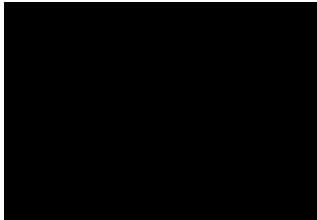
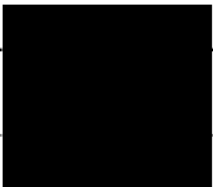
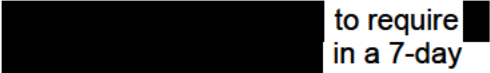
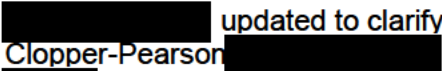
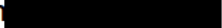
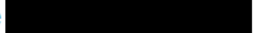


Statistical Analysis Plan

Protocol Title:	A Phase 3, Randomized, 24-week, 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-Horizon)	
Short Protocol Title:	A Phase 3, Placebo-controlled, Double-blind Study Assessing Rocatinlimab (AMG 451) Monotherapy in Moderate-to-severe AD (ROCKET-Horizon)	
Protocol Number:	20210143	
NCT Number:	NCT05651711	
Authors:		
Sponsor:	Amgen One Amgen Center Drive Thousand Oaks, CA, 91320	
SAP Date:	<u>Document Version</u> Original (v1.0) Superseding Original (v2.0) Amendment 1 (v3.0)	<u>Date</u> 

Version Number	Date (DDMMYYYY)	Summary of Changes, including rationale for changes
Original (v1.0)		Original version
Superseding Original (v2.0)		Fixed hyperlinks in Table of Contents
Amendment 1 (v3.0)		- List of abbreviations was updated the definition of rIGA 0/1.  - additional subgroup analyses of race and ethnicity, clarify the endpoints that will be analyzed by subgroups, and explain handling of subgroup when a subgroup level is too small. - Section 5.1.1 was updated the definition of rIGA 0/1. - Section 5.1.2 was updated to clarify that the daily eDiary data collection ends at week 24/EOT visit. - Section 5.3 was updated to clarify the baseline assessment of daily eDiary.  to require  in a 7-day weekly window for the weekly average calculation. The definition of time to initiation of rescue therapy for AD was updated to limit to treatment period only because rescue therapy initiated during safety follow-up after the week 24 assessment will not impact efficacy assessment. - Section 8.3 was updated to clarify how to handle values outside quantification limits in the analysis.  updated to clarify Clopper-Pearson  will be used. - Section 9.5.2.2 was updated to describe  procedure in more details.

Version Number	Date (DDMMYYYY)	Summary of Changes, including rationale for changes
		• Section 9.5.3.3 was updated to reflect the definition of time to event analysis for the 24-week treatment period only, excluding safety follow-up after week 24. • Table 9-3 was updated to clarify the submissions information for EASI75 at W24. • Table 9-4 was updated to remove NRS subgroup analysis and be consistent with Section 4.2. • Section 9.6.3 was updated to include summary of laboratory toxicity grades ≥ 3 and subject incidence of suspected Hy's Law cases. • Section 10 was updated to clarify the coprimary endpoints for ex-US and US submissions per US FDA advice. • Section 11 was updated to add new references. • Appendix B was updated to include additional atopic dermatitis medications in certain Medication of Interest categories. • Table 13-2 in Appendix D was updated to clarify the fixed baseline weekly visit window between Day -6 to Day 1, and the weekly visit window definition for Weeks 22, 23, and 24 used for weekly averages derived from daily eDiary collection. • Appendix E was updated to clarify the submission information for the scheme. • Appendix G was updated to add details about Sample SAS Code. • Editorial/changes were made throughout the document to improve overall clarity.

Table of Contents

1.	Introduction.....	10
2.	Objectives, Endpoints/Estimands and Hypotheses	10
2.1	Objectives and Endpoints/Estimands	10
2.2	Hypotheses	17
3.	Study Overview	18
3.1	Study Design.....	18
3.2	Sample Size.....	18
3.3	Adaptive Design	19
4.	Covariates and Subgroups	19
4.1	Planned Covariates.....	19
4.2	Subgroups.....	19
5.	Definitions.....	20
5.1	Definition of Terms Included in Study Endpoints.....	20
5.1.1	Efficacy Endpoints	20
5.1.2	Efficacy Endpoints Based on Daily Diary Data Collection	24
5.1.3	Safety Endpoints	24
5.2	Study Dates	25
5.3	Study Points of Reference.....	26
5.4	Subject Disposition.....	27
5.5	Arithmetic Calculations.....	27
6.	Analysis Sets	29
6.1	Full Analysis Set.....	29
6.2	Safety Analysis Set	30
6.3	Study-specific Analysis Set(s)	30
7.	Planned Analyses	30
7.1	Interim Analysis and Early Stopping Guidelines	30
7.2	Primary Analysis.....	30
7.3	Final Analysis	30
8.	Data Screening and Acceptance.....	30
8.1	General Principles	30
8.2	Data Handling and Electronic Transfer of Data.....	30
8.3	Handling of Missing and Incomplete Data	31
8.4	Detection of Bias	31
8.5	Outliers	32
8.6	Distributional Characteristics	32
8.7	Validation of Statistical Analyses	32

9.	Statistical Methods of Analysis.....	32
9.1	General Considerations.....	32
9.2	Subject Accountability	33
9.3	Important Protocol Deviations	33
9.4	Demographic and Baseline Characteristics	33
9.5	Efficacy Analyses	34
9.5.1	Primary Estimand and Supportive Estimands	34
9.5.1.1	Primary Estimand	34
9.5.1.2	Supportive Estimand 1	35
9.5.1.3	Supportive Estimand 2	35
	[REDACTED]	
9.5.3	Analysis Methods	38
	[REDACTED]	
9.5.3.3	Time to Event Analysis	39
9.5.4	Analysis of Primary Efficacy Endpoint(s)/Estimand(s).....	39
9.5.5	Analysis of Key Secondary Efficacy Endpoint(s)/Estimand(s).....	39
9.5.6	Analysis of Other Secondary and Exploratory Efficacy Endpoint(s)/Estimand(s).....	40
9.6	Safety Analyses.....	41
9.6.1	Analyses of Primary Safety Endpoint(s).....	41
9.6.2	Adverse Events	41
9.6.3	Laboratory Test Results	41
9.6.4	Vital Signs	42
9.6.5	Electrocardiogram	42
	[REDACTED]	
9.6.8	Exposure to Investigational Product	43
9.6.9	Exposure to Concomitant Medication	43
9.7	Other Analyses.....	43
9.7.1	Analyses of [REDACTED] Endpoints.....	43
10.	Changes From Protocol-specified Analyses.....	43
11.	Literature Citations / References.....	44
12.	Data Not Covered by This Plan.....	44
13.	Appendices.....	46

Appendix A. Reference Values/Toxicity Grades46

Appendix B. Atopic Dermatitis Medications of Interest Category and
Definition47

Appendix C. Handling of Partial or Missing Start Dates of Adverse
Events, Medications, and Procedures51

Appendix D. Analytical Windows.....53

Appendix F. Event of Interest (EOI).....58

Appendix G. Sample SAS Code for Efficacy Analysis61

List of Tables

Table 9-1. Primary and Supportive Estimands35

Table 9-2. Seed Values Used in [REDACTED]37

Table 9-3. Summary of Coprimary Endpoints Analysis.....39

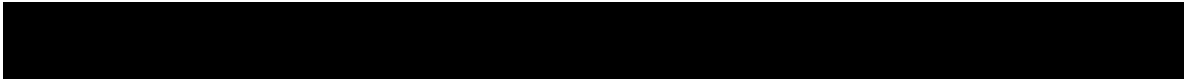
Table 9-4. Summary of Key Secondary Endpoints Analysis40

Table 9-5. Summary of Other Secondary and Exploratory Endpoints Analysis.....40

Table 13-1 Analysis Visit Window for Endpoints Based on Data Collection at
Site Visits.....53

Table 13-2 Weekly Intervals for Efficacy Endpoints Derived From Daily eDiary
Collection.....55

List of Figures



List of Abbreviations

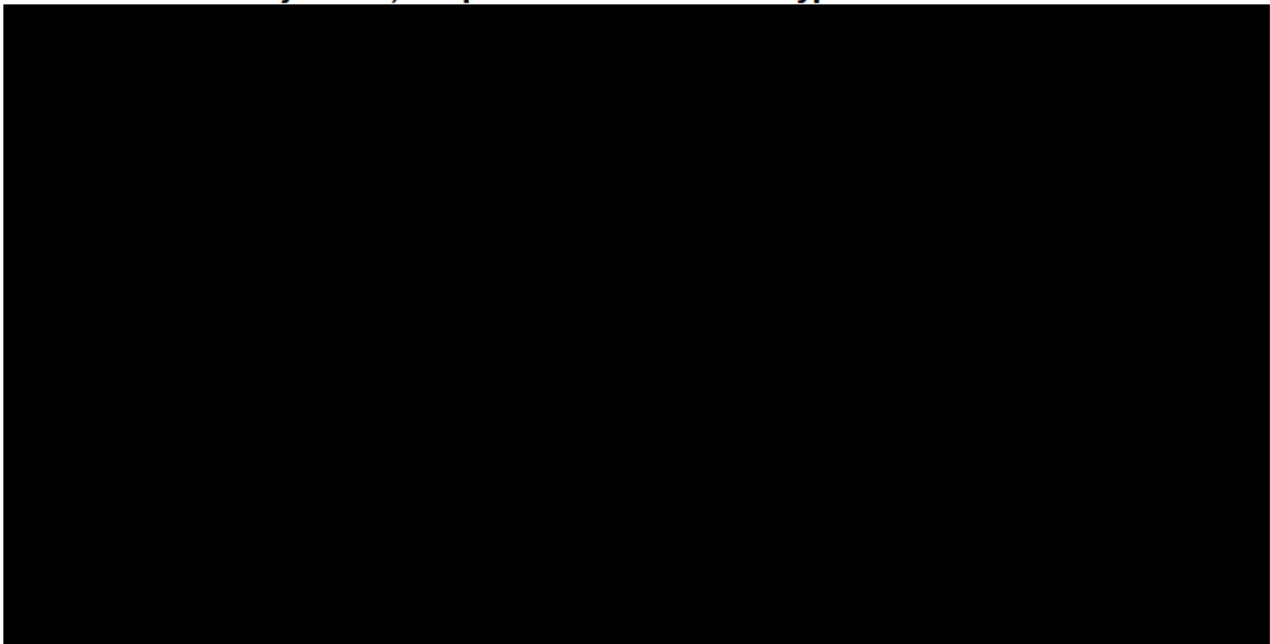
Abbreviation	Explanation
AD	atopic dermatitis
ALP	alkaline phosphatase
ALT	alanine aminotransferase
ANCOVA	analysis of covariance
AST	aspartate aminotransferase
[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]
CRF	case report form
COVID-19	Coronavirus disease 2019
DLQI	Dermatology Life Quality Index
DMC	data monitoring committee
EASI	Eczema Area and Severity Index
[REDACTED]	[REDACTED]
EASI 75	achievement of $\geq 75\%$ reduction from baseline in EASI score
EASI 90	achievement of $\geq 90\%$ reduction from baseline in EASI score
[REDACTED]	[REDACTED]
ECG	electrocardiogram
eCOA	electronic clinical outcome assessments
eCRF	electronic case report form
[REDACTED]	[REDACTED]
EOS	end of study
EOT	end of treatment
[REDACTED]	[REDACTED]
HADS	Hospital Anxiety and Depression Scale
[REDACTED]	[REDACTED]
IBG	Independent Biostatistics Group
ICF	informed consent form
IGA	Investigator's Global Assessment
IPD	Important protocol deviation
IRT	interactive response technology
IV	intravenous
JAK	janus kinase
[REDACTED]	[REDACTED]

Abbreviation	Explanation
[REDACTED]	
NRS	numeric rating scale
PDE4	phosphodiesterase type 4
[REDACTED]	[REDACTED]
POEM	Patient Oriented Eczema Measure
PRO	patient-reported outcome
Q4W	every 4 weeks
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
[REDACTED]	
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

1. Introduction

The purpose of this Statistical Analysis Plan (SAP) is to provide details of the statistical analyses that have been outlined within protocol amendment 5 for Study 20210143, Rocatinlimab (AMG 451) dated 29 November 2023. 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



the US submission.

Objectives	Endpoints
Coprimary	
To evaluate the efficacy of rocatinlimab [REDACTED] mg compared with placebo at week 24, assessed using 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 compared with placebo at week 24, assessed using Eczema Area and Severity Index (EASI)	Achieving $\geq 75\%$ reduction from baseline in EASI score (EASI 75) at week 24
Key Secondary	
To evaluate the efficacy of rocatinlimab [REDACTED] mg compared with	Achieving EASI 75 at week 16

placebo at week 16, assessed using EASI	
To evaluate the efficacy of rocatinlimab [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 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 compared with placebo at week 24, assessed using EASI	Achieving $\geq 90\%$ reduction from baseline in EASI score (EASI 90) at week 24
To evaluate the efficacy of rocatinlimab [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 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 compared with placebo on [REDACTED]	Achieving a [REDACTED]
To evaluate the efficacy of rocatinlimab [REDACTED] mg compared with placebo on [REDACTED]	Achieving a [REDACTED]
Other Secondary	

To evaluate the efficacy of rocatinlimab [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 compared with placebo at week 24 on the PRO measure of 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 compared with placebo at week 24 on the PRO measure of 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 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] - 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 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 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 < 8 at week 24 in subjects with baseline HADS-anxiety subscale score ≥ 8 - Achieving HADS-depression subscale score < 8 at week 24 in subjects with baseline HADS-depression subscale score ≥ 8 - 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 compared with placebo at week 24 on 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

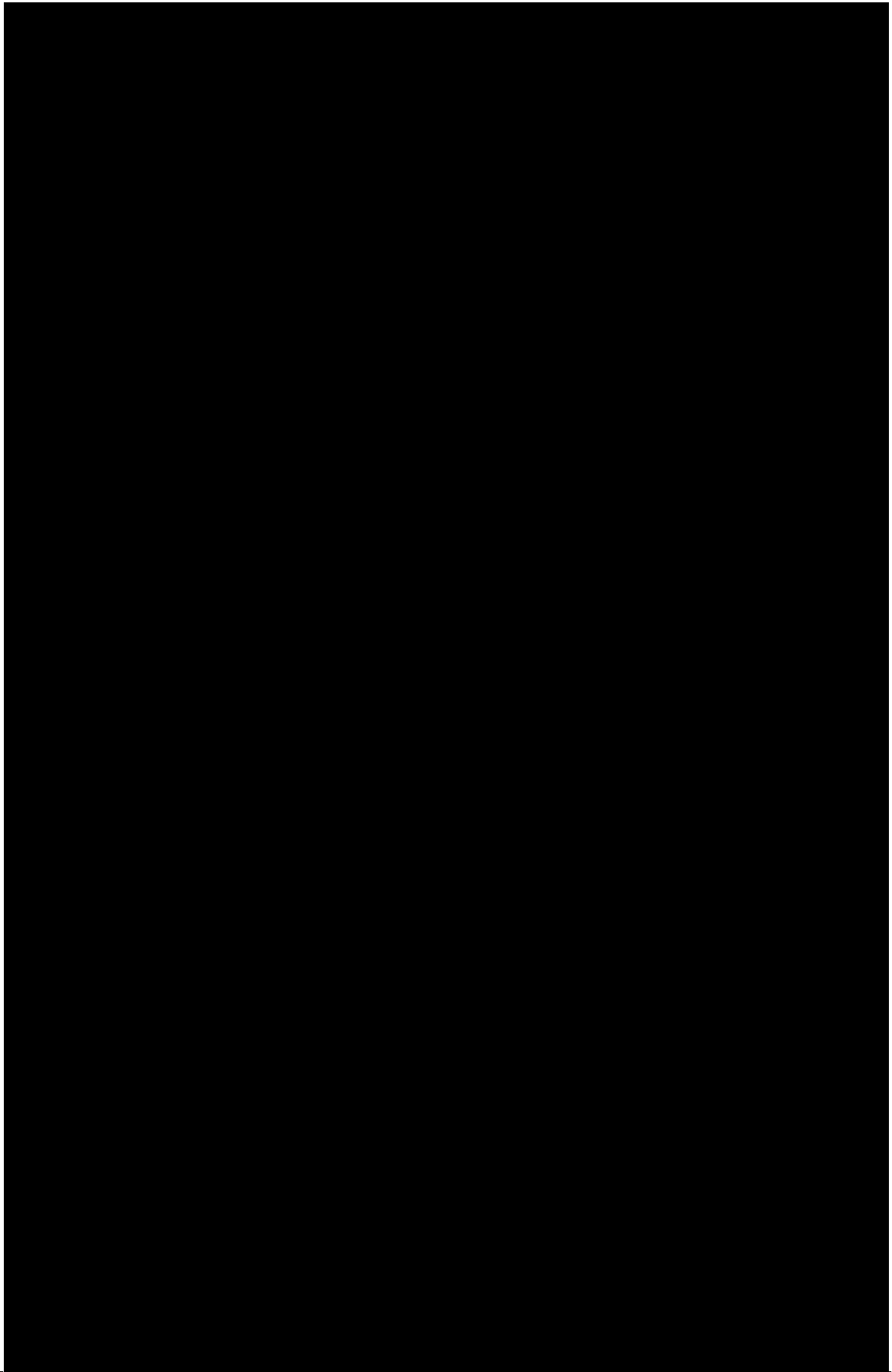
administration and postdose monitoring procedures through week 24 after permanent discontinuation of investigational product for evaluation of endpoints.

The primary estimand for a binary endpoint is to measure the effect of rocatinlimab as assessed by the difference in response rates between rocatinlimab [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 rocatinlimab [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 study treatment, 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
Exploratory	



Exploratory (	)

2.2 Hypotheses

The primary hypothesis is the rocatinlimab [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.

3. Study Overview

3.1 Study Design

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

Approximately [REDACTED] eligible subjects will be randomized in a [REDACTED] ratio to 2 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])
- 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 [REDACTED] 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] weeks 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] weeks 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 (white, non-white)**
- **ethnicity (Hispanic or Latino, Not Hispanic or Latino)**
- geographic region (Japan, non-Japan Asian countries, rest of world; based on stratification factor)
- baseline weight group (< 75 kg, ≥ 75 kg)
- baseline body mass index (< 25 kg/m², ≥ 25 kg/m²)
- baseline vIGA-AD (moderate [REDACTED] based on observed value)
- baseline EASI (< 30, ≥ 30)
- 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)
- [REDACTED]

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 Firth's method (Firth, 1993).**

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 $\geq$ 2-point reduction from baseline.

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

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

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

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.

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

██████████ 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 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.

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 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 electronic case report form (eCRF), and up to the End of Study (EOS) date.

Treatment-emergent Serious Adverse Events (TESAE)

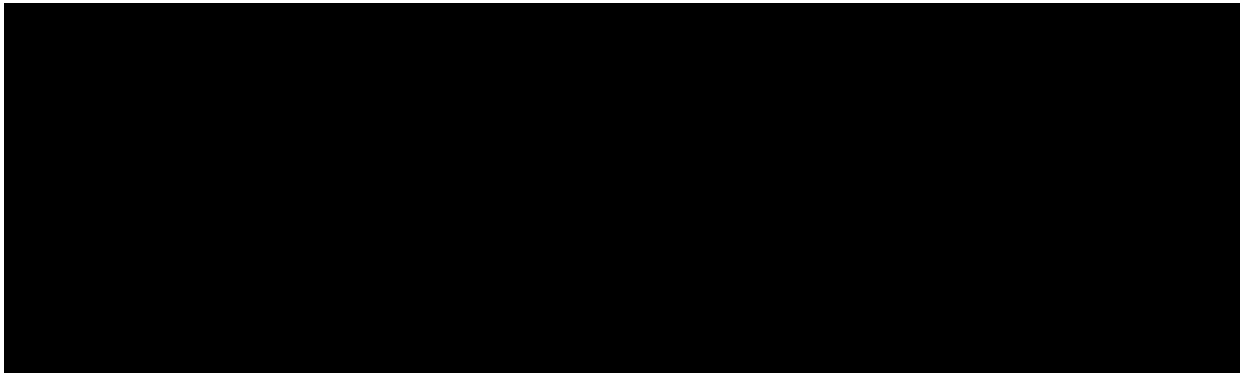
Treatment-emergent adverse events indicated as serious on the Events eCRF.

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.



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, [REDACTED] 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.

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 (i.e., 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 + 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

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 at least 1 dose of rocatinlimab will be analyzed in the rocatinlimab treatment group regardless of the randomized treatment. Analyses of safety endpoints and summary of investigational product administration will utilize this analysis set.

6.3 Study-specific Analysis Set(s)

Not Applicable

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, eDiary data, [REDACTED] 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. [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

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 frequency, percentage, and corresponding **Clopper-Pearson** 95% confidence interval (CI) where appropriate. For continuous endpoints, the descriptive statistics will include the number of observations, mean, standard error, standard deviation, median, first quartile, third quartile, minimum, and maximum.

A prespecified [REDACTED] approach (Bretz et al, 2011) will be used to assess the coprimary and selected secondary endpoints for the overall [REDACTED]

The [REDACTED] for the US submission is depicted in [Appendix E](#). For ex-US submission the [REDACTED] in protocol Section 11.11 (Appendix 11) will be applied.

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/m²)
- Stratification factors (severity of disease and geographical region)
- Medical history

████████████████████ 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 rocatinlimab █████ 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 rocatinlimab █████ 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 rocatinlimab [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 binary endpoints is to measure the effect of rocatinlimab as assessed by the difference in response rates between rocatinlimab [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).

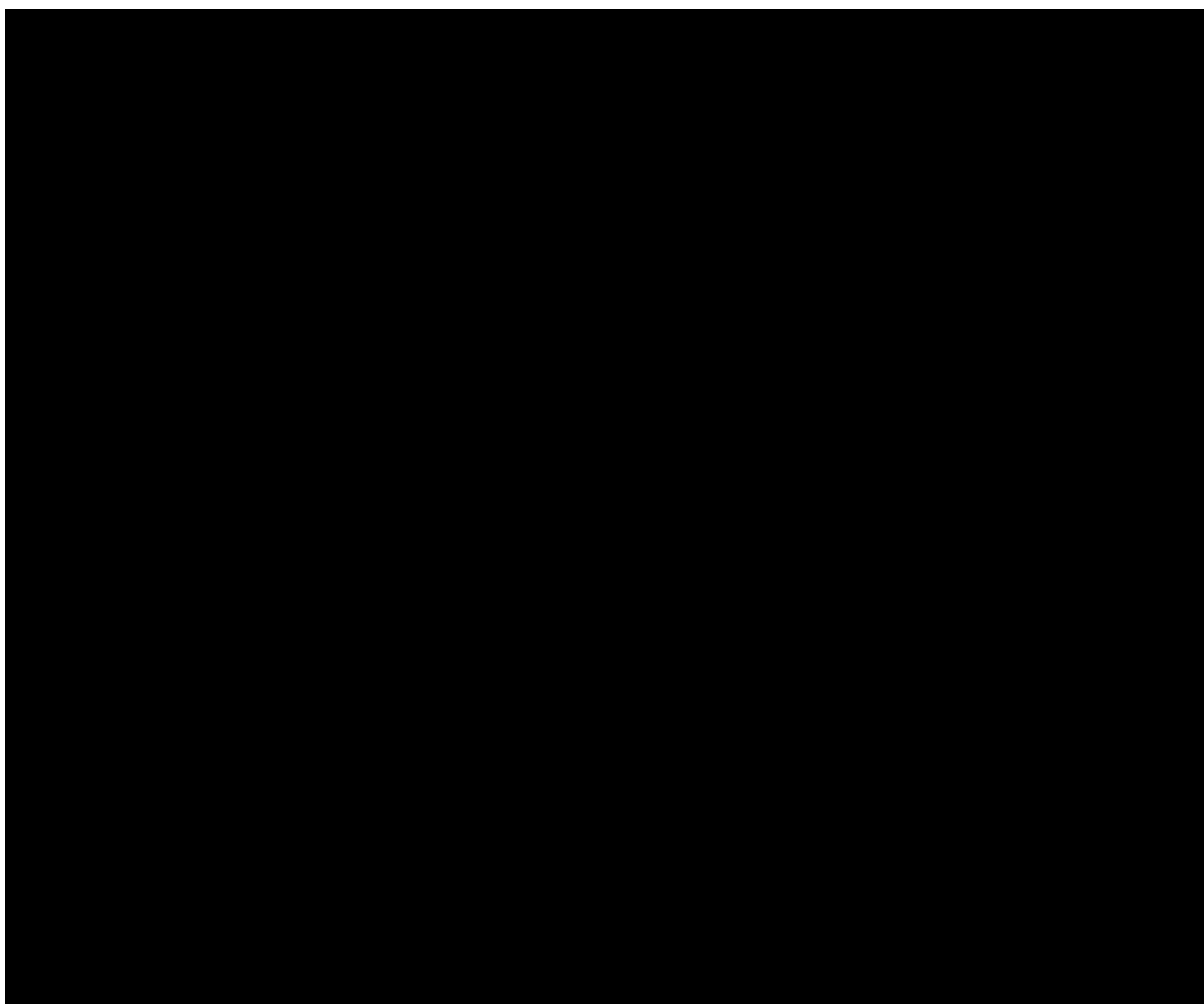
Table 9-1. Primary and Supportive Estimands

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

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.



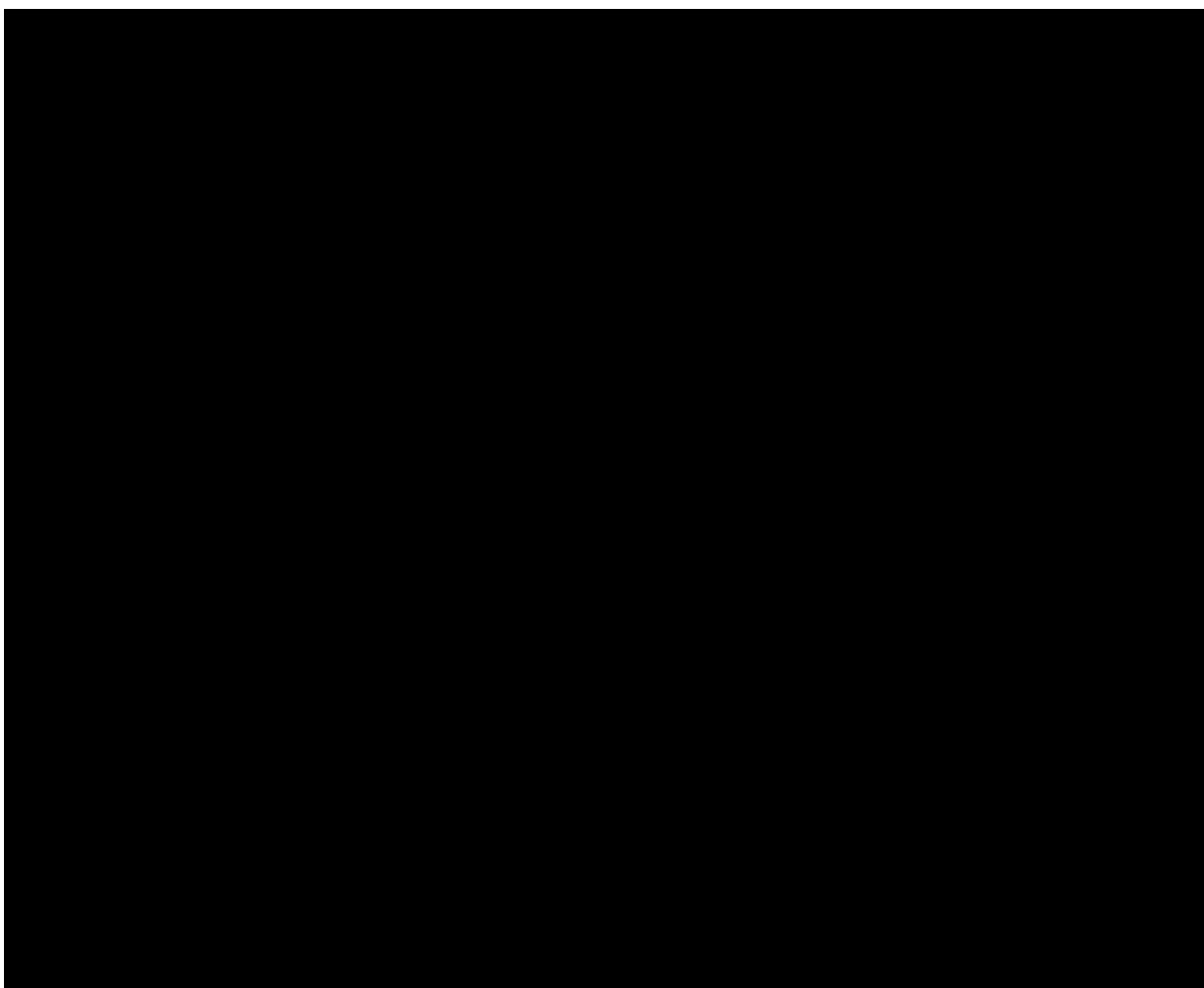
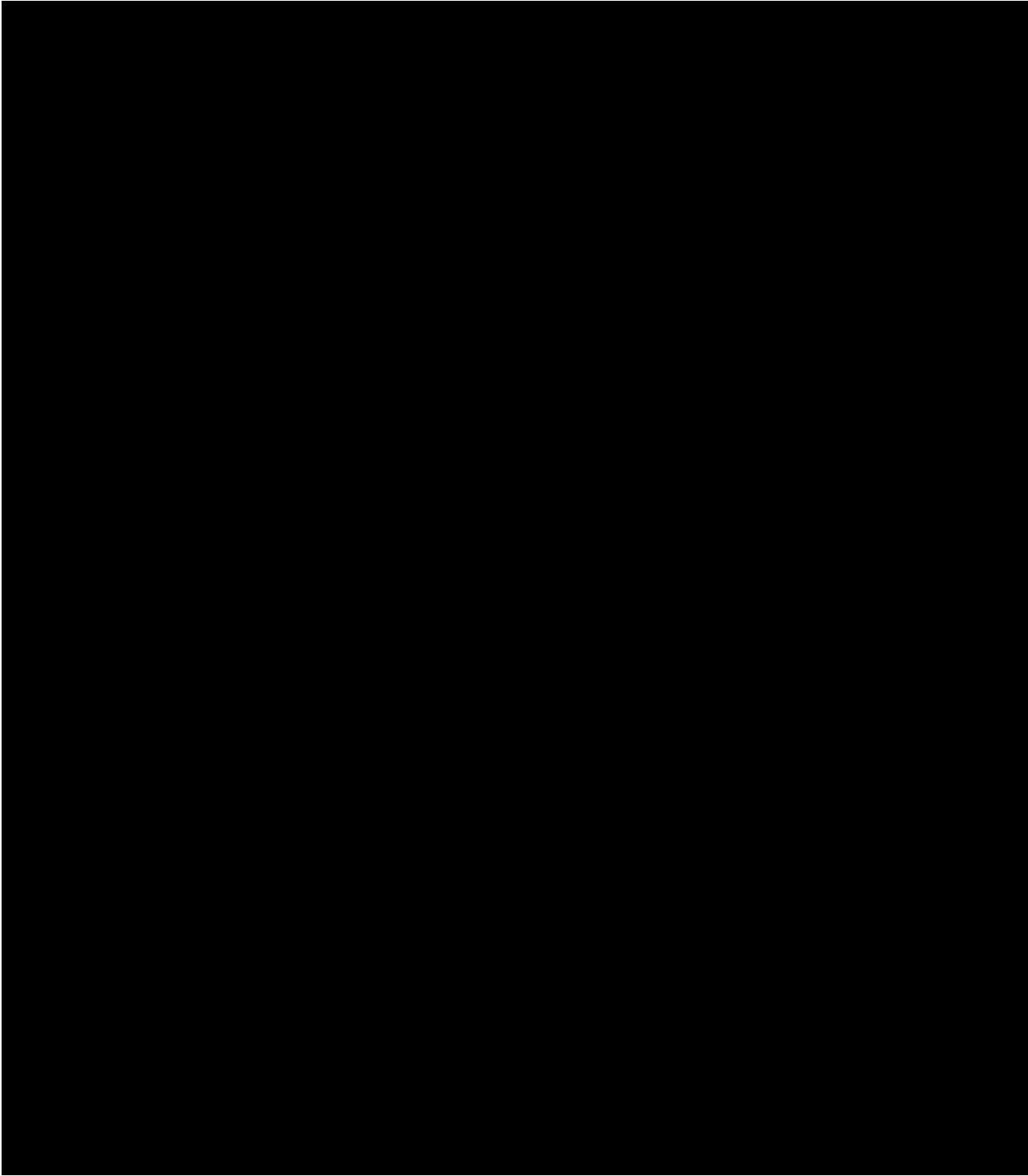


Table 9-2. Seed Values Used in [REDACTED]

Efficacy Measurement	Seed Values for [REDACTED]
viGA-AD score ^a	62496735
EASI total score	99426283
Weekly average worst pruritus NRS score	62909890
Weekly average skin pain NRS score	87172196
Weekly average sleep disturbance NRS score	84221996
Physician-facing SCORAD extent score (A) and intensity score (B); subject-facing SCORAD VAS itch score and VAS sleep score	35251010
DLQI score	68265719
POEM score	44372729
HADS anxiety/depression score	50406144
[REDACTED]	88820620
[REDACTED]	56654521

Efficacy Measurement	Seed Values for
	85840551

^a The vIGA-AD score and the corresponding AD morphologic assessment response when vIGA-AD score = 1 will be incorporated when performing [REDACTED] such that the imputed data can be used to derive both vIGA-AD 0/1 response and rIGA 0/1 response to ensure consistency.



9.5.3.3 Time to Event Analysis

[REDACTED]. Subjects who do not initiate rescue therapy will be censored at the earlier of **week 24 visit date**, EOS date, or data cutoff date. Log-rank test stratified by randomization strata will be used to generate p-value for treatment comparison.

9.5.4 Analysis of Primary Efficacy Endpoint(s)/Estimand(s)

Unless otherwise specified, efficacy analyses will be conducted on the full analysis set defined in [Section 6.1](#). The primary and sensitivity analyses of the coprimary endpoints are listed in [Table 9-3](#). Nominal p-values from the primary analysis will be provided for treatment comparison. Statistical significance will be determined based on the [REDACTED] in [Appendix E](#) of this SAP for US marketing application submission and in Protocol Amendment 5, Appendix 11.11 for ex-US marketing application submissions.

Table 9-3. Summary of Coprimary Endpoints Analysis

Endpoint	Analysis Type	Estimand	[REDACTED]	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]	[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		Analysis Method
vIGA-AD 0/1 at W24 (for US submission)	Primary analysis	Primary estimand (hybrid)		
EASI 75 at W16	Sensitivity analysis	Primary estimand (hybrid)		
vIGA-AD 0/1 at W16		Supportive estimand 1 (composite)		
in weekly average of daily worst pruritus NRS at W24		Supportive estimand 2 (treatment policy)		
in weekly average of daily worst pruritus NRS at W16				
EASI 90 at W24				
in weekly average of daily AD skin pain NRS at W24				
DLQI score at W24 (for ex-US submission)				

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

The primary analysis of other secondary and exploratory endpoints are listed in [Table 9-5](#). There is no sensitivity analysis planned for these endpoints.

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

Endpoint	Estimand		Analysis Method
Binary efficacy endpoints (other secondary/exploratory endpoints)	Primary estimand (hybrid)		
Continuous efficacy endpoints	Primary estimand (hybrid)		
Time to event	Not applicable	Not applicable	The Kaplan-Meier (KM) product limit

Endpoint	Estimand		Analysis Method
			method with log-rank test

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 [REDACTED] 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

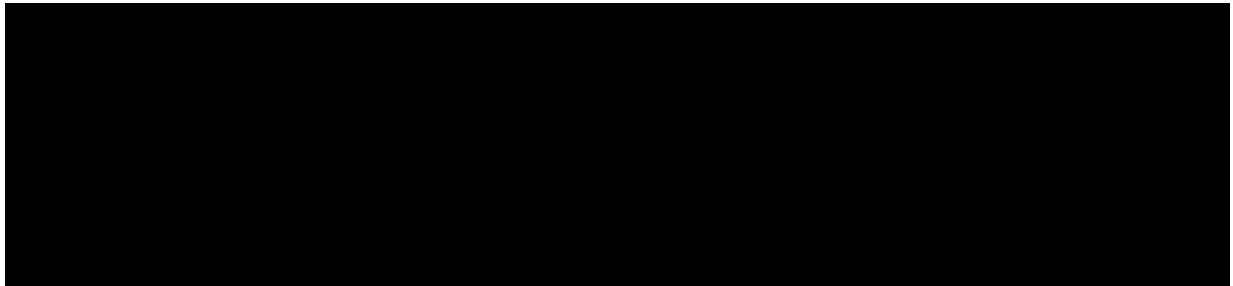
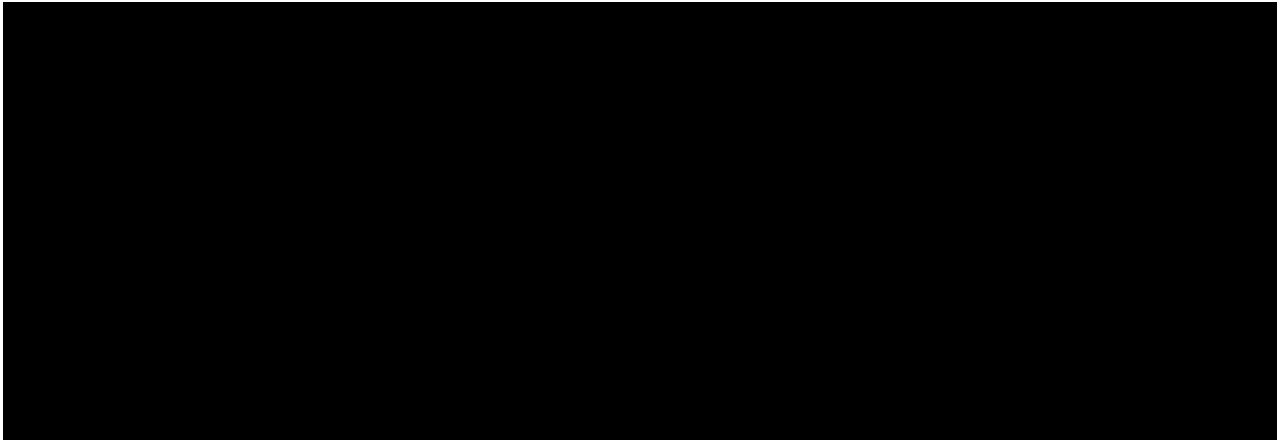
The analyses of vital signs will include summary statistics over time by treatment group. Subject incidence of the following defined categories for systolic/diastolic blood pressure (SBP/DBP) and heart rate (HR) will be provided:

- 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 bpm

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.



9.6.8 Exposure to Investigational Product

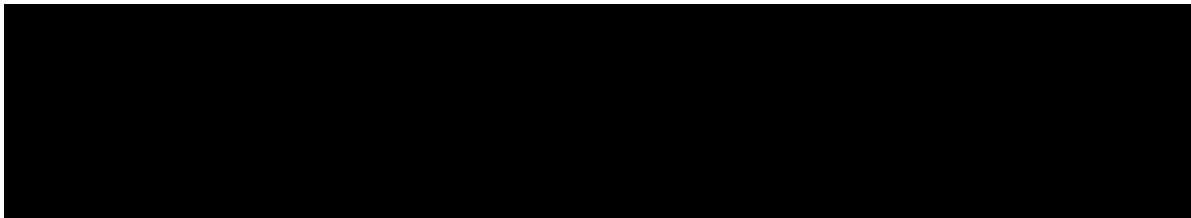
Descriptive statistics will be produced to describe the exposure to investigational product by treatment group.

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 Endpoints



11. Literature Citations / References

Bretz F, Posch M, Glimm E, Klinglmueller F, Maurer W, Rohmeyer K. Graphical approaches for multiple comparison procedures using weighted Bonferroni, Simes, or parametric tests. *Biom J.* 2011;53:894-913.

Firth D. Bias Reduction of maximum likelihood estimates. *Biometrika.* 1993;80:27–38.

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.

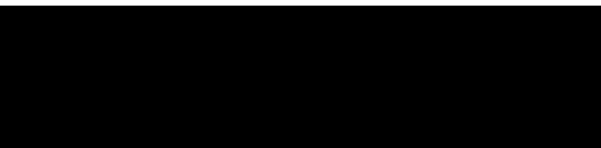
Schenker N, Taylor J. M. G. Partially parametric techniques for multiple imputation. *Computational Statistics and Data Analysis.* 1996;22:425–446.

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



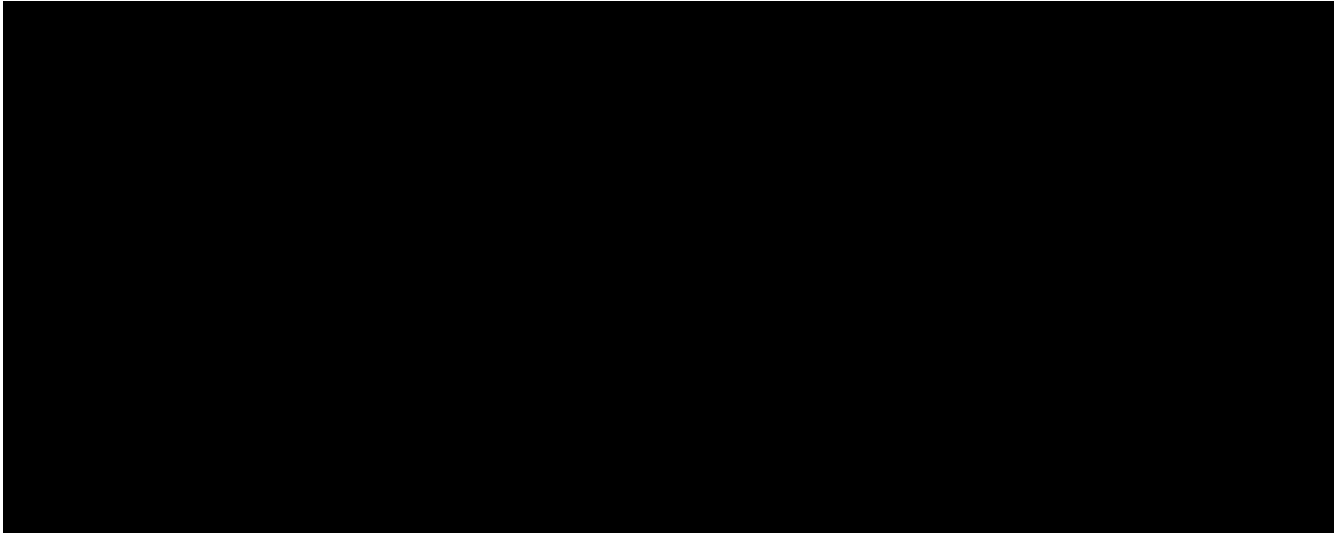


13. Appendices

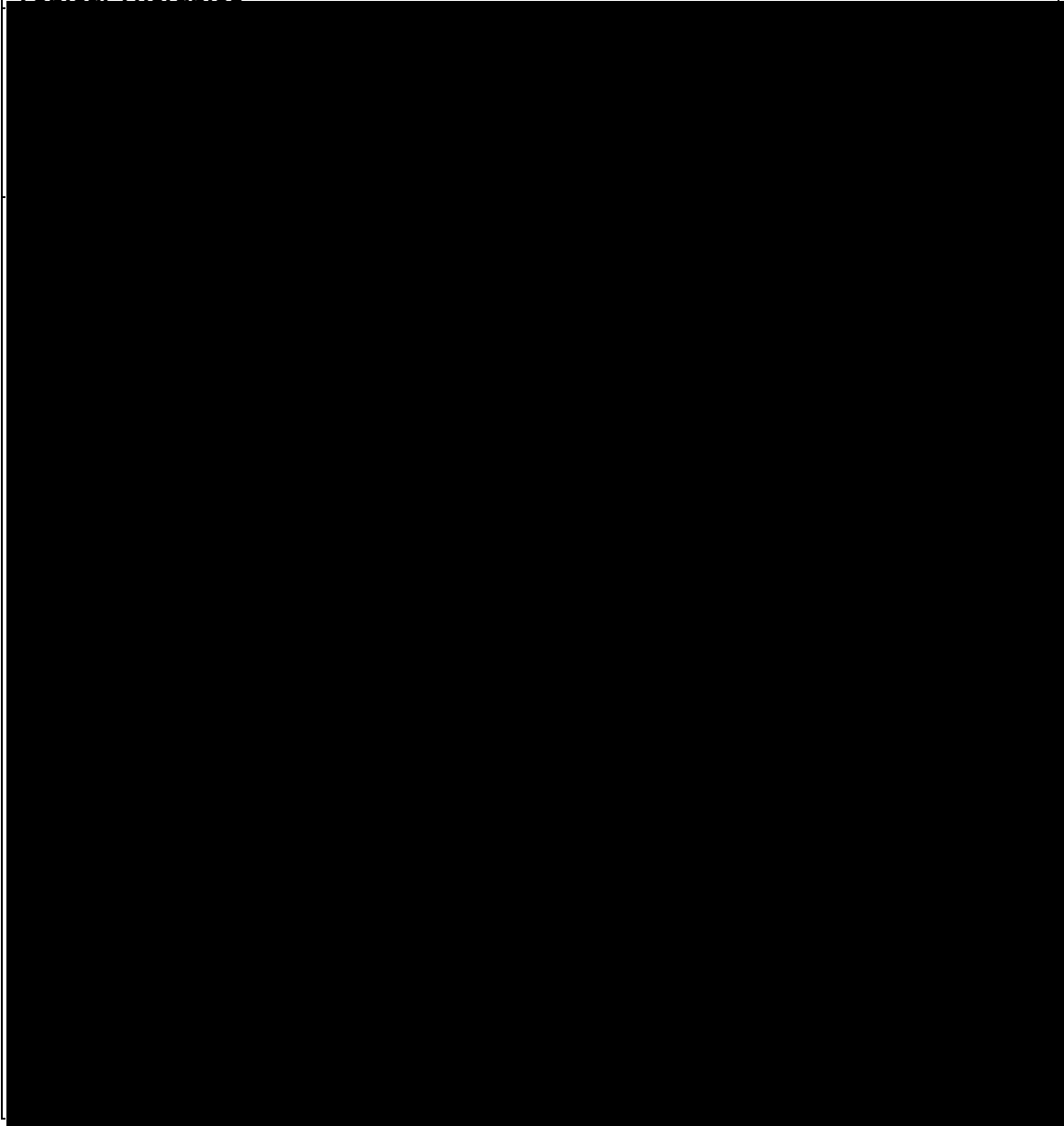
Appendix A. Reference Values/Toxicity Grades

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

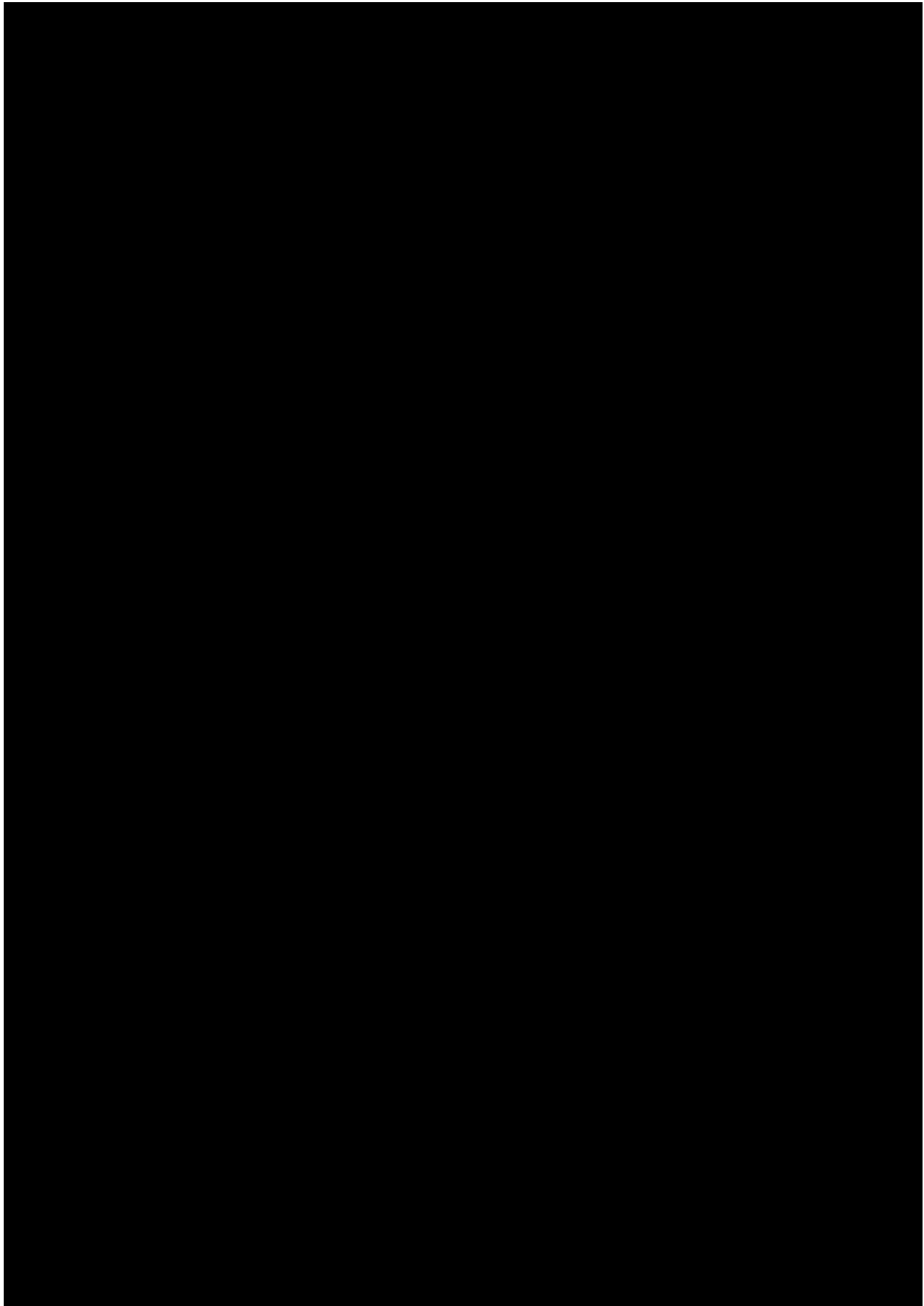


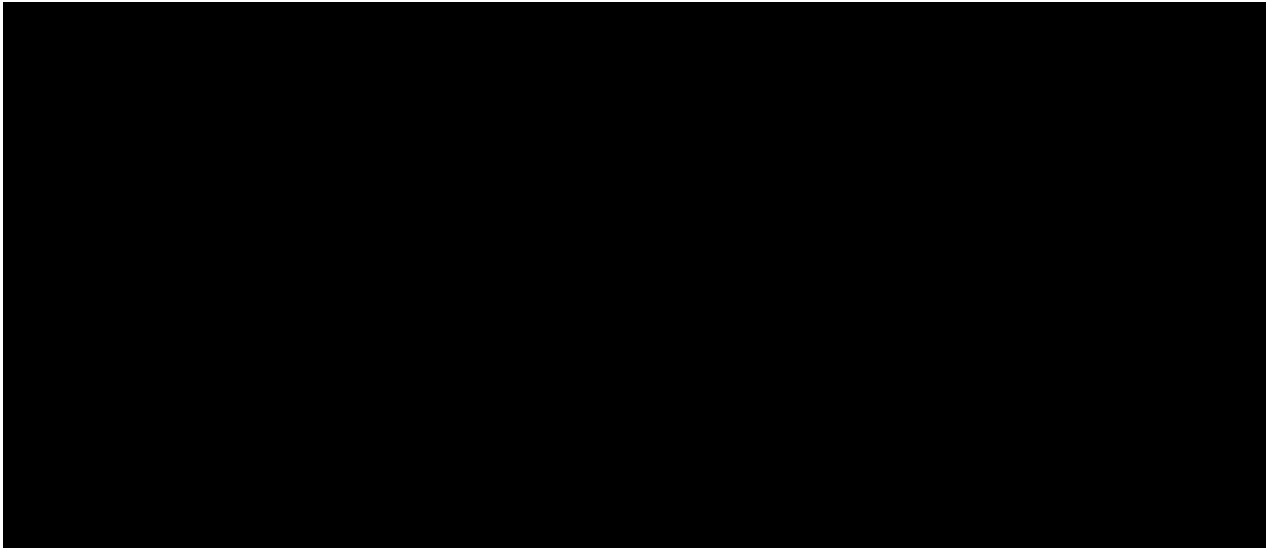
Topical Therapies



Medication of Interest (MOI) Category	Definition
--	------------

[Redacted Content]	
--------------------	--





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 < the 1 st 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

Start Date	Imputation	Description
	The 1 st dose date	Stop date $\geq$ first dose date as determined based on one of the criteria in footnote ^a

^aCriteria to determine stop date $\geq$ first dose date:

- Stop date is missing;
- Stop date (yyyymmdd) $\geq$ first dose date (yyyymmdd);
- Stop date (yyyymm) $\geq$ first dose date (yyyymm) if day is missing in stop date;
- Stop date (yyyy) $\geq$ 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 imputed start 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
	$\geq$ first dose date yyyymm	DOUBLE-BLIND
Partial: yyyy	< first dose date yyyy	BEFORE
	$\geq$ 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 [REDACTED] 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 2	15	2 to 22 for efficacy endpoints, or Day 1 postdose to 22 for safety related endpoints
WEEK 4	29	23 to 43
WEEK 8	57	44 to 71
WEEK 12	85	72 to 99
WEEK 16	113	100 to 127
WEEK 20	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, [REDACTED] 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:

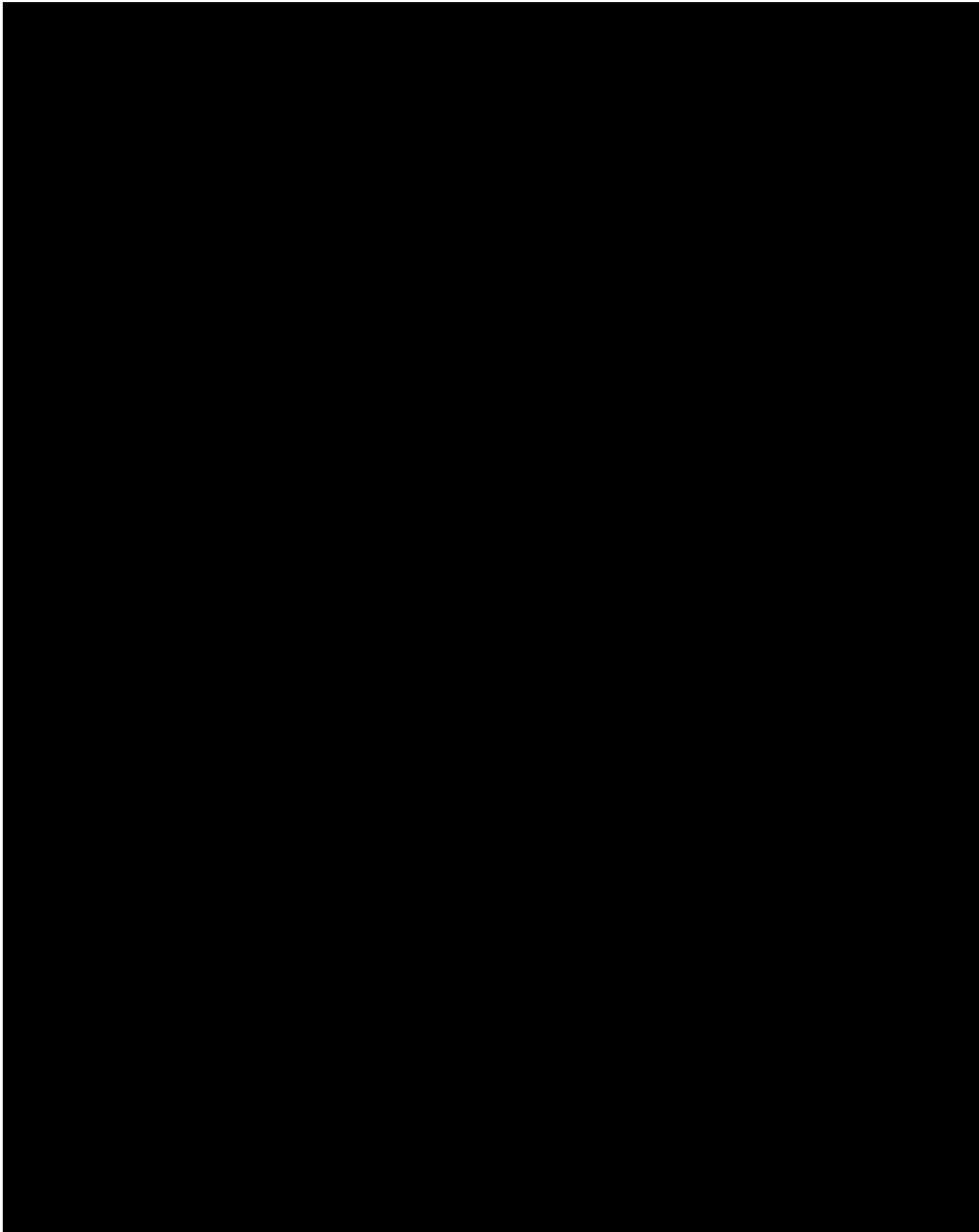
- Worst pruritus NRS
- AD skin pain NRS
- Sleep disturbance NRS

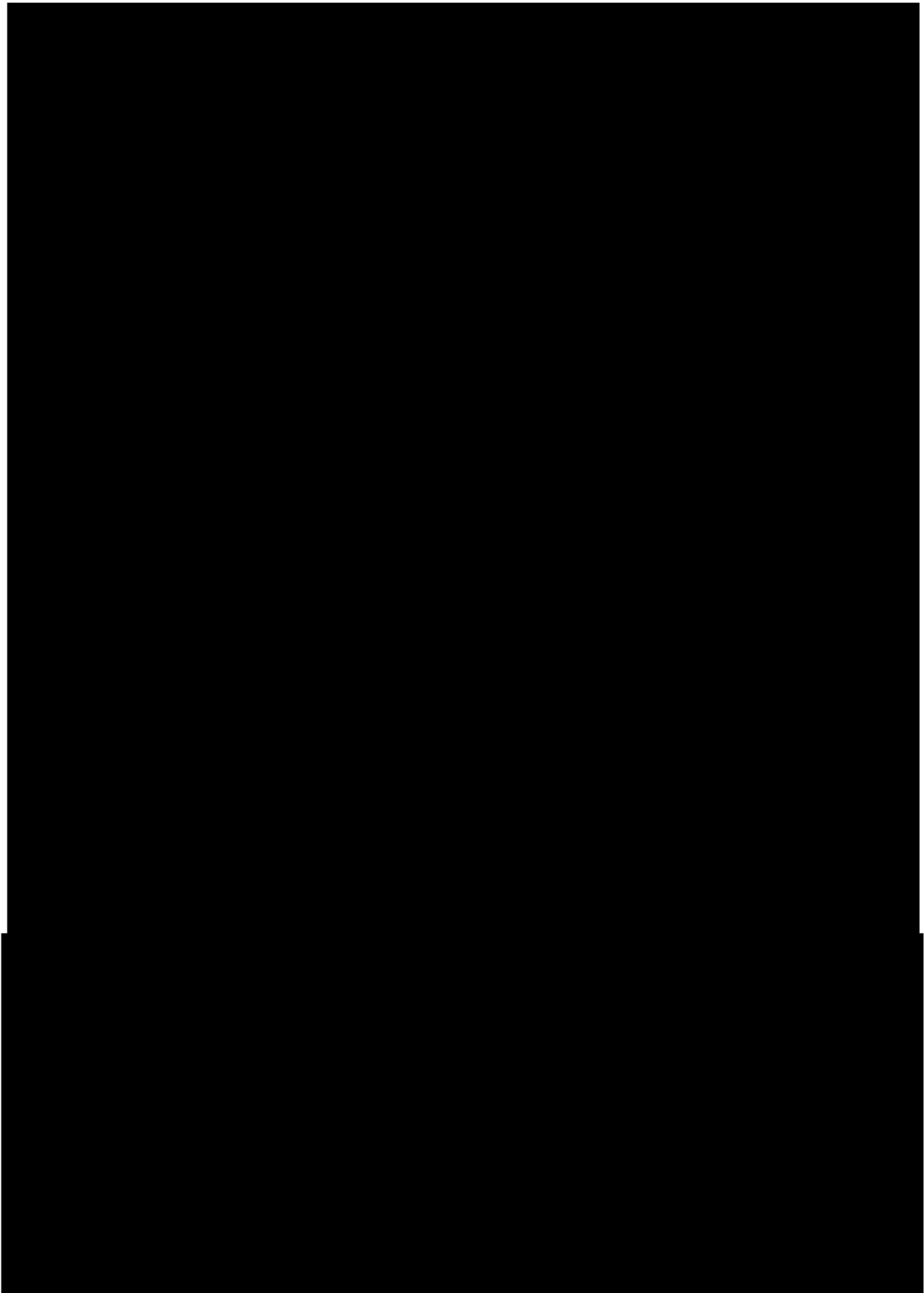
Table 13-2 Weekly Intervals for Efficacy Endpoints Derived From Daily eDiary Collection

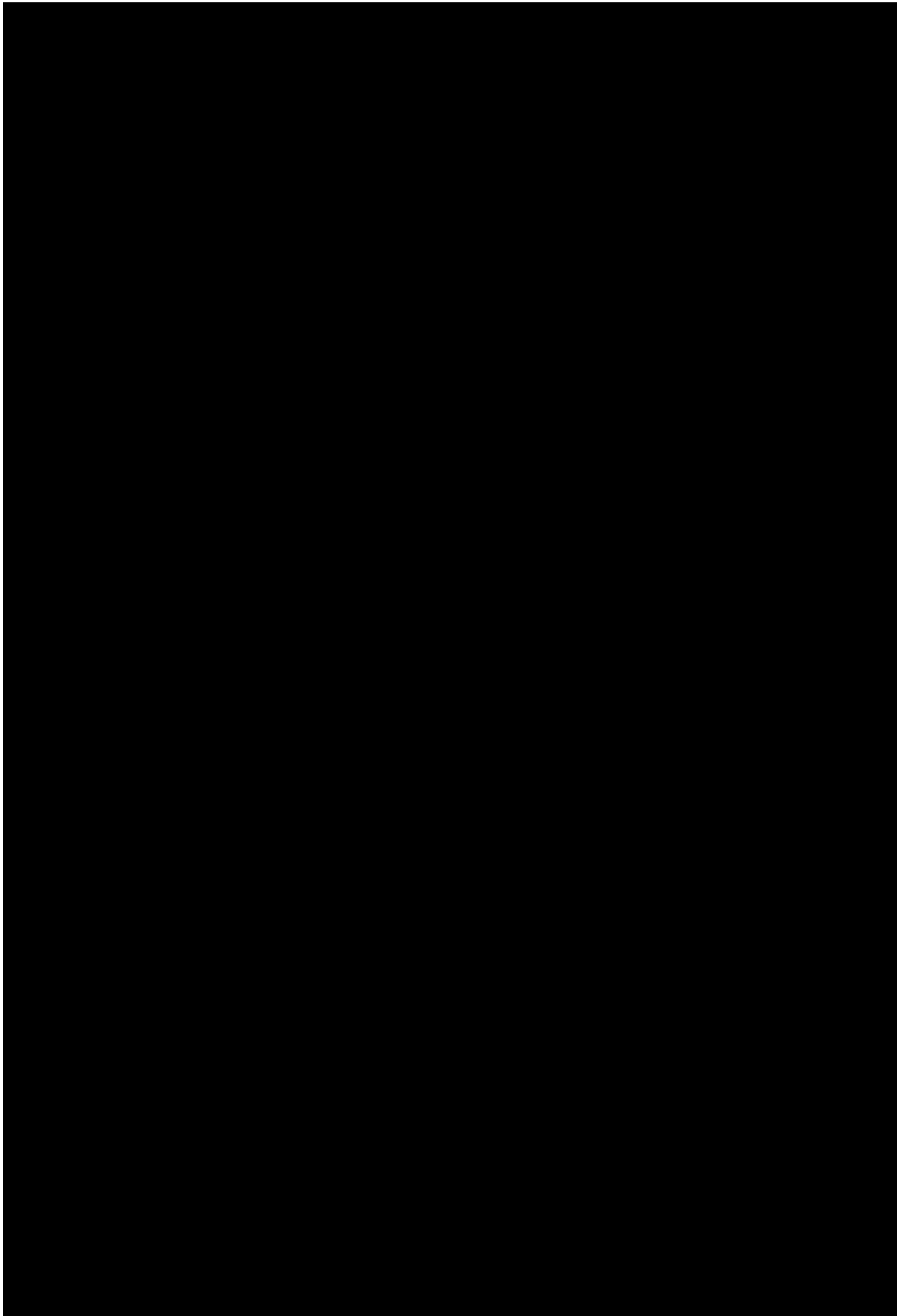
Analysis Visit	Visit Window (Study Day)
BASELINE	7-day window including Day -6 to Day 1
WEEK 1	2 to 8
WEEK 2	9 to 15
WEEK 3	16 to 22
WEEK 4	23 to 29
WEEK 5	30 to 36
WEEK 6	37 to 43
WEEK 7	44 to 50
WEEK 8	51 to 57
WEEK 9	58 to 64
WEEK 10	65 to 71
WEEK 11	72 to 78
WEEK 12	79 to 85
WEEK 13	86 to 92
WEEK 14	93 to 99
WEEK 15	100 to 106
WEEK 16	107 to 113
WEEK 17	114 to 120
WEEK 18	121 to 127
WEEK 19	128 to 134
WEEK 20	135 to 141
WEEK 21	142 to 148
WEEK 22	149 to min (155, the day prior to WEEK 24 weekly visit window refer to WEEK 24 visit window definition below)
WEEK 23	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.

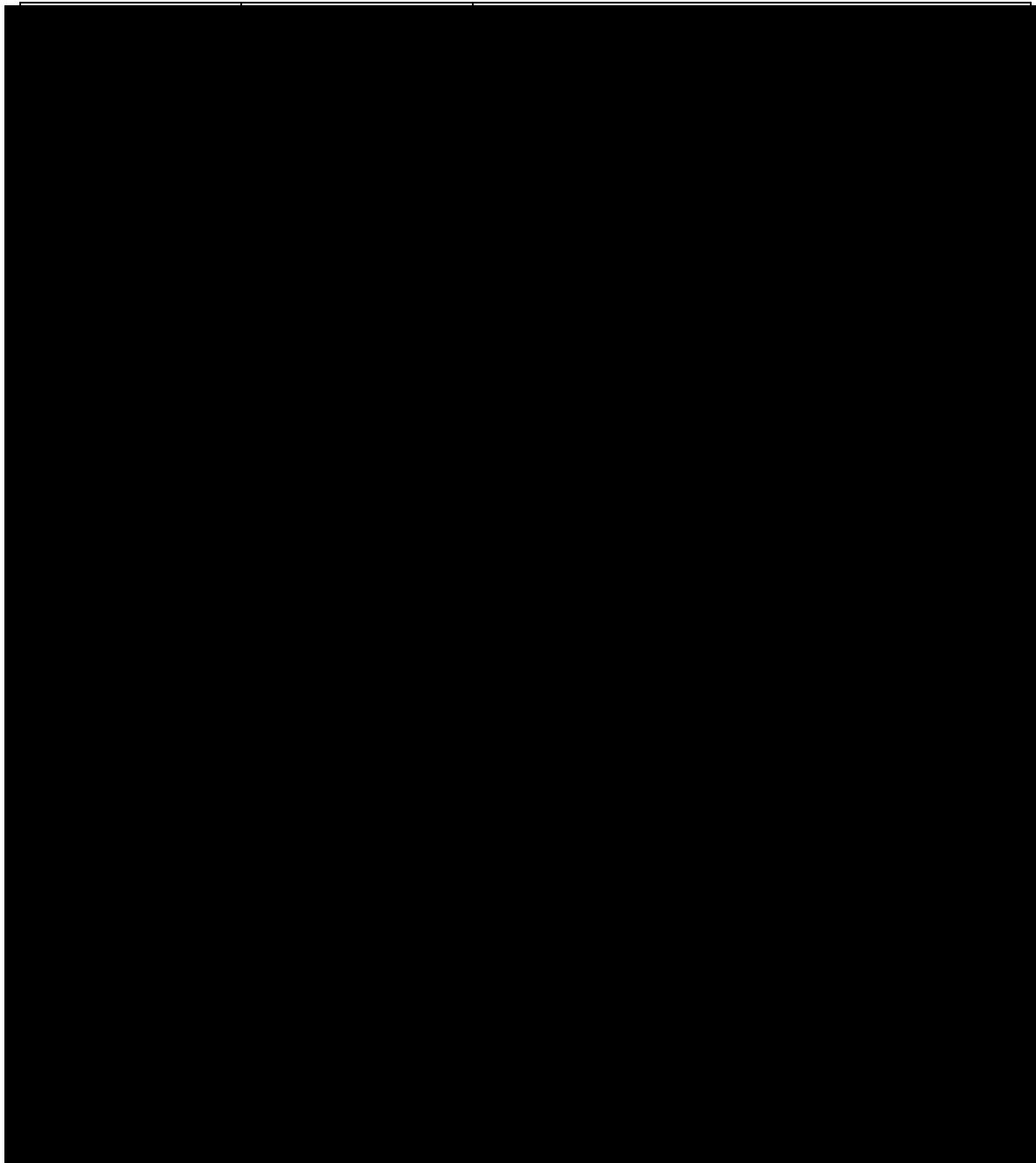
Analysis Visit	Visit Window (Study Day)
	Note: The earliest WEEK 24 visit day is on day 156 (Table 13-1). The corresponding 7-day weekly window can range from study day 150 to 156 inclusively. Therefore, the WEEK 22 and/or WEEK 23 weekly visit window(s) may need to be adjusted to avoid overlapping. Any 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.

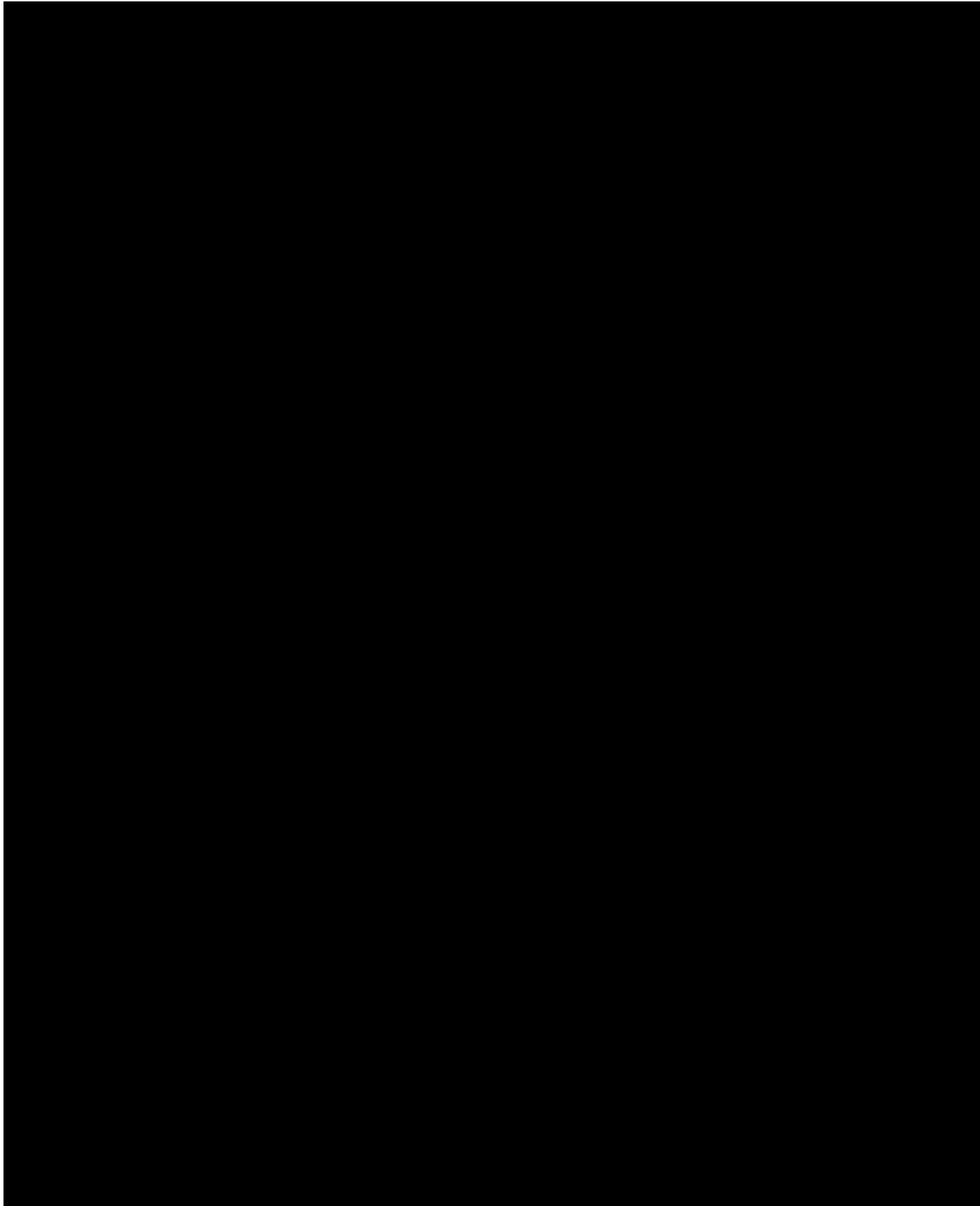


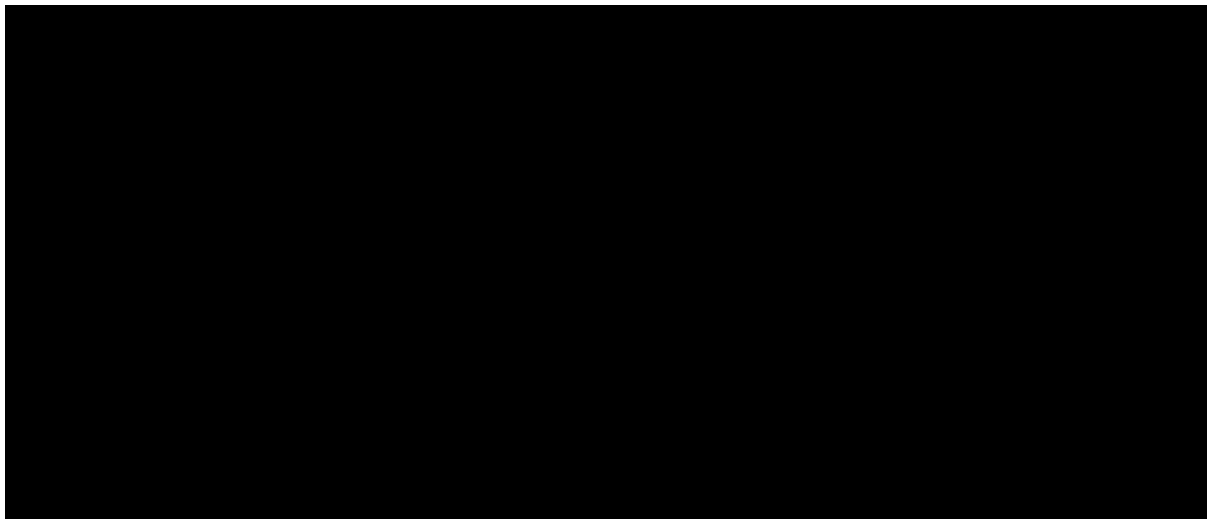
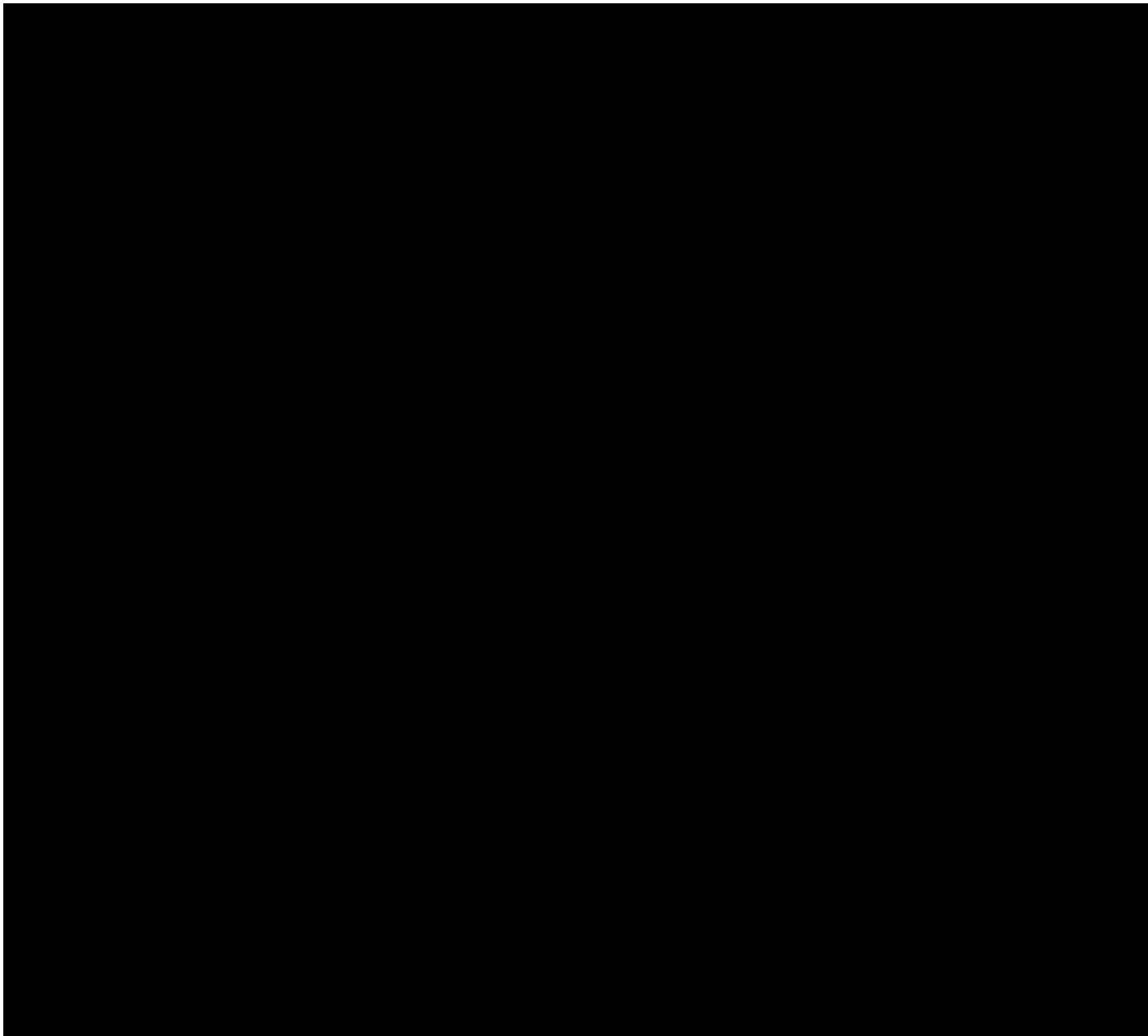


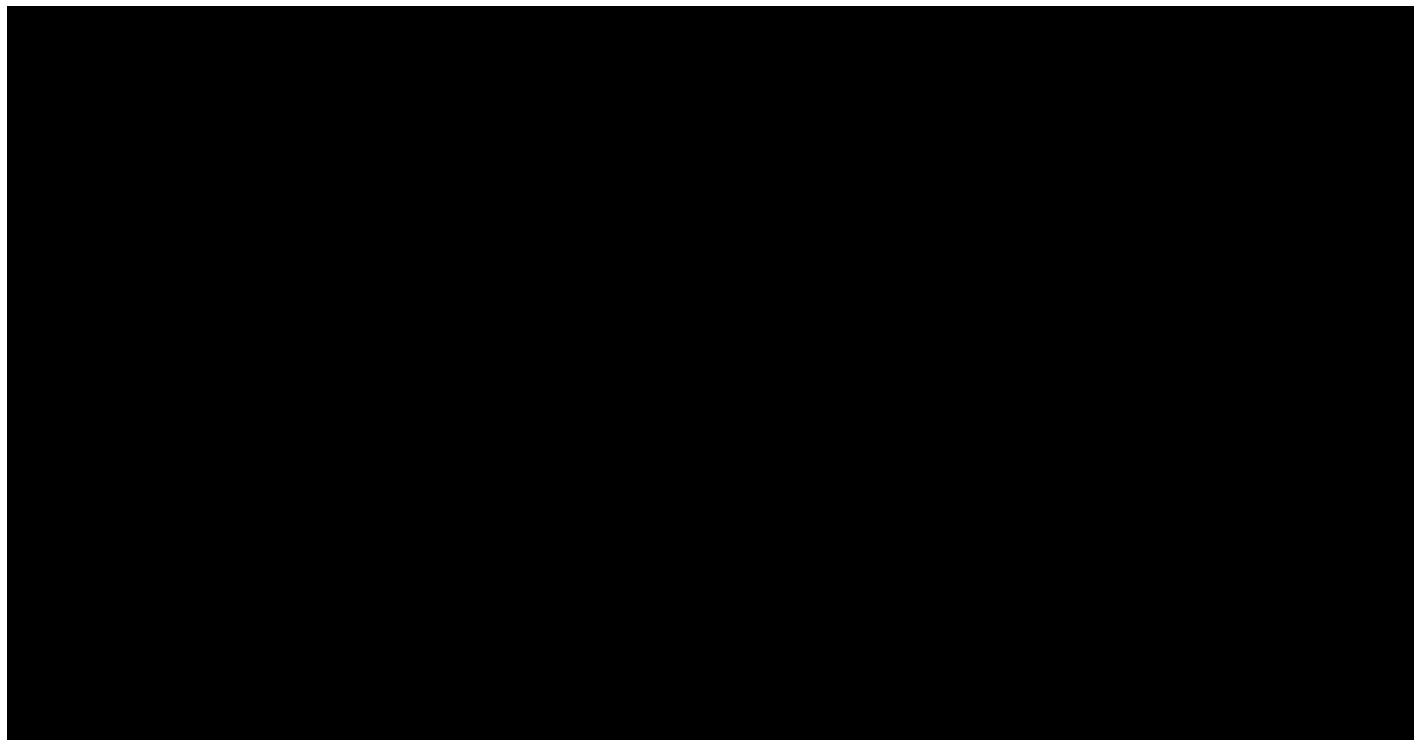
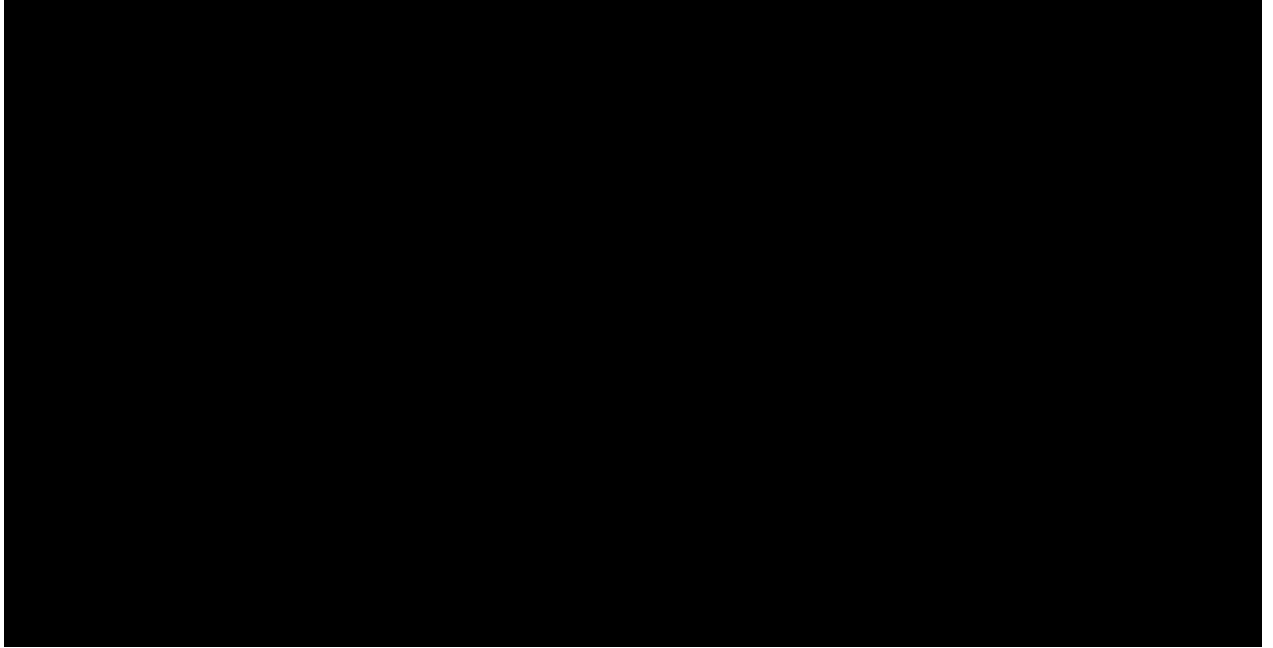
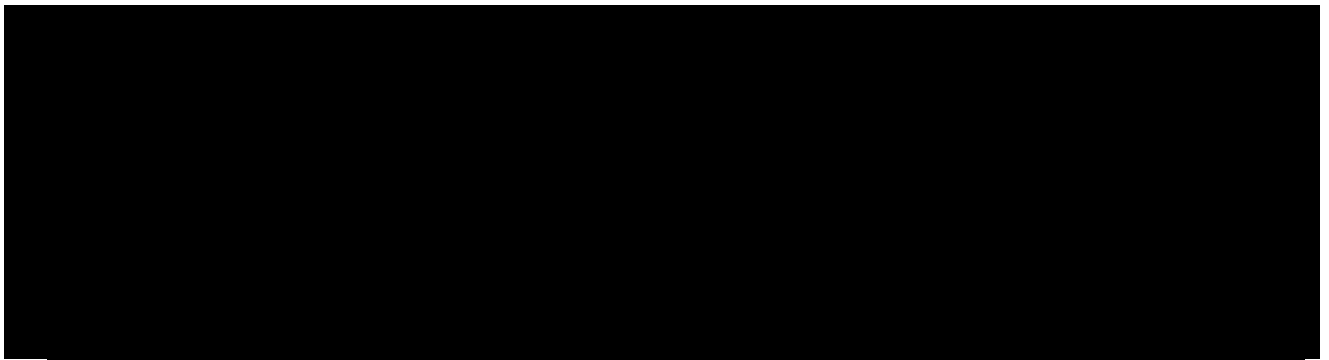


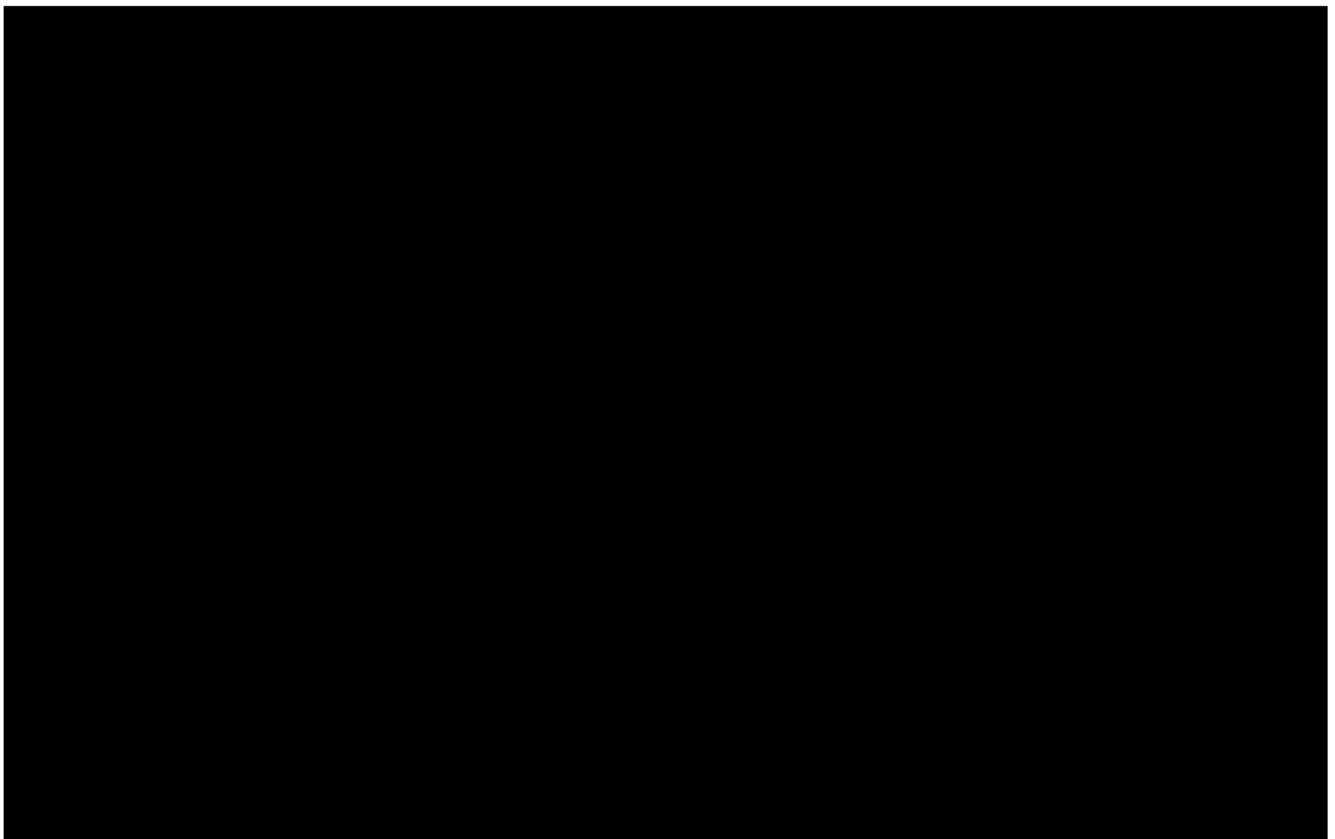
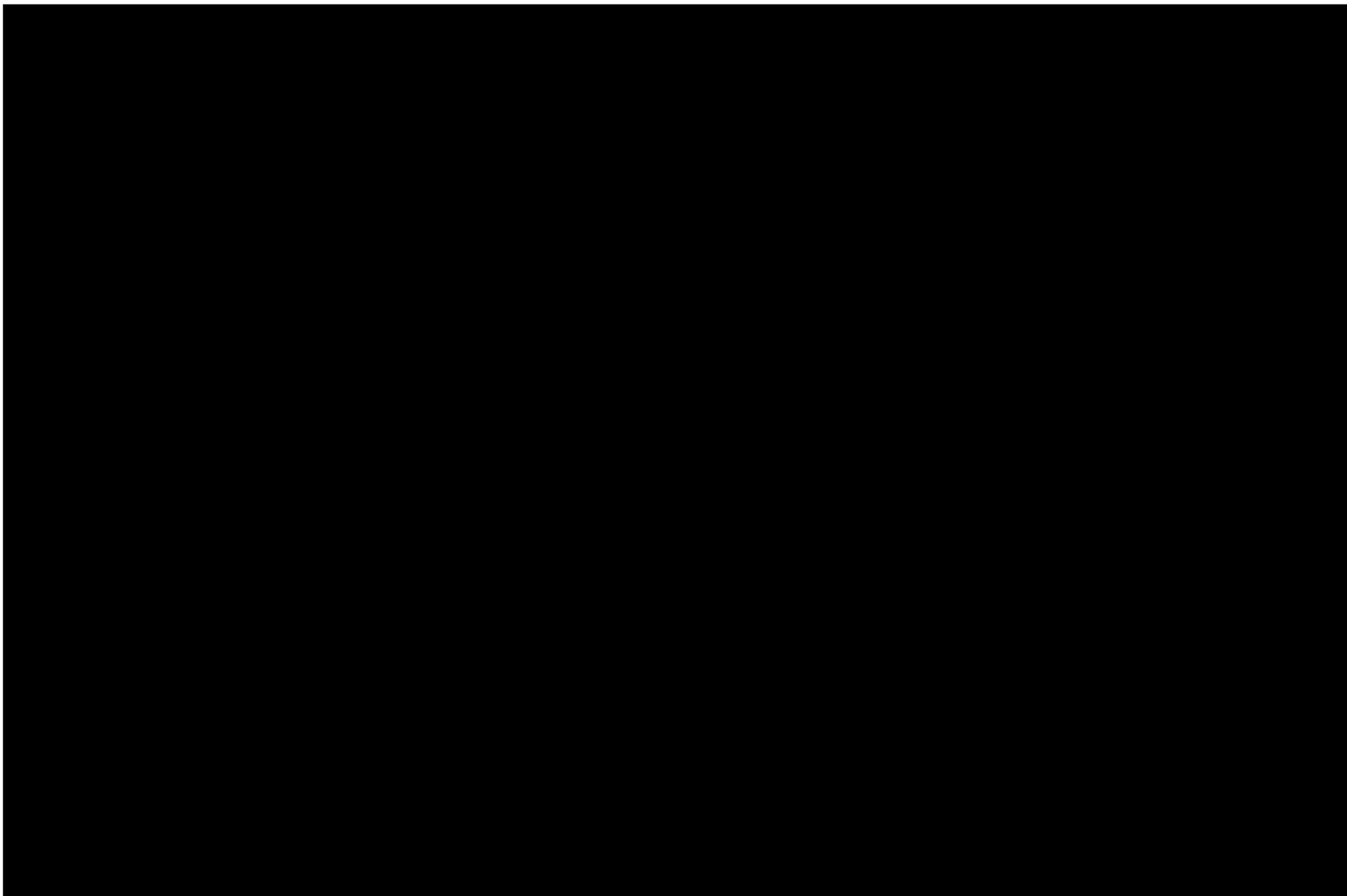


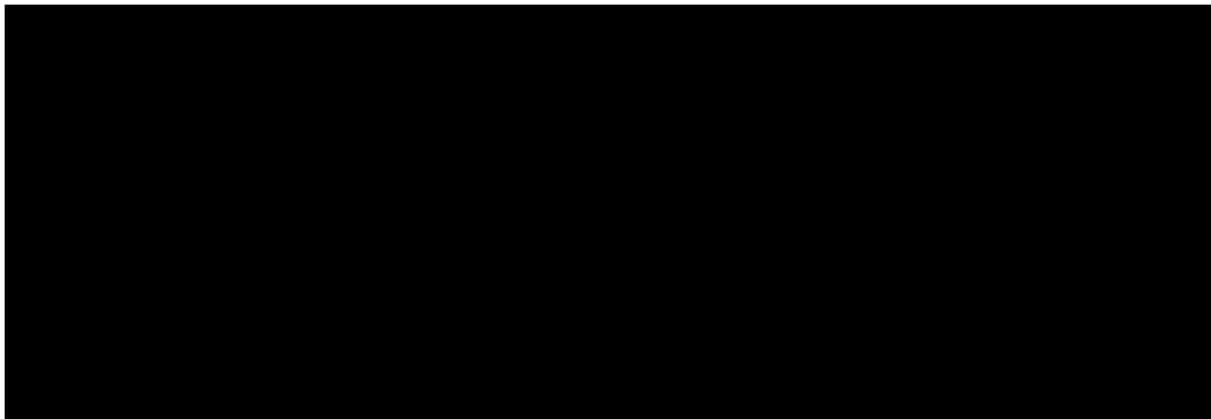
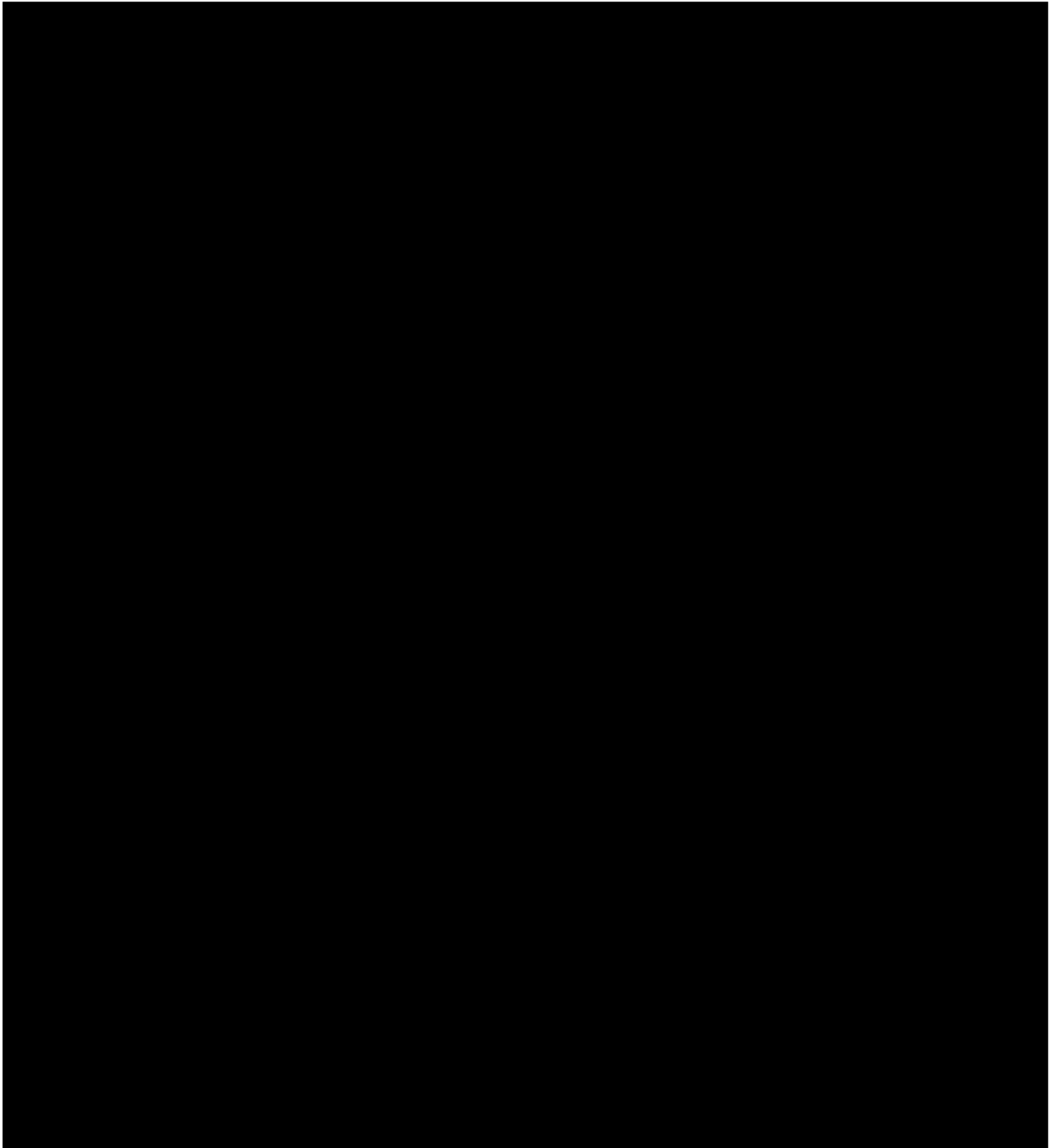
Appendix G. Sample SAS Code for Efficacy Analysis

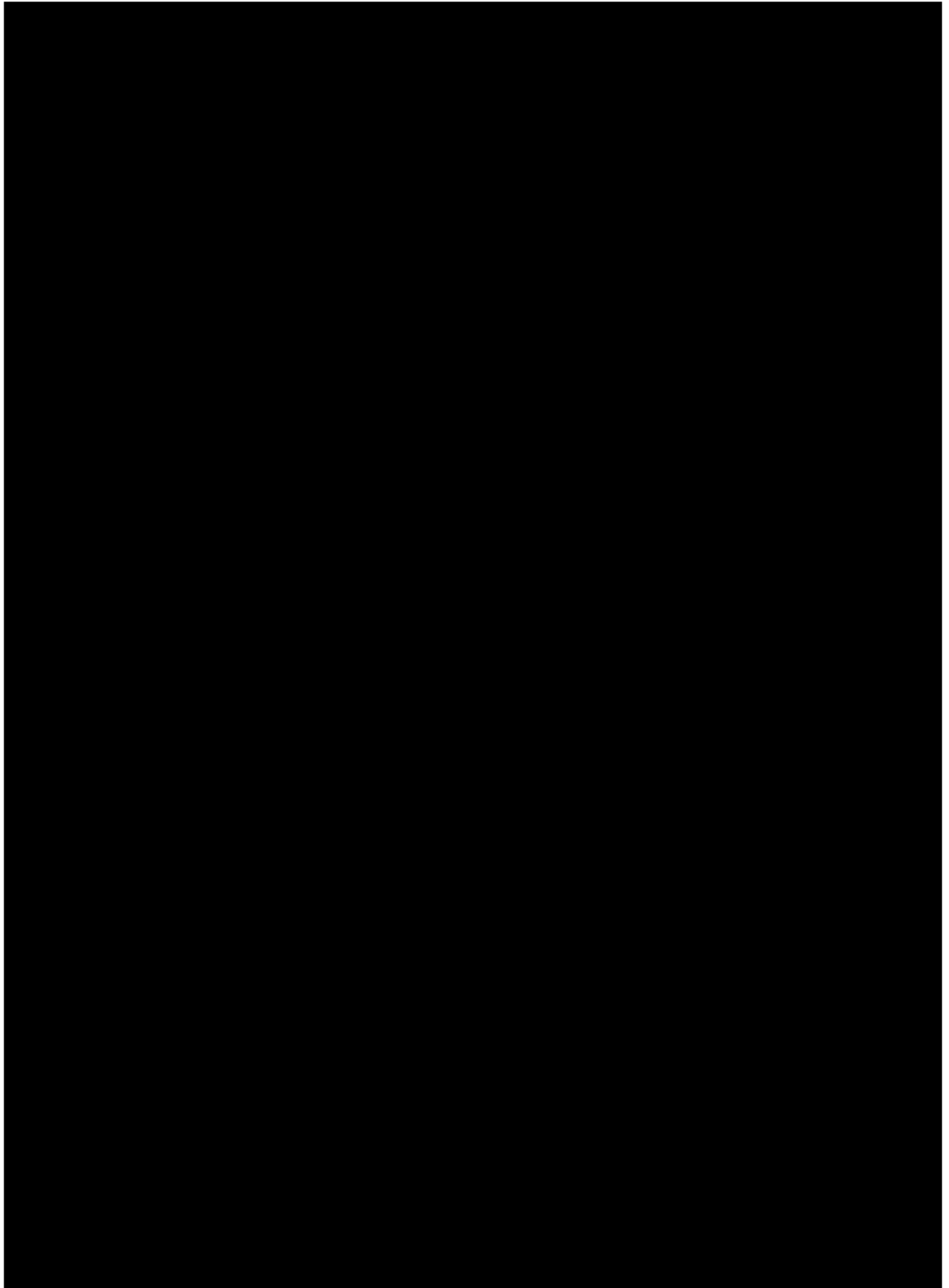


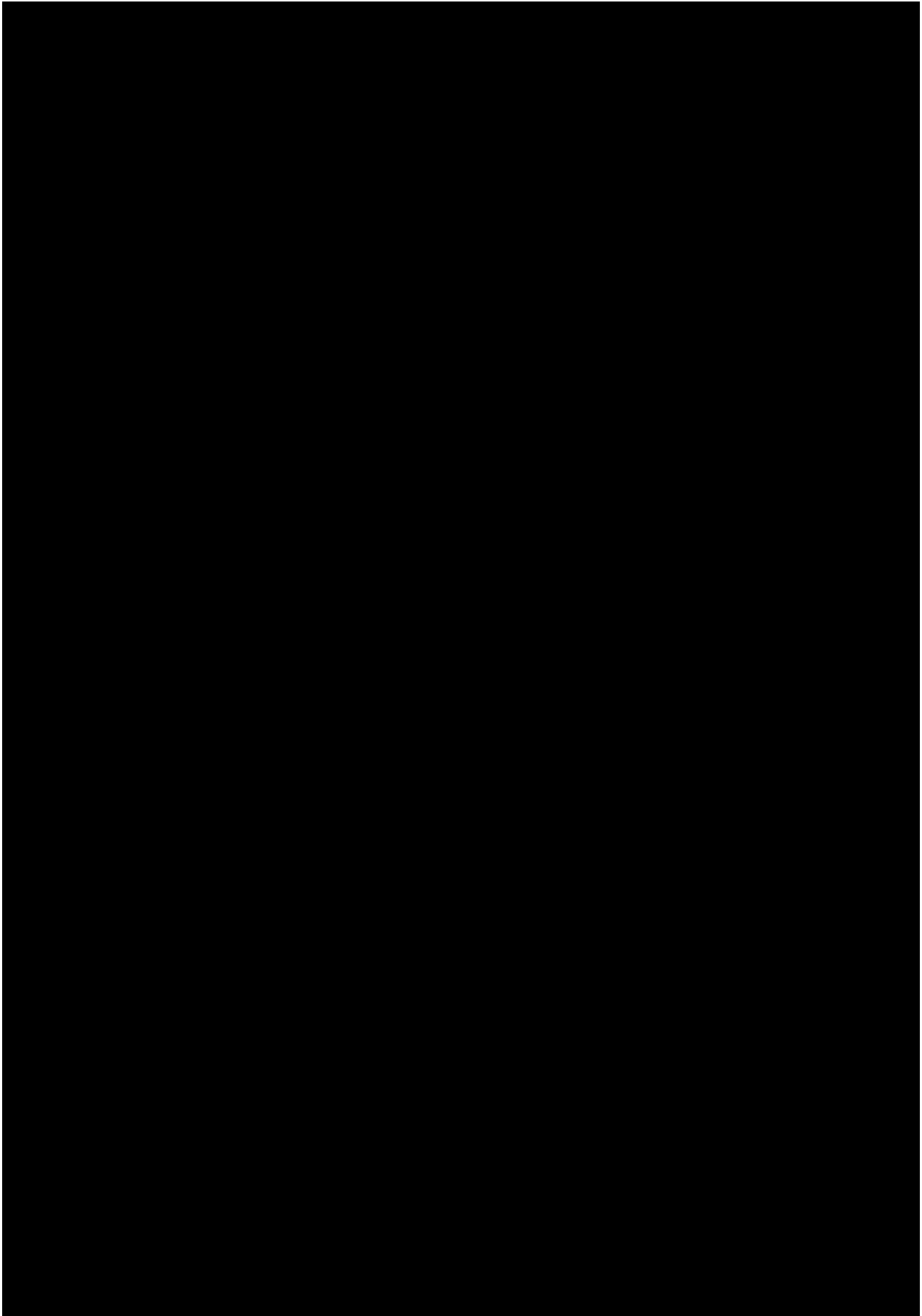












TITLE Statistical Analysis Plan (SAP), Observational Research SAP or Supplementary SAP Additional Information	DOCUMENT NO FORM-492867	VERSION 7.0
	EFFECTIVE DATE 27 Oct 2023	PAGE Page 1 of 1

Product: Rocatinlimab (AMG 451)	Protocol Number: 20210143
Database Snapshot Date: [REDACTED]	
SAP/OR-SAP/SSAP Author: [REDACTED]	[REDACTED]

Changes to Planned Analyses:

Route was inadvertently omitted from Appendix B (Atopic Dermatitis Medications of Interest Category and Definition) for identifying IVARMACITINIB as topical JAK inhibitor (Route = TOPICAL) or systemic JAK inhibitor (Route = ORAL). The omission is noted in bold below.

Topical JAK inhibitors

- IVARMACITINIB with route = **TOPICAL**

Systemic JAK inhibitors

- IVARMACITINIB with route = **ORAL**

Approved by GSL: [REDACTED]

Approved by study lead (*indicate applicable role such as:
EDL/GDL/GMAL/CfOR/HEOR*): [REDACTED] on behalf of [REDACTED]

Effectiv
e

AMGEN[®]			Form		
TITLE Statistical Analysis Plan (SAP), Observational Research SAP or Supplementary SAP Additional Information		DOCUMENT NO FORM-492867		VERSION 7.0	
		EFFECTIVE DATE 27 Oct 2023		PAGE Page 1 of 1	

Product: Rocatinlimab (AMG 451)		Protocol Number: 20210143	
Database Snapshot Date: [REDACTED]			
SAP/OR-SAP/SSAP Author: [REDACTED]		[REDACTED]	

Changes to Planned Analyses: Another set of post hoc sensitivity analyses will be performed on all binary and continuous efficacy endpoints at all assessment points where all observed data regardless of rescue therapy use for AD or permanent discontinuation of investigational product will be included (ie, treatment policy strategy), and missing data will be imputed by [REDACTED] procedure under [REDACTED].
Approved by GSL: [REDACTED]
Approved by study lead (<i>indicate applicable role such as: EDL/GDL/GMAL/CfOR/HEOR</i>): [REDACTED]

Effective
e

AMGEN[®]			Form		
TITLE Statistical Analysis Plan (SAP), Observational Research SAP or Supplementary SAP Additional Information		DOCUMENT NO FORM-492867		VERSION 7.0	
		EFFECTIVE DATE 27 Oct 2023		PAGE Page 1 of 1	

Product: Rocatinlimab (AMG 451)		Protocol Number: 20210143	
Database Snapshot Date: [REDACTED]			
SAP/OR-SAP/SSAP Author: [REDACTED]		[REDACTED]	

Changes to Planned Analyses:
Rationale: Study 20210143 was unblinded for the primary analysis on 28 August 2024. Per Study 20210143 SAP amendment 1 approved on 15 August 2024, NRS weekly average was calculated when there are at least 3 daily NRS scores in a 7-day weekly window. Subsequently, Amgen received an FDA Advice Letter dated 16 September 2024 requesting for sensitivity analysis using alternative missing data rules with respect to NRS endpoints.
Post hoc analysis: Sensitivity analyses with respect to key NRS endpoints will be performed based on weekly average to be calculated when there are at least 4 daily NRS scores in a 7-day weekly window. These post hoc analysis results will be included in the Study 20210143 CSR.
Approved by GSL: [REDACTED]
Approved by study lead (<i>indicate applicable role such as: EDL/GDL/GMAL/CfOR/HEOR</i>): [REDACTED]

Effectiv
e

AMGEN[®]			Form		
TITLE Statistical Analysis Plan (SAP), Observational Research SAP or Supplementary SAP Additional Information		DOCUMENT NO FORM-492867		VERSION 7.0	
		EFFECTIVE DATE 27 Oct 2023		PAGE Page 1 of 1	

Product: Rocatinlimab (AMG 451)		Protocol Number: 20210143	
Database Snapshot Date: [REDACTED]			
SAP/OR-SAP/SSAP Author: [REDACTED]		[REDACTED]	

This FORM provides clarification for [REDACTED] in SAP amendment 1 approved on 15 August 2024.	
When performing [REDACTED] for binary endpoints (eg, vIGA-AD 0/1, EASI 75), if there are no missing values or only a small number of missing values in the original dataset (eg, vIGA-AD response, total EASI score), 100 sets of derived binary endpoints of interest will likely be identical. [REDACTED]. In this scenario, the combined [REDACTED] result is equivalent to single-dataset analysis, and only the first imputed dataset will be analyzed to report the combined [REDACTED] result [REDACTED]	
Approved by GSL: [REDACTED]	
Approved by study lead (indicate applicable role such as: EDL/GDL/GMAL/CfOR/HEOR): [REDACTED]	

Effectiv
e

TITLE Statistical Analysis Plan (SAP), Observational Research SAP or Supplementary SAP Additional Information	DOCUMENT NO FORM-492867	VERSION 7.0
	EFFECTIVE DATE 27 Oct 2023	PAGE Page 1 of 2

Product: Rocatinlimab (AMG 451)	Protocol Number: 20210143
Database Snapshot Date: [REDACTED] [REDACTED]	
SAP/OR-SAP/SSAP Author: [REDACTED]	[REDACTED]

This FORM provides a revision to the safety endpoint definition of Treatment-Emergent Adverse Events (TEAEs) in SAP Amendment [REDACTED]

[REDACTED] The update reflects the need to include 3 non-serious adverse events that were inadvertently omitted from the primary and final analyses due to data issues resulting in incorrect treatment-emergent flag. Specifically, the field "Did event start before first dose of investigational product" was incorrectly marked as "Yes" for these 3 adverse events that occurred after the first dose. The updated TEAE definition will consider these 3 adverse events as treatment-emergent, and all adverse event tables for the primary analysis and final analysis have been updated.

Treatment-emergent Adverse Event (TEAE) definition in SAP amendment 1

Events categorized as Adverse Events (AEs) 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 electronic case report form (eCRF), and up to the End of Study (EOS) date.

Updated Treatment-emergent Adverse Event (TEAE) definition

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

Approved by GSL: [REDACTED]

Effective
e

AMGEN[®]			Form		
TITLE Statistical Analysis Plan (SAP), Observational Research SAP or Supplementary SAP Additional Information		DOCUMENT NO FORM-492867		VERSION 7.0	
		EFFECTIVE DATE 27 Oct 2023		PAGE Page 2 of 2	
Approved by study lead (<i>indicate applicable role such as:</i> <i>EDL/GDL/GMAL/CfOR/HEOR</i>): [REDACTED]					

Effectiv
e