unstable angina
- Peripheral arterial thromboembolism
- Pulmonary hypertension
- Revascularization
- Valvulopathy
- Death (all causes)

A summary table will be generated of all reported cardiovascular events and deaths (all causes). Separate patient profiles will also be provided of participants with each of the above listed events.

4.5.3. Additional Safety Assessments

4.5.3.1. Laboratory Data

Summaries of laboratory data including chemistry and haematology parameters, and liver function test data will be based on GSK Core Data Standards and unless otherwise specified will include on-treatment and post-treatment data. Change from baseline values for clinical chemistry and haematology will be summarised in separate tables using descriptive statistics.

Summaries of the number and percentage of participants with worst-case changes from baseline with respect to both the Normal Range (NR) and Potential Clinical Importance (PCI) criteria (for all lab tests where a PCI ranges are defined in the OPS) will be generated separately. Decreases to low, changes to normal or no changes from baseline,

and increases to high will be summarized for the worst-case post-baseline. If a participant has a decrease to low and an increase to high during the same time interval, then the participant is counted in both the "Decrease to Low" categories and the "Increase to High" categories. Participants with a missing baseline value are assumed to have been normal at baseline and participants with no data for a given lab test post-baseline (following first dose of study treatment) will be excluded from this summary.

Separate summary tables for haematology, and chemistry laboratory tests will be produced. These summaries will also be presented by Background INCS use subgroup (defined within Section 4.11.1).

Liver function laboratory tests will be included with chemistry lab tests.

Any liver stopping or liver monitoring events which occur during the study will be summarised. Summaries of hepatobiliary laboratory events including possible Hy's law cases will also be provided in addition to what has been described above. Possible Hy's law cases are defined as any elevated alanine aminotransferase (ALT)≥3×upper limit of normal (ULN), total bilirubin≥2×ULN and alkaline phosphatase (ALP)<2×ULN/missing. Total bilirubin≥2×ULN can be within 28 days following the ALT elevation and if direct bilirubin is available on the same day, it must be ≥35% of total bilirubin. ALP<2×ULN/missing means it is satisfied unless the ALP is ≥2xULN at the time of bilirubin elevation (within 28 days following the ALT elevation). The summary will be produced for worst-case post-baseline only. A listing will be generated of all chemistry and liver function tests in participants who met the protocol defined liver stopping or liver monitoring criteria.

An e-DISH plot of maximum post-baseline total bilirubin versus maximum post-baseline ALT will be created.

Additionally, the worst-case urinalysis results post-baseline (including both on-treatment and post-treatment data) relative to baseline will be reported. The percentages are based on the number of participants in each treatment group with data for a given test post-baseline. Participants with a missing baseline value are to be assumed to have a negative baseline value.

4.5.3.2. ECG

Summaries of ECG data will be based on GSK Core Data Standards. The number of abnormal and clinically significant 12-lead ECG results post-baseline will be produced by visit, also showing the worst recorded value post-baseline (including on-treatment and post-treatment data). Change from baseline values for each ECG parameter will be summarised using descriptive statistics.

The QTc data analysis will use the collected QTc values based on the Fridericia formula (QTcF). QTcF values will be rounded to the nearest integer and values will be categorized into the following CTCAE grade and ranges: Grade 0 CCI [redacted] Grade 1 CCI [redacted]), Grade 2 CCI [redacted] and Grade 3 CCI [redacted] Summaries of grade increase will be provided. These summaries will display the number and percentage of participants with any grade increase, increase to grade 2 and increase to

grade 3 for the worst-case post-baseline only. Missing baseline grade will be assumed as grade 0. Participants with QTc data post-baseline (following first dose of study treatment) will be excluded from this summary.

The changes in QTc values will be categorized into the clinical concern ranges which are specific to changes in QTc: >30-≤60 and >60 msec. A summary of change in QTc value will display the number and percentage of participants with a change within each range for the worst-case post-baseline only. Participants with missing baseline values or with no QTc data post-baseline (following first dose of study treatment) will be excluded from this summary.

A listing will be generated of all ECG results in participants who met the protocol defined QTc stopping criteria.

4.5.3.3. Vital Signs

Summaries of vital signs data [systolic blood pressure (SBP), diastolic blood pressure (DBP) and heart rate] will be based on GSK Core Data Standards. Change from baseline values for SBP, DBP and heart rate will be summarised using descriptive statistics.

Summaries of grade increase from baseline in SBP and DBP will be provided separately. These summaries will display the number and percentage of participants with any grade increase, increase to Grade 2 and increase to Grade 3 for worst-case post-baseline only. The grade definition for SBP is: Grade 0 CCI Grade 1 CCI, Grade 2 CCI, Grade 3 CCI The grade definition for DBP is: Grade 0 CCI, Grade 1 CCI Grade 2 CCI, Grade 3 CCI. The summaries will be produced for worst-case post-baseline only (including both on-treatment and post-treatment data). A missing baseline grade will be assumed as grade 0. Participants with no SBP or DBP data post-baseline (following first dose of study treatment) will be excluded from this summary.

4.6. Immunogenicity

An immune response to a therapeutic drug can lead to generation of anti-drug antibodies within the blood. For the immunogenicity assessment, two types of antibody assays will be performed, a binding anti-drug antibody (ADA) assay and a neutralizing antibody (NAb) assay.

For the binding ADA assay, there will be three testing steps: screening, confirmation and titration. The screening assay produces a result of positive or negative relative to a screening assay cut point. Positive samples continue with the confirmation assay, which also produces a result of positive or negative relative to a confirmation cut point. For positive confirmation samples, a titre value will also be obtained to quantify the degree of binding in a titration assay. Participants with a positive confirmation result will be positive for the binding ADA assay. Participants who test positive for the binding ADA assay, will be tested for the NAb assay, which also reports results as positive or negative.

Immunogenicity data will be summarised for all participants in the Safety analysis set. For the binding ADA assay, confirmation results at each visit will be categorised as negative or positive and summarised. NAb assay results will also be summarised.

The highest binding ADA assay confirmatory result obtained post-baseline for a participant (including both on-treatment and post-treatment data), will be summarised with a positive result being considered higher than a negative one, participants with both positive and negative results will be identified in the positive category. Descriptive statistics for titre results will also be presented. In addition, the highest NAb assay result during the same period will be summarised (again with a positive result being considered higher than a negative one and participants with both positive and negative results will be identified in the positive category).

A summary of treatment emergent positive confirmatory binding ADA assay results in the subset of participants who did not have a positive confirmatory binding ADA assay result prior to the first dose of study treatment will also be presented.

All immunogenicity results (i.e. binding ADA screening and confirmatory assay results, titre values and NAb results) will be listed.

4.7. Pharmacodynamic Analyses

The pharmacodynamic endpoints for analysis are:

- Absolute and ratio to baseline in blood eosinophil count at Weeks 4, 8, 24 and 52

4.7.1. Definition of Endpoint(s)

Blood eosinophil counts will be measured during the course of the study over the 52 weeks treatment period. Blood eosinophil counts will be log-transformed before analysis. Zero values for the baseline blood eosinophil count as well as for blood eosinophil count during the study will be replaced by 0.005 GI/L prior to log transformation.

4.7.2. Main Analytical Approach

The treatment effect to be estimated will be the comparison of mepolizumab 100 mg SC + SoC to placebo + SoC. The population of interest will be the entire trial population (patients meeting the study inclusion/exclusion criteria with a diagnosis of CRSwNP / ECRS) who are randomised and receiving study treatment (ITT analysis set).

The summary measure of treatment effect will be the difference in mean scores (ratio of the effect) between mepolizumab and placebo at Weeks 4, 8, 24 and 52.

Only data collected while the participant was on-treatment will be used in the analysis (applying a 'while on-treatment strategy'). Blood eosinophil counts taken more than 28 days following the last dose of study treatment will not be included in the analysis. Missing data will be assumed to be missing at random (MAR).

The ratio to baseline in blood eosinophil count will be analysed using mixed models repeated measures adjusting for covariates of treatment group, log of baseline blood eosinophil count, background INCS use, country and time point, plus interaction terms for time point by baseline and time point by treatment group.

An unstructured covariance structure for the R matrix will be used by specifying 'type=UN' on the REPEATED line. The Kenward and Roger method for approximating the denominator degrees of freedom and correcting for bias in the estimated variance-covariance of the fixed effects will be used. The estimated mean for each treatment group will be calculated from the model using the observed marginal distributions of the sample covariates (OM option in LSMEANS statement).

Model-estimates of the absolute mean and mean ratio to baseline will be back-transformed and presented with corresponding standard errors on the log scale for each treatment group by visit. Estimated treatment differences (Mepolizumab/Placebo) with corresponding 95% CIs and p-values will also be presented.

4.8. Population Pharmacokinetic (PopPK) Analyses

The primary goal of this analysis is to characterize the population pharmacokinetics of mepolizumab administered subcutaneously in participants with CRSwNP / ECRS. The analysis will be based on the most recent population pharmacokinetics model established with historical data from various indications. The influence of demographics, baseline characteristics, including disease activity, and co-medication on the pharmacokinetics of mepolizumab in this population will also be investigated. These data will be used to explore potential ethnic differences between Japanese/Chinese and Caucasians in mepolizumab exposure. The individual participant PK parameters will be estimated and documented for the purposes of any subsequent exposure-response (PK/PD) analyses.

Mepolizumab plasma concentration-time data [samples collected at Weeks 4, 28 and 52; at the Early Withdrawal Visit and additional visits (if applicable)] will be analysed by population methods using nonlinear mixed-effects modelling. The analysis will be carried out using appropriate software (e.g. NONMEM ™ or SAS ™).

In support of the analysis described below, a dataset will be generated for which specifications will be provided in a separate document.

4.8.1. Definition of Endpoint(s)

Blood for PK will be collected at Week 4 (pre-dose), Week 28 (pre-dose) and Week 52 (or IP discontinuation/early withdrawal). In addition, one post-dose PK will be collected at Week 29 (one week after dose at Week 28). As for Week 29, a blood sample will be collected from up to approximately the first 30 Japanese and all Chinese participants randomised, no PK samples will be collected outside of Japan and China at Week 29. Individual PK parameters and covariates will be log-transformed and standardized before analysis.

Based on previous knowledge of mepolizumab 100 mg SC pharmacokinetics, concentrations below the limit of quantification (BLQ) of the assay are not anticipated. Any such results will be treated as missing.

4.8.2. Main Analytical Approach

A population PK analysis of sparse concentration data will be performed. The population of interest will be the entire trial population (patients meeting the study inclusion/exclusion criteria with a diagnosis of CRSwNP / ECRS) who are randomised and receiving study treatment with at least 1 non-missing PK assessment (PK analysis set).

The summary measure of interest will be the log-transformed mepolizumab plasma concentration. For participants prematurely discontinuing study treatment, all available data will be included in the analysis (intercurrent event to be handled using a treatment policy strategy).

Sparse blood sampling is implemented in this study for determination of mepolizumab plasma concentration and subsequent data analysis by population PK methods using the

most recent population pharmacokinetics model (meta-analysis PK model of data across indications described in GSK Document Numbers: 2015N238436_00 and 2020N427698_00). The main objectives of this population PK analysis are:

- To evaluate mepolizumab pharmacokinetics in participants with CRSwNP / ECRS following the subcutaneous administration of a 100 mg dose every 4 weeks;
- To investigate the impact of covariates of interest in the studied CRSwNP / ECRS population (such as baseline characteristics, co-medication) on specific parameters (e.g. clearance) to identify potential sources of inter-individual variability in these parameters;
- To obtain individual plasma concentration predictions for the timepoints at which PD is measured to allow the conduct of population PKPD analyses if deemed appropriate.

Outlier data will be assessed for plausibility; however, the aim is to use all available data whenever possible. Any decision to exclude data will be fully documented and specified in the clinical study report.

Further details of the Population PK analyses will be provided in the Output and Programming Specification (OPS) document.

4.8.3. Supplementary Analyses

Complementing the population-PK analysis, an additional analysis of a repeated measures linear mixed-effects model of log-transformed concentration will be used to analyse the sparse PK data. All data will be analysed on an Intent-to-Treat basis using planned Visit, irrespective of discontinuation of study treatment (to be handled using a treatment policy strategy).

The model will adjust for covariates of country, log bodyweight, log creatinine clearance, log serum albumin with a repeated term for Visit. An unstructured covariance structure for the R matrix will be used by specifying ‘type=UN’ on the REPEATED line. The Kenward and Roger method for approximating the denominator degrees of freedom and correcting for bias in the estimated variance-covariance of the fixed effects will be used. Selection of covariates will be based on biological plausibility, parsimony and a significance level of 0.05. Forwards and backwards selection criteria may be applied as per the population PK model.

The estimated means for each visit will be calculated from the model using the observed marginal distributions of the sample covariates (OM option in LSMEANS statement). Model-estimates of the absolute mean and mean ratio to Week 4 (accumulation ratio) will be back-transformed and presented as ratios with corresponding standard errors on the log scale by visit. Estimated exposure ratios with corresponding 95% CIs and p-values will also be presented. The effects of bodyweight, creatinine clearance and serum albumin will be evaluated as a ratio to standard values and back transformed and presented as ratios with corresponding standard errors on the log scale. Estimated exposure ratios with corresponding 95% CIs and p-values will also be presented.

4.9. Population Pharmacokinetic/Pharmacodynamic (PopPKPD) Analyses

If deemed appropriate, a population pharmacokinetic/pharmacodynamic analysis will be conducted. Mepolizumab blood eosinophil count-time data will be analysed by population methods using nonlinear mixed-effects modelling. The analysis will be carried out using appropriate software (e.g. NONMEM or SAS).

In support of the analysis described below, a specific dataset will be generated for which specifications will be provided in a separate document.

4.9.1. Definition of Endpoint(s)

Blood for PK are defined within Section 4.8.1.

Blood eosinophil counts are defined within Section 4.7.1.

4.9.2. Main Analytical Approach

The population of interest will be the PK analysis set and the ITT analysis set (for blood eosinophils). For participants prematurely discontinuing study treatment, all available data will be included in the analysis (intercurrent event to be handled using a treatment policy strategy).

This PopPKPD analysis will be performed by population methods using the most recent population PKPD model [meta-analysis PKPD model of data across indications described in GSK Document Numbers: 2020N427698_00, 2015N238436_00 and 2015N255079_00 (extension of the former report)].

The objectives of the popPKPD analysis are:

- To evaluate mepolizumab pharmacodynamics in participants with CRSwNP / ECRS following subcutaneous administration of a 100 mg dose every 4 weeks;
- To investigate the impact of covariates of interest in the studied CRSwNP / ECRS population (such as baseline characteristics, co-medication) on specific parameters (e.g. maximum blood eosinophil reduction) in order to identify potential sources of inter-individual variability in these parameters.

Outlier data will be assessed for plausibility; however, the aim is to use all available data whenever possible. Any decision to exclude data will be fully documented and specified in the clinical study report.

Further details of the popPKPD analyses will be provided in the Output and Programming Specification (OPS) document.

4.10. Genetics Analysis

A separate analysis plan will be produced for any analysis of genetic data.

CCI

[redacted]

CCI

[redacted]

CCI

[redacted]

4.12. Interim Analyses

No interim analysis of data is planned for this study.

4.13. Changes to Protocol Defined Analyses

The changes from the originally planned statistical analysis specified in the Protocol Amendment 3 [GlaxoSmithKline Document Number: TMF-14790303. Dated: 21-Jul-2022] are described in Table 3.

Table 3 Changes to Protocol Defined Analyses

Protocol Defined Analysis	SAP Defined Analysis	Rationale for Changes
Time to first 'need for surgery' (NP score ≥5 and overall VAS symptom score >7) up to Weeks 4, 8, 24 and 52	Need for surgery (NP score ≥5 and overall VAS symptom score >7) at Week 52	It is expected that patients will enter the study with a 'need for surgery' due to requirement to meet inclusion criteria of severe NP symptoms.
Lund Mackay CT score at Week 52 will use the same population, summary measure and strategies for intercurrent events as for the primary estimands. Analysis methods will be the same as for the co-primary endpoints.	Lund Mackay CT score at Week 52 will use the same population, summary measure and strategies for intercurrent events as for the primary estimands. Analysis methods will based on ANCOVA.	Lund Mackay CT score is recorded once post-baseline at Week 52, therefore assessment of change from baseline via ANCOVA analysis is appropriate.

5. SAMPLE SIZE DETERMINATION

The sample size is based on the co-primary efficacy endpoints of total endoscopic nasal polyp score at Week 52 and nasal obstruction VAS score during the 4 weeks prior to Week 52.

The sample size of 160 participants randomised in a 1:1 ratio to each treatment arm provides at least 90% power to demonstrate a statistically significant result for both co-primary endpoints using a mixed model repeated measures (MMRM) analysis model.

Estimates of residual standard deviation (SD) of 1.82 for total endoscopic nasal polyp score at Week 52 and 3.25 for nasal obstruction VAS score during the 4 weeks prior to Week 52 are taken from a post-hoc MMRM analysis of observed data from study 205687 (SYNAPSE) which assumed a "missing at random" (MAR) assumption for missing data.

With the above assumptions and a sample size of 160 randomized subjects (80 per arm), it is estimated that the null hypothesis will be rejected at the two-sided 5% level of significance for total endoscopic nasal polyp score if the observed difference between treatments is at least 0.57 units and for nasal obstruction VAS score if the observed difference between treatments is at least 1.01.

For total endoscopic nasal polyp score, if the true difference between treatments is 1.00 units, then the study has a probability of 93.2% of observing a difference in total endoscopic nasal polyp score between treatments of at least 0.57 units and therefore 93.2% power for declaring significance on this endpoint. The estimated improvement observed in total endoscopic nasal polyp score for mepolizumab 100 mg SC compared to placebo within study 205687 (SYNAPSE) was 0.99 units (95% CI: 0.61 to 1.36).

For nasal obstruction VAS score, if the true difference between treatments is 2.00, then the study has a probability of 97.2% of observing a difference in nasal obstruction VAS score between treatments of at least 1.01 units and therefore a 97.2% power for declaring significance on this endpoint. The measured improvement observed in nasal obstruction VAS score for mepolizumab 100 mg SC compared to placebo within study 205687 (SYNAPSE) was 1.97 units (95% CI: 1.31 to 2.63).

If the true population distribution of each co-primary endpoint is as described above, a study with sample size of 80 randomised participants randomised in 1:1 ratio to each treatment arm (total 160 participants) is estimated to have at least 90% power regardless of the degree of (positive) correlation of the endpoints.

Assuming a screen failure rate of 40%, approximately 270 participants will need to be screened to achieve approximately 160 randomly assigned to study intervention for a total of approximately 80 participants per intervention group. All participants within the Intent-to-Treat analysis set (randomised participants who receive at least 1 dose of study treatment) will be considered evaluable within the study analysis. See Section 4.2.2 regarding the strategy of handling of intercurrent events for the primary estimands.

6. SUPPORTING DOCUMENTATION

6.1. Appendix 1 Abbreviations and Trademarks

6.1.1. List of Abbreviations

Abbreviation	Description
AE	Adverse Event
AERD	Aspirin-Exacerbated Respiratory Disease
AESI	Adverse events of special interest
ACQ-5	Asthma Control Questionnaire-5
ADA	Anti-drug antibody
ALP	Alkaline phosphatase
ALT	Alanine aminotransferase
BLQ	Below the limit of quantification
CDF	Cumulative distribution function
CI	Confidence interval
COVID-19	Coronavirus Disease-2019
CRSwNP	Chronic rhinosinusitis with nasal polyps
CS	Corticosteroid
CSR	Clinical Study Report
CT	Computed tomography
CV	Cardiovascular
CVT	Cardiac, vascular and thromboembolic
DBP	Diastolic blood pressure
eCRF	Electronic Case Record Form
ECG	Electrocardiogram
ECRS	Eosinophilic chronic rhinosinusitis
ED	Emergency department
eDiary	Electronic diary
FESS	Functional endoscopic sinus surgery
GSK	GlaxoSmithKline
HCP	Healthcare professional
ICE	Intercurrent event
IMCS	Intramuscular corticosteroid
INCS	Intranasal corticosteroid
IP	Investigational Product
ITT	Intent-To-Treat
IWRS	Interactive Web Response System
JESREC	Japanese Epidemiological Survey of Refractory Eosinophilic Chronic Rhinosinusitis
LAR	Legally acceptable representative
MAR	Missing at random
MCID	Minimum clinically important difference
MCS	Mental Component Summary
MedDRA	Medical Dictionary for Regulatory Affairs
MMRM	Mixed Model Repeated Measures
NAb	Neutralizing antibody
NP	Nasal polyps
NR	Normal range
NSAID	Nonsteroidal anti-inflammatory drug
OC	Osteomeatal complex
OM	Observed margins
OPS	Output and Programming Specification

Abbreviation	Description
PCI	Potential Clinical Importance
PCS	Physical Component Summary
PD	Pharmacodynamic
PK	Pharmacokinetic
PopPK	Population PK
PT	Preferred term
QTcF	Frederica's QT Interval Corrected for Heart Rate
SAP	Statistical Analysis Plan
SAE	Serious Adverse Event
SBP	Systolic blood pressure
SC	Subcutaneous
SD	Standard deviation
SF-36	Short Form-36
SMQ	Standardised MedDRA Query
SNOT-22	Sino-Nasal Outcomes Test-22
SOC	System organ class
SoC	Standard of care
SSD	Safety syringe device
TMF	Trial Master File
ULN	Upper limit of normal
VAS	Visual Analogue Scale
WPAI-GH	Work Productivity and Activity Impairment Questionnaire - General Health

6.1.2. Trademarks

Trademarks of the GSK Group of Companies
NONE

Trademarks not owned by the GSK Group of Companies
NONMEM
OPTUM
R
SAS

7. REFERENCES

Bhattacharyya N. Test-retest reliability of computed tomography in the assessment of chronic rhinosinusitis. *Laryngoscope.* 1999;109:1055-58.

Bretz F, Maurer W, Brannath W, Posch M. A graphical approach to sequentially rejective multiple test procedures. *Statistics in Medicine.* 2009;28:586-604.

GlaxoSmithKline Document Number 2015N238436_00 Study ID N/A. A population PK and PKPD meta-analysis of combined intravenous and subcutaneous mepolizumab data. Report Date 27-May-2015.

GlaxoSmithKline Document Number 2015N255079_00 Study ID N/A. Supplementary outputs from a population PK and PKPD meta-analysis of combined intravenous and subcutaneous mepolizumab data. Report Date 04-Oct-2015.

GlaxoSmithKline Document Number 2020N441602_00 Study ID 209692. A randomised, double-blind, placebo-controlled, parallel group Phase III study to assess the clinical efficacy and safety of 100 mg SC Mepolizumab in adults with chronic rhinosinusitis with nasal polyps (CRSwNP) / eosinophilic chronic rhinosinusitis (ECRS) – MERIT: Mepolizumab in Eosinophilic chronic RhinosinusITis study. Report Date 14_Aug-2020.

GlaxoSmithKline Document Number TMF-14790303 Study ID 209692. A randomised, double-blind, placebo-controlled, parallel group Phase III study to assess the clinical efficacy and safety of 100 mg SC Mepolizumab in adults with chronic rhinosinusitis with nasal polyps (CRSwNP) / eosinophilic chronic rhinosinusitis (ECRS) – MERIT: Mepolizumab in Eosinophilic chronic RhinosinusITis study. Report Date 21-Jul-2022.

GlaxoSmithKline Document Number TMF-2114641 Study ID 209692. A randomised, double-blind, placebo-controlled, parallel group Phase III study to assess the clinical efficacy and safety of 100 mg SC Mepolizumab in adults with chronic rhinosinusitis with nasal polyps (CRSwNP) / eosinophilic chronic rhinosinusitis (ECRS) – MERIT: Mepolizumab in Eosinophilic chronic RhinosinusITis study. Report Date 08-Oct-2020.

GlaxoSmithKline Document Number 2020N427698_00 Study ID 205687. A randomised, double-blind, parallel group PhIII study to assess the clinical efficacy and safety of 100 mg SC Mepolizumab as an add on to maintenance treatment in adults with severe bilateral nasal polyps - SYNAPSE (StudY in NAsal Polyps patients to assess the Safety and Efficacy of mepolizumab). Report Date 23-Jun-2020.

Hopkins S, Gillet S, Slack R, Lund V J, Browne J P. Psychometric validity of the 22-item Sinonasal Outcome Test. *Clin. Otolaryngol.* 2009;34:447-454.

Juniper E F, O'Byrne P M, Guyatt G H, Ferrie P J, King D R. Development and validation of a questionnaire to measure asthma control. *Eur. Respir J.* 1999;14:902-207.

Juniper E F, Svensson K, Mork A C, Stahl E. Measurement properties and interpretation of three shortened versions of the asthma control questionnaire. *Respiratory Medicine.* 2005;99:553-558.

Lund V J, Mackay I S. Staging in rhinosinusitis. *Rhinology.* 1993;31(4):183-184.

Rubin D B. *Multiple Imputation for Nonresponse in Surveys.* 1 ed. New York:John Wiley & Sons; 1987.

Sampson H A, Munoz-Furlong A, Campbell R L et al. Second Symposium on the definition and management of anaphylaxis: Summary Report. Second National Institute of Allergy and Infectious Disease / Food Allergy and Anaphylaxis Network Symposium. *Journal of Allergy and Clinical Immunology.* 2006;117:391-397.