Protocol C3291032

A PHASE 3, MULTICENTER, RANDOMIZED, DOUBLE-BLIND, VEHICLE-CONTROLLED STUDY OF THE EFFICACY AND SAFETY OF CRISABOROLE OINTMENT, 2% IN CHINESE AND JAPANESE PEDIATRIC AND ADULT SUBJECTS (AGES 2 YEARS AND OLDER) WITH MILD TO MODERATE ATOPIC DERMATITIS

Statistical Analysis Plan
(SAP)

Version: **3**

Date: **18-Sep-2021**

[redacted]

[redacted]

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

APPENDICES

1. VERSION HISTORY

Table 1. Summary of Changes

Version/ Date	Associated Protocol Amendment	Rationale	Specific Changes
1 20 Oct 2017	Original 17 Oct 2017	N/A	N/A
2 13 Apr 2020	Protocol Amendment 1 15 Nov 2019	Issuance of a protocol amendment, template change, TLF bold moves	Update endpoints to align with global clinical development program, including: • Update the primary and key secondary estimands. • Change primary endpoint from Investigator's Static Global Assessment (ISGA) to Eczema Area and Severity Index (EASI). • Addition of ISGA as a key secondary endpoint. • Update the analysis for Patient Reported Outcomes (PROs). • Update the analysis for Dermatology Life Quality Index (DLQI), Children's Dermatology Life Quality Index (CDLQI), Infant's Dermatitis Quality of Life Index (IDQOL), Dermatitis Family Impact Questionnaire (DFI), Patient Oriented Eczema Measure (POEM), Patient Global Impression of Change (PGIC)/ Observer Reported Global Impression of Change (OGIC). • [redacted]
3 18 Sep 2021	Protocol Amendment 3 18 Dec 2020	To align with the current protocol, specify PPAS criteria, provide further clarification on	• Updated Section 2 to align with the current protocol. • Updated the detailed list of PPAS criteria.

Table 1. Summary of Changes

Version/ Date	Associated Protocol Amendment	Rationale	Specific Changes
		analysis and reporting details, add additional analyses to address COVID-19 pandemic impacts.	• Added detailed descriptive summary statistics for binary and continuous variables. • Deleted supportive analysis for Section 6.1.1 and Section 6.2.3. • Missing data handlings in Section 6.2.4, Section 6.2.5 and Section 6.2.8 are updated to align with the analysis of primary and key secondary estimands. • Updated the analysis set in Section 6.2.2, Section 6.2.4, Section 6.3.1, Section 6.3.2 and Section 6.3.3 to exclude participants who do not have the chance to achieve the event. • Added additional analyses to address COVID-19 pandemic impacts in Section 6.7. • Added by country safety summaries in Section 6.4.

2. INTRODUCTION

Crisaborole, also referred as PF 06930164 and AN2728, is a low molecular weight benzoxaborole anti inflammatory phosphodiesterase 4 (PDE 4) inhibitor that penetrates into the skin to the sites of inflammation. The primary mechanism of the anti-inflammatory effect of crisaborole is through inhibition of PDE 4. Crisaborole has demonstrated in vitro inhibition of a range of cytokines implicated in the pathogenesis of atopic dermatitis (AD). Crisaborole inhibits the release of chemokines that are also important inflammatory mediators. Crisaborole applied to human skin ex vivo or on AD lesions on a subject reduces expression of key drivers of atopic inflammation including T-cell derived cytokines IL-13, IL-31, and interferon gamma (IFNγ) as well as innate markers of inflammation such as matrix metalloproteinase (MMP)12.

Supporting evidence of the safety and efficacy of this product in patients 2 years and older represent a major advancement in the treatment of AD given the challenges of managing this

common, chronic dermatologic condition and the treatment limiting effects of currently available therapies. All primary and secondary efficacy endpoints were statistically significant in the two Phase 3 registration studies. Across the development program, Crisaborole demonstrated an acceptable safety profile, with no crisaborole treatment related Serious Adverse Events (SAEs) and with the majority of Adverse Events (AEs) being mild and deemed unlikely or not related to investigational product.

In this China and Japan study, a similar study design to the two global pivotal studies (AN2728-AD-301 and AN2728-AD-302) will be used to investigate the efficacy and safety of crisaborole ointment, 2% twice daily (BID) in Chinese and Japanese pediatric and adult subjects (ages 2 years and older) with mild to moderate AD, to support the registration of crisaborole in China and Japan.

This SAP provides the detailed methodology for summary and statistical analyses of the data collected in study C3291032. This document may modify the plans outlined in the protocol; however, any major modifications of the primary endpoint definition or its analysis will also be reflected in a protocol amendment.

2.1. Study Objectives, Endpoints, and Estimands

Estimands will be defined for the primary and key secondary efficacy objectives.

Objectives	Endpoints	Estimands
Primary:	**Primary:**	**Primary:**
• To evaluate the efficacy of crisaborole ointment, 2% applied BID versus Vehicle in Chinese and Japanese pediatric and adult subjects (ages 2 years and older) with mild to moderate AD.	• Percent Change from Baseline in EASI total score at Day 29.	The estimand is the hypothetical estimand, which estimates the effect if all participants maintain their randomized treatment and adhere to the protocol. It includes the following 5 attributes: • Treatment: Crisaborole Ointment 2% BID vs Vehicle Ointment BID. • Population: Chinese and Japanese pediatric and adult subjects (ages 2 years and older) with mild to moderate AD as defined by the inclusion and exclusion criteria. • Variable: % change from baseline in EASI total score at Day 29.

Objectives	Endpoints	Estimands
		• Intercurrent events: Had patient not discontinued from treatment prior to Day 29. • Population-level summary: Difference in means between crisaborole ointment 2% BID vs corresponding vehicle.
• To evaluate the safety and tolerability of crisaborole ointment, 2% applied BID versus Vehicle in Chinese and Japanese pediatric and adult subjects (ages 2 years and older) with mild to moderate AD.	• Treatment emergent adverse events (TEAEs) (including application site reactions) & SAEs, and clinically significant changes in vital signs and clinical laboratory parameters.	No estimands will be defined.
Key Secondary:	**Key Secondary:**	**Key Secondary:**
• To evaluate the effect of crisaborole ointment, 2% applied BID versus Vehicle on additional efficacy endpoints in Chinese and Japanese pediatric and adult subjects (Ages 2 years and older) with mild to moderate AD.	• Achievement of Improvement in ISGA (defined as ISGA score of clear (0) or almost clear (1)) at Day 29.	The estimand is the hypothetical estimand, which estimates the effect if all participants maintain their randomized treatment and adhere to the protocol. It includes the following 5 attributes • Treatment: Crisaborole Ointment 2% BID vs Vehicle Ointment BID. • Population: Chinese and Japanese pediatric and adult subjects (ages 2 years and older) with mild to moderate AD as defined by the inclusion and exclusion criteria • Variable: Improvement in ISGA at Day 29. • Intercurrent events: Had patient not discontinued from treatment prior to Day 29.

Objectives	Endpoints	Estimands
		• Population-level summary: Difference in percentage of subjects with Improvement in ISGA between crisaborole ointment 2% BID vs corresponding vehicle.
Secondary:	**Secondary:**	**Secondary:**
• To evaluate the effect of crisaborole ointment, 2% applied BID versus Vehicle on additional efficacy endpoints in Chinese and Japanese pediatric and adult subjects (Ages 2 years and older) with mild to moderate AD.	• Achievement of Success in ISGA (defined as an ISGA score of Clear (0) or Almost Clear (1) with at least a 2 grade improvement from Baseline) at Day 29.	The estimand is the hypothetical estimand, which estimates the effect if all participants maintain their randomized treatment and adhere to the protocol. It includes the following 5 attributes • Treatment: Crisaborole Ointment 2% BID vs Vehicle Ointment BID. • Population: Chinese and Japanese pediatric and adult subjects (ages 2 years and older) with mild to moderate AD as defined by the inclusion and exclusion criteria. • Variable: Success in ISGA at Day 29. • Intercurrent events: Had patient not discontinued from treatment prior to Day 29. • Population-level summary: Difference in percentage of subjects with Success in ISGA between crisaborole ointment 2% BID vs corresponding vehicle.
• To evaluate the effect of crisaborole ointment, 2% applied BID versus Vehicle on patient/observer reported outcomes over time in Chinese and	• Change from Baseline in Peak Pruritus Numeric Rating Scale (NRS) at Week 4 - for subjects ≥12 years.	The estimand is the hypothetical estimand, which estimates the effect if all participants maintain their randomized treatment and adhere to the protocol. It includes the following 5 attributes:

Objectives	Endpoints	Estimands
Japanese pediatric and adult subjects (Ages 2 years and older) with mild to moderate AD.		• Treatment: Crisaborole Ointment 2% BID vs Vehicle Ointment BID. • Population: Chinese and Japanese pediatric and adult subjects (ages 2 years and older) with mild to moderate AD as defined by the inclusion and exclusion criteria. • Variable: Change from Baseline in Peak Pruritus NRS at Week 4 for subjects ≥12 years. • Intercurrent events: Had patient not discontinued from treatment prior to Day 29. • Population-level summary: Difference in means between crisaborole ointment 2% BID vs corresponding vehicle.
Other Secondary:	**Other Secondary:**	**Other Secondary:**
• To evaluate the effect of crisaborole ointment, 2% applied BID versus Vehicle on additional efficacy endpoints in Chinese and Japanese pediatric and adult subjects (Ages 2 years and older) with mild to moderate AD. • To evaluate the effect of crisaborole ointment, 2% applied BID versus Vehicle on patient/observer reported outcomes over time in	• Success in ISGA (defined as an ISGA score of Clear (0) or Almost Clear (1) with at least a 2 grade improvement from Baseline) over time. • Improvement in ISGA (defined as ISGA score of clear (0) or almost clear (1)) over time. • Percent change from Baseline in EASI total score over time. • Change from Baseline in % body surface area (BSA) over time. • Achievement of EASI-50 (≥50% improvement from baseline) over time. • Achievement of EASI-75 (≥75% improvement from baseline) over time. • Change from baseline in Peak Pruritus NRS over time - for subjects≥12 years.	No estimands will be defined.

Objectives	Endpoints	Estimands
Chinese and Japanese pediatric and adult subjects (Ages 2 years and older) with mild to moderate AD.	• Changes from baseline in Patient Reported Itch Severity Scale over time - for subjects ≥6 years and <12 years. • Change from baseline in Observer Reported Itch Severity Scale over time - for subjects <6 years. • DLQI, CDLQI, IDQOL, DFI, POEM, PGIS/OGIS, PGIC/OGIC over time.	
CCI		
CCI		

Objectives	Endpoints	Estimands
	[redacted]	

2.2. Study Design

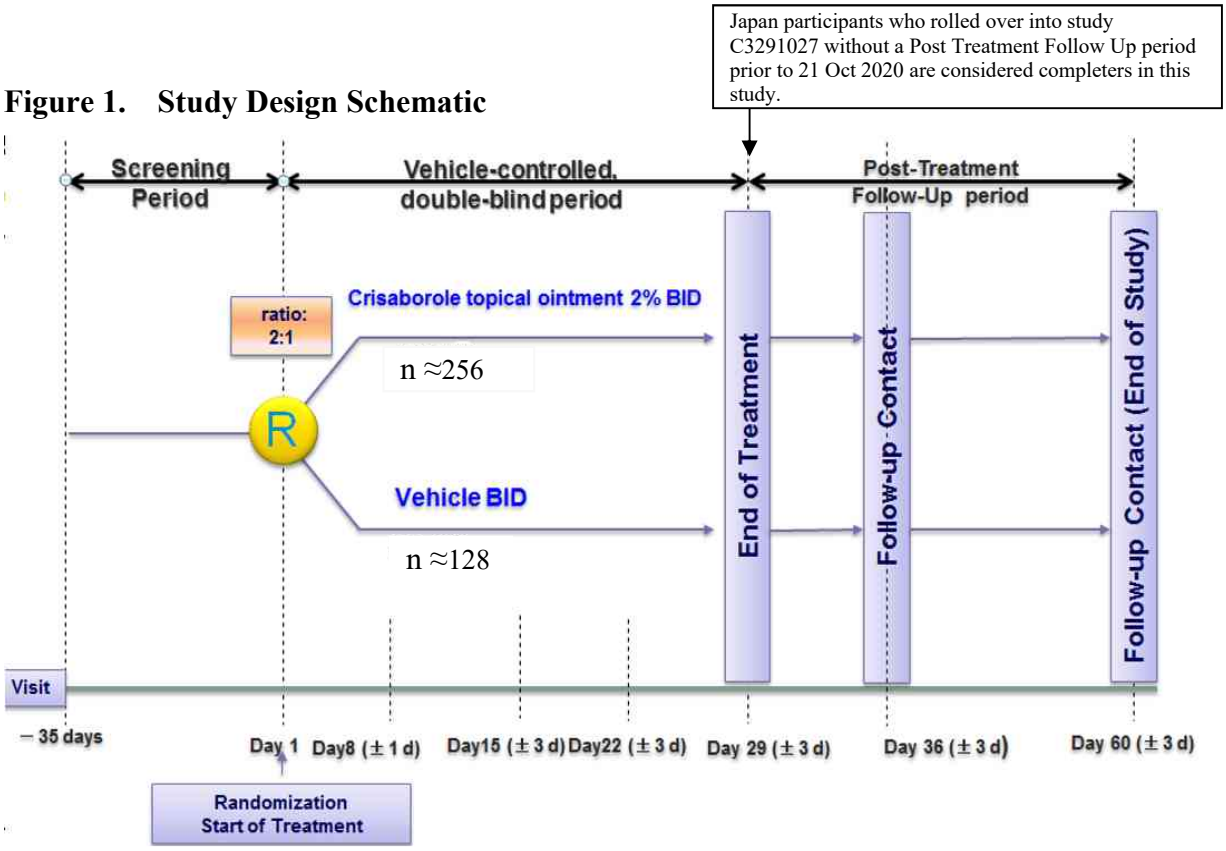
This is a Phase 3, multicenter, randomized, double blind, vehicle controlled study to evaluate the efficacy and safety of crisaborole ointment 2% in Chinese and Japanese pediatric and adult subjects (ages 2 years and older) with mild to moderate atopic dermatitis involving at least 5% treatable BSA.

A total of approximately 384 subjects (approximately 50% for age ≥12 years old and approximately 50% for age <12 years old) will be enrolled in the study from multiple sites in China and Japan. Following the screening period (up to 35 days prior to Baseline/Day 1), eligible subjects will be randomized at the Baseline/Day 1 visit in a 2:1 ratio to one of 2 treatment groups (crisaborole ointment, 2% BID; vehicle BID, respectively), the investigational product will be applied BID for 28 days to the Treatable BSA identified at Baseline/Day 1 and new AD lesions that appear after the Baseline/Day 1. The primary efficacy endpoint, percent change from baseline in EASI total score, will be assessed at Day 29.

Scheduled study visits for all subjects will occur at Screening, Baseline/Day 1, Day 8, Day 15, Day 22, Day 29 (End of treatment/Early termination). A follow up telephone call will be made by site staff to the subjects and/or parents/legal guardians on Day 36 and Day 60.

Japan participants who rolled over into study C3291027 without a Post Treatment Follow Up period prior to 21 Oct 2020 are considered completers in this study. A schematic of the study design is shown in Figure 1.

Figure 1. Study Design Schematic



3. ENDPOINTS AND BASELINE VARIABLES: DEFINITIONS AND CONVENTIONS

3.1. Primary Endpoints

Efficacy endpoint: Percent Change from Baseline in EASI total score at Day 29.

Safety endpoint: TEAEs (including application site reactions) & SAEs and clinically significant changes in vital signs and clinical laboratory parameters.

3.2. Secondary Endpoints

Efficacy endpoints:

- Key secondary efficacy endpoints
    - Achievement of Improvement in ISGA (defined as ISGA score of clear (0) or almost clear (1)) at Day 29.

- Achievement of Success in ISGA (defined as an ISGA score of Clear (0) or Almost Clear (1) with at least a 2 grade improvement from Baseline) at Day 29.
- Change from Baseline in Peak Pruritus Numeric Rating Scale (NRS) at Week 4 - for subjects ≥12 years.

- Other secondary efficacy endpoints
  - Success in ISGA (defined as an ISGA score of Clear (0) or Almost Clear (1) with at least a 2 grade improvement from Baseline) over time.
  - Improvement in ISGA (defined as ISGA score of clear (0) or almost clear (1)) over time.
  - Percent change from Baseline in EASI total score over time.
  - Change from Baseline in %BSA over time.
  - Achievement of EASI-50 (≥50% improvement from baseline) over time.
  - Achievement of EASI-75 (≥75% improvement from baseline) over time.
  - Change from baseline in Peak Pruritus NRS over time - for subjects ≥12 years.
  - Changes from baseline in Patient Reported Itch Severity Scale over time - for subjects ≥6 years and <12 years.
  - Change from baseline in Observer Reported Itch Severity Scale over time - for subjects <6 years.
  - DLQI, CDLQI, IDQOL, DFI, POEM, PGIS/OGIS, PGIC/OGIC over time.

CCI [redacted]

3.4. Baseline Variables

For Peak Pruritus NRS, Patient/Observer Reported Itch Severity Scale and PGIS/OGIS, the average of the latest available 7-day scores immediately prior to Day 1 (including Day 1) will be used as the baseline. The baseline is considered missing if the daily scores are only available for less than four days. If the daily scores are available for more than three days, then the baseline is calculated as the average of the available scores.

Day 1 is defined as the first dosing date.

The baseline values of all other endpoints are defined as the last values collected prior to the first dose of study drug.

Baseline variables include

- Demographics;
- Height and weight;
- ISGA score;
- EASI total score;
- Peak Pruritus NRS, Patient/Observer Reported Itch Severity Scale;
- Treatable %BSA;
- DLQI/CDLQI/IDQOL;
- DFI;

- POEM;
- PGIS/OGIS;
- [redacted]

These data will be summarized as part of the demographic characteristics and baseline characteristics.

3.5. Safety Endpoints

Safety will be assessed by the spontaneous reporting of AEs, SAEs, physical examinations, vital signs, and clinical laboratory results in all subjects who receive at least one dose of the investigational product. Unscheduled safety assessments may be performed at any time during the study to assess any perceived safety concerns.

Safety data will be descriptively summarized and will be presented in tabular. No imputation will be made for missing safety data. The following safety data will be summarized:

- TEAEs, including SAEs;
- Clinically significant changes in vital signs;
- Clinically significant changes in laboratory parameters.

3.5.1. Adverse Events

An adverse event will be considered as a TEAE if the event started after the first dose of treatment regardless of whether a similar event of equal or greater severity existed in the baseline period.

Adverse events will be assessed by the spontaneous reporting of:

- Incidence of TEAEs;
- Incidence of SAEs;
- Incidence of AEs leading to discontinuation.

3.5.2. Laboratory Data

Below is a list of clinical laboratory test parameters.

- Hematology: hemoglobin, hematocrit, red blood cell count, platelet count, white blood cell count (% and absolute), neutrophils, eosinophils, monocytes, basophils and lymphocytes.
- Chemistry: blood urea nitrogen/Urea, glucose, creatinine, sodium, potassium, chloride, Bicarbonate or Total CO_2, alanine aminotransferase, aspartate aminotransferase, total bilirubin, alkaline phosphatase, albumin, and total protein.

3.5.3. Physical Examinations

A physical examination includes, but not limited to the following organ or body systems; head, ears, eyes, nose, mouth, skin, heart and lung examinations, lymph nodes, musculoskeletal, abdomen (liver, spleen), and neurological systems. In addition, an assessment will be made of the condition of all AD-involved skin.

3.5.4. Vital Signs

Vital sign measurements are temperature, respiratory rate, pulse rate, and blood pressures.

4. ANALYSIS SETS

Population	Description
Full Analysis Set (FAS)	All subjects randomized and dispensed study drug, subjects are assigned to the randomized treatment regardless of what treatment was received.
Per Protocol Analysis Set (PPAS)	All subjects in the FAS who complete the Day 29 evaluation without any major protocol deviations.
Safety Analysis Set	All subjects who are randomized and received at least one confirmed dose of investigational product. The subject will be assigned to the randomized treatment group if he received at least one dose of the randomized treatment. The subject will be assigned to the other treatment group if he does not receive any of the randomized treatment.

Specifically, PPAS will include subjects in the FAS who meet all of the following criteria:

- Met all of the Inclusion Criteria and none of the Exclusion Criteria;
- Have not taken any prohibited concomitant medications during the treatment period;
- Completed the Day 29 Visit, including the Day 29 efficacy evaluation of EASI and ISGA, subjects ≥12 years who have missing Week 4 Peak Pruritus NRS evaluation will be excluded;
- Have applied 80%–120% of the total number of expected doses (56 doses) during the treatment period;
- Have not missed 6 or more consecutive doses during the treatment period;
- Were in the visit window (±3 days) for the Day 29 Visit.

Subjects who prematurely discontinue from the study due to lack of efficacy or a treatment related TEAE will be placed back into PPAS.

Data for all participants will be assessed to determine if participants meet the criteria for inclusion in PPAS prior to unblinding and releasing the database and classifications will be documented per standard operating procedures.

5. GENERAL METHODOLOGY AND CONVENTIONS

The final analysis will be performed based on the final released data set after last subject last visit (LSLV).

5.1. Hypotheses and Decision Rules

The study will be declared a success with crisaborole shown to be superior to vehicle with respect to the primary efficacy endpoint, percent change from baseline in EASI total score at Day 29, if the mean percent change in the Crisaborole arm is lower than that in the vehicle arm and the difference is statistically significant at the two-sided level of 0.05.

After crisaborole is shown to be superior to vehicle with respect to the primary efficacy endpoint, it will be compared with vehicle with respect to the key secondary efficacy endpoints using gate-keeping method.

After and only after crisaborole is shown to be superior to vehicle with respect to the percent change from baseline in EASI total score at Day 29, crisaborole will be declared superior to vehicle with respect to the improvement rate in ISGA at Day 29, i.e., the percentage of patients with ISGA score of Clear (0) or Almost Clear (1), at Day 29, if the rate in the crisaborole arm is higher than that in the vehicle arm and the difference is statistically significant at the two-sided level of 0.05.

After and only after crisaborole is shown to be superior to vehicle with respect to the improvement rate in ISGA at Day 29, crisaborole will be declared superior to vehicle with respect to the success rate in ISGA at Day 29, i. e, the percentage of patients with ISGA score of Clear (0) or Almost Clear (1) with at least a 2 grade improvement from Baseline, at Day 29, if the rate in the crisaborole arms is higher than that in the vehicle arm and the difference is statistically significant at the two-sided level of 0.05.

After and only after crisaborole is shown to be superior to vehicle with respect to the success rate in ISGA at Day 29, crisaborole will be declared superior to vehicle with respect to the change from baseline in Peak Pruritus NRS at Week 4 - for subjects ≥12 years, if the mean change of the crisaborole arm is lower (i.e., larger reduction) than that in the vehicle arm and the difference is statistically significant at the two-sided level of 0.05.

5.2. General Methods

5.2.1. Analyses for Binary Endpoints

The binary endpoints will be descriptively summarized using sample size, number, percentage and 95% confidence interval (CI) of percentage at each time point.

Binary data will be analyzed by comparing response rates between treatment groups. Normal approximation to the difference in response rates will be used to obtain p-values and 95% CI.

5.2.2. Analyses for Continuous Endpoints

Continuous endpoints will be descriptively summarized using sample size, mean, standard deviation, median and range.

For the analysis in which missing values are not explicitly imputed, continuous endpoints will be analyzed using a linear Mixed effect Model for Repeated Measures (MMRM, Mallinckrodt et al. 2001) with treatment group, visit, and treatment group-by-visit interaction as factors and baseline value as a covariate. Within-subject variability will be accounted for using a random effect with the first-order autoregressive (AR(1)) covariance matrix if there are no convergence issues, otherwise other structures will be considered.

5.2.3. Analyses for Time-to-Event Endpoints

Time-to-event analysis will be used to analyze time-to-event data.

Kaplan-Meier (product limit) method will be used for estimation of proportion of participants with event, time-to-event curve, median (95% CI) time-to-event.

5.3. Methods to Manage Missing Data

Multiple Imputation based on Markov Chain Monte Carlo (MI-MCMC)

Missing ISGA scores will be derived for the analysis using the method of MCMC multiple imputation. Multiple imputation and subsequent analysis will involve the following four principal tasks:

1. Calculate the number of missing values to be imputed by MCMC (nmiss) for the Day 29 value.
2. Create a data set, one for each treatment group, of subjects with observed values and those needing imputation by MCMC. The missing ISGA values in each data set will be filled in using the MCMC method "10× nmiss" times to generate "10 × nmiss" data sets. The resulting data sets for each treatment group will be combined into one complete data set by imputation. The imputed values will be rounded to the nearest value of 0, 1, 2, 3, or 4.
3. For each complete data set, derive dichotomous variables Success in ISGA (Clear [0] or Almost Clear [1] with a 2-point change from Baseline) and Improvement in ISGA (Clear [0] or Almost Clear [1]). The estimated treatment effect and its standard error for each complete data set will be calculated with the method specified in Section 5.2.1.
4. Combine the estimated treatment effects and standard errors from the above into a single inference using Rubin's formulae as implemented in SAS PROC MIANALYZE.

Success in ISGA and Improvement in ISGA will be derived from the imputed complete ISGA data sets.

For EASI-50 and EASI-75, missing EASI total scores will be imputed in the same way as missing ISGA scores except no rounding will be done to the imputed values. Missing EASI-50 and EASI-75 values will be derived from the imputed complete EASI data sets.

Non-Responder Imputation (NRI) Non-Responder Imputation (NRI) as another alternative way of handling missing data will also be used in Supportive analyses. If a subject has missing data for a binary endpoint at a scheduled visit, this subject will be defined as a non-responder for that endpoint at that visit.

Repeated measure modeling with no explicit imputation For continuous endpoints, missing values are assumed missing at random (MAR) and will be handled by MMRM.

In general, for descriptive statistics and time to event analysis, missing values will not be imputed.

6. ANALYSES AND SUMMARIES

6.1. Primary Endpoints

6.1.1. Percent Change from Baseline in EASI total score at Day 29

6.1.1.1. Main Analysis

- Estimand strategy: Hypothetical.
- Analysis set: FAS.
- Analysis methodology: percent change from baseline in EASI total score at Day 29, will be analyzed using MMRM that includes treatment group, visit, and treatment group-by-visit interaction as factors and baseline value as a covariate. Within-subject variability will be accounted for using a random effect with the first-order autoregressive (AR(1)) covariance matrix.
- Intercurrent events and missing data: data after treatment discontinuation prior to Day 29 will not be considered. Missing values will not be explicitly imputed and will be assumed to be MAR.
- Reporting results:
  - Tables: The sample size, least-squares (LS) mean of percent change from baseline and 95% CI for each treatment, LS mean of difference and the corresponding 95% CI, and p-value will be presented at Day 29.

6.1.1.2. Supplementary Analyses

An analysis using PPAS will be performed. It will use the same methodology and handling of intercurrent events as the main analysis.

6.1.2. TEAEs (including application site reactions) & SAEs and clinically significant changes in vital signs and clinical laboratory parameters

All safety analyses will be conducted on the Safety Analysis Set.

Pfizer standard will be utilized in reporting routine safety variables (Pfizer Safety Rulebook).

6.2. Secondary Endpoints

6.2.1. Achievement of Improvement in ISGA at Day 29

6.2.1.1. Main Analysis

- Estimand strategy: Hypothetical.
- Analysis set: FAS.
- Analysis methodology: The percentage of subjects achieving Improvement in ISGA at Day 29 will be compared between crisaborole arm and vehicle arm and the difference will be tested based on normal approximation to response rates.
- Intercurrent events and missing data: Data after treatment discontinuation prior to Day 29 will not be considered. Missing ISGA scores will be derived for the analysis using MI-MCMC.
- Reporting results:
    - Tables: The sample size, number, percentage and 95% CI of percentage of subjects achieving Improvement in ISGA at Day 29 will be presented for each treatment. Risk difference and the corresponding 95% CI and p-value for the test of difference based on normal approximation to response rates will also be presented at Day 29.

6.2.1.2. Supportive/Supplementary Analysis

Supportive Analyses

Missing Improvement in ISGA (binary variable) will be handled by using NRI in supportive analyses. Reporting results are the same as tables above for the main analysis.

Supplementary Analysis

An analysis using PPAS will be performed. It will use the same methodology and handling of intercurrent events as the main analysis.

6.2.2. Achievement of Success in ISGA at Day 29

6.2.2.1. Main Analysis

- Estimand strategy: Hypothetical.
- Analysis set: FAS. Participants with baseline ISGA score of Clear (0) or Almost Clear (1) will be excluded.
- Analysis methodology: The percentage of subjects achieving Success in ISGA at Day 29 will be compared between crisaborole arm and vehicle arm and the difference will be tested based on normal approximation to response rates.
- Intercurrent events and missing data: Data after treatment discontinuation prior to Day 29 will not be considered. Missing ISGA scores will be derived for the analysis using MI-MCMC.
- Reporting results:
    - Tables: The sample size, number, percentage and 95% CI of percentage of subjects achieving Success in ISGA at Day 29 will be presented for each treatment. Risk difference and the corresponding 95% CI and p-value for the test of difference based on normal approximation to response rates will also be presented at Day 29.

6.2.2.2. Supportive/Supplementary Analysis

Supportive Analyses

Missing Success in ISGA (binary variable) will be handled by using NRI in supportive analyses. Reporting results are the same as tables above for the main analysis.

Supplementary Analysis

An analysis using PPAS will be performed. It will use the same methodology and handling of intercurrent events as the main analysis.

6.2.3. Change from Baseline in Peak Pruritus NRS at Week 4 – for subjects ≥12 years

6.2.3.1. Main Analysis

- Estimand strategy: Hypothetical.
- Analysis set: FAS.
- Analysis methodology: change from baseline to Week 4 (Days 23-29) in weekly average of Peak Pruritus NRS for subjects ≥12 years, will be analyzed using a MMRM that includes treatment group, visit, and treatment group-by-visit interaction as factors and

baseline value as a covariate. Within-subject variability will be accounted for using a random effect with the first-order autoregressive (AR(1)) covariance matrix.

- Intercurrent events and missing data: Data after treatment discontinuation prior to Day 29 will not be considered. The weekly average is considered missing if Peak Pruritus NRS scores are missing for more than three days. If the scores are missing for less than four days, then the weekly average is computed as the mean of the scores for the non-missing days. Missing weekly average values will not be explicitly imputed and will be assumed to be MAR.

- Reporting results:

    - Tables: The sample size, LS mean of change from baseline and 95% CI for each treatment, LS mean of difference and the corresponding 95% CI, and p-value will be presented at Week 4.

6.2.3.2. Supplementary Analyses

An analysis using PPAS will be performed. It will use the same methodology and handling of intercurrent events as the main analysis.

6.2.4. Success in ISGA and Improvement in ISGA over time

- Analysis set: FAS. Participants with baseline ISGA score of Clear (0) or Almost Clear (1) will be excluded for Success in ISGA.

- Analysis methodology: At Day 8, Day 15, Day 22 and Day 29, the percentage of subjects achieving Success or Improvement in ISGA will be compared between crisaborole arm and vehicle arm and the difference will be tested based on normal approximation to response rates.

- Missing data: Data after treatment discontinuation prior to Day 29 will not be considered. Missing data handling will be the same as in Section 6.2.1.1.

- Reporting results:

    - Tables: The sample size, number, percentage and 95% CI of percentage of subjects achieving Success or Improvement in ISGA at Day 8, Day 15, Day 22 and Day 29 will be presented for each treatment. Risk difference and the corresponding 95% CI and p-values for the test of difference based on normal approximation to response rates will be presented.

    - Figures: For two treatments, line plots of the percentage of subjects achieving Success or Improvement in ISGA and 95% CIs at each post-baseline visit will be displayed graphically

6.2.5. Percent Change from baseline in EASI total score over time

- Analysis set: FAS.

- Analysis methodology: At Day 8, Day 15, Day 22 and Day 29, percent change from baseline in EASI total score will be analyzed using a MMRM that includes treatment group, visit, and treatment group-by-visit interaction as factors and baseline value as a covariate. Within-subject variability will be accounted for using a random effect with the first-order autoregressive (AR(1)) covariance matrix.

- Missing data: data after treatment discontinuation prior to Day 29 will not be considered. Missing data handling will be the same as in Section 6.1.1.1.

- Reporting results:

  - Tables: The sample size, LS mean of percent change from baseline and 95% CI for each treatment, LS mean of difference and the corresponding 95% CI, and p-value will be presented for all post-baseline visits.

  - Figures: For two treatments, line plots of the LS means and 95% CIs at each post-baseline visit will be displayed graphically.

6.2.6. Change from baseline in %BSA over time

- Analysis set: FAS.

- Analysis methodology: At Day 8, Day 15, Day 22 and Day 29, change from baseline in %BSA with AD will be analyzed using a MMRM that includes treatment group, visit, and treatment group-by-visit interaction as factors and baseline value as a covariate. Within-subject variability will be accounted for using a random effect with the first-order autoregressive (AR(1)) covariance matrix.

- Missing values will not be explicitly imputed and will be assumed to be MAR.

- Reporting results:

  - Tables: The sample size, LS mean of change from baseline and 95% CI for each treatment, LS mean of difference and the corresponding 95% CI, and p-value will be presented for all post-baseline visits.

6.2.7. Achivevment of EASI-50, EASI-75 over time

- Analysis set: FAS.

- Analysis methodology: At Day 8, Day 15, Day 22 and Day 29, the percentage of subjects achieving EASI-50 or EASI-75 will be compared between crisaborole arm and vehicle arm and the difference will be tested based on normal approximation to response rates.

- Missing data: Missing EASI scores will be derived for the analysis using MI-MCMC.

- Reporting results:

  - Tables: The sample size, number, percentage and 95% CI of percentage of subjects achieving EASI-50 or EASI-75 at Day 8, Day 15, Day 22 and Day 29 will be presented for each treatment. Risk difference and the corresponding 95% CI and p-values for the test of difference based on normal approximation to response rates will be presented.

6.2.8. Change from Baseline in Peak Pruritus NRS over time - for subjects≥12 years

- Analysis set: FAS.

- Analysis methodology: change from baseline in weekly average of Peak Pruritus NRS at Week 1 (Days 2-8), Week 2 (Days 9-15), Week 3 (Days 16-22) and Week 4 (Days 23-29) for subjects ≥12 years, will be analyzed using a MMRM that includes treatment group, visit, and treatment group-by-visit interaction as factors and baseline value as a covariate. Within-subject variability will be accounted for using a random effect with the first-order autoregressive (AR(1)) covariance matrix.

- Missing data: Data after treatment discontinuation prior to Day 29 will not be considered. Missing data handling will be the same as in Section 6.2.3.1.

- Reporting results:

  - Tables: The sample size, LS mean of change from baseline and 95% CI for each treatment, LS mean of difference and the corresponding 95% CI, and p-value will be presented at Week 1, Week 2, Week 3 and Week 4.

  - Figures: For two treatments, line plots of the LS means and 95% CIs at each week will be displayed graphically.

6.2.9. Change from Baseline in Patient Reported Itch Severity Scale over time - for subjects ≥6 years and <12 years

- Analysis set: FAS.

- Analysis methodology: change from baseline in weekly average of Patient Reported Itch Severity Scale at Week 1, Week 2, Week 3 and Week 4, for subjects ≥6 years and <12 years, will be analyzed using a MMRM that includes treatment group, visit, and treatment group-by-visit interaction as factors and baseline value as a covariate. Within-subject variability will be accounted for using a random effect with the first-order autoregressive (AR(1)) covariance matrix.

- Missing data: The weekly average is considered missing if Patient Reported Itch Severity Scales are missing for more than three days. If the scores are missing for less than four

days, then the weekly average is computed as the mean of the scores for the non-missing days. Missing weekly average values will not be explicitly imputed and will be assumed to be MAR.

- Reporting results:
    - Tables: The sample size, LS mean of change from baseline and 95% CI for each treatment, LS mean of difference and the corresponding 95% CI, and p-value will be presented for each week.

6.2.10. Change from Baseline in Observer Reported Itch Severity Scale over time - for subjects <6 years

- Analysis set: FAS.

- Analysis methodology: change from baseline in weekly average of Observer Reported Itch Severity Scale at Week 1, Week 2, Week 3 and Week 4, for subjects<6 years, will be analyzed using a MMRM that includes treatment group, visit, and treatment group-by-visit interaction as factors and baseline value as a covariate. Within-subject variability will be accounted for using a random effect with the first-order autoregressive (AR(1)) covariance matrix.

- Missing data: The weekly average is considered missing if Observer Reported Itch Severity Scales are missing for more than three days. If the scores are missing for less than four days, then the weekly average is computed as the mean of the scores for the non-missing days. Missing weekly average values will not be explicitly imputed and will be assumed to be MAR.

- Reporting results:
    - Tables: The sample size, LS mean of change from baseline and 95% CI for each treatment, LS mean of difference and the corresponding 95% CI, and p-value will be presented for each week.

6.2.11. DLQI, CDLQI, IDQOL, DFI, POEM, PGIS/OGIS, PGIC/OGIC over time

Endpoints:

- Change from baseline in DLQI total score for subjects aged 16 years and older at Day 15 and Day 29.

- Change from baseline in CDLQI total score for subjects aged 4-15 years at Day 15 and Day 29.

- Change from baseline in IDQOL total score for subjects aged 2-3 years at Day 15 and Day 29.

- Change from baseline in DFI total score for subjects aged 2-17 years at Day 15 and Day 29.

- Change from baseline in POEM total score for self-report subjects aged 12 years or older at Day 15 and Day 29.

- Change from baseline in POEM total score for proxy-report subjects ≥2 and <12 years at Day 15 and Day 29.

- Change from baseline in weekly average of PGIS for subjects aged 12 years and older at Week 1, Week 2, Week 3 and Week 4.

- Change from baseline in weekly average of OGIS for subjects ≥2 and <12 years at Week 1, Week 2, Week 3 and Week 4.

- Observed PGIC for subjects aged 12 years and older at Day 8, Day 15, Day 22 and Day 29.

- Observed OGIC for subjects ≥2 and <12 years at Day 8, Day 15, Day 22 and Day 29.

Analysis:

- Analysis set: FAS

- Analysis methodology: Descriptive statistics will be summarized for the above endpoints.

- Missing data:

If two or more response options are checked, the response option with the highest score should be recorded. If there is a response between two boxes, the lower of the two score options should be recorded.

For DLQI, CDLQI, IDQOL, DFI and POEM, the total score is considered missing if more than one item is missing. If one item is missing then total score is computed as the simple algebraic sum of the final item scores for nine non-missing items.

For PGIS/OGIS, missing data for any day will not be imputed. The weekly average is considered missing if PGIS/OGIS scores are missing for more than three days. If the scores are missing for less than four days, then the weekly average is computed as the mean of the scores for the non-missing days.

For PGIC/OGIC, missing data will not be imputed.

- Reporting results:

    - Tables: For change from baseline in DLQI, CDLQI, IDQOL, DFI and POEM, the sample size, mean, standard deviation, median and range, will be presented for each

treatment at Day 15 and Day 29. For change from baseline in weekly average of PGIS and OGIS, the sample size, mean, standard deviation, median and range, will be presented for each treatment at Week 1, Week 2, Week 3 and Week 4. For observed PGIC and OGIC, the sample size, mean, standard deviation, median and range, will be presented for each treatment at Day 8, Day 15, Day 22 and Day 29.

CCI

CCI

CCI

CCI

CCI

CCI

[redacted]

6.4. Subset Analyses

The main analysis of primary efficacy endpoint and key secondary efficacy endpoints will be repeated by subgroups defined through geographical regions, age groups (ages 2–11 years, ages 12–17 years, and ages 18 years and older), and baseline ISGA score (ISGA score 2, ISGA score 3). Subset analyses of these endpoints aim to evaluate the consistency of treatment effect across various subsets. P-values will not be reported here.

Summary of TEAEs and participant discontinuations due to AE will be reported by subgroups defined through geographical regions.

6.5. Baseline and Other Summaries and Analyses

6.5.1. Baseline Summaries

Baseline variables will be summarized. Demographics and baseline characteristics will be summarized according to Pfizer standards.

6.5.2. Study Conduct and Participant Disposition

Subjects evaluation, disposition, discontinuation will be summarized according to Pfizer standards.

6.5.3. Study Treatment Exposure

The extent of exposure to study drug in each treatment group will be summarized by the total number of days of dosing, total number of applications, total amount of study drug applied, and number and percentage of subjects who are compliant with the dosing regimen.

During Days 1–29, a subject will be considered compliant with the dosing regimen if they receive at least 45 but no more than 67 (ie, 80%–120%, inclusive) of the expected 56 total doses to be administered between Baseline/Day 1 and Day 29, and has not missed 6 or more consecutive doses during the treatment period.

6.5.4. Concomitant Medications and Nondrug Treatments

Prior drug and non-drug treatment, concomitant drug and non-drug treatment will be summarized according to Pfizer standards.

6.6. Safety Summaries and Analyses

All safety analyses will be conducted on the Safety Analysis Sset.

Pfizer standard will be utilized in reporting routine safety variables (Pfizer Safety Rulebook).

6.7. Additional Analyses to Address COVID-19 Pandemic Impacts

- Protocol deviations related to COVID-19 pandemic will be summarized and listed separately. Both important and non-important PDs related to COVID-19 pandemic will be reported.
- A separate summary table solely for subject discontinuations related to COVID-19 pandemic, if any, will be provided.
- COVID-19 related AEs, if any, will be reported.

7. INTERIM ANALYSES

7.1. Introduction

This study uses an External Data Monitoring Committee (E-DMC).

The E-DMC is responsible for ongoing monitoring of the safety of subjects in the study according to the charter. The recommendations made by the E-DMC to alter the conduct of the study will be forwarded to Pfizer for final decision. Pfizer will forward such decisions, which may include summaries of aggregate analyses of safety data, to regulatory authorities, as appropriate.

7.2. Interim Analyses and Summaries

No interim analyses of the study data is planned for use outside the E-DMC.

8. REFERENCES

Mallinckrodt, C.H., Clark, W.S. and David, S.R., 2001. Accounting for dropout bias using mixed-effects models. *Journal of biopharmaceutical statistics*, *11*(1-2), pp.9-21.

9. APPENDICES

Appendix 1. Summary of Efficacy Analysis

Table1. List of primary and secondary efficacy analyses

Efficacy Endpoints	Analysis Set	Analysis Method	Missing Data Imputation	Primary Analysis Endpoints
Percent change from baseline in EASI total score at Day 29	FAS	MMRM (model includes Days 8, 15, 22, 29)	Data after treatment discontinuation will not be included, MMRM handles missing data.	Yes
Percent change from baseline in EASI total score at Day 29 (Supplementary)	PPAS	MMRM (model includes Days 8, 15, 22, 29)	Data after treatment discontinuation will not be included, MMRM handles missing data.	Yes
Achievement of Improvement in ISGA at Days 29	FAS	normal approximation to response rates	MI-MCMC	No
Achievement of Improvement in ISGA at Days 29 (Supportive)	FAS	normal approximation to response rates	NRI	No
Achievement of Improvement in ISGA at Days 29 (Supplementary)	PPAS	normal approximation to response rates	MI-MCMC	NO
Achievement of Success in ISGA at Day 29	FAS	Normal approximation to response rates	MI-MCMC	No
Achievement of Success in ISGA at Day 29 (Supportive)	FAS	Normal approximation to response rates	NRI	No
Achievement of Success in ISGA at Day 29 (Supplementary)	PPAS	Normal approximation to response rates	MI-MCMC	No

Table1. List of primary and secondary efficacy analyses

Efficacy Endpoints	Analysis Set	Analysis Method	Missing Data Imputation	Primary Analysis Endpoints
Change from Baseline in Peak Pruritus NRS at Week 4 – for subjects ≥12 years	FAS	MMRM (model includes Week 1, 2, 3, 4)	Data after treatment discontinuation will not be included, MMRM handles missing data.	No
Change from Baseline in Peak Pruritus NRS at Week 4 – for subjects ≥12 years (Supplementary)	PPAS	MMRM (model includes Week 1, 2, 3, 4)	Data after treatment discontinuation will not be included, MMRM handles missing data.	No
Success in ISGA over time	FAS	normal approximation to response rates	MI-MCMC	No
Improvement in ISGA over time	FAS	normal approximation to response rates	MI-MCMC	No
Percent Change from baseline in EASI total score over time	FAS	MMRM (model includes Days 8, 15, 22, 29)	MMRM handles missing data	No
Change from baseline in %BSA over time	FAS	MMRM (model includes Days 8, 15, 22, 29)	MMRM handles missing data	No
Achievement of EASI-50 over time	FAS	normal approximation to response rates	MI-MCMC	No
Achievement of EASI-75 over time	FAS	normal approximation to response rates	MI-MCMC	No
Change from Baseline in Peak Pruritus NRS over time - for subjects ≥12 years	FAS	MMRM (model includes Week 1, 2, 3, 4)	MMRM handles missing data	No

Table1. List of primary and secondary efficacy analyses

Efficacy Endpoints	Analysis Set	Analysis Method	Missing Data Imputation	Primary Analysis Endpoints
Change from Baseline in Patient Reported Itch Severity Scale over time - for subjects ≥6 years and <12 years	FAS	MMRM (model includes Week 1, 2, 3, 4)	MMRM handles missing data	No
Change from Baseline in Observer Reported Itch Severity Scale over time - for subjects <6 years	FAS	MMRM (model includes Week 1, 2, 3, 4)	MMRM handles missing data	No
DLQI, CDLQI, IDQOL, DFI, POEM, PGIS/OGIS, PGIC/OGIC over time	FAS	Descriptive statistics	Specified in Section 6.2.11	No

MMRM= Mixed Effect Model for Repeated Measures; FAS=Full Analysis Set; NRI= Non-responder imputation; MI-MCMC=Multiple Imputation based on Markov Chain Monte Carlo.

Appendix 2. DATA DERIVATION DETAILS

Appendix 2.1. Definition and Use of Visit Windows in Reporting

Visit windows will be used for efficacy summaries, and for any safety summaries that display by visit.

Table 2. Definition of Visit Windows

Visit Label	Target Day	Definition [Day window]
Baseline	Day 1,	≤Day 1
Day 8	Day 8	Days 2-11
Dat 15	Day 15	Days 12-18
Day 22	Day 22	Days 19-25
Day 29	Day 29	Days 26-last dose+7

Table 3. Visit window of Vital Signs/POEM/CDLQI/DLQI/IDLQI/DFI:

Visit Label	Target Day	Definition
Baseline	Day 1	<=Day 1
Day 15	Day 15	Days 2-21
Day 29	Day 29	Days 22 – last dosing date+7

Table 4. Visit window of Lab and limited physical exam:

Visit Label	Target Day	Definition
Baseline	Day 1	<=Day 1
Day 29	Day 29	Days 2 - last dosing date+7

Table 5. Visit window of weekly average of NRS/PGIS/OGIS:

Visit Label	Definition
Baseline	Days -6 - 1
Week 1	Days 2 - 8
Week 2	Days 9 - 15
Week 3	Days 16 - 22
Week 4	Days 23 - 29

If two or more visits fall into the same window, keep the one closest to the Target Day. If two visits are of equal distance from the Target Day in absolute value, the later visit should be used. But if two visits both fall on Day 1, then the first visit will be selected.

(Safety analysis to follow Pfizer standard. See Pfizer Safety Rulebook.)

Appendix 3. Details of Efficacy Assessments

Appendix 3.1. Investigator's Static Global Assessment (ISGA)

The Investigator's Static Global Assessment (ISGA), a five point global static assessment of AD severity (Table 6), will be assessed at times specified in the Schedule of Activities of the protocol to characterize subjects' overall disease severity across all treatable AD lesions.

The assessment will be a static evaluation without regard to the score at a previous visit.

ISGA assessment during the study must be done by the investigator or his/her designee. Every effort should be made to ensure that all ISGA assessments for a given subject are done by the same qualified individual throughout the study.

Table 6. Investigator's Static Global Assessment

Score	Grade	Definition
0	Clear	Minor residual hypo/hyperpigmentation; no erythema or induration/papulation; no oozing/crusting
1	Almost Clear	Trace faint pink erythema, with barely perceptible induration/papulation and no oozing/crusting
2	Mild	Faint pink erythema with mild induration/papulation and no oozing/crusting
3	Moderate	Pink-red erythema with moderate induration/population with or without oozing/crusting
4	Severe	Deep or bright red erythema with severe induration/population and with oozing/crusting

Appendix 3.2. Eczema Area and Severity Index (EASI)

The EASI quantifies the severity of a subject's atopic dermatitis based on both severity of lesion clinical signs and the percent of BSA affected. EASI is a composite scoring of the degree of erythema, induration/papulation, excoriation, and lichenification (each scored separately) for each of four body regions, with adjustment for the percent of BSA involved for each body region and for the proportion of the body region to the whole body.

Percent BSA with Treatable AD: The number of handprints of AD skin in a body region can be used to determine the extent (%) to which a body region is involved with atopic dermatitis (Table 7).

Table 7. Handprint Determination of Body Region Surface Area

Body Region	Total Number of Handprints in Body Region*	Surface Area of Body Region Equivalent of One Handprint	Total Number of Handprints in Body Region*	Surface Area of Body Region Equivalent of One Handprint
	≥8 years of age		2-7 years of age	
Head and Neck	10	10%	20	5%
Upper Limbs	20	5%	20	5%
Trunk (including axillae)	30	3.33%	30	3.33%
Lower Limbs (including buttocks)	40	2.5%	30	3.33%

*The number of handprints will be for the entire body region; these values will not be adjusted for exclusion of scalp, palms, and soles from the BSA assessment.

The extent (%) to which each of the four body regions is involved with AD is categorized using a non-linear scaling method to a numerical area score according to the following BSA scoring criteria. See Table 8 on EASI Area Score Criteria.

Table 8. EASI Area Score Criteria

Treatable Percent Body Surface Area (BSA) with Atopic Dermatitis in a Body Region	Area Score
0%	0
>0-<10%	1
10-<30%	2
30-<50%	3
50-<70%	4
70-<90%	5
90-<100%	6

Lesion Severity by Clinical Signs: The basic characteristics of atopic dermatitis lesions erythema, induration/papulation, excoriation, and lichenification provide a means for assessing the severity of lesions. Assessment of these four main clinical signs is performed separately for four body regions: head and neck, upper limbs, trunk (including axillae and groin) and lower limbs (including buttocks). Average erythema, induration/papulation, excoriation, and lichenification are scored for each body region according to a 4 point scale: 0 = absent; 1 = mild; 2 = moderate; 3 = severe. Morphologic descriptors for each clinical sign severity score are shown in Table 9.

Table 9. Clinical Sign Severity Scoring Criteria for the EASI

Score		Description
Erythema (E)		
0	Absent	None; may have residual discoloration (post-inflammatory hyperpigmentation and/or hypopigmentation).
1	Mild	Light pink to light red
2	Moderate	Red
3	Severe	Deep, dark red
Induration/Papulation (I)		
0	Absent	None
1	Mild	Barely palpable to slight, but definite hard thickened skin and/or papules
2	Moderate	Easily palpable moderate hard thickened skin and/or papules
3	Severe	Severe hard thickened skin and/or papules
Excoriation (Ex)		
0	Absent	None
1	Mild	Slight, but definite linear or picked scratch marks or penetrating surface injury
2	Moderate	Moderate linear or picked scratch marks or penetrating surface injury
3	Severe	Severe linear or picked scratch marks or penetrating surface injury
Lichenification (L)		
0	Absent	None
1	Mild	Barely perceptible to slight, but definite thickened skin, fine skin markings, and lichenoid scale
2	Moderate	Moderate thickened skin, coarse skin markings, and coarse lichenoid scale
3	Severe	Severe thickened skin with very coarse skin markings and lichenoid scale

* The EASI will exclude scalp from the assessment/scoring

In each body region, the sum of the Clinical Signs Severity Scores for erythema, induration/papulation, excoriation, and lichenification is multiplied by the Area Score and by the Body Region Weighting to provide a body region value, which is then summed across all four body regions resulting in an EASI score as described in Equation 1 and Equation 2:

Equation 1 (subject aged 2-7 years old): EASI =0.2Ah(Eh+Ih+Exh+Lh) + 0.2Au(Eu+Iu+Exu+Lu) + 0.3At(Et+It+Ext+Lt) + 0.3Al(El+Il+Exl+Ll)

Equation 2 (subject aged > 7 years old): EASI =0.1Ah(Eh+Ih+Exh+Lh) + 0.2Au(Eu+Iu+Exu+Lu) + 0.3At(Et+It+Ext+Lt) + 0.4Al(El+Il+Exl+Ll)

A = Area Score; E = erythema; I = induration/papulation; Ex = excoriation; L = lichenification; h = head and neck; u = upper limbs; t = trunk; l = lower limbs

The EASI score can vary in increments of range from 0.0 to 72.0, with higher scores representing greater severity of AD. Considering the scalp will be excluded from EASI assessment in the study, the maximum possible score will be less than 72.0 (modified EASI score).

Reference:

Bissonnette R, Papp KA, Poulin Y, et al. Topical tofacitinib for atopic dermatitis: a phase IIa randomized trial. Bit J Dermatol 2016; 175(5):902-911.

Appendix 3.3. Peak Pruritus Numerical Rating Scale (NRS) for subjects ≥12 years, Patient Reported Itch Severity Scale - for subjects age 6-11 years, and Observer Reported Itch Severity Scale - for subjects <6 years.

The severity of itch (pruritus) due to atopic dermatitis will be assessed using the Peak Pruritus Numerical Rating Scale (NRS) for subjects ≥12 years, an 11-category numeric rating scale from 0 to 10, which is subject (12 years and older) reported. A five-category Patient Reported Itch Severity Scale - for subjects age 6-11 years has been developed for subjects ≥6 and <12 years of age. The Observer Reported Itch Severity Scale - for subjects <6 years will be completed by a caregiver for subjects <6 years old. It is preferred that all observer reported outcomes for a given subject are completed by same individual throughout the study.

Subjects will be asked to assess their worst itching due to atopic dermatitis over the past 24 hours on an NRS anchored by the terms "no itch" (0) and "worst itch imaginable" (10). The Peak Pruritus NRS for subjects ≥12 years is presented in Figure 1. The Patient Reported Itch Severity Scale - for subjects age 6-11 years is presented in Figure 2. The Observer Reported Itch Severity Scale - for subjects <6 years is presented in Figure 3. These scales are designed to capture a similar concept.

Figure 1. Peak Pruritus NRS [©Regeneron Pharmaceuticals, Inc. and Sanofi (2017)] for subjects ≥12 years

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

0	1	2	3	4	5	6	7	8	9	10

No itch

Worst imaginable itch

Figure 2. Patient Reported Itch Severity Scale - for subjects age 6-11 years

Circle the face that shows how itchy your skin has been today:

Not Itchy Very Itchy

Figure 3. Observer Reported Itch Severity Scale - for subjects <6 years

On a scale of 0 to 10, with 0 being "no itch" and 10 being "worst itch imaginable", how would you rate your observation of your child's itch (scratching, rubbing) at the worst moment during the previous 24 hours?
0 1 2 3 4 5 6 7 8 9 10 No itch Worst imaginable itch

References:

Yosipovitch G, Reaney M, Mastey V, et al. Validation of the peak pruritus numerical rating scale: Results from clinical studies of dupilumab in adult patients with moderate-to-severe atopic dermatitis. J Am Acad of Dermatol 2017; 76:AB278.

Simpson E, Beck L, Abhijit G, et al. Defining a responder on the Peak Pruritus Numerical Rating Scale (NRS) in patients with moderate-to-severe atopic dermatitis: Detailed analysis from randomized trials of dupilumab J Am Acad of Dermatol 2017; 76:AB93.

Appendix 3.4. Dermatology Life Quality Index (DLQI)

http://sites.cardiff.ac.uk/dermatology/quality-of-life/dermatology-quality-of-life-index-dlqi/dlqi-instructions-for-use-and-scoring/

The Dermatology Life Quality Index questionnaire is designed for use in adults, ie, patients aged 16 years and over. It is self explanatory and can be simply handed to the patient who is asked to fill it in without the need for detailed explanation. It is usually completed in one to two minutes.

The aim of this questionnaire is to measure how much your skin problem has affected your life OVER THE LAST WEEK. Please tick ☑ one box for each question.

1.	Over the last week, how **itchy**, **sore**, **painful** or **stinging** has your skin been?	Very much A lot A little Not at all	❒ ❒ ❒ ❒	
2.	Over the last week, how **embarrassed** or **self conscious** have you been because of your skin?	Very much A lot A little Not at all	❒ ❒ ❒ ❒	
3.	Over the last week, how much has your skin interfered with you going **shopping** or looking after your **home** or **garden**?	Very much A lot A little Not at all	❒ ❒ ❒ ❒	 Not relevant ❒
4.	Over the last week, how much has your skin influenced the **clothes** you wear?	Very much A lot A little Not at all	❒ ❒ ❒ ❒	 Not relevant ❒
5.	Over the last week, how much has your skin affected any **social** or **leisure** activities?	Very much A lot A little Not at all	❒ ❒ ❒ ❒	 Not relevant ❒
6.	Over the last week, how much has your skin made it difficult for you to do any **sport**?	Very much A lot A little Not at all	❒ ❒ ❒ ❒	 Not relevant ❒
7.	Over the last week, has your skin prevented you from **working** or **studying**?	Yes No	❒ ❒	 Not relevant ❒
	If "No", over the last week how much has your skin been a problem at **work** or **studying**?	A lot A little Not at all	❒ ❒ ❒	
8.	Over the last week, how much has your	Very much	❒	

	skin created problems with your **partner** or any of your **close friends** or **relatives**?	A lot A little Not at all	❒ ❒ ❒	 Not relevant ❒
9.	Over the last week, how much has your skin caused any **sexual difficulties**?	Very much A lot A little Not at all	❒ ❒ ❒ ❒	 Not relevant ❒
10.	Over the last week, how much of a problem has the **treatment** for your skin been, for example by making your home messy, or by taking up time?	Very much A lot A little Not at all	❒ ❒ ❒ ❒	 Not relevant ❒

Scoring
The scoring of each question is as follows:

Response	Score
Very much	scored 3
A lot	scored 2
A little	scored 1
Not at all	scored 0
Not relevant	scored 0
Question unanswered	scored 0
Question 7: “prevented work or studying”	scored 3

The DLQI is calculated by summing the score of each question resulting in a maximum of 30 and a minimum of 0. The higher the score, the more quality of life is impaired. The DLQI can also be expressed as a percentage of the maximum possible score of 30.

****Please Note:** That the scores associated with the different answers should not be printed on the DLQI itself, as this might cause bias****

Meaning of DLQI Scores
0-1 = no effect at all on patient’s life
2-5 = small effect on patient’s life
6-10 = moderate effect on patient’s life
11-20 = very large effect on patient’s life
21-30 = extremely large effect on patient’s life

Detailed analysis of the DLQI
The DLQI can be analysed under six headings as follows:

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

The scores for each of these sections can also be expressed as a percentage of either 6 or 3.

Interpretation of incorrectly completed questionnaires
There is a very high success rate of accurate completion of the DLQI. However, sometimes subjects do make mistakes.

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

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

3. If question 7 is answered 'yes' this is scored 3. If question 7 is answered 'no' but then either 'a lot' or 'a little' is ticked this is then scored 2 or 1. If "Not relevant" is ticked, the score for Question 7 is 0. If it is answered 'no', but the second half is left incomplete, the score will remain 0.

4. If two or more response options are ticked, the response option with the highest score should be recorded.

5. If there is a response between two tick boxes, the lower of the two score options should be recorded.

6. The DLQI can be analysed by calculating the score for each of its six sub-scales (see above). When using sub-scales, if the answer to one question in a sub-scale is missing, that sub-scale should not be scored.

Minimal Clinically Important Difference of the DLQI
For **general inflammatory skin conditions** a change in DLQI score of at least 4 points is considered clinically important (Basra et al, 2015, see below). This means that a patient's DLQI score has to either increase or decrease by at least 4 points in order to suggest that

there has actually been a meaningful change in that patient's quality of life since the previous measurement of his/her DLQI scores.

Key References

Original Reference

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

Appendix 3.5. Children's Dermatology Life Quality Index (CDLQI)

http://sites.cardiff.ac.uk/dermatology/quality-of-life/childrens-dermatology-life-quality-index-cdlqi/cdlqi-information-and-instructions/

The Children's Dermatology Life Quality Index questionnaire is designed for use in children, ie, patients from age 4 to age 16. It is self explanatory and can be simply handed to the patient who is asked to fill it in with the help of the child's parent or guardian. It is usually completed in one to two minutes.

The aim of this questionnaire is to measure how much your skin problem has affected you OVER THE LAST WEEK. Please tick ✓ one box for each question.

1.	Over the last week, how **itchy**, **"scratchy"**, **sore** or **painful** has your skin been?	Very much Quite a lot Only a little Not at all	☐ ☐ ☐ ☐
2.	Over the last week, how **embarrassed** or **self conscious**, **upset** or **sad** have you been because of your skin?	Very much Quite a lot Only a little Not at all	☐ ☐ ☐ ☐
3.	Over the last week, how much has your skin affected your **friendships**?	Very much Quite a lot Only a little Not at all	☐ ☐ ☐ ☐
4.	Over the last week, how much have you changed or worn **different** or **special clothes/shoes** because of your skin?	Very much Quite a lot Only a little Not at all	☐ ☐ ☐ ☐
5.	Over the last week, how much has your skin trouble affected **going out**, **playing**, or **doing hobbies**?	Very much Quite a lot Only a little Not at all	☐ ☐ ☐ ☐
6.	Over the last week, how much have you avoided **swimming** or **other sports** because of your skin trouble?	Very much Quite a lot Only a little Not at all	☐ ☐ ☐ ☐
7.	Last week, was it **school time?** **OR**	**If school time**: Over the last week, how much did your skin problem affect your **school work**?	Prevented school Very much Quite a lot Only a little Not at all	☐ ☐ ☐ ☐ ☐

	was it **holiday time?**	⇨	**If holiday time**: How much over the last week, has your skin problem interfered with your enjoyment of the **holiday**?	Very much ☐ Quite a lot ☐ Only a little ☐ Not at all ☐
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**?			Very much ☐ Quite a lot ☐ Only a little ☐ Not at all ☐
9.	Over the last week, how much has your **sleep** been affected by your skin problem?			Very much ☐ Quite a lot ☐ Only a little ☐ Not at all ☐
10.	Over the last week, how much of a problem has the **treatment** for your skin been?			Very much ☐ Quite a lot ☐ Only a little ☐ Not at all ☐

Scoring

The scoring of each question is as follows:

Very much	scored 3
Quite a lot	scored 2
Only a little	scored 1
Not at all	scored 0
Question unanswered	scored 0
Question 7: "Prevented school"	scored 3

The CDLQI is calculated by summing the score of each question resulting in a maximum of 30 and a minimum of 0. The higher the score, the more quality of life is impaired. The CDLQI can also be expressed as a percentage of the maximum possible score of 30.

Detailed analysis of the CDLQI

The CDLQI can be analysed under six headings as follows:

Symptoms and feelings	Questions 1 and 2	Score maximum 6
Leisure	Questions 4, 5 and 6	Score maximum 9
School or holidays	Questions 7	Score maximum 3
Personal relationships	Question 3 and 8	Score maximum 6
Sleep	Questions 9	Score maximum 3

Treatment	Question 10	Score maximum 3

The scores for each of these sections can also be expressed as a percentage of 9, 6 or 3.

The severity banding for CDLQI scores:
0-1 = no effect on child's life
2-6 = small effect
7-12 = moderate effect
13-18 = very large effect
19-30 = extremely large effect

Ref: Waters A, Sandhu D, Beattie P, Ezughah F, Lewis-Jones S. Severity stratification of Children's Dermatology Life Quality Index (CDLQI) scores. Br J Dermatol 2010; 163 (Suppl 1): 121.

Interpretation of incorrectly completed questionnaires
There is a very high success rate of accurate completion of the CDLQI. However, sometimes subjects do make mistakes.

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

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

3. If both parts of question 7 are completed the higher of the two scores should be counted

References
Lewis-Jones MS, Finlay AY. The Children's Dermatology Life Quality Index (CDLQI): Initial validation and practical use. British Journal of Dermatology, 1995; 132: 942-949.

Appendix 3.6. Dermatitis Family Impact Questionnaire (DFI)

http://sites.cardiff.ac.uk/dermatology/quality-of-life/dermatitis-family-impact-questionnaire-dfi/dfi-information-and-instructions/

The aim of this questionnaire is to measure how much your child's skin problem has affected you and your family OVER THE LAST WEEK. Please tick ☑ one box for each question.

1.	Over the last week, how much effect has your child having eczema had on **housework**, eg, washing, cleaning.	Very much A lot A little Not at all	❒ ❒ ❒ ❒
2.	Over the last week, how much effect has your child having eczema had on **food preparation** and **feeding**.	Very much A lot A little Not at all	❒ ❒ ❒ ❒
3.	Over the last week, how much effect has your child having eczema had on the **sleep of others in family**.	Very much A lot A little Not at all	❒ ❒ ❒ ❒
4.	Over the last week, how much effect has your child having eczema had on **family leisure activities**, eg swimming.	Very much A lot A little Not at all	❒ ❒ ❒ ❒
5.	Over the last week, how much effect has your child having eczema had on **time spent on shopping for the family.**	Very much A lot A little Not at all	❒ ❒ ❒ ❒
6.	Over the last week, how much effect has your child having eczema had on your **expenditure**, eg costs related to treatment, clothes, etc.	Very much A lot A little Not at all	❒ ❒ ❒ ❒
7.	Over the last week, how much effect has your child having eczema had on causing **tiredness** or **exhaustion** in your child's parents/carers.	Very much A lot A little Not at all	❒ ❒ ❒ ❒
8.	Over the last week, how much effect has your child having eczema had on causing **emotional distress** such as depression, frustration or guilt in your child's parents/carers.	Very much A lot A little Not at all	❒ ❒ ❒ ❒
9.	Over the last week, how much effect has your child having eczema had on **relationships** between the **main carer and partner** or	Very much A lot A little	❒ ❒ ❒

	between the **main carer and other children** in the family.	Not at all	❒
10.	Over the last week, how much effect has **helping with your child's treatment** had on the main carer's life.	Very much	❒
		A lot	❒
		A little	❒
		Not at all	❒

Instructions for Use and Scoring

The scoring system for the DFI is as follows:

Each question is scored from 0-3.
Not at all = 0
A little = 1
A lot = 2
Very much = 3

The score of each of the 10 questions is summed.

The minimum DFI score is 0 (= no impact on life of family)
The maximum DFI score is 30 (= maximum effect on life of family)
There are no validated score banding descriptors yet published.

Key References

For details of the Dermatitis Family Impact Questionnaire please see the following references:

Lawson V, Lewis-Jones MS, Finlay AY, Reid P, Owens RG. The family impact of childhood atopic dermatitis: the Dermatitis Family Impact.

Appendix 3.7. The Infant's Dermatitis Quality of Life Index (IDQOL)

http://sites.cardiff.ac.uk/dermatology/quality-of-life/the-infants-dermatitis-quality-of-life-index-idqol/idqol-information-and-instructions/

Instructions for Use and Scoring

The Infant's Dermatitis Quality of Life Index questionnaire is designed for use in infants with atopic dermatitis below the age of 4 years. It is self-explanatory and should be completed by the child's parent (s) or regular carer. It can usually be completed within 2 or 3 minutes.

The aim of this chart is to record how your child's dermatitis has been. Each question concerns THE LAST WEEK ONLY. Please could you answer every question.

Dermatitis Severity

	Over the last week, **how severe** do you think your child's dermatitis has been?; ie, how red, scaly, inflamed or widespread.	Extremely severe Severe Average Fairly good None

Life Quality Index

1.	Over the last week, how much has your child been **itching and scratching**?	All the time A lot A little None
2.	Over the last week, what has your child's **mood** been?	Always crying, extremely difficult Very fretful Slightly fretful Happy
3.	Over the last week approximately how much **time** on average has it taken **to get your child off to sleep** each night?	More than 2 hrs 1 - 2 hrs 15mins - 1 hr 0-15mins
4.	Over the last week, what was the **total time** that your child's **sleep was disturbed** on average each night?	5 hrs or more 3 - 4 hrs 1 - 2 hrs Less than 1 hour
5.	Over the last week, has your child's eczema interfered with **playing or swimming**?	Very much A lot A little Not at all
6.	Over the last week, has your child's eczema interfered with your child **taking part in** or **enjoying other family activities**?	Very much A lot A little Not at all
7.	Over the last week, have there been problems with your child at **mealtimes** because of the eczema?	Very much A lot A little

8.	Over the last week, have there been problems with your child caused by the **treatment**?	None Very much A lot A little None
9.	Over the last week, has your child's eczema meant that **dressing and undressing** the child has been **uncomfortable**?	Very much A lot A little None
10.	Over the last week how much has your child having eczema been a problem at **bathtime**?	Very much A lot A little None

Scoring

The scoring of each question is as follows:

Dermatitis Severity (this is scored separately from the Life Quality Index)

Extremely severe	scored 4
Severe	scored 3
Average	scored 2
Fairly good	scored 1
None	scored 0
Maximum score is 4	

Life Quality Index Questions 1 and 5-10

All the time	scored 3
A lot	scored 2
A little	scored 1
None	scored 0

Question 2

Always crying etc.	scored 3
Very fretful	scored 2
Slightly fretful	scored 1
Happy	scored 0

Question 3

More than 2 hrs	scored 3
1-2 hrs	scored 2
15 mins-1hr	scored 1
0-15 mins	scored 0

Question 4

5hrs or more	scored 3
3-4 hours	scored 2
1-2 hours	scored 1
Less than 1 hour	scored 0

Question 5-10

Very much	scored 3
A lot	scored 2
A little	scored 1
Not at all/none	scored 0

The IDQOL is calculated by summing the score of each question resulting in a maximum of 30 and a minimum of 0. The higher the score the more quality of life is impaired. The severity of eczema is scored separately and can be correlated with the IDQOL.

Appendix 3.8. Patient Oriented Eczema Measure (POEM)

The POEM is a tool used for monitoring atopic eczema severity and is to be completed by the subject (POEM for self-completion, by subjects 12 years and older) or by the caregiver (POEM for proxy completion, for subjects <12 years). It is preferred that all observer reported outcomes for a given subject are completed by same individual throughout the study. The POEM contains 7 symptom based questions with responses rating number of days each symptom is experienced over the past week, from 0 (no days) to 4 (every day), with a maximum score of 28. A higher score indicates a worse outcome.

POEM for self-completion and/or proxy completion

Please circle one response for each of the seven questions below about your/your child's eczema. If your child is old enough to understand the questions then please fill in the questionnaire together. Please leave blank any questions you feel unable to answer.

1. Over the last week, on how many days has your/your child's skin been itchy because of the eczema?

No days **1-2 days** **3-4 days** **5-6 days** **Every day**

2. Over the last week, on how many nights has your/your child's sleep been disturbed because of the eczema?

No days **1-2 days** **3-4 days** **5-6 days** **Every day**

3. Over the last week, on how many days has your/your child's skin been bleeding because of the eczema?

No days **1-2 days** **3-4 days** **5-6 days** **Every day**

4. Over the last week, on how many days has your/your child's skin been weeping or oozing clear fluid because of the eczema?

No days **1-2 days** **3-4 days** **5-6 days** **Every day**

5. Over the last week, on how many days has your/your child's skin been cracked because of the eczema?

No days **1-2 days** **3-4 days** **5-6 days** **Every day**

6. Over the last week, on how many days has your /your child's skin been flaking off because of the eczema?

No days **1-2 days** **3-4 days** **5-6 days** **Every day**

7. Over the last week, on how many days has your/your child's skin felt dry or rough because of the eczema?

No days **1-2 days** **3-4 days** **5-6 days** **Every day**

Scoring

Each of the seven questions carries equal weight and is scored from 0 to 4 as follows:

No days	= 0
1-2 days	= 1
3-4 days	= 2
5-6 days	= 3
Every day	= 4

Total scores are associated with the following disease severities:

- 0 to 2 = Clear or almost clear;
- 3 to 7 = Mild eczema;
- 8 to 16 = Moderate eczema;

- 17 to 24 = Severe eczema;
- 25 to 28 = Very severe eczema.

Reference:

Charman CR, Venn AJ, Williams HC. The Patient-Oriented Eczema Measure: Development and Initial Validation of a New Tool for Measuring Atopic Eczema Severity From the Patients' Perspective. Arch Dermatol 2004; 140(12);1513-19.

Appendix 3.9. Patient Global Impression of Change (PGIC) and Observer Reported Global Impression of Change (OGIC)

The scale will be completed by the subject (PGIC, by subjects 12 years and older) or observer (OGIC, for subjects <12 years). It is preferred that all observer reported outcomes for a given subject are completed by same individual throughout the study. The PGIC and OGIC are single-item questions to rate change (retrospectively) in a subjects overall status since the start of the study.

This single item instrument uses a 7-point rating scale, anchored by (1) 'very much improved' to (7) 'very much worse'. The PGIC and OGIC will be used to determine global improvement as assessed by the subject or caregiver. It will be used as an anchor to define a responder definition for the peak pruritus scales for 'clinically important responder' and as a sensitivity analysis for defining a 'clinical important difference' on the peak pruritus scales.

Place an X in the box you feel most closely describes any change which you have experienced. Take into account all change, whether or not you believe it is entirely due to drug treatment. Select only ONE response.

Since starting the study medication, my atopic dermatitis is:

Very much improved ☐

Much improved ☐

Minimally improved ☐

No change ☐

Minimally worse ☐

Much worse ☐

Very much worse ☐

When an observer assesses the state of the atopic dermatitis, the question will be stated as follows:

Since starting the study medication, the child's atopic dermatitis is (select only ONE response).

PGIC or OGIC will be completed at the same day as the ISGA will be assessed post baseline.

Appendix 3.10. Patient Global Impression of Severity (PGIS) and Observer Reported Global Impression of Severity (OGIS)

The scale will be completed by the subject (PGIS, by subjects 12 years and older) or observer (OGIS, for subjects <12 years) every day and preferably at the same time as the peak pruritus NRS/peak pruritus scale or Observer Reported peak pruritus NRS. It is preferred that all

observer reported outcomes for a given subject are completed by same individual throughout the study.

The PGIS and OGIS is a single-item patient or observer-rated measure of the subject's atopic dermatitis condition severity at a given point in time.

This single item instrument uses a 7-point rating scale. The PGIS and OGIS will be used as an anchor for defining a 'clinical important difference' on the peak pruritus scales and can also be used to create severity categorization for peak pruritus scales to enhance interpretation.

Please rate the severity of your atopic dermatitis right now (select only ONE response):

Not present ☐

Very mild ☐

Mild ☐

Moderate ☐

Moderately Severe ☐

Severe ☐

Extremely Severe ☐

The OGIS is similarly structured:

Please rate the severity of the child's atopic dermatitis right now (select only ONE response):

Not present ☐

Very mild ☐

Mild ☐

Moderate ☐

Moderately Severe ☐

Severe ☐

Extremely Severe ☐

Appendix 4. List of Abbreviations

Abbreviation	Term
AD	atopic dermatitis
AE	adverse event
BID	twice daily
BSA	Body Surface Area
CDLQI	Children's Dermatology Life Quality Index
CI	confidence interval
DFI	Dermatitis Family Impact Questionnaire
DLQI	Dermatology Life Quality Index
EASI	Eczema Area and Severity Index
E-DMC	external data monitoring committee
FAS	full analysis set
IDQOL	Infant's Dermatitis Quality of Life Index
IFNγ	interferon gamma
ISGA	Investigator's Static Global Assessment
LS	least-squares
LSLV	last subject last visit
MAR	missing at random
MCMC	Markov Chain Monte Carlo
MI	Multiple Imputation
MMP	matrix metalloproteinase
MMRM	Mixed effect model for repeated measures
NRI	Non-Responder Imputation
NRS	Numeric Rating Scale
OGIC	Observer Reported Global Impression of Change
OGIS	Observer Reported Global Impression of Severity
PDE 4	phosphodiesterase 4
PGIC	Patient Global Impression of Change
PGIS	Patient Global Impression of Severity
POEM	Patient Oriented Eczema Measure
PPAS	Per Protocol Analysis Set
PRO	Patient Reported Outcome
SAE	serious adverse event
SAP	statistical analysis plan
TEAE	treatment emergent adverse event