Statistical Analysis Plan

Protocol Title:	A Phase 3, 24-week, Randomized, Placebo-controlled, Double-blind Study to Assess the Efficacy, Safety and Tolerability of Rocatinlimab (AMG 451) Monotherapy in Adult Subjects With Moderate-to-severe Atopic Dermatitis (AD) (ROCKET-Ignite)
Short Protocol Title:	A Phase 3, Placebo-controlled, Double-blind Study Assessing Rocatinlimab Monotherapy in Moderate-to-severe AD (ROCKET-Ignite)
Protocol Number:	20210142
NCT Number:	NCT05398445
Authors:	[redacted]
Sponsor:	Amgen One Amgen Center Drive Thousand Oaks, CA, 91320
SAP Date:	**Document Version** **Date** **Original (v1.0)** 11 December 2024

Version Number	Date (DDMMMYYYY)	Summary of Changes, including rationale for changes
Original (v1.0)	11DEC2024	Not applicable

Table of Contents

List of Tables

List of Figures

[redacted]

List of Abbreviations

Abbreviation	Explanation
AD	atopic dermatitis
ALP	alkaline phosphatase
ALT	alanine aminotransferase
[redacted]	
AST	aspartate aminotransferase
BSA	body surface area
[redacted]	
COVID-19	coronavirus disease-19
C_{trough}	concentration trough
CRF	Case Report Form
DLQI	Dermatology Life Quality Index
EASI	Eczema Area and Severity Index
[redacted]	
EASI 75	achievement of ≥ 75% reduction from baseline in EASI score
[redacted]	
ECG	electrocardiogram
eCOA	electronic clinical outcome assessments
eCRF	electronic Case Report Form
[redacted]	
EOS	end of study
EOT	end of treatment
EQ-5D-5L	EuroQol 5 Dimension-5 Level
[redacted]	
GCP	Good clinical practice
HADS	Hospital Anxiety and Depression Scale
[redacted]	
IGA	Investigator's Global Assessment
IPD	Important protocol deviation
JAK	Janus kinase
[redacted]	

Abbreviation	Explanation
[redacted]	
NRS	numeric rating scale
NICE	National Institute of Health and Care Excellence
PDE4	phosphodiesterase type 4
PK	pharmacokinetics
POEM	Patient Oriented Eczema Measure
PRO	patient reported outcome
[redacted]	
QTc	corrected QT interval
rIGA™	revised Investigator's Global Assessment
rIGA™ 0/1	achieving vIGA-AD 1 response with presence of only barely perceptible erythema or vIGA-AD 0 response and ≥ 2-point reduction from baseline
SAP	statistical analysis plan
SCORAD	scoring atopic dermatitis
SFU	safety follow-up
TB	tuberculosis
TBL	total bilirubin
TCI	topical calcineurin inhibitors
TCS	topical corticosteroids
ULN	upper limit of normal
US	United States
VAS	visual analogue scale
vIGA-AD™	Validated Investigator's Global Assessment for Atopic Dermatitis
vIGA-AD 0/1	Achieving a vIGA-AD score of 0 (clear) or 1 (almost clear) with ≥ 2-point reduction from baseline
vIGA-AD 0	Achieving a vIGA-AD score of 0 (clear) with ≥ 2-point reduction from baseline
vIGA-AD 1	Achieving a vIGA-AD score of 1 (almost clear) with ≥ 2-point reduction from baseline
WOCF	Worst Observation Carried Forward
WPAI	Work Productivity and Activity Impairment

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 amendment 6 for study 20210142, Rocatinlimab (AMG 451) dated 26 February 2024. The scope of this plan includes 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/Estimands and Hypotheses

2.1 Objectives and Endpoints/Estimands

This table provides the objectives and endpoints for the global protocol. The [redacted] that aligns with the global protocol and this endpoint table for [redacted] is provided in the protocol [redacted] Following advice from the US FDA, an [redacted] scheme is prespecified for the [redacted] in [redacted] This [redacted] scheme incorporates a different endpoint definition for the Investigator's Global Assessment (IGA) as the first endpoint that will be tested in the [redacted], which is called "Revised Investigator's Global Assessment for Atopic Dermatitis (rIGA™) 0/1 Response" (see definition in Section 5.1.1). The vIGA-AD™ 0/1 global primary endpoint will be tested after the rIGA 0/1 and EASI 75 endpoints as a key secondary endpoint for the US submission.

Objectives	Endpoints
Coprimary	
• To evaluate the efficacy of rocatinlimab [redacted] mg and [redacted] mg compared with placebo at week 24, assessed using the Validated Investigator's Global Assessment for Atopic Dermatitis (vIGA-AD)	• Achieving a vIGA-AD score of 0 (clear) or 1 (almost clear) with ≥ 2-point reduction from baseline (vIGA-AD 0/1) at week 24
• To evaluate the efficacy of rocatinlimab [redacted] mg and [redacted] mg compared with placebo at week 24, assessed using Eczema Area and Severity Index (EASI)	• Achieving ≥ 75% reduction from baseline in EASI score (EASI 75) at week 24
Key Secondary	

Objectives	Endpoints
• To evaluate the efficacy of rocatinlimab [redacted] mg and [redacted] mg compared with placebo at week 16, assessed using EASI	• Achieving EASI 75 at week 16
• To evaluate the efficacy of rocatinlimab [redacted] mg and [redacted] mg compared with placebo at week 16, assessed using vIGA-AD	• Achieving vIGA-AD 0/1 at week 16
• To evaluate the efficacy of rocatinlimab [redacted] mg and [redacted] mg compared with placebo at week 16 and week 24 on the patient reported outcome (PRO) measure of pruritus	• Achieving [redacted] from baseline in weekly average of daily worst pruritus numeric rating scale (NRS) score at week 16 in subjects with baseline weekly average of daily worst pruritus NRS score [redacted] • Achieving [redacted] from baseline in weekly average of daily worst pruritus NRS score at week 24 in subjects with baseline weekly average of daily worst pruritus NRS score [redacted]
• To evaluate the efficacy of rocatinlimab [redacted] mg and [redacted] mg compared with placebo at week 24, assessed using EASI	• Achieving [redacted] from baseline in EASI score ([redacted]) at week 24
• To evaluate the efficacy of rocatinlimab [redacted] mg and [redacted] mg compared with placebo at week 24 on the PRO measure of atopic dermatitis AD skin pain	• Achieving [redacted] from baseline in weekly average of daily AD skin pain NRS score at week 24 in subjects with baseline weekly average of daily AD skin pain NRS score [redacted]
• To evaluate the efficacy of rocatinlimab [redacted] mg and [redacted] mg compared with placebo at week 24, assessed using vIGA-AD with an additional assessment of morphology	• Achieving vIGA-AD 1 response with presence of only barely perceptible erythema or vIGA-AD 0 response (revised Investigator's Global Assessment [rIGA] 0/1) at week 24
• To evaluate the efficacy of rocatinlimab [redacted] mg and [redacted] mg compared with placebo on [redacted]	• Achieving a [redacted]
• To evaluate the efficacy of rocatinlimab [redacted] mg and [redacted] mg compared with placebo on [redacted]	• Achieving a [redacted]

Objectives	Endpoints
[redacted]	
Other Secondary	
• To evaluate the efficacy of rocatinlimab [redacted] mg and [redacted] mg compared with placebo at week 16 and week 24 on the PRO measure of pruritus	• Change from baseline in weekly average of daily worst pruritus NRS score at week 16 • Change from baseline in weekly average of daily worst pruritus NRS score at week 24 • Change from baseline in SCORing Atopic Dermatitis (SCORAD) itch visual analogue scale (VAS) score at week 16 • Change from baseline in SCORAD itch VAS score at week 24
• To evaluate the efficacy of rocatinlimab [redacted] mg and [redacted] mg compared with placebo at week 24 on the PRO measure using the Dermatology Life Quality Index (DLQI)	• Achieving [redacted] from baseline in DLQI score at week 24 in subjects with baseline DLQI score [redacted] • Change from baseline in DLQI score at week 24
• To evaluate the efficacy of rocatinlimab [redacted] mg and [redacted] mg compared with placebo at week 24 on the PRO measure using the Patient Oriented Eczema Measure (POEM)	• Achieving [redacted] from baseline in POEM score at week 24 in subjects with baseline POEM score [redacted] • Change from baseline in POEM score at week 24
• To evaluate the efficacy of rocatinlimab [redacted] mg and [redacted] mg compared with placebo at week 16 and week 24 on the PRO measure of AD skin pain	• Achieving [redacted] from baseline in weekly average of daily AD skin pain NRS score at week 16 in subjects with baseline weekly average of daily AD skin pain NRS score [redacted] • Change from baseline in weekly average of daily AD skin pain NRS score at week 24 • Change from baseline in weekly average of daily AD skin pain NRS score at week 16 • Achieving [redacted] from baseline in weekly average of daily AD skin pain NRS score at week 24 in subjects with baseline weekly average of daily AD skin pain NRS score [redacted]

Objectives	Endpoints
	• Achieving [redacted] from baseline in weekly average of daily AD skin pain NRS score at week 16 in subjects with baseline weekly average of daily AD skin pain NRS score [redacted]
• To evaluate the efficacy of rocatinlimab [redacted] mg and [redacted] mg compared with placebo at week 24 on the PRO measure of sleep disturbance	• Change from baseline in weekly average of daily sleep disturbance NRS score at week 24
• To evaluate the efficacy of rocatinlimab [redacted] mg and [redacted] mg compared with placebo at week 24 on the PRO measure of anxiety and depression using the Hospital Anxiety and Depression Scale (HADS)	• Achieving HADS-anxiety subscale score [redacted] at week 24 in subjects with baseline HADS-anxiety subscale score [redacted] • Achieving HADS-depression subscale score [redacted] at week 24 in subjects with baseline HADS-depression subscale score [redacted] • Change from baseline in HADS-anxiety subscale score at week 24 • Change from baseline in HADS-depression subscale score at week 24
• To evaluate the efficacy of rocatinlimab [redacted] mg and [redacted] mg compared with placebo at week 24 using the SCORAD	• Achieving [redacted] from baseline in SCORAD score at week 24 in subjects with baseline SCORAD score [redacted]
Safety and Immunogenicity	
• To characterize the safety and tolerability of rocatinlimab in the proposed population	• Treatment emergent adverse events (including clinically significant changes in laboratory values and vital signs) and serious adverse events
[redacted]	

Primary Estimands

Two intercurrent events will be considered to estimate the treatment effects: initiation of rescue therapy for AD (Appendix B) and permanent discontinuation of investigational product without rescue therapy for AD. Subjects are encouraged to complete all remaining study visits and procedures with the exception of investigational product

administration and postdose monitoring procedures through week 24 after permanent discontinuation of investigational product for evaluation of endpoints.

The primary estimand for a binary endpoint is to measure the effect of rocatinlimab as assessed by the difference in response rates between each of the 2 rocatinlimab treatment groups ([redacted] mg and [redacted] mg) and placebo among randomized subjects within moderate-to-severe AD population without rescue therapy for AD, regardless of permanent discontinuation of investigational product. Subjects initiating rescue therapy for AD will be classified as nonresponders at all subsequent time points (composite strategy reflecting lack of response). For subjects who do not use rescue therapy for AD but permanently discontinue investigational product, data retrieved after permanent discontinuation of investigational product will be included as collected (treatment policy strategy).

The primary estimand for a continuous endpoint is to measure the effect of rocatinlimab as assessed by the difference in means of change or percent change from baseline value between each of the 2 rocatinlimab treatment groups ([redacted] mg and [redacted] mg) and placebo among randomized subjects within moderate-to-severe AD population without rescue therapy for AD, regardless of permanent discontinuation of investigational product. Subjects initiating rescue therapy for AD will be included in the analysis, and each subject's worst observation, including baseline value, before initiation of rescue therapy for AD will be carried forward (worst observation carried forward, WOCF) to all subsequent time points (composite strategy reflecting lack of response). For subjects who do not use rescue therapy for AD but permanently discontinue investigational product, data retrieved after permanent discontinuation of investigational product will be included as collected (treatment policy strategy).

Additional supportive estimands are defined in Section 9.5.1.

Objectives	Endpoints
[redacted]	

Objectives	Endpoints
[redacted]	

Objectives	Endpoints
[redacted]	

Objectives	Endpoints
[redacted]	
Exploratory (Pharmacodynamic and Pharmacokinetics)	
[redacted]	

2.2 Hypotheses

The primary hypothesis is that either dose of rocatinlimab [redacted] mg or [redacted] mg is superior to placebo as treatment in subjects with moderate-to-severe AD, as assessed by the coprimary endpoints, and both coprimary endpoints need to achieve statistical significance to declare study success for each rocatinlimab dose.

3. Study Overview

3.1 Study Design

This is a phase 3, 24-week, randomized, placebo-controlled, double-blind study to assess the efficacy, safety, and tolerability of rocatinlimab monotherapy in adult subjects with moderate-to-severe AD with inadequate response to topical medications.

Approximately [redacted] eligible subjects will be randomized in [redacted] ratio to 3 treatment groups for a duration of 24 weeks as follows:

- Rocatinlimab [redacted] mg every 4 weeks (Q4W) with a loading dose at week 2 (N = [redacted])
- Rocatinlimab [redacted] mg Q4W with a loading dose at week 2 (N = [redacted])

- Placebo Q4W with a loading dose at week 2 (N = [redacted])

Randomization will be stratified by baseline disease severity (moderate Validated Investigator's Global Assessment for Atopic Dermatitis [vIGA-AD score = 3] versus severe [vIGA-AD score = 4]) and geographic region (Japan versus non-Japan Asian countries versus rest of world).

[redacted]

topical phosphodiesterase 4 (PDE4) inhibitors, topical JAK inhibitors, phototherapy, or systemic therapies for AD during the 24-week treatment period will be considered use of rescue therapy in the analysis.

Subjects who permanently discontinue investigational product for any reason during the study are encouraged to complete all remaining study visits and procedures with the exception of investigational product administration and postdose monitoring procedures through week 24 for the evaluation of the coprimary endpoints.

Subjects may be eligible to enroll in the long-term maintenance Study [redacted] at week 24. A safety follow-up (SFU) visit is required [redacted] after the last dose of investigational product for all subjects who do not enroll into the long-term maintenance study. An additional SFU visit will not be required for subjects who discontinue investigational product but who complete the remaining study visits for at least [redacted] after the last dose of investigational product before or at week 24.

3.2 Sample Size

[redacted]

3.3 Adaptive Design

Not Applicable.

4. Covariates and Subgroups

4.1 Planned Covariates

Analyses of efficacy endpoints will be adjusted for the effect of the stratification factors, including baseline disease severity (vIGA-AD score of 3 [moderate] versus 4 [severe] and geographic region (Japan versus non-Japan Asian countries versus rest of world), as well as baseline value of the efficacy endpoint of interest, when applicable.

4.2 Subgroups

The following coprimary and selected key secondary endpoints

- Achieving vIGA 0/1 at week 24 (coprimary endpoint for ex-US and key secondary endpoint for US)
- Achieving rIGA 0/1 at week 24 (coprimary endpoint for US)
- Achieving EASI 75 at week 24 (coprimary endpoint for ex-US and US)

will be analyzed by but not limited to the subgroups listed below:

- Age group (18 to < 40 years, 40 to < 65 years, ≥ 65 years)
- Sex (male, female)
- Race group 1 (white, non-white)
- Race group 2 (white, Asian, black or African American, other)
- Ethnicity (Hispanic or Latino, Not Hispanic or Latino)
- Geographic region (Japan, non-Japan Asian countries, rest of world; based on stratification factor)
- Geographic region 2 (Europe, North America, Asia, Other)
- Baseline weight group (< 75 kg, ≥ 75 kg)
- Baseline body mass index (< 25 kg/m^2, ≥ 25 kg/m^2)
- Baseline vIGA-AD (moderate [vIGA-AD = 3], severe [vIGA-AD = 4]; based on observed value)
- Baseline EASI (< 30, ≥ 30; ≤ 21, > 21)
- Prior use of systemic therapy for the treatment of AD (yes, no)
- Prior discontinuation of systemic therapy for AD due to inadequate response or adverse reaction (yes, no)
- Prior use of biologic therapy or systemic JAK-inhibitor therapy for AD (yes, no)
- Prior discontinuation of biologic therapy or systemic JAK-inhibitor therapy for AD due to inadequate response or adverse reaction (yes, no)

Any level of a subgroup with fewer than 10% of the total population will be combined with the closest level in the same subgroup if the number of levels is more than 2. The p-value for the treatment-by-subgroup interaction will be calculated using a logistic

regression model including treatment, subgroup, treatment-by-subgroup interaction and stratification factor(s) with [redacted]

Some additional subgroup analyses on efficacy may be performed as deemed appropriate and necessary and will not require an amendment to the SAP.

5. Definitions

5.1 Definition of Terms Included in Study Endpoints

5.1.1 Efficacy Endpoints

Validated Investigator's Global Assessment for Atopic Dermatitis (vIGA-AD)

The vIGA-AD is a validated assessment instrument used in clinical studies to rate the severity of AD globally, based on a 5-point scale ranging from 0 to 4: Clear = 0; Almost clear = 1; Mild = 2; Moderate = 3; and Severe = 4.

vIGA-AD 0/1 Response: Achieving vIGA-AD score 0 or 1 with ≥ 2-point reduction from baseline.

vIGA-AD 0 Response: Achieving a vIGA-AD score of 0 with ≥ 2-point reduction from baseline.

vIGA-AD 1 Response: Achieving a vIGA-AD score of 1 with ≥ 2-point reduction from baseline.

Revised Investigator's Global Assessment for Atopic Dermatitis (rIGA) 0/1 Response

The Atopic Dermatitis Morphology Assessment (ADMA) is an additional assessment question of AD morphology completed for those subjects with a vIGA-AD score of '1 – Almost clear ': Barely perceptible erythema, no induration/papulation, no lichenification, and no oozing or crusting (Yes or No)?

The rIGA 0/1 response incorporates vIGA-AD 0/1 response and ADMA response, is defined as achieving vIGA-AD 1 response with presence of only barely perceptible erythema (ie, ADMA response = Yes) or vIGA-AD 0 response and ≥ 2-point reduction from baseline.

Eczema Area and Severity Index (EASI)

The EASI evaluates the extent and severity of 4 key disease signs (Erythema; Induration/Papulation; Excoriation; and Lichenification) for each of the 4 body regions: Head and neck; Trunk; Upper extremities; and Lower extremities. The extent of the disease in each body region (area score $_{region}$) is assessed based on the percentage of

involvement where 0 = 0%; 1 = 1-9%; 2 = 10-29%; 3 = 30-49%; 4 = 50-69%; 5 = 70-89%; and 6 = 90-100%. The intensity score for each sign is scored as 0 = None; 1 = Mild; 1.5 = Between Mild and Moderate; 2 = Moderate; 2.5 = Between Moderate and Severe; and 3 = Severe.

The regional severity score (severity score $_{region}$) is the sum of the 4 intensity scores from the same body region: Erythema intensity $_{region}$ + Edema/Papulation intensity $_{region}$ + Excoriation intensity $_{region}$ + Lichenification intensity $_{region}$.

The regional EASI score (EASI $_{region}$) equals to the severity score $_{region}$ x area score $_{region}$ x multiplier $_{region}$, where the multiplier is 0.1 for Head and neck, 0.3 for Trunk, 0.2 for Upper extremities, and 0.4 for Lower extremities.

The total EASI score is the sum of 4 EASI $_{region}$ scores, with a range from 0 to 72. The higher the EASI score, the more severe the atopic dermatitis.

[redacted]

Dermatology Life Quality Index (DLQI)

The DLQI is a 10-item questionnaire with a one-week recall period that assesses the health-related quality of life of adult patients suffering from skin disease measuring 6 subscales: symptoms and feelings, daily activities, leisure, work and school, personal relationships, and treatment. Responses are as follows except for the question 7: “Very much”, “A lot”, “A little”, “Not at all”, and “Not relevant”. The responses are given a value from 0 to 3: Very much = 3, A lot = 2, A little = 1, Not at all = 0, Not relevant = 0. For the question 7, Yes = 3, A lot = 2, A little = 1, Not at all = 0, Not relevant = 0. The DLQI score is the sum of the responses to the 10 items with a range from 0 to 30. The higher the score, the more quality of life is impaired.

Patient Oriented Eczema Measure (POEM)

The POEM is a 7-items questionnaire that assesses disease symptoms in children and adults. Responses are as follows: “No days”, “1-2 days”, “3-4 days”, “5-6 days”, and “Every day”. The responses are given a value from 0 to 4: No days = 0, 1-2 days = 1, 3-4 days = 2, 5-6 days = 3, Every day = 4. The POEM score is the sum of the responses to the 7 items with a range from 0 to 28. If one question is left unanswered, the response for that question is scored as 0. If more than one question is left unanswered, the questionnaire is not scored. A lower score of POEM represents a better outcome.

Hospital Anxiety and Depression Scale (HADS)

The HADS is a 14-item self-assessment questionnaire with 7 items each for anxiety and depression subscales. Each question has 4 different choices. The responses are given a value from 0 to 3. Each subscale of anxiety and depression is the sum of the responses to the 7 items within the corresponding subscales with a range from 0 to 21. A higher subscale score is indicative of a poorer state. Categorization of each subscale score is as follows: 0 to 7 as normal, 8 to 10 as mild, 11 to 14 as moderate, and 15 to 21 as severe.

Severity Scoring of Atopic Dermatitis (SCORAD)

The SCORAD is a measurement tool for assessing the severity (ie, extent, intensity) of AD using both physician and patient-reported data. The SCORAD consists of three sections (A, B and C) which are then integrated in a formula to calculate the SCORAD score.

The Extent section (A) measures the body surface area affected with handprint units: Anterior Head (range from 0 to 4.5); Posterior Head (0 to 4.5); Anterior Torso (0 to 18); Posterior Torso (0 to 18); Right Anterior Arm (0 to 4.5); Left Anterior Arm (0 to 4.5); Right Posterior Arm (0 to 4.5); Left Posterior Arm (0 to 4.5); Right Anterior Leg (0 to 9); Left Anterior Leg (0 to 9); Right Posterior Leg (0 to 9); Left Posterior Leg (0 to 9); and Genitals (0 to 1). The score for this section ranges from 0 to 100, which is assigned as "A" in the SCORAD score calculation.

The Intensity section (B) assesses the intensity of the 6 clinical signs in a representative area: Erythema, Edema/Papulation, Oozing/crust, Excoriation, Lichenification, and Dryness. Each of the signs is given a value from 0 to 3: Absence = 0; Mild = 1; Moderate = 2; and Severe = 3. The score of Intensity section ranges from 0 to 18 by summing the intensity score of the 6 clinical signs, which is assigned as "B" in the SCORAD score calculation.

The Patient-oriented section consists of itching and trouble sleeping. Each item will be evaluated by averaging the performances of the last 3 days or nights. A VAS is used, with 0 for no symptoms and 10 for most severe symptoms. The score for section C ranges from 0 to 20 by summing the 2 VAS scores, which is assigned as "C" in the SCORAD score calculation.

The SCORAD score is calculated as SCORAD = A/5 + 7 * B/2 + C. The maximal score of the SCORAD is 103. The higher the SCORAD score, the more severe the atopic dermatitis. The SCORAD score will be set to missing if either physician or patient assessments are missing.

EQ-5D-5 Level (EQ-5D-5L)

The EQ-5D-5L questionnaire comprised of a 5-dimension health status measure and a visual analogue scale (VAS). The 5-dimension health status measure evaluates: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression based on a 5-level scale: no problems = 1, slight problems = 2, moderate problems = 3, severe problems = 4, an extreme problem = 5. The VAS records the subject's self-rated

health on a vertical VAS with 100 for 'Best imaginable health state' and 0 for 'Worst imaginable health state'. A unique health state is defined by combining one level from each of the 5 dimensions in the questionnaire. Each EQ-5D-5L health state, combined with subject age (where 1 = <35, 2 = [35, 45), 3 = [45, 55), 4 = [55, 65), 5 = 65 to 100 years old) and sex, (where 1 = Male, 0 = Female), will be mapped to an equivalent EQ-5D-3L index score following the National Institute of Health and Care Excellence (NICE) methods guidance for health technology appraisal using the lookup table provided by the NICE Decision Support Unit based on an English Population Study (Hernandez et al, 2020). The derived index score ranges from -0.532 to 0.989 with higher score indicating better health state.

Work Productivity and Activity Impairment (WPAI) for AD

The WPAI for AD questionnaire consists of six questions: current employment status (Q1), missed work hours due to AD (Q2), missed work hours due to other reason (Q3), total work hours (Q4), impairment while working due to AD (Q5, with 0 for no impairment and 10 for most impairment), activity impairment due to AD (Q6, with 0 for no impairment and 10 for most impairment). Four types of scores were calculated as follows. All scores were expressed as percentages, with higher scores indicating greater impairment.

- percent work time missed due to AD (absenteeism) = Q2/(Q2 + Q4);
- percent impairment while working due to AD (presenteeism) = Q5/10;
- percent overall work impairment due to AD (work productivity loss) = Q2/(Q2 + Q4) + ((1 - Q2/(Q2 + Q4)) * (Q5/10)). For those who did not work in the past seven days (Q4 = 0 and Q5 = missing due to skipped question), the percent overall work impairment due to AD will be equal to the percent work time missed due to AD, ie, Q2/(Q2 + Q4).
- percent activity impairment due to AD = Q6/10.

5.1.2 Efficacy Endpoints Based on Daily Diary Data Collection

Daily diary NRS collection ends at week 24/EOT visit per protocol. Any data collected after week 24/EOT visit will be excluded from the analysis.

Worst Pruritus Numerical Rating Scale (NRS)

Worst Pruritus NRS ranges from 0 to 10, with 10 being the highest amount of itch. Patients report the severity of their worst itch over the past 24 hours in the electronic diary (eDiary).

Atopic Dermatitis Skin Pain NRS

AD skin pain NRS ranges from 0 to 10, with 10 being the highest amount of AD skin pain. Patients report the severity of their worst AD skin pain over the past 24 hours in the eDiary.

Sleep Disturbance NRS

The Sleep Disturbance NRS ranges from 0 to 10, with 10 being the highest amount of sleep disturbance. Patients report the quality of their sleep over the past 24 hours in the eDiary.

The 3 numerical rating scale (NRS) endpoints listed above will be analyzed as the weekly average score (Section 5.5).

5.1.3 Safety Endpoints

Treatment-emergent Adverse Event (TEAE)

Events categorized as Adverse Events (AEs) starting after the first dose of investigational product, or as determined by "Did event start before first dose of investigational product" equal to "No" or missing on the Events electronic case report form (eCRF) for AEs starting on the first dose date, and up to the End of Study (EOS) date.

Treatment-emergent Serious Adverse Events (TESAE)

Treatment-emergent adverse events indicated as serious on the Events eCRF.

Treatment-related TEAE

A treatment-related TEAE is any TEAE with the relationship flag on the Events eCRF indicating there is a reasonable possibility that the event may have been caused by investigational medicinal product per investigator assessment.

TEAE Leading to Discontinuation of Investigational Product

A TEAE leading to discontinuation of investigational product is any TEAE with an action taken with investigational product indicating “Drug withdrawn” on the Events eCRF.

[redacted]

5.2 Study Dates

Informed Consent Date

The date on which subject signs the informed consent form.

Randomization Date

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

First Investigational Product Dose Date

First investigational product dose date is the date on which a subject is administered the first dose of investigational product following randomization, as recorded on the Investigational Product Administration eCRF. The first investigational product dose may be the same day or after the randomization date.

Last Investigational Product Dose Date

Last investigational product dose date is the date on which a subject is administered the last dose of investigational product, as recorded on the Investigational Product Administration eCRF.

End of Investigational Product Administration Date

End of Investigational Product Administration for each subject is defined as the date when the decision was made to end investigational product, as recorded on the End of Investigational Product Administration eCRF page. This date will be used to determine the timing of permanent discontinuation of investigational product as an intercurrent event (Section 9.5.1).

Subject-level End of Study (EOS) Date

End of study for each subject is defined as the date the subject last completed a protocol-specified procedure. The date will be recorded on the End of Study eCRF page.

Primary Completion Date

The primary completion date for the study is the date when the last subject across all sites has completed the assessments for week 24/EOT visit of the study.

End of Study Date

The end of study date for the study is defined as the date when the last subject across all sites has completed the assessments for the EOS/SFU visit of the study.

5.3 Study Points of Reference

Baseline Assessment

Baseline assessment for the endpoint of the interest collected is defined as the last nonmissing measurement taken on or before the first dose of investigational product. In cases where baseline measurements are taken on the same day as the first dose of investigational product, it will be assumed that these measurements are taken prior to first dose of investigational product being administered except for predose vital signs, labs, PK, antibody and ECG where the collection time of the baseline assessment must not be later than the time of the first dose of investigational product administration. For subjects who are randomized but not dosed after the randomization, the baseline of the study is defined as the last nonmissing measurement prior to or on the date of randomization.

Baseline assessment for the endpoint of the interest derived from daily eDiary collection is the weekly average corresponding to the baseline interval defined in Appendix D.

Study Day 1

Study Day 1 is defined as the first investigational product dose date. For subjects who are randomized but not dosed after randomization, the Study Day 1 is defined as the date of randomization.

Study Day

Study Day is defined as the number of days from Study Day 1.
Before Study Day 1: Study Day = (Date of Interest – Date of Study Day 1)
On or after Study Day 1: Study Day = (Date of Interest – Date of Study Day 1) + 1
Therefore, the study day prior to Study Day 1 is -1.

5.4 Subject Disposition

Randomized

Individuals are considered randomized if they have been assigned a randomization number. Randomized individuals are referred to as “subjects”.

On-study

Subjects are considered on-study if they have been randomized and have not yet had their end of study visit.

Exposed to Investigational Product

Subjects are defined as exposed if they receive at least 1 dose of investigational product.

Completing Investigational Product

Subjects are defined as completing the investigational product if their primary reason for ending investigational product is recorded as “Completed” in the End of investigational product Administration eCRF.

Completing Study

Subjects are defined as completing the study if their primary reason for ending study is recorded as “Completed” in the End of Study eCRF.

5.5 Arithmetic Calculations

Duration of Investigational Product Exposure (Days)

If subject received at least 1 dose of investigational product:

Exposure Duration (Days) = (minimum [last dose date + 27, end of study (EOS) date] - first dose date + 1)

Duration of AD

The number of years from the diagnosis date (DXDT) of AD to the date informed consent is signed.

Observed Portion	Missing Portion	Duration of AD (Years)
Year, Month, Day	NA	(Informed Consent Date – DXDT+1) / 365.25
Year, Month	Day	[Year (Informed Consent Date) – Year(DXDT)] + [Month(Informed Consent Date) – Month(DXDT)] / 12 [a]
Year	Month, Day	[Year (Informed Consent Date) – Year(DXDT)] [a]

[a] If it equals 0, add 1/12 years (ie, 1 month) to avoid a disease duration of 0. If the final derived duration of AD is greater than the reported age, the duration of AD will be set to age.

Age at AD Disease Onset

Age at baseline - Duration of AD

Weekly Average Score of Daily NRS

The 3 numerical rating scales (NRS) listed Section 5.1.2 will be analyzed as the weekly average score based on the weekly analysis window in Appendix D. Weekly average score is the sum of the observed daily NRS scores divided by the number of observed daily NRS scores in each weekly window if there are at least 3 days with observed daily NRS. Otherwise, the weekly average score will be set to missing.

Another set of weekly averages requiring at least 4 days with observed daily NRS scores in each weekly window will also be derived to support a sensitivity analysis for the endpoints with respect to achieving [redacted] from baseline in weekly average of daily NRS scores.

Change From Baseline

Postbaseline value – Baseline value, as defined in Section 5.3.

If the baseline or postbaseline value is missing, the change from baseline value will be set to missing.

Percent Change From Baseline

The change from baseline divided by baseline and multiplied by 100:

(Postbaseline value – Baseline value) * 100 / Baseline value

If the baseline value is 0 and the postbaseline value is also 0, the percent change from baseline is set to 0. If the baseline value is 0 and the postbaseline value is non-zero, the percent change from baseline is set to missing.

Subject Incidence

The subject incidence for a given event is defined as the number of subjects with at least one reported occurrence of the event divided by the number of subjects (ie, number of at-risk subjects). Subjects with multiple occurrences of the same event will only be counted once.

Time to Initiation of Rescue Therapy for AD During Treatment Period

Concomitant medications and phototherapies including rescue therapy for AD will be collected during the entire study period including safety follow-up. However, any use of rescue therapy for AD during safety follow-up after Week 24 visit will not impact the efficacy assessment.

For subjects who received rescue therapy for AD during the 24-week treatment period, not including safety follow-up after Week 24 visit, time to event will be calculated as follows:

Time (days) = date of first use rescue therapy for AD –Study Day 1 date + 1

For subjects who did not receive rescue therapy for AD during the 24-week treatment period, time to censorship will be at the end of the 24-week treatment period defined as follows:

Time (days) = Week 24 assessment date– Study Day 1 date + 1 for subjects with last efficacy assessment in WEEK 24 analysis visit window (Appendix D, Table 13-1); or

Time (days) = earlier of (EOS Study Day, data cutoff – Study Day 1 date + 1, or Day 169 [target day of WEEK 24 analysis visit window]) for subjects with last efficacy assessment before WEEK 24 analysis visit window.

6. Analysis Sets

Data from the two subjects enrolled under the original 52-week protocol with Q2W treatment regimen will be excluded from all efficacy and safety analyses because the study has been re-designed with Q4W treatment regimens for 24 weeks. In addition, due to site-wide serious breach and good clinical practice(GCP) violations, data from seven subjects randomized at site 66097 will also be excluded from all efficacy and safety analyses to eliminate potential risk of obscuring the overall study results. All references of “randomized subjects” (including “full analysis set” and “safety analysis set”) hereafter exclude these nine subjects. Given the small number of subjects being excluded compared with the overall population of the study, no safety or efficacy sensitivity

analyses will be planned. Selected data from these nine subjects will be summarized or listed separately.

6.1 Full Analysis Set

The full analysis set (FAS) includes all randomized subjects. Subjects will be analyzed according to the randomized treatment. Analyses of disposition, demographic and baseline data, Important Protocol Deviations (IPDs) or efficacy endpoints will utilize this analysis set.

6.2 Safety Analysis Set

The safety analysis set includes all randomized subjects who receive at least 1 dose of investigational product. Subjects will be analyzed based on the actual treatment received, where subjects who received the highest dose of rocatinlimab will be analyzed in the corresponding rocatinlimab treatment group regardless of the randomized treatment. Analyses of safety endpoints and summary of investigational product administration will utilize this analysis set.

7. Planned Analyses

7.1 Interim Analysis and Early Stopping Guidelines

Not Applicable.

7.2 Primary Analysis

The primary analysis will be conducted when all subjects have completed the week 24/EOT visit, and the data have been entered, cleaned, and locked. The objective of the primary analysis is to assess complete efficacy and safety data through week 24 and additional safety data available only for subjects who do not enroll in the long-term maintenance study.

7.3 Final Analysis

The final analysis will be conducted at the end of the study when all subjects have entered the extension study or completed the EOS visit, and the data have been entered, cleaned, and locked. The objective of the final analysis is to assess complete safety data through EOS on all subjects over the entire study period.

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. This study will use the RAVE database, eCOA data, PK, antibody, ECG and central laboratory data. The database will be subjected to edit checks outlined in the Data Management Plan (DMP). Additional details will be provided in the DMP and Data Acquisition Requirements Specification (DARS).

8.3 Handling of Missing and Incomplete Data

Missing data could occur due to missed visits/assessments intermittently or early discontinuation from study.

Missing data in efficacy endpoints will be imputed using methods specified in Section 9.5.2.

Missing assessments in safety endpoints will not be imputed.

Any laboratory value below the lower limit of quantification or above the upper limit of the quantification will be imputed using the respective limit of quantification for analysis. Any [redacted] assessments below the lower limit of quantification will be imputed using the respective limit of quantification for analysis. For any PK concentration below the lower limit of quantification, a value of zero will be used in the analysis.

Missing and incomplete dates in AE, concomitant medication and procedures will be imputed as outlined in Appendix C.

8.4 Detection of Bias

This study has been designed to minimize potential bias by allocating treatment groups randomly, assessing endpoints and handling withdrawals without knowledge of the treatment. Other factors that may bias the results of the study include:

- Important protocol deviations likely to impact the analysis and interpretation of the efficacy endpoints
- Inadvertent breaking of the blind before formal unblinding
- Investigational product dosing non-compliance
- The timing of and reasons for early withdrawal from treatment and from study

The incidence of these factors may be assessed. Important protocol deviations will be listed and/or tabulated in the clinical study report (CSR). If necessary, the incidence of other factors will be tabulated.

Any breaking of the blind for individual subjects prior to formal unblinding of the study will be documented in the CSR.

The timing of and reasons for early withdrawal from treatment and from study will be tabulated and/or listed.

8.5 Outliers

Histograms will be examined to identify outliers in any of the continuous variables used in the analyses. Unexpected and/or unexplained values in categorical data will be identified by utilizing frequency tables.

Outliers due to data entry errors will be corrected by the study team before final database lock. The validity of any questionable values or outliers will be confirmed. Outliers or any questionable values with confirmed validity will be included in the analyses presented in this statistical analysis plan unless there is sufficient justification to exclude them. However, ad-hoc sensitivity analyses may be conducted to evaluate the influence of extreme values in the data.

8.6 Distributional Characteristics

Continuous endpoints of change or percent change from baseline value will be analyzed under normality assumption. Where appropriate, the assumptions underlying the proposed statistical methodologies will be assessed. If required, data transformation or alternative non-parametric methods of analyses will be utilized.

8.7 Validation of Statistical Analyses

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

Datasets, tables, figures, and listings will be produced and validated in accordance with SOP-430399. Standard macros will be used when available.

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 the randomized treatment. For categorical endpoints, the descriptive statistics will contain [redacted]

[redacted]

Endpoints will be assessed between rocatinlimab [redacted] mg and placebo, and rocatinlimab [redacted] mg and placebo, respectively.

[redacted]

9.2 Subject Accountability

The number and percent of subjects who were randomized, received investigational product, completed investigational product, discontinued investigational product and reasons for discontinuing, completed study, discontinued study and reasons for discontinuing will be summarized. Summary of subjects who discontinue investigational product/study due to COVID-19 control measures may be included.

Key study dates for the first subject enrolled, last subject enrolled, last subject's end of Investigational product and last subject's end of study will be presented.

The number and percent of subjects randomized will be tabulated by study site.

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.

The final IPD list will be used to produce the Summary of IPDs table and the list of subjects with IPDs.

9.4 Demographic and Baseline Characteristics

Demographic and baseline disease characteristics including medical history and prior AD treatments will be summarized using descriptive statistics by randomized treatment

group and overall using FAS. If multiple races have been reported for a subject, the subject will be categorized as multiple race.

The following demographic and baseline characteristics will be summarized but not limited to:

- Age (years)
- Sex
- Ethnicity
- Race
- Height (cm)
- Weight (kg)
- Body Mass Index (BMI, kg/m^2)
- Stratification factors (severity of disease and geographical region)
- Medical history
- Prior treatment for AD with reason for stopping
- Selected clinical outcome assessment (COA)/patient-reported outcome (PRO) at baseline

9.5 Efficacy Analyses

9.5.1 Primary Estimand and Supportive Estimands

Two intercurrent events will be considered to estimate the treatment effects: initiation of rescue therapy for AD (Appendix B) and permanent discontinuation of investigational product without rescue therapy for AD. Subjects are encouraged to complete all remaining study visits and procedures with the exception of investigational product administration and postdose monitoring procedures through week 24 after permanent discontinuation of investigational product for evaluation of endpoints.

The following primary and supportive estimands will be implemented to handle intercurrent events specifically according to the clinical questions of interest.

9.5.1.1 Primary Estimand

The primary estimand for a binary endpoint (Section 2.1) is to measure the effect of rocatinlimab as assessed by the difference in response rates between each of the 2 rocatinlimab treatment groups ([redacted] mg and [redacted] mg) and placebo among randomized subjects within moderate-to-severe AD population without rescue therapy for AD, regardless of permanent discontinuation of investigational product. Subjects initiating rescue therapy for AD will be classified as nonresponders at all subsequent time points

(composite strategy reflecting lack of response). For subjects who do not use rescue therapy for AD but permanently discontinue investigational product, data retrieved after permanent discontinuation of investigational product will be included as collected (treatment policy strategy).

The primary estimand for a continuous endpoint (Section 2.1) is to measure the effect of rocatinlimab as assessed by the difference in means of change or percent change from baseline value between each of the 2 rocatinlimab treatment groups ([redacted] mg and [redacted] mg) and placebo among randomized subjects within moderate-to-severe AD population without rescue therapy for AD, regardless of permanent discontinuation of investigational product. Subjects initiating rescue therapy for AD will be included in the analysis, and each subject's worst observation, including baseline value, before initiation of rescue therapy for AD will be carried forward (WOCF) to all subsequent time points (composite strategy reflecting lack of response). For subjects who do not use rescue therapy for AD but permanently discontinue investigational product, data retrieved after permanent discontinuation of investigational product will be included as collected (treatment policy strategy).

The primary estimand overall endorses a hybrid strategy since different strategies are implemented for the 2 intercurrent events.

9.5.1.2 Supportive Estimand 1

The supportive estimand 1 for the selected binary endpoints is to measure the effect of rocatinlimab as assessed by the difference in response rates between each of the 2 rocatinlimab treatment groups ([redacted] mg and [redacted] mg) and placebo among randomized subjects where subjects who initiate rescue therapy for AD or discontinue investigational product permanently are considered treatment failures and will be classified as nonresponders at all subsequent time points after the intercurrent events (composite strategy reflecting lack of response).

9.5.1.3 Supportive Estimand 2

The supportive estimand 2 for the selected endpoints is to measure the effect of rocatinlimab as assessed by the difference in response rates (binary endpoints), and by the difference in means of change or percent change from baseline value (continuous endpoints) between each of the 2 rocatinlimab treatment groups ([redacted] mg and [redacted] mg) and placebo among randomized subjects regardless of rescue therapy for AD or permanent discontinuation of investigational product. Data retrieved after rescue therapy

initiation or permanent discontinuation of investigational product will be included as collected (treatment policy strategy).

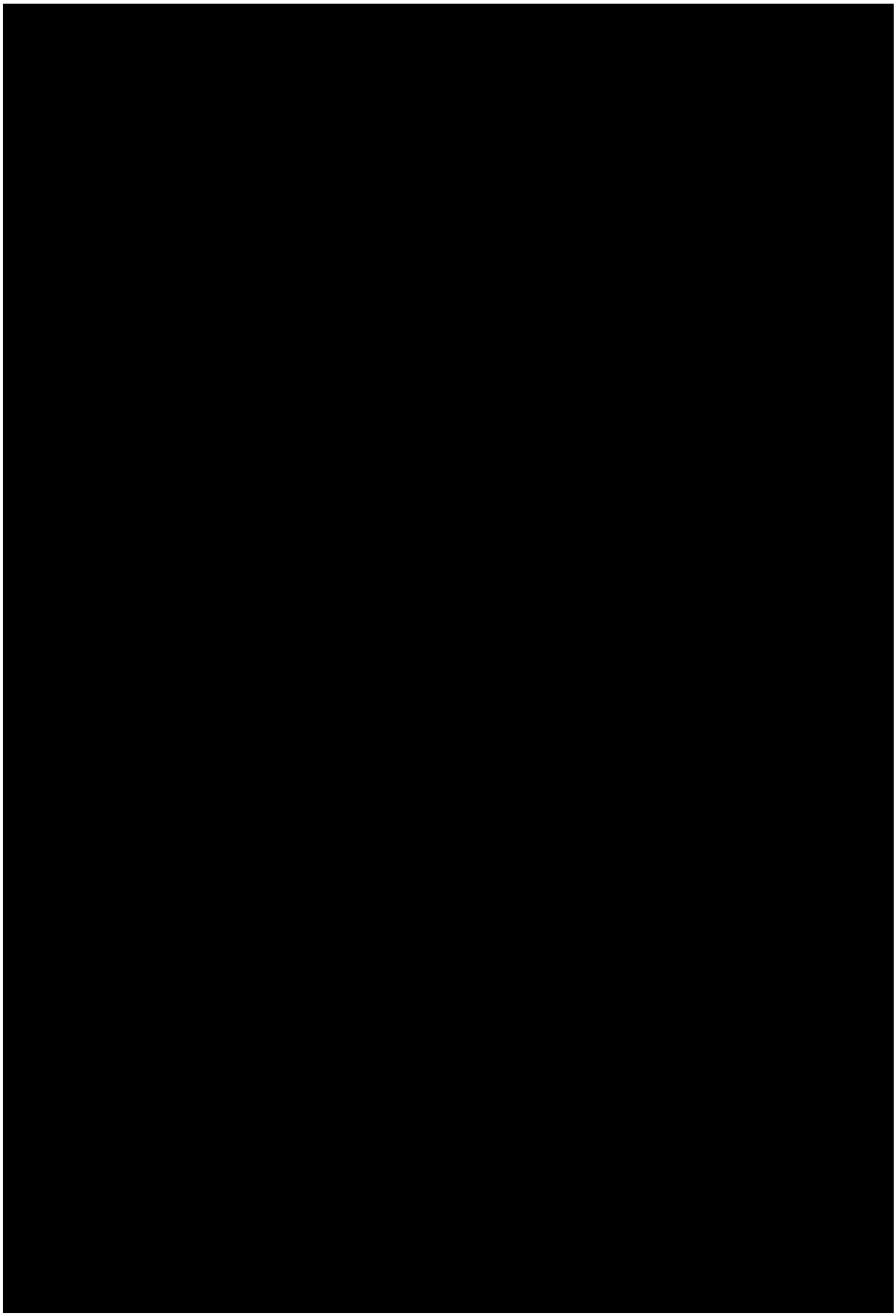
Table 9-1. Primary and Supportive Estimands

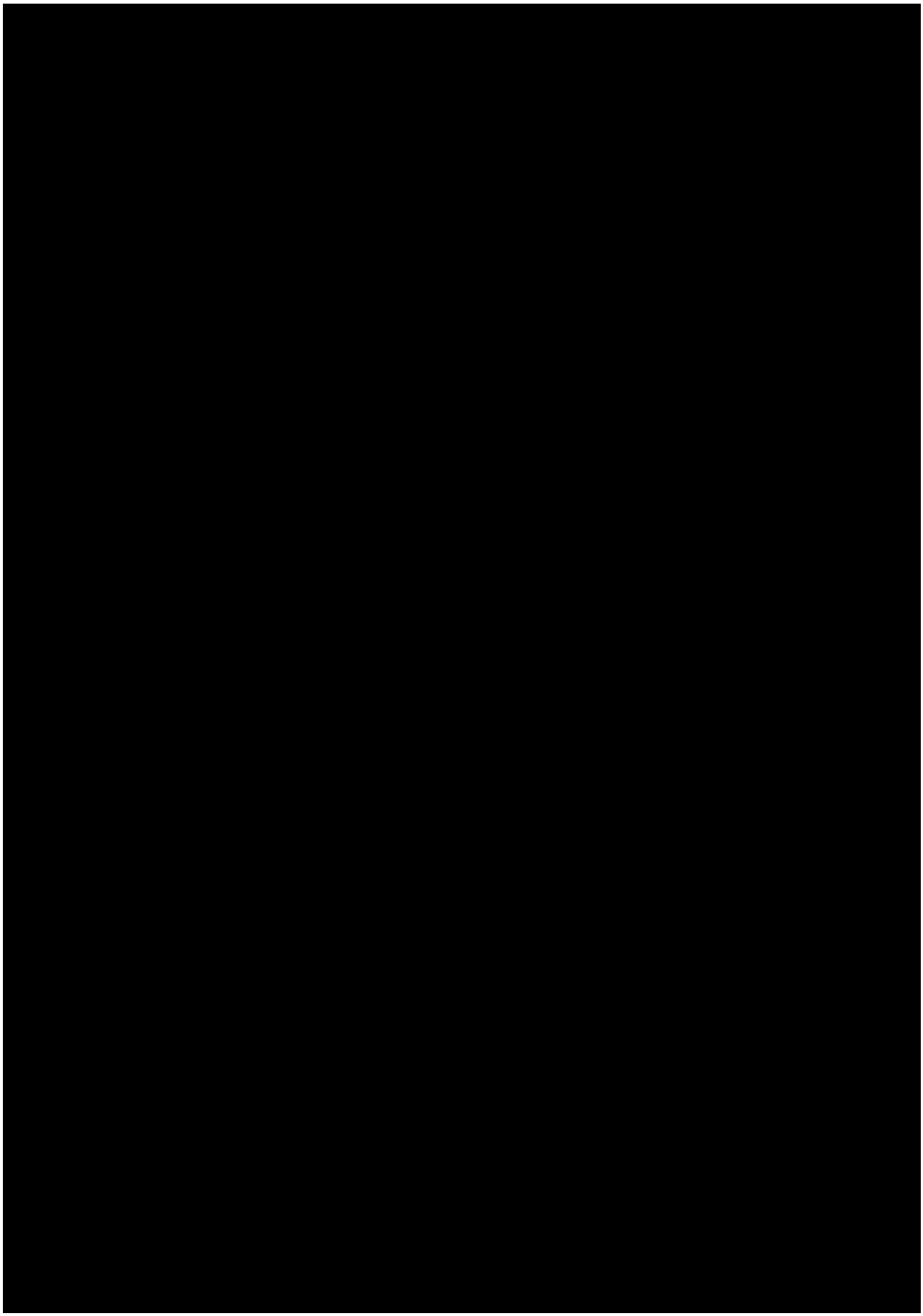
Endpoint Type	Estimand	Intercurrent Events	
		Rescue Therapy Initiation for AD	**Permanent Investigational Product Discontinuation**
Binary	Primary estimand (Hybrid)	Composite strategy: Set to Nonresponder	Treatment policy strategy: Keep observed data
	Supportive estimand 1 (Composite)	Composite strategy: Set to Nonresponder	Composite strategy: Set to Nonresponder
	Supportive estimand 2 (Treatment policy)	Treatment policy strategy: Keep observed data	Treatment policy strategy: Keep observed data
Continuous	Primary estimand (Hybrid)	Composite strategy: WOCF	Treatment policy strategy: Keep observed data
	Supportive estimand 2 (Treatment policy)	Treatment policy strategy: Keep observed data	Treatment policy strategy: Keep observed data

For composite strategy implementation in endpoints collected at each visit, assessments after the day of intercurrent event, corresponding to either rescue therapy start date or end of investigational product administration date, will be set to nonresponse for binary endpoints (or WOCF for continuous endpoints); for weekly endpoints derived from daily diaries, the week when the intercurrent event occurred and thereafter will be set to nonresponse for binary endpoints (or WOCF for continuous endpoints).

Supportive estimands are not planned for continuous endpoints as all coprimary and selected secondary endpoints under multiplicity control to support regulatory submissions are binary endpoints.

[redacted]





[redacted]

9.5.3 Analyses Methods

[redacted]

Table 9-3. Summary of Coprimary Endpoints Analysis

Endpoint	Analysis Type	Estimand	Missing data imputation	Analysis Method
vIGA-AD 0/1 at W24 (for ex-US submission) rIGA 0/1 at W24 (for US submission) EASI 75 at W24 (for ex-US and US submissions)	Primary analysis	Primary estimand (hybrid)	[redacted]	
	Sensitivity analysis	Primary estimand (hybrid)		
		Supportive estimand 1 (composite)		
		Supportive estimand 2 (treatment policy)		
	Subgroup analyses (Section 4.2)	Primary estimand (hybrid)		

9.5.5 Analysis of Key Secondary Efficacy Endpoint(s)/Estimand(s)

The primary and sensitivity analyses of key secondary endpoints are listed in Table 9-4.

Table 9-4. Summary of Key Secondary Endpoints Analysis

Endpoint	Analysis Type	Estimand	Missing data imputation	Analysis Method
• vIGA-AD 0/1 at W24 (for US submission) • EASI 75 at W16 • vIGA-AD 0/1 at W16 • [redacted] in weekly average of daily worst pruritus NRS at W24 • [redacted] in weekly average of daily worst pruritus NRS at W16 • [redacted] at W24 • [redacted] in weekly average of daily AD skin pain NRS at W24 • [redacted] • [redacted] • [redacted] in DLQI score at W24 (for ex-US submission)	Primary analysis	Primary estimand (hybrid)	[redacted]	
	Sensitivity analysis	Primary estimand (hybrid)		
		Supportive estimand 1 (composite)		
		Supportive estimand 2 (treatment policy)		

9.5.6 Analysis of Other Secondary and Exploratory Efficacy Endpoint(s)/Estimand(s)

The primary and sensitivity analysis of other secondary and exploratory endpoints are listed in Table 9-5.

Table 9-5. Summary of Other Secondary and Exploratory Endpoints Analysis

Endpoint	Analysis Type	Estimand	Missing data imputation	Analysis Method
Binary efficacy endpoints (other secondary/exploratory endpoints)	Primary analysis	Primary estimand (hybrid)	[redacted]	
	Sensitivity analysis	Supportive estimand 2 (treatment policy)		
Continuous efficacy endpoints	Primary analysis	Primary estimand (hybrid)		
	Sensitivity analysis	Supportive estimand 2 (treatment policy)		
Time to event	Primary analysis	Not applicable		

9.6 Safety Analyses

9.6.1 Analyses of Primary Safety Endpoint(s)

Not Applicable.

9.6.2 Adverse Events

The Medical Dictionary for Regulatory Activities (MedDRA) version 25.1 or later will be used to code all events categorized as adverse events to a system organ class and a preferred term.

Subject incidence of all TEAE, TESAE, TEAE leading to discontinuation of investigational product, treatment-related TEAE, and fatal TEAE will be tabulated by system organ class and by preferred term in descending order of frequency. Selected TEAE and TESAE summary tables will be tabulated by preferred term in descending order of frequency.

Summaries of all TEAE will be tabulated by system organ class, preferred term, and severity grade.

Adverse event of interest (EOI: Appendix F) will include but not be limited to infections, injection-related reactions, injection site reactions, hypersensitivity reactions, major adverse cardiovascular events (MACE), malignancy, neuropsychiatric events, aphthous ulcers, conjunctivitis, and gastrointestinal ulceration/perforation.

Subject incidence of EOI (standardized MedDRA queries and/or Amgen Medical Queries) will also be summarized according to their categories and preferred term.

9.6.3 Laboratory Test Results

The following analyses will be provided for chemistry and hematology laboratory data:

- Summary statistics over time of selected laboratory data by treatment group
- Shifts in grades of selected safety laboratory values, based on common terminology criteria for adverse events (CTCAE) grade (version 5.0), between baseline and the worst on-study value will be tabulated by treatment group, as applicable
- Summary of grades ≥ 3 laboratory toxicities for selected laboratory parameters, as applicable
- Subject incidence of suspected Hy's law cases
- Summary of liver function abnormalities for alanine aminotransferase (ALT), aspartate aminotransferase (AST), alkaline phosphatase (ALP), and total bilirubin (TBL) or the following categories:
  - ALT (> 3x upper limit of normal (ULN); > 5x ULN; > 10x ULN; > 20x ULN respectively)
  - AST (> 3x ULN; > 5x ULN; > 10x ULN; > 20x ULN respectively)
  - AST or ALT (> 3x ULN; > 5x ULN; > 10x ULN; > 20x ULN respectively)
- Total Bilirubin (> 1x ULN; > 1.5x ULN; > 2x ULN respectively) and ALP (> 1.5x ULN)

9.6.4 Vital Signs

The analyses of vital signs will include summary statistics over time by treatment group. Subject incidence of the following defined categories for systolic/diastolic blood pressure (SBP/DBP) and heart rate (HR) will be provided:

- Increase from baseline ≥ 20 mmHg in SBP with SBP > 140 mmHg
- Increase from baseline ≥ 20 mmHg in SBP with SBP ≤ 140 mmHg
- Increase from baseline ≥ 10 mmHg in DBP with DBP > 90 mmHg
- Increase from baseline ≥ 10 mmHg in DBP with DBP ≤ 90 mmHg
- Change from baseline in HR of ≥ 15 bpm (increase) or HR ≥ 120 bpm
- Change from baseline in HR of ≤ -15 bpm (decrease) or HR ≤ 50 bmp

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.

Shift in ECG diagnosis/interpretation between baseline and postbaseline will be summarized.

9.6.7 Antibody Formation

The incidence and percentage of subjects who develop [redacted] at any time will be tabulated by treatment group.

A listing of antibody results for all timepoints tested for each subject with positive antibodies at any time will be reported.

9.6.8 Exposure to Investigational Product

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

9.6.9 Exposure to Concomitant Medication

Number and percentage of subjects receiving rescue therapy will be summarized by medication of interest category for each treatment group. See Appendix B for a list of selected medications and their groupings.

9.7 Other Analyses

9.7.1 Analyses of Pharmacokinetic or Pharmacokinetic/Pharmacodynamic Endpoints

The PK samples will be analyzed only for those subjects assigned to a rocatinlimab treatment group. All PK-related tables, figures, listings and other deliverables will be generated by [redacted]

10. Changes From Protocol-specified Analyses

There are no changes from the protocol-specified analyses. Following advice from the US FDA, the coprimary endpoints for all regulatory authorities are clarified in this section.

As shown in the first stage of [redacted] the coprimary endpoints for the global protocol as well as for the ex-US submissions include the following:

- Achieving vIGA-AD 0/1 at week 24
- Achieving EASI 75 at week 24

As shown in the first stage of [redacted], the coprimary endpoints for the US submission include the following:

- Achieving rIGA 0/1 at week 24
- Achieving EASI 75 at week 24

In addition to the percent change from baseline in SCORAD score, change from baseline in SCORAD score at assessment time points will also be analyzed.

11. Literature Citations / References

National Institute of Health and Care Excellence health technology evaluations: the manual. NICE process and methods [PMG36] Updated: 31 October 2023. https://www.nice.org.uk/process/pmg36

National Institute of Health and care Excellence Decision Support Unit. Mapping EQ-5D-5L to 3L. https://www.sheffield.ac.uk/nice-dsu/methods-development/mapping-eq-5d-5l-3l

Hernandez Alava, M., Pudney, S., Wailoo, A. Estimating the relationship between EQ-5D-5L and EQ-5D-3L: results from an English Population Study. 2020; REPORT.

[redacted]

Heitjan, F., Little, R. J. A. Multiple Imputation for the Fatal Accident Reporting System. Journal of the Royal Statistical Society. 1991; Series C 40:13–29.

Schenker, N., Taylor, J. M. G. Partially Parametric Techniques for Multiple Imputation. Computational Statistics and Data Analysis 1996;22:425–446.

Ratitch B, Lipkovich I, O'Kelly M. Combining analysis results from multiply imputed categorical data. PharmaSUG 2013; Paper SP03

12. Data Not Covered by This Plan

There are no plans for GBS department to analyze or summarize the following for the Clinical Study Report:

- Quantitative ECG measurements

[redacted]

[redacted]

13. Appendices

Appendix A. Reference Values/Toxicity Grades

Laboratory toxicities are graded based on NCI Common Toxicity Criteria version 5.0, which is available at the following:

https://ctep.cancer.gov/protocoldevelopment/electronic_applications/docs/ctcae_v5_quick_reference_5x7.pdf

Appendix B. Atopic Dermatitis Medications of Interest Category and Definition

Prior and on-study atopic dermatitis therapies will be summarized based on the following categories and definitions. Rescue therapies include a subset of on-study therapies used specifically for atopic dermatitis, as indicated as for “ATOPIC DERMATITIS AND RELATED CONDITIONS” on the Concomitant Medication eCRF or Procedure eCRF (for phototherapies only). Additional preferred terms for newly approved therapies may be added without an amendment to the SAP.

Medication of Interest (MOI) Category	Definition
Topical Therapies	
Topical calcineurin inhibitors (TCI)	[redacted]
Topical corticosteroids (TCS)	

Medication of Interest (MOI) Category	Definition
Topical PDE4 inhibitors	[redacted]
Topical JAK inhibitors	
Other topical therapy	
Emollients	
Systemic Therapies	
Phototherapies	
Systemic PDE4 inhibitors	
Systemic corticosteroids	
Conventional Immunosuppressants	

Medication of Interest (MOI) Category	Definition
	[redacted]
Biologics	
Systemic JAK inhibitors	

Appendix C. Handling of Partial or Missing Start Dates of Adverse Events, Medications, and Procedures

Start Date	Imputation	Description
Missing Day	The first of the start month	Start date (yyyymm) ≠ first dose date (yyyymm) OR Start date (yyyymm) = first dose date (yyyymm) and stop date (yyyymmdd) < first dose date (yyyymmdd)
	First dose date	Start date (yyyymm) = first dose date (yyyymm) and stop date ≥ first dose date as determined based on one of the criteria in footnote [a]
Missing Day/Month	01JAN of the start year	Start date (yyyy) ≠ first dose date (yyyy) OR Start date (yyyy) = first dose date (yyyy) and stop date < first dose date as determined based on one of the criteria below: • Stop date (yyyymmdd) < first dose date (yyyymmdd); • Stop date (yyyymm) < first dose date (yyyymm) if day is missing in stop date
	First dose date	Start date (yyyy) = first dose date (yyyy) and stop date ≥ first dose date as determined based on one of the criteria in footnote [a]
Missing Day/Month/Year (Completely missing)	01JAN of the stop year	Stop date < first dose date as determined based on one of the criteria below: • Stop date (yyyymmdd) < first dose date (yyyymmdd); • Stop date (yyyymm) < first dose date (yyyymm) if day is missing in stop date; • Stop date (yyyy) < first dose date (yyyy) if month and day are missing in stop date
	First dose date	Stop date ≥ first dose date as determined based on one of the criteria in footnote [a]

[a] Criteria to determine stop date ≥ first dose date:

- Stop date is missing;

- Stop date (yyyymmdd) ≥ first dose date (yyyymmdd);
- Stop date (yyyymm) ≥ first dose date (yyyymm) if day is missing in stop date;
- Stop date (yyyy) ≥ first dose date (yyyy) if month and day are missing in stop date

Note: For subjects who were never treated (first dose date is missing), partial start dates will be set to the first day of the month if day is missing or first day of the year if month is also missing.

Derivation Rules for Medication and Procedure Start Interval

Medication and procedure start interval will be set to 'BEFORE' if the imputed start date is prior to the first dose date; else 'DOUBLE-BLIND'. For subjects who did not receive investigational product, the imputed start date will be compared with the randomization date.

Derivation Rules for Medication and Procedure End Interval

Medication and procedure end interval will be set to 'BEFORE' if the end date is prior to the first dose date; else 'DOUBLE-BLIND'. If the medication and procedure end date is partial or missing, the medication and procedure end interval will be derived based on following rule:

End Date		**Interval for End Date**
Partial: yyyymm	< first dose date yyyymm	BEFORE
	≥ first dose date yyyymm	DOUBLE-BLIND
Partial: yyyy	< first dose date yyyy	BEFORE
	≥ first dose date yyyy	DOUBLE-BLIND
Completely missing		ONGOING

For subjects who did not receive investigational product, the end date will be compared with the randomization date.

Appendix D. Analytical Windows

The actual visit for a subject may not exactly coincide with their targeted visit date. Subjects who permanently discontinue investigational product and do not complete the remaining study visits through week 24 should complete the End of Treatment (EOT) visit at the next scheduled visit following last dose of investigational product. Complete efficacy and safety assessments will be collected at the EOT visit and be reported in the nominal week 24 visit in the database. All scheduled visits per protocol including EOT visits and unscheduled visits except for safety follow-up will be mapped to the analysis visits as in Table 13-1.

By-visit data collected at the safety follow-up visit do not belong to treatment period and will not be categorized into analysis visits. Any data occurred after EOS date will not be included in the analysis except antibody data.

Table 13-1 Analysis Visit Window for Endpoints Based on Data Collection at Site Visits

Analysis Visit	Target Day	Visit Window (Study Day)
BASELINE	1	See Baseline Assessment definition in Section 5.3
WEEK [redacted]	15	2 to 22 for efficacy endpoints or Day 1 postdose to 22 for safety related endpoints
WEEK [redacted]	29	23 to 43
WEEK [redacted]	57	44 to 71
WEEK [redacted]	85	72 to 99
WEEK [redacted]	113	100 to 127
WEEK [redacted]	141	128 to 155
WEEK 24	169	≥ 156
SAFETY FOLLOW-UP	Data collected on the safety follow-up visit will be summarized in safety follow-up visit without using analysis visit window above	

If there are multiple sets of the same eCOA/PRO assessment done on the same day (eg, re-assessment is reported in web portal in addition to the original set reported in the eTablet), only one set of assessment per day for baseline and postbaseline will be selected for analysis based on the order described below:

1. Complete set will be chosen over the incomplete set (the complete status at the assessment level is defined as completing all expected questions of the assessment); for multiple incomplete sets on the same day, the one with most questions completed will be chosen.
2. The set with earliest timestamp will be chosen among remaining duplicates on the same day.

For postbaseline eCOA/PRO data, the selected assessment after step 2 will be considered as a scheduled visit if there is at least one set among the multiple sets of assessment with scheduled visit name. Otherwise, it will be considered as an unscheduled visit (ie, CPEVENT='UNSCHED').

Per protocol, all assessments will be collected prior to the dose at each visit except the vital signs collected during 2-hour postdose monitoring, which is for safety monitoring purpose and will not be included in the analysis. For measurements taken on the same day as the investigational product administration, it will be assumed that these measurements are taken prior to the investigational product administration except for predose vital signs, labs, antibody and ECG on day 1, where the collection time of the assessment must be before the time of investigational product administration as defined in Section 5.3 Baseline Assessment.

For each assessment, if more than one visit with nonmissing measurement (including the unscheduled visits, ie, CPEVENT = 'UNSCHED') fall within the same visit window during postbaseline, the following rules will be applied according to the order described below for selecting one visit per visit window for summary by visit:

1. There might be multiple scheduled visits and/or unscheduled visits in the same analysis visit window. The measurement for analysis in each analysis visit window will be selected according to the following order: EOT visit, the nominal scheduled visit corresponding to the analysis window, other scheduled visit and unscheduled visit(s).
2. If multiple unscheduled visits are mapped in the same analysis window without any scheduled visits, the unscheduled visit closest to the target day will be selected for analysis. If two unscheduled visits are equidistant from the target day, the latter visit will be selected for analysis.

Weekly Intervals for Efficacy Endpoints Derived From Daily eDiary Collection

The weekly intervals derived from daily eDiary data collection will be utilized for following efficacy endpoints :

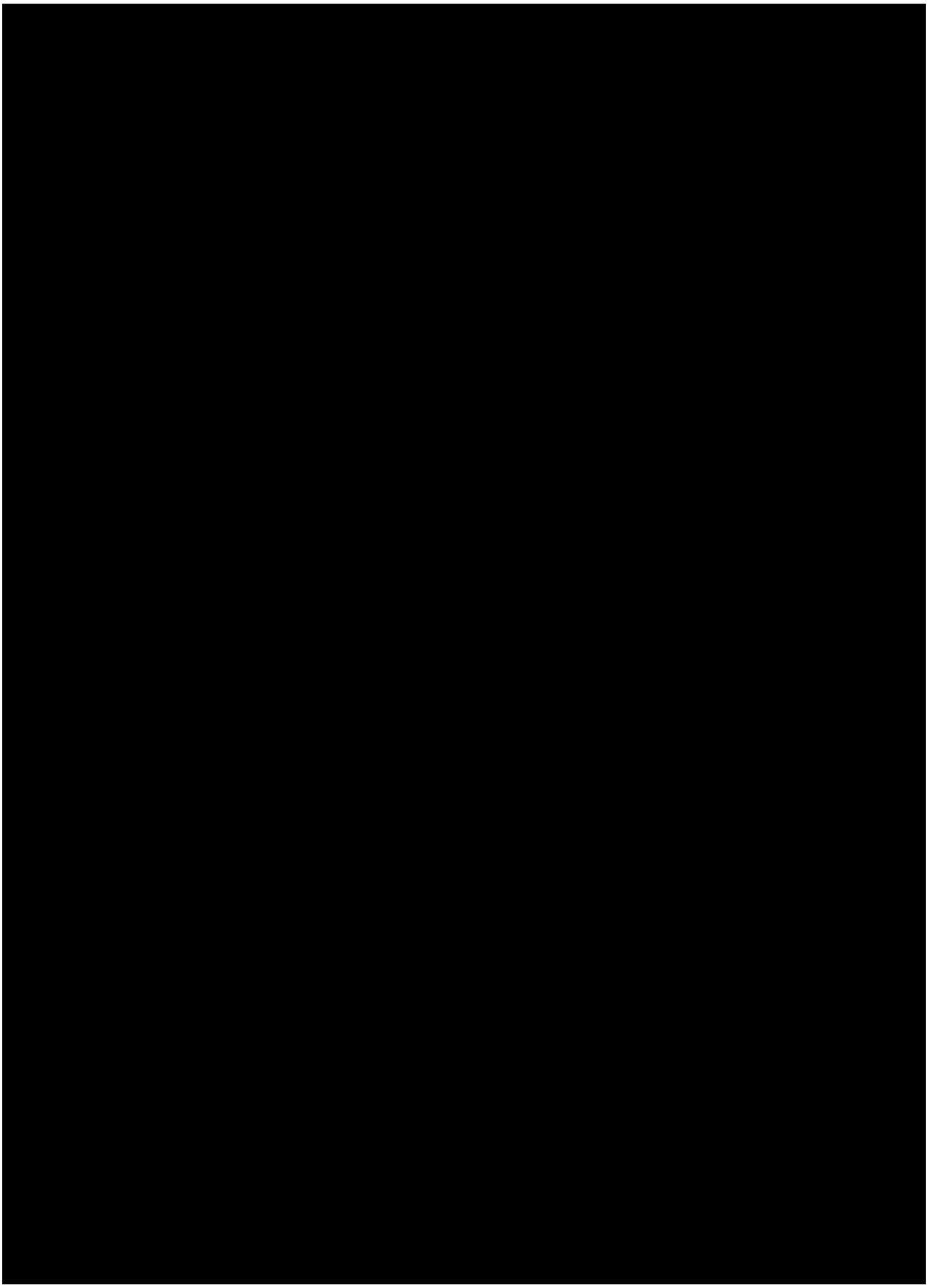
- Worst pruritus NRS
- AD skin pain NRS
- Sleep disturbance NRS

Table 13-2 Weekly Intervals for Efficacy Endpoints Derived From Daily eDiary Collection

Analysis Visit	Visit Window (Study Day)
BASELINE	7-day window including Day -6 to Day 1
WEEK [redacted]	2 to 8
WEEK [redacted]	9 to 15
WEEK [redacted]	16 to 22
WEEK [redacted]	23 to 29
WEEK [redacted]	30 to 36
WEEK [redacted]	37 to 43
WEEK [redacted]	44 to 50
WEEK [redacted]	51 to 57
WEEK [redacted]	58 to 64
WEEK [redacted]	65 to 71
WEEK [redacted]	72 to 78
WEEK [redacted]	79 to 85
WEEK [redacted]	86 to 92
WEEK [redacted]	93 to 99
WEEK [redacted]	100 to 106
WEEK [redacted]	107 to 113
WEEK [redacted]	114 to 120
WEEK [redacted]	121 to 127
WEEK [redacted]	128 to 134
WEEK [redacted]	135 to 141
WEEK [redacted]	142 to 148
WEEK [redacted]	149 to min (155, the day prior to WEEK 24 weekly visit window refer to WEEK 24 visit window definition below)

Analysis Visit	Visit Window (Study Day)
WEEK [redacted]	156 to min (162, the day prior to WEEK 24 weekly visit window refer to WEEK 24 visit window definition below)
WEEK 24	7-day window ending on WEEK 24 visit date, Otherwise, 163 to 169 if no WEEK 24 visit. Note: The earliest WEEK 24 visit day is on day 156 (Table 13-1). The corresponding 7-day weekly window can range from study day 150 to 156 inclusively. Therefore, the WEEK [redacted] and/or WEEK [redacted] weekly visit window(s) may need to be adjusted to avoid overlapping. Any daily eDiary data collected after WEEK 24 analysis visit window will be discarded.

Note: The date on which the daily assessment was reported will be used to determine the visit. The day prior to Day 1 is Day -1. Week 24 visit date is defined by the vIGA-AD assessment date selected for the analysis (AVISIT=Week 24 and ANL01FL='Y'). If two or more assessments are collected on the same day, the first assessment will be selected for the analysis.



Appendix F. Event of Interest (EOI)

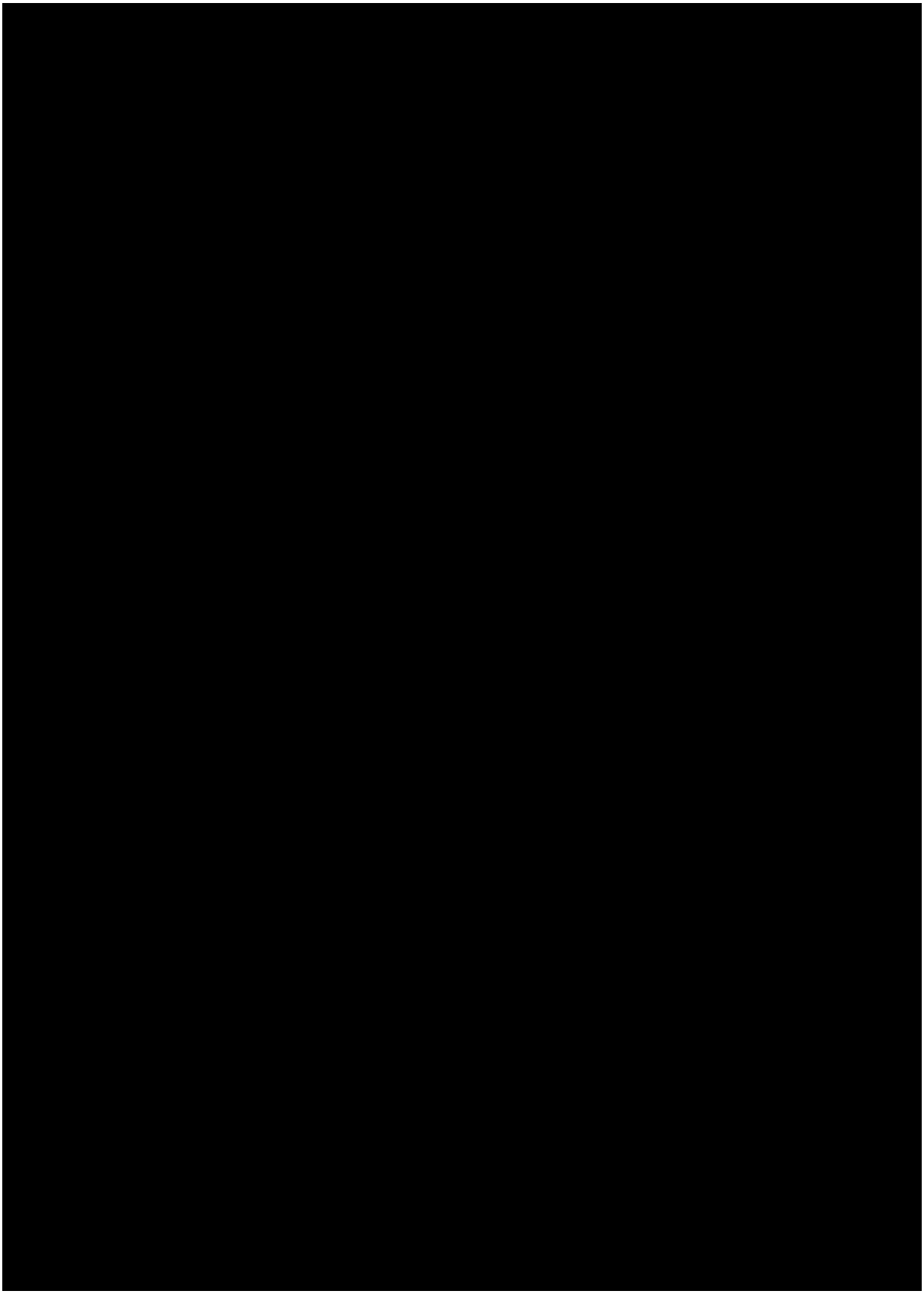
Events of interest are defined based on standardized MedDRA queries (SMQ), Amgen Medical Queries (AMQ), or event classification codes and subcodes (presented in parentheses and brackets, respectively) for positively adjudicated serious cardiovascular events as follows:

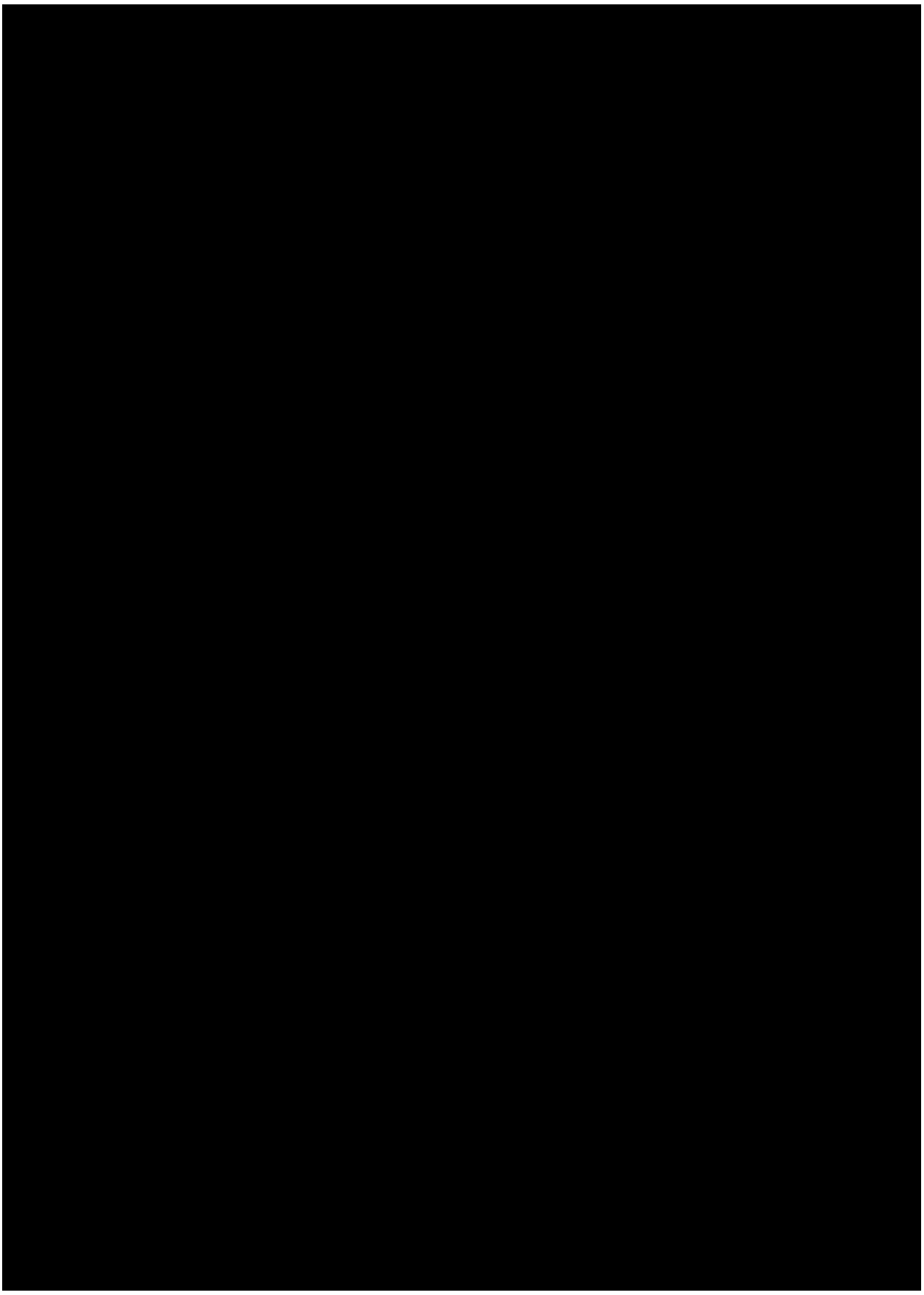
Major Category	Sub-category	Search Strategy or Definition
Injection-related Reactions (IRR)	N/A	[redacted]
Hypersensitivity	N/A	
Anaphylactic reactions	N/A	
Injection site reactions	N/A	
Infections	All infections	
	Serious infections	
	Tuberculosis	
	Opportunistic infections	
	Infections requiring systemic antibiotic therapy	
	COVID-19	
Malignancy	N/A	
Neuropsychiatric events	Depression (excluding suicide and self-injury)	
	Suicidal ideation and behavior	
	Psychosis and psychotic disorders	
Conjunctivitis	N/A	

Major Category	Sub-category	Search Strategy or Definition
Aphthous ulcers	N/A	[redacted]
Gastrointestinal ulceration, hemorrhage, or perforation	N/A	
Cardiovascular events	MACE	
	Extended MACE	
	Other cardiovascular events	

Major Category	Sub-category	
		[redacted]
	All positively adjudicated serious cardiovascular events	

Appendix G. Sample SAS Code for Efficacy Analysis





[redacted]