STATISTICAL ANALYSIS PLAN

FINAL
(Version 1.0)
01 SEPTEMBER 2020

PROTOCOL NUMBER: GB001-2101
(Version 5.0, 16 April 2020)

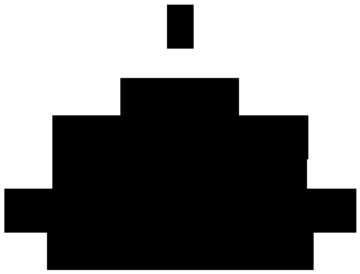
A Phase 2a, randomized, double-blind, placebo-controlled, multi-center study to evaluate the effect of GB001 in patients with chronic rhinosinusitis with or without nasal polyps

SPONSORED BY

GB001, Inc., a wholly owned subsidiary of Gossamer Bio, Inc.
3013 Science Park Road
San Diego, CA 92121
Telephone: (858) 684-1300

PREPARED BY

Gossamer Bio Services, Inc.



VERSION CONTROL

Version Number	Date	Comments/Changes
1.0	01 September 2020	Final

[redacted]

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbreviation	Definition
ACQ-5	Asthma Control Questionnaire
AE	adverse event
ALT	alanine aminotransferase
AM	morning
ANCOVA	analysis of covariance
ANOVA	analysis of variance
AST	aspartate aminotransferase
ATC	anatomical therapeutic chemical
BID	twice daily
BLQ	below limit of quantification
BMI	body mass index
BPM	beats per minute
CI	confidence interval
COVID-19	Coronavirus Disease 2019
CRS	chronic rhinosinusitis
CRSsNP	chronic rhinosinusitis without nasal polyps
CRSwNP	chronic rhinosinusitis with nasal polyps
CSR	clinical study report
CT	computed tomography
CTS	continuous
CV	coefficient of variation
ECG	electrocardiogram
eCRF	electronic case report form
eDISH	evaluation of Drug-Induced Serious Hepatotoxicity
EM	expectation-maximization
EMA	European Medicines Agency
EW	early withdrawal
FDA	Food and Drug Administration
FeNO	fractional exhaled nitric oxide

Abbreviation	Definition
GCP	Good Clinical Practice
GGT	gamma-glutamyl transferase
GINA	Global Initiative for Asthma
ICH	International Conference on Harmonisation
INCS	intranasal corticosteroids
IP	investigational product
ITT	intent-to-treat
LABA	long-acting beta-agonist
LAMA	long-acting muscarinic antagonist
LLOQ	lower limit of quantification
LS	least squares
MCID	minimal clinically important difference
MCMC	Markov Chain Monte Carlo
MedDRA	Medical Dictionary for Regulatory Activities
MFNS	mometasone furoate nasal spray
MI	multiple imputation
MMRM	mixed-effects model with repeated measures
NC	nasal congestion
nNO	nasal nitric oxide
NP	nasal polyps
NPS	nasal polyp score
OR	odds ratio
PD	pharmacodynamics
PGx	pharmacogenetics
PM	evening
PK	pharmacokinetics
PT	preferred term
QD	once daily
QOL	quality of life
QTcF	Fridericia's correction formula for QT interval

Abbreviation	Definition
REML	restricted maximum-likelihood
SAE	serious adverse event
SAP	statistical analysis plan
SD	standard deviation
SE	standard error
SNOT-22	Sino-Nasal Outcome Test
SoA	schedule of activities
SOC	system organ class
TEAE	treatment-emergent adverse event
TSS	total symptom score
ULN	upper limit of normal
UPSIT	University of Pennsylvania Smell Identification Test
VAS	visual analog scale
WHO	World Health Organization

1. PURPOSE OF THE ANALYSIS

This statistical analysis plan (SAP) provides a comprehensive and detailed description of the statistical methodology to be used for reporting of the study results for use in the clinical study report (CSR) and is based on protocol version 5.0. The purpose of this SAP is to describe the methodology, procedures, rules, and conventions to be used for the reporting of results. Results to be reported will include summaries of subject disposition, demographic and baseline characteristics, significant protocol deviations, prior and concomitant medications, study treatment exposure and compliance, primary, secondary, and select tertiary/exploratory efficacy endpoints, pharmacokinetic (PK) concentrations, and safety endpoints, including adverse events (AEs), serious AEs (SAEs), and laboratory, vital sign, and electrocardiogram (ECG) parameters.

If additional analyses are required to supplement the planned analyses described in this SAP, they may be completed and will be identified in the CSR as post hoc.

The last two "Exploratory" objectives and endpoints in Section 2.1 will not be addressed in this SAP.

This SAP was written with due consideration of the recommendations outlined in the most recent International Conference on Harmonisation (ICH) E9 Guideline (ICH E9, 1998) entitled Guidance for Industry E9 Statistical Principles for Clinical Trials and the most recent ICH E3 Guideline (ICH E3, 1995) entitled Guidance for Industry Structure and Content of Clinical Study Reports and was finalized prior to study unblinding.

In this document, investigational product (IP) and study treatment have the same meaning.

2. PROTOCOL SUMMARY

2.1. Objectives and Endpoints

The purpose of this Phase 2a study is to evaluate the efficacy, safety, PK, and pharmacodynamic (PD) of GB001 compared with placebo in patients with chronic rhinosinusitis (CRS) with or without bilateral nasal polyps (NP). The primary, secondary, and exploratory objectives and endpoints are as follows:

Objectives	Endpoints
Primary	
• To evaluate the effect of GB001 on the Sino-Nasal Outcome Test-22 (SNOT-22)	• Change from baseline to Week 16 in SNOT-22
Secondary	
• To evaluate the effect of GB001 on reducing opacification of the sinuses as measured by computed tomography (CT) scan	• Change from baseline to Week 16 in opacification of sinuses as measured by Lund-Mackay score on CT scan
• To evaluate the effect of GB001 in the treatment of bilateral NP by assessment of the endoscopic nasal polyp score (NPS) in a subset of patients with nasal polyps	• Change from baseline to Week 16 in NPS • Time to first response (≥ 1 point improvement from baseline) in NPS
• To evaluate the effect of GB001 in improving patient reported symptoms	• Change from baseline to Week 16 in nasal congestion (NC) • Change from baseline to Week 16 in total symptom score (TSS)
• To evaluate the effect of GB001 in improving sense of smell (University of Pennsylvania Smell Identification Test, UPSIT)	• Change from baseline to Week 16 in UPSIT
• To evaluate the effect of GB001 on chronic rhinosinusitis (CRS) exacerbation	• Time to first CRS exacerbation, defined as deterioration of CRS symptoms requiring treatment with an antibiotic, an anti-inflammatory drug, or a symptom reliever; an Emergency Department visit; or hospitalization
• To evaluate the safety and tolerability of GB001 compared with placebo	• Incidence of treatment-emergent adverse events (TEAEs) • Change from baseline in laboratory, electrocardiogram (ECG), and vital signs parameters

Exploratory	
• To evaluate the effect of GB001 on symptom improvement as measured by a visual analog scale (VAS)	• Change from baseline to Week 16 in VAS
• To evaluate the effect of GB001 on CRS exacerbation rate	• Annualized rate of CRS exacerbations
• To evaluate the effect of GB001 on maxillary CT opacification	• Change from baseline to Week 16 in percentage of maxillary sinus volume occupied by disease on CT scan
• To evaluate the effect of GB001 on worsening of CRS requiring the use of antibiotics	• Proportion of subjects with antibiotic use due to worsening CRS by Week 16
• To evaluate PK of GB001	• Plasma concentration of GB001
• To evaluate the effect of intervention on, as a function of, the PD markers	• Correlation between baseline markers and safety, tolerability, and efficacy parameters • Change in PD markers over time
• To evaluate the relationship between safety, efficacy, and exposure parameters and pharmacogenetics (PGx)	• Correlation between baseline characteristics and single nucleotide polymorphisms • Change from baseline in safety, efficacy, and exposure parameters as a function of single nucleotide polymorphisms

Additional efficacy endpoints not defined in the protocol but defined in this SAP include:

- Proportion of subjects with a ≥ 9 point decrease from baseline at Week 16 in SNOT-22 total score
- Proportion of subjects with a reduction from baseline of ≥ 1 in NPS at Weeks 8 and 16
- Change from baseline to Week 16 in anterior rhinorrhea, posterior rhinorrhea, and loss of sense of smell scores
- Change from baseline to Week 16 in volume of the air, volume of mucosa, and thickness of lateral wall
- Change from baseline to Week 16 in Asthma Control Questionnaire (ACQ-5) score.

2.2. Overall Study Design and Plan

This is a Phase 2a, randomized, double-blind, placebo-controlled, multi-center study to evaluate GB001 administered once a day for 16 weeks in subjects with CRS with or without bilateral NP. All subjects will be maintained on a stable background therapy of mometasone furoate nasal spray (MFNS) for the duration of the study.

The study will commence with a Screening visit (Visit 1), at which time informed consent will be obtained and inclusion and exclusion criteria will be assessed. Informed consent may be

obtained prior to the day of the Screening visit to allow for medication washouts or for other logistical reasons such as obtaining documentation of exacerbations, if necessary. All other screening procedures should be completed, where possible, on the day of the Screening visit. Subjects not meeting the eligibility criteria will be deemed screen failures and will not continue participation in the study. The Run-in period commences with completion of all Screening visit procedures and concludes at the Randomization visit (Visit 2). During the Run-in period of up to 4 weeks, subjects will capture nasal symptoms and MFNS use in an online daily Diary (morning [AM] and evening [PM]):

- Subject's treatment regimen with MFNS should remain stable.
- 2 actuations (50 µg/actuation) in each nostril twice daily (BID; total daily dose of 400 µg), unless subject is intolerant to a BID regimen of MFNS in which case, they can stay on the dose regimen in place prior to screening.
- The only dose modification of MFNS during the Run-in period may be for subjects who were switched from another intranasal corticosteroids (INCS) to MFNS and need to change to a once daily (QD) regimen due to intolerability.

After the Randomization visit, subjects will enter the Double-Blind, Placebo-Controlled period which will be followed by a Follow-up period.

Total duration for study participation per subject is up to 24 weeks (includes a Screening visit; followed by an up to 4-week Run-in period to allow for collection of baseline Diary data and standardized MFNS treatment regimen; a 16-week Double-Blind, Placebo-Controlled period; and a 4-week Follow-up period).

Subjects who permanently discontinue IP or withdraw early from the study will be requested to attend the Early Discontinuation of IP visit or Early Withdrawal (EW) visit, as appropriate, and will be strongly encouraged to complete any remaining study visits as per the Schedule of Activities (SoA).

The first dose of double-blind IP will be administered in the clinic on Day 1. All subsequent doses will be taken orally, in the AM on an empty stomach. Following initiation of IP, subjects will visit the clinic for assessments per protocol. Subjects will continue on mometasone furoate: MFNS two actuations in each nostril BID (or QD in case subject cannot tolerate the high dose) during the course of the study.

All subjects will complete a daily Diary from the Screening visit through the Follow-up visit, responding to the AM and PM CRS symptom questions and documenting twice daily MFNS use.

The protocol contains further details related to the study design and conduct.

2.3. Study Population

The study population is comprised of: symptomatic subjects with CRS without NP (CRSsNP), defined as an absence of NP based on visual examination or an NPS of 0 based on nasal endoscopy, and evidence of an eosinophilic phenotype, defined as a blood eosinophil count of

≥ 250 cells/μL; and symptomatic subjects with CRS with NP (CRSwNP), defined as a minimum NPS of 4 out of a maximum of 8, with at least a score of 2 for each nostril, as assessed by endoscopic diagnosis by a blinded central reader. Subjects will be male or female and ≥ 18 to 75 years of age at the time of the Screening visit. Subjects with historical evidence of neutrophilic predominant polyp disease will be excluded considering that the mechanism of action of GB001 is thought to be linked to eosinophilic Type 2 inflammation.

2.4. Randomization and Treatment Regimens

This study will randomize approximately 100 subjects, with approximately 50 subjects per treatment group, randomized in a 1:1 ratio to the following treatment groups:

- GB001 40 mg QD
- Matching placebo QD

Randomization will be stratified by country, the presence or absence of NP (CRSwNP and CRSsNP), and the presence or absence of comorbid asthma among CRSwNP subjects. Presence of NP is defined by a minimum NPS of 4 out of a maximum score of 8 with a score of at least 2 for each nostril. Absence of NP is defined as an absence of NP based on visual examination or an NPS of 0 based on nasal endoscopy. Presence of comorbid asthma is defined as a diagnosis of asthma by a physician according to Global Initiative for Asthma (GINA) guidelines at screening or prior to study entry. The CRSwNP stratum will consist of approximately 64 subjects, and the CRSsNP stratum will consist of approximately 36 subjects.

2.5. Sample Size Determination

A total sample size of approximately 100 subjects (approximately 50 per treatment group, randomized in a 1:1 ratio) is estimated to provide approximately 90% power to detect a treatment difference of 12.0 between GB001 and placebo at a 0.050 two-sided level of significance for the primary endpoint of change from baseline to Week 16 in SNOT-22 score, assuming a common standard deviation (SD) of 16.8 and a dropout rate of 15% using a two-sided t-test.

A total sample size of approximately 64 subjects in the CRSwNP stratum (approximately 32 per treatment group, randomized in a 1:1 ratio) is estimated to provide approximately 80% power to detect a treatment difference of 1.3 between GB001 and placebo at a 0.050 two-sided level of significance for the secondary endpoint of change from baseline to Week 16 in NPS, assuming a common SD of 1.7 and a dropout rate of 15% using a two-sided t-test.

Therefore, the total sample size of approximately 100 subjects overall will consist of approximately 64 subjects in the CRSwNP stratum and 36 subjects in the CRSsNP stratum.

3. GENERAL ANALYSIS AND REPORTING CONVENTIONS

All analyses will be performed using SAS® System (SAS Institute Inc., Cary, NC) version 9.4 or later.

Categorical variables will be summarized using counts (n) and percentages (%) and will be presented in the format "n (xx.x)". If a count is 0, no percentage will be shown. To ensure completeness, summaries for categorical and discrete variables may include all categories, even if no subjects had a value in a particular category. Additionally, for missing data, a category of "Missing" will be presented as needed.

Continuous variables, unless otherwise stated, will be summarized using descriptive statistics: number of subjects with non-missing data (n), mean, standard deviation (SD), median, 25th and 75th percentiles, minimum, and maximum. Descriptive statistics on select measures may also include the standard error (SE) and 95% confidence interval (CI) for the mean. In general, rounding rules for reporting continuous descriptive summary statistics are as follows:

- If the original values have 0 or 1 decimal places: mean, median, 25th and 75th percentiles will be reported to one more decimal place than the original values, and SD and SE will be reported to 2 more decimal places than the original values
- If the original values have 2 or more decimal places: mean, median, 25th and 75th percentiles, SD, and SE will all be reported to 3 decimal places

Minimum and maximum will always be reported to the same decimal places as the original values, up to a maximum of 3 decimal places. Percent change from baseline will be reported to 1 more decimal place than change from baseline, up to a maximum of 3 decimal places.

Efficacy results will be summarized by various measures including but not limited to least squares (LS) means, differences in LS means, SEs, odds ratios (ORs), hazard ratios, rate ratios, p-values, and two-sided 95% CIs. The SEs, LS means, differences in LS means and corresponding CIs will be reported per the rules above, with the CIs following the same rounding rules as the mean, ORs, hazard ratios, rate ratios, and corresponding CIs will be reported to 3 decimal places, and percentage reductions and corresponding CIs will be reported to 1 decimal place. P-values will be reported to 4 decimal places, with values less than 0.0001 displayed as < 0.0001 and values greater than 0.9999 displayed as > 0.9999. All statistical hypothesis testing will be at an 0.050 two-sided level of significance.

In general, all analyses will be performed for the overall population, with the exception of efficacy endpoints based on NPS, which will be performed only for the randomized CRSwNP subgroup.

Disposition, demographic and baseline characteristics, significant protocol deviations, and medical history tables will be presented by randomized treatment group and will include an overall column combining both treatment groups. Prior and concomitant medication tables will be presented by either randomized treatment group or actual treatment group, depending on the analysis, and will include an overall column combining both treatment groups. Select analyses among these will be performed for the randomized CRSwNP and CRSsNP subgroups.

Efficacy tables will be presented by randomized treatment group only and select efficacy analyses will be performed for the randomized CRSwNP and CRSsNP subgroups. Safety tables will be presented by actual treatment group and select safety analyses will be performed for the actual CRSwNP and CRSsNP subgroups.

When appropriate, sorting in disposition, demographic, and baseline characteristics tables will be based on decreasing incidence within the overall column. In general, AE tables that are displayed by system organ class (SOC) and preferred term (PT) will be sorted by the internationally agreed order for SOC, and by decreasing incidence and then alphabetically for PT within the GB001 40 mg column. Otherwise, safety tables, when applicable, will be sorted by decreasing incidence within the GB001 40 mg column.

Assessments at the Randomization Visit are to be performed prior to the first dose of study treatment. Therefore, for all measures, the baseline value is defined as either the last non-missing value before the first dose of study treatment, the mean over the last 7 days of the Run-in period including the value on Day 1 (AM endpoints collected in the daily Diary), or the mean over the last 7 days of the Run-in period excluding the value on Day 1 (PM endpoints collected in the Diary).

If multiple evaluations occur on the same day, the average of these evaluations will be used for analysis excluding maximum post-baseline, abnormality, and outlier analyses.

For all shift from baseline analyses, percentages will be based on the number of subjects in each baseline category and also having at least one post-baseline measurement.

Values with "<" or ">" signs will be analyzed without the signs in tables and figures. In by-subject data listings, values will be reported as collected with the sign.

Dates in by-subject data listings will be displayed as yyyy-mm-dd (e.g., 2015-01-24). In general, by-subject data listings will be sorted by randomized treatment group and subject number.

3.1. Visit Windowing

Subjects do not always strictly adhere to the visit schedule timing in protocols. Therefore, the designation of visits (or timepoints) will generally be based on the actual day of evaluation relative to the date of first dose of study treatment (Day 1), rather than the nominal visit, for analyses conducted by visit.

Mutually exclusive visit windows containing no gaps will generally be utilized to assign visits for by visit analysis and will correspond to post-baseline visits specified in the protocol. Visits for analysis will be assigned by using a windowing scheme as described below.

The upper bound of the baseline visit window is Day 1 (up to the time of the first dose for assessments with time collected), and the lower bound of the first post-baseline visit window is Day 1 after the time of the first dose for assessments with time collected and Day 2 for assessments without time collected. For all other lower and upper bounds of visit windows, windows will end at the midpoint between scheduled visit timepoints, with the midpoint in the

latter visit window with the exception of the Week 16 visit for the efficacy endpoints, which will include Day 120 as the upper bound of the visit window.

If a subject's last observation is after the date of first dose of study treatment but prior to the first scheduled visit, data from an EW visit will be assigned to the first scheduled visit. If two or more evaluations occur in the same visit window, the evaluation closest to the target visit day will be selected for inclusion in the analysis. If multiple evaluations are equally close to the target visit day, then the latest evaluation will be selected for inclusion in the analysis. The target visit day for Week 1 and on is defined as the week number specified in the SoA of the protocol multiplied by 7 plus 1.

3.2. Standard Calculations

Standard calculations are described as follows:

Age:

If age is not collected on the electronic case report form (eCRF), age at screening will be calculated in years using the date of birth and the date of the Screening visit (Visit 1), rounded down to the nearest integer.

Duration:

Duration between two dates date1 and date2 will be calculated as follows:

date2 – date1 + 1, when expressed in days

(date2 – date1 + 1)/7, when expressed in weeks

(date2 – date1 + 1)/365.25, when expressed in years

Change/Percent Change from Baseline:

Change from baseline will be calculated as: Value at post-baseline visit – value at baseline

Percent change from baseline will be calculated as: (Change from baseline/value at baseline) * 100%

Note: Change from baseline summaries will only include subjects with a baseline value and at least one post-baseline value. If the value at baseline is 0, percent change from baseline will be missing.

Study Day:

Study day will be calculated as follows:

assessment date – date of the first dose of study treatment, where assessment date < date of the first dose of study treatment

assessment date – date of the first dose of study treatment + 1, where assessment date ≥ date of the first dose of study treatment

Date Imputation:

If complete dates are unavailable, July 2 (midpoint of a non-leap year) will be utilized in calculations in cases where both month and day are missing, and the 15th will be utilized in calculations in cases where only day is missing.

3.3. Considerations Related to the COVID-19 Pandemic

Screening of subjects for this study began in April 2019, and the last subject completed their participation in the study in August 2020. As such, a portion of this study's conduct was affected by the global Coronavirus Disease 2019 (COVID-19) pandemic.

At the time of finalization of this SAP in September 2020, several regulatory authorities have issued guidance related to trial conduct and associated methodological issues related to the effect of the COVID-19 pandemic, including, but not limited to, the Food and Drug Administration (FDA) (References: FDA Guidance on Conduct of Clinical Trials of Medical Products during COVID-19 Pandemic, March 2020, updated on July 2, 2020 and Statistical Considerations for Clinical Trials During the COVID-19 Public Health Emergency, June 2020) and European Medicines Agency (EMA) (References: Guidance on the Management of Clinical Trials during the COVID-19 (Coronavirus) pandemic, Version 3 (28/04/2020) and Points to consider on implications of Coronavirus disease (COVID-19) on methodological aspects of ongoing clinical trials, 26 June 2020).

Potential effects on conduct of this study related to the COVID-19 pandemic include, but are not limited to, the following: subject discontinuation of study treatment and/or withdrawal from study; inability or reduced ability for subjects to continue on IP due to disruptions in IP dispensing; missed visits or missed assessments; alternative procedures (e.g., telephone visit or virtual visit) used for collection of critical efficacy and/or safety assessments; and inability for sites to complete entry of data in the electronic case report form (eCRF) and/or collect data via other sources (e.g., eDiary).

A determination will be made at the end of study if additional statistical analyses are warranted as part of the final analysis, should participation of subjects or endpoint ascertainment be affected by the COVID-19 pandemic to a substantial degree. These additional analyses would be undertaken to enable an understanding of the potential impact of the COVID-19 pandemic on study results.

A non-exhaustive list of examples of such additional analyses is as follows:

- Analysis of compliance:
  - By subgroup of subjects who received direct IP shipment due to the COVID-19 pandemic versus those that did not
- Analyses of primary and select secondary efficacy endpoints:

  - By subgroup of subjects whose study participation was affected by the COVID-19 pandemic versus was not affected;
  - Excluding only data that is missing for reasons related to the COVID-19 pandemic (but including imputation of data missing for other reasons); and
  - By subgroup of subjects who had an alternative method of collection for an assessment contributing to a particular endpoint (e.g., SNOT-22 completed by subject at home via virtual visit) versus subjects who did not have an alternative method of collection

- Analysis of AEs:
  - Incidence of AEs by subgroup of subjects whose study participation was affected by the COVID-19 pandemic versus was not affected;
  - Incidence of AEs that were collected via an alternative method versus incidence of AEs that were not collected via an alternative method
- Analysis of laboratory values:
  - By subgroup of subjects whose study participation was affected by the COVID-19 pandemic versus was not affected;
  - Excluding values resulting from local collection due to reasons related to the COVID-19 pandemic; and
  - Including only values resulting from local collection due to reasons related to the COVID-19 pandemic

In addition, by-subject listings indicating subjects whose study participation was affected by the COVID-19 pandemic, subjects with missing data related to the COVID-19 pandemic, and/or subjects with an alternative method of collection for specific assessments may be provided.

4. ANALYSIS POPULATIONS

4.1. All Enrolled Population

The all enrolled population will include all subjects with a non-missing date of informed consent. The all enrolled population will be utilized for disposition summaries.

4.2. Run-in Period Population

The run-in period population will include all enrolled subjects who are not screen failures and enter the Run-in period. The run-in period population will be presented in disposition summaries.

4.3. Intent-to-Treat Population

The intent-to-treat (ITT) population will include all subjects who are randomized and receive at least 1 dose of study treatment, with subjects grouped according to randomized treatment. The ITT population will be utilized for efficacy analyses.

4.4. Safety Population

The safety population will include all subjects who receive at least 1 dose of study treatment, with subjects grouped according to their actual treatment. If GB001 was taken, based on kit and study treatment accountability data, subjects will be included in the GB001 40 mg group. The safety population will be utilized for safety and PK analyses.

5. STUDY PATIENTS

5.1. Disposition of Patients

The disposition of subjects will be summarized for the all enrolled population. The number and percent of subjects who are screened, screen failures, run-in period participants, completed run-in period, run-in period failures, and randomized in error, along with screen failure reasons, run-in period failure reasons, reasons not randomized, and eligibility or randomization criteria not met will be summarized overall. In addition, the number and percent of subjects who are randomized, treated, completed study treatment, discontinued study treatment, completed the study, and withdrew from the study, along with associated reasons, and in each analysis population, will be summarized by treatment group and overall.

Disposition data will also be presented in a by-subject data listing. In addition, a listing of the randomization scheme will be presented.

5.2. Protocol Deviations

Protocol deviations will be identified and reviewed on an ongoing basis by the study team and entered into a Clinical Trial Management System. Significant protocol deviations are defined as those that can affect efficacy and/or safety assessments, the safety or mental integrity of a subject, or the scientific value of the study. Protocol deviations will be classified as significant or non-significant and will be assigned a protocol deviation type prior to unblinding. The number and percent of subjects with at least 1 significant protocol deviation overall and for each protocol deviation type will be summarized for the ITT population. In addition, COVID-19 related and non-COVID-19 related significant protocol deviations will be summarized.

Protocol deviations will also be presented in a by-subject data listing. In addition, subjects who were randomized in error and subjects who were mis-stratified based on the presence or absence of NP (CRSwNP and CRSsNP) or on the presence or absence of comorbid asthma among CRSwNP subjects at randomization will be presented in a by-subject data listing.

6. DEMOGRAPHIC AND OTHER BASELINE CHARACTERISTICS

6.1. Demographic Characteristics

Demographic characteristics will be summarized for the ITT population for the following parameters:

- Age (years; continuous [cts]; ≥ 18-< 50, ≥ 50-< 65, ≥ 65)
- Sex (male, female)
- Race (White, Black or African American, American Indian or Alaska Native, Native Hawaiian or Other Pacific Islander, Asian, Other)
- Ethnicity (Hispanic or Latino, Not Hispanic or Latino, Not Reported)
- Height (cm; cts)
- Weight (kg; cts)
- Body mass index (BMI [kg/m^2; cts; < 18.5, ≥ 18.5-< 25, ≥ 25-< 30, ≥ 30])
- Country (Czech Republic, Ukraine, United States)
- Region (United States, Eastern Europe [Czech Republic, Ukraine])

Demographic characteristics will also be presented in a by-subject data listing.

6.2. Baseline Disease Characteristics

Baseline disease characteristics will be summarized for the ITT population for the following parameters:

- Disease status strata at randomization and actual (CRSwNP [overall, with comorbid asthma, without comorbid asthma], CRSsNP)
- CRS/NP Surgical History
  - Any prior surgery (cts)
  - Prior functional endoscopic sinus surgery (number of surgeries and years since most recent surgery; cts; 0, 1, 2, ≥ 3 for number of surgeries)
  - Prior NP surgery (number of surgeries and years since most recent surgery; cts; 0, 1, 2, ≥ 3 for number of surgeries)
  - Prior stent placement (number of surgeries and years since most recent surgery; cts; 0, 1, ≥ 2 for number of surgeries)
  - Prior nasal septum surgery (number of surgeries and years since most recent surgery; cts; 0, 1, ≥ 2 for number of surgeries)
  - Prior rhinoplasty surgery (number of surgeries and years since most recent surgery; cts; 0, ≥ 1 for number of surgeries)

- CRS/NP History
  - CRS duration (years; cts)
  - History of NP (never had, ongoing, past; NP duration for CRSwNP subjects [years; cts])
  - Comorbid asthma (yes, no)
  - Fungal sinusitis current or prior history (yes, no)
  - Inadequate response to vaccines (yes, no)
  - Nasal blockade/obstruction/congestion (yes, no)
  - Nasal discharge (yes, no)
  - Reduction or loss of smell (yes, no)
  - Facial pain/pressure (yes, no)
  - Loss of sleep (yes, no)
  - Sinus CT scan in past 2 years (yes, no)
  - Sinus X-ray in past 2 years (yes, no)
  - Prior anti-fungal treatment (yes, no)
  - Prior saline nasal irrigations (yes, no)
  - Chronic antibiotic use in the past (yes, no)
  - Chronic antibiotic use currently (yes, no)
- Polyps present (based on visual exam by investigator) in right/left nostril for CRSwNP subjects (yes, no)
- Smoking history (never, former, current; cigarette pack-years for former and current smokers)
- Alcohol history (never, former, current)

Baseline disease characteristics will also be presented in by-subject data listings.

6.3. Atopic/Allergic Conditions

Atopic/allergic conditions will be summarized for the ITT population for the following parameters:

- Chronic Rhinosinusitis, Asthma, Seasonal Allergic Rhinitis, Perennial Allergic Rhinitis, Chronic Spontaneous Urticaria, Eosinophilic granulomatosis with polyangiitis (Churg Strauss Syndrome) (none, ongoing, past)

Atopic/allergic conditions will also be presented in a by-subject data listing.

6.4. Efficacy-Related Baseline Characteristics

Efficacy-related baseline characteristics will be summarized for the ITT population for the following parameters:

- SNOT-22 total score and 5 domain (Nasal, Ear, Sleep, General and Practical, and Emotional) scores (cts; ≤ Median, > Median of ITT population for total score)
- Bilateral Endoscopic NPS at Screening for CRSwNP subjects (cts; 4, 5 - 6, 7 - 8)
- Bilateral Endoscopic NPS at Baseline for CRSwNP subjects (cts; ≤ 4, 5 - 6, 7 - 8)
- TSS (AM, PM; cts; ≤ Median, > Median of ITT population for AM TSS)
- NC score (AM, PM; cts; ≤ Median, > Median of ITT population for AM NC)
- Anterior rhinorrhea score (AM, PM; cts)
- Posterior rhinorrhea score (AM, PM; cts)
- Loss of sense of smell score (AM, PM; cts)
- UPSIT score (cts; Anosmia = 0 – 18, Severe Microsmia = 19 – 25, Moderate Microsmia = 26 – 30, Mild Microsmia = 31 – 34, Normal = 35 - 40)
- VAS score (cts; mild, moderate, severe, using categories defined in Section 8.3.3.1)
- Lund-Mackay score (cts)
- Volume of the air (mL; cts)
- Volume of mucosa (mL; cts)
- Percentage of maxillary sinus volume occupied by disease on CT scan (cts)
- Thickness of lateral wall (mm; cts)
- ACQ-5 score for asthma subjects (cts; < 1.5, ≥ 1.5)

Efficacy-related baseline characteristics will also be presented in a by-subject data listing.

6.5. Biomarker Baseline Characteristics

Biomarker baseline characteristics will be summarized for the ITT population for the following parameters:

- Blood eosinophils at Screening (10^9/L; cts; <0.25, ≥ 0.25)
- Blood eosinophils at Baseline (10^9/L; cts; <0.25, ≥ 0.25)
- Screening and Baseline blood eosinophils (10^9/L; ≥ 0.25 and < 0.25)
- Maximum blood eosinophils value prior to randomization visit for CRSsNP subjects (10^9/L; cts; ≥ 0.25, ≥ 0.30, ≥ 1.50)
- Fractional exhaled nitric oxide (FeNO) at Baseline (ppb; cts; < 25, ≥ 25)

- Chloride at Baseline (mmol/L; cts; ≤ Quartile 1, > Quartile 1 - ≤ Quartile 2, > Quartile 2 - ≤ Quartile 3, > Quartile 3, where quartiles are based on the ITT population)

Biomarker baseline characteristics will also be presented in a by-subject data listing.

6.6. Medical History

Verbatim targeted and general medical history terms will be coded to a SOC and PT using the Medical Dictionary for Regulatory Activities (MedDRA) Version 21.0 or later. Counting will be at the subject level for each level of summarization (e.g., any medical history, SOC, and PT), with subjects experiencing more than one occurrence of a SOC or PT counted only once. Ongoing and past medical history will be summarized for the ITT population.

Medical history will also be presented in a by-subject data listing.

6.7. Prior Medications

Verbatim prior medication terms will be coded to a drug class (anatomical therapeutic chemical [ATC2]) and PT using World Health Organization (WHO)Drug Global B3 March 2018 or later. Prior medications are defined as all medications that started prior to the date of first dose of study treatment. In general, if it is not clear whether a medication is prior due to missing or incomplete medication start date, the medication will be considered to be prior unless the non-missing portions of the start date indicate otherwise. Counting will be at the subject level for each level of summarization (e.g., any medication, ATC2, and PT), with subjects receiving more than one medication counted only once.

Prior medications will be summarized for the safety population.

In addition, MFNS use prior to informed consent, MFNS use between screening and randomization, prior ICS use overall and separately for between screening and randomization, further broken down by ICS alone and ICS + long-acting beta-agonist (LABA) and/or long-acting muscarinic antagonist (LAMA), and prior systemic corticosteroid use will be summarized for the ITT population.

6.8. Concomitant Medications

Verbatim concomitant medication terms will be coded to a drug class (ATC2) and PT using WHODrug Global B3 March 2018 or later.

Concomitant medications are defined as all medications that started on or after the date of first dose of study treatment or that started prior to the date of first dose of study treatment and stopped on or after the date of first dose of study treatment or were ongoing. In general, if it is not clear whether a medication is concomitant due to missing or incomplete medication start

and/or end dates, the medication will be considered to be concomitant unless the non-missing portions of the start and end dates indicate otherwise.

A medication may be considered as both prior and concomitant, i.e. prior and concomitant medications are not mutually exclusive. Counting will be at the subject level for each level of summarization (e.g., any medication, ATC2, and PT), with subjects receiving more than one medication counted only once.

Concomitant medications will be summarized for the safety population. In addition, concomitant MFNS use will be summarized for the ITT population.

Prior and concomitant medications will also be presented in a by-subject data listing and will include an indicator, identifying each medication as prior and/or concomitant.

Concurrent procedures will be presented in a by-subject data listing.

7. EXTENT OF EXPOSURE, COMPLIANCE, AND TIME ON STUDY

7.1. Extent of Exposure

Duration of study treatment (days, weeks, and subject-years) will be summarized for the safety population using descriptive statistics and categorically (every 4 weeks). Duration of study treatment will also be presented in a by-subject data listing.

7.2. Compliance

Compliance with study treatment will be assessed using the following formula:

Treatment compliance (%) = (# of actual tablets taken/# of expected tablets) x 100,
where # of actual tablets taken = (# of tablets dispensed - # of tablets returned) and # of expected tablets is based on actual duration of study treatment

Compliance will be summarized for the safety population using descriptive statistics and categorically (< 80, ≥ 80-100, > 100-< 120, ≥ 120).

Compliance with study treatment will also be presented in a by-subject data listing. In addition, IP accountability including kit information will be presented in a by-subject data listing.

7.3. Time on Study

Time on study (days and weeks, and subject-years) will be summarized for the safety population using descriptive statistics and categorically (every 4 weeks). The latest assessment date for each subject will be used in the calculation of time on study. Time on study will also be presented in a by-subject data listing.

8. EFFICACY EVALUATION

8.1. Overview of Efficacy Analysis Issues

8.1.1. Handling of Dropouts or Missing Data

A variety of methodological approaches will be applied for handling subjects with missing data; full details are provided in Section 8.3.

8.1.2. Multicenter Studies

Data from all countries and sites will be pooled for the purpose of analyses.

8.2. Efficacy Endpoints

Table 1 provides a summary of the statistical models/methods of primary and sensitivity analyses to be used for the primary, secondary, and select exploratory endpoints. For all efficacy analyses, disease status stratification factor values used for randomization (CRSwNP with comorbid asthma, CRSwNP without comorbid asthma, and CRSsNP) and region (United States, Eastern Europe) will be used as covariates and results will be summarized for the ITT population, unless otherwise stated.

Table 1: Efficacy Endpoints and Analysis Methods

Efficacy Endpoints	Missing Data Threshold[a]	MMRM		ANCOVA		Cox Proportional Hazards	Negative Binomial Regression[b]	Logistic Regression
		without MI	with MI	without MI	with MI			
Primary	< 10%	P		S				
Change from baseline to Week 16 in SNOT-22 total score	≥ 10%	S	P		**S**			
Secondary								
Change from baseline to Week 16 in Lund-Mackay score	< 10%			P				
	≥ 10%			S	P			
Change from baseline to Week 16 in NPS[c]	< 10%	P		S				
	≥ 10%	S	P		S			
Time to first response (≥ 1 point improvement from baseline) in NPS[c]						P		
Change from baseline to Week 16 in AM NC score	< 10%	P		S				
	≥ 10%	S	P		S			
Change from baseline to Week 16 in AM TSS	< 10%	P		S				
	≥ 10%	S	P		S			
Change from baseline to Week 16 in UPSIT score	< 10%			P				
	≥ 10%			S	P			
Time to first CRS exacerbation						P		
Exploratory								
Change from baseline to Week 16 in VAS score				P				
Annualized rate of CRS exacerbations							P	
Change from baseline to Week 16 in percentage of maxillary sinus volume occupied by disease on CT scan				P				

Efficacy Endpoints	Missing Data Threshold[a]	MMRM		ANCOVA		Cox Proportional Hazards	Negative Binomial Regression[b]	Logistic Regression
		without MI	with MI	without MI	with MI			
Proportion of subjects with antibiotic use due to worsening CRS by Week 16								P
Change from baseline to Week 16 in FeNO and nNO				P				
Percent change from baseline to Week 16 in FeNO and nNO				P				
Additional								
Change from baseline in SNOT-22 total score including all data	< 10%	P						
	≥ 10%		P					
Proportion of subjects with a ≥ 9 point decrease from baseline at Week 16 in SNOT-22 total score								P
Change from baseline in Lund-Mackay score including all data	< 10%			P				
	≥ 10%				P			
Change from baseline in NPS[c] including all data	< 10%	P						
	≥ 10%		P					
Proportion of subjects with a reduction from baseline of ≥ 1 in NPS[c] at Weeks 8 and 16								P
Change from baseline to Week 16 in PM NC score		P						
Change from baseline to Week 16 in PM TSS		P						

Efficacy Endpoints	Missing Data Threshold[a]	MMRM		ANCOVA		Cox Proportional Hazards	Negative Binomial Regression[b]	Logistic Regression
		without MI	with MI	without MI	with MI			
Change from baseline to Week 16 in AM and PM anterior rhinorrhea, posterior rhinorrhea, and loss of sense of smell scores		P						
Change from baseline to Week 16 in volume of the air, volume of mucosa, and thickness of lateral wall				P				
Change from baseline to Week 16 in ACQ-5 score				P				

ACQ-5 = asthma control questionnaire; AM = morning; ANCOVA = analysis of covariance; CRS = chronic rhinosinusitis; CRSwNP = chronic rhinosinusitis with nasal polyps; CT = computed tomography; FeNO = fractional exhaled nitric oxide; ITT =intent-to-treat; MMRM = mixed-effects model with repeated measures; MI = multiple imputation; NC = nasal congestion; nNO = nasal nitric oxide; NPS = nasal polyp score; P = primary; PM = evening; S = sensitivity; SNOT-22 = Sino-Nasal Outcome Test; TSS = total symptom score; UPSIT = University of Pennsylvania Smell Identification Test; VAS = visual analog scale

[a] The analysis approach depends on the amount of missing data for the endpoint in the ITT population.

[b] If the distribution of severe exacerbation data is underdispersed, or if the negative binomial regression model fails to converge, a Poisson regression model will be used instead.

[c] Assessed in subjects in the randomized CRSwNP stratum.

8.3. Analysis Methods

8.3.1. Primary Efficacy Analyses

The primary endpoint is the change from baseline to Week 16 in SNOT-22 total score.

SNOT-22 is a validated questionnaire to assess the impact of CRS on quality of life (QOL) and utilizes a 2-week recall period. It is a 22-item outcome measure on a 5-point category scale applicable to sinonasal conditions and surgical treatments. The questionnaire consists of 22 items across 5 domains: Nasal, Ear, Sleep, General and Practical, and Emotional. A scale ranging from 0 (no problem) to 5 (problem as bad as it can be) is used to respond to each item in the questionnaire. The total scores range from 0 to 110 with higher total scores implying greater impact of CRS on QOL. A Minimal Clinically Important Difference (MCID) is available: 8.90 (Hopkins, 2009).

Total and domain scores will be derived by taking the sum of the non-missing item responses at each visit.

If the number of subjects overall with a missing change from baseline to Week 16 SNOT-22 total score is < 10%:

- The primary analysis will be mixed-effects model with repeated measures (MMRM) without multiple imputation (MI).
- As a sensitivity analysis, analysis of covariance (ANCOVA) without MI will also be performed.

If the number of subjects overall with a missing change from baseline to Week 16 SNOT-22 total score is ≥ 10%:

- The primary analysis will be MMRM with MI.
- As sensitivity analyses, MMRM without MI and ANCOVA with MI will also be performed.

8.3.1.1. MMRM without MI

MMRM will be performed using a restricted maximum-likelihood (REML)-based approach. The MMRM model will include change from baseline scores through Week 16 as the response variable; the fixed, categorical covariates of treatment group, disease status stratification factor values used for randomization (CRSwNP with comorbid asthma, CRSwNP without comorbid asthma, and CRSsNP), region (United States, Eastern Europe), visit, and treatment group by visit interaction; and the continuous, fixed covariates of baseline SNOT-22 total score and baseline SNOT-22 total score by visit interaction. An unstructured covariance structure will be used to model within-subject error.

If the analysis fails to converge due to the complexity of model specification, the following will be considered to enable model convergence: 1) use of maximum likelihood estimation instead of REML; and/or 2) a compound symmetric or first-order autoregressive covariance structure will

be utilized, based on the covariance structure converging to better fit as determined based on Akaike's information criterion (Akaike, 1973); and/or 3) reduced models. In the case of 3), if the full model does not converge after using maximum likelihood estimation or different covariance structures, reduced models will be fit using REML and an unstructured covariance structure. The reduced model will include treatment group, visit, and treatment group by visit interaction.

Parameters will be estimated using REML with the Newton-Raphson algorithm. Treatment comparisons will be derived from the MMRM. The number and percent of subjects contributing to the MMRM analysis and descriptive statistics along with the LS means, SEs, and corresponding 95% CIs for each treatment group will be provided. The difference in LS means, SE, corresponding 95% CI, and p-value with Kenward-Roger adjustment (Kenward, 1997) will also be provided for comparison for the GB001 40 mg group versus placebo. Results will be presented for the total score only, for change from baseline to Week 16 and separately for change from baseline to all visits through Week 16 (excluding descriptive statistics).

A bar graph of the LS mean change from baseline to Week 16 by treatment group, including 95% CIs, for the total score will be provided. A figure of LS mean change from baseline in the total score, including 95% CIs, by visit will also be provided. The figures reporting model-based results will use the primary analysis approach based on the amount of missing data as described in Table 1. Lastly, a bar graph of the mean change from baseline to Week 16 by treatment group, including SE bars, for each domain score will be provided using as observed data.

Actual, change from baseline, and percent change from baseline values in total and domain scores will be descriptively summarized by visit using as observed data.

SNOT-22 total and domain scores will also be presented in a by-subject data listing.

8.3.1.2. MMRM with MI

Missing SNOT-22 total scores at post-baseline visits will be imputed using an MI approach. Missing SNOT-22 total scores at baseline will not be imputed.

The following steps will be performed:

- Step 1: Intermittent missing values will be imputed using the Markov Chain Monte Carlo (MCMC) method (Schafer, 1997), using a noninformative prior (Jeffreys) for the prior information for the means and covariances and expectation-maximization (EM) for the initial mean and covariance estimates.
- Step 2: Monotone missing values will then be imputed using a linear regression model (Allison, 2005) including the imputed intermittent missing values in step 1. A monotone missing pattern is such that in the event that a variable is missing it is implied that all subsequent variables are also missing (e.g., once a subject withdraws from study, no additional data is expected for that subject).
- For each step, MI will be performed within each treatment group with covariates for disease status stratification factor values used for randomization (CRSwNP with

comorbid asthma, CRSwNP without comorbid asthma, and CRSsNP), region (United States, Eastern Europe), and non-missing values for baseline SNOT-22 total score and Weeks 4, 8, 12, and 16 SNOT-22 total scores. Should convergence issues occur due to small cell size for the categorical covariates at either step, the categorical variable(s) causing convergence issues will be removed from the model.

- Missing change from baseline and percent change from baseline in SNOT-22 total scores will be computed from the imputed SNOT-22 total scores.
- The analysis of the primary endpoint as described in Section 8.3.1 will be performed for each complete, imputed dataset.
- For step 1, fifty imputations will be performed. For step 2, 1 imputation will be performed if 50 imputations were performed for step 1. Fifty imputations will be performed for step 2 if step 1 was skipped (e.g., for assessments only collected once post-baseline or for endpoints without intermittent missing values).
- Results of the MMRM on the multiple imputed data sets will be combined to generate an overall estimate and associated variance using Rubin's rules (Rubin, 1987).

Results will be presented as described in Section 8.3.1.1 but using the MI data and the number and percent of subjects with an MI value or an observed value for each treatment group will also be provided. In addition, actual, change from baseline, and percent change from baseline values in total and domain scores will be descriptively summarized by visit using MI data.

8.3.1.3. ANCOVA without MI

For ANCOVA without MI, the models will include covariates for treatment group, disease status stratification factor values used for randomization (CRSwNP with comorbid asthma, CRSwNP without comorbid asthma, and CRSsNP), region (United States, Eastern Europe), and baseline SNOT-22 total score.

Descriptive statistics for change from baseline to Week 16 along with the LS means, SEs, and corresponding 95% CIs for each treatment group will be provided. The difference in LS means, SE, corresponding 95% CI, and p-value for the GB001 40 mg group versus placebo will also be constructed from the ANCOVA model. Results will be presented for the total score only.

8.3.1.4. ANCOVA with MI

The imputed data described in Section 8.3.1.2 will be used for the ANCOVA with MI analyses. Covariate adjustment will be the same as in Section 8.3.1.3.

Results will be presented as described in Section 8.3.1.3 but using the MI data.

8.3.1.5. Additional Analyses of the Primary Endpoint

The proportion of subjects with a ≥ 9 point decrease from baseline at Week 16 in SNOT-22 total score will be analyzed in the subset of the ITT population who have a non-missing baseline and at least one post-baseline value. Logistic regression modeling will be used to compare the

GB001 40 mg group with placebo using as observed data with covariates for treatment group, disease status stratification factor values used for randomization (CRSwNP with comorbid asthma, CRSwNP without comorbid asthma, and CRSsNP), region (United States, Eastern Europe), and baseline SNOT-22 total score.

The number and proportion of subjects with a ≥ 9 point decrease from baseline at Week 16 in SNOT-22 total score and corresponding 95% Wilson (Score) CIs (Wilson, 1927) will be summarized by treatment group. The OR, corresponding asymptotic 95% CI along with the absolute difference in proportions and corresponding 95% CI, and p-value will be summarized for the GB001 40 mg group versus placebo. The 95% CI for the absolute difference will be Newcombe continuity-corrected CI (Newcombe, 1998). The percentage reduction in odds and corresponding 95% CI will also be provided.

The number and proportion of subjects with a ≥ 9 point decrease from baseline in SNOT-22 total score will also be summarized by visit as described above at Week 16.

8.3.2. Secondary Efficacy Analyses

The secondary efficacy endpoints are listed in Table 1 and will be analyzed as described below.

8.3.2.1. Change from Baseline to Week 16 Endpoints

8.3.2.1.1. Lund-Mackay and UPSIT Scores

Lund-Mackay scores are based on the centralized imaging data assessments and scored by blinded central reading at the imaging core laboratory. The Lund-Mackay system is based on localization with points given for degree of opacification: 0 = normal, 1 = partial opacification, 2 = total opacification. These points are then applied to each sinus (maxillary, anterior ethmoid, posterior ethmoid, sphenoid, and frontal sinus) on each side. The osteomeatal complex on each side is graded as 0 = not occluded, or 2 = occluded, deriving a maximum score of 12 per side, for a total score ranging from 0 to 24 (Lund, 1993). Of note, an aplastic (absent) frontal sinus receives a score of 0.

The UPSIT test consists of four booklets, each containing 10 odorants with one odorant per page (Doty, 1984). The test-time is about 15 minutes. The stimuli are embedded in 10–50 μm diameter plastic microcapsules on brown strips at the bottom of each page. Above each odorant strip is a multiple-choice question with four alternative words to describe the odor. The subject is asked to release the odorant by rubbing the brown-strip with the tip of a pencil and to indicate which of four words best describes the odor. An UPSIT score will only be derived if all 40 responses are non-missing. A higher UPSIT score indicates better olfaction. Anosmia categories for the UPSIT score are defined as:

- Anosmia = 0 - 18
- Severe Microsmia = 19 – 25
- Moderate Microsmia = 26 - 30
- Mild Microsmia = 31 – 34

- Normal = 35 – 40

For the change from baseline to Week 16 in Lund-Mackay and UPSIT scores, the primary analysis will be ANCOVA without MI if the number of subjects overall with a missing change from baseline value to Week 16 is < 10% for the respective endpoint. If the number of subjects overall with a missing change from baseline value to Week 16 is ≥ 10% for the respective endpoint, the primary analysis will be ANCOVA with MI, with an MI procedure similar to that described in Section 8.3.1.2, and ANCOVA without MI will be performed as a sensitivity analysis. For the Lund-Mackay and UPSIT score multiple imputation, step 1 for imputing intermittent missing values will be skipped since Lund-Mackay and UPSIT scores can only be monotone missing, as they are only collected at Weeks 0 and 16. The ANCOVA models will be analogous to those described in Sections 8.3.1.3 and 8.3.1.4, with the baseline score specific to each endpoint included as a covariate.

Bar graphs of the LS mean change from baseline to Week 16 by treatment group, including 95% CIs, for Lund-Mackay and UPSIT scores will be provided, using the primary analysis approach for each endpoint based on the amount of missing data as described in Table 1.

The number and proportion of subjects in each anosmia category will be summarized by visit, including baseline, using as observed data.

CT scan and UPSIT data will also be presented in by-subject data listings.

8.3.2.1.2. NPS and Symptom Scores

The bilateral endoscopic NPS is the sum of the right and left nostril scores, as evaluated by means of blinded, centrally read nasal endoscopy and ranges from 0-8. NP is graded based on polyp size: 0 = No polyps, 1 = Small polyps in the middle meatus not reaching below the inferior border of the middle turbinate, 2 = Polyps reaching below the lower border of the middle turbinate, 3 = Large polyps reaching the lower border of the inferior turbinate or polyps medial to the middle turbinate, and 4 = Large polyps causing complete obstruction of the inferior nasal cavity.

On a daily basis, from Visit 1 and throughout the study, the subject will use a Diary to:
Respond to the AM and PM individual rhinosinusitis symptom questions using a 0–3 categorical scale (where 0 = no symptoms, 1 = mild symptoms, 2 = moderate symptoms and 3 = severe symptoms) for the following symptoms:

- congestion and/or obstruction (i.e. NC score)
- anterior rhinorrhea (runny nose)
- posterior rhinorrhea (post-nasal drip)
- loss of sense of smell

The above symptom scores will be recorded directly in the daily Diary. For post-baseline visits, the average of the daily symptom scores over the prior 4 weeks will be used to determine the value for that visit, if at least 50% of the scores were completed over the prior 4 weeks, and will

be calculated separately for AM and PM scores. The TSS is the sum of the scores from the 4 symptom categories (congestion and/or obstruction, anterior rhinorrhea, posterior rhinorrhea, loss of sense of smell), ranges from 0–12, and will be calculated separately for AM and PM scores.

For the change from baseline to Week 16 in NPS, AM NC score, and AM TSS, the primary analysis approach will depend on the amount of missing data for each endpoint, based on the same criteria for the primary endpoint as described in Section 8.3.1. The MMRM models will be similar to those described in Sections 8.3.1.1 and 8.3.1.2, with the baseline score specific to each endpoint included as a covariate. For NC and TSS MI, Week 20 (Follow-up) values will be included as a covariate in the multiple imputation. For monotone regression, MI is performed sequentially, meaning only the data from weeks prior to the study week with missing data will inform the MI. Therefore, values at Week 20 (Follow-up) will not inform prior weeks. If the analyses fail to converge due to the complexity of model specification, the approaches described in Section 8.3.1.1 will be implemented. The ANCOVA models will be analogous to those described in Sections 8.3.1.3 and 8.3.1.4, with the baseline score specific to each endpoint included as a covariate. For the NPS endpoint, the analysis will be conducted within the randomized CRSwNP stratum.

For the change from baseline to Week 16 and change from baseline to all visits through Week 16 in PM NC score, PM TSS, AM and PM anterior rhinorrhea, AM and PM posterior rhinorrhea, and AM and PM loss of sense of smell, the MMRM without MI models described in Section 8.3.1.1 will be used with the baseline score specific to each endpoint included as a covariate.

Actual, change from baseline, and percent change from baseline values in NPS, AM NC score, and AM TSS will be descriptively summarized by visit using MI data if MI is performed. Actual, change, and percent change from baseline values in NPS, NC scores, anterior rhinorrhea scores, posterior rhinorrhea scores, loss of sense of smell scores, and TSS for AM and PM will be descriptively summarized by visit using as observed data.

Bar graphs of the LS mean change from baseline to Week 16 by treatment group, including 95% CIs, for NPS, AM and PM NC scores, AM and PM anterior rhinorrhea, AM and PM posterior rhinorrhea, AM and PM loss of sense of smell, and AM and PM TSS at Week 16 will be provided. Figures of LS mean change from baseline, including 95% CIs, by visit for each of these endpoints will also be provided. The figures will use the primary analysis approach for each endpoint based on the amount of missing data as described in Table 1.

Nasal endoscopy and diary data will also be presented in by-subject data listings.

8.3.2.2. Time to Event Endpoints

Time to first response in NPS is defined as the time from the date of the first dose of study treatment to the first date with ≥ 1 point improvement in NPS from baseline for subjects in the randomized CRSwNP stratum. Subjects in the randomized CRSwNP stratum who do not experience ≥ 1 point improvement in NPS from baseline will be censored at their last post-baseline NPS assessment, or on Day 1 if no baseline NPS assessment or no post-baseline NPS assessment was performed.

CRS exacerbation is defined as deterioration of CRS symptoms requiring treatment with an antibiotic, an anti-inflammatory drug, or a symptom reliever; an Emergency Department visit; or hospitalization (Kuiper, 2018). Time to first CRS exacerbation is defined as the time from the date of the first dose of study treatment to the first date of treatment with an antibiotic, an anti-inflammatory drug, or a symptom reliever, the date of Emergency Department visit, or the date of hospitalization, whichever occurs first. Subjects without a CRS exacerbation will be censored at their last post-baseline CRS exacerbation assessment, or on Day 1 if no post-baseline CRS exacerbation assessment was performed.

For the primary analysis of both time to event endpoints, all data through Week 16 will be included; a sensitivity analysis of both time to event endpoints will be performed including all data.

The number and proportion of subjects with a response in NPS will be summarized for subjects in the randomized CRSwNP stratum. Time to first response in NPS will be compared between treatment groups using a Cox proportional hazards model and will include covariates for treatment group, disease status stratification factor values used for randomization (CRSwNP with comorbid asthma, CRSwNP without comorbid asthma, and CRSsNP), region (United States, Eastern Europe), and baseline NPS.

The number and proportion of subjects with CRS exacerbations will be summarized. Time to first CRS exacerbation will be analyzed using a Cox proportional hazards model and will include covariates for treatment group, disease status stratification factor values used for randomization (CRSwNP with comorbid asthma, CRSwNP without comorbid asthma, and CRSsNP), and region (United States, Eastern Europe).

The exact method will be used to handle ties in event times.

For both time to event endpoints, estimates of the hazard ratio for the GB001 40 mg group versus placebo, corresponding 95% CIs, and p-values will be provided from the model. Percentage reductions and corresponding 95% CIs will also be provided. Kaplan-Meier summaries of the proportion of subjects with response in NPS and separately for the proportion of subjects with a CRS exacerbation over time will also be presented and displayed graphically. The median event time (and other quartiles), if estimable, and corresponding 95% CI based on the Brookmeyer and Crowley method (Brookmeyer, 1982) using the log-log transformation (Kalbfleisch, 1980) will be provided for each treatment group and endpoint.

An additional analysis of the proportion of subjects with a reduction from baseline of ≥ 1 in NPS at Weeks 8 and 16 in the randomized CRSwNP stratum will be analyzed using a logistic regression model to compare the GB001 40 mg group with placebo, using the approach described in Section 8.3.1.5 with NPS score at baseline included as a covariate.

Time to event data will also be presented in by-subject data listings.

8.3.3. Exploratory and Additional Analyses

The exploratory and additional endpoints are listed in Table 1 and will be analyzed as described below.

8.3.3.1. Change from Baseline to Week 16 Endpoints

The VAS for rhinosinusitis is used to evaluate the total severity and is only validated in adult CRS to date (Fokkens, 2007). The subject is asked to indicate on a VAS the answer to the question: "How troublesome are your symptoms of rhinosinusitis?" The VAS ranks from 0 (Not troublesome) to 10 (Worst thinkable troublesome) (as defined in Protocol Section 10.7). The disease can be divided into Mild, Moderate, and Severe based on total severity VAS score (0 to 10 cm):

- Mild = VAS 0-3
- Moderate = VAS > 3–7
- Severe = VAS > 7–10

Percentage of maxillary sinus volume occupied by disease on CT scan, volume of the air (mL), volume of mucosa (mL), and thickness of lateral wall (mm) will be calculated each for the left and right sides and averaged for analysis, using the non-missing values at each visit.

Change from baseline to Week 16 in VAS score and percentage of maxillary sinus volume occupied by disease on CT scan, volume of the air, volume of mucosa, and thickness of lateral wall will be analyzed using ANCOVA models without MI. The ANCOVA models will be analogous to those described in Section 8.3.1.3, with the baseline value specific to each endpoint included as a covariate. The models will include covariates for treatment group, disease status stratification factor values used for randomization (CRSwNP with comorbid asthma, CRSwNP without comorbid asthma, and CRSsNP), region (United States, Eastern Europe), and baseline value. LS means, SEs, and corresponding 95% CIs for each treatment group, and differences in LS means, SEs, corresponding 95% CIs, and p-values for the GB001 40 mg group versus placebo will be provided.

Actual values and change from baseline values in VAS scores will be descriptively summarized by visit using as observed data.

VAS score, percentage of maxillary sinus volume occupied by disease on CT scan, volume of the air, volume of mucosa, and thickness of lateral wall data will be presented in by-subject data listings.

8.3.3.2. Change from Baseline to Week 16 Including All Data Endpoints

Additional analyses of change from baseline to Week 16 in SNOT-22 total score, NPS, and Lund-Mackay score will be performed including all data. For these analyses, analysis visit windows will be the same as the primary analysis, except no upper bound will be applied to the Week 16 analysis visit window. The primary analysis approach will depend on the amount of missing data for each endpoint, based on the same criteria for the primary endpoint as described

in Section 8.3.1. The MMRM models for SNOT-22 total score and NPS will be similar to those described in Sections 8.3.1.1 and 8.3.1.2, with the baseline score specific to each endpoint included as a covariate. The ANCOVA model for Lund-Mackay score will be analogous to those described in Sections 8.3.1.3 and 8.3.1.4, with the baseline score included as a covariate. For the NPS endpoint, the analysis will be conducted within the randomized CRSwNP stratum.

8.3.3.3. Annualized Rate of CRS Exacerbations

The date of onset of a CRS exacerbation will be considered to be the first date of treatment with an antibiotic, an anti-inflammatory drug, or a symptom reliever, the date of Emergency Department visit, or the date of hospitalization, whichever occurs first. The end date of a CRS exacerbation will be considered to be the last date of treatment with an antibiotic, an anti-inflammatory drug, or a symptom reliever, or 6 days after the date of onset (for a total duration of 7 days), whichever is later. Courses of treatment separated by 7 or more days will be counted as separate exacerbations.

The number and proportion of subjects with 0, 1, and ≥ 2 CRS exacerbations and the adjusted annualized rate of CRS exacerbations will be summarized by treatment group. The unadjusted annualized rate of CRS exacerbations will be calculated as the total number of CRS exacerbations divided by the total number of subject-years of follow-up and will also be summarized by treatment group.

The total number of CRS exacerbations will be analyzed using a negative binomial regression model, with the logarithmic transformation of follow-up time as the offset parameter. The follow-up time is defined as the subject's time on study through Day 120. The model will include terms for treatment group, disease status stratification factor values used for randomization (CRSwNP with comorbid asthma, CRSwNP without comorbid asthma, and CRSsNP), and region (United States, Eastern Europe). If the distribution of CRS exacerbation data is underdispersed, or if the negative binomial regression model fails to converge, a Poisson regression model will be used instead of the negative binomial regression model.

The rate ratio for the GB001 40 mg group versus placebo, corresponding 95% CI, and p-value will be provided from the model. The percentage reduction and corresponding 95% CI will also be provided.

Lastly, the subject CRS exacerbation rate will be calculated as the number of CRS exacerbations for each subject divided by the number of years followed in the study for that subject and will be summarized by treatment group.

Analyses of annualized rate of CRS exacerbations may also be presented by various time periods (Baseline through Week 8, Week 8 through Week 16, and Baseline through Week 20 [Follow-up]) in order to elucidate onset and durability of effect over time.

CRS exacerbation data will also be presented in by-subject data listings.

8.3.3.4. Proportion of Subjects with Antibiotic Use Due to Worsening CRS by Week 16

The proportion of subjects with antibiotic use due to worsening of CRS by Week 16 will be analyzed using a logistic regression model to compare the GB001 40 mg group with placebo, using the approach described in Section 8.3.1.5, excluding baseline SNOT-22 total score as a covariate in the model.

8.3.3.5. Other Endpoints

8.3.3.5.1. Nitric Oxide

Nitric oxide will be measured both by oral (FeNO) and nasal (nasal nitric oxide [nNO]) methods. nNO data will be analyzed for the same nostril, either left or right, per subject at all visits. The primary nNO analysis will include subjects with a change from baseline to Week 16 value for a single nostril and subjects with change from baseline to Week 16 values for both nostrils where the nostril with the lower (i.e., worse) value at baseline is used. In the case of ties at baseline, the nostril with the lower (i.e., worse) change from baseline at Week 16 will be selected for summary. A sensitivity nNO analysis will include subjects with a change from baseline to Week 16 value for a single nostril and subjects with change from baseline to Week 16 values for both nostrils where the nostril with the higher (i.e., better) value at baseline is used. In the case of ties at baseline, the nostril with the higher (i.e., better) change from baseline at Week 16 will be selected for summary.

Change from baseline to Week 16 in both FeNO and nNO and percent change from baseline to Week 16 in both FeNO and nNO will be analyzed using ANCOVA models without imputation. LS means, SEs, and corresponding 95% CIs for each treatment group, and differences in LS means, SEs, corresponding 95% CIs, and p-values for the GB001 40 mg group versus placebo will be provided. The models will include covariates for treatment group and baseline FeNO or nNO. Analyses may be conducted using log transformed data in the case of non-normality.

Actual, change from baseline, and percent change from baseline values in FeNO and nNO will be descriptively summarized by visit using as observed data.

Nitric oxide data will also be presented in a by-subject data listing.

8.3.3.5.2. ACQ-5 Score

The ACQ-5 is a five-item questionnaire which has been developed as a measure of the subject's asthma control that can be quickly and easily completed (Juniper, 2005). The response options for each of these questions consists of a zero (no impairment/limitation) to 6 (total impairment/limitation) scale with the total score ranging from 0 to 30. ACQ-5 score is the average of the 5 non-missing items and will be calculated at each visit for each subject. The MCID is 0.5 points. ACQ-5 will be conducted only in subjects who reported a diagnosis of asthma at the Screening visit.

Change from baseline to Week 16 in ACQ-5 score will be analyzed using an ANCOVA model without MI. The ANCOVA model will be analogous to the model described in Section 8.3.1.3,

and will include covariates for treatment group, region (United States, Eastern Europe), and baseline ACQ-5 score. Actual and change from baseline values in ACQ-5 score will be descriptively summarized by visit using as observed data.

ACQ-5 scores will also be presented in a by-subject data listing.

8.3.3.5.3. PK Concentrations

GB001 plasma concentrations will be summarized by visit/timepoint using the following descriptive statistics: number of subjects with non-missing data, arithmetic mean, SD, arithmetic % coefficient of variation (CV), geometric mean, geometric % CV, median, minimum, and maximum.

For calculation of mean concentrations, all below limit of quantification (BLQ) values will be set to zero. If the number of values that are BLQ at a nominal timepoint exceed 50% of the observations collected at that nominal timepoint, summary statistics for the timepoint will not be calculated. If the mean concentration value is less than the lower limit of quantification (LLOQ) value, then the mean, median, minimum, and maximum concentration values will be set to BLQ.

GB001 plasma concentrations will also be presented in a by-subject data listing.

8.4. Examination of Subgroups

8.4.1. Examination of CRSsNP and CRSwNP Subgroups

Subgroup analyses of the primary and three secondary efficacy endpoints (i.e. change from baseline to Week 16 in SNOT-22 total score, Lund-Mackay score, AM NC score, and AM TSS) will be performed using the primary analysis approach (which is based on the amount of missing data as described in Table 1) for the randomized CRSsNP and CRSwNP subgroups. The CRSwNP subgroup will be further broken down based on the presence or absence of comorbid asthma (CRSwNPsA and CRSwNPwA).

Models used for the subgroup analyses will be similar to those used for the primary and secondary efficacy endpoints as specified in Sections 8.3.1 and 8.3.2.1, with the following additional covariates: treatment group by disease status stratification factor values used for randomization (referred to as 'strata' henceforth in this section, for brevity) interaction, strata by visit interaction, and treatment group by strata by visit interaction for SNOT-22 total score, AM NC score, and AM TSS score; and treatment group by strata interaction for Lund-Mackay score. For SNOT-22 total score, AM NC score, and AM TSS, if the model does not converge, a reduced model excluding region, baseline score, and baseline score by visit interaction will be used. For Lund-Mackay, if the model does not converge, a reduced model including only treatment group, strata, and treatment group by strata interaction will be used.

Heterogeneity of treatment effect across the CRSsNP and CRSwNP subgroups (CRSsNP, CRSwNP, CRSwNPsA, CRSwNPwA) will be evaluated by presenting interaction p-values for the treatment group by strata by visit interaction for Week 16 for SNOT-22 total score, AM NC score and AM TSS score, and treatment group by strata interaction for Lund-Mackay score. The

interaction term will be used to test whether the treatment effects in the CRSsNP and CRSwNP subgroups are significantly different, using CRSsNP as the reference subgroup. Additionally, the interaction term will be used to test whether the treatment effects in CRSwNPsA and CRSwNPwA are significantly different, using CRSwNPsA as the reference group.

Actual, change from baseline, and percent change from baseline values in SNOT-22 domain scores will be descriptively summarized by visit for the randomized CRSsNP and CRSwNP subgroups using as observed data.

8.4.2. Examination of Other Subgroups

Subgroup analyses of the primary and three secondary efficacy endpoints (i.e. change from baseline to Week 16 in SNOT-22 total score, NPS, AM NC score, and AM TSS) will be performed using the primary analysis approach (which is based on the amount of missing data as described in Table 1) for all of the subgroups specified below. For any subgroup level that does not comprise ≥ 20% of the ITT population, the SNOT-22 total score, AM NC score, and AM TSS analyses based on the models will not be performed. For any subgroup level that does not comprise ≥ 20% of subjects in the ITT population with nasal polyps, the NPS analyses based on the models will not be performed. The subgroups of interest are described below.

Demographic characteristic subgroups:

- Age (years; < 50, ≥ 50)
- Sex (male, female)
- BMI (kg/m2; < 30, ≥ 30)
- Region (United States, Eastern Europe)

Baseline disease characteristic and atopic/allergic condition subgroups:

- SNOT-22 total score at Baseline (≤ Median, > Median, where median is based on the ITT population) for the SNOT-22 total score endpoint
- NPS at Baseline (≤ 6, > 6) for the NPS endpoint
- AM NC score at Baseline (≤ Median, > Median, where median is based on the ITT population) for the AM NC score endpoint
- AM TSS score at Baseline (≤ Median, > Median, where median is based on the ITT population) for the AM TSS score endpoint
- Prior NP surgery (yes, no)
- Allergic Rhinitis (seasonal and/or perennial) at Baseline (yes, no)

Biomarker baseline characteristic subgroups:

- Blood eosinophils at Baseline (10^9/L; < 0.25, ≥ 0.25)
- FeNO at Baseline (ppb; < 25, ≥ 25)

- Chloride at Baseline (mmol/L; ≤ Quartile 1, > Quartile 1-≤ Quartile 2, > Quartile 3-≤ Quartile 3, > Quartile 3, where quartiles are based on the ITT population)

Models used for the subgroup analyses will be similar to those used for the primary and secondary efficacy endpoints as specified in Sections 8.3.1 and 8.3.2.1, with the following additional covariates: subgroup (if different than the aforementioned covariates, i.e. region), treatment group by subgroup interaction, subgroup by visit interaction, and treatment group by subgroup by visit interaction. If the model does not converge, a reduced model excluding strata, region (if not the subgroup), baseline score, and baseline score by visit interaction will be used.

Heterogeneity of treatment effect across different levels of each subgroup will be evaluated by presenting interaction p-values for the treatment group by subgroup by visit interaction for Week 16. The interaction term tests whether the treatment effect is significantly different across different levels of the subgroup (between males and females, for example). The least severe or lower subgroup level (where applicable) will be used as the reference subgroup, and the test will assess whether the treatment effect in the more severe level(s) or higher value level(s) differs from that in the reference subgroup. For subgroups without directionality (sex and region), male and United States, respectively, will be used as the reference subgroup.

Forest plots of the treatment effect estimates, corresponding 95% CIs, subgroup p-values, and interaction p-values for the subgroup analyses described in Sections 8.4.1 and 8.4.2 will be provided.

9. SAFETY EVALUATION

9.1. Adverse Events

Verbatim AE terms will be coded to a SOC and PT using MedDRA Version 21.0 or later.

All analyses of AEs will be based on the principle of treatment emergence. An AE is considered to be treatment-emergent if it has a start date on or after the date of the first dose of study treatment. In general, if the treatment emergence of an AE is not clear due to a missing or incomplete AE start date, the AE will be considered to be treatment-emergent unless the non-missing portions of the start date and the end date indicate otherwise.

In general, whenever a tabular summary of AEs is mentioned in this document, it is intended that the tabular summary is in reference to TEAEs, even though "treatment-emergent" may not be explicitly mentioned. By-subject data listings of AEs will include all events regardless of treatment emergence and TEAEs will be identified.

Counting will be at the subject level for each level of summarization (e.g., any AE, SOC, and PT), with subjects experiencing more than one AE counted only once. In the summary of AEs by severity (mild, moderate, severe), subjects will be counted once at the highest severity reported at each level of summarization. AEs that are missing severity will be presented in summary tables as "Missing". AEs that are missing relationship to study treatment will be presented in summary tables as "Related".

The following summaries will be presented for the safety population:

- Overall Summary of AEs - showing the number and percent of subjects with an AE; a moderate or severe AE; a severe AE; a SAE; a treatment-related AE; a treatment-related SAE; an AE leading to discontinuation of study treatment; an AE leading to withdrawal from study; an AE resulting in death; and an AE of interest (as defined in Protocol Section 10.3.3).
- Incidence of AEs by SOC and PT
- Incidence of AEs by PT
- AEs with Incidence ≥ 5% in Either Treatment Group by PT
- AEs with Incidence at Least ≥ 3% Higher in the GB001 40 mg Group Relative to the Placebo Group by Preferred Term
- Incidence of SAEs by SOC and PT
- Incidence of AEs by SOC, PT, and Maximum Severity
- Incidence of Treatment-Related AEs by SOC and PT
- Incidence of AEs Leading to Discontinuation of Study Treatment by SOC and PT
- Incidence of AEs Leading to Withdrawal from Study by SOC and PT
- Summary of Pruritus AEs

- Incidence of Liver Chemistry AEs by PT
- Incidence of Liver Chemistry AEs Leading to Discontinuation of Study Treatment by PT
- Incidence of AEs of Interest by SOC and PT

All AEs, AEs leading to discontinuation of study treatment and/or withdrawal from study, AEs of interest, SAEs, and AEs resulting in death will be presented in by-subject data listings.

9.2. Clinical Laboratory Evaluation

Actual values and change from baseline values for quantitative laboratory parameters (clinical and liver chemistry, hematology, and urinalysis) will be summarized by visit for the safety population.

Shift tables will be presented for clinical and liver chemistry, hematology, and qualitative urinalysis parameters summarizing shifts from baseline to high and low (or to abnormal for urinalysis) at any time post-baseline based on laboratory normal ranges.

A summary of maximum post-baseline values > 1.0, > 1.5, and > 3.0 10^9/L for blood eosinophils will be presented.

Summaries of maximum post-baseline values for liver chemistry parameters (alanine aminotransferase [ALT], aspartate aminotransferase [AST], ALT or AST, concurrent ALT or AST and total bilirubin, total bilirubin, direct bilirubin, indirect bilirubin, alkaline phosphatase, gamma-glutamyl transferase [GGT]), and bile acid according to categories based on multiples of upper limit of normal (ULN) will be presented. Liver chemistry parameters according to categories based on multiples of ULN will also be presented by visit.

For laboratory shift tables and summaries of maximum post-baseline values, only subjects with at least one post-baseline value will be included.

The time to first elevation for ALT will be presented for subjects with an elevation in ALT. An elevation for ALT will be defined as the first post-baseline value ≥ 1.5 x ULN for subjects with a baseline value ≤ ULN, and the first post-baseline value with a > 50% increase from baseline for subjects with a baseline value > ULN. This analysis will also be repeated using a threshold of > 1.0 x ULN. In addition, the time to first elevation of AST will be summarized in the same manner as described above for time to first elevation of ALT.

In addition, the following figures/plots will be provided:

- Mean change from baseline values, including SE bars, by visit figures for liver chemistry parameters
- Mean ratio of actual values to ULN, including SE bars, by visit figures for liver chemistry parameters
- Evaluation of Drug-Induced Serious Hepatotoxicity (eDISH) (Merz, 2014) plots

- Mean actual values and mean change from baseline values, including SE bars, by visit figures for blood eosinophils
- Maximum post-baseline values versus baseline values plot for blood eosinophils

Laboratory parameters will also be presented in by-subject data listings. In addition, liver event information including liver biopsy and liver imaging results will be presented in a by-subject data listing.

9.3. Vital Signs and Other Observations Related to Safety

9.3.1. Vital Signs

Actual values and change from baseline values for vital sign parameters (systolic and diastolic blood pressure, pulse rate, respiratory rate, temperature, and weight) will be summarized by visit for the safety population.

A summary table of the incidence of abnormalities in vital sign parameters will be presented according to the following abnormality criteria.

Table 2: Vital Sign Parameter Abnormality Criteria

Vital Sign Parameter	Abnormality Criteria
Systolic Blood Pressure	High or Increased: > 180 mmHg post-baseline if ≤ 180 mmHg at baseline, or an increase from baseline of > 40 mmHg. Low or Decreased: < 90 mmHg post-baseline if ≥ 90 mmHg at baseline, or a decrease from baseline of > 30 mmHg.
Diastolic Blood Pressure	High or Increased: > 105 mmHg post-baseline if ≤ 105 mmHg at baseline, or an increase from baseline of > 30 mmHg. Low or Decreased: < 50 mmHg post-baseline if ≥ 50 mmHg at baseline, or a decrease from baseline of > 20 mmHg.
Pulse Rate	High or Increased: > 120 bpm post-baseline if ≤ 120 bpm at baseline, or an increase from baseline of > 20 bpm. Low or Decreased: < 50 bpm post-baseline if ≥ 50 bpm at baseline, or a decrease from baseline of > 20 bpm.
Temperature	> 38 degrees C and an increase from baseline of at least 1 degree C.

bpm = beats per minute

Vital sign parameters will also be presented in a by-subject data listing and abnormalities will be flagged.

9.3.2. Other Safety Measures

9.3.2.1. Electrocardiograms

Actual values and change from baseline values will be summarized by visit for the safety population for the following quantitative ECG parameters: heart rate, PR interval, QRS duration,

QT interval (uncorrected), Fridericia's correction formula for QT interval (QTcF), and RR interval.

Outlier analyses for QTcF intervals will be performed. This will consist of a summary of the number and percent of subjects with a post-baseline QTcF interval greater than 450 msec, 480 msec, and 500 msec and the number and percent of subjects with an increase from baseline in QTcF interval of greater than 30 msec and 60 msec.

Each ECG will be assessed with an overall interpretation of normal, abnormal, or unable to evaluate. Shift tables for overall interpretation will be presented, summarizing shifts from baseline to abnormal at any time post-baseline.

ECG parameters will also be presented in a by-subject data listing and outliers will be flagged.

10. REFERENCES

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. *Second International Symposium on Information Theory*, pp. 267-281.

Allison, P. (2005). Imputation of categorical variables with PROC MI. *Presented at 30th Meeting of SAS User Group International, April 10-13*, Philadelphia, PA.

Brookmeyer R, C. J. (1982). A confidence interval for the median survival time. *Biometrics*, 29-41.

Doty R. L. (1984). University of Pennsylvania Smell Identification Test: a rapid quantitative olfactory function test for the clinic. *Laryngoscope,* 94,176-178.

EMA (2020). Guidance on the Management of Clinical Trials during the COVID-19 (Coronavirus) pandemic. Version 3 (28/04/2020) Published 28 April 2020. Amsterdam, The Netherlands. Available from: https://ec.europa.eu/health/sites/health/files/files/eudralex/vol-10/guidanceclinicaltrials_covid19_en.pdf.

EMA (2020). Points to consider on implications of Coronavirus disease (COVID-19) on methodological aspects of ongoing clinical trials. Committee for Medicinal Products for Human Use (CHMP). EMA/158330/2020 Rev. 1. Published 26 June 2020. Amsterdam, The Netherlands. Available from: https://www.ema.europa.eu/en/documents/scientific-guideline/points-consider-implications-coronavirus-disease-covid-19-methodological-aspects-ongoing-clinical_en-0.pdf.

FDA (2020). FDA Guidance on Conduct of Clinical Trials of Medical Products during COVID-19 Pandemic. Guidance for Industry, Investigators, and Institutional Review Boards. Published March 2020, Updated 02 July 2020. Available from: https://www.fda.gov/media/136238/download.

FDA (2020). Statistical Considerations for Clinical Trials During the COVID-19 Public Health Emergency. Guidance for Industry. Published June 2020. Available from: https://www.fda.gov/media/139145/download.

Fokkens W, Lund V, Mullol J. European position paper on rhinosinusitis and nasal polyps 2007. Rhinol Suppl 2007;20:1-136.

Hopkins, C. G. (2009). Psychometric validity of the 22-item Sinonasal Outcome Test. *Clin Otolaryngol*, 34(5):447-454.

ICH Harmonised Tripartite Guideline: Statistical Principles for Clinical Trials E9. Step 4. 1998. (n.d.). Retrieved from https://database.ich.org/sites/default/files/E9_Guideline.pdf

ICH Harmonised Tripartite Guideline: Structure and content of clinical study reports E3. Step 4. 1995. (n.d.). Retrieved from https://database.ich.org/sites/default/files/E3_Guideline.pdf

Kalbfleisch, J. D. and Prentice, R. L. (1980), The Statistical Analysis of Failure Time Data, New York: John Wiley & Sons.

Kenward, M. G. (1997). Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics*, 53, 983-997.

Kuiper J. R. (2018). Prevalence, severity, and risk factors for acute exacerbations of nasal and sinus symptoms by chronic rhinosinusitis status. *Allergy*, 73(6):1244-1253.

Lund V. J. (1993). Staging in rhinosinusitus. *Rhinology,*31,183-184.

Merz (2014). *Drug Saf* , 37 (Suppl 1), S33–S45.

Newcombe, R. G. (1998). Interval Estimation for the Difference Between Independent Proportions: Comparison of Eleven Methods. *Statistics in Medicine*, 17, 873–890.

Rubin, D.B. (1987) Multiple Imputation for Nonresponse in Surveys, New York: John Wiley & Sons Inc.

Schafer, J.L. (1997), *Analysis of Incomplete Multivariate Data*, New York: Chapman and Hall.

Wilson, E. B. (1927). Probable inference, the law of succession, and statistical inference. *J. Amer. Stat. Assoc.,* 22, 209–212.