Statistical Analysis Plan

Protocol Title:	A Dose-Ranging, Double-Blind, Placebo-Controlled Study to Evaluate the Safety and Efficacy of Tezepelumab Alone or Combined With Topical Corticosteroids in Moderate-to-Severe Atopic Dermatitis
Short Protocol Title:	A Dose-Ranging, Placebo-Controlled, Double-Blind Study to Evaluate the Safety and Efficacy of Tezepelumab in Atopic Dermatitis
Protocol Number:	20170755
NCT Number:	**NCT03809663**
Authors:	[redacted]
Sponsor:	Amgen Inc. One Amgen Center Drive, Thousand Oaks, CA, 91320, USA
SAP Date:	**Document Version** — **Date** **Original (v1.0)** — **24 January 2019** **Amendment 1 (v2.0)** — **20 November 2019** **Amendment 2 (v3.0)** — **04 June 2020**

Version Number	Date (DDMMMYYYY)	Summary of Changes, including rationale for changes
Original (v1.0)	24JAN2019	Original version
Amendment 1 (v2.0)	20NOV2019	• Editorial changes were made throughout the document to improve overall clarity • Updated list of sections/sub-sections/appendices (including removal of sections/sub-sections/appendices which were not applicable in this study) • List of abbreviations updated to include all abbreviations used in this document • Added the secondary estimand definition, in section 2.1 • Specified the statistical hypotheses to be tested, in section 2.2 • Updated section 3 to align with the protocol amendment 2, dated 04 September 2019 • Updated section 4.2 to include additional subgroups • Updated section 5 to include all the definitions used in the document and align with the existing definitions with the latest available DES • Updated “Actual treatment received” definition to capture all the scenarios, in section 5.1 • Expanded the definitions of primary/secondary/exploratory endpoints, in section 5.4, section 5.5 and section 5.6 • Defined separate analysis sets for Part A and Part B, in section 6.1 and section 6.2 • Expanded on the details of primary, secondary and exploratory analysis, in section 9.5 • Section 9.6 updated to specify the details of safety analyses • Updated section 10 to clarify that there are no changes to the protocol specified analyses

Version Number	Date (DDMMMYYYY)	Summary of Changes, including rationale for changes
Amendment 2 (v3.0)	04JUN2020	• Editorial changes were made throughout the document to improve overall clarity due to changes in SAP amid COVID19 pandemic • List of abbreviations updated to include all abbreviations used in this document • Updated section 4.2 to clarify regions and subgroups for S. aureus and absolute eosinophil • Updated section 5.1 to clarify definition of actual treatment, TEAE and DRE • Updated section 5.3 to clarify baseline IGA, baseline EASI, improvement from baseline and duration of AD • Updated section 5.4 to modify content pertaining to IGA, EQ-5D-3L, POEM, PGI-S, Pruritus NRS, DLQI and SCORAD • Updated section 6.6 to add clarification on definition of mFAS • Updated section 8.3.2 to clarify descriptive summaries and imputation of missing data for secondary and exploratory endpoints along with missing dates needed to calculate duration of AD • Updated section 9.4 to clarify regions pertaining to demographic information and summary of baseline disease characteristics • Updated section 9.5 to clarify the analysis set to be used for efficacy analyses and statistical tests • Updated section 9.5.2 to clarify covariates of MMRM model for SCORAD and Pruritus NRS at week 16 analysis along with censoring of subjects pertaining to time to achieving EASI 50/75/90 response prior to switching • Updated section 9.5.3.1 to clarify exploratory endpoint summarization and to clarify covariates for MMRM modeling for continuous endpoints, change from baseline in each of IGA, EASI, SCORAD, Pruritus NRS, PGI-S, POEM and DLQI at week 24, week 36 and week 52 and also for change from switching date for part A. Updated part B to clarify censoring for subjects pertaining to time to achieving EASI 50/75/90 response • Updated section 9.6 to clarify two different periods of safety analyses for Part A • Updated section 9.6.1 to clarify summarization of adverse events, MedDRA query to be used for Guillain Barre Syndrome, to clarify EOI for 'severe infection' identification and exposure-adjusted event rate and exposure adjusted incidence rate for final analysis • Updated section 10 to clarify changes from protocol-specified analyses • Updated section 13 to modify analysis window for vitals of week 16 and added analysis window for S. aureus

Table of Contents

List of Figures

List of Abbreviations and Definition of Terms

Abbreviation or Term	Definition/Explanation
AD	Atopic Dermatitis
AE	Adverse Event
CI	**Confidence Interval**
CRF	Case Report Form
CTCAE	**Common Terminology Criteria Adverse Events**
DLQI	Dermatology Life Quality Index
EASI	Eczema Area and Severity Index
EASI 50	50% reduction from baseline in EASI score
EASI 75	75% reduction from baseline in EASI score
EASI 90	90% reduction from baseline in EASI score
ECG	Electrocardiogram
EQ-5D-3L	EuroQOL quality of life 5-dimensions 3-level version
EQ VAS	EuroQOL Visual Analog Scale
HRQoL	Health Related Quality of Life
Ig	Immunoglobulin
IGA	Investigator's Global Assessment
IGA 0/1	response of 0 (clear) or 1 (almost clear) skin on IGA scale
IP	Investigational Product
NRS	Numerical Rating Scale
PD	Pharmacodynamics
PGI-S	Patient Global Impression of Severity
PK	Pharmacokinetics
PRO	Patient-Reported Outcome
POEM	Patient Oriented Eczema Measure
QOL	**Quality Of Life**
Q2W	every 2 weeks
Q4W	every 4 weeks
SAE	**Serious Adverse Event**
SAP	Statistical Analysis Plan
SC	Subcutaneous
SCORAD	Scoring of Atopic Dermatitis
TCS	Topical Corticosteroid
TCI	**Topical Calcineurin Inhibitors**
TEAE	**Treatment-Emergent Adverse Event**
TSLP	Thymic Stromal Lymphopoietin
TTO	**Time trade-off**

1. Introduction

The purpose of this Statistical Analysis Plan (SAP) is to provide details of the statistical analyses that have been outlined within the protocol for study 20170755, AMG 157 dated **04 September 2019**. The scope of this plan includes two interim analyses that will be performed by an independent statistician, the primary analysis and the final analysis, that are planned and will be executed by the Amgen Global Biostatistical Science department unless otherwise specified.

2. Objectives, Endpoints and Hypotheses

2.1 Objectives and Endpoints

Objectives	Endpoints
Primary – Part A	
• To evaluate the effect of tezepelumab compared with placebo, assessed using the Investigator's Global Assessment (IGA) and Eczema Area and Severity Index (EASI)	• IGA score of 0 (clear) or 1 (almost clear) (IGA 0/1) at week 16 • 75% reduction in EASI (EASI 75) at week 16

The first primary estimand consists of:

- Target population – patients with moderate-to-severe atopic dermatitis **(AD).**
- Endpoint – IGA 0/1 at week 16.
- Intercurrent event – there are 2 intercurrent events, discontinuation of IP before week 16 and **concurrent use** of rescue medication **with TCS/TCI** between day **29** and week 16. Treatment effects will be estimated for subjects regardless of whether patients complete 16 weeks of **study treatment**. The use of rescue medication is captured through the endpoint definition **in Section 5.5, and use of TCS/TCI will be classed as a non-response.**
- Summary measure – odds ratio comparing the odds of achieving endpoint on tezepelumab 420 mg Q2W compared to the odds of achieving endpoint on placebo**.**

In summary, the first primary estimand is the odds ratio between tezepelumab 420 mg Q2W and placebo in achieving IGA 0/1 at week 16, in the absence of rescue medication with **TCS/TCI regardless of whether patients complete 16 weeks of study treatment,** in patients with moderate-to-severe AD.

The second primary estimand is the odds ratio between tezepelumab 280 mg Q2W and placebo in achieving IGA 0/1 at week 16, in the absence of rescue medication with **TCS/TCI regardless of whether patients complete 16 weeks of study treatment,** in patients with moderate-to-severe AD.

The third primary estimand is the odds ratio between tezepelumab 420 mg Q2W and placebo in achieving EASI 75 at week 16, in the absence of rescue medication with **TCS/TCI regardless of whether patients complete 16 weeks of study treatment,** in patients with moderate-to-severe AD.

The fourth primary estimand is the odds ratio between tezepelumab 280 mg Q2W and placebo in achieving EASI 75 at week 16, in the absence of rescue medication with **TCS/TCI regardless of whether patients complete 16 weeks of study treatment,** in patients with moderate-to-severe AD.

Secondary – Part A	
• To estimate the effect of tezepelumab compared with placebo on the efficacy measure of EASI	• 50% and 90% reduction in EASI (EASI 50/90) at week 16
• To estimate the time needed to reach EASI 50/75/90	• Time at which EASI 50/75/90 is achieved from day 1
• To estimate the effect of tezepelumab compared with placebo on the efficacy measure of Scoring Atopic Dermatitis (SCORAD)	• SCORAD at week 16
• To estimate the effect of tezepelumab compared with placebo on the patient-reported outcome (PRO) measure of pruritus assessed using a numeric rating scale (NRS)	• Pruritus NRS at week 16
• To characterize the pharmacokinetics (PK) of tezepelumab	• Serum trough concentrations of tezepelumab at scheduled visits

The secondary estimands are described in the same way as primary estimands except the endpoint attribute will change for EASI 50/90 at week 16, SCORAD at week 16 and Pruritus NRS at week 16, where the summary measure changes to be the Least Square Mean difference; and time at which EASI 50/75/90 is achieved from day 1, where the summary measure changes to be the median time to achieve EASI 50/75/90 from day 1.

Safety – Part A and Part B	
• To establish the safety and tolerability of tezepelumab compared with placebo	• Subject incidence of adverse events (including serious adverse events)
Exploratory – Part A and Part B	
• To estimate the long-term effect of tezepelumab on the efficacy measures of EASI, SCORAD, and Pruritus NRS	• IGA 0/1 at week 24, week 36, and week 52 • EASI 75 at week 24, week 36, and week 52 • EASI 50/90 at week 24, week 36, and week 52 • SCORAD at week 24, week 36, and week 52 • Pruritus NRS at week 24, week 36, and week 52
• To investigate potential biomarker development by biochemical analysis of blood or skin samples	• Serum biomarkers may include but are not limited to IgE, CCL17, and CCL22 • RNA transcriptional changes in blood and lesional versus non lesional skin
• To explore the effect of tezepelumab on Patient Global Impression of Severity (PGI-S)	• Change from baseline in PGI-S at scheduled visits
• To evaluate the effect of tezepelumab on health-related quality of life (HRQoL) assessed using the Dermatology Life Quality Index (DLQI)	• Change from baseline in DLQI at scheduled visits
• To explore the effect of tezepelumab on POEM	• Change from baseline in POEM at scheduled visits
• To explore the effect of tezepelumab on EQ-5D-3L	• Change from baseline in EQ-5D-3L at scheduled visits
• To evaluate the immunogenicity of tezepelumab	• Anti-tezepelumab antibodies
• To investigate the effects of drug target genes and AD genes and/or subject response to tezepelumab	• Gene polymorphisms that may influence clinical response to study drug, such as filaggrin and TSLP

Exploratory – Part B	
• To evaluate the effect of tezepelumab compared with placebo, assessed using the IGA and EASI	• IGA score of 0 (clear) or 1 (almost clear) (IGA 0/1) at week 16 • EASI 75 at week 16
• To estimate the effect of tezepelumab compared with placebo on the efficacy measure of EASI	• EASI 50/90 at week 16
• To estimate the time needed to reach EASI 50/75/90	• Time at which EASI 50/75/90 is achieved from day 1
• To estimate the effect of tezepelumab compared with placebo on the efficacy measure of SCORAD	• SCORAD at week 16
• To estimate the effect of tezepelumab compared with placebo on the PRO measure of pruritus assessed using an NRS	• Pruritus NRS at week 16
• To characterize the PK of tezepelumab	• Serum trough concentrations of tezepelumab at scheduled visits
• To establish the safety and tolerability of tezepelumab compared with placebo	• Subject incidence of adverse events (including serious adverse events)

2.2 Hypotheses and/or Estimations

Part A

The primary hypothesis is that tezepelumab effectively increases the probability of an IGA 0/1 response and an EASI 75 response at week 16 in subjects with moderate-to-severe AD.

Specifically, the four primary hypotheses are as follows:

<u>IGA 0/1 at week 16</u>

H_{01}: Odds ratio of IGA 0/1 at week 16 between tezepelumab 420 mg SC Q2W and placebo = 1

Vs.

H_{11}: Odds ratio of IGA 0/1 at week 16 between tezepelumab 420 mg SC Q2W and placebo ≠ 1

H_{02}: Odds ratio of IGA 0/1 at week 16 between tezepelumab 280 mg SC Q2W and placebo = 1

Vs.

H_{12}: Odds ratio of IGA 0/1 at week 16 between tezepelumab 280 mg SC Q2W and placebo ≠ 1

EASI 75 at week 16

H_{03}: Odds ratio of EASI 75 at week 16 between tezepelumab 420 mg SC Q2W and placebo = 1

Vs.

H_{13}: Odds ratio of EASI 75 at week 16 between tezepelumab 420 mg SC Q2W and placebo ≠ 1

H_{04}: Odds ratio of EASI 75 at week 16 between tezepelumab 280 mg SC Q2W and placebo = 1

Vs.

H_{14}: Odds ratio of EASI 75 at week 16 between tezepelumab 280 mg SC Q2W and placebo ≠ 1

An odds ratio greater than one indicates tezepelumab increases response compared to placebo**.**

Family wise type I error rate will be strongly controlled at the 0.05 level (2-sided); the details of testing strategy are described in Section 9.5.1.

Part B

Part B is an estimation study; no hypotheses will be formally tested.

3. Study Overview

3.1 Study Design

This phase 2b study is designed to evaluate the safety and efficacy of tezepelumab as a monotherapy and explore its efficacy as adjunct therapy in subjects with moderate-to-severe AD. This study consists of Part A (the main **dose ranging** study evaluating tezepelumab as a monotherapy) and Part B (a **proof-of-concept** study evaluating tezepelumab as adjunctive therapy when combined with a TCS regimen). In Part A, two futility analyses are planned when approximately **the 80th and 160th subjects are projected to complete week 16,** respectively. Upon a result of “not futile” for the first

analysis **and Part A has completed enrollment of 240 subjects**, enrollment of subjects in Part B of the study will commence. A result of “futile” for either analysis in Part A will result in termination of both Part A and Part B of the study.

Eligible subjects for both Part A and Part B will include adults with a clinical diagnosis of AD at least 2 years prior to screening, with AD that affects ≥ 10% of body surface area as assessed by EASI, an IGA score of ≥ 3, and an EASI score of ≥ 16.

Part A and Part B of this study will consist of a 28-day screening period, a 52-week treatment period, and a **20**-week safety follow-up period. During the 28-day screening period, subjects must discontinue from all topical AD therapies, except for approved moisturizers (see Section 7.1.7 in the protocol), for at least the 7 consecutive days immediately prior to day 1. During the study, approved moisturizers are allowed to be applied, with the exception of 8 hours prior to any scheduled visit.

Part A is a dose-ranging, randomized, placebo-controlled, double-blind study designed to evaluate the safety and efficacy of tezepelumab as monotherapy in subjects with moderate-to-severe AD. Following the screening period, eligible subjects will be randomized 1:1:1:1 to receive 420 mg SC Q2W, 280 mg SC Q2W, 210 mg SC Q4W, or placebo. All subjects will receive a 420 mg SC loading dose (investigational product or placebo) as the first dose. Subjects will then receive the appropriate dose of investigational product at the week 2 visit, depending on which treatment group they are randomized to.

Depending on whether subjects are randomized to Q2W or Q4W dosing, the dose at week 2**, week 6, week 10 and so on,** could be placebo or tezepelumab. Non-responders are defined as those subjects who have not achieved at least a 50% improvement in EASI at week 16 compared to baseline (day 1). Subjects who are determined to be non-responders in Part A will receive tezepelumab 420 mg SC Q2W for the remainder of the study, beginning with the week 18 dose. The randomized treatment assignments for the non-responders will remain blinded until the end of study.

Part B is a randomized, placebo-controlled, double-blind **proof-of-concept** study designed to evaluate the safety and efficacy of tezepelumab when administered with moderate class TCS in adults with moderate-to-severe AD. Enrollment of subjects in Part B of the study will only commence if there is a result of “not futile” for the first interim analysis in Part A **and Part A has completed enrollment of 240 subjects**. Following the screening period, eligible subjects will be randomized 2:1 to receive 420 mg SC

Q2W or placebo; both groups in Part B will be allowed to use TCS after the day 1 visit. Non-responders in Part B will not be switched to 420 mg SC Q2W at week 1**8** and will remain on their initial dose for the duration of the study.

3.2 Sample Size

In Part A of this study, a total of 240 subjects will be randomized in a 1:1:1:1 ratio to tezepelumab 420 mg SC Q2W, 280 mg SC Q2W, 210 mg SC Q4W, and placebo (approximately 60 in each group). There is at least 85% power for each co-primary endpoint, assuming IGA 0/1 response rates of 10% and 40% for placebo and the highest tezepelumab dose group (420 mg SC Q2W), respectively, and assuming EASI response rates of 15% and 50%, respectively. These assumed treatment effects are based on results of other new agents **(recent biologic treatment)** in this indication (DUPIXENT, 2017). These calculations used a 2-sided χ_2 test at a significance level of 0.025 (to reflect the Bonferroni-based gatekeeping procedure described in Section **9.5.1**) and assumed a 10% dropout rate.

Part B of this study is exploratory. A sample size of 60 subjects (40 in the tezepelumab 420 mg SC Q2W group and 20 in the placebo group) was chosen as an adequate sample size to provide initial evidence of tezepelumab as a promising agent for adjunct therapy. **The sample size of 60 in Part B was based on practical considerations rather than statistical considerations. For example, to observe treatment effect differences of tezepelumab 420 Q2W vs. placebo as 30% for IGA 0/1 and 35% for EASI 75 at week 16, the half-width of 95% CI will be approximately 20%.**

3.3 Adaptive Design

For Part A, two interim analyses are planned when approximately **the 80th and 160th subjects are projected to complete their week 16 assessments**, respectively. [redacted]. Further details are described in Section 7.1 **and Supplemental SAP**.

4. Covariates and Subgroups

4.1 Planned Covariates

The analysis of the primary endpoints will be adjusted for the baseline covariate as follows:

- **Primary endpoint – IGA 0/1 at week 16:** Baseline IGA
- **Primary endpoint – EASI 75 at week 16:** Baseline EASI

4.2 Subgroups

Consistency of treatment effect across subgroups on the primary efficacy variables will be explored for Part A only. The following subgroup populations of interest in Part A will be included.

Subgroup analyses are not planned for Part B.

5. Definitions

5.1 Safety Related Definitions

As detailed in Section 3.1, subjects who are determined to be non-responders in Part A should switch to tezepelumab 420 mg SC Q2W for the remainder of the study, beginning with the week 18 dose.

Actual treatment received

Part A:

For Part A of the study, actual treatment will be defined as follows:

- **If a subject either a) did not need to switch to tezepelumab 420 mg during the study, or b) should have switched to tezepelumab 420 mg but did not successfully do so, then the subject is considered a non-switcher. The subject's actual treatment received is defined below:**
  - **from day 1 until EOS for safety analysis of Day 1 to EOS (For safety analysis of Day 1 to week 16, use from day 1 until the day before the next dosing date after week 16 dose),**
    - **If a subject receives placebo only, the actual treatment will be defined as placebo**
    - **If a subject receives two or more doses of tezepelumab, the actual treatment will be defined as the highest dose of tezepelumab the subject receives since the date of the second IP (inclusive)**
    - **If a subject receives only one dose of tezepelumab, the actual treatment will be defined as the dose of tezepelumab the subject receives**
- **If a subject should have switched to tezepelumab 420 mg during the study and received a switching dose, then,**
  - **From day 1 until the day before the switch,**
    - **If a subject receives placebo only throughout this period, the actual treatment will be defined as placebo**
    - **If a subject receives two or more doses of tezepelumab during this period, the actual treatment will be defined as the highest dose of tezepelumab the subject receives since the date of the second IP (inclusive) until the day prior to the switch**
    - **If a subject receives only one dose of tezepelumab during this period, the actual treatment will be defined as the dose of tezepelumab the subject receives**
  - **From the date of switch until EOS,**
    - **If a subject receives at least one dose of tezepelumab 420 mg during this period, the actual treatment will be defined as tezepelumab 420 mg**

Part B:

For Part B of this study, the actual treatment will be defined as follows:

- **If a subject receives placebo throughout the Part B of this study, the actual treatment will be defined as placebo**
- **If a subject receives at least one dose of tezepelumab in Part B of this study, the actual treatment will be defined as the highest dose of tezepelumab the subject receives**

Treatment-Emergent Adverse Events

Events categorized as Adverse Events (AEs) including events reported as **Disease Related Events** starting on or after first dose of investigational product as determined by **"Did event start before first dose of investigational product" equal to "No" or missing** on the Events eCRF **(including either "adverse event" or "disease related event" record on eCRF) and up to End of Study Date.**

Disease Related Events

Disease-related events are events (serious or non-serious) anticipated to occur in the study population due to the underlying disease as recorded on the Events eCRF page. Details on disease related events can be found in protocol Appendix 4.

5.2 Study Dates/Timelines

End of Study (EOS) Date

The End of Study Date **for a subject** is the date **the subject completes the safety follow-up, or the date of the Early Termination (ET) visit if no safety follow-up is completed. The date the subject has ended the study is recorded on the End of Study Electronic Case Report Form (eCRF).**

Primary Completion Date

The primary completion date is the date when the last subject has completed the assessments for week 16 in Part A of the study.

If the study concludes prior to the primary completion date originally planned in the protocol (ie, early termination of the study), then the primary completion date will be the date when the last subject is assessed or receives an intervention for evaluation in the study (ie, last subject last visit).

Randomization Date

Randomization Date is defined as the date subject was allocated to a treatment group.

Study Day

The number of days from Study Day 1, inclusive, given by the following formula:

Study Day = (Date of interest – Date of Study Day 1) + 1, if date of interest is on/after Study Day 1.

Study Day = (Date of interest – Date of Study Day 1), if date of interest is before study day 1.

Study Day 1

Study day 1 is defined as the first day of IP administration after randomization. If a subject is randomized but never dosed, then set study day 1 to randomization date.

Switching Date

If a subject is determined as non-responder at week 16 in Part A as specified by the eCRF, the Switching Date of this subject will be defined as the date that the subject takes the first dose of tezepelumab 420 mg SC Q2W after week 16 in Part A.

5.3 Baseline related definitions

Study Baseline

Study baseline is defined as the last non-missing measurement for the endpoint of interest taken **on or** before **the date of** first dose of investigational product.

For IGA and EASI, the baseline are the IGA or EASI values collected at "Day 1" visit per protocol schedule of assessment. The values collected at "screening visit" per protocol schedule of assessment could not be used as baseline IGA or EASI.

Change from baseline

Change from baseline is defined as (Post-baseline Value – Baseline Value).

Improvement from baseline

For endpoints with higher scores reflecting better clinical results, improvement from baseline is defined as (Post-Baseline Value – Baseline Value).

For endpoints with lower scores as better clinical results, improvement from baseline is defined as (Baseline Value – Post-Baseline Value).

Duration of Atopic Dermatitis (years)

The number of years between the date of diagnosis (DXDT) and study Day 1, rounded to one decimal place, is given by formula below to calculate the duration:

Observed portion	Missing portion	Formula to Calculate Duration
Year, Month, Day		(DAY 1 – DXDT + 1)/365.25
Year, Month	Day	[Year(DAY 1)-Year(DXDT)]+ [Month(DAY 1)-Month(DXDT)]/12 *if duration equals 0, add 1 month or 1/12 years (this is to avoid a disease duration of 0)
Year	Month, Day	[Year(DAY 1)-Year(DXDT)] *if duration equals 0, add 1 month or 1/12 years (this is to avoid a disease duration of 0)

5.4 Study Endpoints

Investigator's Global Assessment (IGA)

The IGA allows investigators to assess overall disease severity at 1 given time point and consists of a **6**-point severity scale from clear to severe disease (0 = clear; 1 = almost clear; 2 = mild disease; 3 = moderate disease; 4 = severe disease; 5 = very severe disease). The IGA uses clinical characteristics of erythema, infiltration, papulation, oozing, and crusting as guidelines for the overall severity assessment (Breuer et al, 2004).

For subjects who take rescue medication of TCS/TCI from day 29, their IGA score will be ignored on and after the date of rescue medication.

Eczema Area and Severity Index (EASI)

The EASI is designed by modifying the Psoriasis Area and Severity Index, which has been widely used in clinical trials and has been established as a well-accepted and standardized instrument for assessing therapeutic response in patients with psoriasis (Schmitt et al, 2007). The EASI evaluates 4 natural anatomical regions for severity and extent of key disease signs and focuses on key acute and chronic signs of inflammation (ie, erythema, induration/papulation, excoriation, and lichenification). The maximum score is 72, with higher values indicating more severe disease.

For subjects who take rescue medication of TCS/TCI from day 29, their EASI score will be ignored on and after the date of rescue medication.

<table>
<tr><th>Body Region</th><th>Severity</th><th>Area</th><th>Score</th></tr>
<tr><td>Head/Neck (H)</td><td rowspan="4">The average degree of severity of each sign (erythema [E], induration/ papulation [I], excoriation [EX] and lichenification [L], is defined on a 4-point ordinal scale:
0 = none
1 = mild
2 = moderate
3 = severe</td><td rowspan="4">The area is defined on a 7-point ordinal scale:
0 = no eruption
1 = <10%
2 = 10% - 29%
3 = 30% - 49%
4 = 50% - 69%
5 = 70% - 89%
6 = 90% - 100%</td><td>(E+I+EX+L) * Area * 0.1</td></tr>
<tr><td>Upper Limbs (UL)</td><td>(E+I+EX+L) * Area * 0.2</td></tr>
<tr><td>Trunk (T)</td><td>(E+I+EX+L) * Area * 0.3</td></tr>
<tr><td>Lower Limbs (LL)</td><td>(E+I+EX+L) * Area * 0.4</td></tr>
<tr><td>EASI Score</td><td></td><td></td><td>Sum of scores of H, UL, T, and LL</td></tr>
</table>

Pruritus NRS via Electronic Diary

The electronic diary includes 1 item to capture subject-reported pruritus. The diary will be completed by the subject at home each morning using an electronic device.

EQ-5D-3L

The EuroQoL 5-Dimension Health Questionnaire 3 Level (EQ-5D-3L) is a standardized measure of health status developed by the EuroQol group in order to provide a simple, generic measure of health for clinical and economic appraisal (Greiner 2005). The EQ-5D-3L is a self-administered questionnaire used to assess health status 'today' that is divided into two sections. The first section includes one item addressing each of five dimensions (mobility, self-care, usual activity, pain/discomfort, and anxiety/depression). Each dimension has 3 ordinal levels of severity: "no problem" (1), "some problems" (2), "severe problems" (3). Each item results into a 1-digit number that expresses the level selected for that dimension. The digits for the five dimensions can be combined into a 5-digit number that describes the patient's health state.

In addition, the EQ-5D score will be calculated based on the five attributes on the EQ-5D-3L questionnaire: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. The answers are recorded as choices of 1, 2, or 3 for each question, with 1 signifying no problem, 2 signifying some problem, and 3 signifying major problem. The calculated EQ-5D score will fall on a scale of 0 to 1, with 0 representing death and 1 representing the perfect health state.

The score will be calculated only if all five attributes are answered. The algorithm for scores starts with everyone having a value of 1, and then points are subtracted from 1 depending on the responses to the five attributes.

If all questions are recorded as 1, the EQ-5D score will stay as 1, nothing will be subtracted. If at least one question is 2 or 3, 0.081 is subtracted. If at least one question is 3, 0.269 is also subtracted. Additional values to be subtracted when the attributes are recorded as 2 or 3 are as follows:

	Score	
Attribute	**2**	**3**
Mobility	**0.069**	**0.314**
Self-care	**0.104**	**0.214**
Usual activity	**0.036**	**0.094**
Pain/discomfort	**0.123**	**0.386**
Anxiety/depression	**0.071**	**0.236**

For example, if a subject recorded mobility as 2, self-care as 3, and the other three items as 1, the EQ-5D score would be 1 – 0.081 – 0.269 – 0.069 – 0.214 = 0.367.

In the second section of the questionnaire, the EQ Visual Analogue Scale (VAS) records the patient's self-rated health on a vertical visual analogue scale where the endpoints are labeled **as 100 representing** 'Best imaginable health state' **and 0 representing** 'Worst imaginable health state'. The VAS can be used as a quantitative measure of health outcome that reflects the patient's own judgment.

For subjects who take rescue medication of TCS/TCI after day 29, their EQ-5D score will be ignored on and after the date of rescue medication.

Patient Oriented Eczema Measure (POEM)

The POEM is a 7-item, validated questionnaire used in clinical practice and clinical trials to assess disease symptoms in children and adults with atopic eczema (Charman 2004). The format is patient response to 7 items (dryness, itching,

flaking, cracking, sleep loss, bleeding, and weeping) based on symptom frequency during the past week and each item is a 5-point scale

(ie., 0 = 'no days', 1 = '1 to 2 days', 2 = '3 to 4 days', 3 = '5 to 6' days, and 4 = 'every day').

The total score is the sum of the 7 items which is ranged from 0 to 28 and reflects disease-related morbidity; a high score is indicative of a poor QOL.

The following POEM banding scores have been established (Charman 2004): 0 to 2 = Clear or almost clear; 3 to 7 = Mild eczema; 8 to 16 = Moderate eczema; 17 to 24 = Severe eczema; 25 to 28 = Very severe eczema. If two or more response options are selected for a question, then the response option with the highest score is recorded. If one question of the seven is left unanswered, then that question is scored as 0 and the scores are summed and expressed as usual out of a maximum of 28. If two or more questions are left unanswered, then the questionnaire is not scored and is set to missing.

For subjects who take rescue medication of TCS/TCI after day 29, their POEM score will be ignored on and after the date of rescue medication.

Patient Global Impression of Severity (PGI-S)

The PGI-S is a single item designed to capture the subject's perception of overall symptom severity at the time of completion on a **6** point categorical response scale **(0 = no symptoms, 1 = very mild, 2 = mild, 3= moderate, 4 = severe, 5 = very severe symptoms).**

For subjects who take rescue medication of TCS/TCI after day 29, their PGI-S score will be ignored on and after the date of rescue medication.

Patient-reported Outcomes (PROs)

The electronic diary (eDiary), which includes the pruritus NRS, will be completed by the subject at home each morning during the treatment and follow-up periods. The remaining PROs will be conducted at the study center and should be completed prior to other study assessments to avoid the possibility of introducing bias to subject responses.

Pruritus Numeric Rating Scale **(NRS)**

The Pruritus NRS is a simple assessment tool that patients will use to report the intensity of their pruritus (itch) during a daily recall period using electronic diary.

Pruritus will be assessed using an NRS (0-10) with 0 = no itch and 10 = worst imaginable itch. Subjects will complete the NRS as part of the daily diary each day in the morning.

Patients will be asked the following questions:

"On a scale of 0 to 10, with 0 being 'no itch' and 10 being the 'worst itch imaginable', how would you rate your itch during the previous 24 hours?"

The mean weekly Pruritus NRS is calculated as the prorated average of the reported daily NRS within the week. For example, if there are 3 scores in a week, the prorated average = (score1 + score2 + score3)/3.

For subjects who take rescue medication of TCS/TCI after day 29, their Pruritus NRS score will be ignored on and after the date of rescue medication.

Dermatology Life Quality Index (DLQI)

The DLQI is a 10-item, subject-completed, HRQoL assessment with content specific to those with dermatology conditions. The recall period is 1 week (Finlay and Kahn, 1994). The DLQI content captures respondent perceptions of dermatology-related symptoms and feelings (embracement), impacts on daily activities, leisure, work or school, personal relationships, and the side effects of treatment. Each item is scored on a 4-point Likert scale: 0 = not at all/not relevant; 1 = a little; 2 = a lot; and 3 = very much (Basra et al, 2008).

Handling missing items from DLQI:

i. If one question is left unanswered this is scored 0 and the scores are summed and expressed as usual out of a maximum of 30.

ii. If two or more questions are left unanswered the questionnaire is not scored.

iii. If question 7 is answered 'yes' this is scored 3 even if in the same question one of the other boxes is ticked.

iv. If question 7 is answered 'no' or 'not relevant' but then either 'a lot' or 'a little' is ticked this is then scored 2 or 1.

v. If two or more response options are ticked for one question, the response option with the highest score should be recorded.

vi. The DLQI can be analyzed by calculating the score for each of its six sub-scales.

When using sub-scales, if the answer to one question in a sub-scale is missing, that sub-scale should not be scored:

Symptoms and feelings	**Questions 1 and 2**	**Score maximum 6**
Daily activities	**Questions 3 and 4**	**Score maximum 6**
Leisure	**Questions 5 and 6**	**Score maximum 6**
Work and School	**Questions 7**	**Score maximum 3**
Personal relationships	**Questions 8 and 9**	**Score maximum 6**
Treatment	**Questions 10**	**Score maximum 3**

For subjects who take rescue medication of TCS/TCI after day 29, their DLQI score will be ignored on and after the date of rescue medication.

Scoring of Atopic Dermatitis (SCORAD)

The SCORAD is a clinical tool with clinician and patient reported for assessing the severity (ie, extent, intensity) of AD. The tool evaluates the extent and intensity of the AD lesions, along with subjective symptoms (Kunz et al, 1997). The maximum total score is 103, with higher values indicating more severe disease.

If any component of SCORAD is missing, then SCORAD will be set to missing.

Component	Definition	Grading	Score
Extent (A)	• Head and neck • Upper limbs (anterior/posterior) • Lower limbs (anterior/posterior) • Anterior trunk • Back • Genitals • Hand (right/left)	• Head and neck: 9% • Upper limbs: 9% each • Lower limbs: 18% each • Anterior trunk: 18% • Back: 18% • Genitals: 1% • Hand: 1% each	Sum of each extent item score to give a total extent score A (range 0-100)
Intensity (B)	• Erythema (Redness) • Papulation/Oedema (Swelling) • Oozing/Crusting • Excoriation (Scratched) • Lichenification (Leathery) • Dryness (Ichthyosis)	Each intensity item is defined on a 4-point ordinal scale: 0 = absent 1 = mild 2 = moderate 3 = severe	Sum of each intensity item score to give a total intensity score B (range 0-18)
Subjective symptoms	• Pruritus • Sleep loss	Each symptom has a NRS (0-10) where 0 is no pruritus/sleep loss and 10 is the worst imaginable pruritus/sleep loss	Sum of each subjective symptom score to give a total symptom score C (0-20)
Total SCORAD score			(A/5) + 7*(B/2) + C

For subjects who take rescue medication of TCS/TCI after day 29, their SCORAD score will be ignored on and after the date of rescue medication.

Target Treatment Effect (TTE)

A Target Treatment Effect (TTE) for IGA 0/1 is based on an odds ratio equivalent to or greater than 5.29 comparing tezepelumab to placebo. The odds ratio of 5.29 is determined by the following assumption: tezepelumab response rate is 37% and placebo rate is 10% (Simpson et al, 2016). The observed placebo rate may be

different from the assumed 10%, therefore, the target tezepelumab response rate may be adjusted fixing the odds ratio at 5.29.

A Target Treatment Effect (TTE) for EASI 75 is based on an odds ratio equivalent to or greater than 5.67 comparing tezepelumab to placebo. The odds ratio of 5.67 is determined by the following assumption: tezepelumab response rate is 50% and placebo rate is 15% (Simpson et al., 2016). The observed placebo rate may be different from the assumed 15%, therefore, the target tezepelumab response rate may be adjusted fixing the odds ratio at 5.67.

5.5 Primary Endpoints

IGA 0/1 at week 16

The primary endpoint, IGA 0/1 at week 16 in the absence of rescue medication of TCS/TCI from day 29 to week 16 in Part A, will be derived as follows:

- **Derive IGA 0/1 at week 16 as 1, if the IGA score is 0 or 1 at week 16 and no rescue medication of TCS/TCI from day 29 to week 16**
- **IGA 0/1 at week 16 will be missing if there is no IGA assessment at week 16 and either no rescue medication of TCS/TCI from day 29 to week 16 or the use of rescue medication is unknown during day 29 to week 16**
- **Otherwise derive IGA 0/1 at week 16 as 0**

EASI 75 at week 16

The primary endpoint, EASI 75 at week 16 in Part A, will be derived as follows

- **Derive EASI 75 at week 16 as 1, if there is greater than or equal to 75% improvement in EASI score from baseline and no rescue medication of TCS/TCI from day 29 to week 16**
- **EASI 75 at week 16 will be missing if there is no EASI assessment at week 16 and either no rescue medication of TCS/TCI from day 29 to week 16 or the use of rescue medication is unknown during day 29 to week 16**
- **Otherwise derive EASI 75 at week 16 as 0.**

5.6 Secondary/Exploratory Endpoints – IGA 0/1 and EASI 50/75/90

Secondary endpoints, EASI 50 and EASI 90 at week 16 in Part A, and Exploratory endpoints, IGA 0/1 at week 24, week 36 and week 52 in Part A and B, and IGA 0/1 at week 16 in Part B, will be defined in the same way as primary endpoints in Section 5.5. For example, the time point of interest is week 24, week 36 and week 52 instead of week 16, the use of TCS/TCI will be considered from day 29, to week 24, week 36 and week 52 respectively.

Time-to-event analysis of EASI 50/75/90 will be handled as described in Section 9.5.2 and Section 9.5.3.

In addition, IGA 0/1 and EASI 75 at time points other than week 16, week 24, week 36 and week 52 in both Part A and B will be defined in the same way as primary endpoints in Section 5.5 at corresponding time points. For example, when the time point of interest is week 4, week 6, week 8, week 10, and week 12 instead of week 16, the use of TCS/TCI will be considered from day 29, to those time points respectively.

6. Analysis Sets

6.1 Full Analysis Set

Full Analysis Set – Part A

The full analysis set – Part A (FAS – Part A) will include all randomized subjects in Part A. Subjects will be analyzed according to their randomized treatment group. Demographics, baseline disease characteristics, and efficacy analyses for Part A will be based on the FAS – Part A unless otherwise specified.

Full Analysis Set – Part B

The full analysis set – Part B (FAS – Part B) will include all randomized subjects in Part B. Subjects will be analyzed according to their randomized treatment group. Demographics, baseline disease characteristics, and efficacy analyses for Part B will be based on the FAS – Part B unless otherwise specified.

6.2 Safety Analysis Set

Subjects will be analyzed according to the actual treatment (defined in Section 5.1) in safety analysis for Part A and B respectively.

Safety Analysis Set – Part A

The safety analysis set for Part A, will consist of all randomized subjects who received at least one dose of investigational product in Part A.

Safety Analysis Set – Part B

The safety analysis set for Part B, will consist of all randomized subjects who received at least one dose of investigational product in Part B.

6.3 Health-related Quality-of-Life or Health Economics Analyses Set(s)

There is no separate Health related Quality of Life or Health Economics Analyses Set(s). The Health-related Quality of Life or Health Economics endpoints will be analyzed using Full Analysis Set unless otherwise specified.

6.4 Pharmacokinetic/Pharmacodynamic Analyses Set(s)

Subjects **who receive IP and have** at least one measurable serum concentration will be included in the PK analysis set and subjects will be analyzed according to **the actual treatment received**.

6.5 Interim Analyses Set(s)

In Part A, interim analysis sets include **first 80 randomized subjects who are projected to complete week 16 assessments and first 160 randomized subjects who are projected to complete week 16 assessments, respectively.**

6.6 Study-specific Analysis Sets

Switcher Analysis Set

Switcher analysis set will include all subjects in Part A, who are identified as non-responders as specified by the eCRF and who actually receive at least one dose of 420mg Q2W from week 18.

Modified Full Analysis Set – Part A

Modified Full Analysis Set – Part A (mFAS – Part A) will include randomized subjects in Part A, who were still able to participate in the study in spite of the COVID-19 2020 outbreak.

7. Planned Analyses

7.1 Interim Analysis and Early Stopping Guidelines

For Part A, two interim analyses are planned when approximately **the 80th and 160th subjects are projected to complete their week 16 assessments,** respectively. [redacted]

The interim analyses will be performed by an independent statistician. The numerical results of the interim analyses will not be accessible to the study team. The study team will only know whether the study will continue or not on the basis of the results.

The details of interim analyses are specified in the Supplement SAP.

For Part B, no interim analyses will be performed.

7.2 Primary Analysis

The primary analysis will occur when the primary completion date described in Section 5.2 has been observed, and the data have been entered, cleaned, and locked. The objective for the primary analysis is to assess safety and efficacy through week 16 for Part A only.

7.3 Final Analysis

The final analysis will occur when the end of study date described in Section 5.2 has been observed, and the data have been entered, cleaned and locked. The final analysis will analyze efficacy and safety over the entire study period for Part A and Part B. For this, Part A and Part B will be analyzed separately.

8. Data Screening and Acceptance

8.1 General Principles

The objective of the data screening is to assess the quantity, quality, and statistical characteristics of the data relative to the requirements of the planned analyses.

8.2 Data Handling and Electronic Transfer of Data

The Amgen Global Study Operations-Data Management (GSO-DM) department will provide all data to be used in the planned analyses.

8.3 Handling of Missing and Incomplete Data

Subjects may be missing specific data points for a variety of reasons. In general, data may be missing due to a subject's early withdrawal from study, a missed visit, or inability to evaluate an endpoint at a particular point in time. The general procedures outlined below describe what will be done when a data point is missing.

8.3.1 Missing Baseline Value

Missing baseline evaluations will not be imputed.

8.3.2 Missing Post-baseline Evaluation

Refer to Section 9.5.1.1 for handling of missing data for primary estimands defined in Section 2.1.

Secondary and exploratory endpoints will be included in descriptive summaries as observed without imputation for missing data. In analyses of secondary/exploratory endpoints involving statistical models detailed in Section 9.5.2 and Section 9.5.3, where endpoints cannot be derived due to missing IGA or EASI assessments, the missing binary endpoints (IGA 0/1, EASI50, EASI75,

EASI90) will be imputed using non-responder imputation (NRI). Missing binary endpoints that are not derived from IGA or EASI data (eg, Pruritus NRS >= 2) and missing continuous endpoints will not be imputed unless otherwise specified.

Missing post-baseline safety data will not be imputed.

8.3.3 Missing and Incomplete Dates

For any listings, missing or incomplete dates will be listed as is.

For safety and concomitant medication tables, if any missing or incomplete date is reported, then, the following imputation rule will be used to impute the date.

	Missing	Imputation	Exception
Start Date	Day	Impute with 01	Default to Study Day 1 if an event starts the same year and month as Study Day 1 and the stop date is after Study Day 1
	Day and Month, or Month	Impute with 01 JAN	Default to Study Day 1 if an event started the same year as Study Day 1 and stop date is after Study Day 1
	Day, Month and Year	No imputation	
End Date	**Day**	**Impute with minimum of last day of the month, EOS date**	
	Day and Month, or Month	**Impute with minimum of December 31 of the year, snapshot date for milestone analysis, EOS date**	
	Day, Month and Year	**No imputation**	

If the imputed start date is after the end date, then the imputed start date will be defaulted to the end date.

If the imputed end date is before the start date, then the imputed end date will be defaulted to the start date.

For missing dates needed to calculate duration of atopic dermatitis, refer to Section 5.3 for formulas to calculate duration.

8.4 Detection of Bias

This study has been designed to minimize potential bias using randomization of subjects into treatment groups and the use of blinding. Other factors that may bias the results of the study include:

- major protocol deviations likely to impact the analysis and interpretation of the efficacy endpoints
- subject level unblinding before final database lock and formal unblinding
- IP dosing non-compliance
- reasons for early withdrawal from treatment or from study

Important protocol deviations likely to impact the analysis and interpretation of the efficacy endpoints, **and would significantly affect subject's right, safety or wellbeing** will be tabulated **by treatment group** in the Clinical Study Report (CSR). Any unblinding of individual subjects prior to formal unblinding of the study will be documented in the CSR.

8.5 Outliers

Descriptive summaries will be examined to identify unexpected values.

Outliers due to data entry errors will be corrected by the study team before data lock. Outliers that are not due to data entry will be included in the analysis. If it is deemed necessary after the team reviews the output from the planned analyses after data lock, a post-hoc sensitivity analysis excluding subjects with outliers may be performed.

8.6 Distributional Characteristics

For categorical endpoints, descriptive summary will be provided. For continuous endpoints, normality will be assumed given the large sample size where central limit theorem applies.

8.7 Validation of Statistical Analyses

Programs will be developed, maintained, and output will be verified in accordance with current risk-based quality control procedures.

Tables, figures, and listings will be produced with validated standard macro programs where standard macros can produce the specified outputs.

The production environment for statistical analyses consists of Amgen-supported versions of statistical analysis software; for example, the SAS System version 9.4 or later.

9. Statistical Methods of Analysis

9.1 General Considerations

Subject disposition, demographics, and baseline disease characteristics will be summarized descriptively by randomized treatment group **using the FAS, by Part A and B respectively.**

All efficacy analyses will be performed using the FAS by Part A and B respectively, unless stated otherwise. All safety analyses will be performed using safety analysis set, for Part A and B separately.

For categorical endpoints, the descriptive statistics will contain frequency and percentage. For continuous endpoints, the descriptive statistics will include the number of observations, mean, standard deviation, 1st quartile, 3rd quartile, median, minimum, and maximum.

9.2 Subject Accountability

Subject disposition will be summarized descriptively for all randomized subjects, by randomized treatment group, **in Part A and B respectively**.

The disposition will include the number of subjects who are randomized, who are dosed with investigational product, **who discontinue IP including reasons of discontinuation,** who complete 16 weeks **of the study**, who complete 52 weeks **of the study**, who complete the whole study, and who withdraw prematurely including their reasons for withdrawal.

The disposition for **non-responders** will include the number of subjects **who switch to 420 mg SC Q2W**, who complete 52 weeks **of the study**, who complete the study, and who withdraw prematurely including their reasons for withdrawal.

9.3 Important Protocol Deviations

Important Protocol Deviations (IPDs) categories are defined by the study team before the first subject's initial visit and updated during the IPD reviews throughout the study prior to database lock. These definitions of IPD categories, subcategory codes, and descriptions will be used during the course of the study. Eligibility deviations are defined in the protocol.

9.4 Demographic and Baseline Characteristics

Demographic and baseline disease characteristics will be summarized by treatment group using descriptive statistics **based on FAS by Part A and B separately**.

Demographic information related to sex, age, race, region (Eastern EU which includes Czech Republic, Estonia, Hungary, Latvia, Poland, and Ukraine, Western EU which includes Germany, UK, Spain and Australia, North America including USA and Canada, Asia including South Korea and Japan), weight, height, and body mass index (BMI) will be presented by treatment group and overall. If multiple races have been reported for a subject, the subject will be categorized as multiple-race, as well as by combination of races.

A summary of baseline disease characteristics includes as follows:

1) Continuous variables to be summarized will include mean, standard deviation, 1st quartile, 3rd quartile, median, minimum, and maximum

- **Screening total serum IgE level (kU/L)**
- **IGA**
- **EASI score**
- **SCORAD, PGI-S, EQ-5D-3L, EQ VAS Score, Pruritus NRS, POEM, DLQI**
- **Absolute eosinophil (cells/µL)**
- **Absolute neutrophil**
- **Basophil**

2) Categorical variables to be summarized will include, the number and % of subjects in each category

- **Screening total serum IgE level (< 150 kU/L vs. ≥ 150 kU/L)**
- **EASI score (≤ 25 points vs. > 25 points)**
- **IGA (Category 3 vs. Category ≥ 4)**
- **Pruritus NRS (<4, >=4 and <7 vs. ≥ 7; ≥2 vs <2, ≥3 vs <3, ≥4 vs. <4)**
- **POEM (0 to 2; 3 to 7; 8 to 16; 17 to 24; 25 to 28)**
- **S. aureus (positive vs. negative) (lesional and non-lesional skin separately)**
- **Age (18-35 years of age vs. 36-75 years of age)**
- **Gender (Male vs. Female)**
- **Race (White vs. non-White)**
- **Absolute eosinophil (< 300 cells/µL vs. ≥ 300 cells/µL)**
- **TCS use at Screening (Low Potency TCS vs. Medium to High Potency TCS vs. No TCS)**
- **Prior use of methotrexate (Yes vs No)**
- **Prior use of systemic immunosuppressants for AD (Yes vs No)**

Additional characteristics may be considered if scientific interest arises.

Medical history and AD history will be summarized by treatment group and for all subjects combined.

9.5 Efficacy Analyses

The multiplicity for analyses of primary estimands will be adjusted for tezepelumab 420 mg vs. placebo and 280 mg vs. placebo comparison using the method described in Section 9.5.1.1. **For sensitivity analyses, subgroup analyses of primary estimands, and analyses of secondary/exploratory endpoints in Part A and B,** significance testing, if performed, will be considered descriptive **at a level of 0.05 (2-sided) with nominal p-values without adjustment for multiplicity**.

When modeling odds ratio based on logistic regression in analyses described in Section 9.5.1, Section 9.5.2 and Section 9.5.3, in case of sparse data in one or more particular treatment arm that results in convergence issues, logistic regression model adjusted by dichotomized baseline score (<=median, >median) will be used. Odds ratio and associated 95% CI and p-value will be reported.

The following descriptive summaries will be provided by treatment at available time points using FAS and observed cases unless otherwise specified

- **Categorical efficacy endpoints: frequency/percentage.**
- **Continuous efficacy endpoints and associated changes from baseline: mean, standard deviation, median, 1st quartile, 3rd quartile, minimum, and maximum.**

The above analyses will be performed using FAS-Part A by truncating the observations collected after switching to 420 mg SC Q2W for EASI-50 non-responders, and will be repeated using Switcher Analysis Set.

9.5.1 Analyses of Primary Efficacy Endpoint(s)

9.5.1.1 Primary Analysis of the Primary Endpoints

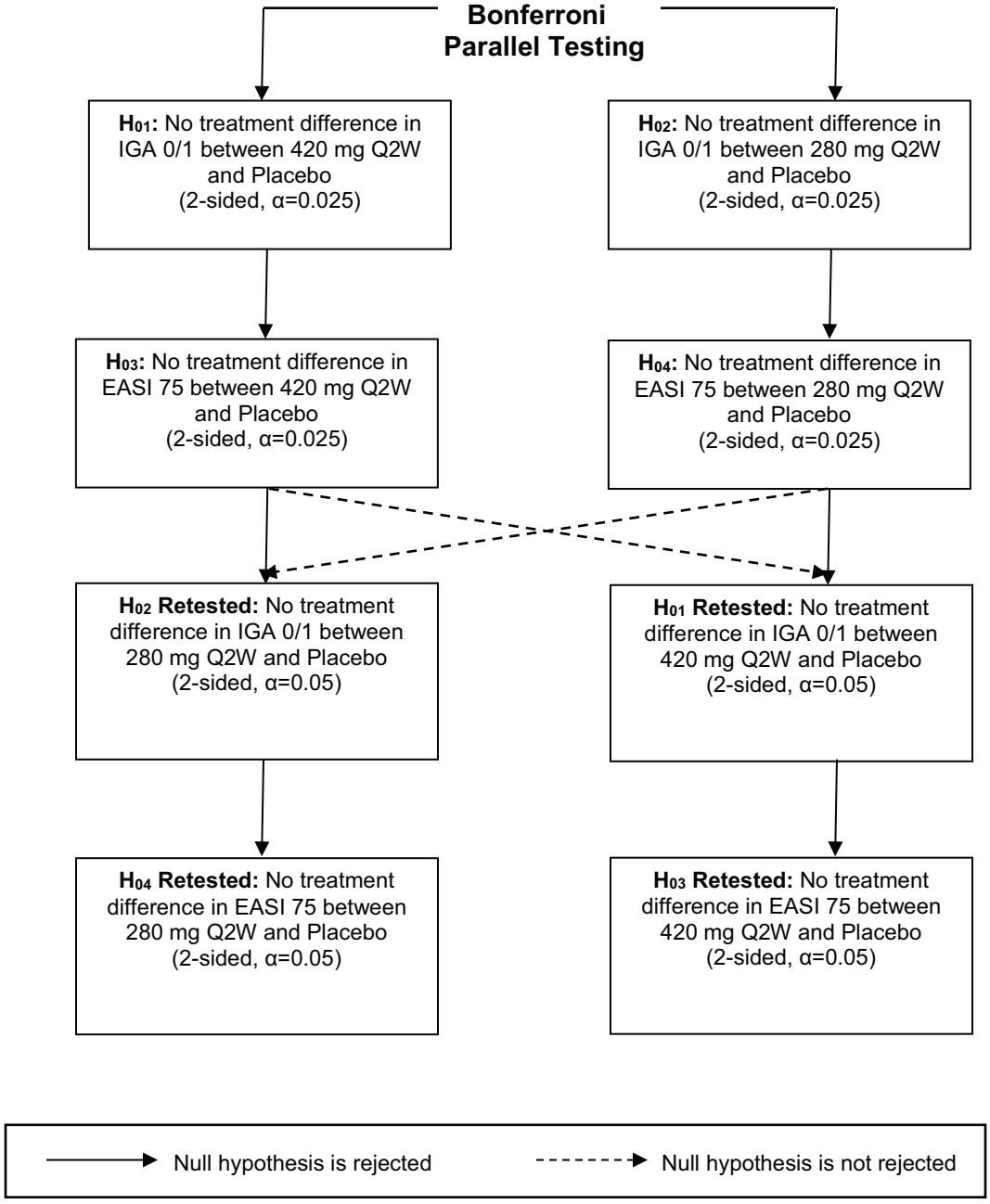
The primary analysis of the primary **estimands as defined in Section 2.1 will be performed using the FAS – Part A. Primary endpoints are defined in Section 5.5. Where an endpoint cannot be derived due to missing IGA or EASI assessments, a patient will be assumed not to have achieved the outcome, that is, Non-Responder Imputation (NRI).**

Treatment effect **for IGA 0/1 at week 16** will be tested using **a** logistic regression model with baseline IGA as covariate. **Similarly, treatment effect for EASI 75 at week 16 will**

be tested using a logistic regression model with baseline EASI as covariate. From **each** model, odds ratio and 95% CIs will be reported, comparing each tezepelumab dose group to placebo.

To control the family-wise type 1 error rate at 0.05 **(2-sided)** for the multiple comparisons of 420 mg vs. placebo and 280 mg vs. placebo, the hypotheses **described in Section 2.2** will be tested using a Bonferroni-based gatekeeping chain procedure **(Dmitrienko 2013)** as shown in Figure 9-1. This procedure will split alpha of 0.05 equally to test in parallel tezepelumab 420 mg SC Q2W versus placebo and tezepelumab 280 mg SC Q2W versus placebo for IGA 0/1 and EASI 75 sequentially within each parallel **dose** path. If one of the parallel **dose** paths rejects **the 2** hypotheses sequentially, and the other **dose** path has at least one hypothesis not rejected, then the unspent α of 0.025 from the successful path will be propagated to the **unsuccessful dose path** to retest the 2 hypotheses sequentially at a level of 0.05.

Figure 9-1. Alpha Split Between Two Doses of Interest



In addition, analysis of primary endpoints may be considered using mFAS-Part A as a supporting analysis to examine the impact of COVID-19 2020 outbreak on outcomes without multiplicity adjustment.

9.5.1.2 Sensitivity Analysis of Primary Estimands

Sensitivity analysis will be carried out using Multiple Imputation (MI) for **handling missing IGA or EASI that prevents the primary endpoints being derived**. The imputation model **for IGA score** will include treatment group, **age** and **IGA values at all**

available time points up to week 16 as covariates. **The imputed IGA scores will be used to derive IGA 0/1 at week 16 per Section 5.5. Similarly, the imputation model for EASI score will include treatment group, age and EASI values at all available time points up to week 16 as covariates. The imputed scores will then be used to derive EASI 75 at week 16 per Section 5.5.**

As a second sensitivity analysis, LOCF imputation will be used for handling missing data that prevents the primary endpoints being derived. The treatment effect for IGA 0/1 at week 16 will be analyzed using a logistic regression model, with baseline IGA scores and age as covariates. Similarly, the treatment effect for EASI 75 at week 16 will be analyzed using a logistic regression model with baseline EASI scores and age as covariates.

Third sensitivity analysis will use the same logistic regression models used in Section 9.5.1.1 with an additional covariate, screening total serum IgE level (< 150 kU/L or ≥ 150 kU/L).

9.5.1.3 Subgroup Analyses of Primary Endpoints

The consistency of the detected overall treatment effect on the primary endpoints will be explored for Part A only. A logistic regression model will include baseline IGA or EASI, the subgroup being tested, treatment, and treatment by subgroup interaction. The odds ratio will be calculated for each category of subgroup populations. The subgroups mentioned in Section 4.2, will be used in this subgroup analyses.

9.5.2 Analyses of Secondary Efficacy Endpoint(s)

The secondary estimands will be analyzed using **FAS – Part A. Secondary endpoints are defined in Section 5.6**.

For the binary endpoint, EASI 50/90 at week 16, the treatment effect will be **analyzed** using a logistic regression model adjusting for baseline EASI. From this model, odds ratio and 95% CIs will be reported comparing each tezepelumab dose group to placebo.

The time to achieving EASI 50/75/90 response prior to switching to tezepelumab 420 mg Q2W will be summarized by Kaplan-Meier (KM) curves, the number of subjects achieving EASI 50/75/90 and the number of subjects censored will be provided at available time points. Subjects who do not achieve EASI 50/75/90 will be censored at the date when formal assessments of non-responder (as defined in protocol, that is, not achieve at least 50% improvement in EASI at week 16

compared to baseline) are recorded or end of study date, whichever is the earliest. Subjects who receive rescue medication from day 29 will be censored at the point when subjects receive rescue medication.

The treatment effect for change from baseline in SCORAD and Pruritus NRS each at week 16 will be **analyzed** using a **Mixed Model for Repeated Measures (MMRM)**. The model **for change from baseline in SCORAD at week 16** will include **visit** and baseline SCORAD as covariates **and an interaction term between arm and visit**. **The model for change from baseline in Pruritus NRS at week 16 will include visit and baseline Pruritus NRS as covariates and an interaction term between arm and visit. AR(1) variance-covariance structure will be used.** Least square means and corresponding 95% CI will be calculated for each treatment group and for the difference between treatment groups.

Serum concentration will be summarized descriptively by treatment group for each sampling time point using the PK analysis set.

9.5.3 Analyses of Exploratory Efficacy Endpoint(s)

9.5.3.1 Exploratory – Part A and Part B

The exploratory endpoints for Part A and Part B per Section 5.6 will be analyzed separately.

Part A

For binary endpoints, IGA 0/1, EASI 75, EASI 50/90, each at week 24, week 36 and week 52, the treatment effect will be analyzed using a logistic regression model adjusting for corresponding baseline values as covariates. From each model, odds ratio and 95% CIs will be reported comparing each tezepelumab dose group to placebo.

SCORAD, DLQI, Pruritus NRS, POEM, and PGI-S scores and changes from baseline will be summarized by visit. Absolute improvement and percent improvement will be summarized for the EQ-5D TTO score. In addition, absolute and percent improvement will be calculated for each of the five EQ-5D-3L attributes individually. The EQ VAS score will be summarized.

For continuous endpoints, change from baseline in each of IGA, EASI, SCORAD, Pruritus NRS, PGI-S, POEM, and DLQI at week 24, week 36 and week 52, the treatment effect will be analyzed using the MMRM models with visit and corresponding baseline scores as covariates and an interaction term between arm

and visit. AR(1) variance-covariance structure will be used. Least square means and corresponding 95% CI will be calculated for each treatment group and for the difference between treatment groups. Efficacy data collected after switching to 420 mg SC Q2W point will be truncated.

The MMRM models will be repeated using the Switcher Analysis Set only, utilizing all available observations for subjects in the set. For binary outcomes of IGA 0/1 and EASI 75 at week 24, week 36 and week 52, descriptive summary of response rates will be provided by randomized treatment arm for the Switcher Analysis Set. For continuous outcomes of IGA, EASI, SCORAD, Pruritus NRS, PGI-S, POEM, and DLQI change from Baseline will be summarized descriptively by randomized treatment arm for the Switcher Analysis Set.

Biomarker analysis, if performed, will be detailed in Section 9.7.4.

Part B

The analysis of exploratory endpoints for Part B will be performed using the **FAS – Part B. The methods of analyses of exploratory endpoints for Part B will be the same as the analysis of exploratory endpoints for Part A. In addition, Bayesian methods may be used to borrow historical control data in estimating treatment effect, in addition to using observed data from the current study.**

The time to achieving EASI 50/75/90 response prior to switching to tezepelumab 420 mg Q2W will be summarized by Kaplan-Meier (KM) curves, the number of subjects achieving EASI 50/75/90 and the number of subjects censored will be provided at available time points. Subjects who do not achieve EASI 50/75/90 will be censored at the date when formal assessments of non-responder (as defined in protocol, that is, not achieve at least 50% improvement in EASI at week 16 compared to baseline) are recorded or end of study date, whichever is the earliest. Subjects who receive rescue medication from day 29 will be censored at the point when subjects receive rescue medication.

9.6 Safety Analyses

The safety analyses will be performed separately for Part A and Part B. Furthermore, the safety analyses for Part A will be performed in two different periods considering that the switching of 420mg Q2W is scheduled on week 18 for EASI 50 non-responders.

- **Day 1 to week 16 for all subjects in Safety Analysis Set – Part A.**
    - **For non-switchers, search window stops at 1 day before the next dosing date after week 16 dose, or end of safety follow up**
    - **For switchers, search window stops at 1 day before switching to 420mg Q2W**
- **Day 1 to EOS, for all subjects in Safety Analysis Set – Part A. For subjects in the Switcher Analysis Set, the actual treatment received is the combination of the actual treatment they received before switch and tezepelumab 420 mg.**

The safety analyses will be performed by actual treatment received (defined in Section 5.1).

9.6.1 Adverse Events

The Medical Dictionary for Regulatory Activities (MedDRA) version 22.0 or later will be used to code all events categorized as adverse events, to a system organ class and a preferred term.

The subject incidence of adverse events will be summarized for all treatment-emergent adverse events, serious adverse events, adverse events leading to withdrawal of investigational product and fatal adverse events.

Subject incidence of all treatment-emergent adverse events**, serious adverse events**, adverse events leading to withdrawal of investigational product, and fatal adverse events will be tabulated by system organ class and preferred term **in descending order of frequency based on total summary**. **Adverse events and disease related events will be summarized together** as defined in **Section 5.1.**

Summary of treatment-emergent **adverse events will be tabulated by system organ class, preferred term, grade, relatedness and seriousness.**

Summary of serious adverse events will be tabulated by system organ class, preferred term **grade and relatedness**.

A separate disease related events listing will be provided.

In addition, subject incidence of Events of Interest (EOI) will also be summarized. EOI will be identified using an Amgen Standardized MedDRA Query (AMQ) for the following:

[redacted]

Exposure-adjusted event rate and exposure adjusted incidence rate will be reported from day 1 to week 52 only in Part A, in the final analysis.

9.6.2 Laboratory Test Results

Absolute value and change from baseline of continuous laboratory parameters will be summarized descriptively by treatment group at each visit.

Grade 0 through 4 toxicities, as defined by the **Common Terminology Criteria Adverse Events** (CTCAE) version 5.0 (or later) will be presented for each laboratory parameter when available. **Central laboratory normal reference ranges will be used for the identification of individual clinically important abnormalities. A shift table will be produced for toxicity grade shifts, not low, normal, high values.** Shift tables will be created for hematology, immunoglobulins and IgE.

In addition, Creatine Kinase (CK) and liver function test (LFT) abnormalities will be assessed by the incidence overall and by visits of the following categories:

- **CK > 5 × ULN**
- **CK > 10 × ULN**
- **ALT or AST > 3 × ULN**
- **ALT or AST > 5 × ULN**
- **INR > 3**
- **Total bilirubin > 2 × ULN**
- **(ALT or AST > 3 × ULN) and Total bilirubin > 2 × ULN and ALP < 2 × ULN**

9.6.3 Vital Signs

Absolute value and change from baseline value will be summarized descriptively by treatment group at each visit. If there is an unexpected safety finding, vital signs will be flagged and summarized descriptively as baseline disease summaries.

9.6.4 Physical Measurements

Height and weight will be summarized descriptively as baseline summaries.

9.6.5 Electrocardiogram

The electrocardiogram (ECG) measurements from this clinical study were performed as per standard of care for routine safety monitoring, rather than for purposes of assessment of potential QT interval corrected (QTc) effect. Because these evaluations may not necessarily be performed under the rigorous conditions expected to lead to meaningful evaluation of QTc data, neither summaries nor statistical analyses will be provided, and these data would not be expected to be useful for meta-analysis with data from other trials.

9.6.6 Antibody Formation

The number and percentage of subjects with binding and neutralizing anti-tezepelumab antibodies will be summarized by treatment group, if applicable.

9.6.7 Exposure to Investigational Product

Descriptive statistics will be produced to describe the exposure to **IP** by treatment group, **for Part A and B respectively**.

9.6.8 Exposure to Non-investigational Product

Descriptive statistics will be produced to describe the exposure to TCS by treatment group.

9.6.9 Exposure to Concomitant Medication

The number and proportion of subjects receiving therapies of interest (TCS and TCI) will be summarized by **preferred term or category for each** treatment **group as coded by the World Health Organization Drug (WHO DRUG) dictionary**.

9.7 Other Analyses

9.7.1 Analyses of Pharmacokinetic or Pharmacokinetic/Pharmacodynamic Endpoints

All subjects randomized to tezepelumab treatment groups will have pharmacokinetic samples assessed. Descriptive statistics of tezepelumab serum concentrations by treatment group and visit will be provided. Summary of mean

and individual serum tezepelumab concentrations by treatment group will be summarized in tables and/or figures and provided. Additional PK/PD analysis and exposure-response analysis may be performed as needed and reported separately.

9.7.2 Analyses of Clinical Outcome Assessments

Analyses of COA endpoints are covered in Section 9.5.2 and Section 9.5.3.

9.7.3 Analyses of Health Economic Endpoints

Analyses of COA endpoints are covered in Section 9.5.2 and Section 9.5.3

9.7.4 Analyses of Biomarker Endpoints

For skin swabs for S. aureus testing, the number of subjects with positive response will be summarized by treatment group and visit.

Other biomarker analysis will be conducted if interest arises.

10. Changes From Protocol-specified Analyses

Exploratory objectives "To explore the effect of tezepelumab on POEM" and "To explore the effect of tezepelumab on EQ-5D-3L" were not originally specified in the protocol, but are added here for completeness in analysis of patient reported outcomes.

In protocol Section 9.2.3.1.5, [redacted] " is removed, because study database does not provide the attribute. Instead, a modification is made in this SAP Section 9.6.1 as the following:

"EOI will be identified using events within the system organ class (SOC) of [redacted]" and which meet the following criteria:

[redacted]

11. Literature Citations / References

Basra MK, Fenech R, Gatt RM, et al. The Dermatology Life Quality Index 1994-2007: a comprehensive review of validation data and clinical results. Br J Dermatol. 2008; 159(5):997-1035.

Breuer K, Braeutigan M, Kapp A, et al. Influence of pimecrolimus cream 1% on different morphological signs of eczema in infants with atopic dermatitis. Dermatology. 2004; 209(4):314-320.

DUPIXENT (Dupilumab) [instructions for use]. Tarrytown, NY: Regeneron Pharmaceuticals, Inc. and Sanofi-Aventis US, LLC; 2017.

Finlay AY, Khan GK. Dermatology Life Quality Index (DLQI) – a simple practical measure for routine clinical use. Clin Exp Dermatol. 1994; 19(3):210-216.

Kunz B, Oranie AP, Labreze L et al. Clinical validation and guidelines for the SCORAD index: consensus report of the European Task Force on Atopic Dermatitis. Dermatology. 1997; 195(1):10-19.

Schmitt J, Langan S, Williams HC; European Dermato-Epidemiology Network. What are the Best Outcome Measurements for Atopic Eczema? A Systematic Review. J Allergy Clin Immunol. 2007; 120:1389-1398.

12. Prioritization of Analyses

Two interim analyses will occur when approximately the 80th and 160th subjects are projected to complete week 16 assessments in Part A. The trial (both Part A and B) will be stopped for futility **based on the results of either interim analyses**.

Primary analysis will be performed when all randomized subjects complete week 16 assessments in Part A.

Exploratory analysis for Part B will occur when all randomized subjects complete week 16 assessments in Part B.

Final analysis will occur at the end of study date.

13. Appendices

Appendix A. Analytical Window for Evaluation

Per protocol, visits are to be performed within 3 days of the protocol-specified study day. To allow for variations in scheduling, the following visit windows will be applied to selected efficacy and safety evaluations (ie, vital signs, laboratory evaluations) to assign a most appropriate nominal visit for analysis. If more than one assigned visit falls within the same defined window, the closest visit to the target day (ie, scheduled visit week × 7 + 1) will be considered for analysis. If two assessment dates are the same distance from the target day, then the latest visit will be considered for analysis. If more than one evaluation falls on the same date and time for laboratory results, then the value with the smallest accession number will be used.

Analysis window for Vitals

Visit Week	Target Day	Window Definition
Baseline	1	Last Evaluation prior to or on Study Day 1
Week 2	15	Study Day 2 to 21
Week 4	29	Study Day 22 to 35
Week 6	43	Study Day 36 to 49
Week 8	57	Study Day 50 to 63
Week 10	71	Study Day 64 to 77
Week 12	85	Study Day 78 to 91
Week 14	99	Study Day 92 to 105
Week 16	113	Study Day 106 to 119
Week 18	127	Study Day 120 to 133
Week 20	141	Study Day 134 to 147
Week 22	155	Study Day 148 to 161
Week 24	169	Study Day 162 to 175
Week 26	183	Study Day 176 to 189
Week 28	197	Study Day 190 to 203
Week 30	211	Study Day 204 to 217
Week 32	225	Study Day 218 to 231
Week 34	239	Study Day 232 to 245
Week 36	253	Study Day 246 to 259
Week 38	267	Study Day 260 to 273
Week 40	281	Study Day 274 to 287
Week 42	295	Study Day 288 to 301
Week 44	309	Study Day 302 to 315

Week 46	323	Study Day 316 to 329
Week 48	337	Study Day 330 to 343
Week 50	351	Study Day 344 to 357
Week 52	365	Study Day 358 to 385
Week 58	407	Study Day 386 to 448
Week 70	491	Study Day 449 to 494

Analysis window for EASI and IGA

Visit Week	Target Day	Window Definition
Baseline	1	Last Evaluation **after Randomization and** prior to or on Study Day 1
Week 4	29	Study Day 2 to **35**
Week 6	43	Study Day **36** to 49
Week 8	57	Study Day 50 to 63
Week 10	71	Study Day 64 to 77
Week 12	85	Study Day 78 to **98**
Week 16	113	Study Day **99** to **140**
Week 24	169	Study Day **141** to **210**
Week 36	253	Study Day **211** to 294
Week 48	337	Study Day 295 to 343
Week 50	351	Study Day 344 to 357
Week 52	365	Study Day 358 to **427**
Week 70	491	Study Day **428** to 494

Analysis Window for SCORAD, EQ-5D-3L, POEM, DLQI and PGI-S

Visit Week	Target Day	Window Definition
Baseline	1	Last Evaluation prior to or on Study Day 1
Week 4	29	Study Day 2 to 42
Week 8	57	Study Day 43 to 70
Week 12	85	Study Day 71 to 98
Week 16	113	Study Day 99 to **140**
Week 24	169	Study Day **141** to **210**
Week 36	253	Study Day **211** to 294
Week 48	337	Study Day 295 to 343
Week 50	351	Study Day 344 to 357
Week 52	365	Study Day 358 to 368

Analysis Window for Serum/Urine Pregnancy Test

Visit Week	Target Day	Window Definition
Baseline	1	
Week 4	29	Study Day 2 to 42
Week 8	57	Study Day 43 to 70
Week 12	85	Study Day 71 to 98
Week 16	113	Study Day 99 to 126
Week 20	141	Study Day 127 to 154
Week 24	169	Study Day 155 to 182
Week 28	197	Study Day 183 to 210
Week 32	225	Study Day 211 to 238
Week 36	253	Study Day 239 to 266
Week 40	281	Study Day 267 to 294
Week 44	309	Study Day 295 to 322
Week 48	337	Study Day 323 to 350
Week 52	365	Study Day 351 to 427
Week 70	491	Study Day 428 to 494

Analysis Window for Coagulation

Visit Week	Target Day	Window Definition
Baseline	1	
Week 14	99	Study Day 2 to 224
Week 50	351	Study Day 225 to 357
Week 52	365	Study Day 358 to 368

Analysis Window for Hematology, Chemistry, Urinalysis (dipstick)

Visit Week	Target Day	Window Definition
Baseline	1	
Week 12	85	Study Day 2 to 126
Week 24	169	Study Day 127 to 210
Week 36	253	Study Day 211 to 294
Week 48	337	Study Day 295 to 350
Week 52	365	Study Day 351 to 385
Week 58	407	Study Day 386 to 448
Week 70	491	Study Day 449 to 494

Analysis Windows for Anti-Tezepelumab Antibodies

Visit Week	Target Day	Window Definition
Baseline	1	
Week 12	85	Study Day 2 to 126
Week 24	169	Study Day 127 to 210
Week 36	253	Study Day 211 to 294
Week 48	337	Study Day 295 to 413
Week 70	491	Study Day 414 to 494

Analysis Windows for Serum Immunoglobulins

Visit Week	Target Day	Window Definition
Baseline	1	
Week 16	113	Study Day 2 to 175
Week 34	239	Study Day 176 to 287
Week 48	337	Study Day 288 to 350
Week 52	365	Study Day 351 to 368

Note: If more than one visit (including the unscheduled visits) falls within the same defined window with non-missing measurements, scheduled visit will be used regardless of the distance from the target day. Unscheduled visit will only be used when there is no measurement from scheduled visit in the defined window.

Analysis Windows for S. aureus

Visit Week	Target Day	Window Definition
Baseline	1	
Week 16	113	Study Day 8 to 127