Statistical Analysis Plan Version 2 J2T-JE-KGAL

A Randomized, Double-Blind, Placebo-Controlled Trial to Evaluate the Efficacy and Safety of Lebrikizumab When Used in Combination with Topical Corticosteroid Treatment in Japanese Patients with Moderate-to-Severe Atopic Dermatitis

NCT04760314

Approval Date: 22-Aug-2022

Title Page

Protocol Title: A Randomized, Double-Blind, Placebo-Controlled Trial to Evaluate the Efficacy and Safety of Lebrikizumab When Used in Combination with Topical Corticosteroid Treatment in Japanese Patients with Moderate-to-Severe Atopic Dermatitis

Protocol Number: J2T-JE-KGAL

Compound Number: LY3650150

Short Title: Safety and Efficacy of Lebrikizumab in Combination with Topical Corticosteroids in Japanese Patients with Moderate-to-Severe Atopic Dermatitis

Acronym: ADhere-J

Sponsor Name: Eli Lilly Japan K.K.

Legal Registered Address: Eli Lilly Japan K.K., Kobe Hyogo Japan

Regulatory Agency Identifier Number(s)

Registry	ID
Not available	

Table of Contents

Title Page **1**

Table of Contents **2**

Version history **4**

1. Introduction **13**
1.1. Objectives, Endpoints, and Estimands 14
1.1.1. Estimands 17
1.2. Study Design 24

2. Statistical Hypotheses **28**
2.1. Multiplicity Adjustment 29

3. Analysis Sets **31**

4. Statistical Analyses **36**
4.1. General Considerations 36
4.1.1. General Considerations for Analyses During Induction Period 36
4.1.2. General Considerations for Analyses During Maintenance Period 38
4.1.3. General Considerations for Safety Analyses for Combined Induction and Maintenance Periods 39
4.1.4. General Considerations for Safety Analyses for Combined Induction and Maintenance Periods Plus Follow-Up Period 39
4.1.5. Adjustments for Covariates 40
4.1.6. Handling of Dropouts or Missing Data 40
4.2. Participant Dispositions 46
4.3. Primary Endpoints Analysis 47
4.3.1. Definition of endpoints 47
4.3.2. Main analytical approach 47
4.3.3. Sensitivity Analyses 48
4.3.4. Supplementary analyses 48
4.4. Secondary Endpoints Analysis 48
4.4.1. Major secondary endpoints 48
4.4.2. Other secondary endpoints 48
4.5. Tertiary/Exploratory Endpoints Analysis 49
4.6. Safety Analyses 49
4.6.1. Extent of Exposure 49
4.6.2. Adverse Events 50
4.6.3. Additional Safety Assessments 52
4.7. Other Analyses 59
4.7.1. Efficacy and Safety Analyses for Participants Who Self-Administer Study Intervention 59
4.7.2. Subgroup analyses 59
4.8. Interim Analyses 60
4.8.1. Data Monitoring Committee (DMC) 61
4.9. Changes to Protocol-Planned Analyses 61

5. Sample Size Determination **62**

6. Supporting Documentation 63
6.1. Appendix 1: Description and Derivation of Efficacy and Health Outcome Endpoints 63
6.2. Appendix 2: Description of Efficacy and Health Outcome Analyses 81
6.3. Appendix 3: Demographic and Baseline Characteristics 99
6.4. Appendix 4: Medical History 101
6.5. Appendix 5: Treatment Compliance 101
6.6. Appendix 6: Prior and Concomitant Therapy 103
6.6.1. Rescue Medication for Atopic Dermatitis (AD) 103
6.7. Appendix 7: Protocol Deviations 104
6.8. Appendix 8: Impact of COVID-19 104
6.9. Appendix 9: Annual Report Analyses 104
6.10. Appendix 10: Clinical Trial Registry Analyses 104
6.11. Appendix 11: Unblinding Plan 105
6.12. Appendix 12: Study Visit Mapping for Itch NRS, Skin Pain NRS and Sleep-loss and POEM 105
6.13. Appendix 13: Definition of Topical and Systemic Atopic Dermatitis Therapy 106
6.14. Appendix 14: Details of Combining Estimates and Test Statistics for Categorial Endpoints with Multiple Imputation 107

7. References 110

Version history

This Statistical Analysis Plan (SAP) for Study J2T-JE-KGAL (KGAL) is based on the protocol amendment (a) dated 22 January 2021 and approved prior to any unblinding.

Table 1 SAP Version History Summary

SAP Version	Approval Date	Change	Rationale
1	12-Jul-2021	Not Applicable	Original version
2		Section 1.1 • Added "Time to loss of EASI-50 during the 52-week Maintenance Period among those originally randomized to lebrikizumab 250 mg Q2W and lebrikizumab 250 mg Q4W who achieved EASI-75 or IGA score of 0 or 1 with a ≥ 2-point improvement from baseline at Week 16" • Added "Time to loss of EASI-75 during the 52-week Maintenance Period among those originally randomized to lebrikizumab 250 mg Q2W and lebrikizumab 250 mg Q4W who achieved EASI-75 at Week 16" • Added "Time to loss of IGA response, i.e.,	To be consistent with global pivotal studies (i.e., KGAB, KGAC, and KGAD)

		developing an IGA score ≥ 2 with 2 points deterioration of achieved IGA response at Week 16, during the 52-week Maintenance Period among those originally randomized to lebrikizumab 250 mg Q2W and lebrikizumab 250 mg Q4W who achieved IGA score of 0 or 1 with a ≥ 2-point improvement from baseline at Week 16” • Added “Proportion of participants with an Itch NRS score of ≥ 4 points at baseline who achieve both EASI-75 and a ≥ 4-point reduction in Itch NRS score from baseline by visit” • Added “Proportion of participants with an Itch NRS score of ≥ 4 points at baseline who achieve both an IGA score of 0 or 1 with a ≥ 2-point improvement from baseline, and a ≥ 4-point reduction in Itch NRS score from baseline by visit”	
		Section 1.1.1	

		• Added summary tables of intercurrent events and missing data	To be consistent with global pivotal studies (i.e., KGAB, KGAC, and KGAD)
		Section 2.1 • Updated the graphical testing scheme for multiplicity control of primary and key secondary endpoints	To prespecify the graphical testing scheme
		Section 3 • Added All Maintenance Population	To summarize important protocol deviations during the Maintenance Period on this population
		Section 4.1.1 • Added more alternative covariance structures for within-participant errors	To help convergence of the covariance matrix
		Section 4.6.1 • Drug interruption time period due to the use of systemic rescue therapies will <u>not</u> be removed from study drug exposure calculations. • Updated categories on duration of exposure	To be consistent with global pivotal studies (i.e., KGAB, KGAC, and KGAD) and compound level safety standards
		Section 4.6.3.1	

		• Added box plots for All Lebrikizumab Safety Population	To be consistent with Safety Population and Maintenance Primary Population
		Section 4.6.3.2 • Added box plots for All Lebrikizumab Safety Population	To be consistent with Safety Population and Maintenance Primary Population
		Section 4.6.3.3 • Updated analysis populations on analysis of immunogenicity	To be consistent with global pivotal studies (i.e., KGAB, KGAC, and KGAD) and compound level safety standards
		Section 4.6.3.4.4 • Added summary tables of TEAE of conjunctivitis cluster by maximum severity	To be consistent with global pivotal studies (i.e., KGAB, KGAC, and KGAD) and compound level safety standards
		Section 4.6.3.4.5 • Removed a listing of participants with hypersensitivity	To be consistent with global pivotal studies (i.e., KGAB, KGAC, and KGAD)
		Section 4.6.3.4.9 • Updated section heading for Suicide/Self-injury	To be consistent with global pivotal studies (i.e., KGAB, KGAC, and KGAD) and compound level safety standards
		Section 4.7.2.1 • Added a subgroup of "Baseline EASI (≥16 to ≤21, >21 to ≤ 50, >50 to ≤72)" • Added a subgroup of "Prior use of systemic treatment (yes, no)"	To prespecify the analysis for this subgroup To be consistent with global pivotal studies (i.e., KGAB, KGAC, and KGAD)

Approved on 22 Aug 2022 GMT

		Section 4.7.2.2 • Updated a description for safety subgroup analysis	This study will not be included in the integrated safety analysis for Japan submission. Therefore, the safety subgroup analysis may be performed for this study to meet regulatory requirement.
		Section 4.8 • Added description of tiered database lock (DBL) for the first DBL	To prespecify the tiered DBL according to internal guidance on the tiered DBL
		Section 6.1 • Updated derivation of "Time to loss of IGA response" variable • Added derivation of "Time to loss of EASI-75" variable	To be consistent with global pivotal studies (i.e., KGAB, KGAC, and KGAD)
		• Added variables and derivations for TARC • Added variables and derivations for each $EASI_{region}$ score and symptom score • Added variables and derivations for Itch NRS daily score • Added variables and derivations for Sleep-loss daily score	To prespecify the analysis for these variables
		• Removed a description of "If more than one	This situation does not happen in this study using eCOA.

Approved on 22 Aug 2022 GMT

		response is selected, then the response with the highest score is used." from Imputation Approach for POEM	
		Section 6.2 • Updated Estimand and Analysis Method of "Proportion of participants achieving IGA (0)" • Added analysis of "Proportion of participants achieving both IGA (0,1) with a ≥ 2-point improvement and a ≥ 4-point improvement in Itch NRS prorated weekly mean score" • Added analysis of "Proportion of participants achieving both EASI-75 and a ≥ 4-point improvement in Itch NRS prorated weekly mean score" • Updated Population for "Time to loss of EASI-50" • Added analysis of "Time to loss of EASI-75"	To be consistent with global pivotal studies (i.e., KGAB, KGAC, and KGAD)

Approved on 22 Aug 2022 GMT

		• Added analysis of each $EASI_{region}$ score and symptom score • Added analysis of Itch NRS daily score-related variables • Added analysis of Sleep-loss daily score-related variables • Added analysis of TARC-related variables	To prespecify the analysis
		• Updated Estimand and Analysis Method of "Proportion of participants achieving EASI-50"	To be consistent with other secondary endpoints
		• Removed "by visit" summary for Proportion of TCS/TCI-free days • Removed "by visit" summary for Mean weight of TCS use by potency (tube weights)	TCS data collection might not be sufficient for the "by visit" analysis.
		Section 6.3 • Added summary on Maintenance W16 Escape Population • Added summary of participant disease characteristics at Week 16 on the MPP	To be consistent with global pivotal studies (i.e., KGAB, KGAC, and KGAD)

		• Added summary on ITT population by participants who experienced or did not experience conjunctivitis adverse events	To prespecify the analysis
		• Updated Atopic Dermatitis treatment used in the past	To be consistent with global pivotal studies (i.e., KGAB, KGAC, and KGAD) and Appendix 13 in Section 6.13
		• Added "Employment status: Employed, Not employed" in WPAI-AD	To be consistent with Appendix 1 in Section 6.1 and Appendix 2 in Section 6.2
		• Added "Thymus and activation-regulated chemokine (TARC)"	
		Section 6.4 • Added summary on MPP and Maintenance W16 Escape Population	To be consistent with global pivotal studies (i.e., KGAB, KGAC, and KGAD)
		Section 6.6.1 • Added summary of rescue medications on Maintenance W16 Escape Population	To be consistent with global pivotal studies (i.e., KGAB, KGAC, and KGAD)
		• Added summary of flare by visit on ITT Population, MPP, and Maintenance W16 Escape Population	

		Section 6.8 • Added a description of how missing data due to pandemic will be handled	To be consistent with global pivotal studies (i.e., KGAB, KGAC, and KGAD)
		Section 6.12 • Added a description of "If an assessment could be mapped to different weeks, it will be mapped to the earlier week."	To be consistent with global pivotal studies (i.e., KGAB, KGAC, and KGAD)
		Section 6.13 • Added "Route of topical treatments includes: Topical and Transdermal." • Added "Topical JAK (Janus kinase) inhibitor: Preferred Term includes: DELGOCITINIB" • Added "Topical PDE4 (phosphodiesterase 4) inhibitor: Preferred Term includes: DIFAMILAST" • Added Baricitinib and Upadacitinib to Immunosuppressant	To be consistent with global pivotal studies (i.e., KGAB, KGAC, and KGAD) This drug has been approved for AD treatment in Japan.

1. Introduction

This SAP includes the analysis plan for efficacy, health outcome, safety, and immunogenicity data.

The table, figure, and listing (TFL) specifications are contained in a separate document.

1.1. Objectives, Endpoints, and Estimands

Objectives	Endpoints
Co-Primary	
• To test the hypothesis that lebrikizumab 250 mg Q2W or lebrikizumab 250 mg Q4W is superior to placebo in reducing signs and symptoms of AD at Week 16 in Japanese participants with moderate to severe AD when used in combination with TCS treatment	• Proportion of participants achieving EASI-75 at Week 16 • Proportion of participants achieving IGA score of 0 or 1 and a reduction of ≥ 2-points from baseline to Week 16
Major Secondary	
• To compare the efficacy and health outcome measures of lebrikizumab 250 mg Q2W or lebrikizumab 250 mg Q4W to placebo during the 16-week Induction Period in Japanese participants with moderate to severe AD when used in combination with TCS treatment	• Percentage change in EASI score from baseline to Week 16 • Proportion of participants achieving EASI-90 at Week 16 • Proportion of participants with an Itch NRS score of ≥ 4-points at baseline who achieve a ≥ 4-point reduction from baseline to Weeks 1, 2, 4 and 16
Other Secondary	
• To measure lebrikizumab exposure and assess the relationship between exposure and immunogenicity	• Average serum lebrikizumab concentration at steady state • Lebrikizumab serum trough concentrations associated with ADA titer
• To evaluate the efficacy and health outcome measures of lebrikizumab 250 mg Q2W and lebrikizumab 250 mg Q4W during the 16-week Induction and 52-week Maintenance periods in Japanese participants with moderate to severe AD when used in combination with TCS treatment	• Percentage change from baseline in EASI score by visit • Proportion of participants with EASI-50, EASI-75, and EASI-90 by visit • Proportion of participants maintaining EASI-75 by visit during the 52-week Maintenance Period among those originally randomized to lebrikizumab 250 mg Q2W and lebrikizumab 250

	mg Q4W who achieved EASI-75 at Week 16 • Time to loss of EASI-50 during the 52-week Maintenance Period among those originally randomized to lebrikizumab 250 mg Q2W and lebrikizumab 250 mg Q4W who achieved EASI-75 or IGA score of 0 or 1 with a ≥ 2-point improvement from baseline at Week 16 • Time to loss of EASI-75 during the 52-week Maintenance Period among those originally randomized to lebrikizumab 250 mg Q2W and lebrikizumab 250 mg Q4W who achieved EASI-75 at Week 16 • Proportion of participants with an IGA score of 0 or 1 and a reduction of ≥ 2 points from baseline by visit • Proportion of participants maintaining an IGA score of 0 or 1 with a ≥ 2-point improvement from baseline by visit during the 52-week Maintenance Period among those originally randomized to lebrikizumab 250 mg Q2W and lebrikizumab 250 mg Q4W who achieved IGA score of 0 or 1 with a ≥ 2-point improvement from baseline at Week 16 • Time to loss of IGA response, i.e., developing an IGA score ≥ 2 with 2 points deterioration of achieved IGA response at Week 16, during the 52-week Maintenance Period among those originally randomized to lebrikizumab 250 mg Q2W and lebrikizumab 250 mg Q4W who achieved IGA score of 0 or 1 with a ≥ 2-point improvement from baseline at Week 16 • Change from baseline in Itch NRS score by visit

	• Proportion of participants with an Itch NRS score of ≥ 4 points at baseline who achieve a ≥ 4-point reduction from baseline by visit • Proportion of participants with an Itch NRS score of ≥ 4 points at baseline who achieve both EASI-75 and a ≥ 4-point reduction in Itch NRS score from baseline by visit • Proportion of participants with an Itch NRS score of ≥ 4 points at baseline who achieve both an IGA score of 0 or 1 with a ≥ 2-point improvement from baseline, and a ≥ 4-point reduction in Itch NRS score from baseline by visit • Change from baseline in Skin Pain NRS score by visit • Proportion of participants with a Skin Pain NRS score of ≥ 4 points at baseline who achieve a ≥ 4-point reduction from baseline by visit • Change from baseline in percent BSA by visit • Change from baseline in Sleep-Loss score by visit • Proportion of participants achieving a ≥ 4-point improvement in DLQI/CDLQI score from baseline by visit • Proportion of participants achieving DLQI/CDLQI score of 0 or 1 by visit • Change from baseline in DLQI/CDLQI by visit • Change from baseline in POEM by visit • Change from baseline in WPAI-AD score by visit • Change from baseline in HADS score by visit • Change from Baseline in ACQ-5 score to Week 16 in participants who have self-reported comorbid asthma

	• Percentage change from Baseline in SCORAD by visit • Proportion of TCS/TCI-free days over the 16 week and 68-week study periods • Total amount of TCS used over the 16-week and 68-week study periods • Time (days) to TCS/TCI-free use over the 16-week and 68-week study periods
• To assess the growth of adolescent participants treated with lebrikizumab	• Mean changes in growth parameters (height, weight, and BMI) over the course of treatment

Abbreviations: ACQ-5 = Asthma Control Questionnaire-5; AD = atopic dermatitis; ADA = anti-drug antibody; BMI = body mass index; BSA = body surface area; DLQI/CDLQI = Dermatology Life Quality Index/Children's Dermatology Life Quality Index; EASI = Eczema Area and Severity Index; EASI-50 = ≥50% reduction from baseline in EASI score; EASI-75 = ≥75% reduction from baseline in EASI score; EASI-90 = ≥90% reduction from baseline in EASI score; HADS = Hospital Anxiety Depression Scale; IGA = Investigator's Global Assessment; NRS = numeric rating scale; POEM = Patient-Oriented Eczema Measure; Q2W = every 2 weeks; Q4W = every 4 weeks; SCORAD = SCORing Atopic Dermatitis; TCS/TCI = topical corticosteroid/topical calcineurin inhibitor; WPAI-AD = Work Productivity and Activity Impairment – Atopic Dermatitis.

1.1.1. Estimands

1.1.1.1. Primary and Supportive Estimands for Induction Period

There will be three estimands of interest in analyzing primary and key secondary endpoints for the Induction Period. Two types of intercurrent events in terms of estimating the treatment effects for the Induction Period will be considered: initiation of rescue medication as defined in Section 6.6.1 and permanent treatment discontinuation.

1.1.1.1.1. Primary Estimand (Hybrid)

The primary estimand is a hybrid estimand representing the primary clinical question of interest: what is the difference between treatment conditions, i.e., Lebrikizumab vs Placebo, in the target participant population, in successful responses or means after 16 weeks achieved without use of rescue medication and if all participants continued with treatment except those who discontinued due to lack of efficacy?

The primary estimand is described by the following attributes:

A. Population: defined through appropriate Inclusion/Exclusion (I/E) criteria to reflect the targeted participant population for approval

B. Endpoint: apply to all primary and key secondary endpoints

C. How to account for intercurrent events (ICEs)

a. Participants who require any use of rescue medication or discontinued treatment due to lack of efficacy prior to Week 16 will be considered as treatment failures, i.e., non-responder, after the ICEs. Therefore, composite strategy is used for these types of ICEs.

b. For participants who discontinue treatment due to reasons other than lack of efficacy prior to Week 16, a hypothetical strategy will be used to estimate what the treatment effect would have been if participants continued with treatment. Therefore, hypothetical strategy is used for these types of ICEs.

D. Population-level summary: difference in response proportions or means between treatment conditions

1.1.1.1.2. Supportive Estimand for Categorical Endpoints (Composite)

The supportive estimand for categorical endpoints is a composite estimand representing the supportive clinical question of interest: what is the difference between treatment conditions in the target participant population, in successful responses after 16 weeks achieved without use of rescue medication or treatment discontinuation?

The supportive estimand is described by the following attributes:

A. Population: defined through appropriate I/E criteria to reflect the targeted participant population for approval

B. Endpoint: apply to all categorical endpoints

C. How to account for ICEs

a. Participants who require any use of rescue medication or discontinued treatment prior to Week 16 will be considered as treatment failures, i.e., non-responder, after the ICEs. Therefore, composite strategy is used for these types of ICEs.

D. Population-level summary: difference in response proportions between treatment conditions

1.1.1.1.3. Supportive Estimand for Continuous Endpoints (Hypothetical)

The supportive estimand for continuous endpoints is a hypothetical estimand representing the supportive clinical question of interest: what is the difference between treatment conditions in the target participant population, in means after 16 weeks if rescue medication was not available and all participants adhered to the treatment?

The supportive estimand is described by the following attributes:

A. Population: defined through appropriate I/E criteria to reflect the targeted participant population for approval

B. Endpoint: apply to all continuous endpoints

C. How to account for ICEs

a. For participants who require any use of rescue medication or discontinued treatment prior to Week 16, a hypothetical strategy will be used to estimate what the treatment effect would have been if rescue medication was not available and all participants adhered to the treatment. Therefore, hypothetical strategy is used for these types of ICEs.

D. Population-level summary: difference in means between treatment conditions

Details on how missing data including those as a result of intercurrent events will be handled can be found in Section 4.1.6.1. Detailed analyses relative to estimands including analysis type, method and imputation, population, time point, and treatment comparisons for efficacy and health outcome analyses can be found in Appendix 2.

The table below summarizes the analytical strategies that will be conducted on the intercurrent events for the three estimands.

Description of Primary and Supportive Estimands for Induction Period

Estimand	Analysis Strategy for Intercurrent Events			Missing Data Imputation Method
	Rescue Medication	Treatment Discontinuation		
		Due to lack of efficacy	Due to any other reasons	
Primary Estimand (Hybrid)	Composite: Set to baseline	Composite: Set to baseline	Hypothetical: Set to missing	Primary analysis: MCMC-MI
	Composite: Set to baseline	Composite: Set to baseline	Hypothetical: Set to missing	Sensitivity analysis Tipping point analysis
Supportive Estimand for Categorical Endpoints (Composite)	Composite: Set to non-responder	Composite: Set to non-responder	Composite: Set to non-responder	Supplementary analysis: NRI
Supportive Estimand for Continuous Endpoints (Hypothetical)	Hypothetical: Set to missing	Hypothetical: Set to missing	Hypothetical: Set to missing	Supplementary analysis MMRM/LOCF

Abbreviations: LOCF = last observation carried forward; MCMC-MI = Markov chain Monte Carlo multiple imputation; MMRM = mixed-effects model for repeated measures; NRI = non-responder imputation.

1.1.1.2. Primary and Supportive Estimands for Maintenance Period

There will be four estimands of interest in analyzing endpoints for the Maintenance Period. Three types of intercurrent events in terms of estimating the treatment effects for the Maintenance Period will be considered, initiation of rescue medication, permanent treatment discontinuation and transfer to escape arm.

1.1.1.2.1. Maintenance Primary Estimand (Hybrid)

The maintenance primary estimand is a hybrid estimand representing the clinical question of interest: what is the response proportions or means for each treatment condition (i.e., Lebrikizumab 250 mg Q2W and Lebrikizumab 250 mg Q4W), in the target participant population, in successful responses or means after 68 weeks achieved without use of systemic rescue medication, without transferring to escape arm, if topical rescue medication were not available and if all participants continued with treatment except those who discontinued due to lack of efficacy?

The maintenance primary estimand is described by the following attributes:

A. Population: Maintenance Primary Population (MPP) as described in Section 3

B. Endpoint: apply to key categorical and continuous endpoints (i.e., Investigator's Global Assessment [IGA], Eczema Area and Severity Index [EASI] and Itch Numeric Rating Scale [NRS] related endpoints) for the Maintenance Period

C. How to account for ICEs

    a. Participants who require any use of systemic rescue medication, discontinue treatment due to lack of efficacy after Week 16, or transfer to escape arm will be considered as treatment failures, i.e., non-responder, after the ICEs. Therefore, composite strategy is used for these types of ICEs.

    b. For participants who require any use of topical rescue medication, a hypothetical strategy will be used to estimate what the response proportions or means for each treatment condition would have been if participants continued with treatment. Therefore, hypothetical strategy is used for these types of ICEs.

    c. For participants who discontinue treatment due to reasons other than lack of efficacy after week 16, a hypothetical strategy will be used to estimate what the response proportions or means for each treatment condition would have been if participants continued with treatment. Therefore, hypothetical strategy is used for these types of ICEs.

D. Population-level summary: response proportions or means for each treatment condition

1.1.1.2.2. Maintenance Supportive Estimand (Hybrid)

The maintenance supportive estimand for both continuous and categorical endpoints is a hybrid estimand representing the clinical question of interest: what is the response proportions or means for each treatment condition (i.e., Lebrikizumab 250 mg Q2W and Lebrikizumab 250 mg Q4W), in the target participant population, in successful responses or means after 68 weeks achieved without use of systemic rescue medication, without transferring to escape arm, regardless of use of topical rescue medication and if all participants continued with treatment except those who discontinued due to lack of efficacy?

The maintenance supportive estimand is described by the following attributes:

A. Population: MPP as described in Section 3

B. Endpoint: apply to key categorical and continuous endpoints (i.e., IGA, EASI and Itch NRS related endpoints) for the Maintenance Period

C. How to account for ICEs

   a. Participants who require any use of systemic rescue medication, discontinue treatment due to lack of efficacy after Week 16, or transfer to escape arm will be considered as treatment failures, i.e., non-responder, after the ICEs. Therefore, composite strategy is used for these types of ICEs.

   b. For participants who require any use of topical rescue medication, observed data will be used. Therefore, treatment policy strategy is used for these types of ICEs.

   c. For participants who discontinue treatment due to reasons other than lack of efficacy after week 16, a hypothetical strategy will be used to estimate what the response proportions or means for each treatment condition would have been if participants continued with treatment. Therefore, hypothetical strategy is used for these types of ICEs.

D. Population-level summary: response proportions or means for each treatment condition

1.1.1.2.3. Maintenance Supportive Estimand for Categorical Endpoints (Composite)

The maintenance supportive estimand for categorical endpoints only is a composite estimand representing the clinical question of interest: what is the response proportions for each treatment condition (i.e., Lebrikizumab 250 mg Q2W and Lebrikizumab 250 mg Q4W), in the target participant population, in successful responses after 68 weeks achieved without use of topical or systemic rescue medication, treatment discontinuation or transferring to escape arm?

The maintenance supportive estimand is described by the following attributes:

A. Population: MPP as described in Section 3

B. Endpoint: apply to all categorical endpoints for the Maintenance Period

C. How to account for ICEs

   a. Participants who require any use of topical or systemic rescue medication, discontinue treatment after Week 16, or transfer to escape arm will be considered as treatment failures, i.e., non-responder, after the ICEs. Therefore, composite strategy is used for these types of ICEs.

D. Population-level summary: response proportions for each treatment condition

1.1.1.2.4. Maintenance Supportive Estimand for Continuous Endpoints (Hypothetical)

The maintenance supportive estimand for continuous endpoints only is a hypothetical estimand representing the clinical question of interest: what is the means for each treatment condition (i.e., Lebrikizumab 250 mg Q2W and Lebrikizumab 250 mg Q4W), in the target participant population, in means after 68 weeks if rescue medication was not available and all participants adhered to the treatment and did not transfer to escape arm?

The maintenance supportive estimand is described by the following attributes:

A. Population: MPP as described in Section 3

B. Endpoint: apply to all continuous endpoints for the Maintenance Period

C. How to account for ICEs

   a. For participants who require any use of rescue medication, discontinue treatment after Week 16, or transfer to escape arm, a hypothetical strategy will be used to estimate what the means for each treatment condition would have been if rescue medication was not available and all participants adhered to the treatment and did not transfer to escape arm. Therefore, hypothetical strategy is used for these types of ICEs.

D. Population-level summary: means for each treatment condition

Details on how missing data including those as a result of intercurrent events will be handled can be found in Section 4.1.6.2. Detailed analyses relative to estimands including analysis type, method and imputation, population, time point, and treatment comparisons for efficacy and health outcome analyses can be found in Appendix 2.

The table below summarizes the analytical strategies that will be conducted on the intercurrent events for the four maintenance estimands. Intercurrent events and missing data will be summarized by treatment group for the Induction Period on the ITT population, and for the Maintenance Blinded Period on the MPP, respectively.

Description of Primary and Supportive Estimands for Maintenance Period

Maintenance Estimand	Analysis Strategy for Intercurrent Events				Transfer to Escape Arm	Missing Data Imputation Method
	Rescue Medication		Treatment Discontinuation			
	Topical rescue medication	Systemic rescue medication	Due to lack of efficacy	Due to any other reasons		
Maintenance Primary Estimand (Hybrid)	Hypothetical: Set to missing	Composite: Set to baseline	Composite: Set to baseline	Hypothetical: Set to missing	Composite: Set to baseline	Primary analysis: MCMC-MI
Maintenance Supportive Estimand (Hybrid)	Treatment policy: As observed	Composite: Set to baseline	Composite: Set to baseline	Hypothetical: Set to missing	Composite: Set to baseline	Supplementary analysis: MCMC-MI
Maintenance Supportive Estimand for Categorical Endpoints (Composite)	Composite: Set to non-responder	Composite: Set to non-responder	Composite: Set to non-responder	Composite: Set to non-responder	Composite: Set to non-responder	Supplementary analysis: NRI
Maintenance Supportive Estimand for Continuous	Hypothetical: Set to missing	Hypothetical: Set to missing	Hypothetical: Set to missing	Hypothetical: Set to missing	Hypothetical: Set to missing	Supplementary analysis: LOCF

Endpoints (Hypothetical)						

Abbreviations: LOCF = last observation carried forward; MCMC-MI = Markov chain Monte Carlo multiple imputation; NRI = non-responder imputation.

1.2. Study Design

Study Design

Study KGAL is a randomized, double-blind, placebo-controlled, parallel-group study, which is 68 weeks in treatment duration. The study is designed to evaluate the safety and efficacy of lebrikizumab when used in combination with topical corticosteroid (TCS) treatment compared with placebo in combination with TCS treatment using a 16-week induction treatment and a 52-week long-term Maintenance Period of lebrikizumab for Japanese participants with moderate-to-severe atopic dermatitis (AD).

Study Population

Participants are eligible for the study if they

- are ≥18 years or an adolescent (≥12 to <18 years of age and weighing ≥40 kg) with moderate-to-severe AD for at least 1 year, defined according to the American Academy of Dermatology Consensus Criteria (Eichenfield et al. 2014)
- have an EASI score ≥16
- have an IGA score ≥3, and
- have an AD involvement in ≥10% of Body Surface Area (BSA)

Investigators should review vaccination status in adolescent participants (≥12 to <18 years) and determine the benefit/risk of their participation. No changes in start or resumption of vaccine is required, and pediatric participants will follow the Japan immunization guidelines, with the exception of live vaccines. If a participant has received a live vaccine within 12 weeks of the baseline visit or intends to receive a live vaccine during the study or up to 125 days after the last dose of investigational product (IP), the participant is not eligible for the study (see exclusion criterion [29] of the protocol).

Induction Period

During the 16-week Induction Period, approximately 280 participants, including approximately 15 adolescent participants (≥12 to <18 years and weighing ≥40 kg), will be randomized in a 3:2:2 ratio to treatment:

- 250 mg lebrikizumab (loading dose of 500 mg given at baseline and Week 2) by SC injection Q2W
- 250 mg lebrikizumab (loading dose of 500 mg given at baseline) by SC injection Q4W, or
- placebo.

Participants will be stratified at randomization according to age (adolescent participants ≥12 to <18 years vs ≥18 years) and disease severity (IGA 3 vs 4). Daily use of mid-potency TCS (low-potency TCS and/or topical calcineurin inhibitor [TCI] for sensitive areas) will be initiated at least 7 days prior to baseline in all participants (not allowed to be tapered or stopped during the Screening Period). Mid-potency TCS, low-potency TCS, and TCIs may be tapered or stopped after baseline, as needed, based on treatment response.

All study intervention injections during the Induction Period will be administered by site staff at the clinic.

Maintenance Period

After completion of the Week 16 visit:

- Participants receiving 250 mg lebrikizumab Q2W
  i. who achieve an IGA score of 0 or 1 and/or a ≥75% reduction in EASI score (≥EASI-75) response at Week 16 will be randomly allocated to receive 250 mg lebrikizumab Q2W or 250 mg lebrikizumab Q4W, in a 1:1 fashion;
  ii. who achieve neither an IGA score of 0 or 1 nor a <75% reduction in EASI score (<EASI-75) response at Week 16 will move to the Escape Arm.

- Participants receiving 250 mg lebrikizumab Q4W

  i. who achieve an IGA score of 0 or 1 and/or an EASI-75 response (≥EASI-75) at Week 16 will continue 250 mg lebrikizumab Q4W;
  ii. who achieve neither an IGA score of 0 or 1 nor an EASI-75 response (<EASI-75) at Week 16 will move to the Escape Arm.

- Participants receiving placebo

  i. who achieve an IGA score of 0 or 1 and/or an EASI-75 response (≥EASI-75) at Week 16 will continue to receive placebo;
  ii. who achieve neither an IGA score of 0 or 1 nor an EASI-75 response (<EASI-75) at Week 16 will move to the Escape Arm and will receive a loading dose of 500 mg lebrikizumab at Week 16 and Week 18.

Treatment regimen assignment will remain blinded during the Maintenance Period. Placebo injections will be utilized in order to maintain the blind and ensure that all participants receive the same number and frequency of injections regardless of treatment regimen assignment.

Escape Arm

In the Escape Arm, participants will receive 250 mg lebrikizumab Q2W.

Escape Arm from Week 16:

- Participants who achieve neither an IGA score of 0 or 1 nor an EASI-75 response (<EASI-75) at Week 16 will move to the Escape Arm.
- Participants who receive rescue treatment in the Induction Period will also be eligible to continue to the Escape Arm at Week 16.
- Participants who required systemic rescue medication in the Induction Period must discontinue study intervention and must wait for the rescue medication washout (≥5 half-lives of the medication) prior to entering the Escape Arm.
- Only participants who receive placebo in the Induction Period will receive a loading dose of 500 mg lebrikizumab at Week 16 and Week 18 followed by 250 mg lebrikizumab Q2W. Participants not achieving an EASI-50 response at 2 consecutive visits in the Escape Arm after Week 32 of treatment will be discontinued from the study.

Escape Arm after Week 20:

- Participants who do not maintain an EASI-50 response at Week 20, 24, 28, 32, 36, 40, 44, 48, 52, 56, 60 or 64 in the Maintenance Period will be assigned to an Escape Arm and receive lebrikizumab 250 mg Q2W as open-label treatment thorough Week 68.
- Participants not achieving an EASI-50 response at 2 consecutive visits in the Escape Arm after 8 weeks of treatment will be discontinued from the study.

During the Maintenance Period, participants will be instructed to self-administer study intervention. Administration by the participant or caregiver is recommended. If the participant or caregiver is not able to administer any dose throughout the study, study site staff may administer the injection.

- At Week 16, site personnel will instruct participants or their caregiver on the proper injection technique.
- At Week 16 and Week 18, participants or their caregiver will administer study intervention under site personnel supervision.
- Subsequent administration will be administered by participants or their caregiver at home.

Efficacy, Health Outcomes, and Safety Assessments

Efficacy and health outcomes will be measured through

- IGA
- EASI
- SCORing Atopic Dermatitis (SCORAD)
- BSA involvement
- Itch NRS
- Sleep-Loss Scoring System
- Skin Pain NRS
- TCS/TCI use
- Columbia-Suicide Severity Rating Scale (C-SSRS) self-harm assessment
- Quality of life and impact of disease (these will be assessed using the Patient-Oriented Eczema Measure [POEM], Dermatology Life Quality Index/Children's Dermatology Life Quality Index [DLQI/CDLQI], Work Productivity and Activity Impairment – Atopic Dermatitis [WPAI-AD], and Hospital Anxiety Depression Scale [HADS]), and
- Asthma Control Questionnaire-5 (ACQ-5) (to be completed by participants reporting comorbid asthma at study entry).

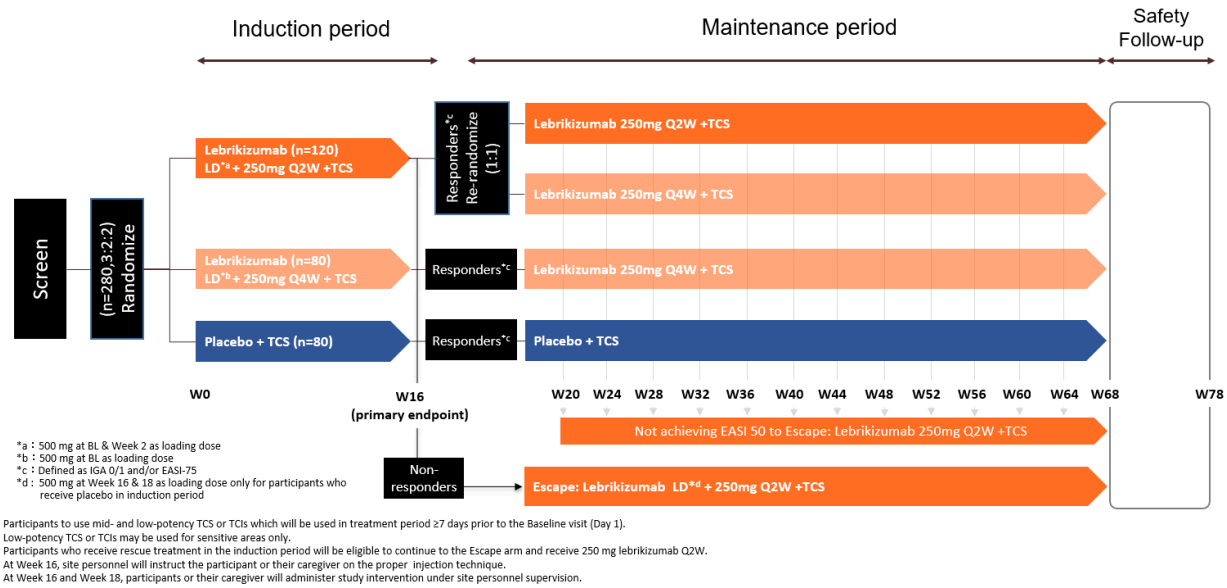
Safety will be assessed in all participants by

- Adverse event (AE) monitoring
- serum chemistry
- hematology
- urinalysis laboratory testing
- physical examination
- pulse and blood pressure (vital signs)

- chest x-ray
- concomitant medications, and
- monitoring of hormone levels and growth parameters (adolescents only).

Serum samples will be collected for pharmacokinetics (PK) analysis and immunogenicity.

Participants who terminate the study early or complete the 68-week treatment period will receive a safety follow-up visit at which vital signs will be measured, C-SSRS self-harm supplement and self-harm follow-up form will be completed, and PK and immunogenicity samples will be collected.



2. Statistical Hypotheses

The following is a list of primary and major secondary endpoints to be tested. The subscript for **H** denotes study intervention arms in the comparisons (l2 = lebrikizumab 250 mg Q2W, l4 = lebrikizumab 250 mg Q4W and p = placebo), the numerical identifier of the endpoint within the comparison, and the type of hypothesis (0 for null, 1 for alternative), respectively.

Co-Primary Null Hypotheses:

- $\mathbf{H_{l2p,1,0}}$: Proportion of lebrikizumab 250 mg Q2W participants achieving EASI-75 at Week 16 is equal to the proportion of placebo participants achieving EASI-75 at Week 16
- $\mathbf{H_{l2p,2,0}}$: Proportion of lebrikizumab 250 mg Q2W participants achieving IGA score of 0 or 1 and a reduction of ≥2-points from baseline to Week 16 is equal to the proportion of placebo participants achieving IGA score of 0 or 1 and a reduction of ≥2-points from baseline to Week 16
- $\mathbf{H_{l4p,1,0}}$: Proportion of lebrikizumab 250 mg Q4W participants achieving EASI-75 at Week 16 is equal to the proportion of placebo participants achieving EASI-75 at Week 16
- $\mathbf{H_{l4p,2,0}}$: Proportion of lebrikizumab 250 mg Q4W participants achieving IGA score of 0 or 1 and a reduction of ≥2-points from baseline to Week 16 is equal to the proportion of placebo participants achieving IGA score of 0 or 1 and a reduction of ≥2-points from baseline to Week 16

Major Secondary Null Hypotheses:

- $\mathbf{H_{l2p,3,0}}$: Percentage change of EASI score from baseline to Week 16 in lebrikizumab 250 mg Q2W is equal to the percentage change of EASI score from baseline to Week 16 in placebo
- $\mathbf{H_{l2p,4,0}}$: Proportion of lebrikizumab 250 mg Q2W participants achieving EASI-90 at Week 16 is equal to the proportion of placebo participants achieving EASI-90 at Week 16
- $\mathbf{H_{l2p,5,0}}$: Proportion of lebrikizumab 250 mg Q2W participants with an Itch NRS score of ≥4-points at baseline who achieve a ≥4-point reduction from baseline to Week 16 is equal to the proportion of placebo participants with an Itch NRS score of ≥4-points at baseline who achieve a ≥4-point reduction from baseline to Week 16
- $\mathbf{H_{l2p,6,0}}$: Proportion of lebrikizumab 250 mg Q2W participants with an Itch NRS score of ≥4-points at baseline who achieve a ≥4-point reduction from baseline to Week 4 is equal to the proportion of placebo participants with an Itch NRS score of ≥4-points at baseline who achieve a ≥4-point reduction from baseline to Week 4
- $\mathbf{H_{l2p,7,0}}$: Proportion of lebrikizumab 250 mg Q2W participants with an Itch NRS score of ≥4-points at baseline who achieve a ≥4-point reduction from baseline to Week 2 is equal to the proportion of placebo participants with an Itch NRS score of ≥4-points at baseline who achieve a ≥4-point reduction from baseline to Week 2
- $\mathbf{H_{l2p,8,0}}$: Proportion of lebrikizumab 250 mg Q2W participants with an Itch NRS score of ≥4-points at baseline who achieve a ≥4-point reduction from baseline to Week 1 is equal to the proportion of placebo participants with an Itch NRS score of ≥4-points at baseline who achieve a ≥4-point reduction from baseline to Week 1
- $\mathbf{H_{l4p,3,0}}$: Percentage change of EASI score from baseline to Week 16 in lebrikizumab 250 mg Q4W is equal to the percentage change of EASI score from baseline to Week 16 in placebo

- **$H_{l4p,4,0}$**: Proportion of lebrikizumab 250 mg Q4W participants achieving EASI-90 at Week 16 is equal to the proportion of placebo participants achieving EASI-90 at Week 16
- **$H_{l4p,5,0}$**: Proportion of lebrikizumab 250 mg Q4W participants with an Itch NRS score of ≥4-points at baseline who achieve a ≥4-point reduction from baseline to Week 16 is equal to the proportion of placebo participants with an Itch NRS score of ≥4-points at baseline who achieve a ≥4-point reduction from baseline to Week 16
- **$H_{l4p,6,0}$**: Proportion of lebrikizumab 250 mg Q4W participants with an Itch NRS score of ≥4-points at baseline who achieve a ≥4-point reduction from baseline to Week 4 is equal to the proportion of placebo participants with an Itch NRS score of ≥4-points at baseline who achieve a ≥4-point reduction from baseline to Week 4
- **$H_{l4p,7,0}$**: Proportion of lebrikizumab 250 mg Q4W participants with an Itch NRS score of ≥4-points at baseline who achieve a ≥4-point reduction from baseline to Week 2 is equal to the proportion of placebo participants with an Itch NRS score of ≥4-points at baseline who achieve a ≥4-point reduction from baseline to Week 2
- **$H_{l4p,8,0}$**: Proportion of lebrikizumab 250 mg Q4W participants with an Itch NRS score of ≥4-points at baseline who achieve a ≥4-point reduction from baseline to Week 1 is equal to the proportion of placebo participants with an Itch NRS score of ≥4-points at baseline who achieve a ≥4-point reduction from baseline to Week 1

2.1. Multiplicity Adjustment

A prespecified graphical multiple testing approach (Bretz et al. 2009, 2011) will be implemented to control the overall type I error rate at a 2-sided alpha of 0.05 for superiority tests of the hypotheses for the co-primary and key secondary endpoints. Multiple testing adjusted p-values using "Algorithm 2" described by Bretz et al. (2009) will be calculated, and any hypothesis tests with a multiple testing adjusted p-value of less than 0.05 will be considered statistically significant. This graphical approach is a closed testing procedure; hence, it strongly controls the family-wise error rate across all endpoints (Alosh et al. 2014). Each hypothesis is represented as a node in a graph. Directed arrows between the nodes with associated weights represent how alpha is passed from its initial allocation to other nodes.

The following is a list of co-primary and key secondary endpoints to be tested.

Co-Primary endpoints:

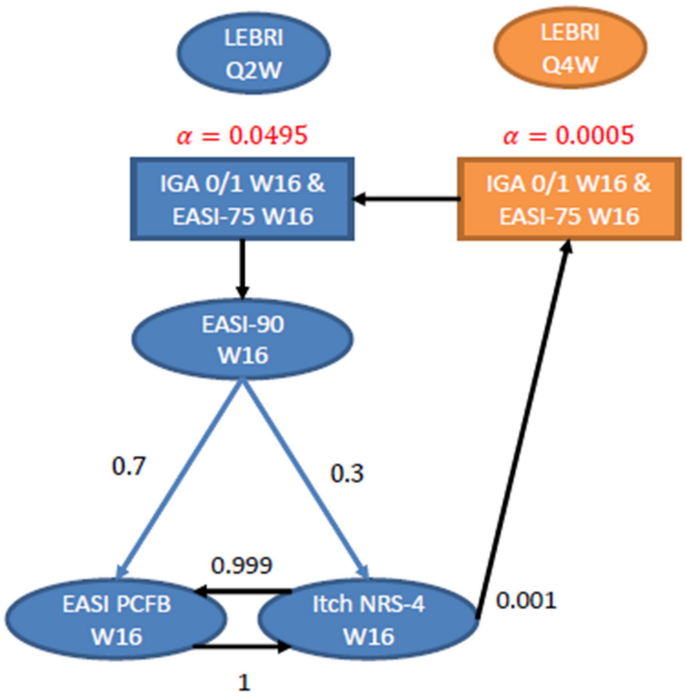
- [EASI-75 W16] Percentage of participants achieving EASI-75 (≥ 75% reduction from Baseline in EASI score) at Week 16 in lebrikizumab 250 mg Q2W versus placebo or lebrikizumab 250 mg Q4W versus placebo.
- [IGA0/1 W16] Percentage of participants with IGA score of 0 or 1 and a reduction ≥ 2 points from Baseline to Week 16 in lebrikizumab 250 mg Q2W versus placebo or lebrikizumab 250 mg Q4W versus placebo.

Key secondary endpoints:

- [EASI PCFB W16] Percentage change in EASI score from Baseline to Week 16 in lebrikizumab 250 mg Q2W versus placebo.
- [EASI-90 W16] Percentage of participants achieving EASI-90 (≥ 90% reduction from Baseline in EASI score) at Week 16 in lebrikizumab 250 mg Q2W versus placebo.

- [Itch NRS-4 W16] Percentage of participants with Itch NRS of ≥ 4-point at Baseline who achieve a ≥ 4-point reduction from Baseline to Week 16 in lebrikizumab 250 mg Q2W versus placebo.

The figure below describes the graphical testing scheme.



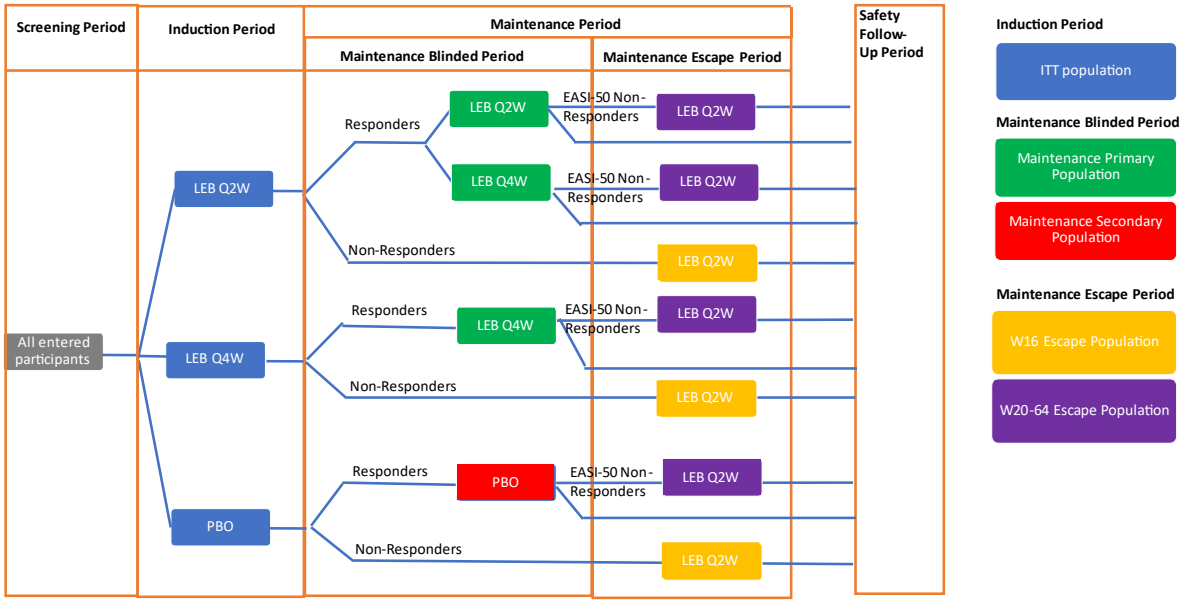
3. Analysis Sets

The following analysis populations are defined:

Population	Description
All Entered Participants	All participants who signed informed consent/assent. Participant flow will be summarized.
Intent-to-Treat (ITT)	All randomized participants, even if the participant does not take the assigned treatment, does not receive the correct treatment, or otherwise does not follow the protocol. Participants will be analyzed according to the treatment group to which they were assigned. Unless otherwise specified, efficacy and health outcome analyses for the Induction Period will be conducted on this population.
Safety Population	All randomized participants who received at least 1 dose of study intervention during the Induction Period. Safety analyses for the Induction Period will be conducted on this population.
Per-Protocol Set (PPS)	All ITT participants who do not have a subset of important protocol deviations that impact the primary efficacy endpoint. Important protocol deviations are defined in a separate document referred to as "The KGAL Trial Issues Management Plan." Primary efficacy analysis for IGA score of 0 or 1 and EASI-75 will be repeated using the PPS.
Maintenance Primary Population (MPP)	All participants who were randomized to lebrikizumab 250 mg Q2W or lebrikizumab 250 mg Q4W at baseline visit (Visit 2, Week 0) and met the response criteria (that is, participants who achieved an IGA score of 0 or 1 and/or EASI-75 at Week 16) and received at least 1 dose of study intervention during the Maintenance Period. Participants will be analyzed according to the treatment to which they were assigned. Only information prior to entering the Escape Arm will be presented. Efficacy, health outcome, and safety analyses for the Maintenance Period will be conducted primarily on the MPP.
Maintenance Secondary Population (MSP)	Participants who were randomized to placebo at baseline visit (Visit 2, Week 0) and continued placebo (that is, participants who achieved an IGA score of 0 or 1 and/or EASI-75 at Week 16) and received at least 1 dose of study intervention during the Maintenance Period. Only information prior to entering the Escape Arm will be presented. Selective efficacy analyses for the Maintenance Period will be conducted on the MSP.

Maintenance W16 Escape Population	Participants who moved to the Escape Arm at Week 16 (either due to not meeting IGA and EASI criteria, or used rescue therapy during the Induction Period), and received at least one dose of study intervention during the Maintenance Period. Selective efficacy analyses for the Maintenance Period will be conducted on the Maintenance W16 Escape Population.
Maintenance W20-64 Escape Population	Participants in MPP or MSP who escaped to lebrikizumab 250 mg Q2W due to EASI-50 non-response at Weeks 20, 24, 28, 32, 36, 40, 44, 48, 52, 56, 60 or 64. Selective efficacy analyses for the Maintenance Period will be conducted on the Maintenance W20-64 Escape Population to assess whether these participants re-gain EASI-50 response or achieve a higher level of response (for example, EASI-75) following re-treatment.
All Maintenance Population	Participants who entered the Maintenance Period and received at least one dose of study intervention during the Maintenance Period.
All Lebrikizumab Safety Population	All randomized participants who received at least 1 dose of lebrikizumab treatment during the Combined Induction and Maintenance Periods including participants who were randomized to placebo at baseline visit (Visit 2, Week 0) and moved to the Escape Arm at Week 16 or later visits in the Maintenance Period. Safety analyses for the Combined Induction and Maintenance Periods will be conducted on the All lebrikizumab Safety Population. Selective safety analyses for the Combined Induction and Maintenance Periods plus the Follow-Up Period will be conducted on the All Lebrikizumab Safety Population.
Pharmacokinetic Analysis	All participants who received at least 1 dose of lebrikizumab and have at least 1 evaluable PK sample.

Abbreviations: EASI-75 = ≥75% reduction from baseline in Eczema Area and Severity Index IGA = Investigator's Global Assessment; Q2W = every 2 weeks; Q4W = every 4 weeks; W = Week.


Screening Period
Induction Period
Maintenance Period
Maintenance Blinded Period
Maintenance Escape Period
Safety Follow-Up Period
All entered participants
LEB Q2W
LEB Q4W
PBO
Responders
Non-Responders
LEB Q2W
LEB Q4W
PBO
EASI-50 Non-Responders
LEB Q2W
Induction Period
ITT population
Maintenance Blinded Period
Maintenance Primary Population
Maintenance Secondary Population
Maintenance Escape Period
W16 Escape Population
W20-64 Escape Population

Treatment Groups and Comparisons for Each Study Period and Analysis Population

Study Period	Analysis Population	Treatment Groups	Abbreviation	Inferential Comparisons When Applicable
Induction Period	ITT; Safety; PPS	Placebo; Lebrikizumab 250 mg Q4W; Lebrikizumab 250 mg Q2W; Total	PBO; LEB250Q4W; LEB250Q2W; Total	LEB250Q4W vs PBO; LEB250Q2W vs PBO
Maintenance Blinded Period	MPP	Lebrikizumab 250 mg Q4W_Res/ Lebrikizumab 250 mg Q4W; Lebrikizumab 250 mg Q2W_Res/ Lebrikizumab 250 mg Q4W; Lebrikizumab 250 mg Q2W_Res/ Lebrikizumab 250 mg Q2W; Total	LEB250Q4W_Res/ LEB250Q4W; LEB250Q2W_Res/ LEB250Q4W; LEB250Q2W_Res/ LEB250Q2W; Total	No Between-Group or Overall Comparisons
Maintenance Blinded Period	MSP	Placebo_Res/Placebo	PBO_Res/PBO	No Between-Group or Overall Comparisons
Maintenance Escape Period	Maintenance W16 Escape Population	Placebo_NonResp/Lebrikizumab 250 mg Q2W; Lebrikizumab 250 mg Q4W_NonResp/Lebrikizumab 250 mg Q2W; Lebrikizumab 250 mg Q2W_NonResp/Lebrikizumab 250 mg Q2W; Total	PBO_NonResp/ LEB250Q2W; LEB250Q4W_NonResp/ LEB250Q2W; LEB250Q2W_NonResp/ LEB250Q2W; Total	No Between-Group or Overall Comparisons

Maintenance Escape Period	Maintenance W20-64 Escape Population	Placebo/Placebo/Lebrikizumab 250 mg Q2W; Lebrikizumab 250 mg Q4W/ Lebrikizumab 250 mg Q4W/ Lebrikizumab 250 mg Q2W; Lebrikizumab 250 mg Q2W/ Lebrikizumab 250 mg Q4W/ Lebrikizumab 250 mg Q2W; Lebrikizumab 250 mg Q2W/ Lebrikizumab 250 mg Q2W/ Lebrikizumab 250 mg Q2W; Total	PBO/PBO/LEB250Q2W; LEB250Q4W/ LEB250Q4W/ LEB250Q2W; LEB250Q2W/ LEB250Q4W/ LEB250Q2W; LEB250Q2W/ LEB250Q2W/ LEB250Q2W;	No Between-Group or Overall Comparisons
Combined Induction and Maintenance Periods	All Lebrikizumab Safety Population	Any Lebrikizumab	N/A	No Between-Group or Overall Comparisons
Combined Induction and Maintenance Periods + FU	All Lebrikizumab Safety Population	Any Lebrikizumab	N/A	No Between-Group or Overall Comparisons

Abbreviations: FU = follow-up; ITT = intent-to-treat; LEB = lebrikizumab; NonResp = non-responder; PBO = placebo; PPS = per protocol set; Q2W = every 2 weeks; Q4W = every 4 weeks; Res = responder.

4. Statistical Analyses

4.1. General Considerations

Statistical analysis of this study will be the responsibility of Eli Lilly Japan K.K. (Lilly). The latest version of the Medical Dictionary for Regulatory Activities (MedDRA) will be used.

Analyses and summaries from assessment of endpoints described in the protocol are planned to be included in a clinical study report (CSR). Analyses and summaries for key safety data are also planned to be included in the CSR. Results from additional efficacy analysis and other safety analyses may also be provided in the CSR as deemed appropriate.

Not all displays described in the SAP will necessarily be included in the CSRs. Any display described and not provided in the CSR would be available upon request. Not all displays will necessarily be created as a "static" display. Some may be incorporated into interactive display tools instead of or in addition to a static display. Any display created interactively will be included in the CSR if deemed relevant to the discussion.

All statistical processing will be performed using SAS® unless otherwise stated. Except where noted, all statistical tests will be two-sided and will be performed at the 0.05 level of significance.

The Schedule of Visits and Procedures outlined in the protocol specifies the allowable windows for assessments. Assessments performed outside these windows will not be excluded from any analysis, unless specified otherwise.

Any change to the data analysis methods described in the protocol will require a protocol amendment ONLY if it changes a principal feature of the protocol. Any other change to the data analysis methods described in the protocol and the justification for making the change will be described in the CSR.

4.1.1. General Considerations for Analyses During Induction Period

Induction Period starts after the first injection of study intervention at Baseline Visit (Visit 2, Week 0) and ends prior to the first injection of study intervention at Week 16 or the early termination visit (ETV) (between Week 0 and Week 16). For participants who were randomized but discontinued the study before the first injection, the Induction Period starts on the date of randomization. For participants who are were randomized but received the first injection after the date of randomization, the Induction Period starts on the date of first injection.

Baseline will be defined as the last available value before the first injection for efficacy and health outcome analyses. In most cases, this will be the measure recorded at Baseline Visit (Week 0). If the participant does not take any injection, the last available value on or prior to randomization date will be used. Change from baseline will be calculated as the visit value of interest minus the baseline value.

For Itch Numeric Rating Scale (NRS), Skin Pain NRS and Sleep-Loss due to Itch collected via electronic Clinical Outcome Assessment (eCOA), the baseline period is the 7-day window prior to the first injection. A participant must have responses on at least 4 of 7 days to calculate a

baseline weekly mean. If a participant has responses for 3 or less days, the baseline mean value will be considered missing. eCOA data for Itch NRS, Skin Pain NRS and Sleep-loss due to Itch are mapped to study visit per Appendix 12.

For the safety analyses, the following baselines will be used. For safety analyses using a baseline period, the baseline period is defined as the time from Screening Visit (Visit 1) to the date/time of the first injection in the Induction Period.

- Treatment-emergent adverse events (TEAEs): baseline will be all results (including medical histories which are ongoing at the date of informed consent) recorded during the baseline period.
- Treatment-emergent abnormal laboratory and vital signs results: baseline will be all results recorded during the baseline period.
- Change from baseline to last post-baseline observation or to each scheduled post-baseline visit for laboratory and vital signs results: baseline will be the last scheduled non-missing assessment recorded during the baseline period.

The randomization to treatment groups is stratified by age (adolescent participants 12 to <18 years versus adults ≥18 years) and baseline disease severity (IGA 3 versus 4). Unless otherwise specified, the statistical analysis models for the Induction Period will adjust for age and baseline disease severity.

For assessments of the primary endpoints and other binary efficacy and health outcome endpoints, the following will be provided:

- Crude proportions for each treatment group along with the 95% two-sided asymptotic (that is, not continuity corrected) confidence intervals (CIs).
- The estimated common risk difference along with 95% CIs: The common risk difference is the difference in proportions adjusted for the stratification factors. SAS® PROC FREQ will be used for the estimates and CIs, where the CIs are calculated by using Mantel-Haenszel-Sato method (Sato 1989).
- Cochran-Mantel-Haenszel (CMH) test will be used to compare the treatment groups while adjusting for the stratification factors. The CMH p-value will be reported, and the CMH adjusted odds ratio (OR) along with the 95% two-sided asymptotic (that is, not continuity corrected) CIs.

Treatment comparisons of key continuous efficacy variables and health outcome variables at each post-baseline time point will be made using analysis of covariance (ANCOVA) with the following in the model: treatment group, baseline value, and stratification factors. Type III tests for least squares (LS) means will be used for statistical comparison between treatment groups. The LS mean difference, standard error (SE), p-value, and 95% CI, unless otherwise specified, will also be reported.

Treatment comparisons of other continuous efficacy variables and health outcome variables with multiple post-baseline measurements will be made using mixed-model for repeated measures (MMRM). When MMRM is used, the model includes treatment, baseline value, visit, the interaction of the baseline value-by-visit, the interaction of treatment-by-visit, and the stratification factors as fixed effects. The covariance structure to model the within-participant

errors will be unstructured. If this analysis fails to converge, the heterogeneous autoregressive [ARH(1)], followed by the heterogeneous compound symmetry (CSH), followed by the heterogeneous Toeplitz (TOEPH), followed by the autoregressive [AR(1)], followed by the compound symmetry (CS) will be used. The restricted maximum likelihood (REML) will be used. The Kenward-Roger method will be used to estimate the denominator degrees of freedom. Type III tests for the LS means will be used for the statistical comparison; the 95% CI will also be reported.

For variables that are not collected at each post-baseline visit, data may exist at visits where the variable was not scheduled to be collected. In these situations, data from the early discontinuation visit that do not correspond to the planned collection schedule will be excluded from the MMRM analysis (Andersen and Millen 2013). Also for by-visit summaries/displays such as boxplots, the weeks when data was not scheduled to be collected may not be displayed. However, unscheduled assessments within any defined study period will still be used in the shift analyses, and for imputing values for the change from baseline to last observation carried forward (LOCF) endpoint analyses.

The Kaplan-Meier product limit method may be used to estimate the survival for time to event analyses. The log-rank test stratified by the stratification factors will be reported. A Kaplan-Meier plot of the time to event by treatment group may be provided.

Unless specified otherwise, Fisher's exact test will be used for AEs and other categorical safety measures. ORs will be created with lebrikizumab treatment as the numerator, and placebo as the denominator. Continuous vital signs and laboratory values will be analyzed by an ANCOVA with treatment and baseline value in the model.

4.1.2. General Considerations for Analyses During Maintenance Period

Maintenance Period starts at the first injection of study intervention at Week 16 and ends on the date of Week 68 or the ETV (between Weeks 16 and 68) unless specified otherwise.

For the efficacy and health outcome analyses, baseline is defined as the last available value before the first injection in the Induction Period and, in most cases, will be the value recorded at Baseline Visit (Visit 2, Week 0).

Unless otherwise specified, efficacy and health outcome variables at Week 16 prior to entering the Maintenance Period will be presented for the visit-wise reports for the Maintenance Period.

Unless specified otherwise, for the safety analyses during the Maintenance Period, baseline is defined as the last available value before the first injection in the Maintenance Period. In most cases, this will be the measure recorded at Week 16. For TEAEs, baseline is the events ongoing just prior to the first injection of the study drug injection at Week 16.

For participants in the MPP and MSP who met escape criteria (EASI-50 non-response) and escaped to lebrikizumab 250 mg Q2W at Weeks 20, 24, 28, 32, 36, 40, 44, 48, 52, 56, 60 and 64, only data in the Maintenance Blinded Period (up to the time of escape) will be included in both efficacy and safety analyses.

4.1.2.1. Maintenance Primary and Secondary Population

The number and percentage of participants achieving or maintaining categorical efficacy and health outcome responses will be summarized by treatment group for all scheduled visits, including Week 68.

Each continuous efficacy and health outcome measure score and the change from baseline (or percentage improvement) will be summarized by treatment group at all scheduled visits during the Maintenance Period, including Week 68 using descriptive statistics (n, mean, SD, median, minimum, and maximum). No inferential statistics will be provided for these populations.

For the MPP, the KM product limit method will be used to estimate the survival for time to event analyses (e.g., time to loss of IGA score of 0 or 1, or loss of EASI-50). A KM plot of the time to event by treatment group may be provided.

4.1.2.2. Maintenance Escape Population

For the Maintenance W16 Escape Population, the number and percentage of participants achieving or maintaining selected categorical efficacy and health outcome responses will be summarized by treatment group for all scheduled visits, including Week 68. Selected continuous secondary efficacy and health outcome measures score and the change from baseline (or percentage improvement) will be summarized by treatment group at all scheduled visits during the Maintenance Period, including Week 68 using descriptive statistics (n, mean, SD, median, minimum, and maximum). No inferential statistics will be provided for this population.

For the Maintenance W20-64 Escape Population who were treated with lebrikizumab 250 mg Q2W following loss of response (EASI-50 non-response), the number and percentage of participants regaining EASI-50 response or achieving EASI-75 will be summarized every 4 weeks after lebrikizumab 250 mg Q2W treatment. No inferential statistics will be provided for this population.

4.1.3. General Considerations for Safety Analyses for Combined Induction and Maintenance Periods

Adverse event, exposure summary, and categorical laboratory/vital signs changes will be provided for the All Lebrikizumab Safety Population during the Combined Induction and Maintenance Periods. For participants who were first exposed to lebrikizumab during the Induction Period, the baseline for TEAE will utilize the baseline for the Induction Period defined in Section 4.1.1; for participants who were first exposed to lebrikizumab during the Maintenance Period, the baseline for TEAE will utilize the baseline for the Maintenance Period defined in Section 4.1.2.

More details on baseline and post-baseline definitions can be found in the Compound Level Safety Standard.

4.1.4. General Considerations for Safety Analyses for Combined Induction and Maintenance Periods Plus Follow-Up Period

Selective AE summaries will be provided for the All Lebrikizumab Safety Population during the Combined Induction and Maintenance Periods plus Follow-Up Period. The baseline definition

for this population is the same as Section 4.1.3. More details on baseline and post-baseline definitions can be found in the Compound Level Safety Standard.

4.1.5. Adjustments for Covariates

Unless otherwise specified, the statistical analysis models for the Induction Period efficacy and health outcome analysis will include the following stratification factors for Baseline randomization: age (adolescent participants 12 to <18 versus adults ≥18 years) and baseline disease severity (IGA 3 versus 4).

In general, when an MMRM is to be used for analyses, baseline value and baseline-by-visit interactions will be included as covariates; when an ANCOVA is to be used for analyses, baseline value will be included as a covariate.

4.1.6. Handling of Dropouts or Missing Data

Depending on the estimand being addressed, different methods will be used to handle missing data. Description of the estimands can be found in Section 1.1.1.

4.1.6.1. Handling of Dropouts or Missing Data for Induction Period

For efficacy analysis relative to the primary estimand, missing data including those as a result of intercurrent events will be imputed based on Markov chain Monte Carlo Multiple Imputation (MCMC-MI). The description of MCMC-MI method can be found in Section 4.1.6.1.1. Tipping point analysis as described in Section 4.1.6.1.2 will serve as the sensitivity analysis for the primary analysis.

For efficacy analysis relative to the supportive estimand for categorical endpoints, missing data including those as a result of intercurrent events will be imputed as non-responder. The description of non-responder imputation (NRI) can be found in Section 4.1.6.1.3.

For efficacy analysis relative to the supportive estimand for continuous endpoints collected multiple times post-baseline, a Mixed-effects Model for Repeated Measures (MMRM) will be performed without explicit imputation. The description of MMRM can be found in Section 4.1.6.1.4.

For efficacy analysis relative to the supportive estimand for continuous endpoints collected only once post-baseline, missing data including those as a result of intercurrent events will be imputed using Last Observation Carried Forward (LOCF). The description of LOCF can be found in Section 4.1.6.1.5.

The table below describes the planned imputation methods for efficacy and health outcome endpoints.

Imputation Techniques for Various Variables during Induction Period

Type of Endpoints	Efficacy and Health Outcome Endpoints	Estimand (Analysis strategy for Intercurrent Events)	Missing Data Imputation Method (Analysis Method)
Categorical		Primary Estimand (Hybrid)	MCMC-MI,

	IGA, EASI, and Itch NRS related categorical endpoints at pre-specified timepoints		Tipping point analysis (CMH)
		Supportive Estimand (Composite)	NRI (CMH)
	Remaining categorical endpoints	Supportive Estimand (Composite)	NRI (CMH)
Continuous	EASI percentage change, Itch NRS percentage change	Primary Estimand (Hybrid)	MCMC-MI (ANCOVA)
		Supportive Estimand (Hypothetical)	No imputation (MMRM)
	Remaining continuous endpoints at multiple post-baseline timepoints	Supportive Estimand (Hypothetical)	No imputation (MMRM)
	Remaining continuous endpoints collected only once post-baseline	Supportive Estimand (Hypothetical)	LOCF (ANCOVA)

Abbreviations: ANCOVA = analysis of covariance; CMH = Cochran-Mantel-Haenszel; EASI = Eczema Area and Severity Index; IGA = Investigator's Global Assessment for AD; LOCF = last observation carried forward; MCMC-MI = Markov chain Monte Carlo multiple imputation; MMRM = mixed-effects model for repeated measures; NRI = Non-Responder Imputation; NRS = Numeric Rating Scale.

4.1.6.1.1. Markov Chain Monte Carlo Multiple Imputation (MCMC-MI)

The primary method of handling missing efficacy data relative to the primary estimand will be as follows for both binary and continuous endpoints:

For participants who receive rescue medication (i.e., high-potency TCS or systemic AD treatment, defined in Appendix 13), or discontinue treatment due to lack of efficacy, set to the participant's baseline value subsequent to this time through Week 16. The MCMC-MI will be used to handle the remaining missing data. Imputation will be conducted within each treatment group independently so the pattern of missing observations in one treatment group cannot influence missing value imputation in another. The SAS PROC MI with MCMC option will be used to conduct the MCMC-MI. The imputation model will include the relevant baseline and post-baseline.

For each imputation process, 25 datasets with imputations will be calculated. The initial seed values are given in the table below. Each complete data set will be analyzed with the specified analysis. The results from these analyses will be combined into a single inference using SAS PROC MIANALYZE.

CMH test statistic will be transformed using the Wilson-Hilferty transformation and then standardized (Ratitch 2013) prior to combining them using SAS PROC MIANALYZE. Details of combining estimates and test statistics for categorial endpoints with multiple imputation can be found in Appendix 14.

For binary responses related to EASI and IGA, the binary response variables will be calculated based on the multiply imputed datasets that have been created. Because the MCMC algorithm is based on the multivariate normal model, imputed values for IGA will not generally be one of the discrete values used in IGA scoring (0, 1, 2, 3, or 4). Therefore, to derive the binary IGA response variable, standard rounding rules will be applied to the imputed values. For example, if a participant has an IGA score imputed as 1.4 (and assuming a Baseline IGA score of 3), the imputed value would be rounded down to 1, and the minimum change from Baseline of 2 would have been met. This participant would be considered a responder.

For derivation of an EASI-75 and EASI-90 response, no rounding will be performed. The imputed Week 16 EASI value will be compared directly to the observed Baseline EASI value to determine whether a reduction of at least 75% or 90% was achieved.

For derivation of the following Itch NRS responses, no rounding will be performed. The imputed Itch NRS value will be compared directly to the observed mean baseline Itch-NRS value to determine whether a response was achieved:

- Percentage of participants with Itch NRS of ≥4-point at Baseline who achieve a ≥4-point reduction from Baseline at Weeks 1, 2, 4, and 16.

Imputation of continuous data will parallel that of binary variable. The imputed values will be used for the following key secondary endpoint:

- Percentage change in EASI score from Baseline to Week 16.

Seed Values for MCMC-MI for Induction Period

Analysis	Seed values Lebrikizumab 250 mg Q2W Lebrikizumab 250 mg Q4W Placebo
Proportion of participants achieving IGA score of 0 or 1 with a ≥2-point improvement from baseline at Week 16	180345
Change and percentage change from baseline in EASI score at Week 16. EASI-75 and EASI-90 will leverage imputation from EASI score and therefore use the same seed number.	177100
Change and percentage change in Itch NRS from baseline to Week 16. Proportion of participants achieving at least a 4-point improvement from baseline at Weeks 1, 2, 4, and 16 will leverage imputation from Itch NRS and therefore use the same seed number.	119549

Abbreviations: EASI = Eczema Area and Severity Index score; IGA = Investigator's Global Assessment for AD; MCMC-MI = Markov chain Monte Carlo multiple imputation; NRS = Numeric Rating Scale; Q2W = every 2 weeks; Q4W = every 4 weeks.

4.1.6.1.2. Tipping Point Analyses

The co-primary endpoints of EASI-75 and IGA score of 0 or 1 with ≥ 2-point improvement from baseline at Week 16 and the following key secondary endpoints: EASI-90 at Week 16 and Itch NRS improvement ≥ 4-point, at Week 16 will be assessed using the tipping point analysis. For

each of these endpoints, the tipping point analysis will only be conducted if its co-primary or key secondary analyses results are statistically significant.

All participants who use rescue medication or discontinue treatment due to lack of efficacy will be imputed as non-responders. Assumptions on missing data as a result of treatment discontinuation due to reasons other than lack of efficacy or any other intermittent missing data will be varied to investigate if there will be any tipping points.

For all the categorical endpoints described above that will be assessed using tipping point analysis, the following process will be used to determine the tipping point:

- Missing responses in the lebrikizumab groups will be imputed with a range of response probabilities, including probabilities of 0, 0.2, 0.4, 0.6, 0.8, and 1.0.
- For missing responses in the placebo group, a range of response probabilities (for example, probability = 0, 0.2 … 1.0) will be used to impute the missing values. Multiple imputed dataset will be generated for each response probability.
- Treatment differences between lebrikizumab and placebo are analyzed for each imputed dataset using CMH test (Section 4.1.1). Results across the imputed datasets are aggregated using SAS PROC MIANALYZE in order to compute a p-value for the treatment comparisons for the given response probability. If the probability values do not allow for any variation between the multiple imputed datasets (for example, all missing responses in the placebo and lebrikizumab groups are imputed as responders and non-responders, respectively, i.e., extreme case), then the p-value from the single imputed dataset will be used.

The tipping point is identified as the response probability value within the placebo group that leads to a loss of statistical significance when evaluating lebrikizumab relative to placebo.

For tipping point analyses, the number of imputed data sets will be m=25 and the seed values to start the pseudorandom number generator of SAS are given in the table below.

Seed Values for Tipping Point Analysis

Analysis	Seed value
Proportion of participants achieving IGA score of 0 or 1 with a ≥ 2-point improvement from baseline at Week 16	123470
Proportion of participants achieving EASI-75 and EASI-90 at Week 16	123471
Proportion of participants achieving at least a 4-point improvement from baseline in Itch NRS at Week 16	123472

4.1.6.1.3. Non-responder Imputation

The non-responder imputation (NRI) method will be used to handle missing data relative to the supportive estimand for categorical endpoints (composite). Participants who receive rescue medication (i.e., high-potency TCS or systemic AD treatment, defined in Appendix 13), or discontinue treatment, will be set to non-response subsequent to this time through Week 16. Intermittent missing values will also be set to non-response.

The non-responder imputation (NRI) method imputes missing values as non-responders and can be justified based on the composite strategy (ICH E9 R1) for handling intercurrent events. In this strategy, participants are defined as responders only if they meet the clinical requirements for

response at the predefined time AND they remain on the assigned study treatment (i.e., not using rescue medications and not having missing values due to other reasons). Failing either criteria by definition makes them non-responders.

Randomized participants without at least 1 post-baseline observation will also be defined as non-responders for all visits for the NRI analysis.

4.1.6.1.4. Mixed-effects Model for Repeated Measures (MMRM)

Mixed-effects model for repeated measures analyses will be performed on continuous endpoints to mitigate the impact of missing data. This approach assumes missing observations are missing-at-random (missingness is related to observed data) and borrows information from participants in the same treatment arm taking into account both the missingness of data through the correlation of the repeated measurements.

The values subsequent to rescue medication use (i.e., high-potency TCS or systemic AD treatment, defined in Appendix 13) or treatment discontinuation will be made missing before applying the MMRM model. The MMRM model is described in Section 4.1.1.

4.1.6.1.5. Last Observation Carried Forward (LOCF)

In this analysis, the values subsequent to rescue medication use (i.e., high-potency TCS or systemic AD treatment, defined in Appendix 13) or treatment discontinuation will be made missing. All missing values will be imputed using LOCF. Baseline value will be used for imputation if there is no post-baseline observation.

4.1.6.2. Handling of Dropouts or Missing Data for Maintenance Period

For maintenance efficacy analysis relative to the maintenance primary estimand (Hybrid), the method of handling missing data including those as a result of intercurrent events will be MCMC-MI. The description of maintenance MCMC-MI method can be found in Section 4.1.6.2.1.

MCMC-MI will also be used to handle missing data relative to the maintenance supportive estimand (Hybrid).

For efficacy analysis relative to the maintenance supportive estimand for categorical endpoints (Composite), missing data including those as a result of intercurrent events will be imputed as non-responder. The description of maintenance NRI can be found in Section 4.1.6.2.2.

For efficacy analysis relative to the maintenance supportive estimand for continuous endpoints (Hypothetical), missing data including those as a result of intercurrent events will be imputed using LOCF. The description of maintenance LOCF can be found in Section 4.1.6.2.3.

The table below describes the planned imputation methods for efficacy and health outcome endpoints for the Maintenance Period.

Imputation Techniques for Various Variables during Maintenance Period

Type of Endpoints	Efficacy and Health Outcome Endpoints	Estimand (Analysis strategy for Intercurrent Events)	Missing Data Imputation Method (Analysis Method)
Categorical	IGA, EASI, and Itch NRS related categorical endpoints at pre-specified timepoints	Maintenance Primary Estimand (Hybrid)	MCMC-MI (Descriptive statistics)
		Maintenance Supportive Estimand (Hybrid)	MCMC-MI (Descriptive statistics)
		Maintenance Supportive Estimand (Composite)	NRI (Descriptive statistics)
	Remaining categorical endpoints	Maintenance Supportive Estimand (Composite)	NRI (Descriptive statistics)
Continuous	EASI percentage change, Itch NRS percentage change	Maintenance Primary Estimand (Hybrid)	MCMC-MI (Descriptive statistics)
		Maintenance Supportive Estimand (Hybrid)	MCMC-MI (Descriptive statistics)
		Maintenance Supportive Estimand (Hypothetical)	LOCF (Descriptive statistics)
	Remaining continuous endpoints	Maintenance Supportive Estimand (Hypothetical)	LOCF (Descriptive statistics)

Abbreviations: EASI = Eczema Area and Severity Index; IGA = Investigator's Global Assessment for AD; LOCF = last observation carried forward; MCMC-MI = Markov chain Monte Carlo multiple imputation; NRI = Non-Responder Imputation; NRS = Numeric Rating Scale.

4.1.6.2.1. Maintenance Period MCMC-MI

The MCMC-MI will be used to handle missing data relative to the maintenance primary estimand (Hybrid) and the maintenance supportive estimand (Hybrid) for both binary and continuous endpoints. Imputation will be conducted within each treatment group independently so the pattern of missing observations in one treatment group cannot influence missing value imputation in another. The SAS PROC MI with MCMC option will be used to conduct the MCMC-MI. The imputation model will include the relevant baseline and post-baseline.

For each imputation process, 25 datasets with imputations will be calculated. The initial seed values are given in the table below. Each complete data set will be analyzed with the specified

analysis. The results from these analyses will be combined into a single inference using SAS PROC MIANALYZE.

The imputation and analysis will be conducted on the MPP only.

The derivation of binary responses related to EASI, IGA and Itch NRS for the Maintenance Period will follow the derivation for the Induction Period. For the derivation of percentage change from baseline in EASI, the imputed values will be used directly to compare with baseline EASI.

Seed Values for MCMC-MI for Maintenance Period

Analysis	Seed values Lebrikizumab 250 mg Q2W Lebrikizumab 250 mg Q4W
IGA	12345
EASI	12346
Itch NRS	12347

Abbreviations: EASI = Eczema Area and Severity Index score; IGA = Investigator's Global Assessment for AD; MCMC-MI = Markov chain Monte Carlo multiple imputation; NRS = Numeric Rating Scale; Q2W = every 2 weeks; Q4W=every 4 weeks.

4.1.6.2.2. Maintenance Period NRI

The NRI method will be used to handle missing data relative to the maintenance supportive estimand for categorical endpoints (Composite). Participants who receive rescue medication (i.e., high-potency TCS or systemic AD treatment, defined in Appendix 13), discontinue treatment, or transfer to escape arm will be set to non-response subsequent to this time through Week 68. Intermittent missing values will also be set to non-response.

Participants without at least 1 post-baseline observation will also be defined as non-responders for all visits for the NRI analysis.

4.1.6.2.3. Maintenance Period LOCF

The LOCF will be used to handle missing data relative to the maintenance supportive estimand for continuous endpoints (Hypothetical). In this analysis, the values subsequent to rescue medication use (i.e., high-potency TCS or systemic AD treatment, defined in Appendix 13), treatment discontinuation or transfer to escape arm will be made missing. All missing values will be imputed using LOCF. Baseline value will be used for imputation if there is no post-baseline observation.

4.2. Participant Dispositions

The following participant disposition summaries will be provided (details of the analysis populations can be found in Section 3):

- Total number and percentage of participants entering each analysis population.

- The number and percentage of participants who entered the study, failed screening, were randomized at Baseline Visit (Visit 2, Week 0), completed Week 16, completed Week 68, and completed the safety Follow-Up Visit. Summary will be provided by the initial randomized treatment group (Analysis population: Intent-to-Treat [ITT]).
- The number and percentage of participants who completed the study, and the number and percentage of participants who discontinued the study at any time, by the initial randomized treatment group and primary reason for discontinuation (Analysis population: ITT).
- The number and percentage of participants who completed the Induction Period and the number and percentage of participants who discontinued from the Induction Period, by treatment group and primary reason for discontinuation (Analysis population: ITT).
- The number and percentage of participants who completed the Maintenance Period and the number and percentage of participants who discontinued from the Maintenance Period, by treatment group and primary reason for discontinuation (Analysis populations: MPP and MSP), in addition, the number and percentage of participants who entered the escape arm will be summarized for the MPP and the MSP.

All participants who were randomized (that is, in the ITT Population) and discontinued from study treatment during any period from the study will be listed together with the discontinuation reason, and the timing of discontinuation from the study will be reported.

Participant allocation by site will be summarized with number of participants who entered the study, number of ITT participants for each treatment group, number of participants discontinued from study treatment, and number of participants discontinued from the study.

4.3. Primary Endpoints Analysis

4.3.1. Definition of endpoints

The co-primary endpoints are comprised of 2 separate endpoints:

- the proportion of participants achieving EASI-75 at Week 16 in lebrikizumab 250 mg Q2W versus placebo or lebrikizumab 250 mg Q4W versus placebo, and
- the proportion of participants achieving IGA score of 0 or 1 and a reduction of ≥2-points from baseline to Week 16 in lebrikizumab 250 mg Q2W versus placebo or lebrikizumab 250 mg Q4W versus placebo.

Descriptions and derivations of these endpoints are shown in Appendix 1.

4.3.2. Main analytical approach

The primary analysis of the study is to test the co-primary null hypotheses described in Section 2 in the ITT Population.

The primary estimand addresses the treatment response as directed. The analysis assumes that treatment response disappears for participants who took rescue medication (i.e., high-potency TCS or systemic AD treatment, defined in Appendix 13) or withdrew from the study due to lack of efficacy, therefore setting to the participant's baseline value subsequent to this time through

Week 16. Other missing values will be imputed using MCMC-MI based on missing at random assumption (Section 4.1.6.1.1).

A CMH test as described in Section 4.1.1 will be used for the comparisons. The OR, the corresponding 95% CIs and p-value, as well as the treatment differences and the corresponding 95% CIs, will be reported.

Multiplicity controlled analyses will be performed on the primary and key secondary endpoints to control the overall Type I error rate at a 2-sided alpha level of 0.05. A graphical approach will be used to perform the multiplicity controlled analyses as described in Section 2.1.

The analyses for primary outcomes of EASI-75 and IGA score of 0 or 1 are described in Appendix 2.

4.3.3. Sensitivity Analyses

Sensitivity analyses are included to demonstrate robustness of analyses methods using different missing data imputations and analyses assumptions. Sensitivity analyses for the co-primary endpoints will be conducted using the tipping point analysis based on missing not at random assumption (Section 4.1.6.1.2) described in Appendix 2.

There will be no adjustment for multiple comparisons for additional analyses of the primary outcomes.

4.3.4. Supplementary analyses

Supplementary analyses of the co-primary endpoints will be conducted using supportive estimand (Section 1.1.1.1.2) and different analysis population of PPS (Section 3) described in .

4.4. Secondary Endpoints Analysis

4.4.1. Major secondary endpoints

4.4.1.1. Definition of endpoints

Major secondary endpoints are listed in Section 1.1 under Major Secondary.

Descriptions and derivations of these endpoints are shown in Appendix 1.

4.4.1.2. Main analytical approach

The analyses for the major secondary outcomes are described in Appendix 2.

4.4.1.3. Sensitivity Analyses

Sensitivity analyses for the major secondary endpoints are described in Appendix 2.

4.4.2. Other secondary endpoints

Other secondary endpoints are listed in Section 1.1 under Other Secondary.

Descriptions and derivations of these endpoints are shown in Appendix 1.

The analyses for the other secondary endpoints are described in Appendix 2.

4.5. Tertiary/Exploratory Endpoints Analysis

These analyses may be described in the Exploratory Analyses and Health Technology Analyses Plan.

4.6. Safety Analyses

The planned analyses of safety data will be performed with an intent to maintain consistency with compound level standard safety analyses. These standards are based on internal standards which were informed by Clinical Data Interchange Standards Consortium (CDISC) standards, regulatory guidance (for example, FDA Clinical Review Template), and cross-industry standardization efforts (for example, Pharmaceutical Users Software Exchange [PhUSE] white papers from the Standard Analyses and Code Sharing Working Group provided in the PhUSE Computational Science Deliverables Catalog).

Safety evaluations will be based upon the following safety analysis populations with their associated study periods, unless specified otherwise:

- Safety Population (Induction Period),
- MPP (Maintenance Blinded Period), and
- All Lebrikizumab Safety Population (Combined Induction and Maintenance Periods, and Combined Induction and Maintenance Periods plus Follow-Up Period [selective analyses]).

These analysis populations, treatment groups, associated study periods, and the comparisons for each analysis population are fully defined in Section 3.

For document writing purposes for safety, tests with two-sided p-values less than 0.05 will be referred to as having strong statistical evidence for a treatment difference, unless otherwise noted. However, p-values should not be over-interpreted for these safety analyses. Except for pre-specified hypotheses, they correspond to data-driven hypotheses and hence are only useful as a flagging mechanism.

4.6.1. Extent of Exposure

Duration of exposure to study intervention will be summarized by treatment group. Drug interruption time period due to the use of systemic rescue therapies will not be removed from study drug exposure calculations as described in compound level safety standards.

The duration of exposure will be calculated as:

$$\begin{aligned} &Duration\ of\ exposure\ (days) \\ &= Date\ of\ last\ visit\ (scheduled\ or\ unscheduled)\ in\ the\ specified\ Treatment\ Period \\ &- Date\ of\ first\ dose\ in\ Treatment\ Period + 1 \end{aligned}$$

Note that date of last visit in the Induction Period will be defined as the first visit date in the Maintenance Period – 1.

The number and percentage of participants in each of the following categories will be included in the summaries:

- >0, ≥7, ≥14, ≥30, ≥60, ≥90, ≥112, ≥120 days for the Induction Period (for the Maintenance Period, use >0, ≥30, ≥60, ≥90, ≥120, ≥150, ≥180, ≥210, ≥240, ≥252, ≥308, ≥365 days, for the Combined Induction and Maintenance Periods, use >0, ≥7, ≥14, ≥30, ≥60, ≥90, ≥112, ≥120, ≥150, ≥183, ≥210, ≥273, ≥365, ≥476 days). Note that participants may be included in more than 1 category.
- >0 to <7, ≥7 to <14, ≥14 to <30, ≥30 to <60, ≥60 to <90, ≥90 to <112, ≥112 to <120, ≥120 days (for the Maintenance Period, use >0 to <30, ≥30 to <60, ≥60 to <90, ≥90 to <120, ≥120 to <150, ≥150 to <180, ≥180 to <210, ≥210 to <240, ≥240 to <252, ≥252 to <308, ≥308 to <365, ≥365 days, for the Combined Induction and Maintenance Periods, use >0 to <7, ≥7 to <14, ≥14 to <30, ≥30 to <60, ≥60 to <90, ≥90 to <112, ≥112 to <120, ≥120 to <150, ≥150 to <183, ≥183 to <210, ≥210 to <273, ≥273 to <365, ≥365 to <476, ≥476 days).

Additional exposure ranges may be considered if necessary. No p-values will be reported.

The summaries will also include the following information:

- Total exposure in participant years, calculated as:

$$Total\ exposure\ in\ participant\ years = \frac{Sum\ of\ duration\ of\ exposures\ for\ all\ participants\ in\ treatment\ group}{365.25}$$

- Mean and median total dose: Total dose (in mg) is calculated by the number of active injections taken during the treatment period multiplied by dose. For participants in the Safety Population who were randomized to lebrikizumab 250 mg Q4W or Q2W, or participants in the MPP who received lebrikizumab 250 mg Q4W or Q2W, the total dose (in mg) taken during the Induction Period or the Maintenance Period will be calculated as follows: *Total lebrikizumab dose=Total number of active injections (including loading doses, if any) received in the Induction Period or the Maintenance Period ×250*.
- Total number of injections received will be derived based on the response to the question "Was dose administered?" on the *Exposure as Collected* eCRF page.

The exposure for the All Lebrikizumab Safety Population during the Combined Induction and Maintenance Periods will be calculated as (Date of last study visit during Treatment Period – Date of first lebrikizumab injection +1 day) calculated for each treatment period where the participant receives lebrikizumab and then summed together (this excludes the duration of time that participants are receiving placebo during the Maintenance Period).

The exposure for the All Lebrikizumab Safety Population during the Combined Induction and Maintenance Periods plus Follow up Period will be calculated as the time between the first dose of lebrikizumab and the study treatment disposition visit plus any follow-up period.

4.6.2. Adverse Events

A TEAE is defined as an event that first occurred or worsened in severity after baseline. The MedDRA Lowest Level Term (LLT) will be used in the treatment-emergent computation. The maximum severity for each LLT during the baseline period defined in Section 4.1.1 for the Induction Period and Section 4.1.2 for the Maintenance Period will be used as baseline. The

Treatment Period will be used as the post-baseline period for the analysis. For events with a missing severity during the baseline period, it will be treated as 'mild' in severity for determining treatment-emergence. Events with a missing severity during the post-baseline period will be treated as 'severe' and treatment-emergence will be determined by comparing to baseline severity. For events occurring on the day of first taking study medication, it will be assumed to be post-treatment.

The planned summaries for adverse events are provided in the table below, and are described more fully in compound level safety standards and in the adverse event-related PhUSE white paper [Analysis and Displays Associated with Adverse Events: Focus on Adverse Events in Phase 2-4 Clinical Trials and Integrated Summary Document (PhUSE 2017)].

Summary tables as described in the table below will be presented for the following periods/analysis populations as indicated. Summary tables will include the number and percentage of participants reporting an event. For events that are gender-specific (as defined by MedDRA), the number of participants at risk will include only participants from the given gender.

- Induction Period (Safety Population, S)
- Maintenance Blinded Period (Maintenance Primary Population, M)
- Combined Induction and Maintenance Periods, Combined Induction and Maintenance Periods Plus Follow-Up Period (selective analyses) (All Lebrikizumab Safety Population, A)

Summary Tables/Listing Related to Adverse Events

Analysis	Population
Overview of AEs	S, MPP, MSP, A
Summary of TEAE by PTs	S, MPP
Summary of TEAE by PTs occurring in ≥1% of participants	S, MPP
Summary of TEAE by PTs within SOC	S, MPP, A
Summary of TEAE PTs by maximum severity	S, MPP
Summary of SAE by PTs within SOC	S, MPP, A
Summary of AEs leading to treatment discontinuation by PTs within SOC	S, MPP, A
Summary of TEAE possibly related to study drug by PTs within SOC	S, MPP
Listing of SAEs (including Death)	ITT
Listing of primary AEs leading to study treatment discontinuation	ITT
Listing of AE (including AEs in the Maintenance Period and the Follow-Up Period)	S

Abbreviations: A = All Lebrikizumab Safety Population; AE = adverse event; ITT = Intent-to-Treat; MPP = Maintenance Primary Population; MSP = Maintenance Secondary Population; PT = Preferred Term; SAE = serious adverse event; S = Safety Population; SOC = System Organ Class; TEAE = treatment-emergent adverse event.

Statistical comparisons will be performed for Safety Population using Fisher's exact test. OR will be provided.

4.6.2.1. Common Adverse Events

The number and percentages of participants with TEAEs will be summarized by treatment using MedDRA Preferred Term (PT) for the common TEAEs (occurred in ≥1% before rounding in total lebrikizumab column in the table).

4.6.2.2. Deaths, Other Serious Adverse Events, and other Notable Adverse Events

The number and percentage of participants reported with an SAE during the treatment period will be summarized by treatment using MedDRA PT. A listing of SAEs will be provided.

The number and percentage of participants who permanently discontinued from study treatment due to an AE (including AEs that led to death) during the treatment period will be summarized by treatment using MedDRA PT. Events will be ordered by decreasing frequency in all treatment groups.

4.6.3. Additional Safety Assessments

4.6.3.1. Clinical Laboratory Evaluation

As described more fully in compound level safety standards and in the laboratory-related PhUSE white papers (PhUSE 2013; PhUSE 2015), the clinical laboratory evaluations will be summarized as described in the table below. Hormone analytes are summarized/plotted similarly for adolescent participants.

Analysis for Clinical Laboratory Evaluations

Analysis	Population
Box plots of observed values by visit Box plots for change values by visit	S, MPP, A
Change from baseline to last observations. ANCOVA model with treatment and baseline value in the model.	S, MPP
Scatter plots of baseline-by-maximum values and baseline-by-minimum values	S, MPP
Treatment-emergent abnormal high lab values (i.e., participants shifting from a normal/low maximum baseline value to a high maximum post-baseline value) or abnormal low lab values (i.e., participants shifting from normal/high minimum baseline value to a low minimum post-baseline value)	S, MPP, A
Shift tables showing the number of participants who shift from each category of maximum (minimum) baseline observation to each category of maximum (minimum) post-baseline observation. Categories may be low, normal, or high with cut-offs defined in the compound level safety standards.	S, MPP
Listing of abnormal findings for laboratory analyte measurements, including qualitative measures	S

Abbreviations: A = All Lebrikizumab Safety Population; ANCOVA = analysis of covariance; MPP = Maintenance Primary Population; S = Safety Population.

4.6.3.2. Vital Signs and Other Physical Findings

As described more fully in compound level safety standards and in the laboratory-related PhUSE white papers (PhUSE 2013; PhUSE 2015), vital signs will be summarized similarly to the

clinical laboratory evaluation as described in the table below. For vital signs, treatment-emergent low and high are based on a combination of a specified value and a change or percentage change for adults and adolescents as defined in the compound level safety standards.

Analysis Related to Vital Signs

Analysis	Population
Box plots for observed values by visit	S, MPP, A
Box plots for change from baseline values by visit	S, MPP, A
Scatterplots of baseline-by-maximum values and baseline-by-minimum values	S, MPP
Tables with the number and percentage of participants who shift from normal/high to low (i.e., treatment-emergent low) and the number and percentage of participants who shift from normal/low to high (i.e., treatment-emergent high). The limits are defined in the compound level safety standards.	S, MPP, A

Abbreviations: A = All Lebrikizumab Safety Population; MPP = Maintenance Primary Population; S = Safety Population.

4.6.3.2.1. Adolescent Standardized Growth

Weight, height, and BMI data will be merged to the Centers for Disease Control and Prevention (CDC) standard growth data (released in 2000) by age and gender in order to compare participants' growth with the standard. Z-score, and standardized percentile of weight, height, and BMI at each visit will be calculated and compared to the 2000 CDC growth charts. Because of the short duration of controlled period and the small number of adolescent participants, only All Lebrikizumab Safety Population will be described during the Combined Induction and Maintenance Periods.

The z-score and percentile calculations are based on algorithms and data provided by the National Center for Health Statistics. The details are provided in the CDC website (https://www.cdc.gov/nccdphp/dnpao/growthcharts/resources/sas.htm) (CDC resources page [WWW]).

Height and weight may also be merged to the Japanese children standard growth data issued by the Japanese Association for Human Auxology to compare adolescent participants' growth with the standard (Isojima et al. 2016).

The following summaries and plots will be provided:

Analysis Related to Adolescent Standardized Growth

Analysis	Population
Summaries for baseline, mean change of actual measure, z-score and standardized percentile of weight, height, and BMI.	A
Scatter plot of participants' mean weight, height, and BMI standardized percentile versus lebrikizumab exposure time	A

Abbreviations: A = All Lebrikizumab Safety Population; BMI = body mass index.

4.6.3.3. Immunogenicity

An individual sample is potentially examined multiple times in a hierarchical procedure to produce a sample anti-drug antibody (ADA) assay result and may yield a sample neutralizing

ADA (NAb) assay result. Treatment-emergent ADA (TE-ADA) are defined as those with a titer 2-fold (1 dilution) greater than the minimum required dilution if no ADAs were detected at baseline (treatment-induced ADA) or those with a 4-fold (2 dilutions) increase in titer compared to baseline if ADAs were detected at baseline (treatment-boosted ADA). A participant is considered TE-ADA positive when at least 1 post-baseline ADA sample meets the definition of TE-ADA.

Compound level safety standards will be followed in the analyses of immunogenicity. Listings of immunogenicity assessments will be provided for the Safety Population. The summary of TE-ADA and NAb status will be produced for the 3 populations (i.e., Safety Population for the Induction Period, MPP for the combined Induction and Maintenance Periods, and All Lebrikizumab Safety Population for the combined Induction and Maintenance Periods plus Follow-Up Period), where the post-baseline period for reporting is the same as described for AEs in Section 4.6.2. Additional assessments of the relationship between immunogenicity and efficacy/safety may be performed in this study.

4.6.3.4. Special Safety Topics including Adverse Events of Special Interest

This section includes areas of interest whether due to observed safety findings, potential findings based on drug class, or safety topics anticipated to be requested by a regulatory agency for any reason. In general, potential adverse events of special interest (AESI) relevant to these special safety topics will be identified by one or more Standardized MedDRA Query(ies) (SMQs), by a Lilly-defined MedDRA PT listing based upon the review of the most current version of MedDRA, or by treatment-emergent relevant laboratory changes, as described below. Additional special safety topics may be added as warranted.

Unless otherwise specified, the special safety topics will be summarized for the Safety Population and All Lebrikizumab Safety Population during their associated study periods as described in Section 4.6.2.

Full details of the search terms and rules for deriving special safety topics in each of the sections below are described in the compound level safety standards along with information about the types of summaries and listings to be provided. In the event that the listing of terms or analysis changes for a special safety topic, it will be documented in the compound level safety standards which will supersede this document; it will not warrant an amendment to the individual study SAP.

4.6.3.4.1. Hepatic Safety

Hepatic labs include alanine aminotransferase (ALT), aspartate transaminase (AST), total bilirubin (TBL), and serum alkaline phosphatase (ALP).

Tables Related to Hepatic Safety

Analysis	Population
ALT and AST: The number and percentage of participants with a measurement greater than or equal to 3 times (3X), 5 times (5X), and 10 times (10X) the performing lab upper limit of normal (ULN) during the treatment period for all participants with a post-baseline value and for subsets based on various levels of baseline value	S, A

TBL and ALP: The number and percentage of participants with a measurement greater than or equal to 2 times (2X) the performing lab ULN during the treatment period will be summarized for all participants with a post-baseline value and for subsets based on various levels of baseline value	
Plot of maximum post-baseline ALT vs. maximum post-baseline total bilirubin	Safety Population for All Periods: ever on lebri and never on lebri

Abbreviations: A = All Lebrikizumab Safety Population; ALP = serum alkaline phosphatase; ALT = alanine aminotransferase; AST = aspartate transaminase; lebri = lebrikizumab; S = Safety Population; TBL = total bilirubin.

4.6.3.4.2. Eosinophilia and Eosinophil-Related Disorders

In addition to the standard laboratory analysis (Section 4.6.3.1), eosinophilia and eosinophil-related AE will be summarized. Details regarding eosinophil-related PTs are in Compound Level Safety Standard.

Tables Related to Eosinophilia and Eosinophil-Related AE

Analysis	Population
Shift table summarizing the number and percentage of participants within each maximum baseline category versus each maximum post-baseline category by treatment	S, A
Summary of eosinophil-related TEAE by PT	S, A

Abbreviations: A = All Lebrikizumab Safety Population; AE = adverse event; PT = Preferred Term; S = Safety Population; TEAE = treatment-emergent adverse event.

4.6.3.4.3. Infections, including herpes infections, and relevant parasitic infections

Infections will be defined using the PTs from the MedDRA Infections and Infestations System Organ Class (SOC). The MedDRA terms used to identify infections considered to be opportunistic infections (OI) in participants with immune mediated inflammatory conditions treated with immunomodulatory drugs are based on Winthrop et al. (2015) and are listed in the compound level safety standards. The list contains narrow (more specific) and broad (less specific) PTs with respect to these prospectively defined OIs. Definitions of herpes infections, parasitic infections and skin infections are listed in the compound level safety standards.

Summary Tables/Listing Related to Infection Related AE

Analysis	Population
Summary of treatment-emergent infections by maximum severity	S, A
Summary of serious infections by PT	S, A
Summary of infection AEs resulting in permanent study drug discontinuation by PT	S, A
Summary of treatment-emergent potential OI by PT nested with categories for narrow terms and broad terms separately	S, MPP, A

Summary of treatment-emergent adverse events - herpes and parasitic infections	S, A
Summary of treatment-emergent adverse events - skin infections by maximum severity	S, A
Summary and/or listing of infection follow-up form	S
A listing of participants with potential OI, serious infection, herpes and parasitic infections (including events in the Maintenance Period and the Follow-Up Period)	S

Abbreviations: A = All Lebrikizumab Safety Population; AE = adverse event; MPP = Maintenance Primary Population; OI = opportunistic infections; PT = Preferred Term; S = Safety Population; TEAE = treatment-emergent adverse event.

4.6.3.4.4. Conjunctivitis

Conjunctivitis are events of special interest and will be identified using PTs nested within the categories of conjunctivitis and Keratitis as described in the Compound Level Safety Standards.

Summary Tables/Listing Related to Conjunctivitis

Analysis	Population
Summary of TEAE of conjunctivitis within categories by maximum severity	S, MPP, A
Summary of TEAE of conjunctivitis cluster by maximum severity	S, MPP, A
Summary and/or listing of conjunctivitis and eye inflammation follow-up form	S
A listing of participants with conjunctivitis (including events in the Maintenance Period and the Follow-Up Period)	S

Abbreviations: A = All Lebrikizumab Safety Population; MPP = Maintenance Primary Population; S = Safety Population; TEAE = treatment-emergent adverse event.

4.6.3.4.5. Hypersensitivity

Potential hypersensitivity reactions will be determined using the following SMQs: anaphylactic reaction, hypersensitivity, and angioedema. Potential hypersensitivity will be categorized as immediate (i.e., occurring the same day as drug administration) and non-immediate (i.e., occurring after the day of study drug administration but prior to subsequent drug administration).

Summary Tables/Listing Related to Hypersensitivity

Analysis	Population
For immediate hypersensitivity: (1) combined narrow/algorithmic search (that is, any narrow term from any one of the SMQs, or anaphylaxis algorithm); (2) narrow search (that is, any narrow term) by SMQ; (3) broad search (that is, any narrow or broad term) by SMQ; and (4) TEAEs (occurring on the day of study drug administration) by PT not in any of the 3 SMQs	S, A
For nonimmediate hypersensitivity: (1) combined narrow search (that is, any narrow term from any one of the SMQs); (2) narrow search (that is, any narrow term) by SMQ; and (3) broad search (that is, any narrow or broad term) by SMQ	S, A

Abbreviations: A = All Lebrikizumab Safety Population; MedDRA = Medical Dictionary for Regulatory Activities; PT = Preferred Term; S = Safety Population; SMQ = Standardized MedDRA Query; TEAE = treatment-emergent adverse events.

4.6.3.4.6. Injection Site Reactions (ISR)

Injection site reactions (ISRs) are AEs localized to the immediate site of the administration of a drug. The evaluation of study drug related ISRs will be through the unsolicited reporting of ISR TEAEs. Injection site reactions will be defined using the MedDRA High Level Term (HLT) of Injection Site Reaction, excluding certain PTs related to joints as described in the Compound Level Safety Standards.

Tables Related to Injection Site Reactions

Analysis	Population
Summary of TEAE of ISR overall, and by PT	S, MPP, A

Abbreviations: A = All Lebrikizumab Safety Population; ISR = injection site reaction; MPP = Maintenance Primary Population; PT = Preferred Term; S = Safety Population; TEAE = treatment-emergent adverse event.

4.6.3.4.7. Malignancies

Malignancies will be defined using PTs from the Malignant tumors SMQ and summarized separately for the 2 categories: Non-melanoma skin cancer (NMSC) and Malignancies excluding NMSC as below.

Summary Tables Related to Malignancies

Analysis	Population
Summary of TEAE of malignancies within categories of NMSC and malignancy excluding NMSC	S, A

Abbreviations: A = All Lebrikizumab Safety Population; NMSC = non-melanoma skin cancer; TEAE = treatment-emergent adverse event; S = Safety Population.

4.6.3.4.8. Atopic Dermatitis Exacerbation

Atopic dermatitis exacerbation will be defined using PTs specified in the Compound Level Safety Standards and summarized below.

Summary Tables Related to Atopic Dermatitis Exacerbation

Analysis	Population
Summary of TEAE of atopic dermatitis exacerbation	S, A

Abbreviations: A = All Lebrikizumab Safety Population; S = Safety Population; TEAE = treatment-emergent adverse event.

4.6.3.4.9. Suicide/Self-Injury

Suicide/self-injury will be defined as described in the Compound Level Safety Standards and summarized below.

Summary Tables Related to Suicide/self-injury Standardized Medical Dictionary for Regulatory Activities Query

Analysis	Population
Summary of TEAE of Suicide/self-injury SMQ	S, A

Abbreviations: A = All Lebrikizumab Safety Population; MedDRA = Medical Dictionary for Regulatory Activities; S = Safety Population; SMQ = Standardized MedDRA Query; TEAE = treatment-emergent adverse event.

4.6.3.4.9.1. Columbia-Suicide Severity Rating Scale (C-SSRS)

The C-SSRS is an assessment tool that evaluates suicidal ideation and behavior. Information on the C-SSRS scale can be found through the following link: http://www.cssrs.columbia.edu.

Specifically, the following outcomes are C-SSRS categories and have binary responses (yes/no). The categories have been re-ordered from the actual scale to facilitate the definitions of the composite and comparative endpoints, and to enable clarity in the presentation of the results.

Category 1 – Wish to be Dead
Category 2 – Non-specific Active Suicidal Thoughts
Category 3 – Active Suicidal Ideation with Any Methods (Not Plan) without Intent to Act
Category 4 – Active Suicidal Ideation with Some Intent to Act, without Specific Plan
Category 5 – Active Suicidal Ideation with Specific Plan and Intent
Category 6 – Preparatory Acts or Behavior
Category 7 – Aborted Attempt
Category 8 – Interrupted Attempt
Category 9 – Actual Attempt (non-fatal)
Category 10 – Completed Suicide

Self-injurious behavior without suicidal intent is also a C-SSRS outcome (although not suicide-related) and has a binary response (yes/no).

Composite endpoints based on the above categories are defined below.

- **Suicidal ideation:** A "yes" answer at any time during treatment to any one of the five suicidal ideation questions (Categories 1-5) on the C-SSRS
- **Suicidal behavior:** A "yes" answer at any time during treatment to any one of the five suicidal behavior questions (Categories 6-10) on the C-SSRS
- **Suicidal ideation or behavior:** A "yes" answer at any time during treatment to any one of the ten suicidal ideation and behavior questions (Categories 1-10) on the C-SSRS

By-participant listings of C-SSRS and Self-Harm supplement and follow-up data will be provided (Analysis population: Safety Population). Given that few or no suicidal ideation or behaviors are anticipated, C-SSRS will be listed by participant and visit. Only participants that show suicidal ideation/behavior or self-injurious behavior without suicidal intent will be displayed (that is, if a participant's answers are all 'no' for the C-SSRS, then that participant will not be displayed). However, if a participant reported any suicidal ideation/behavior or self-injurious behavior without suicidal intent at any time point then all their ideation and behavior will be displayed, even if not positive. Note, missing data should not be imputed.

The *Self-Harm Supplement Form* in the eCRF is a one-question form that is completed, at any visit, including baseline visit, that asks for the number of suicidal behaviors, possible suicidal behaviors or non-suicidal self-injurious behaviors the participant has experienced since the last assessment. For each unique event identified, a questionnaire (*Self-Harm Follow-Up Form* in the

eCRF) which collects supplemental information on the self-injurious behavior is to be completed. The Self-Harm data will be listed by participant and visit if number of events on Self-Harm Supplement Form is not zero.

4.7. Other Analyses

4.7.1. Efficacy and Safety Analyses for Participants Who Self-Administer Study Intervention

During the Maintenance Period, participants will be instructed to self-administer study intervention. Administration by the participants or caregiver is recommended. If the participant or caregiver is not able to administer any dose throughout the study, study site staff may administer the injection.

In support of the regulatory submission in Japan, efficacy and safety data will be summarized for participants who self-administered lebrikizumab or placebo during the Maintenance Blinded Period. Efficacy and safety analyses will be conducted in the Maintenance Blinded Period on the MPP for participants who self-administered study intervention, defined as participants who received at least one dose of study intervention by self-injection (i.e., Study Subject or Caregiver on the *Exposure as Collected* eCRF page).

The number of self-injections will be summarized using descriptive statistics by treatment group for self-injected participants on the MPP in the Maintenance Blinded Period. The number and percentage of participants falling into the following categories on the number of self-injections during the Maintenance Blinded Period will be summarized by treatment group: >0 to <3, ≥3 to <6, ≥6 to <9, ≥9 to <12, ≥12 to <15, ≥15 to <18, ≥18 to <21, ≥21 to <24, ≥24 to <27, and ≥27.

The number and percentage of participants maintaining or achieving EASI-75 and IGA score of 0 or 1 with ≥ 2-point improvement responses from baseline at all scheduled visits during the Maintenance Blinded Period including Week 68 (NRI) will be presented by treatment group.

The following summary tables will be provided by treatment group for the Maintenance Blinded Period on the MPP:

- TEAEs, by SOC and PT.
- TEAEs possibly related to study treatment, by SOC and PT.

A by-participant listing of self-injection will be provided, including age, gender, treatment, and the number and percentages of self-injections.

4.7.2. Subgroup analyses

4.7.2.1. Efficacy Subgroup Analyses

Subgroup analyses will be conducted for the co-primary endpoints of IGA score of 0 or 1, EASI-75, EASI-90 and 4-point improvement in Itch NRS at Week 16 in the ITT Population using MCMC-MI approach as in the primary analysis (Section 4.3). A logistic regression analysis with treatment, subgroup, and treatment-by-subgroup interaction as factors will be used. The treatment-by-subgroup interaction will be tested using the Firth correction (Firth 1993) at the 10% significance level. Treatment group differences will be evaluated within each subgroup

using the chi-square test, regardless of whether the interaction is statistically significant. If any group within the subgroup (for example, yes, no) is <10% of the total population, only descriptive statistics will be provided for that subgroup (that is, no inferential testing).

Forest plots may be created to illustrate the treatment differences with 95% CIs between each of the lebrikizumab treatment groups and placebo group, by each subgroup category.

The following subgroups will be analyzed:

- Age group (Adolescents (12<18), Adults ≥18)
- Age group (Adolescents (12<18), Adults ≥18 to < 65, ≥65 to < 75, ≥75)
- Sex (male, female)
- Weight category (<60 kg, ≥60 to <100 kg, ≥100 kg)
- BMI category (Underweight (<18.5 kg/m^2), Normal (≥18.5 and <25 kg/m^2), Overweight (≥25 and <30 kg/m^2), Obese (≥30 and <40 kg/m^2), Extreme obese (≥40 kg/m^2))
- Duration since AD onset category (0 to <2 years, 2 to <5 years, 5 to <10 years, 10 to <20 years, ≥20 years)
- Baseline IGA (3 versus 4)
- Baseline EASI (≥16 to ≤21, >21 to ≤ 50, >50 to ≤72)
- Baseline Itch NRS (<4 versus ≥4)
- Prior use of systemic treatment (yes, no)

Some additional subgroup analyses may be added to meet regulatory requirement. The analysis of additional subgroups will not require an amendment to the SAP.

4.7.2.2. Safety Subgroup Analyses

Subgroup analysis for safety related endpoints will be performed within the context of the integrated safety analysis. Subgroup analyses may be added to meet regulatory requirement. The analysis of additional subgroups will not require an amendment to the SAP.

4.8. Interim Analyses

The first DBL and unblinding will occur, and the interim analysis, including the Maintenance Period, will be performed at the time the last participant completes Week 52 or the ET visit (that is, a cut-off date). This DBL will include all data collected by the cut-off date. Because the study will be ongoing for the Maintenance and Follow-Up periods at the time of this DBL, the analysis will be referred to as an interim analysis. The analyses from the Week 52 DBL will be treated as a primary analysis for an initial regulatory submission in Japan because all primary and major secondary study objectives will be assessed at this time. The study will not be terminated early on the basis of efficacy following this interim analysis. The sponsor or designee could unblind a small team, including but not limited to medical, statistics, data management, regulatory to prepare for regulatory interactions, safety updates, and disclosures if needed, while the investigator, study-site personnel, and Lilly study members who are site facing and the participants will be blinded to treatment assignment until the final DBL. In the first DBL, a tiered

DBL approach will be employed. Two tiers of raw data transfers will be performed for this tiered DBL. The first tier raw data transfer includes all data except for some laboratory data (e.g., thymus and activation-regulated chemokine, pharmacokinetics, and immunogenicity data) while only the remaining laboratory data will be added to the database in the second tier raw data transfer. All the planned efficacy and safety analyses will be conducted after the first tier raw data transfer in order to start preparation of the relevant sections of regulatory submission documents in Japan earlier than the remaining laboratory-related sections of the documents which will be started after the second tier raw data transfer. The detailed plan of the tiered DBL is described in the Data Management Plan. Mitigation plan of perceived bias and validation/verification activities for the tiered DBL are described in the Blinding and Unblinding Plan (BUP).

The second DBL will occur, and the interim analysis, including the Maintenance Period, will be performed at the time the last participant completes Week 68 or the ET visit (that is, a cut-off date). This DBL will include all data collected by the cut-off date. The additional results from the Week 68 DBL will be submitted to the Japan Regulatory Agency during the review period 6 months before the approval timing of lebrikizumab for the AD indication in Japan.

The final DBL will then be conducted after all participants have completed the Follow-Up Period.

Depending on the regulatory submission timeline, the second DBL and the final DBL may be combined, (that is, 1 final DBL will occur after all participants have either completed the Follow-Up Period or discontinued the study early).

4.8.1. Data Monitoring Committee (DMC)

An independent DMC composed of members who are independent of the study sponsor and study investigators will monitor patient safety by conducting formal reviews of accumulated safety data that is blinded by treatment group; if requested, the DMC may have access to the treatment allocation code or any other data requested for the purposes of a risk-benefit assessment.

The DMC will provide the sponsor with appropriate recommendations on the conduct of the clinical study to ensure the protection and safety of the participants enrolled in the study. The DMC will also institute any measures that may be required for ensuring the integrity of the study results during the study execution.

All activities and responsibilities of the DMC are described in the DMC charter. Details of the planned data analyses for the DMC are also specified in a separate DMC SAP.

4.9. Changes to Protocol-Planned Analyses

- Removed analysis for other secondary endpoint of “Number of Skin Pain-Free (Skin Pain NRS = 0) days”
- Updated definition of analysis population for the Maintenance Period

5. Sample Size Determination

Approximately 280 participants will be randomized at a 3:2:2 ratio to lebrikizumab 250 mg Q2W, lebrikizumab 250 mg Q4W, or placebo (120 participants: 80 participants: 80 participants). The inclusion of approximately 15 adolescents is based on enrollment feasibility in Japan.

The assumed IGA score of 0 or 1 at Week 16 response rates are 38% for lebrikizumab 250 mg Q2W, 33% for lebrikizumab 250 mg Q4W, and 13% for placebo. The assumed EASI-75 response rates at Week 16 are 58% for lebrikizumab 250 mg Q2W, 53% for lebrikizumab 250 mg Q4W, and 20% for placebo. The assumptions for lebrikizumab are based on the DRM06-AD01 Phase 2b study, and the proportion of participants who achieved an IGA score of 0 or 1 and proportion of participants who achieved EASI-75 response at Week 16 using the rescue medication non-response sensitivity analysis, adjusting for the allowed use of TCS. The placebo response rate is based on the review of historical TCS clinical studies in AD (Simpson et al. 2016). This study has >95% and >80% power to test the superiority of lebrikizumab 250 mg Q2W to placebo and lebrikizumab 250 mg Q4W to placebo in the co-primary endpoints based on a two-sided Fisher exact test with alpha of 0.05.

6. Supporting Documentation

6.1. Appendix 1: Description and Derivation of Efficacy and Health Outcome Endpoints

Description and Derivation of Efficacy/Health Outcome Measures and Endpoints

Measure	Description	Variable	Derivation/Comment	Imputation Approach if Missing Components
Investigator's Global Assessment (IGA)	The IGA is a static assessment and rates the severity of the participant's AD. The IGA is comprised of a 5-point scale ranging from 0 (clear) to 4 (severe) and a score is selected using descriptors that best describe the overall appearance of the lesions at a given time point.	IGA score	Single item. Range: 0 to 4 0 represents "clear" 4 represents "severe"	Single item, missing if missing.
		Change from baseline in IGA score	Change from baseline: observed IGA score – baseline IGA score	Missing if baseline or observed value is missing.
		IGA (0,1) with ≥ 2-point improvement IGA (0)	Observed score of 0 or 1 and change from baseline ≤ -2 Observed score of 0	Missing if baseline or observed value is missing. Single item, missing if missing.
		Time to loss of IGA response, i.e., developing an IGA score ≥ 2 with 2 points deterioration of achieved IGA response at Week 16	Date of first time developing an IGA score ≥ 2 with 2 points deterioration of achieved IGA response at Week 16 – date of Week 16 + 1	If a participant has not experienced loss of response by completion or early discontinuation of the Maintenance Blinded Period, the participant will be censored at the date of their last visit during the Maintenance Blinded Period. If a participant has not experienced loss of response by the time of systemic rescue, the participant will be censored at the date of systemic rescue.

Eczema Area and Severity Index (EASI)	The EASI scoring system uses a defined process (Steps 1-5 below) to grade the severity of the signs of eczema and the extent affected. The extent of disease (percentage of skin affected: 0 = 0%; 1 = 1-9%; 2 = 10-29%; 3 = 30-49%; 4 = 50-69%; 5 = 70-89%; 6 = 90-100%) and the severity of 4 clinical signs (erythema, edema/papulation, excoriation, and lichenification) each on a scale of 0 to 3 (0 = none, absent; 1 = mild; 2 = moderate; 3 = severe) at 4 body sites (head and neck, trunk, upper limbs, and lower limbs). Half scores are allowed between severities 1, 2 and 3. Each body site will have a score that ranges from 0 to 72, and the final EASI score will be obtained by weight-averaging these 4 scores. Hence, the final EASI score will range from 0 to 72 for each time point.	EASI score	Derive EASI region score for each of head and neck, trunk, upper limbs, and lower limbs as follows: $EASI_{region}$ = r*(erythema + edema/papulation + excoriation + lichenification)*(value from percentage involvement), where erythema, edema/papulation, excoriation, and lichenification are evaluated on a scale of 0 to 3, and value from percentage involvement is on a scale of 0 to 6, r = 0.1 for head and neck, r = 0.2 for upper limbs, r = 0.3 for trunk, and r = 0.4 for lower limbs. Then, total EASI score is as follows: $EASI = EASI_{head\ and\ neck} + EASI_{trunk} + EASI_{upper\ limbs} + EASI_{lower\ limbs}$	N/A – partial assessments cannot be saved.
		Change from baseline in EASI ($EASI_{region}$) score Percentage change from baseline EASI score	Change from baseline: observed EASI ($EASI_{region}$) score – baseline EASI ($EASI_{region}$) score % change from baseline: $100 \times \frac{Observed\ score - Baseline}{Baseline}$	Missing if baseline or observed value is missing.
		Each symptom score and percentage involvement score by body regions in EASI	The following scores by body regions (i.e., head and neck, trunk, upper limbs, and lower limbs): Erythema score Edema/papulation score Excoriation score Lichenification score Percentage involvement score	Missing if missing.
		Change from baseline in each symptom score and percentage involvement	Change from baseline: observed score – baseline score	Missing if baseline or observed value is missing.

		score by body regions in EASI		
		EASI-50	% Improvement in EASI score from baseline ≥ 50%: % change from baseline ≤ -50	Missing if baseline or observed value is missing.
		EASI-75	% Improvement in EASI score from baseline ≥ 75%: % change from baseline ≤ -75	Missing if baseline or observed value is missing.
		EASI-90	% Improvement in EASI score from baseline ≥ 90%: % change from baseline ≤ -90	Missing if baseline or observed value is missing.
		Time to loss of EASI-50	Date of first time % change from baseline in EASI score > -50 – date of Week 16 + 1	If a participant has not experienced loss of response by completion or early discontinuation of the Maintenance Blinded Period, the participant will be censored at the date of their last visit during the Maintenance Blinded Period. If a participant has not experienced loss of response by the time of systemic rescue, the participant will be censored at the date of systemic rescue.

		Time to loss of EASI-75	Date of first time % change from baseline in EASI score > -75 – date of Week 16 + 1	If a participant has not experienced loss of response by completion or early discontinuation of the Maintenance Blinded Period, the participant will be censored at the date of their last visit during the Maintenance Blinded Period. If a participant has not experienced loss of response by the time of systemic rescue, the participant will be censored at the date of systemic rescue.
Body Surface Area (BSA)	The BSA assessment estimates the extent of disease or skin involvement with respect to AD and is expressed as a percentage of total body surface. BSA will be determined by the Investigator or designee using the participant palm = 1% rule.	BSA	Use the percentage of skin affected for each region (0 to 100%) in EASI as follows: BSA Total = $0.1*BSA_{\text{head and neck}} + 0.3*BSA_{\text{trunk}} + 0.2*BSA_{\text{upper limbs}} + 0.4*BSA_{\text{lower limbs}}$	N/A – partial assessments cannot be saved.
		Change from baseline in BSA	Change from baseline: observed BSA score – baseline BSA score	Missing if baseline or observed value is missing.
SCORing Atopic Dermatitis (SCORAD)	SCORAD is a validated clinical tool for assessing the extent and intensity of atopic dermatitis. There are 3 components to the assessment: • The extent of AD is assessed as a percentage of each defined body area and reported as the sum of all areas, with a maximum score of 100%	SCORAD	SCORAD = A/5 + 7B/2 + C, where A is extent of disease, range 0-100 B is disease severity, range 0-18 C is subjective symptoms, range 0-20	Missing if components A and B are missing or if component C is missing. Partial assessments performed by physician cannot be saved and partial assessments performed by participant cannot be saved.

	(assigned as "A" in the overall SCORAD calculation).			
	• The severity of 6 specific symptoms of AD (redness, swelling, oozing/crusting, excoriation, skin thickening/lichenification, dryness) is assessed using the following scale: none (0), mild (1), moderate (2), or severe (3) (for a maximum of 18 total points, assigned as "B" in the overall SCORAD calculation). • Subjective assessment of itch and of sleeplessness is recorded for each symptom by the participant or relative on a VAS, where 0 is no itch (or sleeplessness) and 10 is the worst imaginable itch (or sleeplessness), with a maximum possible score of 20 (assigned as "C" in the overall SCORAD calculation).	Change from baseline in SCORAD Percentage change from baseline in SCORAD	Change from baseline: observed SCORAD score – baseline SCORAD score % change from baseline: $100 \times \frac{Observed\ score - Baseline}{Baseline}$	Missing if baseline or observed value is missing.
		SCORAD75	% Improvement in SCORAD from baseline ≥ 75%: % change from baseline ≤ -75	Missing if baseline or observed value is missing.
		SCORAD90	% Improvement in SCORAD from baseline ≥ 90%: % change from baseline ≤ -90	Missing if baseline or observed value is missing.

Itch Numeric Rating Scale (NRS)	The Itch Numeric Rating Scale (NRS) is a an 11-point scale used by participants to rate their worst itch severity over the past 24 hours with 0 indicating "No itch" and 10 indicating "Worst itch imaginable." Assessments will be recorded daily by the participant using an electronic diary.	Itch NRS prorated weekly mean score	The prorated weekly mean is based on previous 7 days. If a participant has at least one daily score, the weekly mean is the prorated average of daily scores within the given week. Single item; range 0-10. eCOA data are mapped to study visit per Appendix 12.	Weekly mean score missing if the participant has no Itch NRS responses within the week.
		Change from baseline in Itch NRS proratedd weekly mean score Percentage change from baseline in Itch NRS prorated weekly mean score	Change from baseline: observed Itch NRS prorated weekly mean score – baseline Itch NRS weekly mean score % change from baseline: $100 \times \frac{Observed\ score - Baseline}{Baseline}$	Missing if baseline or observed value is missing.
		4-point Itch improvement in Itch NRS prorated weekly mean score	Change from baseline ≤ -4 in Itch NRS prorated weekly mean score	Missing if baseline or observed value is missing.
		Itch NRS daily score for Day 1 through Day 15	Observed Itch NRS daily score	Missing if missing.
		Change from baseline in Itch NRS daily score for Day 1 through Day 15 Percentage change from baseline in Itch NRS daily score for Day 1 through Day 15	Change from baseline: observed Itch NRS daily score – baseline Itch NRS weekly mean score % change from baseline: $100 \times \frac{Observed\ score - Baseline}{Baseline}$	Missing if baseline or observed value is missing.
		4-point Itch improvement in Itch NRS daily score for Day 1 through Day 15	Change from baseline ≤ -4 in Itch NRS daily score	Missing if baseline or observed value is missing.

Skin Pain Numeric Rating Scale (NRS)	Skin Pain NRS is a participant-administered, 11-point horizontal scale anchored at 0 and 10, with 0 representing "no pain" and 10 representing "worst pain imaginable." Overall severity of a participant's skin pain is indicated by selecting the number that best describes the worst level of skin pain in the past 24 hours.	Skin Pain NRS prorated weekly mean score	The prorated weekly mean is based on previous 7 days. If a participant has at least one daily score, the weekly mean is the prorated average of daily scores within the given week. Single item; range 0-10. eCOA data are mapped to study visit per Appendix 12.	Weekly mean score missing if the participant has no Skin Pain NRS responses within the week.
		Change from baseline in Skin Pain NRS prorated weekly mean score Percentage change from baseline in Skin Pain NRS prorated weekly mean score	Change from baseline: observed Skin Pain NRS prorated weekly mean score – baseline Skin Pain NRS weekly mean score % change from baseline: $100 \times \frac{Observed\ score - Baseline}{Baseline}$	Missing if baseline or observed value is missing.
		4-point Skin Pain NRS improvement in Skin Pain NRS prorated weekly mean score	Change from baseline ≤ -4 in Skin Pain NRS prorated weekly mean score	Missing if baseline observed value is missing.
Sleep-loss due to itch	Sleep-loss due to itch will be assessed by the patient. Patients rate their sleep based on a 5-point Likert scale [0 (not at all) to 4 (unable to sleep at all)]. Assessments will be recorded daily by the patient using an electronic diary.	Sleep-loss prorated weekly mean score	The prorated weekly mean is based on previous 7 days. If a participant has at least one daily score within the week, the weekly mean is the prorated average of daily scores within the given week. Single item; range 0 to 4. eCOA data are mapped to study visit per Appendix 12.	Weekly mean score missing if the participant has no Sleep-loss responses within the week.
		Change from baseline in Sleep-loss prorated weekly mean score Percentage change from baseline in Sleep-loss prorated weekly mean score	Change from baseline: observed Sleep-loss prorated weekly mean score – baseline Sleep-loss weekly mean score % change from baseline: $100 \times \frac{Observed\ score - Baseline}{Baseline}$	Missing if baseline or observed value is missing.

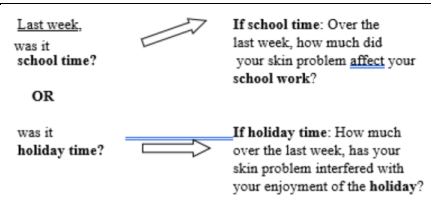
		2-point improvement in Sleep-loss prorated weekly mean score	Change from baseline in Sleep-loss prorated weekly mean score ≤ -2	Missing if baseline or observed value is missing.

		Sleep-loss daily score for Day 1 through Day 15	Observed Sleep-loss daily score	Missing if missing.
		Change from baseline in Sleep-loss daily score for Day 1 through Day 15 Percentage change from baseline in Sleep-loss daily score for Day 1 through Day 15	Change from baseline: observed Sleep-loss daily score – baseline Sleep-loss weekly mean score % change from baseline: $100 \times \frac{Observed\ score - Baseline}{Baseline}$	Missing if baseline or observed value is missing.
		2-point improvement in Sleep-loss daily score for Day 1 through Day 15	Change from baseline in Sleep-loss daily score ≤ -2	Missing if baseline or observed value is missing.
Patient-Oriented Eczema Measure (POEM)	The POEM is a 7-item, validated, questionnaire used by the participant to assess disease symptoms over the last week. The participant is asked to respond to 7 questions on skin dryness, itching, flaking, cracking, sleep loss, bleeding and weeping. All 7 answers carry equal weight with a total possible score from 0 to 28 (answers scored as: No days = 0; 1– 2 days = 1; 3-4 days = 2; 5–6 days = 3; everyday = 4). A high score is indicative of a poor quality of life. POEM responses will be captured using an electronic diary and transferred into the clinical database.	POEM	POEM total score: sum of questions 1 to 7, Range 0 to 28.	If a single question is left unanswered, then that question is scored as 0. If more than one question is unanswered, then the tool is not scored.
		Change from baseline in POEM	Change from baseline: observed POEM – baseline POEM	Missing if baseline or observed value is missing.
		4-point improvement	Change from baseline ≤ -4	Missing if baseline or observed value is missing.

Dermatology Life Quality Index (DLQI)	DLQI is a validated, dermatology-specific, patient-reported measure that evaluates participant's health-related QoL. This questionnaire has 10 items that are grouped in 6 domains, including symptoms and feelings, daily activities, leisure, work and school, personal relationships, and treatment. The recall period of this scale is over the "last week". Response categories and corresponding scores are: • Very much = 3 • A lot = 2 • A little = 1 • Not at all = 0 • Not relevant = 0 Scores range from 0-30 with higher scores indicating greater impairment of quality of life. A DLQI total score of 0 to 1 is considered as having no effect on a participant's health-related QoL (Hongbo et al. 2005), and a 4-point change from baseline is considered as the minimal clinically important difference threshold (Khilji et al. 2002; Basra et al. 2015)	DLQI total score	A DLQI total score is calculated by summing all 10 question responses and has a range of 0-30 (less to more impairment) (Finlay and Khan 1994; Basra et al. 2008).	Score of 1 unanswered question = 0; If 2 or more questions are missing, the total score is missing. Note: #7B could be a valid missing while #7A is not "No." That is, #7 should be considered as 1 question.
		DLQI (0,1)	A DLQI (0,1) response is defined as a post-baseline DLQI total score of 0 or 1. A DLQI total score of 0 to 1 is considered as having no effect on a participant's health-related QoL (Khilji et al. 2002; Hongbo et al. 2005).	Missing if DLQI total score is missing.
		4-point improvement	Change from baseline ≤ -4	Missing if baseline or observed value is missing.
		DLQI total score and domain scores change from baseline	Calculated as: observed DLQI (total score or domain scores) – baseline DLQI (total score or domain scores)	Missing if baseline or observed value is missing.
		DLQI symptoms and feelings domain	Sum of responses of questions #1 and #2: #1. How itchy, sore, painful or stinging has your skin been? #2. How embarrassed or self-conscious have you been because of your skin?	If 1 question in a domain is missing, that domain is missing.

		DLQI daily activities domain	Sum of responses of questions #3 and #4: #3. How much has your skin interfered with you going shopping or looking after your home or garden? #4. How much has your skin influenced the clothes you wear?	If 1 question in a domain is missing, that domain is missing.
		DLQI leisure domain	Sum of responses of questions #5 and #6: #5. How much has your skin affected any social or leisure activities? #6. How much has your skin make it difficult for you to do any sport?	If 1 question in a domain is missing, that domain is missing.
		DLQI work and school domain	Sum of responses of questions #7A and #7B: #7A. Has your skin prevented you from working or studying? #7B. If No: how much has your skin been a problem at work or studying?	If the answer to question #7A is missing, this domain is missing. If #7A is No, and #7B is missing, this domain is missing.
		DLQI personal relationships domain	Sum of responses of questions #8 and #9: #8. How much has your skin created problems with your partner or any of your close friends or relatives? #9. How much has your skin caused any sexual difficulties?	If 1 question in a domain is missing, that domain is missing.
		DLQI treatment domain	Response of question #10: #10. How much of a problem has the treatment for your skin been, for example by making your home messy, or by taking up time?	If 1 question in a domain is missing, that domain is missing.
Children's Dermatology Life Quality	The CDLQI is designed to measure the impact of any skin disease on the lives of children. Participants ≤16 years will complete the CDLQI and should	CDLQI total score	A CDLQI total score is calculated by summing all 10 question responses and has a range of 0-30 (less to more impairment) (Waters et al. 2010).	Score of 1 unanswered question = 0; If 2 or more questions are missing, the total score is missing.

Index (CDLQI)	continue to complete the CDLQI for the duration of the study. The scoring of each question is: • Very much = 3 • Quite a lot = 2 • Only a little = 1 • Not at all = 0 • Question unanswered = 0 • Question 7: 'Prevented school' (text-only questionnaire) = 3	CDLQI (0,1)	A CDLQI (0,1) response is defined as a post-baseline CDLQI total score of 0 or 1.	Missing if CLQI total score is missing.
		4-point improvement	Change from baseline ≤ -4	Missing if baseline or observed value is missing.
		CDLQI total score and domain scores change from baseline	Calculated as: observed CDLQI (total score or domain scores) – baseline CDLQI (total score or domain scores)	Missing if baseline or observed value is missing.
		CDLQI symptoms and feelings domain	Sum of responses of questions #1 and #2: #1. Over the last week, how itchy, "scratchy", sore, or painful has your skin been? #2. Over the last week, how embarrassed or self-conscious, upset, or sad have you been because of your skin?	If 1 question in a domain is missing, that domain is missing.
		CDLQI sleep	Responses of questions 9 #9. Over the last week, how much has your sleep been affected by your skin problem?	Single item, missing if missing.
		CDLQI leisure domain	Sum of responses of questions #4, #5 and #6: #4. Over the last week, how much have you changed or worn different or special clothes/shoes because of your skin? #5. Over the last week, how much has your skin trouble affected going out, playing, or doing hobbies? #6. Over the last week, how much have you avoided swimming or other sports because of your skin trouble?	If 1 question in a domain is missing, that domain is missing.
		CDLQI school or holiday domain	Responses of questions 7: If select 'Prevented school,' score = 3	Single item, missing if missing.

				
		CDLQI personal relationships domain	Sum of responses of questions #3 and #8: #3: Over the last week, how much has your skin affected your friendships? #8. Over the last week, how much trouble have you had because of your skin with other people calling you names, teasing, bullying, asking questions or avoiding you?	If 1 question in a domain is missing, that domain is missing.
		CDLQI treatment domain	Response of question #10: #10. How much of a problem has the treatment for your skin been?	Single item, missing if missing.

Work Productivity and Activity Impairment: Atopic Dermatitis (WPAI-AD)	The Work Productivity and Activity Impairment Questionnaire – Atopic Dermatitis (WPAI-AD) records impairment due to AD during the past 7 days. The WPAI-AD consists of 6 items grouped into 4 domains: absenteeism (work time missed), presenteeism (impairment at work/reduced on the job effectiveness), work productivity loss (overall work impairment/absenteeism plus presenteeism), and activity impairment. Scores are calculated as impairment percentages (Reilly et al. 1993), with higher scores indicating greater impairment and less productivity.	Employment status	Question (Q)1	Single item, missing if missing.
		Change in employment status	Employed at baseline and remained employed: Q1 = 1 at post-baseline visit and at baseline visit. Not employed at baseline and remain unemployed: Q1 = 0 at post-baseline visit and at baseline visit.	Missing if baseline or observed value is missing.
		Percentage of absenteeism	Percent work time missed due to problem: (Q2/(Q2 + Q4))*100	If Q2 or Q4 is missing, then missing.
		Change from baseline in absenteeism	Change from baseline: observed absenteeism – baseline absenteeism	Missing if baseline or observed value is missing.
		Percentage of presenteeism	Percent impairment (reduced productivity while at work) while working due to problem: (Q5/10)*100	If Q5 is missing, then missing.
		Change from baseline in presenteeism	Change from baseline: observed presenteeism – baseline absenteeism	Missing if baseline or observed value is missing.
		Overall work impairment	Percent overall work impairment (combines absenteeism and presenteeism) due to problem: (Q2/(Q2 + Q4) + [(1 - Q2/(Q2+Q4))*(Q5/10)])*100	If Q2, Q4, or Q5 is missing, then missing.

		Change from baseline in work impairment	Change from baseline: observed overall work impairment – baseline overall work impairment	Missing if baseline or observed value is missing.
		Percentage of activity impairment	Percent activity impairment (performed outside of work) due to problem: (Q6/10)*100	If Q6 is missing, then missing.
		Change from baseline in activity impairment	Change from baseline: observed activity impairment – baseline activity impairment	Missing if baseline or observed value is missing.
Hospital Anxiety Depression Scale (HADS)	The Hospital Anxiety Depression Scale (HADS) is a 14-item self-assessment scale that determines the levels of anxiety and depression that a participant is experiencing over the past week. The HADS utilizes a 4-point Likert scale (e.g., 0 to 3) for each question and is intended for ages 12 to 65 years (Zigmond and Snaith 1983; White et al. 1999). Scores for each domain (anxiety and depression) can range from 0 to 21, with higher scores indicating greater anxiety or depression (Zigmond and Snaith 1983; Snaith 2003).	HADS domain scores for anxiety and depression	Anxiety domain score is sum of the seven anxiety questions, range 0 to 21; Depression domain score is sum of the seven depression questions, range 0 to 21.	N/A – partial assessments cannot be saved.
		Change from baseline in HADS total score, anxiety and depression domain scores	Change from baseline: observed HADS total/domain score – baseline HADS total/domain score	Missing if baseline or observed value is missing.
Asthma Control Questionnaire (ACQ-5)		ACQ-5 total score	An ACQ-5 total score is the mean score of all 5 questions.	If more than 1 question is missing, the ACQ-5 total score is missing.

	Participants who report comorbid asthma prior to enrollment will complete the Asthma Control Questionnaire in addition to other patient reported outcomes in this trial. The ACQ-5 has been shown to reliably measure asthma control and distinguish participants with well-controlled asthma (score ≤ 0.75 points) from those with uncontrolled asthma (score ≥ 1.5 points). It consists of 5 questions that are scored on a 7-point Likert scale with a recall period of 1 week. The total ACQ-5 score is the mean score of all questions; a lower score represents better asthma control.	Change from baseline in ACQ-5 score	Change from baseline: observed ACQ-5 total score – baseline ACQ-5 total score	Missing if baseline or observed value is missing.
		Minimal Clinically Important Difference (MCID) of 0.5	Change from baseline ≤ -0.5	Missing if baseline or observed value is missing.
Topical corticosteroid (TCS) or topical calcineurin inhibitor (TCI) Use	A mid-potency TCS, locoid ointment 0.1%, and a low-potency TCS, prednisolone cream 0.5% (for use on sensitive skin areas) will be provided by the Sponsor for use in this trial. Participants are to be instructed to return all used and unused TCS medication (tubes) to the study site for accountability purposes.	Time (days) to TCS/TCI-free use from Baseline to Week 16 for the Induction Period, and from Baseline to Week 68 for the Combined Induction and Maintenance Blinded Periods	Days from the first study drug injection to the day participant stop using all TCS/TCI (if a participant starts and stops using low or mid potency TCS/TCI multiple times, use the last stop date as the stop date for this participant.)	If do not stop using the TCS/TCI, the participant will be censored at the date of their last visit
		Proportion of TCS/TCI-free days from Baseline to Week 16 for the Induction Period, and from Baseline to Week 68 for the Combined Induction and Maintenance Blinded Periods	100*(Number of the total TCS/TCI free days divided by total number of days during the treatment period)	N/A

		Mean gram quantity (tube weights) of low and moderate-potency TCS used from Baseline to Week 16 for the Induction Period, and from Baseline to Week 68 for the Combined Induction and Maintenance Blinded Periods	Weight each tube after participants returned. The weight of TCS use calculated as standard weight of each tube by supplier minus the weight of the used tube. If a returned tube is not weighed or not returned, then the tube can be classified as partially used, fully used, unused, or unknown. Partially used will be defined as 50% used whereas fully used and unused tubes will be defined as 100% and 0% used, respectively. Unknown will be treated as missing.	Missing data will be treated as missing.
Thymus and activation-regulated chemokine (TARC)		Observed TARC value	Single value	Missing if missing.
		Change from baseline in TARC	Change from baseline: observed TARC value – baseline TARC value	Missing if baseline or observed value is missing.

6.2. Appendix 2: Description of Efficacy and Health Outcome Analyses

Description of Efficacy and Health Outcome Analyses

Measure	Variable	Estimand (Section 1.1.1)	Analysis Method (Section 4.1)	Population (Section 3)	Comparison/Time Point	Analysis Type
Investigator's Global Assessment (IGA)	Proportion of participants achieving IGA (0,1) with a ≥ 2-point improvement	Primary Estimand (Hybrid)	CMH analysis with MCMC-MI	ITT	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; Week 16 and all scheduled visits in the Induction Period	Primary analysis: Week 16; Secondary analysis: other timepoints
				PPS	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; Week 16	Supplementary analysis
			CMH analysis with tipping point analysis	ITT	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; Week 16	Sensitivity analysis
		Supportive Estimand (Composite)	CMH analysis with NRI	ITT	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; Week 16 and all scheduled visits in the Induction Period	Supplementary analysis
	Proportion of participants achieving IGA (0)	Supportive Estimand (Composite)	CMH analysis with NRI	ITT	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; Week 16 and all scheduled visits in the Induction Period	Secondary analysis
	Proportion of participants achieving both IGA (0,1) with a ≥ 2-point improvement and a ≥ 4-point improvement in Itch NRS prorated weekly mean score	Supportive Estimand (Composite)	CMH analysis with NRI	ITT with baseline Itch NRS ≥ 4	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; Week 16 and all scheduled visits in the Induction Period	Secondary analysis
	Maintenance of IGA (0,1):	Maintenance Primary Estimand (Hybrid)	Descriptive statistics with MCMC-MI		No comparisons. All scheduled visits in the	Secondary analysis

Measure	Variable	Estimand (Section 1.1.1)	Analysis Method (Section 4.1)	Population (Section 3)	Comparison/Time Point	Analysis Type
	Proportion of participants maintaining IGA (0,1) with a ≥ 2-point improvement from baseline among those re-randomized participants who achieved IGA (0,1) with a ≥ 2-point improvement from baseline at Week 16	Maintenance Supportive Estimand (Hybrid)	Descriptive statistics with MCMC-MI	MPP who have achieved IGA (0,1) with a ≥ 2-point improvement from baseline at Week 16	Maintenance Period	Supplementary analysis
		Maintenance Supportive Estimand (Composite)	Descriptive statistics with NRI			Supplementary analysis
	Time to loss of IGA (0,1)	N/A	KM method	MPP who have achieved IGA (0,1) with a ≥ 2-point improvement from baseline at Week 16	No comparisons.	Secondary analysis
	Proportion of participants with IGA (0,1) with a ≥ 2-point improvement from baseline	N/A	Descriptive statistics	MSP	No comparisons. All scheduled visits in the Maintenance Period	Secondary analysis
	Proportion of participants with IGA (0,1) with a ≥ 2-point improvement from baseline	N/A	Descriptive statistics	Maintenance W16 Escape Population	No comparisons. All scheduled visits in the Maintenance Period	Secondary analysis
	Proportion of participants with IGA (0,1) with a ≥ 2-point improvement from baseline after lebrikizumab re-treatment	N/A	Descriptive statistics	Maintenance Period W20-64 Escape Population	No comparisons. Every 4 weeks after escape and re-treated by lebrikizumab 250mg Q2W	Secondary analysis

Measure	Variable	Estimand (Section 1.1.1)	Analysis Method (Section 4.1)	Population (Section 3)	Comparison/Time Point	Analysis Type
Eczema Area and Severity Index (EASI)	Change from baseline in EASI ($EASI_{region}$) score Percentage change from baseline in EASI score	Primary Estimand (Hybrid)	ANCOVA with MCMC-MI for EASI score only	ITT	Leb 250 mg Q4W vs PBO; Leb 250 mg Q2W vs PBO; Week 16 and all scheduled visits in the Induction Period	Key secondary analysis: percentage change at Week 16; Secondary analysis: other timepoints
		Supportive Estimand (Hypothetical)	MMRM for EASI score and $EASI_{region}$ score, respectively	ITT	Leb 250 mg Q4W vs PBO; Leb 250 mg Q2W vs PBO; Week 16 and all scheduled visits in the Induction Period	Supplementary analysis for EASI score Secondary analysis for $EASI_{region}$ score
		Maintenance Primary Estimand (Hybrid)	Descriptive statistics with MCMC-MI for EASI score only	MPP	No comparisons. All scheduled visits in the Maintenance Period	Secondary analysis
		Maintenance Supportive Estimand (Hybrid)	Descriptive statistics with MCMC-MI for EASI score only			Supplementary analysis
		Maintenance Supportive Estimand (Hypothetical)	Descriptive statistics with LOCF for EASI score and $EASI_{region}$ score, respectively			Supplementary analysis for EASI score Secondary analysis for $EASI_{region}$ score
		N/A	Descriptive statistics for EASI score only	Maintenance W16 Escape Population	No comparisons. All scheduled visits in the Maintenance Period	Secondary analysis
		Supportive Estimand (Hypothetical)	MMRM	ITT	Leb 250 mg Q4W vs PBO; Leb 250 mg Q2W vs PBO; Week 16 and all scheduled visits in the Induction Period	Secondary analysis

Measure	Variable	Estimand (Section 1.1.1)	Analysis Method (Section 4.1)	Population (Section 3)	Comparison/Time Point	Analysis Type
	Change from baseline in each symptom score (i.e., erythema, edema/papulation, excoriation, lichenification) and percentage involvement score by body regions (i.e., head and neck, trunk, upper limbs, and lower limbs)	Maintenance Supportive Estimand (Hypothetical)	Descriptive statistics with LOCF	MPP	No comparisons. All scheduled visits in the Maintenance Period	Secondary analysis
	Proportion of participants achieving EASI-75 Proportion of participants achieving EASI-90	Primary Estimand (Hybrid)	CMH analysis with MCMC-MI	ITT	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; Week 16 and all scheduled visits in the Induction Period	Primary analysis: EASI-75 at Week 16; Key secondary analysis: EASI-90 at Week 16; Secondary analysis: other timepoints
				PPS	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; Week 16	Supplementary analysis (EASI-75 only)
			CMH analysis with tipping point analysis	ITT	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; Week 16	Sensitivity analysis
		Supportive Estimand (Composite)	CMH analysis with NRI	ITT	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; Week 16 and all scheduled visits in the Induction Period	Supplementary analysis

Measure	Variable	Estimand (Section 1.1.1)	Analysis Method (Section 4.1)	Population (Section 3)	Comparison/Time Point	Analysis Type
	Proportion of participants achieving EASI-50	Supportive Estimand (Composite)	CMH analysis with NRI	ITT	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; Week 16 and all scheduled visits in the Induction Period	Secondary analysis
	Proportion of participants achieving both EASI-75 and a ≥ 4-point improvement in Itch NRS prorated weekly mean score	Supportive Estimand (Composite)	CMH analysis with NRI	ITT with baseline Itch NRS ≥ 4	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; Week 16 and all scheduled visits in the Induction Period	Secondary analysis
	Proportion of participants maintaining EASI-50 (only for Maintenance Supportive Estimand [Composite]) and EASI-75 or achieving EASI-90 among those re-randomized participants who achieved EASI-75 at Week 16	Maintenance Primary Estimand (Hybrid)	Descriptive statistics with MCMC-MI	MPP who have achieved EASI-75 at Week 16	No comparisons. All scheduled visits in the Maintenance Period	Secondary analysis
		Maintenance Supportive Estimand (Hybrid)	Descriptive statistics with MCMC-MI			Supplementary analysis
		Maintenance Supportive Estimand (Composite)	Descriptive statistics with NRI			Supplementary analysis
	Time to loss of EASI-50	N/A	KM method	MPP	No comparisons.	Secondary analysis
	Time to loss of EASI-75	N/A	KM method	MPP who have achieved EASI-75 at Week 16	No comparisons.	Secondary analysis
	Proportion of participants with EASI-75	N/A	Descriptive statistics	MSP	No comparisons. All scheduled visits in the Maintenance Period	Secondary analysis
	Proportion of participants with EASI-75	N/A	Descriptive statistics	Maintenance W16 Escape Population	No comparisons. All scheduled visits in the Maintenance Period	Secondary analysis

Measure	Variable	Estimand (Section 1.1.1)	Analysis Method (Section 4.1)	Population (Section 3)	Comparison/Time Point	Analysis Type
	Proportion of participants with EASI-75 after lebrikizumab re-treatment	N/A	Descriptive statistics	Maintenance Period W20-64 Escape Population	No comparisons. Every 4 weeks after escape and re-treated by lebrikizumab 250 mg Q2W	Secondary analysis
Body Surface Area (BSA) Affected by AD	Change from baseline in BSA	Supportive Estimand (Hypothetical)	MMRM	ITT	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; Week 16 and all scheduled visits in the Induction Period	Secondary analysis
		Maintenance Supportive Estimand (Hypothetical)	Descriptive statistics with LOCF	MPP	No comparisons. All scheduled visits in the Maintenance Period	Secondary analysis
		N/A	Descriptive statistics	Maintenance W16 Escape Population	No comparisons. All scheduled visits in the Maintenance Period	Secondary analysis
Itch Numeric Rating Scale (NRS)	Change from baseline in Itch NRS prorated weekly mean score Percentage change from baseline in Itch NRS prorated weekly mean score	Primary Estimand (Hybrid)	ANCOVA with MCMC-MI	ITT	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; Week 16 and all scheduled visits in the Induction Period	Secondary analysis
		Supportive Estimand (Hypothetical)	MMRM	ITT	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; Week 16 and all scheduled visits in the Induction Period	Supplementary analysis
		Maintenance Primary Estimand (Hybrid)	Descriptive statistics with MCMC-MI	MPP	No comparisons. All scheduled visits in the Maintenance Period	Secondary analysis
		Maintenance Supportive Estimand (Hybrid)	Descriptive statistics with MCMC-MI			Supplementary analysis

Approved on 22 Aug 2022 GMT

Measure	Variable	Estimand (Section 1.1.1)	Analysis Method (Section 4.1)	Population (Section 3)	Comparison/Time Point	Analysis Type
		Maintenance Supportive Estimand (Hypothetical)	Descriptive statistics with LOCF			Supplementary analysis
		N/A	Descriptive statistics	Maintenance W16 Escape Population	No comparisons. All scheduled visits in the Maintenance Period	Secondary analysis
	Change from baseline in Itch NRS daily score for Day 1 through Day 15 Percentage change from baseline in Itch NRS daily score for Day 1 through Day 15	Supportive Estimand (Hypothetical)	MMRM	ITT	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; Day 1 through Day 15 in the Induction Period	Secondary analysis
	Proportion of participants achieving at least 4-point improvement in Itch NRS prorated weekly mean score in participants who had baseline Itch NRS ≥ 4	Primary Estimand (Hybrid)	CMH analysis with MCMC-MI	ITT with baseline Itch NRS ≥ 4	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; Week 16 and all scheduled visits in the Induction Period	Key secondary analysis: Weeks 1, 2, 4, and 16; Secondary analysis: other timepoints
			CMH analysis with tipping point analysis	ITT with baseline itch NRS ≥ 4	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; Weeks 1, 2, 4 and 16	Sensitivity analysis
		Supportive Estimand (Composite)	CMH analysis with NRI	ITT with baseline Itch NRS ≥ 4	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; Week 16 and all scheduled visits in the Induction Period	Supplementary analysis

Measure	Variable	Estimand (Section 1.1.1)	Analysis Method (Section 4.1)	Population (Section 3)	Comparison/Time Point	Analysis Type
	Proportion of participants achieving at least 4-point improvement in Itch NRS daily score in participants who had baseline Itch NRS ≥ 4 for Day 1 through Day 15	Supportive Estimand (Composite)	CMH analysis with NRI	ITT with baseline Itch NRS ≥ 4	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; Day 1 through Day 15	Secondary analysis
	Proportion of participants maintaining ≥ 4-point reduction from baseline among those participants with Itch NRS of ≥ 4-point at baseline who achieved ≥ 4-point reduction from baseline at Week 16	Maintenance Primary Estimand (Hybrid)	Descriptive statistics with MCMC-MI	MPP with Itch NRS of ≥ 4-points at baseline and who achieved ≥ 4-point reduction from baseline at Week 16	No comparisons. All scheduled visits in the Maintenance Period	Secondary analysis
		Maintenance Supportive Estimand (Hybrid)	Descriptive statistics with MCMC-MI			Supplementary analysis
		Maintenance Supportive Estimand (Composite)	Descriptive statistics with NRI			Supplementary analysis
	Proportion of participants with ≥ 4-point reduction from baseline among those participants with Itch NRS of ≥ 4-point at baseline	N/A	Descriptive statistics	MSP with Itch NRS of ≥ 4-points at baseline	No comparisons. All scheduled visits in the Maintenance Period	Secondary analysis
	Proportion of participants with ≥ 4-point reduction from baseline among those participants with Itch NRS of ≥ 4-point at baseline	N/A	Descriptive statistics	Maintenance W16 Escape Population with Itch NRS of ≥ 4-points at baseline	No comparisons. All scheduled visits in the Maintenance Period	Secondary analysis

Approved on 22 Aug 2022 GMT

Measure	Variable	Estimand (Section 1.1.1)	Analysis Method (Section 4.1)	Population (Section 3)	Comparison/Time Point	Analysis Type
	Proportion of participants with ≥ 4-point reduction from baseline after lebrikizumab re-treatment among those participants with Itch NRS of ≥ 4-point at baseline	N/A	Descriptive statistics	Maintenance W20-64 Period Escape Population with Itch NRS of ≥ 4-points at baseline	No comparisons. Every 4 weeks after escape and re-treated by lebrikizumab 250mg Q2W	Secondary analysis
Skin Pain Numeric Rating Scale (NRS)	Change from baseline in Skin Pain NRS Percentage change from baseline in Skin Pain NRS score	Supportive Estimand (Hypothetical)	MMRM	ITT	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; Week 16 and all scheduled visits in the Induction Period	Secondary analysis
		Maintenance Supportive Estimand (Hypothetical)	Descriptive statistics with LOCF	MPP	No comparisons. All scheduled visits in the Maintenance Period	Supplementary analysis
	Proportion of participants with ≥ 4-point reduction from baseline among those participants with Skin Pain NRS of ≥ 4-point at baseline	Supportive Estimand (Composite)	CMH analysis with NRI	ITT with baseline Skin Pain NRS score ≥ 4	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; Week 16 and all scheduled visits in the Induction Period	Secondary analysis
	Proportion of participants maintaining ≥ 4-point reduction from baseline among those participants with Skin Pain NRS of ≥ 4-point at baseline and who achieved ≥ 4-point reduction from baseline at Week 16	Maintenance Supportive Estimand (Composite)	Descriptive statistics with NRI	MPP with Skin Pain NRS of ≥ 4-points at baseline and who achieved ≥ 4-point reduction from baseline at Week 16	No comparisons. All scheduled visits in the Maintenance Period	Secondary analysis

Measure	Variable	Estimand (Section 1.1.1)	Analysis Method (Section 4.1)	Population (Section 3)	Comparison/Time Point	Analysis Type
	Proportion of participants with ≥ 4-point reduction from baseline among those participants with Skin Pain NRS of ≥ 4-point at baseline	N/A	Descriptive statistics	Maintenance W16 Escape Population with baseline Skin Pain NRS score ≥ 4	No comparisons. All scheduled visits in the Maintenance Period	Secondary analysis
Sleep-loss	Change from baseline in Sleep-loss prorated weekly mean score Percentage change from baseline in Sleep-loss prorated weekly mean score	Supportive Estimand (Hypothetical)	MMRM	ITT	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; Week 16 and all scheduled visits in the Induction Period	Secondary analysis
		Maintenance Supportive Estimand (Hypothetical)	Descriptive statistics with LOCF	MPP	No comparisons. All scheduled visits in the Maintenance Period	Secondary analysis
	Change from baseline in Sleep-loss daily score for Day 1 through Day 15 Percentage change from baseline in Sleep-loss daily score for Day 1 through Day 15	Supportive Estimand (Hypothetical)	MMRM	ITT	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; Day 1 through Day 15 in the Induction Period	Secondary analysis
	Proportion of participants achieving at least 2-point improvement in Sleep-loss prorated weekly mean score in participants who had baseline Sleep-loss ≥ 2	Supportive Estimand (Composite)	CMH analysis with NRI	ITT with baseline Sleep-loss ≥ 2	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; Week 16 and all scheduled visits in the Induction Period	Secondary analysis
	Proportion of participants achieving at least 2-point improvement in Sleep-loss prorated weekly mean score in participants who had baseline Sleep-loss ≥ 2	Maintenance Supportive Estimand (Composite)	Descriptive statistics with NRI	MPP with baseline Sleep-loss ≥ 2	No comparisons. All scheduled visits in the Maintenance Period	Secondary analysis

Measure	Variable	Estimand (Section 1.1.1)	Analysis Method (Section 4.1)	Population (Section 3)	Comparison/Time Point	Analysis Type
	Proportion of participants achieving at least 2-point improvement in Sleep-loss prorated weekly mean score in participants who had baseline Sleep-loss ≥ 2	N/A	Descriptive statistics	Maintenance W16 Escape Population with baseline Sleep-loss ≥ 2	No comparisons. All scheduled visits in the Maintenance Period	Secondary analysis
	Proportion of participants achieving at least 2-point improvement in Sleep-loss daily score in participants who had baseline Sleep-loss ≥ 2 for Day 1 through Day 15	Supportive Estimand (Composite)	CMH analysis with NRI	ITT with baseline Sleep-loss ≥ 2	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; Day 1 through Day 15 in the Induction Period	Secondary analysis
(Children) Dermatology Life Quality Index (DLQI/ CDLQI)	Change from baseline in DLQI total and domain scores Change from baseline in CDLQI total and domain scores	Supportive Estimand (Hypothetical)	MMRM	ITT	Leb 250 mg Q4W vs PBO; Leb 250 mg Q2W vs PBO; Week 16 and all scheduled visits in the Induction Period	Secondary analysis
		Maintenance Supportive Estimand (Hypothetical)	Descriptive statistics with LOCF	MPP	No comparisons. All scheduled visits in the Maintenance Period	Secondary analysis
		N/A	Descriptive statistics	Maintenance W16 Escape Population	No comparisons. All scheduled visits in the Maintenance Period	Secondary analysis
		Supportive Estimand (Composite)	CMH analysis with NRI	ITT with baseline DLQI/CDLQI total score ≥ 4	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; Week 16 and all scheduled visits in the Induction Period	Secondary analysis

Measure	Variable	Estimand (Section 1.1.1)	Analysis Method (Section 4.1)	Population (Section 3)	Comparison/Time Point	Analysis Type
	Proportion of participants achieving at least 4-point improvement in DLQI/CDLQI total score in participants who had baseline DLQI/CDLQI total score ≥ 4	Maintenance Supportive Estimand (Composite)	Descriptive statistics with NRI	MPP with baseline DLQI/CDLQI total score ≥ 4 who have achieved ≥ 4-point improvement in DLQI/CDLQI total score at Week 16	No comparisons. All scheduled visits in the Maintenance Period	Secondary analysis
		N/A	Descriptive statistics	Maintenance W16 Escape Population with baseline DLQI/CDLQI total score ≥ 4	No comparisons. All scheduled visits in the Maintenance Period	Secondary analysis
	Proportion of participants achieving DLQI (0,1)/CDLQI (0,1)	Supportive Estimand (Composite)	CMH analysis with NRI	ITT	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; Week 16 and all scheduled visits in the Induction Period	Secondary analysis
		Maintenance Supportive Estimand (Composite)	Descriptive statistics with NRI	MPP	No comparisons. All scheduled visits in the Maintenance Period	Secondary analysis
		N/A	Descriptive statistics	Maintenance W16 Escape Population	No comparisons. All scheduled visits in the Maintenance Period	No comparisons. All scheduled visits in the Maintenance Period

Measure	Variable	Estimand (Section 1.1.1)	Analysis Method (Section 4.1)	Population (Section 3)	Comparison/Time Point	Analysis Type
SCORing Atopic Dermatitis (SCORAD)	Change from baseline in SCORAD Percentage change from baseline in SCORAD	Supportive Estimand (Hypothetical)	MMRM	ITT	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; Week 16 and all scheduled visits in the Induction Period	Secondary analysis
		Maintenance Supportive Estimand (Hypothetical)	Descriptive statistics with LOCF	MPP	No comparisons. All scheduled visits in the Maintenance Period	Secondary analysis
		N/A	Descriptive statistics	Maintenance W16 Escape Population	No comparisons. All scheduled visits in the Maintenance Period	Secondary analysis
	Proportion of participants achieving SCORAD75 Proportion of participants achieving SCORAD90	Supportive Estimand (Composite)	CMH analysis with NRI	ITT	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; Week 16 and all scheduled visits in the Induction Period	Secondary analysis
		Maintenance Supportive Estimand (Composite)	Descriptive statistics with NRI	MPP	No comparisons. All scheduled visits in the Maintenance Period	Secondary analysis
		N/A	Descriptive statistics	Maintenance W16 Escape Population	No comparisons. All scheduled visits in the Maintenance Period	Secondary analysis
	Percentage change in SCORAD from baseline in participants who achieved EASI-75 at Week 16	Maintenance Supportive Estimand (Hypothetical)	Descriptive statistics with LOCF	MPP who have achieved EASI-75 at Week 16	No comparisons. All scheduled visits in the Maintenance Period	Secondary analysis
	Change from baseline in POEM	Supportive Estimand (Hypothetical)	MMRM	ITT	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; Week 16 and all scheduled visits in the Induction Period	Secondary analysis

Measure	Variable	Estimand (Section 1.1.1)	Analysis Method (Section 4.1)	Population (Section 3)	Comparison/Time Point	Analysis Type
Patient-Oriented Eczema Measure (POEM)		Maintenance Supportive Estimand (Hypothetical)	Descriptive statistics with LOCF	MPP	No comparisons. All scheduled visits in the Maintenance Period	Secondary analysis
		N/A	Descriptive statistics	Maintenance W16 Escape Population	No comparisons. All scheduled visits in the Maintenance Period	Secondary analysis
Work Productivity and Activity Impairment: Atopic Dermatitis (WPAI-AD)	Observed and change from baseline in employment status	N/A	Descriptive statistics with observed data	ITT	No comparisons. All scheduled visits in the Induction Period	Secondary analysis
		N/A	Descriptive statistics with observed data	MPP	No comparisons. All scheduled visits in the Maintenance Period	Secondary analysis
	Change from baseline in: • absenteeism • presenteeism • overall work impairment • impairment in activities	Supportive Estimand (Hypothetical)	MMRM	ITT	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; Week 16 and all scheduled visits in the Induction Period	Secondary analysis
		Maintenance Supportive Estimand (Hypothetical)	Descriptive statistics with LOCF	MPP	No comparisons. All scheduled visits in the Maintenance Period	Secondary analysis
		N/A	Descriptive statistics	Maintenance W16 Escape Population	No comparisons. All scheduled visits in the Maintenance Period	Secondary analysis
	Change from baseline in HADS total score, anxiety and depression domain scores	Supportive Estimand (Hypothetical)	MMRM	ITT	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; Week 16 and all scheduled visits in the Induction Period	Secondary analysis

Measure	Variable	Estimand (Section 1.1.1)	Analysis Method (Section 4.1)	Population (Section 3)	Comparison/Time Point	Analysis Type
Hospital Anxiety Depression Scale (HADS)		Maintenance Supportive Estimand (Hypothetical)	Descriptive statistics with LOCF	MPP	No comparisons. All scheduled visits in the Maintenance Period	Secondary analysis
		N/A	Descriptive statistics	Maintenance W16 Escape Population	No comparisons. All scheduled visits in the Maintenance Period	Secondary analysis
Asthma Control Questionnaire (ACQ-5)	Change from baseline in ACQ-5 score	Supportive Estimand (Hypothetical)	ANCOVA with LOCF	ITT with self-reported comorbid asthma	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; Week 16 in the Induction Period	Secondary analysis
		Maintenance Supportive Estimand (Hypothetical)	Descriptive statistics with LOCF	MPP with self-reported comorbid asthma	No comparisons. All scheduled visits in the Maintenance Period	Secondary analysis
Topical corticosteroid (TCS) or topical calcineurin inhibitor (TCI) Use	Time (days) to TCS/TCI-free use from Baseline to Week 16	N/A	KM method with log-rank test	ITT	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO;	Secondary analysis
	Time (days) to TCS/TCI-free use from Baseline to Week 68	N/A	KM method	MPP	No comparisons.	Secondary analysis
	Proportion of TCS/TCI-free days from Baseline to Week 16	N/A	Descriptive statistics with observed data Two-sample t-test and ANOVA including treatment group and stratification factors with observed data	ITT	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; Baseline through Week 16 in the Induction Period	Secondary analysis

Measure	Variable	Estimand (Section 1.1.1)	Analysis Method (Section 4.1)	Population (Section 3)	Comparison/Time Point	Analysis Type
	Proportion of TCS/TCI-free days from Baseline to Week 68	N/A	Descriptive statistics with observed data	MPP	No comparisons. Baseline through Week 68 in the Combined Induction and Maintenance Periods	Secondary analysis
	Mean weight of TCS use by potency (tube weights) from Baseline to Week 16	N/A	Descriptive statistics with observed data Two-sample t-test and ANOVA including treatment group and stratification factors with observed data	ITT	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; Baseline through Week 16 in the Induction Period	Secondary analysis
	Mean weight of TCS use by potency (tube weights) from Baseline to Week 68	N/A	Descriptive statistics with observed data	MPP	No comparisons. Baseline through Week 68 in the Combined Induction and Maintenance Periods	Secondary analysis
Thymus and activation-regulated chemokine (TARC)	Observed TARC value	N/A	Descriptive statistics and box plots with observed data	ITT	No comparisons. All scheduled visits in the Induction Period	Secondary analysis
				MPP	No comparisons. All scheduled visits in the Maintenance Period	Secondary analysis
	Change from baseline in TARC	N/A	Descriptive statistics and box plots with observed data ANCOVA with LOCF	ITT	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; Week 16 and all scheduled visits in the Induction Period	Secondary analysis

Measure	Variable	Estimand (Section 1.1.1)	Analysis Method (Section 4.1)	Population (Section 3)	Comparison/Time Point	Analysis Type
		N/A	Descriptive statistics and box plots with observed data	MPP	No comparisons. All scheduled visits in the Maintenance Period	Secondary analysis

Abbreviations: ANCOVA = analysis of covariance; ANOVA = analysis of variance; CMH = Cochran-Mantel-Haenszel; ITT = intent-to-treat; KM = Kaplan-Meier; Leb = lebrikizumab; LOCF = last observation carried forward; MCMC-MI = Markov chain Monte Carlo multiple imputation; MMRM = mixed-effects model for repeated measures; N/A = not applicable; NRI = non-responder imputation; PBO = placebo; PPS = per protocol set; Q2W = every 2 weeks; Q4W = every 4 weeks; VAS = Visual Analog Scale; W = week.

6.3. Appendix 3: Demographic and Baseline Characteristics

Participant demographic variables and baseline characteristics will be summarized by treatment group for the ITT Population, the MPP, and the Maintenance W16 Escape Population. The summary will also be created by treatment group in the ITT Population by participants who experienced or did not experience conjunctivitis adverse events defined using Customized MedDRA Query (CMQ) PTs as described in the Compound Level Safety Standards during the Induction Period. In addition, participant disease characteristics at Week 16 will be summarized by treatment group for the MPP. The continuous variables will be summarized using descriptive statistics and the categorical variables will be summarized using frequency counts and percentages. No formal statistical comparisons will be made between treatment groups unless otherwise specified.

The following demographic information will be included:

- Age
- Age group (Adolescents (12<18), Adults ≥ 18)
- Age group (Adolescents (12<18), Adults ≥ 18 - < 65, ≥ 65 - < 75, ≥ 75)
- Sex (male, female)
- Race (American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, White, Multiple, Other, Not Reported)
- Weight (kg)
- Weight category (< 60 kg, ≥ 60 to < 100 kg, ≥ 100 kg)
- Height (cm)
- Body mass index (BMI) (kg/m^2)
- BMI category: Underweight (< 18.5 kg/m^2), Normal (≥ 18.5 and < 25 kg/m^2), Overweight (≥ 25 and < 30 kg/m^2), Obese (≥ 30 and < 40 kg/m^2), Extreme obese (≥ 40 kg/m^2)
- Alcohol use (Never, Current, Former)
- Tobacco use (Never, Current, Former)

By-participant listings of basic demographic information for the ITT Population will be provided.

The following baseline disease/clinical characteristics will be included:

- Age at onset (years): calculated as the difference between date of onset of AD and the date of birth
- Duration since AD onset (years): calculated as the difference between date of Informed Consent and the date of onset of AD
- Duration since AD onset category (0 to < 2 years, 2 to < 5 years, 5 to < 10 years, 10 to < 20 years, ≥ 20 years)

- Anatomical area affected by atopic dermatitis:
  - Head
  - Trunk (internal/medial axillae and groin)
  - Upper extremities (includes external axillae)
  - Lower extremities (includes buttocks and feet)
  - At least 2 areas
- Atopic Dermatitis treatment used in the past defined in Appendix 13
  - None
  - Topical corticosteroids
  - Topical calcineurin inhibitors
  - Crisaborole
  - Topical Janus kinase inhibitor
  - Topical phosphodiesterase 4 inhibitor
  - Systemic corticosteroids
  - Immunosuppressant
  - Biologics
  - Phototherapy
  - Photochemotherapy
- Investigator's Global Assessment for AD (IGA) score: 3 versus 4
- Eczema Area and Severity Index (EASI) score
- SCORing Atopic Dermatitis (SCORAD)
- Body Surface Area (BSA)
- Itch NRS
- Itch NRS: $< 4, \geq 4$
- Sleep loss due to itch
- Sleep loss due to itch $< 2, \geq 2$
- Skin pain NRS
- Patient-Oriented Eczema Measure (POEM)
- Dermatology Life Quality Index (DLQI)
- Children Dermatology Life Quality Index (CDLQI)

- Work Productivity and Activity Impairment – Atopic Dermatitis (WPAI-AD)
  – Employment status: Employed, Not employed
  – Absenteeism
  – Presenteeism
  – Overall work impairment
  – Activity impairment
- Hospital Anxiety Depression Scale (HADS)
  – Total score
  – Anxiety domain score
  – Depression domain score
- Asthma Control Questionnaire (ACQ-5) (among participants who report comorbid asthma)
- Thymus and activation-regulated chemokine (TARC)
- Baseline Columbia-Suicide Severity Rating Scale (C-SSRS) in lifetime: Category 1 – Wish to be Dead, Category 2 – Non-specific Active Suicidal Thoughts, Category 3 – Active Suicidal Ideation with Any Methods (Not Plan) without Intent to Act, Category 4 – Active Suicidal Ideation with Some Intent to Act, without Specific Plan, Category 5 – Active Suicidal Ideation with Specific Plan and Intent, Category 6 – Preparatory Acts or Behavior, Category 7 – Aborted Attempt, Category 8 – Interrupted Attempt, Category 9 – Actual Attempt (non-fatal), and Self-injurious behavior without suicidal intent

6.4. Appendix 4: Medical History

Medical histories are defined as the conditions/events recorded on the *Pre-Existing Conditions and Medical History Details* eCRF with a start date prior to the first study drug injection.

The number and percentage of participants with medical histories will be summarized for the overall ITT Population by treatment group and by treatment and age groups for the ITT Population using the MedDRA PT nested within SOC.

The number and percentage of participants with specific medical history events of interest pre-specified on the *Prespecified Medical History* eCRF (hand dermatitis, facial dermatitis, conjunctivitis, herpes zoster, and others) will be summarized for the overall ITT Population, the MPP, and the Maintenance W16 Escape Population by treatment group and by treatment and age groups.

6.5. Appendix 5: Treatment Compliance

Treatment compliance with investigational product will be summarized for participants who have at least one dose for the Safety Population in the Induction Period and for the MPP during the Maintenance Period. Treatment compliance for each participant will be calculated as:

$$Treatment\ compliance\ (\%) = 100\ \times \frac{Total\ number\ of\ injections\ administered}{Total\ number\ of\ injections\ expected}$$

- The number of injections expected can be derived from the study drug dispense dataset.

- The total number of injections administered will be based on the response to the question "Was dose administered?" on the *Exposure as Collected* eCRF page.

The number of injections expected at each visit and total number of injections up to each visit during the Induction Period are as follows:

Visit	**W0**	**W2**	**W4**	**W6**	**W8**	**W10**	**W12**	**W14[a]**
# injections at each visit	2	2	1	1	1	1	1	1
Total # injections up to each visit	2	4	5	6	7	8	9	10

Abbreviation: W = week.

[a] Last injection during the Induction Period occurs on Week 14.

The number of injections expected at each visit and total number of injections up to each visit during the Maintenance Period are as follows:

Timepoint	**W16**	**W18**	**W20**	**W22**	**W24**	**W26**	**W28**	**W30**	**W32**	**W34**
Visit	**W16**	**W18**	**W20**		**W24**		**W28**		**W32**	
# injections at each visit	2	2	1	1	1	1	1	1	1	1
Total # injections up to each visit	2	4	5	6	7	8	9	10	11	12

Abbreviation: W = week.

Timepoint	**W36**	**W38**	**W40**	**W42**	**W44**	**W46**	**W48**	**W50**	**W52**	**W54**
Visit	**W36**		**W40**		**W44**		**W48**		**W52**	
# injections at each visit	1	1	1	1	1	1	1	1	1	1
Total # injections up to each visit	13	14	15	16	17	18	19	20	21	22

Abbreviation: W = week.

Timepoint	**W56**	**W58**	**W60**	**W62**	**W64**	**W66[a]**	**W68**
Visit	**W56**		**W60**		**W64**		**W68**
# injections at each visit	1	1	1	1	1	1	0
Total # injections up to each visit	23	24	25	26	27	28	28

Abbreviation: W = week.

[a] Last injection during the Maintenance Period occurs on Week 66.

A participant will be considered compliant if he or she received ≥75% of the expected number of injections in the respective treatment period while enrolled in the study. Descriptive statistics for

percent compliance will be summarized. Sub-intervals of interest, such as compliance between visits, may also be presented.

6.6. Appendix 6: Prior and Concomitant Therapy

Medications will be classified into anatomical therapeutic chemical (ATC) drug classes using the latest version of the World Health Organization (WHO) drug dictionary. Medication start and stop dates will be compared to the date of first dose of treatment in each treatment period to allow medications to be classified as concomitant for each treatment period.

Prior medications are those medications that start prior to the date of first dose and stop prior to or on the date of first dose of study intervention. *Concomitant medications* are those medications that start before, on, or after the first day of study intervention of the defined treatment period and continue into the treatment period. Concomitant medications are assigned to the treatment period in which they are actually ongoing. For example, if a participant is receiving a concomitant medication during the Induction Period but has a stop date during the Induction Period, the same medication would not be listed as a concomitant medication during the Maintenance Period unless participant has a new start date.

Prior medication will be summarized for the ITT Population. Concomitant medication during the Induction Period and the Maintenance Period will be presented separately for the ITT Population and the MPP.

Specific atopic dermatitis (AD) treatment during the Induction and the Maintenance Periods will be presented by treatment groups for the ITT Population and the MPP separately based on the information collected on *Concomitant Therapy* eCRF page. This will include: (1) topical AD treatment (including TCS, TCI and crisaborole), (2) systemic AD treatment (including oral corticosteroids, immunosuppressant, biologics and phototherapy). The TCS will be presented by potency. Definition of these drugs of interest is described in Appendix 13.

6.6.1. Rescue Medication for Atopic Dermatitis (AD)

Rescue medication for AD is defined as:

- any high-potency TCS as defined in Appendix 13, and the response to the question “For what indication was the medication / therapy taken?” on the Concomitant Therapy eCRF page is “Rescue Therapy”
- any systemic medication as defined in Appendix 13, and the response to the question “For what indication was the medication / therapy taken?” on the Concomitant Therapy eCRF page is “Rescue Therapy”

Participants who use these rescue medications will be summarized. The summary will be provided for any rescue medication use, with high-potency TCS and systemic therapy summarized separately for the Induction Period on the ITT Population, the Maintenance Blinded Period on the MPP, and the Maintenance Escape Period on the Maintenance W16 Escape Population, respectively.

Disease flares will be assessed based on rescue therapy usage. Flare is defined as initiation or intensification of rescue therapy. A summary of percentage of participants in the ITT Population, the MPP, and the Maintenance W16 Escape Population rescued by visit will be provided for the

Induction Period, Maintenance Blinded Period, and Maintenance Escape period, respectively. Kaplan-Meier curves for time to first rescue use may be generated.

6.7. Appendix 7: Protocol Deviations

Protocol deviations will be identified throughout the study. Important protocol deviations are defined as those deviations from the protocol likely to have a significant impact on the completeness, accuracy, and/or reliability of the study data or that may significantly affect a participant's rights, safety, or well-being. Out of all important protocol deviations (IPDs) identified, a subset occurring during the Induction Period with the potential to affect primary efficacy analysis will result in exclusion from the PPS.

Potential examples of important protocol deviations include participants who violated the inclusion/exclusion criteria, used an interfering concomitant medication, significant non-compliance with study intervention (<75% of expected injections). Refer to a separate document called "KGAL Trial Issues Management Plan" for the important protocol deviations with categorizations and whether or not these deviations will result in the exclusion of participants from the PPS.

The number and percentage of participants having IPD(s) will be summarized within category and subcategory of deviation by treatment group for the Induction Period on the ITT Population and for the Maintenance Period on All Maintenance Population.

A by-participant listing of important protocol deviations will be provided for the ITT population.

6.8. Appendix 8: Impact of COVID-19

Impact of pandemic (e.g., COVID-19) on analyses will be systemically addressed prior to study unblinding at Week 16 DBL, once the impact on study conducts is fully understood. In general, any missing assessments/visit window will be documented as protocol deviations. For participants who have missing assessments at Week 16 due to COVID-19, these patients may enter the escape arm. A summary or listing may be provided to summarize missing visits due to COVID-19.

Treatment discontinuation due to pandemic will be treated the same type of intercurrent event as treatment discontinuation due to reasons other than lack of efficacy. Strategies of how this type of intercurrent event will be handled are described in Section 1.1.1. Intermittent missing assessment due to pandemic will be treated the same as any other intermittent missing values. Details of how missing data will be handled are described in Section 4.1.6.

6.9. Appendix 9: Annual Report Analyses

Based on regulatory requirements for the Development Safety Update Report (DSUR), reports will be produced (if not already available from the study CSR) for the reporting period covered by the DSUR.

6.10. Appendix 10: Clinical Trial Registry Analyses

Additional analyses will be performed (if not already available from the study CSR) for the purpose of fulfilling the Clinical Trial Registry (CTR) requirements.

Analyses provided for the CTR requirements include the following:

Summary of AEs, provided as a dataset, will be converted to an XML file. Both serious adverse events (SAEs) and 'Other' AEs are summarized by treatment group and by MedDRA PT.

- An AE is considered 'Serious' whether or not it is a TEAE.
- An AE is considered in the 'Other' category if it is both a TEAE and is not serious. For each SAE and 'Other' AE, for each term and treatment group, the following are provided:
    - the number of participants at risk of an event,
    - the number of participants who experienced each event term, and
    - the number of events experienced.
- Consistent with www.ClinicalTrials.gov requirements, 'Other' AEs that occur in fewer than 5% of participants in every treatment group may not be included if a 5% threshold is chosen (5% is the minimum threshold).
- AE reporting is consistent with other document disclosures for example, the CSR, manuscripts, and so forth.

6.11. Appendix 11: Unblinding Plan

Unblinding details are specified in the BUP.

6.12. Appendix 12: Study Visit Mapping for Itch NRS, Skin Pain NRS and Sleep-loss and POEM

Itch NRS, Skin Pain NRS and sleep loss are collected via eCOA; entries will be mapped to study week by the following:

Week	Start Day	End Day
Baseline	Date of First Injection[a] - 7	Date of First Injection – 1
Week 1	Max (Date of First Injection, Week 2 Assessment Date – 14)	Week 2 Assessment Date – 8
Week 2	Week 2 Assessment Date – 7	Week 2 Assessment Date – 1
Week 4	Week 4 Assessment Date – 7	Week 4 Assessment Date – 1
Week 6	Week 6 Assessment Date – 7	Week 6 Assessment Date – 1
Week 8	Week 8 Assessment Date – 7	Week 8 Assessment Date – 1
Week 10	Week 10 Assessment Date – 7	Week 10 Assessment Date – 1
Week 12	Week 12 Assessment Date – 7	Week 12 Assessment Date – 1
Week 14	Week 14 Assessment Date – 7	Week 14 Assessment Date – 1
Week 16	Week 16 Assessment Date – 7	Week 16 Assessment Date – 1
Week 20	Week 20 Assessment Date – 7	Week 20 Assessment Date – 1
Week 24	Week 24 Assessment Date – 7	Week 24 Assessment Date – 1
Week 28	Week 28 Assessment Date – 7	Week 28 Assessment Date – 1

Week 32	Week 32 Assessment Date – 7	Week 32 Assessment Date – 1
Week 36	Week 36 Assessment Date – 7	Week 36 Assessment Date – 1
Week 40	Week 40 Assessment Date – 7	Week 40 Assessment Date – 1
Week 44	Week 44 Assessment Date – 7	Week 44 Assessment Date – 1
Week 48	Week 48 Assessment Date – 7	Week 48 Assessment Date – 1
Week 52	Week 52 Assessment Date – 7	Week 52 Assessment Date – 1
Week 56	Week 56 Assessment Date – 7	Week 56 Assessment Date – 1
Week 60	Week 60 Assessment Date – 7	Week 60 Assessment Date – 1
Week 64	Week 64 Assessment Date – 7	Week 64 Assessment Date – 1
Week 68	Week 68 Assessment Date – 7	Week 68 Assessment Date – 1

[a] If date of first injection is missing, the randomization date will be used.

If multiple assessments on a single day are present, use the first assessment. If an assessment could be mapped to different weeks, it will be mapped to the earlier week. Derivation of the weekly mean scores for Itch NRS, Skin Pain NRS and Sleep-loss could be found in Appendix 1. If at least 1 of the 7 days contains non-missing daily assessments, post-baseline weekly score will be calculated using prorated weekly average. If the range of 7 days are all missing daily assessments, then the weekly score is missing.

POEM is collected every week via eCOA, the visit week mapping will follow the following rule: the last collected POEM data before or on the visit date would be used, the evaluation window is injection date – 7 to injection date – 1 for baseline and assessment date – 7 to assessment date – 1 for post-baseline. For example, if a participant gets an injection/assessment on the 14th, the scale completed between the 13th and the 7th would be used.

6.13. Appendix 13: Definition of Topical and Systemic Atopic Dermatitis Therapy

The atopic dermatitis therapy in this study is defined as: high-potency TCS and systemic atopic dermatitis therapy. The topical treatments and systemic treatments are defined as following:

1. **Topical Atopic Dermatitis Treatment (including corticosteroids, TCI and crisaborole)**

Route of topical treatments includes: Topical and Transdermal.

Corticosteroids (TCS):

High-potency TCS: ATC code is D07, and the response to the item “If topical, collect Potency” on the Concomitant Therapy eCRF page is “High”.

Low or Moderate-potency TCS: ATC code is D07, and the response to the item “If topical, collect Potency” on the Concomitant Therapy eCRF page is “Low” or “Moderate”.

Topical calcineurin inhibitor (TCI): Preferred Term includes: TACROLIMUS, PIMECROLIMUS

Crisaborole: Preferred Term includes: CRISABOROLE

Topical JAK (Janus kinase) inhibitor: Preferred Term includes: DELGOCITINIB

Topical PDE4 (phosphodiesterase 4) inhibitor: Preferred Term includes: DIFAMILAST

2. Systemic Atopic Dermatitis Treatment (including oral corticosteroids, immunosuppressant, biologics and phototherapy/photochemotherapy)

Route of systemic treatments administration includes: Oral, Intra-Arterial, Intramuscular, Intraperitoneal, Intravenous, Subcutaneous, Transdermal. (This condition applies to the following categories except for phototherapies.)

Systemic Corticosteroids: ATC code is H02

Immunosuppressant: Defined as: ATC2 is L04 or Preferred terms of Abrocitinib, Baricitinib, Upadacitinib or Ruxolitinib

Biologics: Defined as following Preferred terms:

Infliximab, Infliximabum, Etanercept, Etanerceptum, Adalimumab, Adalimumabum, Certolizumab, Certolizumabum, Certolizumab pegol,Golimumab, Golimumabum, Ozoralizumab, Afelimomab, Afelimomabum, Tumor Necrosis Factor Alpha (TNF-) Inhibitors, Tabalumab, Tregalizumab, Anakinra, Basiliximab, Basiliximabum, Daclizumab, Daclizumabum, Tocilizumab, Tocilizumabum, Mepolizumab, Mepolizumabum, Rilonacept, Rilonaceptum, Ustekinumab, Canakinumab, Briakinumab, Fezakinumab, Sirukumab, Sarilumab, Lebrikizumab, Secukinumab, Olokizumab, Gevokizumab, Brodalumab, Ladarixin, Ixekizumab, Dupilumab, Tildrakizumab, Tildrakizumabum, Reslizumab, Reslizumabum, Guselkumab, Guselkumabum, Olamkicept, Fletikumab, Bimekizumab, Mirikizumab, Risankizumab, Abatacept, Ligelizumab, Vedolizumab, Belimumab, Nemolizumab, Tralokinumab, Omalizumab

Phototherapy or Photochemotherapy:

Programming search of medication name (actual term or preferred term) contains ‘photo’ or ‘UV’ then medicals to manually review to confirm whether the medication in question is indeed ‘Phototherapy’ or ‘Photochemotherapy’

6.14. Appendix 14: Details of Combining Estimates and Test Statistics for Categorial Endpoints with Multiple Imputation

Following the implementation of MCMC-MI imputation as specified in Section 4.1.6.1.1, the 25 data sets with imputations should be set together and sorted by imputation number. The following sections describe the processes for combining inferences for the individual imputed data sets into one inference for reporting. All calculations are performed in SAS software version 9.4.

Summarize Unadjusted Response Rate

The response rates, overall and by treatment arm, and their associated SEs are computed for each imputed data set using PROC FREQ with the *riskdiff* option specified for the appropriate column in the TABLES statement. The response rates and SEs from the resulting output are combined across the 25 imputed data sets using PROC MIANALYZE, separately for each arm and the overall group.

Note that the estimate and 95% CI bounds output by PROC MIANALYZE are percents (i.e., they are in terms of the response rate). To obtain the number of responders, the estimated percent is multiplied by the number of individuals in the analysis population and rounded to the nearest integer.

Compute Stratified Measures of Association

The common risk difference, common OR, and CMH test statistic are computed for each imputed data set using PROC FREQ with the *riskdiff* option for the appropriate column (for risk difference) and the *cmh* option (for OR and CMH test statistic) specified in the TABLES statement. Each of these analyses are stratified by age group, and baseline disease severity via inclusion of these variables in the TABLES statement with the treatment and outcome variables.

Note that the PROC FREQ output corresponding to the Mantel-Haenszel method is used for the risk difference, and the output corresponding to the General Association statistic is used for the CMH statistic. PROC MIANALYZE is then called separately for each of these measures, with further details in the sections below.

Common Risk Difference

No transformation is necessary before using PROC MIANALYZE to combine the risk difference estimates and their associated SEs across the 25 imputed data sets. This procedure outputs an estimate of the common risk difference and the associated 95% CI bounds.

Common Odds Ratio

The OR from each imputed data set is first transformed using the natural logarithm. The SE for each log OR (SE_{lOR}) is derived from the OR 95% CI bounds (LB_{OR}, UB_{OR}) according to the following equation: $SE_{lOR} = (\ln(UB_{OR}) - \ln(LB_{OR}))/(2 \times 1.96)$. The log OR and derived SE are then combined using PROC MIANALYZE, which outputs a combined estimate of the log OR and the associated 95% CI. Finally, these measures can be exponentiated to transform them back to the OR scale.

Cochran-Mantel-Haenszel Test

The CMH test statistic (CMH) from each imputed data set is transformed using the Wilson-Hilferty transformation and standardized so that it has approximately a standard Normal distribution (Ratitch 2013). In particular, the transformed CMH statistic is computed as follows: $CMH_{WH} = \frac{\left(\frac{CMH}{df}\right)^{\frac{1}{3}} - \left(1 - \frac{2}{9 \times df}\right)}{\sqrt{\frac{2}{9 \times df}}}$, where df is the degrees of freedom of the CMH statistic. Then the SE for each CMH_{WH} is 1, and PROC MIANALYZE is used to output a combined estimate of the transformed CMH statistic. Note that the two-sided p-value output by PROC MIANALYZE is not used directly, but instead the one-sided p-value is computed manually using both the t statistic and two-sided p-value output by PROC MIANALYZE: if t statistic is greater than 0,

then one-sided p-value is computed as half of the two-sided p-value; otherwise, the one-sided p-value is computed as 1 - half of the two-sided p-value. The resulting one-sided p-value is reported as the pooled p-value for the CMH test.

7. References

Alosh M, Bretz F, Huque M. Advanced multiplicity adjustment methods in clinical trials. *Stat Med*. 2014;33(4):693–713. https://doi.org/10.1002/sim.5974

Andersen SW, Millen BA. On the practical application of mixed effects models for repeated measures to clinical trial data. *Pharm Stat*. 2013;12(1):7–16. https://doi.org/10.1002/pst.1548

Basra MKA, Fenech R, Gatt RM, et al. The Dermatology Life Quality Index 1994-2007: a comprehensive review of validation data and clinical results .*Br J Dermatol*. 2008;159(5):997-1035. https://doi.org/10.1111/j.1365-2133.2008.08832.x

Basra MK, Salek MS, Camilleri L, Sturkey R, Finlay AY. Determining the minimal clinically important difference and responsiveness of the Dermatology Life Quality Index (DLQI): further data. *Dermatology*. 2015;230(1):27-33.

Bretz F, Maurer W, Brannath W, *et al.* A graphical approach to sequentially rejective multiple test procedures. *Stat Med.* 2009;28:586–604.

Bretz F, Posch M, Glimm E, et al. Graphical approaches for multiple comparison procedures using weighted Bonferroni, Simes, or parametric tests. *Biom J*. 2011;53(6):894-913.

[CDC] Centers for Disease Control and Prevention. Growth Chart Training resources page. Available at: https://www.cdc.gov/nccdphp/dnpao/growthcharts/resources/sas.htm. Accessed June 09, 2020.

Columbia University Medical Center. Columbia–Suicide Severity Rating Scale (C-SSRS) web site. Available at: http://www.cssrs.columbia.edu. Accessed April 21, 2011.

Eichenfield LF, Tom WL, Chamlin SL, et al. Guidelines of care for the management of atopic dermatitis: section 1. diagnosis and assessment of atopic dermatitis. *J Am Acad Dermatol*. 2014;70(2):338-351. https://doi.org/10.1016/j.jaad.2013.10.010

Finlay AY, Khan GK. Dermatology Quality of Life Index (DLQI)--a simple practical measure for routine clinical use. *Clin Exp Dermatol*. 1994;19(3):210-216. https://doi.org/10.1111/j.1365-2230.1994.tb01167.x

Firth D. Bias reduction of maximum likelihood estimates. *Biometrika*. 1993;80(1):27-38. https://doi.org/10.2307/2336755

Hongbo Y, Thomas CL, Harrison MA, et al. Translating the science of quality of life into practice: what do dermatology life quality index scores mean? *J Invest Dermatol*. 2005;125(4):569-664. https://10.1111/j.0022-202X.2005.23621.x

Isojima T, Kato N, Ito Y, Kanzaki K, Murata M. Growth standard charts for Japanese children with mean and standard deviation (SD) values based on the year 2000 national survey. Clin Pediatr Endocrinol. 2016;25(2):71-76.

Khilji FA, Gonzalez M, Finlay AY. Clinical meaning of change in Dermatology Life Quality Index scores. *Br J Dermatol*. 2002;147(suppl 62):50.

PhUSE resources page. Analyses and Displays Associated with Measures of Central Tendency – Focus on Vital Sign, Electrocardiogram, and Laboratory Analyte Measurements in Phase 2-4 Clinical Trials and Integrated Submission Documents. 2013. Available at:

http://www.phusewiki.org/docs/CSS%20White%20Papers%202016/CSS_WhitePaper_Central Tendency_v1.0.pdf. Accessed June 09, 2020.

PhUSE resources page. Analyses and Displays Associated with Outliers or Shifts from Normal to Abnormal: Focus on Vital Signs, Electrocardiogram, and Laboratory Analyte Measurements in Phase 2-4 Clinical Trials and Integrated Summary Documents. 2015. Available at http://www.phusewiki.org/docs/CSS%20White%20Papers%202016/CS_WhitePaper_Outliers Shifts_v1.0.pdf. Accessed June 09, 2020.

[PhUSE]. Pharmaceutical Users Software Exchange. PhUSE Computational Science Standard Analyses and Code Sharing Working Group. Analysis and Display White Papers Project Team. Analysis and displays associated with adverse events: focus on adverse events in Phase 2-4 clinical trials and integrated summary documents. Version 1.0. Available at: https://www.phuse.eu/documents/working-groups/cs-whitepaper-adverseevents-v10-4442.pdf. Published 03 February 2017. Accessed June 09, 2020.

Ratitch B. 2013. Combining Analysis Results from Multiply Imputed Categorical Data. PharmaSUG 2013. Paper SP03.

Reilly MC, Zbrozek AS, Dukes EM. The validity and reproducibility of a work productivity and activity impairment instrument. *Pharmacoeconomics*. 1993;4(5):353-365. https://doi.org/10.2165/00019053-199304050-00006

Sato T. On the variance estimator of the Mantel-Haenszel risk difference. *Biometrics*. 1989;45(4):1323–1324.

Simpson EL, Bieber T, Guttman-Yassky E, et al. Two phase 3 trials of dupilumab versus placebo in atopic dermatitis. *N Engl J Med*. 2016;375(24):2335-2348. https://doi.org/10.1056/NEJMoa1610020

Snaith RP. The Hospital Anxiety and Depression Scale. *Health Qual Life Outcomes*. 2003;1:29. https://doi.org/10.1186/1477-7525-1-29

Waters A, Sandhu D, Beattie P, et al. Severity stratification of Children's Dermatology Life Quality Index (CDLQI) scores. *Br J Dermatol*. 2010;163 (suppl 1):121.

White D, Leach C, Sims R, et al. Validation of the Hospital Anxiety and Depression Scale for use with adolescents. *Br J Psychiatry*. 1999;175(5):452-454. https://doi.org/10.1192/bjp.175.5.45

Winthrop KL, Novosad SA, Baddley JW, et al. Opportunistic infections and biologic therapies in immune-mediated inflammatory diseases: consensus recommendations for infection reporting during clinical trials and postmarketing surveillance. *Ann Rheum Dis*. 2015;74(12):2107-2116. https://doi.org/10.1136/annrheumdis-2015-207841

Zigmond AS, Snaith RP. The Hospital Anxiety and Depression Scale. *Acta Psychiatr Scand*. 1983;67(6):361-370. https://doi.org/10.1111/j.1600-0447.1983.tb09716.x

Signature Page for VV-CLIN-002396 v1.0

Approval	PPD 22-Aug-2022 09:12:42 GMT+0000

Signature Page for VV-CLIN-002396 v1.0