Statistical Analysis Plan Version 5 J2T-DM-KGAC

A randomized, double-blind, placebo-controlled trial to evaluate the efficacy and safety of lebrikizumab in patients with moderateto- severe atopic dermatitis

NCT04178967

Approval Date: 19-May-2022

1. Statistical Analysis Plan: J2T-DM-KGAC (DRM06-AD05): A RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED TRIAL TO EVALUATE THE EFFICACY AND SAFETY OF LEBRIKIZUMAB IN PATIENTS WITH MODERATE-TO-SEVERE ATOPIC DERMATITIS

Lebrikizumab (LY3650150)

Eli Lilly and Company
Indianapolis, Indiana USA 46285
Protocol J2T-DM-KGAC
Phase 3

Statistical Analysis Plan Version 2 electronically signed and approved by Lilly:
12 May 2021
Statistical Analysis Plan Version 3 electronically signed and approved by Lilly:
05 Aug 2021
Statistical Analysis Plan Version 4 electronically signed and approved by Lilly:
22 Oct 2021
Statistical Analysis Plan Version 5 electronically signed and approved by Lilly on date provided below.

Approved on 19 Mar 2022 GMT

2. Table of Contents

Section **Page**

Table of Contents

Table **Page**

Table of Contents

Figure **Page**

Table of Contents

Appendix **Page**

3. Revision History

Statistical Analysis Plan (SAP) Version 1 was approved prior to any unblinding.

Statistical Analysis Plan (SAP) Version 2 was approved prior to any unblinding and includes the following changes. Minor corrections/additions may not be included.

Revisions in SAP Version 2

Section	Description of Change	Rationale
Section 4, Section 6.6	• Added responder definition for sleep-loss. • Removed all maintenance endpoints and percentage of patients achieving EASI-50 at Week 2 from the list of multiplicity controlled major secondary endpoints for FDA.	Per feedback from FDA via advice letter, 1) mere change might not translate to a clinical meaningful improvement and a responder definition for sleep-loss is required; 2) formal statistical testing against placebo for maintenance of response is not meaningful or required for inclusion in labeling, 3) EASI-50 is not considered as a clinically meaningful improvement.
Section 4, Section 6.6	• Removed percentage of patients achieving at least 4-point improvement in pruritus NRS in patients who had baseline pruritus NRS ≥5 at Week 16, 4, 2 and 1 from the list of multiplicity controlled major secondary endpoints for Induction Period for EMA. • Removed Percentage of patients from those with a Pruritus NRS of ≥5-points at baseline re-randomized having achieved ≥4-point reduction from baseline at Week 16 who continue to exhibit ≥4-point reduction from baseline at Week 52 from the list of multiplicity controlled major secondary endpoints for Maintenance Period for EMA.	Removed because pruritus NRS 4-point improvement has been primarily investigated in patients who had baseline pruritus NRS≥4.

	• Changed "Percentage of patients who achieve a ≥4-point improvement from baseline to Week 16" to "Percentage of patients <u>with a DLQI total score of ≥4-points at Baseline</u> who achieve a ≥4-point improvement from baseline to Week 16".	Clarification that the evaluation of DLQI 4-point improvement will be conducted in patients who have DLQI total score of ≥4-points at Baseline only.
	• Added "Percentage of patients with a DLQI total score of ≥4-points at Baseline who achieve a ≥4-point improvement from baseline by visit" to the list of other secondary endpoints.	Added to allow for an evaluation of DLQI 4-point improvement by visit.
	• Added "Percentage of patients with a Sleep-loss score ≥2 points at Baseline who achieve a ≥2 point improvement by visit" to the list of other secondary endpoints.	Added a responder definition with meaningful improvement to allow for an evaluation of response of Sleep-loss.
	• Added "Time to loss of EASI-50 in the subset of patients who were re-randomized and achieved EASI-75 at Week 16 (EASI-50 and EASI-75 calculated relative to baseline EASI score)" and "Time to loss of IGA response, i.e., developing an IGA score ≥2 with 2 points deterioration of achieved IGA response at Week 16, in the subset of patients who were re-randomized and achieved IGA 0 or 1 and a ≥2-point improvement from Baseline at Week 16" to the list of other secondary endpoints for Maintenance Period	Added to allow analysis on time to relapse from difference aspects.
	• Removed "Percentage of patients with Pruritus NRS change of ≥4 from Baseline by visit."	Removed because pruritus NRS 4-point improvement has been primarily investigated in patients who had baseline pruritus NRS≥4.
	• Added time to first use of rescue medication for both Induction Period and Maintenance Blinded Period.	Added per clinical request.

	• Added percentage of patients rescued by visit. • Added analysis on SQAAQ for patients who complete SQAAQ at any visit.	Added per clinical request. Added to allow for analysis on SQAAQ for patients who complete SQAAQ at any visit.
Section 5.1.3	• Added definition of maintenance blinded period and maintenance escape period.	Added because efficacy analyses of maintenance primary population will be focused on maintenance blinded period.
Section 5.2	• Added statistical test that has been used to calculate sample size and power.	Clarification
Section 6.1.1	• This section has been amended to implement updated definition of analysis population for Maintenance Period. There is no change to the primary analysis population as ITT population remained as the primary analysis population for Induction Period. • Removed per protocol set (PPS) from analysis population.	To pre-specify and clarify different analysis population. PPS has been removed as it is not related to any estimand and hard to interpret under the estimand framework.
Section 6.1.2	• Added "For patients who are randomized but not dosed, the Induction Period starts on the date of randomization." • For Pruritus Numeric Rating Scale (NRS) and Sleep-Loss due to Pruritus collected via eDiary, the baseline period has been updated to the 7-day window on or prior to the first injection.	Clarification To be consistent with Appendix 1.
Section 6.1.3	• For Maintenance W24-48 Escape Population, efficacy results will be summarized every 4 weeks after lebrikizumab 250 mg Q2W treatment.	Clarification

Section 6.2	• This section has been amended to implement the definition of primary and supportive estimands for both Induction Period and Maintenance Period following ICH E9(R1) addendum. • Added the definition of supportive estimands for both categorical endpoints and continuous endpoints. • Added the missing data imputation methods relative to each estimand.	Following ICH E9(R1) addendum, details on how each type of intercurrent events will be handled for different estimands has been provided and the methods of handling missing data relative to estimands have been specified.
Section 6.4, 6.11	• This section has been amended to align with the definition of estimands. • Removed all missing values MCMC-MI from sensitivity analyses, keeping tipping point analyses as the only sensitivity analyses for the primary estimand.	To describe in details how missing data will be handled for each endpoint. Per ICH E9 (R1) addendum, sensitivity analyses have been redefined. All missing MCMC-MI do not qualify for sensitivity analyses as they handle intercurrent events differently from primary estimand.
Section 6.4.1.2	Updated tipping point analysis.	Per feedback from FDA via advice letter, all subjects who use rescue medication need to be imputed as nonresponders prior to varying the response and non-response rates for those with missing data.
Section 6.6	• Updated graphical testing scheme for multiplicity control of primary and major secondary endpoints for US. • Modified multiplicity strategy for Induction Period for EMA, replacing serial gatekeeping procedure with graphical testing scheme. • Updated testing hierarchy for Maintenance Period for EMA.	To fully specify the graphical testing scheme with arrows and weights among all endpoints to be adjusted for multiplicity for US. Updated because a couple of endpoints have been removed from the list of multiplicity

		controlled major secondary endpoints for Maintenance Period for EMA.
Section 6.8.1	• Added following to baseline disease characteristics: Sleep loss due to pruritus: <2, ≥2; EQ-5D US Population-based index score; EQ-5D UK Population-based index score. • Separated DLQI and CDLQI. • Ethnicity <u>for US</u> (Hispanic or Latino, Not Hispanic or Latino, Not reported, Unknown)	The percentage of patients with Sleep-loss 2-point reduction will be evaluated in patients with a Sleep-loss score ≥2 points at Baseline. DLQI and CDLQI are two different questionnaires anchoring different populations. Clarification that baseline ethnicity will be reported for US sites only.
Section 6.10	• *Prior medications* are those medications that start <u>prior to the date of first dose</u> and stop prior to <u>or on</u> the date of first dose of study treatment. • Removed the description of summary of Atopic Dermatitis treatment of interest. • Consolidated the summary of Atopic Dermatitis treatment of interest with the summary of rescue medications. • Added definition of flare.	Clarification Removed because this is covered by the Section of Rescue Medication. To avoid redundancy. Added to allow for the analysis on flares.
Section 6.11	• Removed analyses for itch-free days and no sleep loss days. • Added analyses for time to loss of IGA response, i.e., developing an IGA score ≥2 with 2 points deterioration of achieved IGA response at Week 16, in the subset of patients who were re-randomized and achieved IGA 0 or 1	Other exploratory endpoints related to Pruritus and Sleep loss eDiary score were added in supplementary analyses. Added to allow analysis on time to relapse from difference aspects.

	and a ≥2-point improvement from Baseline at Week 16 • Updated the definition of censoring for the analysis of time to loss of IGA response. • Updated the derivation of BSA Total. • Updated the derivation of post-baseline weekly mean for Pruritus and Sleep loss to prorated weekly mean. • Separated analyses for DLQI and CDLQI total scores. • Table KGAC.6.12 has been updated to be in alignment with the definition of estimands and the specification of methods of missing data imputation.	Clarification Clarification To mitigate potential bias introduced by inadequate eDiary entries and improve efficiency for multiple imputation. DLQI and CDLQI are two different questionnaires anchoring different populations. To be consistent with the definition of estimands and the specification of methods of missing data imputation.
Section 6.11.2	• This section has been updated to reflect the change in the sensitivity analyses for primary outcomes.	To ensure consistency.
Section 6.14	• This section has been updated to be in alignment with compound level safety standard. o Added "Drug interruption time period due to the use of systemic rescue therapies will be removed from study drug exposure calculations as described in compound level safety standards." o Added Section of Atopic Dermatitis Exacerbation and	To ensure consistency between SAP and compound level safety standard.

	Section of Suicidal Ideation and Behavior. • Removed listing of exposure.	Removed because listing of exposure is not required for CSR.
Section 6.15.1	• Added subgroup analyses for EASI-90 and 4-point improvement in Pruritus NRS at Week 16. • Removed subgroup analysis of efficacy by TE-ADA status. • Removed subgroup analysis of efficacy by ethnicity. • Updated the statistical test that will be used to evaluate treatment group differences within each subgroup from fisher's exact test to chi-square test.	To be consistent with protocol. Removed because the impact of TE-ADA status will be better evaluated in integrated database due to small sample size. Removed because ethnicity will be reported for US sites only. To allow for the use of PROC MIANALYZE to combine results from multiply imputed dataset.
Section 6.16.1	• Added "A summary or listing may be provided to summarize missing visits due to COVID-19".	To allow for the investigation of missing data due to COVID-19.
Appendix 1	• Replaced "assessment date" with "visit date". • If multiple assessments on a single day are present, use the <u>first</u> assessment. • Clarified the derivation of weekly mean for Pruritus NRS and sleep loss score. • Added visit mapping for PEOM data analysis.	Clarification
Appendix 2	• Added details of combining estimates and test statistics for categorial endpoints with multiple imputation.	To provide detailed instructions on how to combine estimates and test statistics for categorical endpoints from multiply imputed datasets.
Appendix 3	• Added definition of rescue medications.	To provide detailed instructions on how to determine rescue medications for this study.

Statistical Analysis Plan (SAP) Version 3 was approved prior to any unblinding and before Week 16 interim database lock and includes the following changes. Minor corrections/additions may not be included.

Revisions in SAP Version 3

Section	Description of Change	Rationale
Section 4, Section 6.11	• Added back several endpoints as other secondary endpoints. • Added following other secondary endpoints o Percentage of patients with a Pruritus NRS score of ≥4 points at Baseline who achieve both an IGA score of 0 or 1 and a reduction of ≥2 points in IGA score from Baseline, and a ≥4-point reduction in Pruritus NRS score from Baseline by visit o Percentage of patients with a Pruritus NRS score of ≥4 points at Baseline who achieve both EASI-75 and a ≥4-point reduction in Pruritus NRS score from Baseline by visit	To be consistent with protocol and CT.gov To allow the analysis on composite endpoints
Section 4, Section 6.6	• Removed "Percentage of patients with a Pruritus NRS of ≥4-points at Baseline who achieve a ≥4-point reduction from Baseline to Week 1" from the list of multiplicity controlled major secondary endpoints for the FDA and EMA. • Removed "Percentage of patients with an IGA score of 0 or 1 and a reduction ≥2 points at Week 2." from the list of multiplicity controlled major secondary endpoints for FDA.	Strategy change in multiplicity control

Section 6.6	• Updated the graphical testing scheme for multiplicity control of primary and major secondary endpoints for the FDA. • Added the graphical testing scheme for multiplicity control of primary and major secondary endpoints for the Induction Period for the EMA.	To reflect the change in the strategy of multiplicity control To prespecify the graphical testing scheme for EMA
Section 6.8.1	• Updated the subcategories for Atopic Dermatitis treatment used in the past. • Added prior use of systemic treatment (yes, no).	Clarification
Section 6.14.6.5	• Removed listing of patients with hypersensitivity.	Listing of patients with hypersensitivity will be provided in the context of evaluating immunogenicity.
Section 6.14.6.9	• Updated the section heading for Suicide/Self-injury.	To reflect the search strategy using SMQ code
Section 6.15.1	• Added a subgroup "Prior use of systemic treatment (yes, no)" for efficacy subgroup analysis.	To prespecify the analysis for this subgroup
Section 6.16.1	• Added a description of how missing data due to pandemic will be handled.	Clarification
Appendix 1	• Added "If an assessment could be mapped to different weeks, it will be mapped to the earlier week."	Clarification
Appendix 2	• Revised the formula for the transformed CMH statistic.	Correction
Appendix 3	• Added "Route of topical treatments includes: Topical and Transdermal."	Clarification

Abbreviations: CMH = Cochran-Mantel-Haenszel; EASI = Eczema Area and Severity Index; EMA = European Medicines Agency; FDA = Food and Drug Administration; IGA = Investigator Global Assessment; NRS = Numerical Rating Scale; SMQ = Standardized MedDRA Query.

After primary database lock, a site audit with a critical finding necessitated revision of the Statistical Analysis Plan (SAP). SAP Version 4 was prepared and approved by statisticians who at the time of SAP amendment have been independent from the study team and blinded to

patient-level data of J2T-DM-KGAD and J2T-DM-KGAC. The limited number of study team statisticians who were unblinded to patient-level data did not participate in revision of the SAP. The re-randomization into the Maintenance Period remains blinded to all study team members at the time of SAP amendment.

Changes in Version 4 are documented in the following table. Minor corrections/additions may not be included.

Revisions in SAP Version 4

Section	Description of Change	Rationale
Section 6.1.1	• Added 7 new analysis populations (Modified ITT, Modified Safety, Modified Maintenance Primary, Modified Maintenance Secondary, Modified Maintenance W16 Escape, Modified Maintenance W24-48 Escape, All Lebrikizumab Modified Safety). • Removed "Unless otherwise specified, efficacy and health outcomes analyses for the Induction period will be conducted on this population." from the ITT Population and added to the mITT Population. • Removed "Safety analyses for the Induction period will be conducted on this population." from the Safety Population and added to the Modified Safety Population. • Updated that the efficacy, health outcomes, and/or safety analyses during the maintenance period will be conducted on the Modified Maintenance Primary Population and/or Modified Maintenance Secondary Population and removed relevant languages from the Maintenance Primary and Secondary Populations.	A directed site audit was triggered by statistically implausible data in study J2T-DM-KGAD at one study site, and the same site was also included in this study (J2T-DM-KGAC) with similarly implausible data. It was determined by the audit that some or all of the study participants at the site did not meet the eligibility criterion of having moderate-to-severe atopic dermatitis, and associated data was unreliable.

	• Updated that safety analyses for the Combined Induction and Maintenance Periods and the Combined Induction and Maintenance Periods plus the follow-up Period will be conducted on All Lebrikizumab Modified Safety Population, and removed relevant languages from All Lebrikizumab Safety Population. • In Table KGAC.6.2., - Removed "ITT" and added "mITT" and "Modified Safety" for Induction Period. - Added "Modified Maintenance Primary" and "Modified Maintenance Secondary" for Maintenance Blinded Period, and specified only safety analysis will be conducted on the Maintenance Primary Population. - Updated Maintenance Escape Populations to Modified Maintenance Escape Populations for the Maintenance Escape Period. - Added "All Lebrikizumab Modified Safety" for the Combined Induction and Maintenance Periods and the Combined Induction and Maintenance Periods + FU.	
Section 6.1.3	• Specified that for patients in the mMPP and mMSP, who met escape criteria and escaped to lebrikizumab 250 mg Q2W at Weeks 24, 32, 40 and 48, only data in the Maintenance Blinded Period (up to the time of escape)	See above.

	will be included in both efficacy and safety analyses. • Added mMPP to Section 6.1.3.1. • Updated from MSP to mMSP in Section 6.1.3.2. • Updated from Maintenance Escape Populations to Modified Maintenance Escape Populations in Section 6.1.3.3.	
Sections 6.1.4 and 6.1.5	• Added All Lebrikizumab Modified Safety Population	See above.
Section 6.2.2	• Updated population for maintenance period estimands from MPP to mMPP	See above.
Section 6.4.2	• Updated that missing data imputation for Maintenance Period will be conducted on mMPP only.	See above.
Section 6.7	• Added patient disposition summaries for the mITT Population. • Added patient disposition summaries for Maintenance Period for mMPP and mMSP.	See above.
Section 6.8	• Updated analysis population from ITT to mITT for a summary of - demographic and baseline characteristics, and - medical histories.	See above.
Section 6.9	• Added treatment compliance for the Modified Safety Population in Induction Period and mMPP for the Maintenance Blinded Period.	See above.
Section 6.10	• Updated analysis population from ITT to mITT for a summary of	See above.

	- prior medications, - concomitant medications, - rescue medications.	
Section 6.11	• Updated analysis population for all efficacy and health outcome analyses - from ITT to mITT, - from MPP to mMPP, - from MSP to mMSP, - from Maintenance W16 Escape Population to Modified Maintenance W16 Escape Population, and - from Maintenance W24-48 Escape Population to Modified Maintenance W24-48 Escape Population	See above.
Section 6.11.1	• Updated analysis population from ITT to mITT for primary analysis of the primary outcome (IGA of 0 or 1 at Week 16) and the additional EMA primary outcome (EASI-75 at Week 16).	See above.
Section 6.14	• Updated the Modified Safety Population as primary analysis population for safety evaluations (exposure, adverse events, clinical laboratory data, vital signs, immunogenicity, adverse events of special interest) in Induction Period. • Update safety evaluations based on the Safety Population in Induction Period as sensitivity analysis. • Added selective safety evaluations summaries based on Modified Maintenance Primary Population for Maintenance Blinded Period, and based on All Lebrikizumab Modified Safety	See above.

	Population for Combined Induction and Maintenance Periods and Combined Induction and Maintenance Periods + FU.	
Section 6.15	• Updated analysis population from ITT to mITT for efficacy subgroup analyses.	See above.
Section 6.16	• Removed languages related to the per-protocol set. • Clarified a listing of IPDs will be provided for the ITT Population.	• Per-protocol set analyses not planned. • Clarification.

Abbreviations: FU = follow-up; IPD = important protocol deviation; mITT = modified intent-to-treat; mMPP = modified maintenance primary population; mMSP = modified maintenance secondary population.

Statistical Analysis Plan (SAP) Version 5 was approved prior to Week 52 interim database lock and prior to the unblinding of the re-randomized maintenance treatment but after the Lilly study team was unblinded to the induction treatment. This version of the SAP includes the following changes. Minor corrections/additions may not be included.

Revisions in SAP Version 5

Section	Description of Change	Rationale
Section 4, Section 6.11	• Added "Time to loss of EASI-75 in the subset of patients who were re-randomized and achieved EASI-75 at Week 16 (EASI-75 calculated relative to baseline EASI score)" • Updated "Time to loss of EASI-50 in the subset of patients who were re-randomized and achieved EASI-75 at Week 16 (EASI-50 and EASI-75 calculated relative to baseline EASI score)" to "Time to loss of EASI-50 in the subset of patients who were re-randomized at Week 16 (EASI-50 calculated relative to baseline EASI score)" • Added "Percentage change in EASI score from Baseline at Week 52 in the subset of patients who were re-randomized at Week 16 " as a major secondary endpoint for EMA • Moved "Percentage change in SCORAD (having achieved EASI-75 at Week 16) from baseline at Week 52" from major secondary endpoints to other secondary endpoints	Added to allow analysis on time to loss of EASI-75 Modified to allow analysis on time to loss of EASI-50 on patients who were re-randomized at Week 16 as an overall evaluation of how soon those patients will move to escape arm To reflect the change in the strategy of multiplicity control for EMA
Section 6.3	• Added specification of covariates to be adjusted for maintenance period analysis	Clarification
Section 6.8	• Patient demographic variables and baseline characteristics will be summarized by treatment group for the mITT Population, the Modified Maintenance Primary Population and the Modified Maintenance W16 Escape Population.	Added to allow the comparison between responders and non-responders at Week 16 in terms of patient characteristics

	• The number and percentage of patients with specific medical history events of interest pre-specified on the History Assessment eCRF (hand dermatitis, facial dermatitis, conjunctivitis, herpes Zoster, and others) will be summarized for the mITT Population, the Modified Maintenance Primary Population and the Modified Maintenance W16 Escape Population by treatment group and by treatment and age groups.	
Section 6.9	• Removed analysis of treatment compliance on safety population and maintenance primary population. Treatment compliance will be summarized in modified safety population and modified maintenance populations.	Due to limited number of patients from site 5042, treatment compliance summaries in nonmodified populations were expected to be similar to modified populations.
Section 6.10.1	• Added analysis of Rescue Medication use on Modified Maintenance W16 Escape Population	Added to allow the analysis of Rescue Medication use on Escape arm
Section 6.14	• Removed most of the safety sensitivity analyses on safety population, maintenance primary population, and all lebrikizumab safety population. Instead, a limited number of safety summaries were selected on these populations, including overview of AEs, Summary of TEAE PTs by maximum severity and a listing of TEAEs occurred in safety population but not in modified safety population	Due to limited number of patients from site 5042, safety summaries in nonmodified populations were expected to be similar to modified populations. The adverse events from site 5042 will be included in the AE listing.
Section 6.14.1	• 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.	To be consistent with compound level safety standards.
Section 6.14.5	• Removed immunogenicity analyses on all lebrikizumab modified safety population and all lebrikizumab safety population. Removed the summary of specified	Immunogenicity analyses on all lebrikizumab modified safety population and all lebrikizumab safety

	TEAEs by TE-ADA status. Added immunogenicity analyses on modified maintenance primary population.	population will be evaluated in integrated database, as well as the summary of TEAEs by TE-ADA status.

4. Study Objectives

Table KGAC.4.1 shows the objectives and endpoints of the study. In addition, the analysis of some exploratory endpoints is described in Section 6.11 to provide supportive evidence of efficacy.

Table KGAC.4.1. Objectives and Endpoints

Study Objective: To evaluate the safety and efficacy of lebrikizumab compared with placebo in patients with moderate-to-severe AD	
FDA Endpoints	**EMA Endpoints**
Primary percentage of patients with an IGA score of 0 or 1 and a reduction ≥2 points from Baseline to Week 16.	**Co-primary** Percentage of patients with an IGA score of 0 or 1 and a reduction ≥2 points from baseline to Week 16. Percentage of patients achieving EASI-75 (≥75% reduction from Baseline in EASI score) at Week 16
Major Secondary • Percentage of patients achieving EASI-75 (≥75% reduction from Baseline in EASI score) at Week 16 • Percentage of patients achieving EASI-90 (≥90% reduction from Baseline in EASI score) at Week 16 • Percentage of patients with a Pruritus Numerical Rating Scale (NRS) of ≥4-points at Baseline who achieve a ≥4-point reduction from Baseline to Week 16 • Percentage of patients with a Pruritus NRS of ≥4-points at Baseline who achieve a ≥4-point reduction from Baseline to Week 4 • Percentage of patients with a Pruritus NRS of ≥4-points at Baseline who achieve a ≥4-point reduction from Baseline to Week 2 • Percentage of patients with an IGA score of 0 or 1 and a reduction ≥2 points at Week 4. • Percentage of patients with an IGA score of 0 or 1 and a reduction ≥2 points at Week 16 in adults. • Percentage of patients with a Sleep-loss score ≥2 points at Baseline who achieve a ≥2 points reduction from Baseline at Week 16	**Major Secondary Endpoints Specific for Induction Period** • Percentage of patients achieving EASI-90 at Week 16 • Percentage of patients achieving EASI-90 at Week 4 • Percentage change in EASI score from Baseline to Week 16 • Percentage change in Pruritus NRS score from Baseline to Week 16 • Percentage of patients with a Pruritus NRS of ≥4-points at Baseline who achieve a ≥4-point reduction from Baseline to Week 16 • Percentage of patients with a Pruritus NRS of ≥4-points at Baseline who achieve a ≥4-point reduction from Baseline to Week 4 • Percentage of patients with a Pruritus NRS of ≥4-points at Baseline who achieve a ≥4-point reduction from Baseline to Week 2 • Change from baseline in DLQI total score at Week 16 • Percentage of patients with a DLQI total score of ≥4-points at Baseline who achieve a ≥4-point improvement from baseline to Week 16 • Change from Baseline in Sleep-loss score at Week 16 • Percentage of patients with a Sleep-loss score ≥2 points at Baseline who achieve a ≥2 points reduction from Baseline at Week 16

Abbreviations: AD = atopic dermatitis; DLQI = Dermatology Life Quality Index; EASI = Eczema Area and Severity Index; Investigator Global Assessment (IGA)

Objectives and Endpoints

Study Objective: To evaluate the safety and efficacy of lebrikizumab compared with placebo in patients with moderate-to-severe AD	
FDA Endpoints	**EMA Endpoints**
Other Secondary Endpoints Specific for Maintenance Period: • Percentage of patients from those re-randomized having achieved EASI-75 at Week 16 who continue to exhibit EASI-75 at Week 52 (EASI-75 calculated relative to baseline EASI score) • Percentage of patients from those re-randomized having achieved IGA 0 or 1 and a ≥2-point improvement from Baseline at Week 16 who continue to exhibit an IGA 0 or 1 and a ≥2-point improvement from Baseline at Week 52 • Time to loss of EASI-50 in the subset of patients who were re-randomized at Week 16 (EASI-50 calculated relative to baseline EASI score) • Time to loss of EASI-75 in the subset of patients who were re-randomized and achieved EASI-75 at Week 16 (EASI-75 calculated relative to baseline EASI score) • Time to loss of IGA response, ie, developing an IGA score ≥2 with 2 points deterioration of achieved IGA response at Week 16, in the subset of patients who were re-randomized and achieved IGA 0 or 1 and a ≥2-point improvement from Baseline at Week 16	**Major Secondary Endpoints Specific for Maintenance Period:** • Percentage of patients from those re-randomized having achieved EASI-75 at Week 16 who continue to exhibit EASI-75 at Week 52 (EASI-75 calculated relative to baseline EASI score) • Percentage of patients from those re-randomized having achieved IGA 0 or 1 and a ≥2-point improvement from Baseline at Week 16 who continue to exhibit an IGA 0 or 1 and a ≥2-point improvement from Baseline at Week 52 • Percentage of patients from those with a Pruritus NRS of ≥4-points at baseline re-randomized having achieved ≥4-point reduction from baseline at Week 16 who continue to exhibit ≥4-point reduction from baseline at Week 52 • Percentage change in EASI score from Baseline at Week 52 in the subset of patients who were re-randomized at Week 16 **Other Secondary Endpoints Specific for Maintenance Period:** • Percentage change in SCORAD (having achieved EASI-75 at Week 16) from baseline at Week 52 • Percentage of patients from those with a Pruritus NRS of ≥5-points at baseline re-randomized having achieved ≥4-point reduction from baseline at Week 16 who continue to exhibit ≥4-point reduction from baseline at Week 52 • Time to loss of EASI-50 in the subset of patients who were re-randomized at Week 16 (EASI-50 calculated relative to baseline EASI score) • Time to loss of EASI-75 in the subset of patients who were re-randomized and achieved EASI-75 at Week 16 (EASI-75 calculated relative to baseline EASI score) • Time to loss of IGA response, ie, developing an IGA score ≥2 with 2 points deterioration of achieved IGA response at Week 16, in the subset of patients who were re-randomized and achieved IGA 0 or 1 and a ≥2-point improvement from Baseline at Week 16
Evaluate the pharmacokinetics of lebrikizumab. • Average serum lebrikizumab concentration	Evaluate the pharmacokinetics of lebrikizumab. • Average serum lebrikizumab concentration

Objectives and Endpoints

Other Secondary Endpoints
• Percentage of patients with EASI-75, EASI-90 and EASI-50 by visit • Percentage of patients with IGA Score of 0 or 1 and a reduction ≥2 points from Baseline by visit • Percentage change from Baseline in EASI Score by visit • Percentage change from Baseline in Pruritus NRS by visit • Percentage of patients with a Pruritus NRS score of ≥4 points at Baseline who achieve a ≥4-point reduction from Baseline by visit • Percentage of patients with a Pruritus NRS score of ≥5 points at Baseline who achieve a ≥4-point reduction from Baseline by visit • Percentage of patients with Pruritus NRS change of ≥4 from Baseline by visit • Change from Baseline in Sleep-Loss score by visit • Percent change from Baseline in Sleep-Loss score by visit • Percentage of patients with a Sleep-loss score ≥2 points at Baseline who achieve a ≥2 points by visit • Change from Baseline in DLQI by visit • Change from baseline in CDLQI by visit • Percentage of patients with a DLQI total score of ≥4-points at Baseline who achieve a ≥4-point improvement from baseline by visit • Percentage of patients who achieve ≥4-point improvement in DLQI from baseline to Week 16 • Change from Baseline in EQ-5D by visit • Change from Baseline in POEM by visit • Change from Baseline in PROMIS Anxiety measure by visit • Change from Baseline in PROMIS Depression measure by visit • Change in ACQ-5 score from Baseline to Week 16 in patients who have self-reported comorbid asthma • Percentage change from Baseline to Week 16 in SCORAD • Change from baseline in BSA by visit • Time to first use of rescue medication during Induction Period/Maintenance Blinded Period • Percentage of patients rescued by visit • Percentage of patients who respond “Strongly Agree” or “Agree” for each item of the modified SQAAQ by data collection sequence • Percentage of patients with a Pruritus NRS score of ≥4 points at Baseline who achieve both an IGA score of 0 or 1 and a reduction ≥2 points from Baseline, and a ≥4-point reduction in Pruritus NRS score from Baseline by visit • Percentage of patients with a Pruritus NRS score of ≥4 points at Baseline who achieve both EASI-75 and a ≥4-point reduction in Pruritus NRS score from Baseline by visit

Abbreviations: ACQ-5 = Asthma Control Questionnaire 5-item version; AD = atopic dermatitis; BSA = body surface area; CDLQI = Children’s Dermatology Life Quality Index; DLQI = Dermatology Life Quality Index; EASI = Eczema Area and Severity Index; EQ-5D = standardized instrument developed by the EuroQol Group; EMA = European Medicines Agency; FDA = Food and Drug Administration; IGA = Investigator Global Assessment; POEM =Patient Oriented Eczema Measure; PROMIS = Patient-Reported Outcomes Measurement Information System; SCORAD = SCORing Atopic Dermatitis; SQAAQ= subcutaneous administration assessment questionnaire.

For Food and Drug Administration (FDA), primary and major secondary endpoints for Induction Period will be adjusted for multiplicity. For European Medicines Agency (EMA), primary and major secondary endpoints for Induction Period and major secondary endpoints for Maintenance Period will be adjusted for multiplicity separately (ie, induction and maintenance endpoints will be tested separately). Details can be found in Section 6.6.

5. Study Design

5.1. Summary of Study Design

Study J2T-DM-KGAC (KGAC) [aka DRM06-AD05] is a randomized, double-blind, placebo-controlled, parallel-group study in adult and adolescent (≥12 to <18 years weighing ≥40 kg) patients with moderate-to-severe atopic dermatitis (AD). Approximately 400 patients will be enrolled into the study. The study is comprised of 2 treatment periods (16-week Induction and 36-week Maintenance). Patients completing this 52-week study will be offered continued treatment in a separate long-term extension study J2T-DM-KGAA (DRM06-AD07). Patients who early terminate or choose not to enter the long-term extension study will undergo a follow-up visit approximately 12 weeks after the last study drug injection for safety follow-up.

5.1.1. Screening Period

Screening Period: Patients will be evaluated for study eligibility before the baseline visit (Day 1). Electronic diary collection will begin at screening.

5.1.2. Baseline and Double-Blinded Induction Period (Week 0 to Week 16)

At baseline visit (Day 1), patients who meet the study eligibility criteria will be 2:1 randomly assigned to their induction treatments with stratification based on geographic region (United States [US] versus European Union [EU] versus rest of world), age (adolescent patients 12 to <18 versus adults ≥18 years) and disease severity (IGA 3 versus 4). The treatment groups in the Blinded Induction Period are:

- Lebrikizumab 250 mg every 2 weeks (Q2W): 500 mg lebrikizumab administered at Baseline and Week 2 (loading dose; 2 pre-filled syringes with a pre-assembled needle safety device [PFS-NSD]) and 250 mg Q2W through Week 14.
- Placebo: 4 mL (2 PFS-NSD) administered at Baseline and Week 2 and 2 mL Q2W through Week 14.

5.1.3. Maintenance Period (Week 16 to Week 52 [36 Weeks])

5.1.3.1. Maintenance Blinded Period

After completion of the Week 16 visit, patients who have responded to treatment (defined as having an IGA of 0 or 1 or a 75% reduction in EASI from Baseline to Week 16 [EASI-75] according to IWRS system) will enter the Maintenance Period and will be re-randomized 2:2:1 to one of the following treatment groups: lebrikizumab 250 mg Q2W, lebrikizumab 250 mg Q4W, or placebo Q2W. Throughout the maintenance blinded period, patients will receive placebo, as appropriate, to maintain the study blind across treatment groups.

5.1.3.2. Maintenance Escape Period

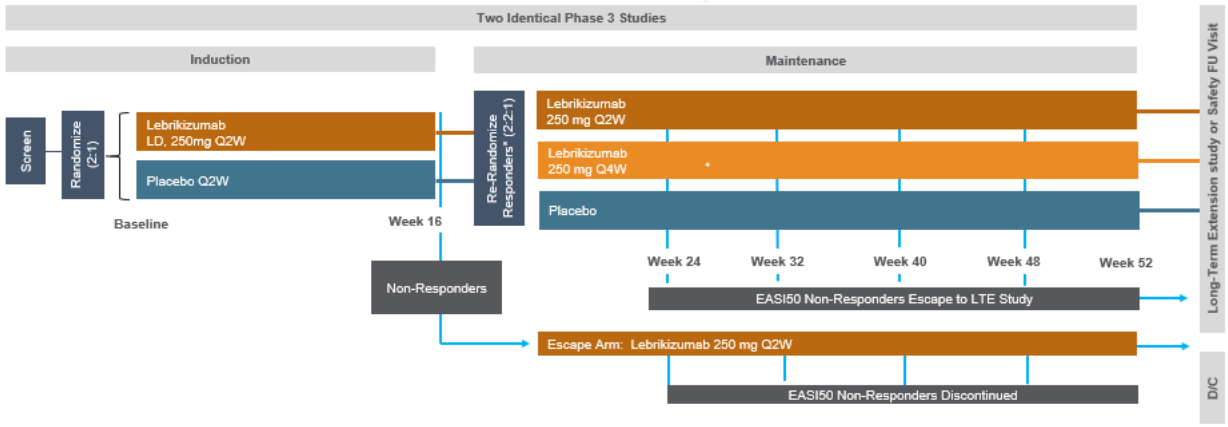
Patients who do not achieve an IGA of 0 or 1 or an EASI-75 at Week 16, patients received topical or systemic rescue therapy between baseline to Week 16 and those patients not

maintaining an EASI-50 response following re-randomization at Week 24, 32, 40, or 48 will be assigned to an Escape Arm and receive lebrikizumab 250 mg as open-label treatment Q2W through Week 52. Patients not achieving an EASI-50 response in the Escape Arm after 8 weeks of treatment will be terminated from the study.

5.1.4. Safety Follow-up Visit

Patients who terminate early from the study or do not enroll in the long-term extension study, J2T-DM-KGAA (DRM06-AD07), will undergo a follow up visit approximately 12 weeks after the last study drug injection.

Figure KGAC.5.1 illustrates the study design.



* Responder is defined as having an IGA of 0 or 1 or a 75% reduction in EASI from Baseline to Week 16 (EASI-75)

Figure KGAC.5.1. Illustration of study design for Clinical Protocol KGAC.

5.2. Determination of Sample Size

For FDA: In the DRM06-AD01 Phase 2b study (J2T-DM-KGAF), the proportion of patients who achieved an IGA score of 0 or 1 at Week 16 using the rescue medication non-response sensitivity analysis was approximately 34.7% for lebrikizumab 250 mg Q2W versus 7.7% for placebo. A sample size of 96 for lebrikizumab 250 mg Q2W versus 48 for placebo will have more than 95% power to detect a statistically significant difference based on a two group continuity corrected chi-square test with a two-sided significance level of 0.05. However, to ensure sufficient safety information is collected and to ensure sufficient responders for the Maintenance Period, the sample size will be increased to approximately 400 in total with a randomization ratio of 2:1 lebrikizumab:placebo.

For European Medicines Agency (EMA): In the DRM06-AD01 Phase 2b study (J2T-DM-KGAF), the proportion of patients who achieved an IGA score of 0 or 1 at Week 16 using the rescue medication non-response sensitivity analysis was approximately 34.7% for lebrikizumab 250 mg Q2W versus 7.7% for placebo, and the proportion of patients who achieved an EASI-75 at Week 16 using the rescue medication non-response sensitivity analysis was approximately

48.0% for lebrikizumab 250 mg Q2W versus 11.5% for placebo. A sample size of 96 for lebrikizumab 250 mg Q2W versus 48 for placebo will have more than 95% power to detect a statistically significant difference based on a two-group continuity corrected chi-square test with a two-sided significance level of 0.05 for each of the co-primary endpoints, which imply and overall power of at least 90%. However, to ensure sufficient safety information is collected and to ensure sufficient responders for the Maintenance Period, the sample size will be increased to 400 in total with a randomization ratio of 2:1 lebrikizumab:placebo.

5.3. Method of Assignment to Treatment

All patients will be randomly allocated to receive the study treatment using an electronic data capture (EDC) system at the Baseline visit. The allocation to treatment will be prospectively stratified by geographic region (US versus EU versus rest of world), age (adolescent patients 12 to <18 years versus adults ≥18 years) and disease severity (IGA 3 versus 4). At the Baseline visit (Day 1), once a patient is considered eligible to participate in the study, demographic and stratification information will be entered into the EDC system to receive a medication number assigning a kit to a patient.

During the Maintenance Period, the EDC will be used to re-randomize a patient to a maintenance treatment based on the IGA or EASI score at Week 16 and rescue therapy usage during induction period.

6. A Priori Statistical Methods

6.1. General Considerations

Statistical analysis of this study will be the responsibility of Eli Lilly and Company (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 (eg, described in KGAC Protocol Table 1) 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.

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.

All statistical processing will be performed using SAS® unless otherwise stated. Some of the analyses described in this document will be incorporated into interactive display tools instead of or in addition to static displays. 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.

6.1.1. Analysis Populations

Analysis populations are defined in Table KGAC.6.1 along with the analysis they will be used to conduct. Table KGAC.6.2 describes the treatment groups and the comparisons for each study period and the analysis population.

Figure KGAC.6.1 shows a pictorial description of analysis populations.

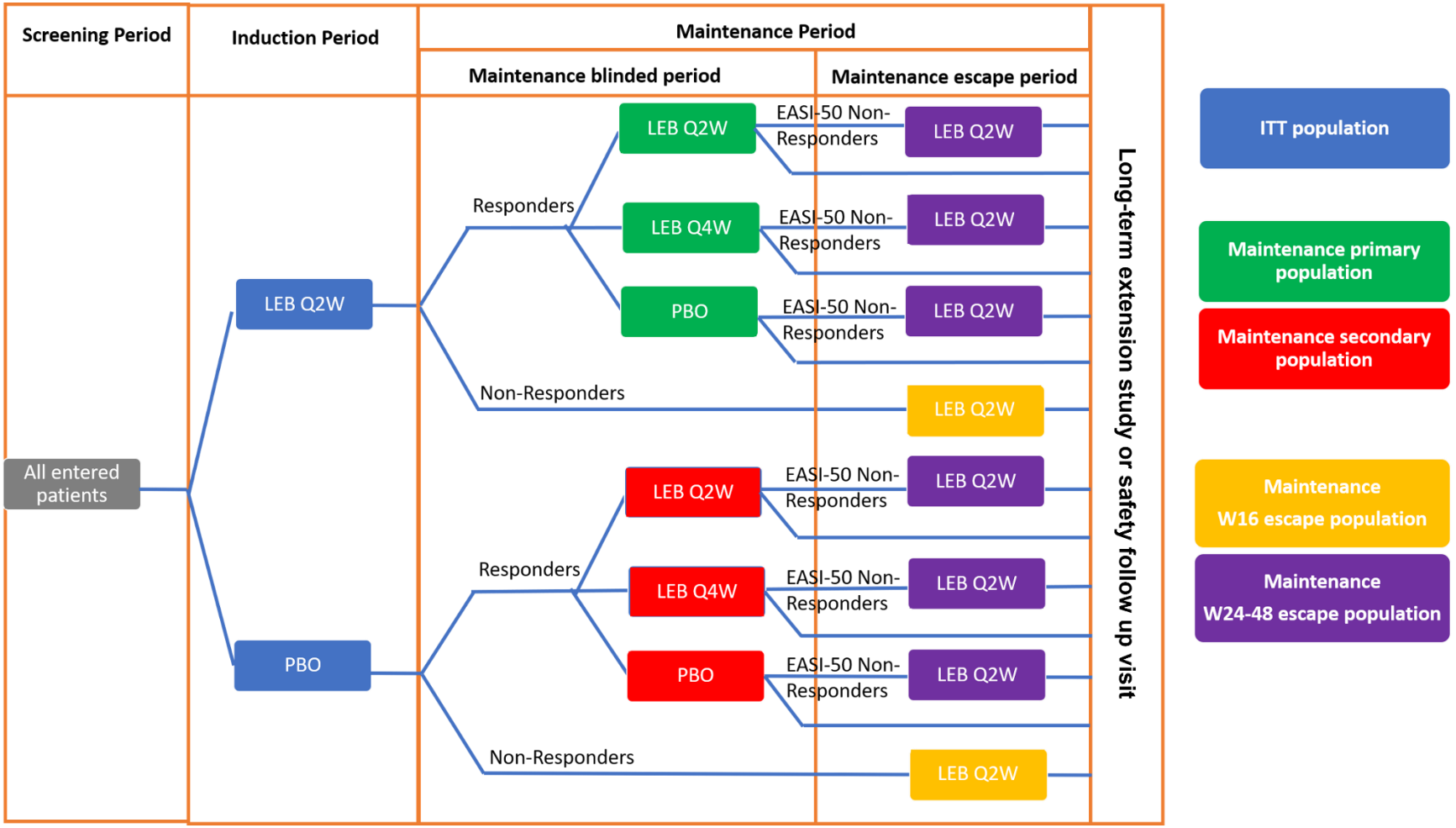


Figure KGAC.6.1. Study periods and analysis populations.

Table KGAC.6.1. Analysis Populations

Population	Description
All Entered Patients	All patients who signed informed consent. Patient flow will be summarized.
Intent-to-Treat (ITT) Population	All randomized patients, even if the patient does not take the assigned treatment, does not receive the correct treatment, or otherwise does not follow the protocol. Patients will be analyzed according to the treatment to which they were assigned.
Modified ITT (mITT) Population	ITT Population *excluding* all patients from Site 5042. Patients will be analyzed according to the treatment to which they were assigned. Unless otherwise specified, efficacy and health outcomes analyses for the Induction period will be conducted on this population.
Safety Population	All randomized patients who received at least 1 dose of study treatment during Induction Period.
Modified Safety Population	Safety Population *excluding* all patients from Site 5042. Safety analyses for Induction period will be conducted on this population.
Maintenance Primary Population (MPP)	All patients who were randomized to Lebrikizumab 250 mg Q2W at Baseline Visit and re-randomized to Lebrikizumab 250 mg Q2W, Lebrikizumab 250 mg Q4W or placebo at Week 16 and received at least 1 dose of study treatment during the maintenance period. Patients will be analyzed according to the treatment to which they were re-randomized. Only information prior to escape will be presented.
Modified Maintenance Primary Population (mMPP)	MPP Population *excluding* all patients from Site 5042. Patients will be analyzed according to the treatment to which they were re-randomized. Only information prior to escape will be presented. Efficacy, health outcomes, and safety analyses for the maintenance period will be conducted on the Modified Maintenance Primary Population
Maintenance Secondary Population (MSP)	All patients who were randomized to placebo at Baseline Visit and re-randomized to Lebrikizumab 250 mg Q2W, Lebrikizumab 250 mg Q4W or placebo at Week 16, and received at least one dose of study treatment during the maintenance period.
Modified Maintenance Secondary Population (mMSP)	MSP Population *excluding* all patients from Site 5042. Patients will be analyzed according to the treatment to which they were re-randomized. Only information prior to escape will be presented. Selective efficacy analyses for the maintenance period will be conducted on the Modified Maintenance Secondary Population.
Maintenance W16 Escape Population	All patients who were NOT re-randomized to Lebrikizumab 250 mg Q2W, Lebrikizumab 250 mg Q4W or placebo but assigned to escape arm at Week 16, and received at least one dose of study treatment during the maintenance period.
Modified Maintenance W16 Escape Population	Maintenance W16 Escape Population *excluding* all patients from Site 5042. Selective efficacy analyses for the maintenance period will be conducted on the Maintenance W16 Escape Population.
Maintenance W24-48 Escape Population	All patients from Maintenance Primary and Secondary Population who escaped to Lebrikizumab 250 mg Q2W due to EASI-50 non-response at Week 24, 32, 40 or 48.
Modified Maintenance W24-48 Escape Population	Maintenance W24-48 Escape Population *excluding* all patients from Site 5042. Selective efficacy analyses for the maintenance period will be conducted on the Maintenance W24-48 Escape Population
All Lebrikizumab Safety Population	All randomized patients who received at least 1 dose of lebrikizumab treatment during Combined Induction and Maintenance Periods.

All Lebrikizumab Modified Safety Population	All Lebrikizumab Safety Population *excluding* all patients from Site 5042. Safety analyses for the Combined Induction and Maintenance Periods will be conducted on All Lebrikizumab Modified Safety Population. Selective safety analyses for the Combined Induction and Maintenance Periods plus the follow-up Period will be conducted on All Lebrikizumab Modified Safety Population.

Abbreviations: EASI = Eczema Area and Severity Index; IGA = Investigator Global Assessment.

Table KGAC.6.2. Treatment Groups and Comparisons for Each Study Period and Analysis Population

Study Period	Analysis Population	Treatment Groups	Abbreviation	Inferential Comparisons When Applicable
Induction Period	mITT; Modified Safety; Safety	Placebo; Lebrikizumab 250 mg Q2W	PBO; LEB250Q2W	LEB250Q2W vs PBO
Maintenance Blinded Period	Modified Maintenance Primary; Maintenance Primary (Safety analysis only)	Lebrikizumab_Res/Placebo; Lebrikizumab_Res/Lebrikizumab 250 mg Q4W; Lebrikizumab_Res/Lebrikizumab 250 mg Q2W; Total Lebrikizumab_Res/ Lebrikizumab (Safety analysis only)	LEB_Res/PBO; LEB_Res/LEB250Q4W; LEB_Res/LEB250Q2W; Total LEB_Res/LEB (Safety analysis only)	LEB_Res/LEB250Q4W vs LEB_Res/PBO; LEB_Res/LEB250Q2W vs LEB_Res/PBO
Maintenance Blinded Period	Modified Maintenance Secondary	Placebo_Res/Placebo; Placebo_Res/Lebrikizumab 250 mg Q4W; Placebo_Res/Lebrikizumab 250 mg Q2W	PBO_Res/PBO; PBO_Res/LEB250Q4W; PBO_Res/LEB250Q2W	No Between-Group or Overall Comparisons
Maintenance Escape Period	Modified Maintenance W16 Escape	Lebrikizumab_NonResp/ Lebrikizumab 250 mg Q2W; Placebo_NonResp/Lebrikizumab 250 Q2W	LEB_NonResp/ LEB250Q2W; PBO_NonResp/ LEB250Q2W	No Between-Group or Overall Comparisons
Maintenance Escape Period	Modified Maintenance W24-48 Escape	Lebrikizumab 250 mg Q2W/Placebo/ 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; Placebo/Placebo/Lebrikizumab 250 mg Q2W; Placebo/Lebrikizumab 250 mg Q4W/Lebrikizumab 250 mg Q2W; Placebo/Lebrikizumab 250 mg Q2W/Lebrikizumab 250 mg Q2W	LEB250Q2W/PBO/LEB250Q2W; LEB250Q2W/ LEB250Q4W/ LEB250Q2W; LEB250Q2W/ LEB250Q2W/ LEB250Q2W; PBO/PBO/LEB250Q2W; PBO/LEB250Q4W/LEB250Q2W; PBO/LEB250Q2W/LEB250Q2W	No Between-Group or Overall Comparisons

Study Period	Analysis Population	Treatment Groups	Abbreviation	Inferential Comparisons When Applicable
Combined Induction and Maintenance Periods	All Lebrikizumab Modified Safety; All Lebrikizumab Safety	Any Lebrikizumab	N/A	No Between-Group or Overall Comparisons
Combined Induction and Maintenance Periods + FU	All Lebrikizumab Modified Safety; All Lebrikizumab Safety	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; Q2W = every 2 weeks; Q4W = every 4 weeks; Res = responder.

6.1.2. General Considerations for Analyses During Induction Period

Induction Period starts after the first injection of study treatment at Baseline Visit (Day 1) and ends prior to the first injection of study treatment at Week 16 or the early termination visit (ETV) (between Day 1 and Week 16). For patients who are randomized but not dosed, the Induction Period starts on the date of randomization.

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 (Day 1). If the patient 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 Pruritus Numeric Rating Scale (NRS) and Sleep-Loss due to Pruritus collected via eDiary, the baseline period is the 7-day window prior to the first injection. A patient must have responses on at least 4 of 7 days to calculate a baseline weekly mean. If a patient has 3 or fewer responses, the baseline mean value will be considered missing. eDiary data for Pruritus NRS and Sleep-loss due to Pruritus are mapped to study visit per Appendix 1.

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 to the date/time of the first injection in Induction Period.

- Treatment-emergent adverse events (TEAEs): baseline will be all results 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 geographical region (US versus EU versus rest of world), age (adolescent patients 12 to <18 years versus adults ≥18 years) and baseline disease severity (IGA 3 versus 4) as described in Section 5.3. The countries will be categorized into geographic regions for analysis (Section 6.3). Unless otherwise specified, the statistical analysis models for Induction Period will adjust for geographic region, age and baseline disease severity.

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

- Crude proportions for each treatment group along with the 95% two-sided asymptotic (ie, not continuity corrected) confidence intervals (CIs) will be provided.

- The estimated common risk difference along with 95% CIs. The common risk difference is the difference in proportions adjusted for the stratification factors as mentioned in Section 6.3. 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 along with the 95% two-sided asymptotic (ie, not continuity corrected) CIs.

Treatment comparisons of key continuous efficacy variables and health outcome variables at each postbaseline time point will be made using analysis of covariance (ANCOVA) with the following in the model: treatment group, baseline value, and stratification factors mentioned in Section 6.3. Type III tests for least squares (LS) means will be used for statistical comparison between treatment groups. The LS mean difference, standard error, 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 postbaseline 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 mentioned in Section 6.3 as fixed factors. The covariance structure to model the within-patient errors will be unstructured. If the unstructured covariance matrix results in a lack of convergence, the heterogeneous Toeplitz covariance structure, followed by the heterogeneous autoregressive covariance structure 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 postbaseline 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 (KM) product limit method maybe used to estimate the survival for time to event analyses. The log-rank test stratified by the stratification factors mentioned in Section 6.3 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 adverse events (AEs) and other categorical safety measures. Odds ratios will be created with lebrikizumab treatment as the numerator, and placebo as the denominator. Continuous vital sign and laboratory values will be analyzed by an ANCOVA with treatment and baseline value in the model.

6.1.3. General Considerations for Analyses During Maintenance Period

Maintenance Period starts at the first injection of study treatment at Week 16 and ends on the date of Week 52 or the ETV (between Weeks 16 and 52) unless specified otherwise.

For the efficacy and health outcome analyses, baseline is defined as the last available value before the first injection in Induction Period and, in most cases, will be the value recorded at Baseline Visit (Day 1).

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

Unless specified otherwise, for the safety analyses during Maintenance Period, baseline is defined as the last available value before first injection in 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 patients in the Modified Maintenance Primary Population, Modified Maintenance Secondary Population, Maintenance Primary Population, and Maintenance Secondary Population who met escape criteria (EASI-50 nonresponse) and escaped to lebrikizumab 250 mg Q2W at Weeks 24, 32, 40 and 48, only data in the Maintenance Blinded Period (up to the time of escape) will be included in both efficacy and safety analyses.

6.1.3.1. Modified Maintenance Primary and Maintenance Primary Populations

Unless otherwise specified, treatment comparisons of categorical efficacy and health outcomes variables will be analyzed using CMH test with treatment group and covariates as mentioned in Section 6.3 in the model. The CMH p-value will be reported, and the CMH adjusted odds ratio along with the 95% two-sided asymptotic (ie, not continuity corrected) CIs.

Each continuous efficacy and health outcomes measure score, change from baseline and percent improvement from baseline will be summarized by treatment group at all scheduled visits during Maintenance Period including Week 52, using descriptive statistics (n, mean, standard deviation (SD), median, minimum and maximum).

Treatment comparisons for continuous efficacy and health outcome variables will be made using ANCOVA model as specified.

When the ANCOVA is used, the model will include treatment, baseline value and covariates as mentioned in Section 6.3. The ANCOVA analysis will be conducted as described in Section 6.1.2.

The KM product limit method will be used to estimate the survival for time to event analyses (eg, time to loss of IGA response or time to loss of EASI-50 or time to loss of EASI-75). The stratified log-rank test will be performed with treatment group and covariates as mentioned in Section 6.3 in the model. A KM 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. Odds ratios will be created with lebrikizumab treatment as the numerator and placebo as the denominator. Continuous vital sign and laboratory values will be analyzed by an ANCOVA model with treatment and baseline value as independent variables.

6.1.3.2. Modified Maintenance Secondary Population

The number and percentage of patients achieving or maintaining a categorical efficacy and health outcome responses will be summarized by treatment group for all scheduled visits, including Week 52.

Selected continuous secondary efficacy and health outcomes measure score and change from baseline (or percent improvement) will be summarized by treatment group at all scheduled visits during Maintenance Period, including Week 52 using descriptive statistics (n, mean, SD, median, minimum, and maximum). No inferential statistics will be provided for this population.

6.1.3.3. Modified Maintenance Escape Populations

For the Modified Maintenance W16 Escape Population, the number and percentage of patients achieving or maintaining a categorical efficacy and health outcome responses will be summarized by treatment group for all scheduled visits, including Week 52. Selected continuous secondary efficacy and health outcomes measure score and change from baseline (or percent improvement) will be summarized by treatment group at all scheduled visits during Maintenance Period, including Week 52 using descriptive statistics (n, mean, SD, median, minimum, and maximum). No inferential statistics will be provided for this population.

For the Modified Maintenance W24-48 Escape Population who were treated with lebrikizumab 250 mg Q2W following loss of response (EASI-50 nonresponse), the number and percentage of patients 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.

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

Adverse event, exposure summary, and categorical laboratory/vital sign changes will be provided for the All Lebrikizumab Modified Safety Population and the All Lebrikizumab Safety Population during the Combined Induction and Maintenance Periods. For patients who were first exposed to lebrikizumab during Induction Period, the baseline for TEAE will utilize the baseline for Induction Period defined in Section 6.1.2; for patients who were first exposed to lebrikizumab during Maintenance Period, the baseline for TEAE will utilize the baseline for Maintenance Period defined in Section 6.1.3.

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

6.1.5. 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 Modified Safety Population and 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 6.1.4. More details on baseline and postbaseline definitions can be found in the Compound Level Safety Standard.

6.2. Primary and Supportive Estimands

There will be three estimands addressing different clinical questions of interest and intercurrent events for Induction Period. The estimands for Maintenance Period will be defined separately addressing different clinical questions of interest and intercurrent events for Maintenance Period.

6.2.1. Primary and Supportive Estimands for Induction Period

There will be three estimands of interest in analyzing primary and secondary endpoints for Induction Period. Two types of intercurrent events in terms of estimating the treatment effects for Induction Period will be considered, initiation of rescue medication as defined in Protocol Section 6.3 and permanent treatment discontinuation.

6.2.1.1. Primary Estimand (Hybrid)

The primary estimand is a hybrid estimand representing the primary clinical question of interest: what is the difference between treatment conditions, ie, Lebrikizumab vs Placebo, in the target patient population, in successful responses or means after 16 weeks achieved without use of rescue medication and if all patients 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 I/E criteria to reflect the targeted patient population for approval

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

C. How to account for intercurrent events (ICEs)

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

   b. For subjects 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 subjects 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

6.2.1.2. 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 patient 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 patient population for approval

B. Endpoint: apply to categorical endpoints

C. How to account for intercurrent events (ICEs)

    a. Subjects who require any use of rescue medication or discontinue treatment prior to week 16 will be considered as treatment failures, ie, 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

6.2.1.3. 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 patient population, in means after 16 weeks if rescue medication was not available and all patients 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 patient population for approval

B. Endpoint: apply to continuous endpoints

C. How to account for intercurrent events (ICEs)

    a. For subjects who require any use of rescue medication or discontinue 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 subjects adhered to the treatment. Therefore, hypothetical strategy is used for these types of ICEs.

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

Analytical details on how missing data including those as a result of intercurrent events will be handled for Induction Period can be found in Section 6.4.1. Detailed analyses relative to

estimands including analysis type, method and imputation, population, time point, and treatment comparisons for efficacy/health outcomes analyses can be found in Section 6.11.

The following table (Table KGAC.6.3) summarizes the analytical strategies that will be conducted on the intercurrent events for the three estimands.

Table KGAC.6.3. 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 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	NRI
Supportive Estimand for Continuous Endpoints (Hypothetical)	Hypothetical: Set to missing	Hypothetical: Set to missing	Hypothetical: Set to missing	MMRM, LOCF

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

6.2.2. Primary and Supportive Estimands for Maintenance Period

There will be four estimands of interest in analyzing endpoints for Maintenance Period. Three types of intercurrent events in terms of estimating the treatment effects for Maintenance Period will be considered, initiation of rescue medication as defined in Protocol Section 6.3, permanent treatment discontinuation and transfer to escape arm.

6.2.2.1. Maintenance Primary Estimand (Hybrid)

The maintenance primary estimand is a hybrid estimand representing the clinical question of interest: what is the difference between treatment conditions, ie, Lebrikizumab vs Placebo, in the target patient population, in successful responses or means after 52 weeks achieved without use of systemic rescue medication, without transferring to escape arm, if topical rescue medication were not available and if all patients continued with treatment except those who discontinued due to lack of efficacy?

The maintenance primary estimand is described by the following attributes:

A. Population: Modified Maintenance Primary Population as described in Section 6.1.1.

B. Endpoint: apply to all major and other secondary endpoints for Maintenance Period

C. How to account for intercurrent events (ICEs)

    a. Subjects 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, ie, non-responder, after the ICEs. Therefore, composite strategy is used for these types of ICEs.

b. For subjects who require any use of topical rescue medication, a hypothetical strategy will be used to estimate what the treatment effect would have been if subjects continued with treatment. Therefore, hypothetical strategy is used for these types of ICEs.

c. For subjects who discontinue treatment due to reasons other than lack of efficacy after week 16, a hypothetical strategy will be used to estimate what the treatment effect would have been if subjects 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

6.2.2.2. 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 difference between treatment conditions, ie, Lebrikizumab vs Placebo, in the target patient population, in successful responses or means after 52 weeks achieved without use of systemic rescue medication, without transferring to escape arm, regardless of use of topical rescue medication and if all patients continued with treatment except those who discontinued due to lack of efficacy?

The maintenance primary estimand is described by the following attributes:

A. Population: Modified Maintenance Primary Population as described in Section 6.1.1.

B. Endpoint: apply to all major and other secondary endpoints for Maintenance Period

C. How to account for intercurrent events (ICEs)

a. Subjects 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, ie, non-responder, after the ICEs. Therefore, composite strategy is used for these types of ICEs.

b. For subjects 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 subjects who discontinue treatment due to reasons other than lack of efficacy after week 16, a hypothetical strategy will be used to estimate what the treatment effect would have been if subjects 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

6.2.2.3. 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 difference between treatment conditions, ie, Lebrikizumab vs Placebo, in the target patient population, in successful responses after 52 weeks achieved without use of topical or systemic rescue medication, treatment discontinuation or transferring to escape arm?

The maintenance supportive estimand for categorical endpoints is described by the following attributes:

A. Population: Modified Maintenance Primary Population as described in Section 6.1.1.

B. Endpoint: apply to all major and other secondary categorical endpoints for Maintenance Period

C. How to account for intercurrent events (ICEs)

    a. Subjects 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, ie, 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

6.2.2.4. 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 difference between treatment conditions, ie, Lebrikizumab vs Placebo, in the target patient population, in means after 52 weeks if rescue medication was not available and all patients adhered to the treatment and did not transfer to escape arm?

The maintenance supportive estimand for continuous endpoints is described by the following attributes:

A. Population: Modified Maintenance Primary Population as described in Section 6.1.1.

B. Endpoint: apply to all major and other secondary continuous endpoints for Maintenance Period

C. How to account for intercurrent events (ICEs)

    a. For subjects 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 treatment effect would have been if rescue medication was not available and all subjects 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: difference in means between treatment conditions

Analytical details on how missing data including those as a result of intercurrent events will be handled for Maintenance Period can be found in Section 6.4.2. Detailed analyses relative to estimands including analysis type, method and imputation, population, time point, and treatment comparisons for efficacy/health outcomes analyses can be found in Section 6.11.

The following table (Table KGAC.6.4) summarizes the analytical strategies that will be conducted on the intercurrent events for the four maintenance estimands.

Table KGAC.6.4. Analysis of Primary and Supportive Estimands for Maintenance Period

Maintenance Estimand	Analysis Strategy for Intercurrent Events					Missing Data Imputation Method
	Rescue Medication		Treatment Discontinuation		Transfer to escape arm	
	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	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	MCMC-MI
Maintenance Supportive Estimand for Categorical Endpoints (Composite)	Composite: Set to nonresponder	Composite: Set to nonresponder	Composite: Set to nonresponder	Composite: Set to nonresponder	Composite: Set to nonresponder	NRI
Maintenance Supportive Estimand for Continuous Endpoints (Hypothetical)	Hypothetical: Set to missing	Hypothetical: Set to missing	Hypothetical: Set to missing	Hypothetical: Set to missing	Hypothetical: Set to missing	LOCF

6.3. 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: geographic region (US versus EU versus rest of world), age (adolescent patients 12 to <18 versus adults ≥18 years) and baseline disease severity (IGA 3 versus 4).

The statistical analysis models for the Modified Maintenance Primary Population (Maintenance Blinded Period) efficacy and health outcome analysis will include geographic region (US versus EU versus rest of world).Below are the country allocations within each geographic region.

Below are the country allocations within each geographic region.

Table KGAC.6.5. Geographic Regions for Statistical Analysis

Geographic Region	Country or Countries
US	United States
Europe	Bulgaria, Germany, Ukraine
Rest of world	Canada, Mexico, Singapore, Taiwan

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.

6.4. Handling of Dropouts or Missing Data

Depending on the estimands being addressed, different methods will be used to handle missing data. Description of the estimands can be found in Section 6.2.

6.4.1. Handling of Dropouts or Missing Data for Induction Period

For efficacy analysis relative to the primary estimand , the primary method of handling missing data including those as a result of intercurrent events will be based on Markov Chain Monte Carlo Multiple Imputation (MCMC-MI). The description of MCMC-MI method can be found in Section 6.4.1.1. Tipping point analysis as described in Section 6.4.1.2 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 6.4.1.3.

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 6.4.1.4.

For efficacy analysis relative to the supportive estimand for continuous endpoints collected multiple times post-baseline, a Mixed Model for Repeated Measures (MMRM) will be performed without explicit imputation. The description of MMRM can be found in Section 6.4.1.5.

Table KGAC.6.6 describes the planned imputation methods for efficacy and health outcome endpoints for Induction Period.

Table KGAC.6.6. 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	IGA, EASI, Pruritus NRS, sleep loss and DLQI related categorical endpoints at pre-specified timepoints	Primary Estimand (Hybrid)	MCMC-MI, Tipping point analysis (CMH)
		Supportive Estimand (Composite)	NRI (CMH)
	Remaining categorical endpoints	Supportive Estimand (Composite)	NRI (CMH)
Continuous	EASI percent change, Pruritus NRS percent change, Sleep loss change from baseline, DLQI change from baseline	Primary Estimand (Hybrid)	MCMC-MI (ANCOVA)
		Supportive Estimand (Hypothetical)	No imputation (MMRM)
	Remaining continuous endpoints collected at multiple post-baseline timepoints including BSA, POEM and CDLQI	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-model repeated measures; NRI = Nonresponder Imputation; NRS = Numeric Rating Scale.

6.4.1.1. 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 patients who receive topical rescue medication (per Protocol Section 6.3), receive systemic rescue medication, or discontinue treatment due to lack of efficacy, set to the patient'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 Table KGAC.6.7. 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.

Cochran-Mantel-Haenszel (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 2.

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 patient 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 patient 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 Pruritus NRS responses, no rounding will be performed. The imputed Pruritus NRS value will be compared directly to the observed mean baseline Pruritus-NRS value to determine whether a response was achieved:

- Percentage of patients with a Pruritus NRS of ≥4-points 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 variables. The imputed values will be used for the following secondary endpoints:

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

Table KGAC.6.7. Seed Values for MCMC-MI for Induction Period

Analysis	Seed values Lebrikizumab 250 mg Q2W Placebo
Proportion of patients achieving IGA of 0 or 1 with a ≥2-point improvement from baseline at Week 16	671970387 1339715635
Change and percent change from baseline in EASI score at 16 weeks. EASI-75 and EASI-90 will leverage imputation from EASI and therefore use the same seed numbers.	1015171075 1806114500
Change and percentage change in Pruritus NRS score from Baseline to Week 16. Proportion of patients achieving at least a 4-point improvement from baseline at Weeks 1, 2, 4, and 16 will leverage imputation from Pruritus NRS and therefore use the same seed numbers.	1461173528 1492214362
Change and percent change in Sleep loss from Baseline to Week 16. Proportion of patients achieving at least a 2-point improvement from baseline at Weeks 16 will leverage imputation from Sleep loss and therefore use the same seed numbers.	321568 765982
Change DLQI from Baseline to Week 16. Proportion of patients achieving at least a 4-point improvement from baseline at Weeks 16 will leverage imputation from DLQI and therefore use the same seed numbers.	458734 525683

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

6.4.1.2. Tipping Point Analysis

Tipping point analysis will be conducted as sensitivity analysis for the primary endpoint of an IGA 0 or 1 and a ≥2-point improvement from Baseline at Week 16 and the following secondary endpoints: EASI-75 and EASI-90 at Week 16 and Pruritus-NRS improvement ≥4-points, at Weeks 1, 2, 4 and 16. For each of these endpoint, the tipping point analysis will only be conducted if its primary or key secondary analyses results are statistically significant.

All subjects who use rescue medication or discontinue treatment due to lack of efficacy will be imputed as nonresponders. 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 the 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 responses probabilities (eg, probability = 0, 0.2 … 1) 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 6.1.2). 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 (eg, all missing responses in the placebo and lebrikizumab groups are imputed as responders and nonresponders, respectively, ie extreme case), then the p-value from the single imputed dataset will be used.

For each imputed response probability of Lebrikizumab, 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 in SAS are given in Table KGAC.6.8.

Table KGAC.6.8. Seed Values for Tipping Point Analysis

Analysis	Seed value
Proportion of patients achieving IGA of 0 or 1 with a ≥2-point improvement from baseline at Week 16	123470
Proportion of patients achieving EASI-75 and EASI-90 at Week 16	123471
Proportion of patients achieving at least a 4-point improvement from baseline at Weeks 1, 2, 4, and 16	123472, 123473, 123474, 123475 for 4 time points

6.4.1.3. Nonresponder Imputation

The nonresponder imputation (NRI) method will be used to handle missing data relative to the supportive estimand for categorical endpoints (composite). Patients who receive rescue medication (per Protocol Section 6.3), 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 nonresponder imputation (NRI) method imputes missing values as non-responders and can be justified based on the composite strategy (ICH E9R1) for handling intercurrent events. In this strategy patients 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 (ie not using rescue medications and not having missing values due to other reasons). Failing either criteria by definition makes them nonresponders.

Randomized patients without at least 1 postbaseline observation will also be defined as nonresponders for all visits for the NRI analysis.

6.4.1.4. Last Observation Carried Forward (LOCF)

In this analysis, the values subsequent to rescue medication use (per Protocol Section 6.3) 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 postbaseline observation.

6.4.1.5. Mixed-effects Model for Repeated Measures

Mixed 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 patients in the same treatment arm taking into account both the missingness of data through the correlation of the repeated measurements.

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 mentioned in Section 6.3 as fixed factors. The covariance structure to model the within-patient errors will be unstructured.

6.4.2. Handling of Dropouts or Missing Data for Maintenance Period

For maintenance efficacy analysis relative to the maintenance primary estimand, the method of handling missing data including those as a result of intercurrent events will be Markov Chain Monte Carlo Multiple Imputation (MCMC-MI). The description of maintenance MCMC-MI method can be found in Section 6.4.2.1.

MCMC-MI will also be used to handle missing data relative to the maintenance supportive estimand (Hybrid) as described in Section 6.4.2.1.

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 non-responder imputation (NRI) can be found in Section 6.4.2.2.

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 Last Observation Carried Forward (LOCF).The description of maintenance LOCF can be found in Section 6.4.2.3.

Table KGAC.6.9 describes the planned imputation methods for efficacy and health outcome endpoints for Maintenance Period.

Table KGAC.6.9. 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 Pruritus NRS related categorical endpoints at pre-specified timepoints	Maintenance Primary Estimand (Hybrid)	MCMC-MI (CMH)
		Maintenance Supportive Estimand (Hybrid)	MCMC-MI (CMH)
		Maintenance Supportive Estimand (Composite)	NRI (CMH)
	Remaining categorical endpoints	Maintenance Supportive Estimand (Composite)	NRI (CMH)
Continuous	EASI percent change, Pruritus NRS percent change	Maintenance Primary Estimand (Hybrid)	MCMC-MI (ANCOVA)
		Maintenance Supportive Estimand (Hybrid)	MCMC-MI (ANCOVA)
		Maintenance Supportive Estimand (Hypothetical)	LOCF (ANCOVA)
	Remaining continuous endpoints	Maintenance 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-model repeated measures; NRI = Nonresponder Imputation; NRS = Numeric Rating Scale.

6.4.2.1. Maintenance Period MCMC-MI

The MCMC-MI will be used to handle missing data relative to the maintenance primary estimand (Hybrid) and 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 Table KGAC.6.7. 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.

Cochran-Mantel-Haenszel (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 2.

The imputation and analysis will be conducted on the Modified Maintenance Primary Population only.

The derivation of binary responses related to EASI, IGA and Pruritus NRS for Maintenance Period will follow the derivation for Induction Period.

Table KGAC.6.10. Seed Values for MCMC-MI for Maintenance Period

Analysis	Seed values* Placebo/ Lebrikizumab 250 mg Q2W/Q4W
IGA	12345
EASI	12346
Pruritus NRS score	12347

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

6.4.2.2. Maintenance Period NRI

The nonresponder imputation (NRI) method will be used to handle missing data relative to the maintenance supportive estimand for categorical endpoints (composite). Patients who receive rescue medication (per Protocol Section 6.3), discontinue treatment, or transfer to escape arm will be set to non-response subsequent to this time through Week 52 Intermittent missing values will also be set to non-response.

Re-randomized patients without at least 1 postbaseline observation will also be defined as nonresponders for all visits for the NRI analysis.

6.4.2.3. Maintenance Period LOCF

Maintenance LOCF will be used to handle missing data relative to maintenance supportive estimands for continuous endpoints (Hypothetical). In this analysis, the values subsequent to rescue medication use (per Protocol Section 6.3), 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 postbaseline observation.

6.5. Multicenter Studies

This study will be conducted by multiple investigators at multiple sites internationally. Typically, a logistic regression with treatment, site, and treatment-by-site may be used to assess the consistence of treatment effect in sites. However, due to a large number of sites and countries and relative small sample size in the study, this logistic regression model will not likely converge. The site will not be adjusted as a covariate. Instead, the subgroup analysis on the region will be evaluated. The countries will be categorized into geographic regions as in Section 6.3. Subgroup analysis details are provided in Section 6.15.1.

For the analysis of the primary endpoint, the presence of a treatment-by-geographic region interaction will be tested at 10% significance level. Treatment group comparisons for the primary endpoint will be presented separately for each geographic region. When there is evidence of an interaction ($p<.10$), descriptive statistics may be used to assess whether the interaction is quantitative (ie, the treatment effect is consistent in direction but not size of effect) or qualitative (the treatment is beneficial for some but not other geographic regions or countries).

6.6. Multiple Comparisons/Multiplicity

6.6.1. Multiplicity Control for FDA

A prespecified graphical multiple testing approach (Bretz et al. 2009, 2011) will be implemented to control the overall Type I error rate at two-sided alpha of 0.05, for all primary and major secondary endpoints for FDA. 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 (Bretz et al. 2009, 2011; 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 primary and major secondary endpoints to be tested for FDA.

Primary endpoint:

- [IGA01 W16] Percentage of patients with an Investigator Global Assessment (IGA) score of 0 or 1 and a reduction ≥2 points from Baseline to Week 16.

Major secondary endpoints:

- [EASI-75 W16] Percentage of patients achieving EASI-75 (≥75% reduction from Baseline in Eczema Area and Severity Index [EASI] score) at Week 16.
- [EASI-90 W16] Percentage of patients achieving EASI-90 (≥90% reduction from Baseline in EASI score) at Week 16.
- [Pruritus NRS-4 W16] Percentage of patients with a Pruritus Numerical Rating Scale (NRS) of ≥4-points at Baseline who achieve a ≥4-point reduction from Baseline to Week 16.
- [Pruritus NRS-4 W4] Percentage of patients with a Pruritus NRS of ≥4-points at Baseline who achieve a ≥4-point reduction from Baseline to Week 4.
- [Pruritus NRS-4 W2] Percentage of patients with a Pruritus NRS of ≥4-points at Baseline who achieve a ≥4-point reduction from Baseline to Week 2.
- [IGA01 W4] Percentage of patients with an IGA score of 0 or 1 and a reduction ≥2 points at Week 4.
- [IGA01 Adult W16] Percentage of patients with an IGA score of 0 or 1 and a reduction ≥2 points at Week 16 in adults.

- [Sleep loss W16] Percentage of patients with a Sleep-loss score ≥2 points at Baseline who achieve a ≥2 points reduction from Baseline at Week 16.

Figure KGAC.6.2 describes the graphical testing scheme for FDA.

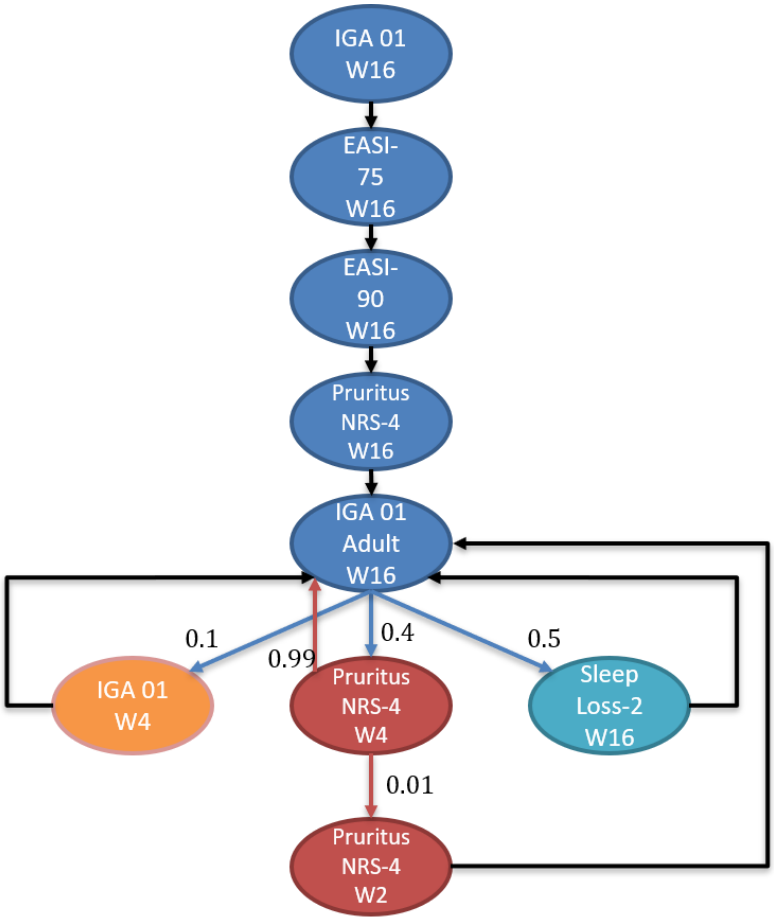


Figure KGAC.6.2. Graphical approach to control type 1 error rate for Study J2T-DM-KGAC for FDA purposes.

6.6.2. Multiplicity Control for EMA

Only for EMA purposes, two families for alpha control will be defined: 1 for induction and 1 for maintenance with each family-wise error rate at 0.05. So, different testing schemes will be used, 1 for the induction and another separate one for the maintenance period.

For all primary and major secondary endpoints for Induction Period, a prespecified graphical multiple testing approach (Bretz et al. 2009, 2011) will be implemented to control the overall Type I error rate at two-sided alpha of 0.05.

The following is a list of primary and major secondary endpoints to be tested for EMA for Induction Period.

Co-primary endpoints:

- [IGA01 W16] Percentage of patients with an IGA 0 or 1 and a ≥2-point improvement from Baseline to Week 16.
- [EASI-75 W16] Percentage of patients achieving EASI-75 (≥75% reduction from Baseline in EASI score) at Week 16.

Major secondary endpoints for Induction Period:

- [EASI-90 W16] Percentage of patients achieving EASI-90 (≥90% reduction from Baseline in EASI score) at Week 16.
- [EASI PCFB W16] Percentage change in EASI score from Baseline to Week 16.
- [EASI-90 W4] Percentage of patients achieving EASI-90 at Week 4.
- [Pruritus PCFB W16] Percentage change in Pruritus NRS score from Baseline to Week 16.
- [Pruritus NRS-4 W16] Percentage of patients with a Pruritus NRS of ≥4-points at Baseline who achieve a ≥4-point reduction from Baseline to Week 16.
- [Pruritus NRS-4 W4] Percentage of patients with a Pruritus NRS of ≥4-points at Baseline who achieve a ≥4-point reduction from Baseline to Week 4.
- [Pruritus NRS-4 W2] Percentage of patients with a Pruritus NRS of ≥4-points at Baseline who achieve a ≥4-point reduction from Baseline to Week 2.
- [DLQI W16] Percentage of patients with a DLQI total score of ≥4-points at Baseline who achieve a ≥4-point improvement from baseline to Week 16.
- [DLQI CFB W16] Change from baseline in DLQI at Week 16.
- [Sleep loss W16] Percentage of patients with a Sleep-loss score ≥2 points at Baseline who achieve a ≥2 points reduction from Baseline at Week 16.
- [Sleep loss CFB W16] Change from Baseline in Sleep-loss score at Week 16.

Figure KGAC.6.3 describes the graphical testing scheme for Induction Period for EMA.

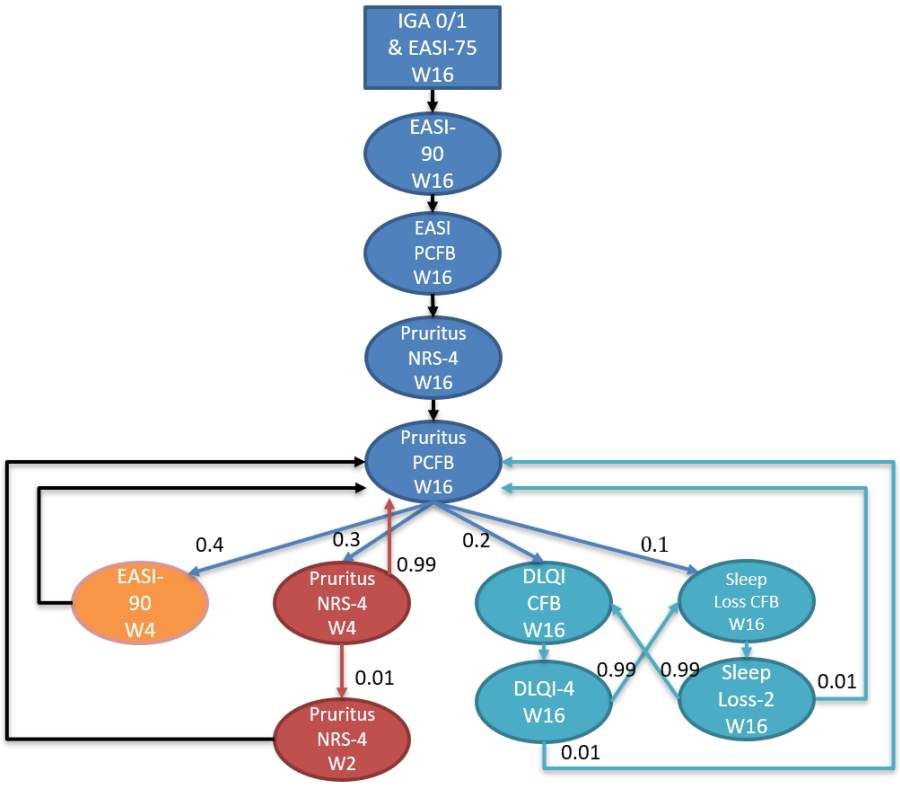


Figure KGAC.6.3. Graphical approach to control type 1 error rate for Study J2T-DM-KGAC for EMA purposes.

A separate set of major secondary endpoints will be considered at Week 52 (End of Maintenance). These secondary endpoints across the 2 different regimens will be tested following the hierarchical testing procedure with a pre-specified order, ie, inferential conclusions about secondary endpoints require statistical significance at the 0.05 significance level.

The hierarchy for the major secondary endpoints at Week 52 is as follows

Q2W Maintenance therapy:

- Percentage of patients from those re-randomized to Q2W maintenance therapy having achieved EASI-75 at Week 16 who continue to exhibit EASI-75 at Week 52 (EASI-75 calculated relative to baseline EASI score).
- Percentage of patients from those re-randomized to Q2W maintenance therapy having achieved IGA 0 or 1 and a ≥2-point improvement from Baseline at Week 16 who continue to exhibit an IGA 0 or 1 and a ≥2-point improvement from Baseline at Week 52.

- Percentage of patients from those with a Pruritus NRS of ≥4-points at baseline and re-randomized to Q2W maintenance therapy having achieved ≥4-point reduction from baseline at Week 16 who continue to exhibit ≥4-point reduction from baseline at Week 52.

Q4W Maintenance therapy:

- Percentage of patients from those re-randomized to Q4W maintenance therapy having achieved EASI-75 at Week 16 who continue to exhibit EASI-75 at Week 52 (EASI-75 calculated relative to baseline EASI score).
- Percentage of patients from those re-randomized to Q4W maintenance therapy having achieved IGA 0 or 1 and a ≥2-point improvement from Baseline at Week 16 who continue to exhibit an IGA 0 or 1 and a ≥2-point improvement from Baseline at Week 52.
- Percentage of patients from those with a Pruritus NRS of ≥4-points at baseline and re-randomized to Q4W maintenance therapy having achieved ≥4-point reduction from baseline at Week 16 who continue to exhibit ≥4-point reduction from baseline at Week 52.

Q2W Maintenance therapy:

- Percentage change in EASI Score from baseline at Week 52 for those patients re-randomized to Q2W maintenance therapy at Week 16.

Q4W Maintenance therapy:

- Percentage change in EASI Score from baseline at Week 52 for those patients re-randomized to Q4W maintenance therapy at Week 16.

6.7. Patient Disposition

The following patient disposition summaries will be provided (details of the analysis populations can be found in Section 6.1.1):

- Total number and percentage of patients entering each statistical analyses population defined in Section 6.1.1.
- The number and percentage of patients who entered the study, failed screening, were randomized at Baseline Visit (Day 1), completed Week 16, completed Week 52, completed the safety Follow-Up Visit and entered long-term extension study. Summary will be provided by the initial randomized treatment group (Analysis population: modified intent-to-treat [mITT]; intent-to-treat [ITT]).
- The number and percentage of patients who completed the study, and the number and percentage of patients who discontinued the study at any time, by the initial randomized treatment group and primary reason for discontinuation (Analysis population: mITT; ITT).
- The number and percentage of patients who completed Induction Period and the number and percentage of patients who discontinued from Induction Period, by treatment group and primary reason for discontinuation (Analysis population: mITT; ITT).

- The number and percentage of patients who completed Maintenance Period and the number and percentage of patients who discontinued from Maintenance Period, by treatment group and primary reason for discontinuation (Analysis populations: Modified Maintenance Primary Population [mMPP]; Modified Maintenance Safety Population [mMSP], Maintenance Primary Population [MPP] and Maintenance Secondary Population [MSP]), in addition, the number and percentage of patients who entered the escape arm will be summarized.

All patients who were randomized (ie, 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.

Patient allocation by region, country, and center/site will be summarized with number of patients who entered the study, number of ITT patients for each treatment group, number of patients discontinued from study treatment, and number of patients discontinued from the study.

6.8. Patient Characteristics

6.8.1. Demographics and Baseline Characteristics

Patient demographic variables and baseline characteristics will be summarized by treatment group for the mITT Population, the Modified Maintenance Primary Population, and the Modified Maintenance W16 Escape Population. 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. By-patient listings of basic demographic information for the ITT population will be provided.

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)
- Ethnicity for US (Hispanic or Latino, Not Hispanic or Latino, Not reported, Unknown)
- Region (as defined in Section 6.3)
- Country
- 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)

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 collected on the CRF.
- Duration since AD onset (years): calculated as the difference between date of Informed Consent and the date of onset of AD collected on the CRF.
- 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
    - None
    - Topical corticosteroids
    - Topical calcineurin inhibitors
    - Immunosuppressive/immunomodulating drugs: systemic corticosteroids; cyclosporine; mycophenolate-mofetil; IFN-γ; Janus kinase inhibitors; azathioprine; methotrexate
    - Phototherapy
    - Photochemotherapy (PUVA)
    - Other Biologics (eg, cell depleting biologics)
    - Other non-Biologic medication/treatment
- Prior use of systemic treatment (yes, no) 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)
- Pruritus NRS

- Pruritus NRS: <4, ≥4
- Pruritus NRS: <5, ≥5
- Sleep loss due to pruritus
- Sleep loss due to pruritus: <2, ≥2
- Patient-Oriented Eczema Measure (POEM)
- Dermatology Life Quality Index (DLQI)
- Children Dermatology Life Quality Index (CDLQI)
- EQ-5D Visual Analog Score (VAS) score
- EQ-5D US Population-based index score
- EQ-5D United Kingdom (UK) Population-based index score
- Patient-Reported Outcomes Measurement Information System (PROMIS®) Anxiety and Depression scores
- Asthma Control Questionnaire (ACQ-5) (among patients who report comorbid asthma)

6.8.2. Medical History

Medical histories are defined as the conditions/events recorded on the *Medical History* electronic case report form (eCRF) with a start date prior to the first study drug injection.

The number and percentage of patients with medical histories will be summarized for the mITT Population by treatment group and by treatment and age groups using the MedDRA Preferred Term (PT) nested within System Organ Class (SOC).

The number and percentage of patients with specific medical history events of interest pre-specified on the *History Assessment* eCRF (hand dermatitis, facial dermatitis, conjunctivitis, herpes Zoster, and others) will be summarized for the mITT Population, the Modified Maintenance Primary Population, and the Modified Maintenance W16 Escape Population by treatment group and by treatment and age groups.

6.9. Treatment Compliance

Treatment compliance with investigational product will be summarized for patients who have at least 1 dose for the Modified Safety Population in Induction Period and for all the Modified Maintenance Populations (including Modified Maintenance Primary Population, Modified Maintenance Secondary Population and Modified Maintenance W16 Escape Population) in the Maintenance Period. Treatment compliance for each patient 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 related datasets.

- The total number of injections administered will be based on the *Study Drug Administration* eCRF page and the information from the Dosing Diary.

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

Visit	Day 1	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 Induction Period occurs on Week 14.

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

Timepoint	W16	W18	W20	W22	W24	W26	W28	W30	W32	W34
Visit	**W16**		**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[a]	W52
Visit	**W36**		**W40**		**W44**		**W48**		**W52**
# injections at each visit	1	1	1	1	1	1	1	1	0
Total # injections up to each visit	13	14	15	16	17	18	19	20	20

Abbreviation: W = week.

[a] Last injection during Maintenance Period occurs on Week 50.

A patient 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.

6.10. 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 treatment. *Concomitant medications* are those medications

that start before, on, or after the first day of study treatment 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 patient 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 patient has a new start date.

Prior medication will be summarized for the mITT population. Concomitant medication during the Induction Period and Maintenance Period will be presented separately for the mITT Population and Modified Maintenance Primary Population.

6.10.1. Rescue Medication

Rescue medications during the Induction Period, Maintenance Blinded Period, and Maintenance Escape Period will be presented by the treatment groups for the mITT Population, Modified Maintenance Primary Population, and Modified Maintenance W16 Escape Population, respectively. This will include: (1) topical AD treatment (including TCS, TCI and crisaborole), (2) systemic AD treatment (including systemic corticosteroids, immunosuppressant, biologics, and phototherapy). TCS will be presented by potency. Definition of rescue medications is provided in Appendix 3.

Flare

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 patients in the mITT Population, Modified Maintenance Primary Population, and Modified 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.

6.11. Efficacy Analyses

Table KGAC.6.11 includes the description and derivation of the efficacy/health outcomes measures and endpoints.

Table KGAC.6.12 provides the detailed analyses relative to estimands including analysis type, method and imputation, population, time point, and treatment comparisons for efficacy/health outcomes analyses.

Table KGAC.6.11. Description and Derivation of Efficacy/Health Outcomes 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 patient'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.
		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 increase in IGA score ≥2 compared to Week 16 - date of W16 re-randomization +1	If a patient has not experienced loss of IGA response by completion or early discontinuation of Maintenance Blinded Period or transfer to escape arm, the patient will be censored at the date of their last visit during Maintenance Blinded Period. If a patient has not experience loss of response by the time of systemic rescue during Maintenance Blinded Period, the patient will be censored at the date of systemic rescue.

Measure	Description	Variable	Derivation / Comment	Imputation Approach if Missing Components
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: $EASI_{region}$ = (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. Then total EASI score is as follows: $EASI = 0.1*EASI_{head\ and\ neck} + 0.3*EASI_{trunk} + 0.2*EASI_{upper\ limbs} + 0.4*EASI_{lower\ limbs}$	If value of percentage involvement is 0 for any region, then severity scores of that region could be missing. Otherwise missing if any component is missing.
		Change from baseline in EASI score Percent change from baseline EASI score	Change from baseline: observed EASI score – baseline EASI score % change from baseline: $\overset{100}{\times}\frac{Observed\ score - Baseline}{Baseline}$	Missing if baseline or observed value is missing.
		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.

Measure	Description	Variable	Derivation / Comment	Imputation Approach if Missing Components
		Time to loss of EASI-50 Time to loss of EASI-75	Date of first time % change from baseline in EASI score >-50 - date of W16 re-randomization +1 Date of first time % change from baseline in EASI score >-75 - date of W16 re-randomization +1	If a patient has not experienced loss of EASI-50 or EASI-75 by completion or early discontinuation of Maintenance Blinded Period or transfer to escape arm, the patient will be censored at the date of their last visit during Maintenance Blinded Period. If a patient has not experience loss of response by the time of systemic rescue during Maintenance Blinded Period, the patient 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 patient palm = 1% rule	BSA score	BSA Total = $BSA_{head\ and\ neck} + BSA_{trunk} + BSA_{upper\ extremities} + BSA_{lower\ extremities}$	N/A – partial assessments cannot be saved.
		Change from baseline in BSA score	Change from baseline: observed BSA score – baseline BSA score	Missing if baseline or observed value is missing.
SCORing Atopic Dermatitis (SCORAD)		SCORAD score	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 component A or B or C is missing.
		Change from baseline in SCORAD score Percent change from baseline in SCORAD score	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.

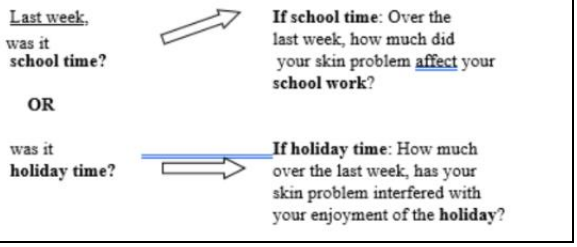
Measure	Description	Variable	Derivation / Comment	Imputation Approach if Missing Components
	SCORAD is a validated clinical tool for assessing the extent and intensity of atopic dermatitis. There are 3 components to the assessment: • The extent of AD is assessed as a percentage of each defined body area and reported as the sum of all areas, with a maximum score of 100% (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 patient 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.	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.

Measure	Description	Variable	Derivation / Comment	Imputation Approach if Missing Components
Pruritus Numeric Rating Scale (NRS)	The Pruritus Numeric Rating Scale (NRS) is a an 11-point scale used by patients 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 patient using an electronic diary.	Pruritus NRS prorated weekly mean score	The prorated weekly mean is based on previous 7 days. If the patient 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. eDiary data are mapped to study visit per Appendix 1.	Weekly mean score missing if the patient has no Pruritus-NRS responses within the week.
		Change from baseline in Pruritus NRS prorated weekly mean score Percent change from baseline in Pruritus NRS prorated weekly mean score	Change from baseline: observed Pruritus prorated weekly mean score – baseline Pruritus weekly mean score % change from baseline: $100 \times \frac{Observed\ score - Baseline}{Baseline}$	Missing if baseline or observed value is missing.
		4-point Pruritus improvement in Pruritus NRS prorated weekly mean score	Change from baseline in Pruritus NRS prorated weekly mean score ≤-4	Missing if baseline is missing or observed value is missing.
Sleep-loss due to pruritus	Sleep-loss due to pruritus 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 the patient 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. eDiary data are mapped to study visit per Appendix 1.	Weekly mean score missing if the patient has no Sleep-loss responses within the week.
		Change from baseline in Sleep-loss prorated weekly mean score Percent change from baseline in Sleep-loss prorated weekly mean score	Change from baseline: observed sleep loss prorated weekly mean score – baseline sleep loss score % change from baseline: $100 \times \frac{Observed\ score - Baseline}{Baseline}$	Missing if baseline or observed value is missing.

Measure	Description	Variable	Derivation / Comment	Imputation Approach if Missing Components
		2-point improvement in Sleep-loss prorated weekly mean score	Change from baseline in Sleep-loss prorated weekly mean score ≤-2	Missing if baseline is missing or observed value is missing.
Patient-Oriented Eczema Measure (POEM)	The POEM is a 7-item, validated, questionnaire used by the patient to assess disease symptoms over the last week. The patient 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 score	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. If more than one response is selected, then the response with the highest score is used.
		Change from baseline in POEM score	Change from baseline: observed POEM score – baseline POEM score	Missing if baseline or observed value is missing.
		4-point improvement	Change from baseline ≤-4	Missing if baseline is missing or observed value is missing.
Dermatology Life Quality Index (DLQI)	DLQI is a validated, dermatology-specific, patient-reported measure that evaluates patient'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	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 postbaseline DLQI total score of 0 or 1. A DLQI total score of 0 to 1 is considered as having no effect on a patient's HRQoL (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 is missing or observed value is missing.

Measure	Description	Variable	Derivation / Comment	Imputation Approach if Missing Components
	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 patient'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. 2008)	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 question #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.

Measure	Description	Variable	Derivation / Comment	Imputation Approach if Missing Components
		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 Index (CDLQI)	The CDLQI is designed to measure the impact of any skin disease on the lives of children. Patients ≤16 years will complete the CDLQI and should 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 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.
		CDLQI (0,1)	A CDLQI (0,1) response is defined as a postbaseline CDLQI total score of 0 or 1.	Missing if CDLQI total score is missing
		4-point improvement	Change from baseline ≤-4	Missing if baseline is missing 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.

Measure	Description	Variable	Derivation / Comment	Imputation Approach if Missing Components
		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.
				

Measure	Description	Variable	Derivation / Comment	Imputation Approach if Missing Components
		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.
European Quality of Life–5 Dimensions–5 Levels (EQ-5D-5L)	EQ-5D comprises five dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. The EQ VAS records the patient's self-rated health on a vertical visual analogue scale. The scores on these five dimensions can be presented as a health profile or can be converted to a single summary index number (utility) reflecting preferability compared to other health profiles	▪ EQ-5D mobility ▪ EQ-5D self-care ▪ EQ-5D usual activities ▪ EQ-5D pain/ discomfort ▪ EQ-5D anxiety/ depression	Five health profile dimensions, each dimension has 5 levels: 1 = no problems 2 = slight problems 3 = moderate problems 4 = severe problems 5 = extreme problems It should be noted that the numerals 1 to 5 have no arithmetic properties and should not be used as a primary score.	Each dimension is a single item, missing if missing.
		▪ EQ-5D VAS	Single item. Range 0 to 100. 0 represents "worst health you can imagine" 100 represents "best health you can imagine"	Single item, missing if missing.
		Change from baseline in EQ-5D VAS	Change from baseline: observed EQ-5D VAS score – baseline EQ-5D VAS score	Missing if baseline or observed value is missing.

Measure	Description	Variable	Derivation / Comment	Imputation Approach if Missing Components
		EQ-5D-5L UK Population-based index score (health state index)	Derive EQ-5D-5L UK Population-based index score according to the link by using the UK algorithm to produce a patient-level index score between -0.59 and 1.0 (continuous variable).	N/A – partial assessments cannot be saved on the eCOA tablet.
		Change from baseline in EQ-5D-5L UK Population-based index score	Change from baseline: observed EQ-5D-5L UK score – baseline EQ-5D-5L UK score	Missing if baseline or observed value is missing.
		EQ-5D-5L US Population-based index score (health state index)	Derive EQ-5D-5L US Population-based index score according to the link by using the US algorithm to produce a patient-level index score between -0.11 and 1.0 (continuous variable).	N/A – partial assessments cannot be saved on the eCOA tablet.
		Change from baseline in EQ-5D-5L US Population-based index score	Change from baseline: observed EQ-5D-5L US score – baseline EQ-5D-5L US score	Missing if baseline or observed value is missing.
Patient-Reported Outcomes Measurement Information System (PROMIS®)	PROMIS® is a set of person-centered measures that evaluates and monitors physical, mental, and social health in adults and children. Pediatric and tools for anxiety and depression. Patients ≤17 years will complete pediatric versions for the duration of the study.	PROMIS anxiety total score PROMIS depression total score	A PROMIS anxiety has 8 questions on Emotion Distress-Anxiety (or Pediatric Anxiety) -Short Form 8a. Each ranges 1 to 5. Total raw scores are converted to T-Scores with higher scores representing greater anxiety. A PROMIS depression has 8 questions on Emotion Distress-Depression (or Pediatric Depressive Symptom) -Short Form 8a. Each ranges 1 to 5. Total raw scores are converted to T-score with higher scores representing greater depression.	Total score can be derived even with partial response as instrument use item response theory method.

Measure	Description	Variable	Derivation / Comment	Imputation Approach if Missing Components
		Change from baseline in PROMIS anxiety total score Change from baseline in PROMIS depression total score	Change from baseline: observed score – baseline PROMIS anxiety total score Change from baseline: observed score – baseline PROMIS depression total score	Missing if baseline or observed value is missing.
Asthma Control Questionnaire (ACQ-5)	Patients 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 patients 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.	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.
		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.
		MCID of 0.5	Change from baseline ≤-0.5	Missing if baseline is missing or observed value is missing.
Modified Subcutaneous Administration Assessment Questionnaire (SQAAQ)	Adolescent patients from EU may complete the modified SQAAQ uses 10 questions to assess the acceptability and tolerability with using a device to administer a subcutaneous injection. The person who administered the dose (adolescent patient or their parent/caregiver) should complete a 7-point Likert scale (from "Strongly Disagree" to "Strongly Agree") shortly after completing the injection.	Respond "Strongly Agree" or "Agree" for each self/caregiver administration of the study drug	For each EU adolescent patient have SQAAQ scale completed, the proportion of patients who answer "Strongly Agree" or "Agree" in each of the 10 questions	Missing data will be treated as missing;

Abbreviations: eCOA = electronic clinical outcome assessment

Table KGAC.6.12. Description of Efficacy/Health Outcome Analyses

Measure	Variable	Estimand (Section 6.2)	Analysis Method (Section 6.1)	Population (Section 6.1.1)	Comparison/Time Point	Analysis Type
Investigator's Global Assessment (IGA)	Proportion of patients achieving IGA [0,1] with a ≥2-point improvement	Primary Estimand (Hybrid)	CMH analysis with MCMC-MI	mITT	Leb 250 mg Q2W vs PBO; Week 16 and all scheduled visits in Induction Period	Primary analysis: W16; Key secondary W4 (for FDA only); Secondary analysis: other timepoints
			CMH analysis with tipping point analysis	mITT	Leb 250 mg Q2W vs PBO; Week 16	Sensitivity analysis
		Supportive Estimand (Composite)	CMH analysis with NRI	mITT	Leb 250 mg Q2W vs PBO; Week 16 and all scheduled visits in Induction Period	Supplementary analysis
	Proportion of patients achieving both IGA [0,1] with a ≥2-point improvement and a ≥4-point improvement in Pruritus Numeric Rating Scale (NRS)	Primary Estimand (Hybrid)	CMH analysis with MCMC-MI	mITT with Baseline Pruritus NRS score at least 4	Leb 250 mg Q2W vs PBO; Week 16 and all scheduled visits in Induction Period	Secondary analysis
		Supportive Estimand (Composite)	CMH analysis with NRI	mITT with Baseline Pruritus NRS score at least 4	Leb 250 mg Q2W vs PBO; Week 16 and all scheduled visits in Induction Period	Secondary analysis
	Proportion of patients achieving IGA [0]	Primary Estimand (Hybrid)	CMH analysis with MCMC-MI	mITT	Leb 250 mg Q2W vs PBO; Week 16 and all scheduled visits in Induction Period	Secondary analysis
	Maintenance of IGA [0,1]:	Maintenance Primary Estimand (Hybrid)	CMH analysis with MCMC-MI		Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; All scheduled visits in Maintenance	Key secondary analysis: Week 52; Secondary analysis: other timepoints

Measure	Variable	Estimand (Section 6.2)	Analysis Method (Section 6.1)	Population (Section 6.1.1)	Comparison/Time Point	Analysis Type
	Proportion of patients maintaining IGA [0,1] with a ≥2-point improvement from baseline among those re-randomized patients who achieved IGA [0,1] with a ≥2-point improvement from Baseline at Week 16	Maintenance Supportive Estimand (Hybrid)	CMH analysis with MCMC-MI	mMPP who have achieved IGA [0,1] with a ≥2-point improvement from Baseline at Week 16	Period	Supplementary analysis
		Maintenance Supportive Estimand (Composite)	CMH analysis with NRI			Supplementary analysis
	Time to loss of IGA response	NA	KM method with log-rank test	mMPP who have achieved IGA [0,1] with a ≥2-point improvement from Baseline at Week 16	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO	Secondary analysis
	Proportion of patients with IGA [0,1] with a ≥2-point improvement from baseline	NA	Descriptive statistics	mMSP	No comparisons. All scheduled visits in Maintenance Period	Secondary analysis
	Proportion of patients with IGA [0,1] with a ≥2-point improvement from baseline	NA	Descriptive statistics	Modified Maintenance W16 Escape Population	No comparisons. All scheduled visits in Maintenance Period	Secondary analysis

Measure	Variable	Estimand (Section 6.2)	Analysis Method (Section 6.1)	Population (Section 6.1.1)	Comparison/Time Point	Analysis Type
	Proportion of patients with IGA[0,1] with a ≥2-point improvement from baseline after lebrikizumab retreatment	NA	Descriptive statistics	Modified Maintenance W24-48 Escape Population	No comparisons. Every 4 weeks after escape and re-treated by lebrikizumab 250mg Q2W	Secondary analysis
Eczema Area and Severity Index (EASI)	Change from baseline in EASI score Percent change from baseline in EASI score	Primary Estimand (Hybrid)	ANCOVA with MCMC-MI	mITT	Leb 250mg Q2W vs PBO; Week 16 and all scheduled visits in Induction Period	Key secondary analysis: percent change at Week 16; Secondary analysis: other timepoints
		Supportive Estimand (Hypothetical)	MMRM with observed data	mITT	Leb 250 mg Q2W vs PBO; Week 16 and all scheduled visits in Induction Period	Supplementary analysis
		Maintenance Primary Estimand (Hybrid)	ANCOVA with MCMC-MI	mMPP	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; All scheduled visits in Maintenance Period	Key secondary analysis: percent change at Week 52; Secondary analysis: other timepoints
		Maintenance Supportive Estimand (Hybrid)	ANCOVA with MCMC-MI			Supplementary analysis
		Maintenance Supportive Estimand (Hypothetical)	ANCOVA with LOCF			Supplementary analysis
		NA	Descriptive statistics	Modified Maintenance W16 Escape Population	No comparisons. All scheduled visits in Maintenance Period	Secondary analysis

Measure	Variable	Estimand (Section 6.2)	Analysis Method (Section 6.1)	Population (Section 6.1.1)	Comparison/Time Point	Analysis Type
	Proportion of patients achieving EASI-75 Proportion of patients achieving EASI-90	Primary Estimand (Hybrid)	CMH analysis with MCMC-MI	mITT	Leb 250 mg Q2W vs PBO; Week 16 and all scheduled visits in Induction Period	Primary analysis (for EMA only): EASI-75, W16; Key secondary analysis: EASI-90, W16, W4 (for EMA only); Secondary analysis: other timepoints
			CMH analysis with tipping point analysis	mITT	Leb 250 mg Q2W vs PBO; Week 16	Sensitivity analysis
		Supportive Estimand (Composite)	CMH analysis with NRI	mITT	Leb 250 mg Q2W vs PBO; Week 16 and all scheduled visits in Induction Period	Supplementary analysis
	Proportion of patients achieving EASI-50	Primary Estimand (Hybrid)	CMH analysis with MCMC-MI	mITT	Leb 250 mg Q2W vs PBO; Week 16 and all scheduled visits in Induction Period	Secondary analysis
	Proportion of patients achieving both EASI-75 and a ≥4-point improvement in Pruritus NRS	Primary Estimand (Hybrid)	CMH analysis with MCMC-MI	mITT with Baseline Pruritus NRS score at least 4	Leb 250 mg Q2W vs PBO; Week 16 and all scheduled visits in Induction Period	Secondary analysis
		Supportive Estimand (Composite)	CMH analysis with NRI	mITT with Baseline Pruritus NRS score at least 4	Leb 250 mg Q2W vs PBO; Week 16 and all scheduled visits in Induction Period	Secondary analysis
	Proportion of patients maintaining EASI-75 among those re-randomized patients who achieved EASI-75 at Week 16	Maintenance Primary Estimand (Hybrid)	CMH with MCMC-MI	mMPP who have achieved EASI-75 at Week 16	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; All scheduled visits in Maintenance Period	Key secondary analysis: Week 52; Secondary analysis: other timepoints
		Maintenance Supportive Estimand (Hybrid)	CMH with MCMC-MI			Supplementary analysis

Measure	Variable	Estimand (Section 6.2)	Analysis Method (Section 6.1)	Population (Section 6.1.1)	Comparison/Time Point	Analysis Type
		Maintenance Supportive Estimand (Composite)	CMH with NRI			Supplementary analysis
	Time to loss of EASI-50	NA	KM method with log-rank test	mMPP	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO	Secondary analysis
	Time to loss of EASI-75	NA	KM method with log-rank test	mMPP who have achieved EASI-75 at Week 16	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO	Secondary analysis
	Proportion of patients with EASI-75	NA	Descriptive statistics	mMSP	No comparisons. All scheduled visits in Maintenance Period	Secondary analysis
	Proportion of patients with EASI-75	NA	Descriptive statistics	Modified Maintenance W16 Escape Population	No comparisons. All scheduled visits in Maintenance Period	Secondary analysis
	Proportion of patients with EASI-75 after lebrikizumab retreatment	NA	Descriptive statistics	Modified Maintenance W24-48 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 score	Supportive Estimand (Hypothetical)	MMRM with observed data	mITT	Leb 250 mg Q2W vs PBO; Week 16 and all scheduled visits in Induction Period	Secondary analysis
		Maintenance Supportive Estimand (Hypothetical)	ANCOVA with LOCF	mMPP	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; All scheduled visits in Maintenance Period	Secondary analysis
		NA	Descriptive statistics	Modified Maintenance W16 Escape Population	No comparisons. All scheduled visits in Maintenance Period	Secondary analysis

Measure	Variable	Estimand (Section 6.2)	Analysis Method (Section 6.1)	Population (Section 6.1.1)	Comparison/Time Point	Analysis Type
Pruritus NRS	Change from baseline in Pruritus NRS Percent Change from baseline in Pruritus NRS	Primary Estimand (Hybrid)	ANCOVA with MCMC-MI	mITT	Leb 250 mg Q2W vs PBO; Week 16 and all scheduled visits in Induction Period	Key secondary analysis: W16; Secondary analysis: other timepoints
		Supportive Estimand (Hypothetical)	MMRM with observed data	mITT	Leb 250 mg Q2W vs PBO; Week 16 and all scheduled visits in Induction Period	Supplementary analysis
		Maintenance Primary Estimand (Hybrid)	ANCOVA with MCMC-MI	mMPP	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; All scheduled visits in Maintenance Period	Secondary analysis
		Maintenance Supportive Estimand (Hybrid)	ANCOVA with MCMC-MI			Supplementary analysis
		Maintenance Supportive Estimand (Hypothetical)	ANCOVA with LOCF			Supplementary analysis
		NA	Descriptive statistics	Modified Maintenance W16 Escape Population	No comparisons. All scheduled visits in Maintenance Period	Secondary analysis
	Proportion of patients achieving at least 4-point improvement in pruritus NRS	Primary Estimand (Hybrid)	CMH analysis with MCMC-MI	mITT, mITT with Baseline Pruritus NRS score at least 4, mITT with Baseline Pruritus NRS score at least 5	Leb 250 mg Q2W vs PBO; Week 16 and all scheduled visits in Induction Period	Key secondary analysis: 2, 4, and 16 for mITT with Baseline Pruritus NRS score at least 4. Secondary analysis: other timepoints and population
			CMH analysis with tipping point analysis	mITT with Baseline Pruritus NRS score at least 4	Leb 250 mg Q2W vs PBO; Weeks 1, 2, 4 and 16	Sensitivity analysis
		Supportive Estimand (Composite)	CMH analysis with NRI	mITT with Baseline Pruritus NRS score at least 4	Leb 250 mg Q2W vs PBO; Week 16 and all scheduled visits in Induction Period	Supplementary analysis

Measure	Variable	Estimand (Section 6.2)	Analysis Method (Section 6.1)	Population (Section 6.1.1)	Comparison/Time Point	Analysis Type
	Proportion of patients maintaining ≥4-point reduction from baseline among those patients with Pruritus NRS of **≥4**-point at baseline and re-randomized and who achieved ≥4-point reduction from baseline at Week 16	Maintenance Primary Estimand (Hybrid)	CMH with MCMC-MI	mMPP with Pruritus NRS of ≥4-points at baseline and who achieved ≥4-point reduction from baseline at Week 16	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; All scheduled visits in Maintenance Period	Key secondary analysis: Week 52; Secondary analysis: other timepoints
		Maintenance Supportive Estimand (Hybrid)	CMH with MCMC-MI			Supplementary analysis
		Maintenance Supportive Estimand (Composite)	CMH with NRI			Supplementary analysis
	Proportion of patients maintaining ≥4-point reduction from baseline among those patients with Pruritus NRS of **≥5**-point at baseline and re-randomized and who achieved ≥4-point reduction from baseline at Week 16	Maintenance Primary Estimand (Hybrid)	CMH with MCMC-MI	mMPP with Pruritus NRS of ≥5-points at baseline and who achieved ≥4-point reduction from baseline at Week 16	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; All scheduled visits in Maintenance Period	Secondary analysis
	Proportion of patients with ≥4-point reduction from baseline among those patients with Pruritus NRS of **≥4**-point at baseline	NA	Descriptive statistics	mMSP with Baseline Pruritus NRS score at least 4	No comparisons. all scheduled visits in Maintenance Period	Secondary analysis

Measure	Variable	Estimand (Section 6.2)	Analysis Method (Section 6.1)	Population (Section 6.1.1)	Comparison/Time Point	Analysis Type
	Proportion of patients with ≥4-point reduction from baseline among those patients with Pruritus NRS of ≥4-point at baseline	NA	Descriptive statistics	Modified Maintenance W16 Escape Population with Baseline Pruritus NRS score at least 4	No comparisons. all scheduled visits in Maintenance Period	Secondary analysis
	Proportion of patients with ≥4-point reduction from baseline after lebrikizumab retreatment among those patients with Pruritus NRS of ≥4-point at baseline	NA	Descriptive statistics	Modified Maintenance W24-48 Escape Population with Baseline Pruritus NRS score at least 4	No comparisons. Every 4 weeks after escape and re-treated by lebrikizumab 250mg Q2W	Secondary analysis
Sleep-loss Score	Percent Change from baseline in Sleep-loss Change from baseline in Sleep-loss	Primary Estimand (Hybrid)	ANCOVA with MCMC-MI	mITT	Leb 250 mg Q2W vs PBO; Week 16 and all scheduled visits in Induction Period	Key secondary analysis: percent change and change, W16; Secondary analysis: other timepoints
		Supportive Estimand (Hypothetical)	MMRM with observed data	mITT	Leb 250 mg Q2W vs PBO; Week 16 and all scheduled visits in Induction Period	Key secondary analysis: percent change and change, W16; Secondary analysis: other timepoints
		Maintenance Supportive Estimand (Hypothetical)	ANCOVA with LOCF	mMPP	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; All scheduled visits in Maintenance Period	Secondary analysis

Measure	Variable	Estimand (Section 6.2)	Analysis Method (Section 6.1)	Population (Section 6.1.1)	Comparison/Time Point	Analysis Type
	Proportion of patients achieving at least 2-point improvement in Sleep-loss in patients who had baseline Sleep-loss ≥2	Primary Estimand (Hybrid)	CMH analysis with MCMC-MI	mITT with Baseline Sleep-loss score at least 2	Leb 250 mg Q2W vs PBO; Week 16 and all scheduled visits in Induction Period	Key secondary analysis: Weeks 16; Secondary analysis: other timepoints
		Supportive Estimand (Composite)	CMH analysis with NRI	mITT with Baseline Sleep-loss score at least 2	Leb 250 mg Q2W vs PBO; Week 16 and all scheduled visits in Induction Period	Supplementary analysis
	Proportion of patients achieving at least 2-point improvement in Sleep-loss in patients who had baseline Sleep-loss ≥2	Maintenance Supportive Estimand (Composite)	CMH with NRI	mMPP with Baseline Sleep-loss score at least 2	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; All scheduled visits in Maintenance Period	Secondary analysis
	Proportion of patients achieving at least 2-point improvement in Sleep-loss in patients who had baseline Sleep-loss ≥2	NA	Descriptive statistics	Modified Maintenance W16 Escape Population with Baseline Sleep-loss score at least 2	No comparisons. All scheduled visits in Maintenance Period	Secondary analysis
(Children) Dermatology Life Quality Index (DLQI/ CDLQI)	Change from baseline in DLQI total score	Primary Estimand (Hybrid)	ANCOVA with MCMC-MI	mITT	Leb 250 mg Q2W vs PBO; Week 16 and all scheduled visits in Induction Period	Key secondary analysis: W16 for DLQI chg; Secondary analysis: other timepoints
	Change from baseline in DLQI total score Change from baseline in CDLQI total score	Supportive Estimand (Hypothetical)	MMRM with observed data	mITT	Leb 250 mg Q2W vs PBO; Week 16 and all scheduled visits in Induction Period	Supplementary analysis for DLQI; Secondary analysis for CDLQI

Measure	Variable	Estimand (Section 6.2)	Analysis Method (Section 6.1)	Population (Section 6.1.1)	Comparison/Time Point	Analysis Type
	Change from baseline in DLQI total score Change from baseline in CDLQI total score	Maintenance Supportive Estimand (Hypothetical)	ANCOVA with LOCF	mMPP	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; All scheduled visits in Maintenance Period	Secondary analysis
	Change from baseline in DLQI total score Change from baseline in CDLQI total score	NA	Descriptive stats	Modified Maintenance W16 Escape Population	No comparisons. all scheduled visits in Maintenance Period	Secondary analysis
	Proportion of patients achieving at least 4-point improvement in DLQI	Primary Estimand (Hybrid)	CMH analysis with MCMC-MI	mITT, mITT with Baseline DLQI score at least 4	Leb 250 mg Q2W vs PBO; Week 16 and all scheduled visits in Induction Period	Key secondary analysis: W16 for mITT with Baseline DLQI score at least 4, Secondary analysis: other timepoints and population
		Supportive Estimand (Composite)	CMH analysis with NRI	mITT with Baseline DLQI score at least 4	Leb 250 mg Q2W vs PBO; Week 16 and all scheduled visits in Induction Period	Supplementary analysis
	Proportion of patients achieving ≥ 4 point improvement in DLQI in patients who had baseline DLQI score ≥4	Maintenance Supportive Estimand (Composite)	CMH with NRI	mMPP with baseline DLQI score at least 4	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; All scheduled visits in Maintenance Period	Secondary analysis

Measure	Variable	Estimand (Section 6.2)	Analysis Method (Section 6.1)	Population (Section 6.1.1)	Comparison/Time Point	Analysis Type
	Proportion of patients achieving ≥ 4 point improvement in DLQI in patients who had baseline DLQI score ≥4	NA	Descriptive statistics	Modified Maintenance W16 Escape Population with Baseline DLQI score at least 4	No comparisons. All scheduled visits in Maintenance Period	Secondary analysis
SCORing Atopic Dermatitis (SCORAD)	Change from baseline in SCORAD score Percent change from baseline in SCORAD score	Supportive Estimand (Hypothetical)	ANCOVA with LOCF	mITT	Leb 250 mg Q2W vs PBO; Week 16 and all scheduled visits in Induction Period	Secondary analysis
		Maintenance Supportive Estimand (Hypothetical)	ANCOVA with LOCF	mMPP	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; All scheduled visits in Maintenance Period	Secondary analysis
		NA	Descriptive stats	Modified Maintenance W16 Escape Population	No comparisons. all scheduled visits in Maintenance Period	Secondary analysis
	Proportion of patients achieving SCORAD75 Proportion of patients achieving SCORAD90	Supportive Estimand (Composite)	CMH analysis with NRI	mITT	Leb 250 mg Q2W vs PBO; Week 16 and all scheduled visits in Induction Period	Secondary analysis
		Maintenance Supportive Estimand (Composite)	CMH with NRI	mMPP	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; All scheduled visits in Maintenance Period	Secondary analysis
		NA	Descriptive stats	Modified Maintenance W16 Escape Population	No comparisons. all scheduled visits in Maintenance Period	Secondary analysis

Measure	Variable	Estimand (Section 6.2)	Analysis Method (Section 6.1)	Population (Section 6.1.1)	Comparison/Time Point	Analysis Type
	Percentage change in SCORAD (having achieved EASI-75 at W16) from baseline	Maintenance Supportive Estimand (Hypothetical)	ANCOVA with LOCF	mMPP who have achieved EASI-75 at Week 16	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; All scheduled visits in Maintenance Period	Secondary analysis
Patient-Oriented Eczema Measure (POEM)	Change from baseline in POEM score	Supportive Estimand (Hypothetical)	MMRM with observed data	mITT	Leb 250 mg Q2W vs PBO; Week 16 and all scheduled visits in Induction Period	Secondary analysis
		Maintenance Supportive Estimand (Hypothetical)	ANCOVA with LOCF	mMPP	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; all scheduled visits in Maintenance Period	Secondary analysis
		NA	Descriptive stats	Modified Maintenance W16 Escape Population	No comparisons. All scheduled visits in Maintenance Period	Secondary analysis
		Supportive Estimand (Composite)	CMH analysis with NRI	mITT	Leb 250 mg Q2W vs PBO; Week 16 in Induction Period	Secondary analysis

Measure	Variable	Estimand (Section 6.2)	Analysis Method (Section 6.1)	Population (Section 6.1.1)	Comparison/Time Point	Analysis Type
European Quality of Life–5 Dimensions–5 Levels (EQ-5D-5L)	Proportion of patients having no problem in each domain: • EQ-5D mobility • EQ-5D self-care • EQ-5D usual activities • EQ-5D pain/ discomfort EQ-5D anxiety/ depression	Maintenance Supportive Estimand (Composite)	CMH analysis with NRI	mMPP	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; Week 52 in Maintenance Period	Secondary analysis
	Change from baseline in • EQ-5D VAS • EQ-5D-5L UK Population-based index score • EQ-5D-5L US Population-based index score	Supportive Estimand (Hypothetical)	ANCOVA with LOCF	mITT	Leb 250 mg Q2W vs PBO; Week 16 and all scheduled visits in Induction Period	Secondary analysis
		Maintenance Supportive Estimand (Hypothetical)	ANCOVA with LOCF	mMPP	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; All scheduled visits in Maintenance Period	Secondary analysis
		NA	Descriptive stats	Modified Maintenance W16 Escape Population	No comparisons. All scheduled visits in Maintenance Period	Secondary analysis
	Change from baseline in PROMIS Anxiety score	Supportive Estimand (Hypothetical)	ANCOVA with LOCF	mITT	Leb 250 mg Q2W vs PBO; Week 16 and all scheduled visits in Induction Period	Secondary analysis

Measure	Variable	Estimand (Section 6.2)	Analysis Method (Section 6.1)	Population (Section 6.1.1)	Comparison/Time Point	Analysis Type
Patient-Reported Outcomes Measurement Information System (PROMIS®)	Change from baseline in PROMIS Depression score	Maintenance Supportive Estimand (Hypothetical)	ANCOVA with LOCF	mMPP	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; All scheduled visits in Maintenance Period	Secondary analysis
		NA	Descriptive stats	Modified Maintenance W16 Escape Population	No comparisons. All scheduled visits in Maintenance Period	Secondary analysis
Asthma Control Questionnaire (ACQ-5)	Change from baseline in ACQ-5 score	Supportive Estimand (Hypothetical)	ANCOVA with LOCF	mITT with self-reported comorbid asthma	Leb 250 mg Q2W vs PBO; Week 16 and all scheduled visits in Induction Period	Secondary analysis
		Maintenance Supportive Estimand (Hypothetical)	ANCOVA with LOCF	mMPP with self-reported comorbid asthma	Leb 250 mg Q2W vs PBO; Leb 250 mg Q4W vs PBO; All scheduled visits in Maintenance Period	Secondary analysis
Modified Subcutaneous Administration Assessment Questionnaire (SQAAQ)	Proportion of patients who answer "Strongly Agree" or "Agree" in each of 10 questions in a visit	NA	Descriptive stats	Patients who complete SQAAQ at any visit	Leb 250 mg Q2W vs PBO; by sequence of each self/caregiver injection; (note, patient could start self/caregiver injection at any visit, the visits will be aligned as: first self/caregiver injection, second self/caregiver injection…)	Secondary analysis

Abbreviations: ANCOVA = analysis of covariance; CMH = Cochran-Mantel-Haenszel; EASI = Eczema Area and Severity Index; ITT = intent-to-treat; KM = Kaplan-Meier; Leb = lebrikizumab; LOCF = last observation carried forward; MCMC-MI = Markov chain Monte Carlo multiple imputation; mITT = modified intent-to-treat; mMPP = modified maintenance primary population; mMSP = modified maintenance secondary population; MMRM = mixed model repeated measures; NRI = nonresponder imputation; PBO = placebo; Q2W = every 2 weeks; Q4W = every 4 weeks; SCORAD = SCORing Atopic Dermatitis; SQAAQ= subcutaneous administration assessment questionnaire; VAS = Visual Analog Scale; W = week.

6.11.1. Primary Outcome and Methodology

The primary analysis of the study is to test the null hypotheses that lebrikizumab 250 mg Q2W is the same as placebo when evaluating the proportion of patients achieving IGA of 0 or 1 at Week 16 in the mITT population. For EMA, an additional null hypothesis is that lebrikizumab 250 mg Q2W is the same as placebo when evaluating the proportion of patients achieving EASI-75 at Week 16 in the mITT population.

The primary estimand for the primary analysis is described in Section 6.2.1.1. The missing data including those as a result of intercurrent events will be imputed using MCMC-MI based on missing at random assumption (Section 6.4.1.1).

A CMH test as described in Section 6.1.2 will be used for the comparisons. The odds ratio, 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 major secondary objectives 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 6.6.

Primary outcome IGA 0/1 ad EASI-75 and their analysis are described in Table KGAC.6.11.

6.11.2. Sensitivity Analyses of Primary Outcome

Sensitivity analyses are included to demonstrate robustness of analyses. Tipping point analysis as described in Section 6.4.1.2 will serve as the sensitivity analyses for primary outcomes.

Sensitivity and supplementary analyses for both primary and secondary endpoints are described in Table KGAC.6.11 and Table KGAC.6.12.

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

6.11.3. Major Secondary Efficacy Analyses

Major secondary outcomes and their analyses are described in Table KGAC.6.11 and Table KGAC.6.12.

6.11.4. Other Secondary Efficacy Analyses

Other secondary outcomes and their analyses are described in Table KGAC.6.11 and Table KGAC.6.12.

6.12. Health Outcomes/Quality-of-Life Analyses

Analyses of POEM, DLQI, EQ-5D-5L, PROMIS, and ACQ-5 are described in Table KGAC.6.11 and Table KGAC.6.12.

6.13. Bioanalytical and Pharmacokinetic/Pharmacodynamic Methods

Details of PK/pharmacodynamic (PD) analyses can be found in a separate PK/PD analysis plan.

6.14. 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 (eg, FDA Clinical Review Template), and cross-industry standardization efforts (eg, 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:

- Modified Safety Population (Induction Period),
- Safety Population (Induction Period) as a sensitivity analysis,
- Modified Maintenance Primary Population (Maintenance Blinded Period),
- Maintenance Primary Population (Maintenance Blinded Period) as a sensitivity analysis,
- All Lebrikizumab Modified Safety Population (Combined Induction and Maintenance Periods, and Combined Induction and Maintenance Periods plus Follow Up Period [selective analysis])
- All Lebrikizumab Safety Population (Combined Induction and Maintenance Periods, and Combined Induction and Maintenance Periods plus Follow Up Period [selective analysis]) as a sensitivity analysis.

These analysis populations are fully defined in Table KGAC.6.1 while Table KGAC.6.2 describes the treatment groups, associated study periods, and the comparisons for each analysis population.

Selected safety summaries as sensitivity analysis will be conducted on Safety Population, Maintenance Primary Population, and All Lebrikizumab Safety Population, including overview of AEs, Summary of TEAE PTs by maximum severity and a listing of TEAEs that occurred in the safety population but not in modified safety population.

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.

Not all displays described in this section 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.

6.14.1. Extent of Exposure

Duration of exposure to study treatment 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:

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

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

- >0, ≥7, ≥14, ≥30, ≥60, ≥90, ≥112, ≥120 days for Induction Period (for Maintenance Period, use ≥30, ≥60, ≥90, ≥120, ≥150, ≥180, ≥210, ≥240, and ≥252 days). Note that patients 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 <120, ≥120 days (for 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).

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 patient years, calculated as:

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

- 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 patients in Safety Population randomized to lebrikizumab 250 mg Q2W or patients in Maintenance Primary Population re-randomized to lebrikizumab 250 mg Q2W or Q4W, the total dose (in mg) taken during Induction Period or Maintenance Period will be calculated as follows: *Total lebrikizumab dose=Total number of active injections (including loading doses, if any) received in Induction Period or Maintenance Period ×250*.
- Total number of injections received will be derived based on the *Study Drug Administration* eCRF page and the response to the question "Did you or a caregiver successfully inject the study drug?" on the Dosing Diary eCRF page.

The exposure for the All Lebrikizumab Modified 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 patient receives lebrikizumab and then summed together (this excludes the duration of time that patients are receiving placebo during the Maintenance Period). If a patient was randomized to lebrikizumab during Induction Period, then to placebo during Maintenance

Period and later on entered escape arm following loss of response (<EASI-50), the patient's exposure on lebrikizumab during Maintenance Period will be calculated (Date of last study visit during Maintenance Period– Date of first injection to resume lebrikizumab + 1 day) and will be added to the exposure in the Induction Period.

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

6.14.2. 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 will be used as baseline. The treatment period will be included as postbaseline 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 postbaseline 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 posttreatment.

The planned summaries for adverse events are provided in Table KGAC.6.13, 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 Table KGAC.6.13 will be presented for the following periods/analysis populations as indicated. Summary tables will include the number and percentage of patients reporting an event. For events that are gender-specific (as defined by MedDRA), the number of participants at risk will include only patients from the given gender.

- Induction Period (Modified Safety Population, mS)
- Induction Period (Safety Population, S) as a sensitivity analysis
- Maintenance Blinded Period (Modified Maintenance Primary Population, mM)
- Maintenance Blinded Period (Maintenance Primary Population, M) as a sensitivity analysis
- Combined Induction Period and Maintenance Period, Combined Induction and Maintenance Periods Plus Follow-Up Period [selective analysis] (All Lebrikizumab Modified Safety Population, mA)
- Combined Induction Period and Maintenance Period, Combined Induction and Maintenance Periods Plus Follow-Up Period [selective analysis] (All Lebrikizumab Safety Population, A) as a sensitivity analysis

Table KGAC.6.13. Summary Tables/Listing Related to Adverse Events

Analysis	Population (Section 6.1.1)
Overview of AEs	mS, S, mM, M, mA, A
Summary of TEAE by PTs	mS, mM
Summary of TEAE by PTs occurring in ≥1% of patients	mS, mM
Summary of TEAE by PTs within SOC	mS, mM, mA
Summary of TEAE PTs by maximum severity	mS, S, mM, M
Summary of SAE by PT within SOC	mS, mM, mA
Summary of AEs leading to treatment discontinuation by PT with SOC	mS, mM, mA
Summary of TEAEs possibly related to study drug by PTs within SOC	mS, mM
Listing of SAEs (including Death)	ITT
Listing of primary AEs leading to study treatment discontinuation	ITT
Listing of TEAEs (for Japan submission only)	S
Listing of TEAEs occurred in safety population but not in modified safety population	

Abbreviations: A = All Lebrikizumab Safety Population; AE = adverse event; ITT = Intent-to-Treat; M = Maintenance Primary Population; mA = Modified All Lebrikizumab Safety Population; mM = Modified maintenance Primary Population, mS = Modified Safety Population; PT = Preferred Term; S = Safety Population; SAE = serious adverse event; SOC = System Organ Class; TEAE = treatment-emergent adverse event.

Statistical comparisons will be performed using Fisher's exact test. Odds ratio will be provided.

6.14.2.1. Common Adverse Events

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

6.14.2.2. Deaths, Other Serious Adverse Events and Other Notable Adverse Events

The number and percentage of patients 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 patients 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.

6.14.3. 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 Table KGAC.6.14. Hormone analytes are summarized/plotted similarly for adolescent patients.

Table KGAC.6.14. Analysis for Clinical Laboratory Evaluations

Analysis	Population
Box plots of observed values by visit Box plots for change values by visit	mS, mM
Change from baseline to last observations. ANCOVA model with treatment and baseline value in the model.	mS, mM
Scatter plots of baseline-by-maximum values and baseline-by-minimum values	mS, mM
Treatment-emergent abnormal high lab values (ie, patients shifting from a normal/low maximum baseline value to a high maximum postbaseline value) or abnormal low lab values (ie, patients shifting from normal/high minimum baseline value to a low minimum postbaseline value)	mS, mM, mA
Shift tables showing the number of patients who shift from each category of maximum (minimum) baseline observation to each category of maximum (minimum) postbaseline observation. Here categories may be low, normal, or high with cut-offs defined in the compound level safety standards.	mS, mM
Listing of abnormal findings for laboratory analyte measurements, including qualitative measures	All Enrolled

Abbreviations: ANCOVA = analysis of covariance; mA = Modified All Lebrikizumab Safety Population; mM = Modified maintenance Primary Population, mS = Modified Safety Population.

6.14.4. 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 (Table KGAC 6.15). 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.

Table KGAC 6.15. Analysis Related to Vital Signs

Analysis	Population
Box plots for observed values by visit	mS, mM
Box plots for change from baseline values by visit	mS, mM
Scatterplots of baseline-by-maximum values and baseline-by-minimum values	mS, mM
Tables with the number and percentage of subjects who shift from normal/high to low (ie, treatment-emergent low) and the number and percentage of subjects who shift from normal/low to high (ie, treatment-emergent high); the limits are defined in the compound level safety standards	mS, mM, mA

Abbreviations: mA = Modified All Lebrikizumab Safety Population; mM = Modified maintenance Primary Population, mS = Modified Safety Population.

6.14.4.1. 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 patients' 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, only All Lebrikizumab Safety Population will be described during 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].

The following summaries and plots will be provided:

Table KGAC.6.16. 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.	mA
Scatter plot of patients' mean weight, height, and BMI standardized percentile versus lebrikizumab exposure time	mA

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

6.14.5. 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 patient is considered TE-ADA positive when at least 1 postbaseline 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 Modified Safety Population during the Induction Period. The summary of the Modified Maintenance Primary Population will be provided for the Combined Induction and Maintenance Periods. For the Modified Maintenance Primary Population, the immunogenicity analysis will be cumulative across both the Induction and Maintenance Periods. Additional assessments of the relationship between immunogenicity and efficacy and TEAE by TE-ADA status will be performed as part of the integrated analysis including other Phase 3 lebrikizumab AD trials.

6.14.6. 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 Modified Safety Population and the All Lebrikizumab Modified Safety Population during their associated study periods as described in Section 6.1.4. Additional safety analysis may be added as needed.

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.

6.14.6.1. Hepatic Safety

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

Table KGAC.6.17. Summary Tables Related to Hepatic Safety

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

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

6.14.6.2. Eosinophilia and Eosinophil-Related Disorders

In addition to the standard laboratory analysis (Section 6.14.3), eosinophilia and eosinophil-related AE will be summarized. Details regarding eosinophil-related PTs are in Compound Level Safety Standard.

Table KGAC.6.18. Summary 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 postbaseline category by treatment	mS, mA
Summary of eosinophilia and eosinophil-related TEAE by PT	mS, mA

Abbreviations: AE = adverse event; PT = Preferred Term; mA = All Lebrikizumab Modified Safety Population; mS = Modified Safety Population; TEAE = treatment-emergent adverse event.

6.14.6.3. Infections, Including Herpes Infections and Relevant Parasitic Infections

Infections will be defined using the PTs from the MedDRA Infections and Infestations SOC. The MedDRA terms used to identify infections considered to be opportunistic infections (OI) in patients 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.

Table KGAC.6.19. Summary Tables/Listing Related to Infection Related AE

Analysis	Population
Summary of treatment-emergent infections by PT by maximum severity	mS, mA
Summary of serious infections by PT	mS, mA
Summary of infection AEs resulting in permanent study drug discontinuation	mS, mA
Treatment-emergent potential OI by PT nested with categories for narrow terms and broad terms separately	mS, mA
Treatment-emergent adverse events, herpes and parasitic infections	mS, mA
Treatment-emergent adverse events, skin infection	mS, mA
Summary and/or listing of Infection follow-up form	mS (summary only)
A listing of patients with potential OI, Serious Infection, herpes and parasitic infections	S

Abbreviations: AE = adverse event; OI = opportunistic infections; mA = All Lebrikizumab Modified Safety Population; mS = Modified Safety Population; PT = Preferred Term.

6.14.6.4. 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:

Table KGAC.6.20. Summary Tables/Listing Related to Conjunctivitis

Analysis	Population
Summary of TEAE of conjunctivitis within categories	mS, mA
Summary and/or listing of conjunctivitis and eye inflammation follow-up form	mS (summary only)
A listing of patients with conjunctivitis	S

Abbreviations: mA = All Lebrikizumab Modified Safety Population; mS = Modified Safety Population; TEAE = treatment-emergent adverse event.

6.14.6.5. Hypersensitivity

Potential hypersensitivity reactions will be determined using the following SMQs: anaphylactic reaction, hypersensitivity, and angioedema. Potential hypersensitivity will be categorized as immediate (ie, occurring the same day as drug administration) and non-immediate (ie, occurring after the day of study drug administration but prior to subsequent drug administration). The planned summaries are provided in Table KGAC.6.21.

Table KGAC.6.21. Summary Tables Related to Hypersensitivity

Analysis	Population
for immediate hypersensitivity: (1) combined narrow/algorithmic search (ie, any narrow term from any one of the SMQs, or anaphylaxis algorithm); (2) narrow search (ie, any narrow term) by SMQ; (3) broad search (ie, 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	mS, mA
for nonimmediate hypersensitivity: (1) combined narrow search (ie, any narrow term from any one of the SMQs); (2) narrow search (ie, any narrow term) by SMQ; and (3) broad search (ie, any narrow or broad term) by SMQ	mS, mA

Abbreviations: mA = All Lebrikizumab Modified Safety Population; MedDRA = Medical Dictionary for Regulatory Activities; mS = Modified Safety Population; PT = Preferred Term; SMQ = Standardised MedDRA Query; TEAE = treatment-emergent adverse events.

6.14.6.6. 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.

Table KGAC.6.22. Summary Tables Related to Injection Site Reactions

Analysis	Population
Summary of TEAE of ISR overall and by PT	mS, mA

Abbreviations: HLT = High Level Term; ISR = injection site reaction; mA = All Lebrikizumab Modified Safety Population; mS = Modified Safety Population; PT = Preferred Term; TEAE = treatment-emergent adverse event.

6.14.6.7. Malignancies

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

Table KGAC.6.23. Summary Tables Related to Malignancies

Analysis	Population
Summary of TEAE of malignancies within categories of NMSC and malignancy excluding NMSC	mS, mA

Abbreviations: mA = All Lebrikizumab Modified Safety Population; mS = Modified Safety Population; NMSC = nonmelanoma skin cancer; TEAE = treatment-emergent adverse event.

6.14.6.8. Atopic Dermatitis Exacerbation

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

Table KGAC.6.24. Summary Tables Related to Atopic Dermatitis Exacerbation

Analysis	Population
Summary of TEAE of atopic dermatitis exacerbation	mS, mA

Abbreviations: mA = All Lebrikizumab Modified Safety Population; mS = Modified Safety Population; TEAE = treatment-emergent adverse event.

6.14.6.9. Suicide/Self-Injury Standardised Medical Dictionary for Regulatory Activities Query

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

Table KGAC.6.25. Summary Tables Related to Suicide/Self-Injury Standardised Medical Dictionary for Regulatory Activities Query

Analysis	Population
Summary of TEAE of Suicide/self-injury SMQ	mS, mA

Abbreviations: mA = All Lebrikizumab Modified Safety Population; mS = Modified Safety Population; TEAE = treatment-emergent adverse event.

6.15. Subgroup Analyses

6.15.1. Efficacy Subgroup Analyses

Subgroup analyses will be conducted for the primary endpoints IGA 0/1, EASI-75, EASI-90, and 4-point improvement in Pruritus NRS at Week 16 in the mITT Population using MCMC-MI approach as in primary analysis (Section 6.4.1.1). 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 (eg, yes, no) is <10% of the total population, only descriptive statistics will be provided for that subgroup (ie, 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 to <18), Adults ≥18)
- Age group (Adolescents (12 to <18), Adults ≥18 to < 65, ≥65 to < 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)
- Ethnicity (Hispanic or Latino, Not Hispanic or Latino, Not reported, Unknown)
- Region (as defined in Section 6.5)
- 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 pruritus <4 versus ≥4
- Prior use of systemic treatment (yes, no)

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

6.15.2. Safety Subgroup Analyses

Subgroup analysis for safety related endpoints will be performed within the context of the integrated safety analysis. No safety subgroup analysis will be performed specifically for this study unless there is a potentially relevant finding during the periodic study safety reviews.

6.16. 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 patient's rights, safety, or well-being.

Potential examples of important protocol deviations include patients who violated the inclusion/exclusion criteria, used an interfering concomitant medication, significant non-compliance with study medication (<75% of expected injections). Refer to a separate document called "KGAC Trial Issues Management Plan" for the important protocol deviations with categorizations.

The number and percentage of patients having IPD(s) will be summarized within category and subcategory of deviation by treatment group for Induction Period using the ITT population and for Maintenance Period using all the Maintenance Populations (including Maintenance Primary Population, Maintenance Secondary Population and Maintenance W16 Escape Population).

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

6.16.1. Impact of COVID-19

Impact of pandemic (eg, COVID-19) on analyses may be addressed prior to study unblinding at Week 16 DBL, once the impact on study conducts are fully understood. In general, any missing assessments/visit window will be documented as protocol deviations. For patients 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 6.2. 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 6.4.

6.17. Interim Analyses and Data Monitoring

Data Monitoring Committee/Data Safety Monitoring Board (DMC/DSMB): The lebrikizumab Phase 3 AD programs' DSMB is an independent expert advisory group commissioned and charged with the responsibility of evaluating cumulative safety at regular intervals, as well as on an ad hoc basis, as needed. The DSMB will consist of members external to Lilly and follow the rules defined in the DSMB charter, focusing on potential and identified risks for this molecule. Data Monitoring Committee membership will include, at a minimum, a physician with expertise in dermatology and a statistician. No member of the DSMB may have contact with study sites. This committee will make recommendations as to a) continue the clinical studies without modification; or b) continue the clinical studies with modifications; or c) terminate one or more of the clinical studies. Details outlining the roles and responsibilities of the DMC are documented in the "Dermira DRM06 DSMB Program Charter" and the planned analyses are outlined in the DMC analysis plan prior to the first unblinded assessment.

Access to the unblinded safety data will be limited to the DSMB. The study team will not have access to the unblinded data. Only the DSMB is authorized to evaluate unblinded data. The purpose of the DSMB is to advise Lilly regarding patient safety; however, the DSMB may request key efficacy data to put safety observations into context and to confirm a reasonable benefit/risk profile for ongoing patients in the study. Hence, there will be no alpha adjustment for these interim assessments.

Week 16 Database lock (DBL): An unblinded interim analysis will be performed at the time (ie, a cut-off date) the last patient completes Week 16 or the ETV from the study. This database lock will include all data collected by the cut-off date. Only the Induction Period treatment assignment will be unblinded at the time of this interim lock. Maintenance Period treatment assignment will remain blinded.

Week 52 DBL: Another unblinded interim analysis will be performed at the time (ie, a cut-off date) the last patient completes Week 52 or the ETV from the study. This database lock will

include all data collected by the cut-off date and is the final analysis for the efficacy endpoints up to Week 52.

The study will not be terminated early on the basis of efficacy following these interim analyses.

Final DBL: A final DBL will occur after all patients have completed the safety follow-up period of the study, discontinued current study, or enrolled into the long-term extension study DRM06-AD07.

Depending on the regulatory submission timeline, the Week 52 DBL and the final DBL may be combined, ie, one final DBL will occur after all patients have either completed the follow-up period of the study discontinued the study early, or entered the long-term extension study DRM06-AD07.

6.18. 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.

6.19. 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 patients/subjects 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 eg, the CSR, manuscripts, and so forth.

7. Unblinding Plan

Unblinding details are specified in a separated unblinding plan.

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

Bretz F, Maurer W, Brannath W, et al. A graphical approach to sequentially rejective multiple test procedures. *Stat Med.* 2009;28(4):586–604. https://doi.org/10.1002/sim.3495

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. https://doi.org/10.1002/bimi.201000239

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

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):659-664. https://doi.org/10.1111/j.0022-202X.2005.23621.x

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. Accessed June 09, 2020. http://www.phusewiki.org/docs/CSS%20White%20Papers%202016/CSS_WhitePaper_CentralTendency_v1.0.pdf

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. Accessed June 09, 2020. http://www.phusewiki.org/docs/CSS%20White%20Papers%202016/CS_WhitePaper_OutliersShifts_v1.0.pdf

[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. Published 03 February

2017. Accessed June 09, 2020. https://www.phuse.eu/documents/working-groups/cs-whitepaper-adverseevents-v10-4442.pdf

Ratitch B. Combining analysis results from multiple imputed categorical data. PharmaSUG. 2013. Paper SP03.

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

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.

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

9. Appendices

Appendix 1. Study Visit Mapping for Pruritus NRS and Sleep-loss Diary and POEM

Pruritus NRS and sleep loss are collected as a daily diary; 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 Visit Date – 14)	Week 2 Visit Date – 8
Week 2	Week 2 Visit Date – 7	Week 2 Visit Date - 1
Week 4	Week 4 Visit Date – 7	Week 4 Visit Date – 1
Week 6	Week 6 Visit Date – 7	Week 6 Visit Date - 1
Week 8	Week 8 Visit Date – 7	Week 8 Visit Date - 1
Week 10	Week 10 Visit Date – 7	Week 10 Visit Date - 1
Week 12	Week 12 Visit Date – 7	Week 12 Visit Date – 1
Week 14	Week 14 Visit Date – 7	Week 14 Visit Date - 1
Week 16	Week 16 Visit Date – 7	Week 16 Visit Date – 1
Week 20	Week 20 Visit Date – 7	Week 20 Visit Date – 1
Week 24	Week 24 Visit Date – 7	Week 24 Visit Date – 1
Week 28	Week 28 Visit Date – 7	Week 28 Visit Date – 1
Week 32	Week 32 Visit Date – 7	Week 32 Visit Date – 1
Week 36	Week 36 Visit Date – 7	Week 36 Visit Date – 1
Week 40	Week 40 Visit Date – 7	Week 40 Visit Date – 1
Week 44	Week 44 Visit Date – 7	Week 44 Visit Date – 1
Week 48	Week 48 Visit Date – 7	Week 48 Visit Date – 1
Week 52	Week 52 Visit Date – 7	Week 52 Visit 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 Pruritus NRS and Sleep-loss could be found in Table KGAC.6.11. 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 are collected every week via eDiary, the visit week mapping will follow the following rule: the last collected POEM data before 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 patient gets an injection/assessment on the 14th, we would use the scale completed in between the 13th and the 7th.

Appendix 2. Details of Combining Estimates and Test Statistics for Categorial Endpoints with Multiple Imputation

Following the implementation of MCMC-MI imputation as specified in Section 6.4.1.1, 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 standard errors (SE) 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% confidence interval (CI) bounds output by PROC MIANALYZE are percents (ie, they are in terms of the response rate). To obtain the number of responders, the estimated percentage 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 odds ratio (OR), and Cochran-Mantel-Haenszel (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 odds ratio and CMH test statistic) specified in the TABLES statement. Each of these analyses are stratified by geographic region, 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: $SE_{lOR} = (\ln(UB_{OR}) - \ln(LB_{OR}))/(2 * 1.96)$. 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: $CMH_{WH} = \frac{(\frac{CMH}{df})^{\frac{1}{3}} - (1 - \frac{2}{9*df})}{\sqrt{\frac{2}{9*df}}}$, 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-side p-value is reported as the pooled p-value for the CMH test.

Appendix 3. Definition of Rescue Medications

This appendix provides the definition of rescue medications for this study, including topical and systemic treatments defined as follows:

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

Route of topical treatments includes: Topical and Transdermal.

Topical Corticosteroids (TCS): ATC code is D07

High Potency TCS: ATC codes: D07AC or D07AD

Low or moderate potency TCS: ATC code is D07, excluding D07AC or D07AD

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

Crisaborole: Preferred Term includes: CRISABOROLE

2. **Systemic Atopic Dermatitis Treatment (including systemic 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 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’ then medicals to manually review to confirm whether the medication in question is indeed ‘Phototherapy’ or ‘Photochemotherapy’

Signature Page for VV-CLIN-015694 v1.0

Approval	PPD an 19-Mar-2022 13:13:42 GMT+0000

Signature Page for VV-CLIN-015694 v1.0