Statistical Analysis Plan	
Study Code	D5242C00001
Edition Number	4.0
Date	21 Aug 2024

A Multicentre, Randomised, Double-Blind, Parallel-Group, Placebo-Controlled Phase 3 Efficacy and Safety Study of Tezepelumab in Participants with Severe Chronic Rhinosinusitis with Nasal Polyposis (WAYPOINT)

Redacted for Public Disclosure

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Abbreviation or special term	Explanation
ACQ-6	Asthma Control Questionnaire 6
ADA	Anti-Drug Antibody
AE	Adverse Event
AERD	Aspirin Exacerbated Respiratory Disease
AESI	Adverse Event of Special Interest
ALT	Alanine Aminotransferase
ANCOVA	Analysis of Covariance
AST	Aspartate Aminotransferase
ATC	Anatomical Therapeutic Classification
BD	Bronchodilator
BMI	Body Mass Index
BP	Blood Pressure
CDISC	Clinical Data Interchange Standards Consortium
CI	Confidence Interval
CMH	Cochran-Mantel-Haenszel
COVID-19	Corona Virus Disease 2019
CRO	Contract Research Organization
CRSwNP	Chronic Rhinosinusitis with Nasal Polyposis
CSR	Clinical Study Report
CSP	Clinical Study Protocol
CT	Computed Tomography
CV	Coefficient of Variation
DAE	AE Leading to Discontinuation of Investigation Product
DBL	Database Lock
DNA	Deoxyribonucleic Acid
DRMI	Dropout Reason-based Multiple Imputation
EAIR	Exposure-adjusted incidence rate
ECG	Electrocardiogram
eCRF	Electronic Case Report Form
EOT	End of Treatment
EQ-5D-5L	European Quality of Life-5 Dimensions-5 Levels
FAS	Full Analysis Set
FeNO	Fractional Exhaled Nitric Oxide
FEF	Forced Expiratory Flow

Abbreviation or special term	Explanation
FEV_1	Forced Expiratory volume in 1 second
FU	Follow-up
FVC	Forced Vital Capacity
HR	Hazard Ratio
HRU	Health Resource Utilisation
HRQoL	Health-Related Quality of Life
IAC	Independent Adjudication Committee
IgE FEIA	Allergen-Specific Immunoglobulin E
INCS	Inhaled Corticosteroids
IP	Investigational Product
IPD	Investigation Product Discontinuation
IWRS	Integrated Web Response System
LLOQ	Lower Limit of Quantification
LMK	Lund Mackay Score
LS	Least Squares
MACE	Major Adverse Cardiac Events
MAR	Missing at Random
MCID	Minimal Clinically Important Difference
MCMC	Markov Chain Monte Carlo
MedDRA	Medical Dictionary for Regulatory Activities
MFNS	Mometasone Furoate Nasal Spray
MI	Multiple Imputation
MNAR	Missing Not At Random
nAb	Neutralising Antibodies
NC	Nasal Congestion
NCS	Nasal Congestion Score
NPIF	Nasal Peak Inspiratory Flow
NPS	Nasal Polyp Score
NPSD	Nasal Polyposis Symptom Diary
NQ	Non-Quantifiable
NSAID-ERD	Nonsteroidal Anti-Inflammatory Drug Exacerbated Respiratory Disease
OCS	Oral Corticosteroid
PD	Protocol Deviation
PGI-C	Patient Global Impression of Change
PGI-S	Patient Global Impression of Severity
PK	Pharmacokinetic

Abbreviation or special term	Explanation
PKS	Pharmacokinetic Analysis Set
ePRO	electronic Patient-Reported Outcome
PT	Preferred Term
QTc	Corrected QT
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SCS	Systemic Corticosteroids
SD	Standard Deviation
SNOT-22	SinoNasal Outcome Test, 22 item
SF-36v2	Short Form 36-Item Health Survey Version 2
SOC	System Organ Class
SS	Safety Set
TBL	Total Bilirubin
TSS	Total Symptom Score
ULN	Upper Limit of Normal Reference Range
ULOQ	Upper Limit of Quantification
UPSIT	University of Pennsylvania Smell Identification Test
VAS	Visual Acuity Score
WHO	World Health Organization
WOCF	Worst Observation Carried Forward
WPAI	Work Productivity and Activity Impairment
WPS	Worst Possible Score

AMENDMENT HISTORY

Category: Change refers to	Date	Description of change	In line with the CSP?	Rationale
Primary or key secondary endpoints	21Aug2024	Removed the text "For the NCS, if there are sufficient evaluable data prior to the intercurrent event in a 14-day period, the bi-weekly mean NCS for that period will be based on the collected data. Otherwise, the NCS for that period will be set to missing" from MI Step 1. Added clarification that biweekly mean NCS will be set to missing if the intercurrent events occurred before the last day of the same 14-day interval. (Section 4.2.6.1)	N/A	To streamline analysis.
Primary or key secondary endpoints	21Aug2024	Clarified baseline value will be used to replace WOCF for participants whose post-baseline values are all missing (Section 4.2.6.1)	N/A	To provide clarification.
Primary or key secondary endpoints	21Aug2024	Clarified COVID-19 infection is based on COVID-19 MedDRA terms, and if COVID-19 infection started prior to randomisation, data prior to the end date of the infection will be set to missing.(Section 4.2.6.3.2)	N/A	To provide clarification.
Primary or key secondary endpoints	21Aug2024	Added the missing parenthesis in the st_wh_cmh algorithm (Section 4.2.6.4.1)	N/A	Correction.
Primary or key secondary endpoints	21Aug2024	Added clarification in tipping point analysis (Section 4.2.6.4)	N/A	To provide clarification.
Primary or key secondary endpoints	21Aug2024	Clarified the presentation on the cumulative distribution function curves (Section 4.2.6.5)	N/A	To clarify data presentation.
Primary or key secondary endpoints	21Aug2024	Added more details on model for the subgroup analyses (Section 4.2.6.6)	N/A	To clarify analysis.
Other secondary endpoints	21Aug2024	Clarified the maximum total Zinreich score is 50 because Osteomeatal complex score is not reported for Zinreich score (Section 3.4.4).	N/A	To provide clarification.

Exploratory endpoints	21Aug2024	Removed the analysis for timepoints after Week 52 from the co-primary, key secondary, and other secondary efficacy endpoint sections (Section 4.2.6.1, 4.2.7.1.8, 4.2.9.1, 4.2.9.4, 4.2.9.5, 4.2.9.6, and 6) Removed the text "For participants whose last scheduled visit is Week 64, data after Week 64 will be set to missing after the imputation" from MI Step 4 (Section 4.2.6.1) Updated the analysis method on change from baseline in NPS and NCS up to Week 76 (Section 4.2.9.1)	N/A	To address potential model convergency issue and clarify the analysis.
Data Presentation	21Aug2024	Added information about GCP significant deviations at PPD PPD and a critical data quality issue identified at PPD (Section 2.1.4)	N/A	To add findings on GCP significant deviations and a critical data quality issue
Data Presentation	21Aug2024	Removed the text about baseline not to be imputed. Missing baseline may be imputed by multiple imputation (Section 3.1.1)	N/A	Correction
Data Presentation	21Aug2024	Added clarification that Day 0 in CSP is Day 1 in analysis following CDISC standard (Section 3.1.1)	Yes CSP v6.0	Clarification
Data Presentation	21Aug2024	Updated baseline eosinophil count from (<150, ≥150 to ≤300, >300) to (<150, ≥150 to <300, ≥300) to align with tezepelumab asthma studies, updated time since last NP surgery at screening visit from (<6 months, ≥6 months to <3 years, ≥3 years) to (<3 years, ≥3 years), Clarified the subgroup variable AERD/NSAID-ERD is at screening visit (Section 3.1.6)	N/A	Correction and clarification.
Data Presentation	21Aug2024	Updated the derivation on time to premature study withdrawal for ongoing participants (change from last contact date to data cut-off date) at the time of primary DBL (Section 3.1.8)	N/A	To keep consistency with on-study period derivation for ongoing participants at the time of primary DBL.
Data Presentation	21Aug2024	Clarified the MedDRA v27.0 and WHODrug Mar 2024 will be used (Section 3.6.2 and 4.2.4)	N/A	To provide clarification.
Data Presentation	21Aug2024	Clarified the details of urinalysis for China and non-China participants, and that urinalysis shift tables will be separate for China and non-China participants (Section 3.6.4 and 4.2.10.2)	N/A	To provide clarification.

Data Presentation	21Aug2024	Clarified imputed date will be capped by Date of Birth when partially missing medical history date is imputed (Section 3.6.8)	N/A	To provide clarification.
Data Presentation	21Aug2024	Clarified presentation on standard of care medications at study entry and disallowed concomitant medications (Section 4.2.4)	Yes CSP v6.0	To provide clarification.
Data Presentation	21Aug2024	Minor update on the categories of cumulative extent of exposure from >0 - ≤ 4, >4 - ≤8, >8 - ≤12, etc. >48 - ≤52, and >52 weeks to >0 - < 4, ≥4 - <8, ≥8 - <12, etc. ≥48 - <52, and ≥52 weeks (Section 4.2.5)	N/A	For consistency.
Data Presentation	21Aug2024	Clarified the SNOT-22, NPS, NCS responders will be presented on the number and proportion of responders, not non-responders or missing data (Section 4.2.7.2.1, 4.2.8.7, 4.2.9.7)	N/A	To provide clarification.
Data Presentation	21Aug2024	Updated the safety subgroup variable from Race (White, Black or African American, Asian, Other) to Race (White, non-White) (Section 4.2.10.1)	N/A	To keep consistency with efficacy.
Data Presentation	21Aug2024	Removed the locally estimated scatterplot smoothing (LOESS) plot for biomarkers; Removed the conventional unit presentations on blood eosinophils (10^9/L) and total serum IgE (mg/L) (Section 4.2.12).	N/A	To streamline analyses.
Data Presentation	21Aug2024	Simplified presentation of disruption due to global/country situation. Corrected the IPD summaries due to global/country situation are performed using Randomised analysis set (Section 4.2.13).	N/A	To streamline analyses and align with corporate TFL standard.
Data Presentation	21Aug2024	Clarified no p-values will be presented for subpopulation analyses performed for the participants enrolled in the region and used in China and Japan registration (Section 4.2.14).	N/A	To provide clarification.
Data Presentation	21Aug2024	Clarified the Cox regression will not be performed for the time to event type endpoints if there are <10 events across the treatment groups and <2 events in each treatment group in China/Japan subpopulation analyses (Section 4.2.14).	N/A	To provide clarification.

Other	21Aug2024	Fix the formatting issue on inserted reference in SAP v3 (Section 3.2.2)	N/A	Correction.
Other	21Aug2024	Added another reference (Bannick et al., 2024) for robust variance estimation (Section 4.2.7.2.1 and Section 7).	N/A	To provide clarification.
Other	21Aug2024	Updated the text 'treatment-emergent AEs (TEAEs)' to 'AEs' (Appendix 8.3)	N/A	To align with tezepelumab standard.
Primary or key secondary endpoints	6Feb2024	Endpoints related to proportion of participants achieving resolution/near complete resolution of NP and achieving resolution/near complete resolution of NP and NPSD TSS response moved from key secondary objectives (Section 1.1.2) to exploratory objectives (Section 1.1.5). Endpoints removed from MTP (Section 4.1.1). All other sections related to those endpoint moved from SAP to exploratory SAP (Sections in SAP v2.0: 3.3.6, 3.3.7, 4.2.7.1.8, 4.2.7.1.9, 4.2.7.3.5, 4.2.7.3.6, and 4.2.9.11).	Yes CSP v6.0	To align with the CSP v6.0, and to streamline the analysis included in the SAP and CSR.
Primary or key secondary endpoints	6Feb2024	Added the clarification of total NPS score derivation (Section 3.2.1).	N/A	To provide clarification of NPS derivation.
Primary or key secondary endpoints	6Feb2024	Updated Holm procedure in MTP to truncated Holm procedure. Text and Figure 2 updated accordingly (Section 4.1.2).	Yes CSP v6.0	To align with the CSP v6.0.
Primary or key secondary endpoints	6Feb2024	US FDA Primary Estimand (Section 4.1.3.2) and Supplementary Estimand (Section 4.1.3.3) description added.	Yes CSP v6.0	To align with the CSP v6.0.
Primary or key secondary endpoints	6Feb2024	Clarification provided for the analysis of co-primary endpoints following primary estimand analysis (Section 4.2.6.1), US FDA primary estimand analysis (Section 4.2.6.2), and supplementary estimand analysis (Section 4.2.6.3).	Yes CSP v6.0	To align with the CSP v6.0.
Primary or key secondary endpoints	6Feb2024	Additional information provided for rank ANCOVA analysis (Section 4.2.6.4.1).	N/A	To provide clarification of the analysis.
Primary or key secondary endpoints	6Feb2024	Description of ANCOVA model added to control based imputation (Section 4.2.6.4.2).	N/A	Missing information provided.

Primary or key secondary endpoints	6Feb2024	Analysis for US FDA primary estimand added to selected key secondary endpoints (Sections 4.2.7.1.1, 4.2.7.1.2, 4.2.7.1.3, 4.2.7.1.7, 4.2.8.1, 4.2.8.2, and 4.2.8.3).	Yes CSP v6.0	To align with the CSP v6.0.
Primary or key secondary endpoints	6Feb2024	Co-morbid asthma removed from the model of pre-BD FEV_1, data after biologic use for NP will be used in analysis, prior NP surgery status and region will be based on eCRF (Section 4.2.7.1.8).	N/A	To clarify covariates in the model.
Primary or key secondary endpoints	6Feb2024	Section describing supplementary analyses for key secondary endpoints (Section 4.2.7.2 SAP v2.0) removed. The analysis has been replaced by US FDA primary analysis.	Yes CSP v6.0	To align with the CSP v6.0.
Data presentation	6Feb2024	"Seasonal allergic rhinitis" replaced with "allergic rhinitis" (Section 3.1.6).	Yes CSP v6.0	To align with the CSP v6.0.
Data presentation	6Feb2024	Additional region categories provided (Section 3.1.6).	N/A	To provide information of interests.
Data presentation	6Feb2024	JESREC scores added to the list of other baseline characteristics, baseline ratio of ethmoid/maxillary sinus score removed, NPIF separated from assessments done for co-morbid asthma/AERD/NSAID-ERD subset (Section 3.1.7).	N/A	To provide information of interests and clarification.
Data presentation	6Feb2024	Added NCS responder derivation (Section 3.2.2) and analysis (Section 4.2.9.7).		To support the NSC co-primary endpoint.
Data presentation	6Feb2024	Presentation of disruption due to global/country situation clarified (Section 4.2.1).	N/A	To provide clarification.
Data presentation	6Feb2024	Clarified that percentage of participants in analysis sets will not be provided (Section 4.2.2).	N/A	To provide clarification.
Data presentation	6Feb2024	Clarification that IP compliance is presented using FAS. Categories of compliance updated. Compliance for daily diary and with MFNS/INCS will be provided bi-weekly. (Section 4.2.5).	N/A	To provide clarification.

Data presentation	6Feb2024	Minimum number of participants in subgroup provided, and clarified that stratification factors as captured on eCRF will be used in analysis (Section 4.2.6.6).	N/A	To provide clarification.
Data presentation	6Feb2024	Exploratory endpoints: Health Resource Utilisation (Sections 3.5.2 and 4.2.9.3 SAP v2.0), Short Form 36 Item Health Survey (Sections 3.5.3 and 4.2.9.5 SAP v2.0), Work Productivity and Activity Impairment Questionnaire (Sections 3.5.5 and 4.2.9.7 SAP v2.0), Proportion of participants requiring NP surgery decision and/or SCS for NP up to week 76 (Section 3.5.1 and Section 4.2.9.2 SAP v2.0) moved from the SAP to exploratory SAP.	N/A	To streamline the analysis included in the SAP and CSR.
Data presentation	6Feb2024	Clarified how the analysis results of change from baseline in SNOT-22, loss of smell, UPSIT, NPSD TSS and NPIF overtime up to Week 76 will be presented Section (4.2.9.4).	N/A	To provide clarification.
Data presentation	6Feb2024	Pre-BD FEV_1 and ACQ-6 overtime will not be presented graphically (4.2.9.5).	N/A	To streamline the analysis included in the SAP and CSR.
Data presentation	6Feb2024	Presentation of exposure adjusted AEs clarified. Added subgroups summaries of AEs (Section 4.2.10.1).	N/A	To provide clarification.
Data presentation	6Feb2024	Added information of presenting Immunogenicity data for China and non-China participants separately. Clarification provided that ADA results will be presented in CSR regardless of the number of ADA positive participants. Association of ADA status with other endpoints will be listed, and not presented in tables. (Section 4.2.11.2)	N/A	To provide clarification.
Other	6Feb2024	Information provided that in case of severe site misconduct or suspected fraud, the data may not be used in analyses. Decision has to be taken before unblinding. (Section 2.1)	N/A	To clarify how the data will be used in case of potential severe misconduct or fraud.
Other	6Feb2024	Reference to Intention To Treat removed (Section 2.1.1).	N/A	To correct inconsistency across the document.
Other	6Feb2024	Definition of Additional FU Analysis Set updated to provide more clarity of who should be assigned to this set (Section 2.1.3).		

Other	6Feb2024	"Co-morbid Asthma and AERD/NSAID-ERD" updated to "Co-morbid Asthma/AERD/NSAID-ERD" to clarify who should be included in the subset (update made throughout the document except in tables with objectives).	N/A	To provide clarification.
Other	6Feb2024	"Safety Analysis Set" updated to "Safety Set" (Section 2.1.2, 4.2.5 and 4.2.11.2).	Yes CSP v6.0	To align with the CSP v6.0.
Other	6Feb2024	Clarify SCS for NP derivation (Section 3.4.5).	N/A	To provide clarification of derivation.
Other	6Feb2024	Clarify that the highest NPIF value will be used in analysis (Section 3.4.7). Clarification of the analysis provided (Section 4.2.8.5).	N/A	To provide clarification of derivation.
Other	6Feb2024	Clarified that stratification variables as captured on eCRF will be used as covariates in the analysis (Section 4.2.7.1.4).	N/A	To provide clarification of analysis.
Other	6Feb2024	Section describing supplementary analyses for key secondary endpoints (Section 4.2.8.9 SAP v2.0) removed.	N/A	To align with Health Authority feedback.
Other	6Feb2024	Clarified that analysis of change from baseline in NPS and NCS up to Week 76 will be performed based on both primary and US FDA estimand (Section 4.2.9.1).	N/A	To provide clarification.
Other	6Feb2024	Changed "NPSD component" to "NPSD domain" in Sections 1.1.3, 3.4.6, and 4.2.8.4 to be consistent with the wording in the CSP. Added a sentence in Section 4.2.8.4 to clarify the analysis will be performed for each item in the NPSD.	N/A	To correct the terminology.
Other	6Feb2024	Time at risk defined, and clarification provided that stratification variables as captured on eCRF will be used in analysis (Section 4.2.8.8).	N/A	To provide clarification.
Other	6Feb2024	Responder derivation, presentation and analysis clarified for SNOT-22 (Section 4.2.7.2.1), ACQ-6 (Section 4.2.8.6), NPS (Section 4.2.8.7).	N/A	To align with Health Authority feedback.
Other	6Feb2024	The analysis of pre-BD FEV_1 is consistent in CSP and SAP. The text referring to discrepancy removed from Section 6.	Yes CSP v6.0	To align with the CSP v6.0.

Other	6Feb2024	Definition of serious hypersensitivity reactions (AESI) clarified (Section 8.1)	N/A	To provide clarification.
Other	6Feb2024	Appendix 8.3 added with partial dates imputation for prior and concomitant medication (Section 3.1.5) and for Adverse Events (Section 3.6.2).	N/A	To provide clarification.
Primary or key secondary endpoints	08Feb2023	The text "bi-weekly mean" added to NCS endpoint (Section 1.1.1), loss of smell, NPSD TSS (Section 1.1.2) and NCS (Section 1.1.3) "Time to surgery" updated to "Time to surgery decision (Sections 1.1.2, 3.3.3, 4.1.1, 4.1.2, 4.2.7.1.4, and 4.2.7.1.5).	Yes CSP v5.0	To align with the CSP v5.0.
Primary or key secondary endpoints	08Feb2023	Time to actual NP surgery (Section 3.3.3) will be used in the supportive analysis (Section 4.2.7.2.3).	N/A	Added to support the analysis of key secondary endpoints.
Primary or key secondary endpoints	08Feb2023	Steps 1 and 4 updated for NCS analysis to use observed data, if possible, for the bi-weekly mean NCS prior to NP surgery/SCS for NP (Section 4.2.6.1).	N/A	To simplify the algorithm.
Primary or key secondary endpoints	08Feb2023	Clarification provided that the focus of presenting the results should be on Week 52 (Section 4.2.6.1).	N/A	To provide clarification how the co-primary endpoints should be presented.
Primary or key secondary endpoints	08Feb2023	Clarification added when the data after COVID-19 infection should be set to missing (Section 4.2.6.3.2).	N/A	To provide clarification for the analysis.
Primary or key secondary endpoints	08Feb2023	The tipping point analysis updated to be 2-dimentional (Section 4.2.6.4.3).	N/A	To be in line with tezepelumab project.
Primary or key secondary endpoints	08Feb2023	It was clarified that only ANCOVA model has to be repeated for subgroup analysis (Section 4.2.6.6).	N/A	To provide clarification for subgroup analyses of co-primary endpoints.
Primary or key secondary endpoints	08Feb2023	Additional covariate (baseline NPSD TSS) added to the model specification (Section 4.2.7.1.9 SAP v2.0).	N/A	To provide clarification for the analysis.
Primary or key secondary endpoints	08Feb2023	Added additional supplementary analyses for the key secondary endpoints (Section 4.2.7.2 SAP v2.0).	N/A	To evaluate the robustness of the analysis results.

Primary or key secondary endpoints	08Feb2023	The pre-BD FEV_1 added to supportive analysis of change from baseline analysis trough Week 52 (Section 4.2.7.2.2).	N/A	To provide clarification of how the pre-BD FEV_1 will be analysed.
Primary or key secondary endpoints	08Feb2023	Supportive analysis added presenting proportion of participant achieving a maximum NPS ≤1 in each nostril trough Week 52 (Section 4.2.7.3.2 SAP v2.0).	N/A	To provide summaries of clinical interests.
Data presentation	08Feb2023	PK analysis set definition updated to clarify that at least one detectable tezepelumab serum concentration for a sample collected post-treatment is required (Section 2.1.3).	Yes CSP v5.0	To align with the CSP v5.0.
Data presentation	08Feb2023	Additional FU analysis set has been added (Section 2.1.3).	N/A	To be used for AE summaries.
Data presentation	08Feb2023	The title of Section 2.2 and categories of important PDs updated (Section 2.2).	N/A	To be align with Non-compliance handling plan v4.0.
Data presentation	08Feb2023	Reference to the date of first dose of IP for safety analysis purpose removed from planned treatment period (Section 3.1.3).	N/A	Planned treatment period is used only for efficacy analyses.
Data presentation	08Feb2023	Additional FU period has been added (Section 3.1.3).		To be used for AE summaries.
Data presentation	08Feb2023	Clarification provided that Day 0 in the CSP is equal the Day 1 according to the CDISC standard (Section 3.1.4).	N/A	To align with CDISC standards.
Data presentation	08Feb2023	The upper limit for CT scan visit window updated from 378 to 490 (Section 3.1.4).	N/A	To allow for CT scans done by the end of Week 64 to be included in the analysis.
Data presentation	08Feb2023	To clarify that subgroups are based on all data at the study entry or at baseline (Section 3.1.6).	N/A	To provide clarification of subgroups definition.
Data presentation	08Feb2023	Perennial specific IgE status will be based on 7 (and not 8) panels (Section 3.1.6).	N/A	IgE Oriental Cockroach panel will not be assessed.
Data presentation	08Feb2023	Baseline Ethmoid / Maxillary ratio has been defined as the additional baseline characteristic (Section 3.1.7).	N/A	To provide information of interests.
Data presentation	08Feb2023	The FEV_1 assessment flagged as the best value according to DTS will be used in analysis (Section 3.3.6).	N/A	To provide clarification which FEV_1 assessment should be used in analysis.

Data presentation	08Feb2023	The Health State Index has been added as part of EQ-5D-5L presentation (Sections 3.5.2 and 4.2.9.3).	N/A	To be align with EQ-5D-5L presentation across the tezepelumab project.
Data presentation	08Feb2023	The denominator for presenteeism, work productivity loss and activity impairment updated from 6 to 10 (Section 3.5.5 SAP v2.0).	N/A	To correct the derivation of WPAI outcome measures.
Data presentation	08Feb2023	The wording of PGI-C categories updated to be consistent with the responses to individual items. Category "Not better" added (Section 3.5.3).	N/A	To provide clarification of how PGI-C categories should be presented.
Data presentation	08Feb2023	The derivation for AEs during the additional FU period has been added (Section 3.6.2).	N/A	To be used for AE summaries.
Data presentation	08Feb2023	Disposition and study disruption due to global/country situation will be presented (Section 4.2.1).	N/A	To summarise the impact of global/country situation to the study.
Data presentation	08Feb2023	Stratification factors will also be summarised based on data collected on eCRF (Section 4.2.2).	N/A	To provide additional summary of stratification factors.
Data presentation	08Feb2023	Presentation of the cumulative extent of exposure and compliance has been updated (Section 4.2.5).	N/A	To align with the integrated summary for asthma studies.
Data presentation	08Feb2023	Remove the LMK and modified LMK summary by sinus (Section 4.2.7.1.3 and Section 4.2.8.3).	N/A	Not required for the CSR.
Data presentation	08Feb2023	Clarified the UPSIT test (Section 4.2.8.2) and LMK score (Section 4.2.8.3) will be analysed trough Week 52.	N/A	To provide clarification of the analysis.
Data presentation	08Feb2023	Added information of how ACQ-6 response and asthma control status will be presented (Section 4.2.8.6).	N/A	To provide clarification of ACQ-6 presentation.
Data presentation	08Feb2023	Definition of NPS responder was clarified. Information provided that the endpoint will be also analysed trough Week 52 (Section 4.2.8.7) and trough Week 76 (Section 4.2.9.6).	N/A	To clarify the analysis of the endpoint.
Data presentation	08Feb2023	Clarification provided to SCS exposure presentation (Section 4.2.8.8). Additional analysis added.	N/A	To provide clarification of SCS presentation and support labelling claim.

Data presentation	08Feb2023	Added supplementary analyses for other secondary endpoints (Section 4.2.8.9 SAP v2.0).	N/A	To evaluate the robustness of the results.
Data presentation	08Feb2023	Information added how the results up to Week 72 should be presented (Section 4.2.8.9 SAP v2.0).	N/A	To provide clarification of how the exploratory endpoints should be presented.
Data presentation	08Feb2023	Additional information provided regarding the analysis of FEV_1 and ACQ-6 up to Week 76 (Section 4.2.9.5).	N/A	To provide clarification of the analysis.
Data presentation	08Feb2023	Additional analysis added presenting proportion of participant achieving a maximum NPS ≤1 in each nostril trough Week 76 (Section 4.2.9.11 SAP v2.0).	N/A	To provide summaries of clinical interests.
Data presentation	08Feb2023	Clarification provided of how of observed AEs/SAEs/DAEs/AESIs as well as adjudicated AEs will be presented (Section 4.2.10.1).	Yes CSP v5.0	To clarify AEs presentation.
Data presentation	08Feb2023	Updated the time at risk calculation for the EAIR (Section 4.2.10.1) and deleted the duplicated text in Section 3.6.2.	N/A	To align with the new guidelines.
Data presentation	08Feb2023	Presentation of laboratory data updated (Section 4.2.10.2). Graphs with mean changes from baseline over time was replaced by shift plots. The shift table presenting baseline and last post-baseline value added. Presentation of potential Hy's Law cases and participants who meet the criteria for Hy's Law clarified.	N/A	To be in line with presenting laboratory data across tezepelumab project.
Data presentation	08Feb2023	Shift plots presenting baseline vital signs and max/min/last post-baseline observation as well as shift table presenting baseline and last post-baseline value added (Section 4.2.10.2).	N/A	To be in line with presenting laboratory data across tezepelumab project.
Data presentation	08Feb2023	Immunogenicity presentation updated (Section 4.2.11.2). Treatment emergent ADA definition updated. Definition of nAb prevalence and nAb incidence added. Additional analysis added (conditional on number of treatment emergent ADA positive subjects).	N/A	To be in line with presenting immunogenicity data across tezepelumab project.
Other	08Feb2023	Changed "NPSD domain" to "NPSD component" (Sections 1.1.3, 3.4.6, and 4.2.8.4).	N/A	To correct the terminology.

Other	08Feb2023	The text “bi-weekly mean” added to NCS endpoint and “Nasal secretions” updated to “Nasal epithelial lining fluid” (Section 1.1.5).	Yes CSP v5.0	To align with the CSP v5.0.
Other	08Feb2023	Clarification provided regarding follow up time (Section 1.2).	Yes CSP v5.0	To align with the CSP v5.0.
Other	08Feb2023	Clarification provided that nasal endoscopy and CT is not required if a participant discontinues IP due to NP surgery and follows option 1 (Section 1.2). Visit windows for EOT updated from +/- 3 days to +/- 5 days (Section 1.2).	Yes CSP v5.0	To align with the CSP v5.0.
Other	08Feb2023	Removed “In the above, no imputations will be performed for missing values” and added “If response to any of the questions is missing, the ACQ-6 will be missing” in Section 3.4.8.	N/A	To clarify ACQ-6 calculation.
Other	08Feb2023	Clarification provided of how NPIF is collected (Section 3.4.7).	Yes CSP v5.0	To align with the CSP v5.0.
Other	08Feb2023	Clarification provided (Section 4.1.3) that intercurrent event “adherence to background MFNS or an equivalent INCS and IP” include switching to other non-biologic/non-surgery/non-steroid treatments. “SCS” replaced with “biologic for NP” for intercurrent event “biologic use for NP”.	Yes CSP v5.0	To provide clarification to intercurrent event definitions.
Other	08Feb2023	Table 12 in Section 4.2.6.4.2 updated to match reason of study withdrawal with reason of IP discontinuation.	N/A	To provide clarification.
Other	08Feb2023	Change “additional analysis due to COVID-19 pandemic” to “ additional analysis due to global/country situation study disruption” (Section 4.2.13).	N/A	To be aligned with the new standards.
Other	08Feb2023	Added the analysis of pre-BD FEV1 in Section 6.	N/A	To clarify the difference with the CSP.
Other	08Feb2023	List of AESI updated (Section 8)	Yes CSP v5.0	To align with the CSP v5.0.

1 STUDY DETAILS

This is the statistical analysis plan (SAP) for study D5242C00001. The SAP describes the statistical analyses specified in the latest version of the clinical study protocol (CSP) in more detail. Any changes to what is specified in the CSP will be described.

1.1 Study Objectives

1.1.1 Co-Primary Objectives

Objective:	**Endpoint/Variable:**
To evaluate the effect of tezepelumab on:	
Nasal polyp score (NPS)	• Co-Primary: Change from baseline in total NPS evaluated by nasal endoscopy at Week 52.
Participant reported Nasal Congestion (NC)	• Co-Primary: Change from baseline in bi-weekly mean Nasal Congestion score (NCS) evaluated as part of the Nasal Polyposis Symptom Diary (NPSD) at Week 52

1.1.2 Key Secondary Objectives

Objective:	**Endpoint/Variable**
To evaluate the effect of tezepelumab on:	
Loss of smell	• Change from baseline in bi-weekly mean loss of smell evaluated as part of the NPSD at Week 52.
Nasal polyp-quality of life compared with placebo	• Change from baseline in SNOT-22 scores at Week 52.
NP surgery and/or receiving systemic corticosteroids (SCS) for NP	• Time to surgery decision and/or SCS for NP up to Week 52. • Time to NP surgery decision up to Week 52. • Time to SCS for NP up to Week 52.
Sinus opacification	• Change from baseline in Lund Mackay score (LMK) evaluated by computed tomography (CT) at Week 52.

Objective:	Endpoint/Variable
NPSD total symptom score (TSS)	• Change from baseline in bi-weekly mean NPSD TSS at Week 52.
Lung function in participants with co-morbid asthma and aspirin exacerbated respiratory disease (AERD) /nonsteroidal anti-inflammatory drug exacerbated respiratory disease (NSAID-ERD)	• Change from baseline in pre-bronchodilator (BD) FEV1 at Week 52.

1.1.3 Other Secondary Objectives

Objective:	Endpoint/Variable
To evaluate the effect of tezepelumab on:	
NPS	• Change from baseline over time in total NPS evaluated by nasal endoscopy through Week 52. • Proportion of participants with (i) ≥1 point reduction and (ii) ≥2 points reduction in NPS at Week 52.
Participant reported NC	• Change from baseline over time in bi-weekly mean NCS evaluated by NPSD through Week 52.
Loss of smell	• Change from baseline in loss of smell evaluated by UPSIT test at Week 52.
Sinus opacification	• Change from baseline in modified LMK score evaluated by CT at Week 52. • Sinus severity score by quantitative CT assessment at Week 52
Systemic corticosteroid use	• Exposure of SCS over 52 Weeks.
NPSD	• Change from baseline by domain of NPSD through Week 52.
Nasal peak inspiratory flow (NPIF)	• Change from baseline in NPIF through Week 52.

Objective:	Endpoint/Variable
Asthma control in participants with co-morbid asthma and AERD/NSAID-ERD	• Change from baseline in ACQ-6 at Week 52.
To evaluate the PK and immunogenicity of tezepelumab	• PK: Serum Concentration • Immunogenicity: Anti-drug antibody (ADA)

1.1.4 Safety Objectives

Objective:	Endpoint/Variable
To evaluate the safety and tolerability of tezepelumab	• Adverse events (AEs)/serious adverse events (SAEs) • Clinical chemistry/hematology/urinalysis • Vital signs • Electrocardiograms (ECG)

1.1.5 Exploratory Objectives

Objective:	Endpoint/Variable
To explore the off-treatment effect and the recurrence rate of nasal polyps after discontinuation of treatment with tezepelumab	• Change from baseline in NPS • Change from baseline in bi-weekly mean NCS • Participants requiring surgery and SCS for NP
To explore the effect of tezepelumab on exploratory biomarkers of inflammation and nasal polyposis disease and investigate biomarkers for predicting response to tezepelumab	Exploratory biomarker parameters: • Serum and plasma for protein biomarkers • Whole blood for transcriptomic profiling • Nasal epithelial lining fluid for protein biomarkers • Nasal polyp biopsies for inflammatory cells by histology and immunohistochemistry, and transcriptomic profiling
To explore tezepelumab PK in nasal epithelial lining fluid	• Exploratory PK concentrations in nasal epithelial lining fluid
Optional exploratory genomics sample	• A blood sample for DNA isolation will be collected from participants who have

Objective:	Endpoint/Variable
	consented to participate in the genetic analysis component of the study.
To evaluate the effect of tezepelumab on unplanned health care resource utilization	• Hospitalisations, emergency room, unplanned out participant visits, and urgent care visits.
To evaluate the participant reported quality of life outcomes	• Short form 36-item health survey version 2 (SF-36 v2) • European quality of life-5 dimensions-5 levels (EQ-5D-5L) • Work productivity and activity impairment (WPAI)
Participant perception of overall symptom severity and change	• Patient Global Impression of Severity (PGI-S) • Patient Global Impression of Change (PGI-C)
To evaluate airway inflammation in participants with co-morbid asthma/AERD/NSAID-ERD	• Change from baseline in fractional exhaled nitric oxide (FeNO).
Resolution/near complete resolution of nasal polyps (defined as maximum NPS of 1 in each nostril)	• Proportion of participants who achieve a maximum NPS of 1 in each nostril at Week 52.
Resolution/near complete resolution of nasal polyps (defined as maximum NPS of 1 in each nostril) and NPSD TSS response	• Proportion of participants who achieve a maximum NPS of 1 in each nostril and NPSD TSS response at Week 52.

1.2 Study Design

This is a Phase 3, multicenter, randomised, double-blind, placebo-controlled, parallel-group, study to evaluate the efficacy and safety of repeat dosing of tezepelumab 210 mg SC versus placebo in participants with severe chronic rhinosinusitis with nasal polyposis (CRSwNP).

Approximately 400 participants will be randomised globally in a 1:1 ratio to receive tezepelumab 210 mg or matching placebo.

Participants will be stratified by region (China, Japan, and Rest of World [ROW]), prior nasal polyp surgery and co-morbid asthma/AERD/NSAID-ERD. Randomisation will be monitored to ensure that 50%-70% of the study population will have co-morbid asthma/AERD/NSAID-ERD, and at least 50% will have had prior surgery for CRSwNP. When the target percentage of participants in a stratum in a region is reached, consideration will be given to closing the IWRS randomisation for that subgroup, which may be done either overall or within a specific region. Once a subgroup is closed, participants in the screening/run-in period in the closed subgroup will not be allowed to be randomised and will be screen failed.

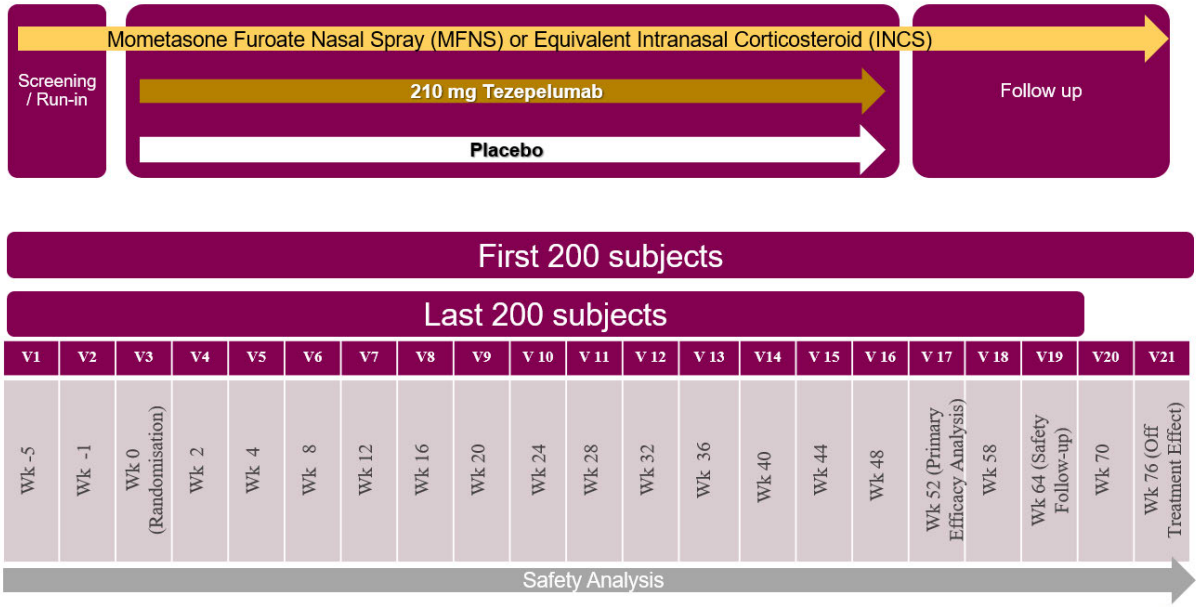
After enrolment, eligible participants will enter a screening/run in period and will have their current inhaled corticosteroid (INCS) therapy standardized to mometasone furoate nasal spray (MFNS) or equivalent INCS at Visit 1 and throughout the screening and study period.

Participants who continue to meet eligibility criteria will be randomised at Visit 3 (Day 0) to receive either placebo or tezepelumab 210 mg SC at randomisation (Week 0) and every 4 weeks thereafter (Weeks 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48). A total of 13 doses will be administered. An end of treatment (EOT) Visit will be conducted at Week 52.

The first approximate 200 randomised participants will have a 24-week follow up (FU) period without investigational product (IP) administration to assess durability of benefit. These participants will have the final safety assessment at Week 76. The remaining participants will be followed for 12 weeks after the end of treatment visit, with the final safety assessments at Week 64.

An overview of the study design is provided below in Figure 1.

Figure 1 Study Design Schema



Participants who prematurely discontinue IP will be encouraged to remain in the study and undergo appropriate study visits/procedures for the full 52-week planned treatment period, despite no longer receiving treatment. At the IP Discontinuation (IPD) visit the participant will be given the following 3 options (further details are given in the CSP Section 7.1.1):

1. The participant should be encouraged to return for all regular clinic visits and perform all scheduled assessments (excluding IP administration and, in the case of IP discontinuation due to NP surgery, nasal endoscopy and CT) until the EOT visit at Week 52 (+/-5 days).

2. The participant will be offered follow-up on a monthly basis via telephone calls while continuing ePRO data collection on the handheld device at home until the participant completes the EOT visit at Week 52 (+/-5 days). The UPSIT assessment will not be completed at home after IPD visit until the participant returns back to the site at the EOT visit.

3. If the participant cannot or does not wish to comply with any of the options above, (or any component of them such as only telephone-based visits without completion of the eDiary), the Investigator will contact the participant at 52 weeks post-randomisation. No other study assessments will be performed prior to this contact.

1.3 Number of Participants

Approximately 400 participants will be randomised to tezepelumab 210 mg Q4W or matching placebo in a 1:1 ratio.

The study is sized to provide compelling statistical evidence for the co-primary endpoints (NPS and NCS) and key secondary endpoints of change from baseline in loss of smell, LMK, and SNOT-22 (considering a significance level of 0.01 in the sample size calculations), as well as sufficient power to assess the composite endpoint of time to SCS and/or surgery (2-sided level of 1%). In addition, this sample size allows assessment of the effect of tezepelumab vs placebo on NPS and NCS in key subgroups and provides a reasonably-sized safety database (CSP Section 9.2).

Assuming a population standard deviation (SD) of 2.25 in total NPS change from baseline and 1.22 in NCS change from baseline, this sample size of 200 participants per arm will provide at least 95% total power to observe a statistically significant difference at a 2-sided 1% level on both co-primary endpoints if the true effect of tezepelumab is -1.8 in total NPS change from baseline and -0.87 in NCS change from baseline. The assumption of true effects is based on reported estimates in the dupilumab Phase III nasal polyp studies. It is expected that treatment with tezepelumab will demonstrate similar or better treatment effects over a 52-week period. The assumed SD was estimated based on the CIs of the corresponding endpoints in the dupilumab Phase III nasal polyp studies (Bachert et al 2019).

Two hundred participants per arm also provides at least 95% power to detect a statistically significant difference at a 2-sided 1% level in secondary endpoints of change from baseline in loss of smell, LMK, and SNOT-22 (assuming mean [SD] of -1 [1.3], -5 [5.45], and -17.4 [26.8], respectively). The size of the study also provides at least 95% power to detect a statistically significant difference at the 2-sided 1% level in the key secondary endpoint time to SCS and/or surgery for NP, assuming a true reduction in risk of 67% (hazard ratio [HR] of 0.33) with a placebo rate at Week 52 of 35% (placebo data estimated from dupilumab SINUS-52). In addition, this provides at least 80% power to detect a difference (at a nominal 5% significance level) in the key secondary endpoint of time to surgery, assuming a true HR of 0.33 and a placebo rate at Week 52 of 11% (placebo data estimated from pooled dupilumab studies SINUS-24 and SINUS-52). Assuming 50-70% of participants will have co-morbid asthma, the sample size will also provide >80% power for the co-primary endpoints at a 2-sided 5% significance level in the co-morbid asthma subgroups.

2 ANALYSIS SETS

2.1 Definition of Analysis Sets

Participants must have provided their informed consent. If no signed informed consent is collected (major protocol deviation), then the participant will be excluded from all analysis sets defined below. In the case of severe site misconduct or suspected fraud, it may be appropriate to remove sites, participants, or repeated participant information from the reporting of the study. This must be assessed dependent on the individual circumstances, and would be detailed prior to unblinding in a formal file note.

Enrolled Analysis Set

This analysis set comprises all enrolled participants who signed the informed consent form, including screening failures, and will be used for the reporting of disposition.

Randomised Analysis Set

This analysis set comprises all participants randomised to study treatment, irrespective of whether IP was subsequently taken, and will also be used for the reporting of disposition.

2.1.1 Efficacy Analysis Set

Full Analysis Set (FAS)

This analysis set comprises all participants randomised to study treatment who received at least one dose of IP, irrespective of their protocol adherence and continued participation in the study.

Efficacy analyses will be performed using all participants in the FAS. Participants will be analysed according to their randomised treatment (including in the case of any discrepancies between randomised and actual treatment).

The FAS specifies which participants are included in efficacy analyses. Details of which data are included in efficacy analyses for these participants are given in the respective sections, notably in Section 3.1.3 and Section 4.2.

For consistency with efficacy analyses, demographics and baseline characteristics will be summarised using the FAS.

Co-morbid Asthma/AERD/NSAID-ERD Subset

All participants in FAS who have comorbid asthma/AERD/NSAID-ERD at baseline will be included in this subset, irrespective of their protocol adherence and continued participation in the study. Participants will be analysed according to their randomised treatment. ACQ-6, FEV_1, and FeNO will be analysed based on the co-morbid asthma/AERD/NSAID-ERD subset.

2.1.2 Safety Set (SS)

This analysis set comprises all participants who received at least one dose of IP.

Safety analyses will be performed using all participants in the safety analysis set. Participants will be analysed according to their actual treatment in the case of any discrepancies between randomised and actual treatment. Specifically, a participant who has on one or more occasion actually received active (tezepelumab) treatment will be assigned to the tezepelumab group, regardless of the randomised treatment assignment. A participant who has on no occasion actually received any active (tezepelumab) treatment will be assigned to the placebo group, regardless of the randomised treatment assignment.

Safety data will also be listed separately and discussed in the clinical study report (CSR) for any participant who received a treatment that was not the randomised treatment at one or more visits.

Summaries of ADA will also be based on the safety analysis set, using the same approach to handle treatment dispensing errors.

2.1.3 Other Analysis Sets

PK Analysis Set (PKS)

This analysis set comprises all participants in the FAS who received active (tezepelumab) treatment and had at least one detectable tezepelumab serum concentration from a sample

collected post-treatment that is assumed not to be affected by factors such as protocol deviations (PDs) (e.g. disallowed medication or incorrect study medication received). Summaries of PK will be based on the PK analysis set.

Additional FU Analysis Set

Additional FU analysis set comprises approximately the first 200 participants who have a 24-week follow-up period post EOT. Only participants who complete Week 64 visit are included in this analysis set, regardless of completion of Week 76 visit or withdrawal from the study after Week 64 visit.

2.1.4 Handling of Other Issues Impacting Analysis Sets

If it is found that any participant has been randomised on more than one occasion (contrary to the protocol) under different participant numbers, either at the same site or at different sites, then data corresponding to the first participant participation will be used in the analyses. Data associated with the second (and any subsequent) participation of the same participant will be listed and discussed in the CSR. All data associated with duplicate randomisations will be reviewed, and decisions regarding the analysis and reporting of these data will be documented, prior to unblinding at the primary database lock.

The above analysis set definitions assume the integrity of data captured from all participating sites in the trial. If it is deemed necessary to exclude participants from analysis sets due to suspected fraud/other serious non-compliance at a particular site, or to perform sensitivity analyses with participants from such a site removed for the same reason, the SAP will document this (amended if necessary) where possible prior to primary database lock. Otherwise, it will be fully described in the CSR. The SAP will not be updated after primary database lock.

In August 2023, the Japanese Ministry of Health, Labour and Welfare performed an inspection of an Site Management Organization (SMO) in Japan (Tokyo). This inspection identified significant deviations from GCP at the investigator sites supported by this SMO. Any Japanese participants enrolled from sites managed by the SMO (i.e. PPD PPD will be excluded from all analysis sets.

In May 2024, a critical data quality issue was identified at PPD The NPSD data for the PPD randomized participants PPD will be excluded from the analyses. Details will be described in the CSR.

2.2 Protocol Deviations

Only important protocol deviations (PDs) will be listed and tabulated in the CSR, and only for randomised participants (not screening failures). These are defined as PDs which may significantly affect the completeness, accuracy and/or reliability of the study data, or which may significantly affect a participant's rights, safety, or well-being. Important PDs in this trial will be grouped under one of the following categories:

- Did not fulfill key eligibility criteria
- Discontinuation criteria for study product met, but participant not withdrawn from study treatment
- Discontinuation criteria for overall study withdrawal met but patient not withdrawn from study
- Investigational product (IP) deviation
- Excluded medications taken
- Deviations related to study procedure
- Other important deviations

Important protocol deviations associated with the global/county situation will also be summarised separately (see Section 4.2.13).

Important PDs will be identified and documented by the study team prior to unblinding of the trial at the primary database lock. As far as possible, the occurrence of important PDs will be monitored (blinded) during the trial, with the emphasis on their future prevention.

With the exception of the PK analyses, important PDs will not be used to exclude any participant from any analysis set, nor to exclude any data from participants included in an analysis set.

The study non-compliance handling plan outlines the management of PDs and includes the proposed specific categories of PDs in this trial. Any PDs which are not defined as important will not be reported and discussed in the CSR.

3 PRIMARY AND SECONDARY VARIABLES

3.1 General Definitions

3.1.1 Definition of Baseline

In general, the last non-missing measurement on or prior to the date of randomisation will serve as the baseline measurement for efficacy variables. If there is no value on or prior to the date of randomisation, then the baseline value will be set to missing.

In general, the last non-missing measurement prior to first dose of study treatment will serve as the baseline measurement for safety variables. If there is no value prior to first dose of study treatment, then the baseline value will be set to missing.

Where unscheduled/repeat assessments are relevant and exist for any participant at a particular visit specified below, they will also be considered in the baseline definitions, provided they remain prior to the date of randomisation (efficacy) or the date of first dose of study treatment (safety).

For the NPSD (NCS included), the baseline is the average of daily responses from Day –13 to Day 0. The mean is calculated as the sum of all non-missing daily scores over these 14 sequential days divided by the number of non-missing daily scores. If more than 7 daily scores (>50%) within the baseline period are missing, then baseline is set to missing.

Of note, Day 0 is Day 1 in the analysis following the Clinical Data Interchange Standards Consortium (CDISC) standards (Section 3.1.4), although it is labelled as Day 0 in the CSP.

3.1.2 Absolute Change from Baseline

Absolute change from baseline is defined as *(post-baseline value – baseline value)*.

If either the post-baseline value or the baseline value is missing, then the absolute change from baseline will also be missing.

Unless otherwise specified, “change from baseline” is assumed to be the absolute change from baseline.

3.1.3 Study Periods

The following study periods are defined for analysis purposes:

- **Screening/run-in period**: starting on the date of the first study procedure and ending one day prior to randomisation (for randomised participants) or on the date of the last study procedure (for screening failures). If any participant is re-screened, the latest available screening will be used for this purpose.
- **Planned treatment period**: starting on the date of randomisation (efficacy) and ending on the date of the Week 52 visit or earlier study withdrawal date (for participants not followed up until Week 52). Unless noted otherwise, efficacy analyses will be based on planned treatment period.
- **On-treatment period**: starting on the date of first dose of IP and ending on the earlier (date of last dose of IP + 33 days, date of study withdrawal). Note: The 33 days is derived as dose frequency (28 days) plus 5 days as specified in the CSP Section 1.3.

- **Post-treatment period**: starting on the day after the end of the on-treatment period until the end of the study (Week 76 for the first ~200 participants and Week 64 for the remaining participants) or early withdrawal date.
- **On-study period** (planned treatment and follow-up): starting on the date of randomisation (efficacy) / date of first dose of IP (safety) and ending on the study completion or withdrawal date. For analyses performed at primary database lock, the on-study period is understood to include all data recorded up until the date of the data cut-off for the primary database lock (which includes all follow-up data available at the time of the data cut-off).
- **Additional FU period**: starting on the day after Week 64 visit until the end of the study (Week 76) or early withdrawal date. Only the first ~ 200 participants will continue follow-up beyond Week 64.

If a participant dies, the end dates of the study periods defined above will be capped at the date of death.

3.1.4 Visit Windows

All summaries and analyses, both efficacy and safety, which are presented by time point (e.g. “Week 52”) will use a visit window to classify the data record, which is derived from the assessment date relative to the reference start date. This approach allows appropriate classification of visits which may have occurred significantly earlier or later than the protocol assessment schedule, as well as the use of data captured at visits which have no fixed timing (notably the IPD visit), and the handling of data captured at visits for which the database label is incorrect and unresolvable.

Nominal database visit numbers will not be used in any summary or analysis by visit.

For Screening/run-in period, the relative day is defined as (Date of assessment – Date of randomization).

For periods after randomization:

For efficacy variables, the reference start date is the date of randomisation, and relative day is therefore defined as (Date of assessment – Date of randomisation) + 1.

For safety variables, the reference start date is the date of first dose of IP, and relative day is therefore defined as (Date of assessment – Date of first dose of IP) + 1.

Any data collected at unscheduled or repeat visits will be listed and will be included in baseline definitions (see Section 3.1.1), and in any definitions of maximum value, minimum value or last value within the relevant study period.

Data collected at unscheduled or repeat visits will also be included in visit windows, and therefore may be included in summaries or analyses by visit or used in any sensitivity analyses which involve imputation of data from participants with non-missing values to participants with missing values. In the case of a missing value at a scheduled visit, which is then followed by a non-missing value at an unscheduled or repeat assessment within the same visit window, the non-missing value at the unscheduled/repeat assessment will replace the missing value at the scheduled visit.

If a participant has more than one non-missing value within the same visit window, the following rules will apply:

- The non-missing value closest to the target day will be selected for analysis at that visit.
- If two non-missing values are equidistant from the target day, the earlier of the two values will be selected for the analysis at that visit.
- If two non-missing values are recorded on the same day and have a different assessment time, the value with the earliest assessment time will be selected for analysis at that visit.
- If two non-missing values (for continuous variables) are recorded on the same day and have no assessment time associated with at least one of them, or the same assessment time associated with both of them, the average of the two values will be selected for analysis at that visit. For categorical variables in this situation, the worst case will be used.
- If there are multiple ADA samples in the same visit window with both positive and negative results, the sample with a positive result and the highest titre value will be selected.

If a participant has no value within a particular visit window, then the participant will have a missing value at that visit in summaries and analysis.

The same visit window definitions below will be used regardless of whether the planned treatment period or the on-treatment period is used for analysis (see Section 3.1.3). In practice, each data record in the planned treatment period will be first identified, and then further flagged according to whether it is on-treatment or off-treatment. This flag will be used to select all eligible records for subsequent visit windowing, according to whether the derived visits are to be used in a planned treatment period or an on-treatment period analysis. It should be noted that, if treatment was discontinued within a particular visit window, the rules above for handling multiple values within the same visit window could select a different record according to whether a planned treatment period analysis or an on-treatment period analysis is needed.

Table 1 summarises the visit windows to be used for all variables unless specified otherwise. It corresponds to the full (mostly 4-weekly) protocol scheduling for clinic visits and will be

used for all variables by default, including those variables which are not capture at every clinic visit, unless indicated below that the visit windows in Table 2 or Table 3 should be used. Of note, Week 0 is Day 1 in the analysis following the Clinical Data Interchange Standards Consortium (CDISC) standards, although it is labelled as Day 0 in the CSP.

Table 1. Visit Windows – All Variables

Time Point	Target Day	Visit Window
Baseline (Week 0)	See Section 3.1.1 for baseline definitions	
Week 2	15	2-21
Week 4	29	22-42
Week 8	57	43-70
Week 12	85	71-98
Week 16	113	99-126
Week 20	141	127-154
Week 24	169	155-182
Week 28	197	183-210
Week 32	225	211-238
Week 36	253	239-266
Week 40	281	267-294
Week 44	309	295-322
Week 48	337	323-350
Week 52	365	351-378
Follow-up Week 58	407	379-427
Follow-up Week 64	449	428-469
Follow-up Week 70	491	470-511
Follow-up Week 76	533	512-553

Table 2 summarises the adjusted analysis-defined visit windows for NPS and variables in which the sparsest scheduling is planned in the protocol. These variables include: WPAI, ECG, spirometry, FeNO, ACQ-6, and total IgE.

Table 2 Visit Windows – NPS and Sparse protocol schedule

Time Point	Target Day	Visit Window
Baseline (Week 0)	See Section 3.1.1 for baseline definitions	
Week 4	29	2-56
Week 12	85	57-126
Week 24	169	127-210
Week 36	253	211-308
Week 52	365	309-406
Follow-up Week 64	449	407-490
Follow-up Week 76 (only NPS, ECG, spirometry, ACQ-6)	533	491-574

Table 3 summarises the visit windows which will be used for NCS and other items in the NPSD which will be summarised every two weeks (bi-weekly).

Table 3 Visit Windows – NPSD (including NCS)

Time Point	Target Day	Visit Window
Baseline (Week 0)	See Section 3.1.1 for baseline definitions	
Week 2	15	2-15
Week 4	29	16-29
Week 6	43	30-43
…	…	…
Week 52	365	352-365
…	…	…
Follow-up Week 76	533	520-533

In all cases above, no time points will be presented in summary tables or included in statistical analysis which do not correspond to the time points scheduled in the protocol for the variable in question.

For CT scan, the visit window for Week 52 will include Days 1 – 490.

3.1.5 Prior and Concomitant Medications

Medications taken by any participant at any time during the study will be coded using the anatomical therapeutic classification (ATC) system within the World Health Organization (WHO) Drug Dictionary.

Medications will be categorised for analysis according to their onset and end dates as follows:

- Prior medications:
    - end date ≤ date of first dose of IP
- Concomitant medications during on-treatment period:
    - end date > date of first dose of IP and start date ≤ earliest (date of last dose of IP + 33 days, date of death, date of study withdrawal), or
    - end date ongoing and start date ≤ earliest (date of last dose of IP + 33 days, date of death, date of study withdrawal)
- Concomitant medications during post-treatment period (for participants still being followed up then):
    - start date > date of last dose of IP + 33 days.

The handling of partial/missing dates for prior/concomitant medications is detailed in Appendix 8.3.

Essentially the above says that:

- Prior and concomitant medications are mutually exclusive.
- Concomitant medications on-treatment and post-treatment are also mutually exclusive (here, the word “concomitant” means concomitant with study procedures, irrespective of whether IP was still being taken). Specifically, a concomitant medication which started on-treatment and ended post-treatment will only be considered on-treatment.

If the medication record has a completely missing onset date, the participant will be assumed to have been on the medication on the date of the first study procedure. If the medication record has a partially missing onset date (month/year or year only) which is the same as that for the end of IP treatment, it will be assumed to have started on-treatment. If the medication record has a partially missing onset date (month/year or year only) which is the same as that for the start of IP treatment, it will be assumed to have started before treatment.

If the medication record has a completely missing end date, the participant will be assumed to have been on the medication on the date of study completion or withdrawal. If the medication record has a partially missing end date (month/year or year only) which is the same as that for start of IP treatment, it will be assumed to have ended on-treatment. If the medication record has a partially missing end date (month/year or year only) which is the same as that for end of IP treatment, it will be assumed to have ended post-treatment.

Background INCS/MFNS

The background INCS medications should be maintained at a stable dose from Visit 1 and throughout the screening and study period and no other medications intended to prevent polyp recurrence should be used. All participants will use INCS for a minimum of 4 weeks prior to V3 and continued throughout the study period.

Standardized MFNS (50 µg/actuation) nasal spray is contained in a bottle with 120 actuations for US and 140 actuations for all other countries. Two doses of MFNS (50µg/actuation) in each nostril twice daily (total 400µg daily) or equivalent INCS will be administered unless there is a medical rationale to use the lower dose (QD) regimen.

3.1.6 Subgroups

The following subgroups are defined for the purposes of efficacy subgroup analysis.

- Age category at screening visit (<65, ≥65 years)
- Gender at screening visit (male, female)
- Race at screening visit (white, non-white)
- Prior surgery for NP status (yes, no)
- Allergic rhinitis at screening visit (yes, no)
- Co-morbid asthma/AERD/NSAID-ERD status at screening visit (yes, no)
- AERD/NSAID-ERD at screening visit (yes, no)
- Baseline eosinophil count (<150, ≥150 to <300, ≥300 eosinophils/µL)
- Baseline BMI (< 30 kg/m2, ≥ 30 kg/m2)
- Region (China, Japan, ROW)
- Region (Asia Pacific including China and Japan, Europe, North America)
- Baseline NPS (<median, ≥ median)
- Baseline NCS (< median, ≥ median)
- Time since last NP surgery at screening visit (<3 years, ≥3 years)
- Number of previous NP surgeries (0, 1, 2 or more)
- Prior SCS use for NP (yes, no)

- Baseline weight quartiles
- Baseline staphylococcus aureus colonization in nasal culture (yes, no, not applicable)
- Baseline eosinophil count quartiles
- Baseline total serum IgE quartiles
- Baseline perennial specific IgE status:
    - Any perennial FEIA positive: requires 1 or more specific IgE (FEIA) panels to be positive. Provided that at least one IgE panel is positive, no further requirement is made for data on all 7 panels for all countries, excluding China, and all 6 panels for China only to be available.
    - All perennial FEIA negative: requires all 7 specific IgE panels to be negative for all countries, excluding China, and all 6 specific IgE panels to be negative for China.
    - If there are fewer than 7 panels for all countries, excluding China and less than 6 panels for China with data available and none of these is positive, then IgE status is considered “Unknown perennial FEIA”.
    - Positive is defined as a value ≥0.35 kU/L.
    - The 7 panels for all countries, excluding China include: IgE American Cockroach, IgE Cat Dander, IgE D. farina, IgE D. pteronyssinus, IgE Dog Dander, IgE German Cockroach, and IgE Mould Mix 1..
    - The 6 panels for China only include: IgE Dog Dander, IgE Cat Dander, IgE D. farinae, D. pteronyssinus, IgE German Cockroach, and IgE Mould Mix 1.
- Baseline seasonal specific IgE status:
    - Any seasonal FEIA positive: requires 1 or more specific IgE (FEIA) panels to be positive. Provided that at least one IgE panel is positive, no further requirement is made for data on all 4 panels for all countries, excluding China and all 4 panels for China only to be available.
    - All seasonal FEIA negative: requires all 4 specific IgE panels for all countries, excluding China and all 4 specific IgE panels for China only to be negative.
    - If there are fewer than 4 panels for all countries, excluding China and China only with data available and none of these is positive, then IgE status is considered “Unknown seasonal FEIA”.
    - Positive is defined as a value ≥0.35 kU/L.
    - The 4 panels for all countries, excluding China include: IgE Grass Mix Pollen, IgE Silver Birch Pollen, IgE Weed Mix Pollen, and IgE Japanese Cedar.
    - The 4 panels for China only include: IgE Common Ragweed, IgE Common Silver Birch, IgE Mugwort, and IgE Goosefoot.

3.1.7 Other Baseline Characteristics

In addition to the subgroups in Section 3.1.6, the following baseline characteristics will be provided:

- Weight
- Height
- Baseline NPS, NCS, loss of smell, SNOT-22, LMK, and NPSD TSS
- Baseline lung function (FEV_1, FEF 25-75%, FVC), ACQ-6, and FeNO for the asthma/AERD/NSAID-ERD subset
- Baseline NPIF
- Baseline eosinophils and IgE
- JESREC scores (<11, ≥11) including participants from China only (Tokunaga et al., 2015)

3.1.8 Disposition

The following definitions will be used for time to event variables in Kaplan-Meier disposition plots:

Time to last dose of IP

Time to last dose of IP will be defined as follows:

Time to last dose (days) = [Date of last dose of IP from eCRF – date of first dose of IP] + 1.

Date of last dose of IP will be the date of last dose taken from the "Discontinuation of Investigational Product" electronic case report form (eCRF) page for all participants, including those who prematurely discontinue IP as well as those who complete IP dosing as per protocol.

Time to premature study withdrawal

Time to premature study withdrawal will be defined as follows:

Time to premature study withdrawal (days) = [study withdrawal date from eCRF – date of randomisation] + 1.

Study withdrawal date will be the completion or discontinuation date from the "Disposition" eCRF page, where any participant status other than "Completed" has been entered.

Participants who did not prematurely withdraw from study will be censored at one of the following dates:

- Completion or discontinuation date from the "Disposition" eCRF page, where participant status of "Completed" has been entered.
- At primary database lock (DBL), ongoing participants will be censored at data cut-off date.

3.2 Primary Efficacy Variables

3.2.1 Change from baseline in total nasal polyp score at Week 52

The total NPS will be used to evaluate participant's NP size throughout the study by the nasal endoscopy. The total NPS and the changes from baseline at Week 52 will be calculated for the co-primary efficacy endpoint.

The total NPS is the sum of the right and left nostril scores (maximum of 8), as evaluated by nasal endoscopy. The left and right score will be based on a central read with a scale from 0 to 4 as listed in Table 4. Each nasal endoscopy is evaluated by two independent physician reviewers. The process of evaluation differs between confirmation of eligibility criteria and collection of scores for efficacy analysis. As a consequence, the NPS scores are derived differently for eligibility and efficacy use.

NPS score derivation to confirm eligibility criteria at Visit 2: If polyp scores for left and right nostril (maximum of 4 per each nostril) provided by two reviewers are:

- exactly the same, then the NPS score is the sum of right and left nostril scores provided by one reviewer,
- different for right and/or left nostril but the sum for each reviewer is >5, then the NPS score is calculated as the mean of right and left nostril scores provided by both reviewers,
- different for right and/or left nostril and the sum for one reviewer is <5, then the nasal endoscopy is evaluated by an independent adjudicator. The NPS score is the sum of left and right nostril scores provided by the reviewer selected by adjudicator.

NPS score derivation for efficacy analysis use: If polyp scores for left and right nostril provided by two reviewers at the same visit are:

- exactly the same, then the NPS score is the sum of right and left nostril scores provided by one reviewer at this visit,
- different for right and/or left nostril, then the nasal endoscopy is evaluated by an independent adjudicator. The NPS score at this visit is calculated as the sum of scores for left and/or right nostril selected by adjudicator, ie. there is no requirement the right and left nostril scores selected by adjudicator are provided by the same reviewer.

Table 4 Endoscopic Nasal Polyp Score Within a Nostril

Polyp score	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	Polyps reaching the lower border of the inferior turbinate or a middle meatal polyp with a score of 2 with any additional polyp medial to the middle turbinate
4	Large polyps causing complete or near-complete obstruction of the inferior nasal cavity i.e. touching the floor of the nose

3.2.2 Change from baseline in nasal congestion score at Week 52

Participant reported nasal congestion (NC) will be evaluated as part of the NPSD. The NCS (nasal congestion score) is captured by one item in the NPSD asking participants to rate the severity of their worst NC over the past 24 hours using the following response options: 0 – None; 1 – Mild; 2 – Moderate; 3 – Severe. Baseline will be the mean of daily responses from Day -13 to Day 0. Bi-weekly (14-day) mean NCS will be calculated if at least 8 days in each 14-day period has evaluable data; otherwise the bi-weekly mean is set to missing. The NCS and the changes from baseline at Week 52 will be calculated for the co-primary efficacy endpoint.

Additional, based on the Meaningful Change Thresholds (MCT) reported in Shih et al., 2023, participants with change from baseline in bi-weekly mean NCS ≤ -1.0 will be defined as NCS responders.

3.3 Key Secondary Efficacy Variables

3.3.1 Change from baseline in loss of smell at Week 52

Participant reported sense of smell will be evaluated as part of the NPSD. Loss of smell is captured by the DSS item (difficulty with sense of smell) in the NPSD asking participants to rate the severity of their worst difficulty with sense of smell over the past 24 hours using the following response options: 0 – None; 1 – Mild; 2 – Moderate; 3 – Severe. Baseline will be the mean of daily responses to the from Day -13 to Day 0. Bi-weekly (14-day) mean loss of smell will be calculated if at least 8 days in each 14-day period has evaluable data; otherwise the bi-weekly mean is set to missing.

3.3.2 Change from baseline in sinonasal outcome test score at Week 52

The SNOT-22 is a condition-specific HRQoL assessment which captures participant-reported physical problems, functional limitations, and emotional consequences of sinonasal

conditions. Patient-reported symptom severity and symptom impact over the past 2 weeks are captured via a 6-point scale (0 = No Problem to 5 = Problem as bad as it can be). The total score is the sum of all item scores and has a range from 0 to 110 (higher scores indicate poorer outcomes). The total score and the changes from baseline will be calculated.

3.3.3 Time to (and proportion of participants with) nasal polyp surgery decision and/or systemic corticosteroids for nasal polyposis up to Week 52

Surgery is defined as any procedure involving instruments resulting in incision and removal of tissue (e.g., polypectomy, endoscopic sinus surgery). Rescue treatment of NP is defined as requiring treatment with systemic corticosteroids (SCS) for at least 3 consecutive days (a single depo-injectable dose of corticosteroids will be considered equivalent to a 3-day course of systemic corticosteroids).

The following variables up to Week 52 will be derived for all participants:

- Time to first NP surgery decision or SCS for NP

  Time to first NP surgery decision or SCS for NP = (date of the first NP surgery decision or start date of first SCS for NP use – date of randomisation)+1

  Participants who do not require NP surgery or do not receive SCS for NP will be censored on the date of the end of the planned treatment period defined in Section 3.1.3.

- Time to first NP surgery decision

  Time to first NP surgery decision = (date of the first NP surgery decision – date of randomisation)+1

  Participants who do not require any NP surgery will be censored on the date of the end of the planned treatment period defined in Section 3.1.3.

- Time to first SCS use for NP

  Time to first SCS use for NP = (start date of the first SCS for NP use – date of randomisation)+1

Participants who do not receive SCS for NP will be censored on the date of the end of the planned treatment period defined in Section 3.1.3.

Kaplan-Meier methods will be used to estimate the cumulative proportion of participants who had NP surgery decision, SCS for NP, and NP surgery decision or SCS for NP up to Week 52.

Time to first NP surgery and time to first NP surgery or SCS for NP will be derived in a similar way but based on the actual surgery date.

3.3.4 Change from baseline in Lund Mackay score (LMK) at Week 52

The LMK score provides a semi-quantitative assessment of nasal sinuses on sinus CT scans and will be used to assess sinus opacification. Based on the sinus CT images, the five sinuses (maxillary, anterior ethmoid, posterior ethmoid, sphenoid and frontal) on each side are scored by central radiologist as follows:

Table 5 Lund-Mackay score

Score	CT scan assessment
0	No abnormality
1	Partial opacification
2	Total opacification

The osteomeatal complex (Table 6) is scored for the right and left sides as follows:

Table 6 Osteomeatal complex score

Score	CT scan assessment
0	Not occluded
2	Occluded

The maximum total LMK score is 24. The change from baseline in LMK score evaluated by CT at Week 52 will be presented as a key secondary endpoint.

3.3.5 Change from baseline in NPSD total symptom score at Week 52

The NPSD TSS consists of 8 out of the 11-items on the NPSD symptom diary, which evaluates a participant's experience with NP over the past 24 hours. The 8 items supporting the TSS ask participants to report their experience with nasal blockage, NC, runny nose, postnasal drip (mucus drainage down the throat), headache, facial pain, facial pressure, and difficulty with sense of smell. Participants report the severity of each symptom at its worst using a 4-point verbal rating scale (0-None to 3-Severe). The TSS is calculated by taking the sum of the 8 equally weighted symptom items. The total TSS and the changes from baseline will be calculated.

3.3.6 Change from baseline in pre-BD FEV_1 at Week 52 (in participants with co-morbid asthma/AERD/NSAID-ERD)

Lung function in participants with co-morbid asthma/AERD/NSAID-ERD will be measured by spirometry and assessed using FEV_1. Only those spirometry tracings determined to be acceptable or borderline will be used. The measurement used in the analysis is the one marked with Best Result Flag (not necessarily the highest) equal to "Yes" as per Data Transfer Specification. The change from baseline in pre-BD FEV_1 at Week 52 will be presented.

3.4 Other Secondary Efficacy Variables

3.4.1 Change from baseline in nasal polyp score through Week 52

The total NPS evaluated by nasal endoscopy over time through Week 52 (Baseline, Weeks 4, 12, 24, 36, 52) will be evaluated as described in Section 3.2.1 above. Additionally, the proportion of participants with a ≥1 point reduction and a ≥2 points reduction in NPS through Week 52 will be presented.

3.4.2 Change from baseline in nasal congestion score through Week 52

Participant reported NC evaluated as part of the NPSD over time through Week 52 (every 14-days) will be evaluated as described in Section 3.2.2 above.

3.4.3 Change from baseline in UPSIT test at Week 52

Participant reported loss of smell will be evaluated using the University of Pennsylvania Smell Identification Test (UPSIT) test of olfactory function which uses microencapsulated odorants that are released by scratching standardized odour-impregnated test booklets (Doty et al 1984). Four booklets with 10 odorants each are used for the test. Participants are asked to identify the odour using multiple choice format which lists different possibilities. Scores are based on number of correctly identified odours with a total score ranging from 0 to 40. The olfactory diagnosis will be classified based on the test scores by gender (male vs female) as listed in Table 7.

Table 7 UPSIT Olfactory Diagnosis

Olfactory Diagnosis	Test Scores (Male)	Test Scores (Female)
Anosmia	00 - 18	00 - 18
Severe Microsmia	19 - 25	19 – 25
Moderate Microsmia	26 - 29	26 – 30
Mild Microsmia	30 - 33	31 – 34
Normosmia	34 - 40	35 - 40

3.4.4 Change from baseline in modified LMK score at Week 52

The modified Lund-Mackay or Zinreich scoring system will be utilized to score the same CT images scored using the LMK to assess sinus opacification. Based on the sinus CT images, the five sinuses (maxillary, anterior ethmoid, posterior ethmoid, sphenoid and frontal) on each side will be scored based on the percentage of opacification from mucosal thickening according to Table 8.

Table 8 Modified Lund Mackay (Zinreich) Score

Score	Percent opacification
0	0%
1	1% - 25%
2	26% - 50%
3	51% - 75%
4	76% -99%
5	100%

The maximum total Zinreich score is 50 without the Osteomeatal complex score described in Section 3.3.4 above.

3.4.5 Exposure of systemic corticosteroid use over Week 52

Rescue treatment for NP or other reasons is defined as requiring treatment with SCS for at least 3 consecutive days (a single depo-injectable dose of corticosteroids will be considered equivalent to a 3-day course of SCS). A course of SCS is considered continuous if treatment episodes are separated by less than 7 days. The following variables over 52 weeks will be derived for all participants:

- The number of courses of SCS for NP.
- Total SCS for NP dose (converted to prednisolone equivalents according to Appendix 8.2) = SCS dose* duration of SCS for NP. If multiple courses are taken, the total is the sum of all courses. Any SCS treatment given for less than 3 consecutive days will not be included.
- Total duration of SCS for NP = the sum of days a participant receives SCS for NP. Any SCS treatment given for less than 3 consecutive days will not be included.

SCS for NP is recorded on Medication of Interests (MINT) form with the ATC codes H02AB and H02BX, and Therapy Reason Chronic rhinosinusitis with nasal polyps.

The same analysis above will be repeated for SCS for any reason and SCS for any reason excluding NP.

3.4.6 Change from baseline by domain of nasal polyposis symptom diary through Week 52

The NPSD is an 11-item daily questionnaire in which the participant is asked to consider their experience with NP over the past 24 hours. Participants are asked to report their experience with 8 NP symptoms (nasal blockage, nasal congestion, runny nose, postnasal drip (mucus drainage down the throat), headache, facial pain, facial pressure, and difficulty with sense of smell) and 2 symptom impacts (difficulty with sleeping due to nasal symptoms and difficulty with daily activities due to nasal symptoms). Participants report the severity of each symptom and symptom impact at its worst using a 4-point verbal rating scale (0-None to 3-Severe). A single item to capture INCS compliance (yes or no) will be administered after the symptom and symptom impact items. The score and changes from baseline through Week 52 (every 14-days) by component will be calculated.

3.4.7 Change from baseline in nasal peak inspiratory flow (NPIF) through Week 52

Nasal peak inspiratory flow evaluation represents a physiologic measure of the air flow through both nasal cavities during forced inspiration expressed in liters per minute. The NPIF is the best validated technique for the evaluation of nasal flow through the nose. Nasal inspiration correlates most with the participative feeling of obstruction and is the best validated technique for monitoring nasal flow in clinical trials.

Up to a maximum of eight NPIF efforts will be performed by the participant; all values will be recorded by the participant on the MasterScope, and the highest value will be used for evaluation. Total NPIF values and changes from baseline through Week 52 (Weeks 0 (Baseline), 12, 24, 36, 52) will be calculated.

3.4.8 Change from baseline in asthma control questionnaire (ACQ-6) at Week 52 for participants with co-morbid asthma/AERD/NSAID-ERD

The ACQ-6 is an assessment of asthma symptoms (night-time waking, symptoms on waking, activity limitation, shortness of breath, wheezing, and short acting beta-agonist use). Participants are asked to recall their level of asthma control during the previous week by responding to one BD use question and 5 symptom questions.

Questions are weighted equally and scored from 0 (totally controlled) to 6 (severely uncontrolled). The ACQ-6 total score is computed as the unweighted mean of the responses to the 6 questions. If response to any of the questions is missing, the ACQ-6 will be missing. Mean scores of ≤0.75 indicate well-controlled asthma, scores between 0.75 and <1.5 indicate

partly controlled asthma, and a score ≥1.5 indicates not well controlled asthma (Juniper et al 2006). Individual changes of at least 0.5 are considered to be clinically meaningful. Total ACQ-6 score and changes from baseline through Week 52 will be calculated.

Additional variables based on ACQ-6 will be reported at each time point include:

- ACQ-6 responder:
    - Responder: Change from baseline ACQ-6 score ≤ -0.5
    - Non-responder: Change from baseline ACQ-6 score > -0.5
- ACQ-6 response:
    - Improvement: Change from baseline ACQ-6 score ≤ -0.5
    - No change: -0.5 < Change from baseline ACQ-6 score < 0.5
    - Deterioration: Change from baseline ACQ-6 score ≥ 0.5
- Participant's asthma control as measured by ACQ-6 score:
    - Well controlled: ACQ-6 score ≤ 0.75
    - Partly controlled: 0.75 < ACQ-6 score < 1.5
    - Not well controlled: ACQ-6 score ≥ 1.5.

3.4.9 Sinus severity score

Quantitative assessment of sinus CT image data will be used to derive an objective measure of sinus disease burden called sinus severity score.

The sinus severity score is defined as: *(sinus mucosal volume)/(sinus mucosal volume + sinus air volume)×100%.*

Both observed values and the changes from baseline will be calculated.

3.5 Exploratory Efficacy Variables

3.5.1 Post-treatment effect and recurrence rate of nasal polyposis

The off-treatment effect and recurrence rate of nasal polyps after discontinuation of treatment with tezepelumab will be explored. Total NPS and the changes from baseline during the post-treatment period will be calculated. Total NCS and the changes from baseline during the post-treatment period will also be calculated.

3.5.2 European Quality of Life-5 Dimensions-5 Levels (EQ-5D-5L)

The European Quality of Life-5 Dimensions questionnaire assesses 5 dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Each dimension has 5 response options (no problems, slight problems, moderate problems, severe problems, and

extreme problems) that reflect increasing levels of difficulty. The participant will be asked to indicate his/her current health state by selecting the most appropriate level in each of the 5 dimensions. Total score and changes from baseline for an overall utility score of all dimensions as well as individual dimension scores will be calculated.

The questionnaire also includes a visual acuity scale (VAS), where the participant will be asked to rate current health status on a scale of 0 to 100, with 0 being the worst imaginable health state. Total VAS score and changes from baseline will be calculated.

The health state valuation (an index-based value) for the EQ-5D-5L will be derived from the 5 dimensions using the UK population-based preference weights. Further details are given in van Hout et al., 2012 and Devlin et al., 2017.

3.5.3 Patient Global Impression of Severity and Patient Global Impression of Change

The PGI-S is a single item designed to capture the participant's perception of overall symptom severity at the time of completion using a 6-point categorical response scale of no symptoms to very severe symptoms (i.e., 0 = no symptoms, 1 = very mild, 2 = mild, 3 = moderate, 4 = severe, 5 = very severe). The PGI-C captures the participant's overall evaluation of response to treatment. The participants are asked to report the degree to which their health status has changed since entering the treatment period using a 7-point scale of very much improved to no changes to very much worse (i.e., 1 = much better, 2 = moderately better, 3 = a little better, 4 = about the same, 5 = a little worse, 6 = moderately worse, and 7 = much worse).

For the PGI-C, participants will also be categorized as Not better, At least a little better, At least moderately better, and Much better according to the following post-baseline responses.

- Much better: participants in this category will include those with a response of 'Much better'.
- At least moderately better: participants in this category will include those with responses of 'Much better' and 'Moderately better'.
- At least a little better: participants in this category will include those with responses of 'Much better', 'Moderately better', and 'A little better'.
- Not better: participants in this category will include those with responses of 'About the same', 'A little worse', 'Moderately worse', and 'Much worse'.

Participants can be counted in more than one category at a given time point.

Calculation of percentages will be based on the number of participants in the FAS with a completed assessment. There will be no imputation for missing values.

3.5.4 Airway Inflammation in Participants with Co-morbid Asthma/AERD/NSAID-ERD

A standardized single-breath FeNO test will be used to evaluate airway inflammation in the co-morbid asthma/AERD/NSAID-ERD subset.

It is expected that one technically acceptable FeNO measurement will be performed at each relevant visit. If more than one technically acceptable FeNO measurement is available on the same date at the clinic, all data will be transferred, and the first available technically acceptable FeNO measurement on that date will be used. Multiple FeNO measurements on different dates will be handled according to the rules for unscheduled/repeat visits (see Section 3.1.4).

FeNO results and changes from baseline will be calculated.

3.6 Safety Variable Derivations

3.6.1 Exposure and Treatment Compliance

Extent of exposure to IP is defined as the number of days between the date of first dose of IP and the date of last dose of IP inclusive, plus the number of days allowance for the dosing interval, that is:

Extent of exposure (days) = earliest (date of last dose of IP + 33 days; date of death; date of study withdrawal) – date of first dose of IP + 1

This calculation does not consider any gaps in exposure caused by the participant missing one or more intermediate scheduled 4-weekly doses. Such cases will be identified in the CSR if they occur but will not explicitly be accounted for in any analysis.

The total participant-years exposure for a treatment group will be derived as the sum of the individual participant extents of exposure (days) for that treatment group divided by 365.25.

Treatment compliance will be calculated as follows:

Treatment compliance (%) = [(Total number of actual dosing occasions/total number of expected dosing occasions) × 100%

In order to allow for participants who discontinue IP early in the compliance calculation, the number of expected dosing occasions will be calculated as the number of scheduled dosing visits up to and including the last available dosing visit for that participant.

INCS/MFNS compliance are recorded by the participant in the ePRO diary and checked by Investigators during the study period. Compliance for background INCS/MFNS will be calculated as follows for the planned treatment period:

INCS/MFNS compliance (%) = number of days when the participant is compliant to the background INCS/MFNS / days with diary × 100%.

3.6.2 General Adverse Events

Adverse events experienced by any participant at any time during the entire study will be coded using Medical Dictionary for Regulatory Activities (MedDRA V27.0). AEs will be categorised for analysis according to their onset date into the following study periods:

- AEs occurring during screening/run-in period: date of Visit 1 ≤ AE onset date < date of first dose of IP
- AEs occurring during on-treatment period: date of first dose of IP ≤ AE onset date ≤ earliest (date of last dose of IP + 33 days, date of death, date of study withdrawal)
- AEs occurring during post-treatment period: date of last dose of IP + 33 days < AE onset date ≤ study completion or withdrawal date
- AEs occurring during on-study period: date of first dose of IP ≤ AE onset date ≤ study completion or withdrawal date.
- AEs occurring during the additional FU period: date of Visit 19 ≤ AE onset date ≤ study completion or withdrawal date.

If the AE has a completely missing (and unresolvable) onset date, then the AE will be assumed to have occurred during the on-treatment period, unless the end date indicates unambiguously that the AE resolved before treatment started. If the AE has a partially missing (and unresolvable) onset date, then the AE will also be assumed to have occurred during the on-treatment period, unless either the end date indicates unambiguously that the AE resolved before treatment started, or the partial onset date is in the month/year prior to start of treatment.

The handling of partial/missing dates for AEs is detailed in Appendix 8.3.

3.6.3 Adverse Events of Special Interest

The protocol specifies adverse events of special interest (AESIs) as those which merit special attention in this trial, and for which derivation details (for those derived from the eCRF), or a statement when the derivation needs to be referenced externally to the SAP (for those derived from MedDRA dictionary terms), are given in Appendix 8.1.

3.6.4 Laboratory Variables

Clinical chemistry and haematology will be performed by a central laboratory according to the schedule and the variable specifications described in the CSP Table 12. Urinalysis for China participants will also be performed by central laboratory. Urinalysis (dipstick) for Non-China

participantswill be completed and assessed at site and sent for analysis at the central laboratory only if a positive dipstick result for any parameter is observed.

Changes from baseline in continuous laboratory variables will be calculated at relevant visits as specified in Section 3.1.1 and Section 3.1.2.

In all analyses of continuous laboratory variables, any value recorded as below the Lower Limit of Quantification (LLOQ) will be set to LLOQ/2 and included in the analysis. Any value recorded as above the Upper Limit of Quantification (ULOQ) will be set to ULOQ and included in the analysis.

Absolute values will be compared to the relevant normal reference range, as provided by the central laboratory, and classified as low (below range), normal (within range or on the limits) or high (above range). All values falling outside the normal reference ranges will be flagged. These classifications will also be used for shift tables.

For the purpose of shift tables, baseline will be defined as specified in Section 3.1.1. Minimum, maximum and last values calculated across all visits in the relevant study period will use all available values, including those from unscheduled and repeat visits, and irrespective of whether the values have been selected for use in summaries using visit windows (see Section 3.1.4).

Liver function tests will also be evaluated as multiples of the upper limit of the normal reference range (ULN). Participants who meet any of the following criteria at any time during the study will be flagged:

- Aspartate aminotransferase (AST) ≥ 3 x ULN
- Alanine aminotransferase (ALT) ≥ 3 x ULN
- Total bilirubin (TBL) ≥ 2 x ULN

Other multiples of ULN will also be used in the display of liver function tests.

3.6.5 Vital Signs

Changes from baseline in vital signs (pulse rate, systolic blood pressure (BP), diastolic BP, respiratory rate, body temperature, body weight and BMI) will be calculated at relevant visits as specified in Section 3.1.1 and Section 3.1.2.

BMI is calculated as:

$$BMI = Weight\ (kg) / [Height\ (m)]^2$$

Absolute values and changes from baseline (where applicable) will be compared to the relevant reference range tabulated below, and classified as low (below range), normal (within range or on the limits) or high (above range). All values falling outside the reference ranges will be flagged.

Table 9 Vital Signs Reference Ranges

Parameter	Standard Unit	Lower Limit	Upper Limit	Change from Baseline Criteria
Diastolic BP (sitting)	mmHg	60	100	±15
Systolic BP (sitting)	mmHg	90	160	±30
Pulse Rate (sitting)	beats/min	50	100	±20
Respiratory Rate	breaths/min	8	20	
Body Temperature	Celsius	36.0	37.5	
Weight	Kg	40	150	

3.6.6 ECG

The outcome of the overall evaluation (normal, abnormal or borderline) will be taken directly from the eCRF, as will the assessment of clinical significance.

Changes from baseline in continuous ECG variables (data provided external to the eCRF) will be calculated at relevant visits as specified in Section 3.1.1 and Section 3.1.2.

3.6.7 Physical Exam

A complete physical examination will be performed and include assessments of the following body systems: general appearance, respiratory, cardiovascular, abdomen, skin, head and neck (including ears, eyes, nose and throat), lymph nodes, thyroid, muscular-skeletal (including spine and extremities) and neurological systems.

A brief physical examination will include, at a minimum, assessments of the skin, lungs, cardiovascular system, and abdomen (liver and spleen). For the brief physical assessment only, information on whether the assessment was performed or not is to be recorded.

Clinical signs related to previous serious illnesses as new or worsening abnormalities may be recorded as AEs.

3.6.8 Medical History

If a partial diagnosis date is available only, the following rules will be used to impute a complete date (eg, for derivation of time since diagnosis):

If both the month and the year are available, the first of the recorded month will be imputed. If only the year is available, January 1st will be imputed. Imputed date will be capped by Date of Birth if applicable.

3.7 Pharmacokinetic and Immunogenicity Variables

Serum samples for determination of tezepelumab concentrations and the presence of ADA and neutralising antibodies (nAb) will be collected at baseline prior to first IP administration, at multiple time points before IP administration during the treatment period, and at selected timepoints in the follow-up period, according to the CSP schedule of assessments.

Samples will be used to determine tezepelumab concentrations, and to measure the presence of ADA and nAb using validated assays performed by a designated third-party vendor. Details of the bioanalytical methods used will be described in a separate bioanalytical report.

For immunogenicity, tiered analysis will be performed to include screening, confirmatory, and ADA titre assays. Samples that are confirmed positive for ADAs will be further analysed for the presence of nAb.

The third-party vendor analysing the PK samples will be unblinded to the randomised treatment assignments of all participants; no one from the study team will have access to the PK data until after the study has been unblinded. The assay for determination of tezepelumab concentrations will only be performed using samples for participants randomised to tezepelumab. Participants who are randomised to placebo will not have their PK samples analysed by the vendor laboratory. The ADA samples from all participants, regardless of treatment assignment, will be analysed.

4 ANALYSIS METHODS

4.1 General Principles

4.1.1 Statistical Hypotheses for Confirmatory Endpoints

The following 2-sided hypotheses will be evaluated in this trial:

Endpoint	Null Hypotheses	Alternative Hypotheses	Direction of superiority of tezepelumab is indicated by
Co-primary endpoints			
NPS at Week 52	H01a: Difference in mean change from baseline in NPS at 52 weeks (tezepelumab minus placebo) = 0	H11a: Difference in mean change from baseline in NPS at 52 weeks (tezepelumab minus placebo) ≠ 0	A difference in means less than 0.

Endpoint	Null Hypotheses	Alternative Hypotheses	Direction of superiority of tezepelumab is indicated by
NCS at Week 52	H01b: Difference in mean change from baseline in NCS at 52 weeks (tezepelumab minus placebo) = 0	H11b: Difference in mean change from baseline in NCS at 52 weeks (tezepelumab minus placebo) ≠ 0	A difference in means less than 0.
Key secondary endpoints			
Loss of smell at Week 52	H02a: Difference in mean change from baseline in loss of smell at 52 weeks (tezepelumab minus placebo) = 0	H12a: Difference in mean change from baseline in loss of smell at 52 weeks (tezepelumab minus placebo) ≠ 0	A difference in means less than 0.
SNOT-22 at Week 52	H02b: Difference in mean change from baseline in SNOT-22 at 52 weeks (tezepelumab minus placebo) = 0	H12b: Difference in mean change from baseline in SNOT-22 at 52 weeks (tezepelumab minus placebo) ≠ 0	A difference in means less than 0.
LMK at Week 52	H02c: Difference in mean change from baseline in LMK at 52 weeks (tezepelumab minus placebo) = 0	H12c: Difference in mean change from baseline in LMK score at 52 weeks (tezepelumab minus placebo) ≠ 0	A difference in means less than 0.
Time to SCS and/or surgery decision up to Week 52	H02d: HR of time to first SCS and/or surgery decision (tezepelumab/placebo) = 1	H12d: HR of time to first SCS and/or surgery decision (tezepelumab/placebo) ≠ 1	A HR less than 1.
Time to surgery decision up to Week 52	H03: HR of time to first surgery decision (tezepelumab/placebo) = 1	H13: HR of time to first surgery decision (tezepelumab/placebo) ≠ 1	A HR less than 1.
Time to SCS up to Week 52	H04: HR of time to first SCS (tezepelumab/placebo) = 1	H14: HR of time to first SCS (tezepelumab/placebo) ≠ 1	A HR less than 1.
NPSD TSS	H05: Difference in mean change from baseline in NPSD TSS at 52 weeks (tezepelumab minus placebo) = 0	H15: Difference in mean change from baseline in NPSD TSS at 52 weeks (tezepelumab minus placebo) ≠ 0	A difference in means less than 0.
FEV_1 at Week 52 *Subset of participants with co-morbid asthma/AERD/ NSAID-ERD*	H06: Difference in mean change from baseline in pre-BD FEV_1 at 52 weeks (tezepelumab minus placebo) = 0	H16: Difference in mean change from baseline in pre-BD FEV_1 at 52 weeks (tezepelumab minus placebo) ≠ 0	A difference in means greater than 0.

4.1.2 Testing Strategy for Confirmatory Endpoints

The overall type I error rate will be strongly controlled at the 5% level across primary and key secondary endpoints. The NPS and NCS are co-primary endpoints and, as such, both primary endpoints will be tested at the 5% level to determine the success of the study. If both are significant at 5%, then testing will proceed to the key secondary endpoints at 5%. In addition, the co-primary and key secondary endpoints will also be tested at the 1% level to further demonstrate persuasive statistical significance for this single Phase 3 study.

The following hierarchical testing strategy will be applied:

Level 1
The null hypotheses H01a and H01b (for endpoints NPS and NCS at Week 52) will be tested at a 2-sided 5% significance level

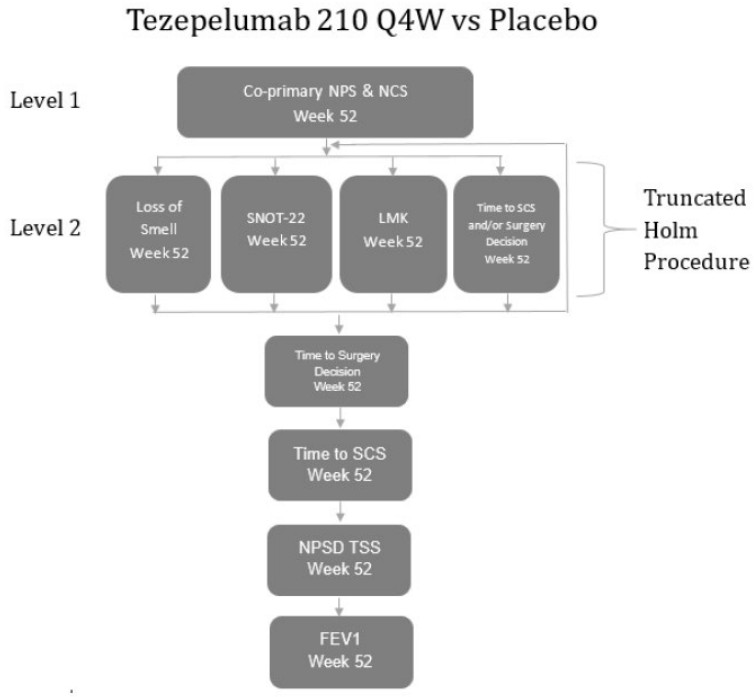
Level 2
If both H01a and H01b are rejected at the 2-sided 5% significance level, then the null hypotheses H02a, H02b, H02c, and H02d will be tested using the truncated Holm procedure (truncation parameter gamma =0.8) at an overall 2-sided 5% significance level for endpoints:

- Change from baseline in loss of smell at Week 52
- Change from baseline in SNOT-22 at Week 52
- Change from baseline in LMK at Week 52
- Time to SCS and/or surgery decision for NP up to Week 52

Under this approach, the 4 p-values are first ordered from smallest to the largest If the smallest p-values (2-sided) within level 2 is < 0.0125, the treatment effect for the endpoint associated with this p-value is considered significant. The test then evaluates the next smallest p-value at (0.05 + 0.8/3)*5% significance level (2-sided) and if significant, the third smallest at (0.05 + 0.8/2)*5% significance level (2-sided), and if significant, the fourth at (0.05 + 0.8)*5% significance level (2-sided). If none of the level 2 hypotheses are rejected, the procedure will stop.
If one, two, three, or four of H02a, H02b, H02c, and H02d is rejected the null hypotheses H03, H04, H05, and H06 will be tested in a sequential order at 0.25%, 0.5%, 0.75%, or 5% significance level (2-sided), respectively.

The testing procedure is summarised graphically in the following figure (Figure 2). For the US FDA, the described testing procedure will be applied to the US FDA primary estimand described in Section 4.1.3.2.

Figure 2 Testing Procedure



FEV1, forced expiratory volume in 1 second; LMK, Lund-Mackay score; NCS, nasal congestion score; NPSD, nasal polyp symptom diary; NPS, nasal polyp score; Q4W, every four weeks; SCS, systemic corticosteroids; SNOT-22, Sino-Nasal Outcome Test, 22 item; TSS, total nasal symptom score.

4.1.3 Estimands

4.1.3.1 Primary Estimand

- **Treatment:** Randomised treatment of tezepelumab 210 mg Q4W or placebo
- **Population of interest:** Adult participants with severe CRSwNP (total NPS ≥ 5) and an inadequate response to standard of care therapy, based on their randomised treatment and receiving at least one dose of IP
- **Endpoints of interest:** Change from baseline in co-primary endpoints: NPS and mean daily NCS at Week 52
- **Population level summary for the endpoint:** Difference in means between tezepelumab and placebo treatment groups
- **Handling of intercurrent events (ICEs):**
  - **Nasal polyp surgery:** The participants who had actual nasal polyp surgery would be considered poor outcome. The composite variable strategy will be applied where the worst possible score (ie. 8 for NPS and 3 for NCS) will be used for the post-surgery scores.
  - **SCS for NP:** The participants who received SCS (at least 3 consecutive days of OCS or equivalent 1 intramuscular injection for NP) would be considered poor outcome but less severe than surgery. The composite variable strategy will be applied where the

worst observation prior to the SCS will be carried forward (WOCF) for the post-SCS scores.

- **Treatment discontinuation:** The treatment policy strategy will be applied.

- **Adherence to background MFNS or an equivalent INCS and IP (including switching to other non-biologic/non-surgery/non-steroid treatments):** The treatment policy strategy will be applied. Data collected for participants who do not have rescue NP surgery or SCS are used regardless of the adherence of background MFNS or an equivalent INCS and IP and regardless of whether they switched to another non-biologic/non-surgery/non-steroid treatment.

- **Biologic use for NP:** The participants who received another biologic for NP would be considered poor outcome but less severe than surgery. The composite variable strategy will be applied where the worst observation prior to the biologic for NP will be carried forward (WOCF) for the post-biologic for NP scores.

- **Steroids and biologic use for comorbid conditions:** The treatment policy strategy will be applied for the participants who received another biologic for comorbid conditions.

- **COVID-19 related:** intercurrent events caused by COVID-19 include protocol deviations caused by COVID-19, COVID-19 vaccination, and COVID-19 infection. The treatment policy strategy will be applied.

4.1.3.2 US FDA Primary Estimand

Following the US FDA Chronic Rhinosinusitis with Nasal Polyps: Developing Drugs for Treatment Guidance for Industry (FDA 2023), different strategies for the ICEs of SCS rescue for NP and biologic use for NP will be used than for primary estimand, i.e.:

- **SCS for NP:** The treatment policy strategy will be applied and data after SCS rescue will be included in the analysis.

- **Biologic use for NP:** The composite variable strategy will be applied where the worst possible score (i.e. 8 for NPS and 3 for NCS) will be used for the post-biologic scores.

4.1.3.3 Supplementary Estimand

Composite variable strategy will be used for SCS for NP, biologic use for NP, and treatment discontinuation: the worst possible score will be used after the ICEs.

The key secondary and other secondary efficacy endpoints will use the same strategies for the intercurrent events unless noted otherwise.

4.2 Analysis Methods

4.2.1 Participant Disposition

Participant disposition will be summarised using the All participants analysis set. The number of enrolled participants will be summarised. The number and percentage of participants within each treatment group will be presented by the following categories; randomised, not randomised (and reason), received IP, did not receive IP (and reason), completed treatment, discontinued treatment (and reason), completed study (participants who completed IP and study, and participants who discontinued IP but completed study assessments), discontinued study (including reason), completed the short-term follow-up period, and completed the long-term follow-up period. Participant recruitment by country and centre will also be summarised.

The number and percentage of participants who discontinued IP, but remained in the study, will be presented by treatment group and option of follow up (Section 1.2).

Kaplan-Meier plots will be produced summarising the time (in days) to last dose of IP and premature withdrawal from the study separately. Participants without the premature event will be censored as described in Section 3.1.8.

The disposition due to the global/country situation will be summarized using the Randomised analysis set. The number and percentage of participants within each treatment group will be presented by the following categories: subjects who completed treatment, subjects who discontinued treatment due to the global/country situation, subjects who completed the study, and subjects who withdrew from the study due to the global/country situation. The global/country situation study disruptions will also be summarized for the Randomised analysis set. The number and percentage of participants within each treatment group will be presented by the following categories: participants with at least one disruption due to global/country situation (including type of disruption), and participants who discontinued treatment or withdrew from study due to global/country situation.

4.2.2 Demography and Other Baseline Characteristics

All demographic data and baseline characteristics will be summarised by treatment group and total for the FAS.

- Demographic data (eg, age, age group, gender, ethnicity, and race)
- Stratification factors from the IWRS and eCRF
- Subgroups defined in Section 3.1.6
- Other baseline characteristics defined in Section 3.1.7

- Important PDs

The number of participants in each of the analysis sets defined in Section 2.1 will be summarised.

Disease characteristics collected on the Respiratory Disease History, Nasal Polyposis Respiratory Disease History, and Nasal Polyposis Surgery eCRFs will also be summarised. The time since diagnosis, age at diagnosis, and time since last surgery will be calculated and summarised.

4.2.3 Medical and Surgical History

Medical and surgical histories will be summarised by MedDRA Preferred Term (PT) within the System Organ Class (SOC) level of MedDRA.

4.2.4 Prior and Concomitant Medications

The number and percentage of participants receiving each medication (by ATC code and generic name) will be presented by treatment group and total for the FAS. Separate tables will be presented for all medications received during each of the following periods as defined in Section 3.1.3: Prior, Concomitant (on-treatment), and Concomitant (post-treatment).

A separate table for prior biologics taken for asthma and/or NP, either marketed or in a clinical study, will be presented by treatment group and total for the FAS.

A table for standard of care medications at study entry will be produced displaying any standard of care medication for chronic rhinosinusitis with nasal polyps stable for at least 30 days prior to V1.

A separate table will be presented for participants who took disallowed concomitant medications. Disallowed medications will include medications defined as prohibited according to Section 6.4 of the CSP. They will be defined following a physician review (prior to primary database lock) of the unique combinations of ATC code and generic terms captured and are important protocol deviations.

Medications will be classified using WHO Drug Dictionary Mar 2024.

4.2.5 Exposure and Compliance

Exposure and treatment compliance derivation details are defined in Section 3.6.1.

Extent of exposure to IP, and total number of dosing occasions will be summarised by treatment group, for the Safety Set. Additionally, cumulative extent of exposure will be summarised for the Safety Set:

- any participant exposed, ≥4, ≥8, ≥12, ≥16, etc., and ≥52 weeks,
- >0 - < 4, ≥4 - <8, ≥8 - <12, etc. ≥48 - <52, and ≥52 weeks.

The date and time of IP administrations will be listed using the Safety Set.

Compliance with taking the regularly scheduled MFNS or equivalent INCS medication as recorded in the daily diary will be summarised bi-weekly by treatment group. The compliance with IP (%) and the number and percentage of subjects in each study treatment compliance (%) category (<80, 80-100, >100) will also be summarized. Compliance with the use of the daily diary will also be summarised bi-weekly by treatment group. Compliance with MFNS/INCS, IP and daily diary will be presented for Full Analysis Set.

4.2.6 Co-Primary Endpoints

4.2.6.1 Primary Estimand Analysis

For the primary estimand the primary analysis of the co-primary efficacy endpoints (mean change from baseline in NPS at 52 weeks and mean change from baseline in NCS at 52 weeks) uses a composite variable strategy in which data collected after nasal polyp surgery will be replaced with the worst possible score (WPS). Data collected after SCS for NP will be replaced with the worst post-baseline observation prior to the SCS for NP (WOCF). Data collected after treatment discontinuation without surgery or SCS for NP will be included in the analysis. For participants whose post-baseline data are all missing, the baseline value will be used in the place of WOCF. For participants who discontinue the study without NP surgery or SCS/biologic for NP, a multiple imputation (MI) approach will be used to impute missing Week 52 value assuming missing at random (MAR). A random seed of 539612 will be used for both monotone and non-monotone imputations.

Of note, biologics are prohibited medications per the CSP Section 6.4. However, if biologics are used for NP, the same strategy used for the SCS for NP will be applied.

The following steps will be taken to generate completed datasets up to Week 52 for analysis:

Step 1: Data collected after date of surgery, SCS for NP, or biologic for NP treatment will be set to missing for the purpose of imputation. For the NCS, if the intercurrent event occurred prior to the last day of a bi-weekly interval, the bi-weekly mean NCS for that period and all periods afterwards will be set to missing.

Step 2: Based on the dataset from Step 1, the intermittent missing data will be imputed using the MCMC method in SAS PROC MI with an IMPUTE=MONOTONE statement for each treatment group separately. The imputation model will include baseline co-morbid asthma/AERD/NSAID-ERD status, prior surgery status, region, baseline value, and data at each timepoint. 100 imputed datasets will be generated.

Step 3: The remaining monotone missing data will be imputed using the MI procedure with the 'MONOTONE REG' option for each treatment group separately. The imputation model will include baseline co-morbid asthma/AERD/NSAID-ERD status, prior surgery status, region, baseline value, and data at each timepoint.

Step 4: In the datasets from Step 3, replace post-surgery score with WPS, and replace post-SCS/biologic for NP score with WOCF. If a participant received both NP surgery and SCS for NP/biologic for NP, scores post-surgery will be set to WPS.

Each of the imputed datasets at Week 52 will be analysed using the analysis of covariance (ANCOVA) model with treatment group, baseline co-morbid asthma/AERD/NSAID-ERD status, prior surgery status, and region as factors (based on eCRF), and the baseline value as a covariate. Rubin's rule will be applied to combine analysis results (point estimates and standard errors) from the imputed datasets using PROC MIANALYZE.

Descriptive statistics including number of participants, mean, standard error, and least squares (LS) means will be provided at Week 52. In addition, difference in LS means (using the observed marginal distribution of the sample covariates) and the corresponding 95% confidence intervals (CIs) will be provided along with the p-value. A single asterisk (*) will be used to denote a result that is statistically significant at alpha = 0.05 and a double asterisk (**) will be used to denote a result that is statistically significant at alpha = 0.01. Nominal p-values will be presented unless specified otherwise.

4.2.6.2 US FDA Primary Estimand Analysis

For FDA estimand different strategies will be applied for ICEs of SCS rescue for NP and biologic use for NP (Section 4.1.3.2).

In Step 1 data collected after SCS rescue for NP will be utilised in the analysis. Data collected after biologic use for NP as well as after NP surgery will be set to missing for the purpose of imputation. Steps 2 and 3 will be repeated as described in Section 4.2.6.1. In Step 4, both post-surgery score and post-biologic for NP data will be replaced by WPS. The same ANCOVA model as described in Section 4.2.6.1 will be used for each imputed dataset. Rubin's rule will be applied to combine analysis results (point estimates and standard errors) from the imputed datasets using PROC MIANALYZE. The results will be presented as shown in Section 4.2.6.1.

4.2.6.3 Supplementary Estimand Analyses

4.2.6.3.1 Composite strategy for NP surgery, SCS/biologic for NP and IP discontinuation using WPS

Supplementary analysis will be performed using different strategies (comparing to primary estimand) for SCS rescue for NP, biologic use for NP and treatment discontinuation.

In Step 1 data collected after NP surgery, SCS rescue for NP, biologic use for NP and treatment discontinuation will be set to missing. Steps 2 and 3 will be repeated. In Step 4, data after all 4 intercurrent events mentioned in Step 1 will be replaced by WPS. The same ANCOVA model as described in Section 4.2.6.1 will be used for each imputed dataset. Rubin's rule will be applied to combine analysis results (point estimates and standard errors) from the imputed datasets using PROC MIANALYZE. The results will be presented as shown in Section 4.2.6.1.

4.2.6.3.2 Hypothetical strategy for COVID-19 infections

If more than 5% of participants in either treatment group have reported COVID-19 infections (based on the COVID-19 MedDRA terms), an additional supplementary analysis will be performed for the co-primary endpoints using a hypothetical strategy. Data after the first infection will be set to missing if NP surgery/SCS for NP/biologic for NP didn't occur prior to the COVID-19 infection. If COVID-19 infection started prior to the randomisation, data prior to the end date of the infection will also be set to missing. The analyses described in Section 4.2.6.1 will be repeated.

4.2.6.4 Sensitivity Analyses

4.2.6.4.1 Rank ANCOVA

To investigate the robustness of primary efficacy analyses for the co-primary endpoints, a rank ANCOVA model will be performed as described by Stokes et al., 2000. More specifically, the ranks of post-baseline score and baseline score in the combined treatment group will be ranked using PROC RANK. The residuals from the linear regression of the response variable ranks vs ranks of the baseline score using PROC REG. The residuals in each treatment group will be compared using the CMH means score from PROC FREQ. The standardized transformed CMH statistic for the *m*th imputed dataset ($st\_wh\_cmh^{(m)}$) will be calculated as follows and then combined using PROC MIANALYZE.

$$st\_wh\_cmh^{(m)} = \left(\sqrt[3]{\frac{cmh^{(m)}}{df}} - \left(1 - \frac{2}{9 \times df}\right)\right) / \sqrt[2]{2/(9 \times df)}$$

where $cmh^{(m)}$ is CMH statistic for the *m*th imputed dataset, df is the number of associated degrees of freedom (O'Kelly and Ratitch, 2014). The difference in median and 95% CI will be estimated using PROC QUANTREG (Mehrotra, 2017). The model will include the same covariates as in the ANCOVA model in Section 4.2.6.1 The median and standard error for each treatment group will also be estimated using PROC QUANTREG. The results from each MI dataset will be combined using PROC MIANALYZE.

4.2.6.4.2 Control-based imputation

Control-based imputation (copy reference) will be performed to examine the robustness of the MAR assumption.

The analysis follows similar 4 steps as described in Section 4.2.6.1. Imputed dataset from step 2 in Section 4.2.6.1 will be used in step 3, where missing data from participants in the tezepelumab group who dropped out for treatment related reasons will be assumed missing not at random (MNAR) /dropout reason-based (DRMI) and will be imputed using the placebo group, whereas the remaining participants who dropped out will be imputed assuming MAR. Table 10 summarises how tezepelumab participants withdrawing from study will be handled in the DRMI analyses. A random seed of 539612 will be used. Step 4 will be repeated as in Section 4.2.6.1.

The same ANCOVA model as described in Section 4.2.6.1 will be used for each imputed datasets. Rubin's rule will be applied to combine analysis results (point estimates and standard errors) from the imputed datasets using PROC MIANALYZE. The results will be presented as shown in Section 4.2.6.1.

Table 10 Treatment arms for imputation of tezepelumab participants under DRMI

Reason for withdrawing from study	Reason for discontinuing IP	DRMI
Site terminated by sponsor		Tezepelumab
Study terminated by sponsor		Tezepelumab
Death		Placebo
Loss to follow-up	Subject lost to follow-up	Placebo by default (pending blinded review of any further information)
Withdrawal by subject	Subject decision	Placebo by default (pending blinded review of any further information)
Other	Severe non-compliance to protocol	Placebo
	Adverse event	Placebo
	Development of study-specific discontinuation criteria	Placebo
	Pregnancy	Placebo
	Other	Placebo by default (pending blinded review of any further information)

A blinded review of subjects who discontinued IP for reasons of "Subject lost to follow-up", "Subject decision" or "Other" will be performed prior to unblinding at the primary database lock. A listing of these subjects and the assumptions made under DRMI will be documented. If any recorded comments (on either of the "Discontinuation of Investigational Product" or "Disposition" eCRF pages) indicate clearly that the reason for study withdrawal was not related to treatment, then the "Placebo" default for DRMI in the above table may be changed for that subject.

4.2.6.4.3 Tipping point analysis

If more than 5% of participants in either treatment group have missing data at Week 52, a tipping point analysis will also be performed. After Step 4 in Section 4.2.6.1, the following steps will be performed:

Step 5. The imputed values (for the missing values assuming MAR) in the tezepelumab group at Week 52 are added by a small positive amount δ_t for each imputed dataset. The imputed values in the placebo group are subtracted by a small positive amount δ_p for each imputed dataset.

Step 6. The change from baseline in the endpoint will be analysed using the same ANCOVA model used in the primary analysis. Then the SAS MIANALYZE procedure will be used to generate statistical inferences by combining results from the 100 analyses using Rubin's formula.

Step 5 to Step 6 will be repeated iteratively until the p-value for treatment effect of tezepelumab compared to placebo estimated in Step 6 is >0.01 and then >0.05. The LS mean difference between tezepelumab and placebo in change from baseline in co-primary endpoints at Week 52 and the corresponding p values will be provided for each combination of shift parameters. The value of δ_t and δ_p will be adjusted to avoid reaching the tipping point at the first iteration.

4.2.6.5 Supportive analyses

To facilitate the understanding of the co-primary endpoints, the empirical cumulative distribution function curves will be created for NPS and NCS at Week 52 by treatment group. The x-axis is the change from baseline and the y-axis is the percentage of participants. As described in Section 4.2.6.1, the post-surgery score is replaced by WPS and the post-SCS/biologic for NP score is replaced by WOCF. Participants who have missing change from baseline at WK52 without a preceeding surgery or SCS for NP will be excluded from this analysis.

4.2.6.6 Subgroup Analyses

To explore the treatment effect across categorical baseline or demographic variables as detailed in Section 3.1.6, a similar ANCOVA model, as described in Sections 4.2.6.1 and

4.2.6.2, will be fitted, using imputed datasets, with additional factors for the relevant subgroup variable (based on eCRF, if data collected on an eCRF form) and its interaction with treatment. The estimate of the treatment effect and its 95% CI within each of the subgroup categories will be tabulated and also summarised graphically using a forest plot. A p-value for the treatment by subgroup interaction will not be presented for each of these models due to the exploratory nature of the analyses. If the subgroup variable is the same as a covariate, only the interaction term will be added. If the subgroup variable is a categorisation of the continuous baseline value (or re-classification of a categorical covariate), the categorical subgroup variable and interaction term will be added in the model, and the continuous baseline value (or categorical covariate) will be dropped from the model.

If any of the subgroups have fewer than 10 participants in one or both treatment groups, this subgroup level is not included in the model. If that leaves only one subgroup level, the model is not fitted for that categorical variable.

4.2.7 Key Secondary Endpoints

4.2.7.1 Primary Analysis

For all the key secondary endpoints, a single asterisk (*) will be used to denote a result that is statistically significant at alpha = 0.05 and a double asterisk (**) will be used to denote a result that is statistically significant at alpha = 0.01. Nominal p-values will be presented unless specified otherwise. Analyses based on the primary estimand and US FDA primary estimand will be provided for key secondary endpoints when applicable, unless noted otherwise.

4.2.7.1.1 Change from baseline in loss of smell at Week 52

The change from baseline in loss of smell at Week 52 will be analysed using the methods as described in Sections 4.2.6.1 (primary estimand) and 4.2.6.2 (US FDA primary estimand).

4.2.7.1.2 Change from baseline in SNOT-22 at Week 52

The same methods as described in Sections 4.2.6.1 (primary estimand) and 4.2.6.2 (US FDA primary estimand) will be applied.

4.2.7.1.3 Change from baseline in LMK at Week 52

The same methods as described in Sections 4.2.6.1 (primary estimand) and 4.2.6.2 (US FDA primary estimand) will be applied.

4.2.7.1.4 Time to first nasal polyp surgery decision and/or SCS for NP up to Week 52

Time to first SCS for NP or decision to have surgery for NP during the 52-week treatment period will be analysed using a Cox proportional hazards model, adjusting for treatment, baseline co-morbid asthma/AERD/NSAID-ERD status, prior surgery status, and region (based on eCRF). The event of interest is the first SCS for NP or nasal polyp surgery decision, whichever is earlier if both occur. Hazard ratios, corresponding 95% CIs, and p-values will be

presented. The proportion of participants with the decision of NP surgery or who took SCS for NP up to week 52 will be estimated using the Kaplan-Meier method. The point estimate and 95% CI will be presented. Kaplan-Meier curves will be generated.

4.2.7.1.5 Time to first surgery decision for NP up to Week 52

The same method described in Section 4.2.7.1.4 will be used to analyse the time to first surgery decision for NP.

4.2.7.1.6 Time to first SCS for NP up to Week 52

The same method described in Section 4.2.7.1.4 will be used to analyse the time to first SCS for NP.

4.2.7.1.7 Change from baseline in NPSD TSS at Week 52

The same methods as described in Sections 4.2.6.1 (primary estimand) and 4.2.6.2 (US FDA primary estimand) will be applied.

4.2.7.1.8 Change from baseline in pre-BD FEV_1 at Week 52 (comorbid asthma/AERD/NSAID-ERD subset)

Observed pre-BD FEV_1 data will be used in the analysis regardless of NP surgery and SCS/biologic for NP. Missing data up to Week 52 will be imputed assuming MAR. The intermittent missing data will be imputed using the MCMC method in SAS PROC MI with an IMPUTE=MONOTONE statement for each treatment group separately. The imputation model will include prior NP surgery status, region, baseline pre-BD FEV1 value, and data at each timepoint (both prior NP surgery status and region will be based on eCRF). 100 imputed datasets will be generated. The remaining monotone missing data will be imputed using the MI procedure with the 'MONOTONE REG' option for each treatment group separately. The imputation model will include baseline value, prior surgery status, region, and data at each timepoint. A random seed of 539612 will be used for both monotone and non-monotone imputations.

The same ANCOVA model described in Section 4.2.6.1 will be used at Week 52. Descriptive statistics including number of participants, mean, standard error, and LS means will be provided. In addition, the difference in LS means and the corresponding 95% CI will be provided along with the p-value.

4.2.7.2 Supportive Analyses

4.2.7.2.1 Responder analysis for SNOT-22 at Week 52

An MCID of 8.90 has been established for SNOT-22 total score (Hopkins et al., 2009).

A participant who has a change from baseline ≤ -8.9 at Week 52 in the absence of SCS for NP, biologic for NP, or NP surgery at or prior to that time point will be defined as a

responder, otherwise the participant will be defined as a non-responder. The number and proportion of responders at Week 52 will be summarised descriptively.

For the treatment comparison, the imputed datasets as described in Section 4.2.6.1 Step 1 to Step 4 will be used. A subject will be assigned to responder or non-responder after dichotomizing the imputed values in each imputed dataset. The proportion of responders in each imputed dataset will be analysed using the logistic regression model with treatment group, baseline co-morbid asthma/AERD/NSAID-ERD status, prior NP surgery status and region (based on eCRF) as factor and baseline value as a covariate. The response variable in the model will be the binary responder status at Week 52. The analysis results from the imputed datasets will be combined using PROC MIANALYZE. Odds ratios, corresponding 95% CIs, and nominal p-value will be presented.

In addition, the same model as described above will be used in g-computation method to estimate the unconditional treatment effect. The variance will be estimated using the method proposed in Ye et al., 2023 and Bannick et al., 2024 for covariate adjusted randomisation. The analysis results from the imputed datasets will be combined using PROC MIANALYZE. Unconditional difference in response rates and corresponding 95% CIs will be presented.

The same methods will be applied to each timepoint through Week 52.

4.2.7.2.2 Change from baseline over time through Week 52

Change from baseline in SNOT-22, loss of smell, NPSD TSS, and pre-BD FEV_1 for each timepoint through Week 52 will be analysed using the same methods and imputed datasets as described in Sections 4.2.6.1 (primary estimand) and 4.2.6.2 (US FDA estimand). Descriptive statistics including number of participants, mean, standard error, and LS means will be provided for each timepoint. In addition, the difference in LS means and the corresponding 95% CI will be provided along with the p-values for each timepoint.

4.2.7.2.3 Time to first nasal polyp surgery and/or SCS for NP up to Week 52

Time to first nasal polyp surgery and/or SCS for NP up to Week 52 as well as time to first nasal polyp surgery will be analysed as described in Sections 4.2.7.1.4 and 4.2.7.1.5 respectively using the actual date of surgery for NP as per Section 3.3.3.

4.2.7.2.4 Time to first SCS for any reason up to Week 52

Time to first SCS for any reason will be analysed using the Cox proportional hazards model described in Section 4.2.7.1.6. The cumulative proportion of participants receiving SCS for any reason will be estimated by Kaplan-Meier methods.

4.2.8 Other Secondary Endpoints

For other secondary endpoints, nominal p-values will be presented when appropriate.

4.2.8.1 Change from baseline over time in NPS and NCS through Week 52

The same methods described in the primary analyses for the co-primary endpoints in Sections 4.2.6.1 (primary estimand) and 4.2.6.2 (US FDA primary estimand) will be used each timepoint through Week 52. Descriptive statistics including number of participants, mean, standard error, and LS means will be provided for each timepoint. In addition, the difference in LS means and the corresponding 95% CI will be provided along with the nominal p-values for each timepoint.

4.2.8.2 Change from baseline in loss of smell evaluated by UPSIT test at Week 52

The UPSIT scores will be analysed overall at Week 52. The same methods as described in Section 4.2.6.1 (primary estimand) will be applied.

Additionally, the number (%) of participants with anosmia will be summarised by treatment group at baseline and Week 52 separately for men and women according to Table 7.

The same methods will be applied to each timepoint through Week 52.

4.2.8.3 Change from baseline in modified LMK score and sinus severity score at Week 52

The same methods as described in Section 4.2.6.1 (primary estimand) will be applied for modified LMK score and sinus severity score at Week 52. The same method for LMK score will be applied to each timepoint through Week 52.

4.2.8.4 Change from baseline in NPSD domain scores through Week 52

The same methods as described in Section 4.2.6.1 (primary estimand) will be applied to each item in the NPSD.

4.2.8.5 Change from baseline in NPIF through Week 52

The change from baseline in NPIF trough Week 52 will be analysed using similar method as described in Section 4.2.6.1 (primary estimand). Step 1-3 will be repeated with no changes. In Step 4 post-surgery assessments will be replaced with WOCF (WPS not known). Each of the imputed datasets will be analysed using the analysis of covariance (ANCOVA) model shown in Section 4.2.6.1 through 52 weeks.

4.2.8.6 Change from baseline in ACQ-6 at Week 52 (co-morbid asthma/AERD/NSAID-ERD subset)

The same methods described in Section 4.2.7.1.8 will be applied to ACQ-6 at Week 52. The same methods described in Section 4.2.7.2.2 trough Week 52 will be applied.

The number and proportion of ACQ-6 response status and asthma control status as defined in Section 3.4.8 will be summarized descriptively. The proportion of responders will be analyses as described in 4.2.7.2.1.

The same methods will be applied to each timepoint through Week 52.

4.2.8.7 Proportion of participants with ≥1 point reduction or ≥2 points reduction in NPS at Week 52

A participant with ≥1 point reduction or ≥2 points reduction in NPS at Week 52 in the absence of SCS for NP, biologic for NP, or NP surgery at or prior to that time point will be defined as a responder, otherwise the participant will be defined as a non-responder. The number and proportion of responders at Week 52 will be summarised descriptively. The proportion of responders will be analysed as described in 4.2.7.2.1.

The same methods will be applied to each timepoint through Week 52.

4.2.8.8 Exposure of SCS

Exposure of SCS over 52 Weeks will be summarised descriptively by treatment group. The number of days of SCS as well as the number of courses taken per participant will also be presented. Total SCS dose for NP (converted to prednisolone equivalents) will be summarised by treatment group. The same summary will also be provided for SCS for any reasons and for SCS for any reason excluding NP.

The number of courses of SCS for NP per year will be analysed using a negative binomial model. The response variable will be the number of courses of SCS for NP received by a participant over the planned treatment period. The model incudes treatment group, baseline co-morbid asthma/AERD/NSAID-ERD status, prior NP surgery status, and region (based on eCRF) as factors. The logarithm of the time at risk (in years) will be used as an offset variable, to adjust different follow-up times. Time during a course will not be included in the calculation of time at risk.

The time at risk over 52 weeks (year) for a participant will be derived as follows:

The time at risk over 52 weeks (years) = [planned treatment period (days) as presented in Section 3.1.3 – sum of duration of all SCS courses over planned treatment period (days)] / 365.25.

Duration of SCS course (days) = end date of SCS course – start date of SCS course + 1.

4.2.9 Exploratory Endpoints

4.2.9.1 Change from baseline in NPS and NCS over time up to Week 76

To explore the off-treatment effect, change from baseline in NPS and NCS over time up to Week 76 will be analysed using the Additional FU Analysis Set. All the following steps will be performed on Additional FU Analysis Set.

Step I: For the data after Week 52 and up to Week 76, set data collected after date of surgery, SCS for NP, or biologic for NP treatment to missing.

Step II: Merge the dataset from Step I with each MI dataset from Step 3 in Section 4.2.6.1.

Step III: Based on the dataset from Step II, the intermittent missing data after Week 52 and up to Week 76 will be imputed using the MCMC method in SAS PROC MI with an IMPUTE=MONOTONE statement for each treatment group separately. For the NPS, the imputation model will include baseline co-morbid asthma/AERD/NSAID-ERD status, prior surgery status, region, baseline value, all post-baseline timepoints up to Week 76. For the NCS, the imputation model will include baseline co-morbid asthma/AERD/NSAID-ERD status, prior surgery status, region, baseline value, data at Weeks 12, 24, 36, 48, 50, 52, and all timepoints after Week 52 and up to Week 76. Timepoints prior to Week 52 may be further reduced from the model in case of convergence problems. NIMPUTE=1 and MAXITER=5000.

Step IV: The remaining monotone missing data will be imputed using the MI procedure with the 'MONOTONE REG' option for each treatment group separately. The same imputation model as in Step III will be used. NIMPUTE=1 and MAXITER=5000.

Step V: In the datasets from Step IV, replace post-surgery score with WPS, and replace post-SCS/biologic for NP score with WOCF. If a participant received both NP surgery and SCS for NP/biologic for NP, scores post-surgery will be set to WPS.

The same ANCOVA model described in the primary analyses for the co-primary endpoints in Section 4.2.6.1 (primary estimand) will be used for the timepoints up to Week 76. Descriptive statistics including number of participants, mean, standard error, and LS means will be provided for each timepoint. In addition, the difference in LS means and the corresponding 95% CI will be provided for each timepoint. The LS means and 95% CI over time will be displayed graphically by treatment group.

4.2.9.2 PGI-S and PGI-C

PGI-C response categories are described in Section 3.5.3. The number and percentage of participants in each category in PGI-S and PGI-C, as well as PGI-C improvement responses, will be summarised descriptively by treatment group and visit.

4.2.9.3 EQ-5D-5L

The total score and change from baseline in EQ-5D-5L overall utility score of all dimensions as well as individual dimension scores will be summarised descriptively by treatment group and visit. VAS and Health State Index will also be summarised descriptively by treatment group and visit.

4.2.9.4 Change from baseline in SNOT-22, loss of smell, UPSIT, NPSD TSS, and NPIF overtime up to Week 52

The change from baseline in SNOT-22, loss of smell, and NPSD TSS will be analysed as described in Section 4.2.9.1.

The change from baseline in UPSIT will be analysed as described in Section 4.2.8.2. Descriptive statistics including number of participants, mean, standard error, and LS means will be provided for each timepoint. In addition, the difference in LS means and the corresponding 95% CI will be provided for each timepoint. The LS means and 95% CI over time will be displayed graphically.

The change from baseline in NPIF will be analysed as described in Section 4.2.8.5. Descriptive statistics including number of participants, mean, standard error, and LS means will be provided for each timepoint. In addition, the difference in LS means and the corresponding 95% CI will be provided for each timepoint.

4.2.9.5 Change from baseline in pre-BD FEV_1 and ACQ-6 overtime up to Week 52 (co-morbid asthma/AERD/NSAID-ERD subset)

The same methods described in Section 4.2.7.2.2 for pre-BD FEV_1 and Section 4.2.8.6 for ACQ-6 score will be applied trough Week 76. Descriptive statistics including number of participants, mean, standard error, and LS means will be provided for each timepoint. In addition, the difference in LS means and the corresponding 95% CI will be provided for each timepoint.

4.2.9.6 Proportion of participants with ≥1 point reduction and ≥2 points reduction in total NPS up to Week 52

The same methods described in Section 4.2.8.7 will be applied through Week 52.

4.2.9.7 Proportion of NCS responders up to Week 52

A participant with >=1 point reduction in NCS at Week 52 in the absence of SCS for NP, biologic for NP, or NP surgery at or prior to that timepoint will be defined as a responder, otherwise the participant will be defined as a non-responder. The number and proportion of NCS responders will be analysed as described in 4.2.7.2.1.

The same methods will be applied to each timepoint through Week 52.

4.2.10 Safety and Tolerability

4.2.10.1 Adverse Events

All AE summaries will be presented by treatment group. AEs occurring during the screening/run-in period will be listed, but not summarised separately.

An overall summary table will be produced for the on-treatment and on-study periods showing the number and percentage of participants with at least one AE in each of the following categories: any AEs, SAEs, any AE with outcome death, and AEs leading to discontinuation of IP (DAEs). The total number of AEs in the different AE categories will also be presented. Key information of SAEs, AEs with outcome death, AEs leading to discontinuation of IP, and AESIs during the on-study period will be tabulated.

All AEs during the on-treatment and on-study periods will be summarised separately by SOC and PT assigned to the event using the MedDRA dictionary. An AE will be assigned into a period based on the onset date. For each PT, the number and percentage of participants reporting at least one occurrence of the event will be presented (ie, participants with multiple occurrences of the same PT will only be counted once).

Similar summaries by SOC and PT will be presented for the following during the on-treatment and on-study period:

- SAEs
- AE with outcome death
- AEs leading to discontinuation of IP (DAEs)
- Each AESI category separately
- The most common AEs (defined as those occurring in >3% of participants in either treatment group) – by PT only
- AEs by investigator's causality
- Each AESI category separately by investigator's causality

AEs are summarised by PT and maximum intensity during the on-treatment and on-study period. If a participant reports multiple occurrences of an AE within each PT, the maximum intensity will be taken as the highest recorded (the order being mild, moderate, and severe), respectively.

Exposure-adjusted AE (EAIR) summaries, and the difference in EAIR between treatment groups with 95% CIs calculated during the on-treatment and on-study period will be presented for each of the following summaries:

- overall AE categories: any AEs, SAEs, any AE with outcome death, and DAEs,
- All AEs by SOC and PT,
- SAEs by SOC and PT,

- DAEs by SOC and PT,
- Each AESI category by SOC and PT separately.

In the above summaries, the exposure-adjusted incidence rate (EAIR) will be calculated for each treatment as the number of participants in that treatment group reporting the AE divided by the total time at risk in that treatment group, where time at risk is the time to the first event for a participant who experienced the event during the analysis period and time during the analysis period for a participant who didn't experience the event. EAIR will be reported as events per 100 participant-years. The 95CI% for EAIR difference will be calculated based on Miettinen and Nurminen's score method.

The overall AE categories summaries for on-treatment and on-study period will be repeated for the following subgroups:

- Age group (<65, ≥65 years)
- Gender (Male, Female)
- Race (White, non-White)
- BMI (<25, ≥25 to <30, ≥30 kg/m^2)
- Region (Asia Pacific including China and Japan, Europe, North America)

The EAIR (95% CIs) by subgroups for all AEs and SAEs will be presented on the forest plots.

Overall AEs as well as AEs and SAEs by SOC and PT summaries will be produced for the AEs with onset date during the additional follow-up period for the Additional FU analysis set.

The number and percentage of participant reporting COVID-19 AEs (as defined based on the COVID-19 MedDRA terms) is summarised by SOC and PT for the on-treatment and on-study periods. In addition, if there are more than 10 participants reporting COVID-19 AEs, then the AE listing is repeated including only these participants, with details of all AEs reported by these participants.

The adjudicated adverse events will also be summarised at subject level and event level for placebo-controlled on-treatment and placebo-controlled on-study period. In addition, the adjudicated MACE events, serious cardiac events, death, and malignancy will be presented by SOC, PT and causality (as determined by the IAC). The adverse events sent for adjudication will be listed together with the outcome of adjudication.

4.2.10.2 Laboratory Data

All continuous laboratory variables will be summarised by absolute value at each visit by treatment group, together with the corresponding changes from baseline. These summaries will be produced for the on-study period, as defined in Section 3.1.3. The summary statistics

presented will be the minimum, 1st quartile, median, 3rd quartile, maximum, mean and SD. Shift plots showing each individual participant's laboratory value at baseline and at maximum/minimum/last value post-baseline are produced for each continuous laboratory variable. If any laboratory variables show any unusual features (high or low values or a general shift in the data points) at other time points, then shift plots of these data may be produced. The diagonal line of no change is also displayed on the shift plots.

Central laboratory normal reference ranges will be used for the identification of individual clinically important abnormalities. A shift table will be produced for each laboratory variable to display low, normal, high, and missing values. The shift tables will present baseline and maximum/minimum/last post-baseline values for each variable.

The frequencies of clinically noteworthy values (using normal reference ranges) occurring during the study will also be given.

In order to identify potential Hy's Law cases, maximum post-baseline TBL is plotted separately against both maximum post-baseline ALT and AST, expressed as multiples of ULN. These plots are produced on a log scale, with reference lines included at 2xULN for TBL, and at 3xULN for both ALT and AST.

For all participants who meet the biochemical criteria for Hy's Law (potential Hy's Law cases), the relevant laboratory variables will be tabulated showing all visits for these participants. Participants with elevated ALT or AST in addition to elevated TBL at any time may be explored further graphically using individual participant profile plots, if indicated.

For urinalysis data, a shift table (separate for China and non-China participants) will be generated to present changes from baseline to maximum. All data for the on-study period will be used.

4.2.10.3 Vital Signs

All vital signs variables will be summarised by absolute value at each visit by treatment group, together with the corresponding changes from baseline. This will also include weight, BMI, and height. These summaries will be produced for the on-study period, as defined in Section 3.1.3. The summary statistics presented will be the minimum, 1st quartile, median, 3rd quartile, maximum, mean and SD.

AZ-defined reference ranges (see Section 3.6.5) will be used for the identification of individual abnormalities. A shift table will be produced for each vital signs variable to display low, normal, high and missing values. The shift tables will present baseline and maximum/minimum/last observation values for each variable.

Shift plots showing each individual participant's vital signs value at baseline and at maximum/minimum/last value post-baseline will be produced for each continuous vital signs variable.

Participants who have changes from baseline outside the pre-defined AZ clinically important change criteria defined in Section 3.6.5 will be summarised.

4.2.10.4 Electrocardiogram

ECG variables will be summarised by absolute value at each visit by treatment group, together with the corresponding changes from baseline. These summaries will be produced for the on-study period, as defined in Section 3.1.3. The summary statistics presented will be the minimum, 1st quartile, median, 3rd quartile, maximum, mean and SD.

A shift table will be produced to display the investigator assessment of normal, abnormal – not clinically significant, abnormal – clinically significant and not done between baseline and end of study. For this purpose, borderline (also recorded on the eCRF) will be grouped with normal.

A frequency table showing participants with Frederician corrected QT (QTc) values and increases from baseline at any time during the on-study period using standard pre-specified thresholds will be produced.

4.2.10.5 Physical Examination

Clinical signs related to previous serious illnesses as new or worsening abnormalities will be recorded and summarised as AEs. No separate summaries of physical examination findings will be produced.

4.2.11 Pharmacokinetics and Immunogenicity

4.2.11.1 Analysis of Pharmacokinetics

All analyses of PK variables will be based on the PK analysis set as defined in Section 2.1.3.

Serum tezepelumab concentrations will be summarised over time for the on-study period using descriptive statistics (for the tezepelumab group only).

Serum samples for PK are scheduled to be collected at Weeks 0, 4, 12 24, 36, 52, 64 and at the premature IP discontinuation visit, where appropriate. Data will be assigned to weeks based on the windows defined in Section 3.1.4.

The following criteria will also apply for data to be included in the summary table:

- Only pre-dose samples at Week 0.

- Only pre-dose samples at Weeks 4, 12, 24, and 36 and samples at Week 52 that were also between ≥21 and ≤35 days post the previous dose.
- All samples for Week 64 that were taken within the visit window defined in Section 3.1.4.

For descriptive statistics of tezepelumab concentrations:

- If, at a given time point, 50% or less of the concentrations are non-quantifiable (NQ), the geometric mean, coefficient of variation (CV), arithmetic mean and SD will be calculated by substituting the lower limit of quantification (LLOQ) divided by 2 for values which are NQ.
- If more than 50%, but not all, of the concentrations are NQ, the geometric mean, CV, arithmetic mean and SD will be reported as not calculable
- If all the concentrations are NQ, the geometric mean and arithmetic mean will be reported as NQ and the CV and SD as not calculable
- The median, minimum and maximum will also be reported

The LLOQ of tezepelumab in serum will be 0.010 µg/mL.

4.2.11.2 Analysis of Immunogenicity

All analyses of immunogenicity variables will be based on the safety set as defined in Section 2.1.2. Results for China will be separated from results for all other countries (non-China).

The number of ADA positive participants at each visit will be summarised by treatment group for the on-study period. Descriptive statistics including number of participants, median, lower and upper quartile and range of the actual ADA titres by treatment group and visit, where possible, will be provided.

The ADA status across the study for each participant will also be classified and summarised by treatment group. Specifically, the following ADA results will be evaluated as number and proportion of participants in cohorts together with corresponding titre summaries:

- Participants who are ADA positive at any time including baseline (ADA prevalence)
- Participants who are ADA positive at baseline only
- Participants who are ADA positive at baseline and positive in at least one post baseline measurement
- Participants who are ADA positive at baseline regardless of post-baseline result
- Participants who are ADA positive post-baseline
- Participants who are ADA positive post-baseline and ADA negative at baseline (treatment induced ADA)

- Participants who are persistently positive; persistently positive is defined as having at least 2 post-baseline ADA positive measurements (with ≥ 16 weeks between first and last positive) or an ADA positive result at the last available post-baseline assessment
- Participants who are transiently positive; transiently positive is defined as having at least one post-baseline ADA positive measurement and not fulfilling the conditions for persistently positive
- Participants with treatment boosted ADA, defined as baseline positive ADA titre that was boosted to a 4-fold or higher level following IP administration
- Participants with treatment emergent ADA (ADA incidence): defined as either treatment induced ADA or treatment boosted ADA
- Participants who are nAb positive at baseline and/or post-baseline (nAb prevalence).
- Participants who are treatment-induced nAb positive (nAb incidence)

For ADA summaries at a single time point (eg, baseline ADA or by visit) the corresponding titre summary will be based on the titre of the positive sample for that particular visit. For summaries across visits (eg, ADA positive in an ADA category) the corresponding titre summaries will be based on the maximum titre of all positive samples for each participant.

Neutralizing ADA evaluations will be conducted on confirmed ADA positive samples. The test sample is deemed positive or negative for the presence of nAb to tezepelumab relative to a pre-determined (in assay validation) statistically derived cut point. The number and proportion of participants who are nAb positive at any time will be evaluated.

The ADA status across the study will be listed together with primary endpoints, biomarkers and serum concentrations over time, as well as with AEs/SAEs.

4.2.11.3 Exposure-Response Analysis

Tezepelumab participants will be grouped according to drug concentration quartiles. The treatment groups referred to in the following analyses will be the drug concentration quartiles and placebo. Each drug concentration quartile will be compared against the overall placebo group. The same methods described in the primary analyses for the co-primary endpoints in Section 4.2.6.1 will be used.

The four, mutually exclusive, drug concentration quartiles will be derived using the median trough steady state drug concentration data. The median trough steady state is calculated as the median of the observed drug concentration data at weeks 24, 36 and 52 for each individual participant.

4.2.12 Biomarkers

Biomarker values and change from baseline in these values will be summarised descriptively and presented graphically by treatment group up to Week 76:

- FeNO (co-morbid asthma/AERD/NSAID-ERD subset)
- Peripheral blood eosinophils (Cells/μL)
- Total serum IgE (IU/mL)

Analyses for the other exploratory biomarkers will be described in a separate exploratory analysis plan, which will be finalized before the database lock. These results will be reported outside the main CSR, in an addendum, or separately in a scientific report or publication.

4.2.13 Additional Analyses Due to Global/Country Situation Study Disruption

Efforts are ongoing to collect outstanding data via alternative means where possible when onsite visits cannot be performed due to global/country situation study disruption. All related protocol deviations (important PDs) will be listed and summarised separately for the Randomised analysis set.

4.2.14 Analyses for China and Japan Registration

To support the registration in China and Japan, selected analyses will be performed for the participants enrolled in the region. The analyses may include but are not limited to disposition, demographic and baseline characteristics, extent of exposure, prior and concomitant medications, efficacy and safety analyses. Details of the analyses may be documented separately. No p-values will be presented. The Cox regression will not be performed for the time to event type endpoints, if there are <10 events across the treatment groups and <2 events in each treatment group.

5 INTERIM ANALYSES

No interim analyses are planned in this trial.

There will be two database locks (DBLs) in this study. The primary DBL will be conducted after the last participant completes Week 52, and the final DBL will be conducted once the last participant has completed the last safety follow-up visit (Week 76 for the first 200 participants completing the treatment period and Week 64 for the remaining participants). All analyses of the primary and key secondary endpoints will be performed based on the primary DBL data.

All personnel involved with the analysis and conduct of the study will remain blinded until primary database lock and important protocol deviations during the double-blind treatment

period identified. After primary DBL, treatment allocation for participants during this study will become known to the Sponsor staff and/or designated CRO. The blind will be maintained for the Investigator, investigational site staff, and for the participant.

6 CHANGES OF ANALYSIS FROM PROTOCOL

Added the following exploratory analyses:

- Change from baseline in SNOT-22, loss of smell, UPSIT, NPSD TSS, and NPIF overtime up to Week 52
- Change from baseline in FEV_1 and ACQ-6 overtime up to Week 52 (co-morbid asthma/AERD/NSAID-ERD)

7 REFERENCES

Bachert et al., 2019
Bachert C, Han JK, Desrosiers M, Hellings PW, Amin N, Lee SE, et al. Efficacy and safety of dupilumab in participants with severe chronic rhinosinusitis with nasal polyps (LIBERTY NP SINUS-24 and LIBERTY NP SINUS-52): results from two multicentre, randomised, doubleblind, placebo-controlled, parallel-group phase 3 trials. Lancet 2019. Available from: https://doi.org/10.1016/S0140-6736(19)31881-1.

Bannick et al., 2024
Bannick M, Shao J, Liu J, Du Y, Yi Y, Ye T. A General Form of Covariate Adjustment in Randomized Clinical Trials. Available from: https://arxiv.org/pdf/2306.10213.pdf.

Devlin et al., 2017
Devlin N, Shah K, Feng Y, Mulhern B, van Hout B. Valuing health-related quality of life: an EQ-5D-5L value set for England. Health Economics 2017; 27(1): 7-22. Available from: https://euroqol.org/eq-5d-instruments/eq-5d-5l-about/valuation/crosswalk-index-valuecalculator/

Doty et al., 1984
Doty RL, Shaman P, Dann M. Development of the University of Pennsylvania Smell Identification Test: A standardized microencapsulated test of olfactory function. Physiology & Behavior (Monograph) 1984; 32:489-502.

FDA 2023
U.S. Department of Health and Human Services Food and Drug Administration (FDA). (2023). Chronic Rhinosinusitis with Nasal Polyps: Developing Drugs for Treatment. Available from: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/chronic-rhinosinusitis-nasal-polyps-developing-drugs-treatment.

Hopkins et al., 2009
Hopkins C, Gillett S, Slack R, Lund VJ, Browne JP. Psychometric validity of the 22-item Sinonasal Outcome Test. Clinical otolaryngology 2009; 34(5):447-5.

Juniper et al., 2006
Juniper EF, Bousquet J, Abetz L, Bateman ED. Identifying ‘well-controlled’ and ‘not well-controlled’ asthma using the Asthma Control Questionnaire. Respir Med 2006; 100: 616- 621.

O’Kelly and Ratitch, 2014
O’Kelly M, Ratitch B. Clinical trials with missing data: A guide for practitioners. 2014.

Stokes et al., 2000
Stokes ME, Davis CS, Koch GG. Categorical data analysis using the SAS system (2 ed.). 2000.

Ye et al., 2023
Ye T, Bannick M, Yi Y, Shao J. Robust variance estimation for covariate-adjusted unconditional treatment effect in randomized clinical trials with binary outcomes. Stat. Theory Relat. Fields 2023; 7(2): 159–163. Available from: https://www.tandfonline.com/doi/epdf/10.1080/24754269.2023.2205802

van Hout et al., 2012
van Hout B, Janssen M, Feng Y et al. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. Value Health 2012; 15(5): 708-15.

Tokunaga et al., 2015
Tokunaga T et al., Novel scoring system and algorithm for classifying chronic rhinosinusitis: the JESREC Study. Allergy 2015; 70(8): 995–1003.

Sampson et al., 2006
Sampson H.A, A. Munoz-Furlong, R.L. Campbell, N.F. Adkinson Jr., S.A. Bock, A. Branum 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. J Allergy Clin Immunol, 117 (2006), pp. 391–397

Shih et al., 2023
Shih V et al, Psychometric validation and meaningful change thresholds of the new nasal polyposis symptom diary. Ann. Otol. Rhinol. Laryngol. 2023; 132 (12); Available from: https://doi.org/10.1177/000348942311777

8 APPENDIX

8.1 Adverse Events of Special Interest

AESIs and related definitions based on MedDRA terms are not included in this SAP to facilitate their maintenance (e.g. management of MedDRA version changes), and for convenience in using them directly in SAS programming. These detailed definitions will be finalised by the study team prior to the primary database lock and provided together with the study datasets at the time of submission.

Helminth infection

Helminth infection is directly determined from what is entered on the eCRF. A participant is considered to have this AESI if the participant has at least one preferred term where the

dedicated Helminth Infection eCRF page was also completed for that event (linked by AE number), with AE onset date during the relevant study period for analysis.

Other AESIs are defined by MedDRA term selection criteria:

Serious hypersensitivity reactions: defined as a combination of anaphylactic reaction algorithm with 72 hours time restriction and events meeting regulatory serious criteria within the 'Hypersensitivity' narrow SMQ (no time restriction). A subject will be considered to have this AESI if the subject has at least one adverse event with onset date during the relevant study period for analysis, which satisfies either of the following:

- Anaphylactic reactions:
  - Potential anaphylactic reactions will be defined on the basis of Sampson's criteria (see Sampson et al., 2006) identified using a modified Standardised MedDRA Query (SMQ), with additional constraints on the timing of the AE onset date relative to the timing of the injection.
  - Confirmed anaphylactic reactions will be those defined following medical review of the preferred terms identified as potential anaphylactic reactions, as well as any relevant supporting data.
- Hypersensitivity: events meeting regulatory serious criteria within the 'Hypersensitivity' narrow SMQ.

Serious infections: serious adverse events with MedDRA System Organ Class of "Infections and infestations"

Malignancy: based on SMQ "Malignant or unspecified tumours"

Guillain Barré Syndrome: based on SMQ "Guillain Barré syndrome"

Serious cardiac events: serious adverse events with MedDRA System Organ Class of "Cardiac disorders"

8.2 OCS conversion factors for prednisone equivalents

Total daily OCS dose will be converted to a prednisone equivalent using the following table:

Table 11 Estimated OCS dose therapy equivalence

Oral Corticosteroid	Approximate equivalence dose
Prednisone	10 mg
Prednisolone	10 mg
Cortisone	50 mg
Hydrocortisone	40 mg

Oral Corticosteroid	Approximate equivalence dose
Methylprednisolone	8 mg
Triamcinolone	8 mg
Betamethasone	1.2 mg
Dexamethasone	1.5 mg
Deflazacort	12 mg

For example, to convert a cortisone total daily dose to a prednisone equivalent total daily dose, a multiplication factor of 0.2 = 10/50 should be used.

8.3 Partial dates for adverse events and prior/concomitant medications

Dates missing the day, or both the day and month of the year adhere to the following conventions to classify AEs and to classify prior/concomitant medications:

- The missing start day is set to:
  - First day of the month of occurrence, if the start YYYY-MM is after the YYYY-MM of first study treatment
  - The day of the first study treatment, if the start YYYY-MM is the same as YYYY-MM of the first study treatment
  - The date of informed consent, if the onset YYYY-MM is before the YYYYMM of the first study treatment.
- The missing end day is set to:
  - The last day of the month of the occurrence, if the end YYYY-MM is after the YYYY-MM of the first study treatment.
  - Death date if the participant died in the same month.
  - The day of last study treatment if the YYYY-MM of occurrence is the same as the last study treatment.
- If the start date is missing both the day and month, the start date is set to:
  - January 1 of the year of occurrence.
  - The date of the first study treatment, if the start year is the same as the year of the first study treatment
- If the end date is missing both the day and month, the date is set to:
  - December 31 of the year of occurrence.
  - Date of death if the participant died in the same year.
  - Last study treatment date if the year of occurrence is the same as the last study treatment date.

- If the start date is null, the date is set to:
    - The date of first study treatment.
    - January 1 of the same year as the end date, if the end date suggests that the start date could be prior to the date of first study treatment.
- If the end date is null and not recorded as ongoing, the date is set to:
    - The date of the first study treatment, if the start date is prior to the date of first study treatment.
    - The date of last visit, if the start date is on or after the date of first study treatment.
- If the end date is null and recorded as ongoing, the end date is not imputed.

Signature Page for VV-RIM-02702313 System v5.0
D5242C00001 SAP Edition 4

Approve: Document Level Task Verdict: Approved	PPD Content Approval 21-Aug-2024 20:36:52 GMT+0000

Approve: Document Level Task Verdict: Approved	PPD Content Approval 21-Aug-2024 20:42:05 GMT+0000

Signature Page for VV-RIM-02702313 System v5.0
D5242C00001 SAP Edition 4