Protocol C3291037

A PHASE 3B/4, MULTICENTER, RANDOMIZED, ASSESSOR BLINDED, VEHICLE AND ACTIVE (TOPICAL CORTICOSTEROID AND CALCINEURIN INHIBITOR) CONTROLLED, PARALLEL GROUP STUDY OF THE EFFICACY, SAFETY, AND LOCAL TOLERABILITY OF CRISABOROLE OINTMENT, 2% IN PEDIATRIC AND ADULT SUBJECTS (AGES 2 YEARS AND OLDER) WITH MILD TO MODERATE ATOPIC DERMATITIS

Statistical Analysis Plan
(SAP)

Version: **3**

Date: 10-NOV-2020

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. VERSION HISTORY

This SAP for Study C3291037 is based on the protocol amendment 3 dated 12AUG2019.

Table 1. Summary of Major Changes in SAP Amendments

Version/Date	Associated Protocol Amendment	Specific Changes	Rationale
1 16MAY2018	Original 16MAR2018	Not Applicable	Not Applicable
2 27SEP2019	Protocol Amendment 3 12 Aug 2019	Removal of telephone follow-up at Day 36; addition of a clinic visit at Day 43; addition of restrictions to concurrent medications during the follow-up phase; and changes to the pruritis scale descriptions, endpoints, and analysis.	Per Protocol Amendment 3
3 10NOV2020	Protocol Amendment 3 12 Aug 2019	Modifications of efficacy and patient reported outcomes endpoints: • Only descriptive summary will be provided for primary efficacy and selected secondary endpoints. • No summary/analysis will be done for tertiary/exploratory endpoints. Appendix 8 Listing of Abbreviations was added	Study termination

2. INTRODUCTION

This SAP provides the detailed methodology for summary and statistical analyses of the data collected in Study C3291037. 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

Study objectives and corresponding endpoints are provided in the Table 2 below.

Table 2. Study Objectives and Endpoints

Primary Efficacy Objective	Primary Efficacy Endpoint
• To compare the efficacy of crisaborole ointment, 2% applied twice daily (BID) versus vehicle in pediatric and adult subjects (ages 2 years and older) with mild to moderate Atopic Dermatitis (AD).	• Percent change from baseline in the Eczema Area and Severity Index (EASI) total score at Day 29.
Primary Safety Objectives	**Primary Safety Endpoints**
• To evaluate the safety and local tolerability of crisaborole ointment, 2% applied BID versus vehicle in pediatric and adult subjects (ages 2 years and older) with mild to moderate AD. • To evaluate the safety and local tolerability of hydrocortisone butyrate cream 0.1% and pimecrolimus cream 1% applied BID in pediatric and adult subjects (ages 2 years and older) with mild to moderate AD.	• Adverse Events (AEs), Serious Adverse Events (SAEs), local tolerability, discontinuations and clinically significant changes in vital signs and clinical laboratory parameters.
Secondary Objectives	**Secondary Endpoints**
• To evaluate the effect of crisaborole ointment, 2% applied BID versus vehicle on additional efficacy endpoints over time in pediatric and adult subjects (ages 2 years and older) with mild to moderate AD. • To evaluate the efficacy of crisaborole ointment, 2% BID versus hydrocortisone butyrate cream 0.1% and pimecrolimus cream 1% applied BID in pediatric and adult subjects (ages 2 years and older) with mild to moderate AD.	Efficacy endpoints: • Percent change from baseline in EASI total score by scheduled time points except Day 29. • Achievement of success in the Investigator's Static Global Assessment (ISGA) (defined as an ISGA score of Clear (0) or Almost Clear (1) with at least a 2-grade improvement from baseline) by scheduled time points. • Achievement of ISGA score of clear (0) or almost clear (1) by scheduled time points. • Achievement of EASI75 (≥75% improvement from baseline) by scheduled time points. • Time to EASI75. • Change from baseline in %BSA by scheduled time points.
• To evaluate the effect of crisaborole ointment, 2% applied BID versus vehicle, hydrocortisone butyrate cream 0.1% and pimecrolimus cream 1% applied BID on patient/observer reported outcomes over time in pediatric and adult subjects (ages 2 years and older) with mild to moderate AD.	Patient/observer reported outcomes (PRO) endpoints: • Change from Baseline in Peak Pruritus Numerical Rating Scale (NRS) – for subjects >12 years by scheduled time points. • Change from Baseline in Patient Reported Itch Severity Scale - for subjects age 6-11 years Scale by scheduled time points. • Change from Baseline in Observer Reported Itch Severity Scale – for subjects <6 years by scheduled time points. • Time to ≥2-point improvement from Baseline in Peak Pruritus NRS for subjects >12 years. • Time to ≥3-point improvement from Baseline in Peak Pruritus NRS for subjects >12 years. • Time ≥2 point to improvement from Baseline in Observer Reported Itch Severity Scale - for subjects <6 years. • Time to ≥3-point improvement from Baseline in Observer Reported Itch Severity Scale - for subjects <6 years. • Achievement of ≥2-point improvement from Baseline in Peak Pruritus NRS for subjects >12 years. • Achievement of ≥3-point improvement from Baseline in Peak Pruritus NRS for subjects >12 years. • Achievement of ≥2-point improvement from Baseline in Observer Reported Itch Severity Scale - for subjects <6 years. • Achievement of ≥3-point improvement from Baseline in Observer Reported Itch Severity Scale - for subjects <6 years. • Change from baseline in Dermatology Life Quality Index (DLQI) (for Subjects 16 years and older), Children's Dermatology Life Quality Index (CDLQI) (for Subjects 4-15 years), and Dermatitis Family Impact Questionnaire (DFI) (Completed by parent/caregiver

Table 2. Study Objectives and Endpoints

	of Subjects 2-17 years) by scheduled time points.
CCI	[redacted]
[redacted]	[redacted]

2.2. Study Design

This is a Phase 3b/4, multicenter, randomized, assessor blinded, vehicle and active (topical corticosteroid [TCS] and topical calcineurin inhibitor [TCI]) controlled study of the efficacy, safety and local tolerability of crisaborole ointment, 2% in pediatric and adult subjects, ages 2 years and older, with mild to moderate AD involving at least 5% treatable body surface area (%BSA). Treatment will be clinical assessor blinded for all treatment arms and double blinded for crisaborole ointment, 2% and vehicle treatment arms.

A total of approximately 600 subjects will be enrolled in the study, of which at least 150 subjects aged 2-6; at least 140 subjects aged 7-11; at least 120 subjects aged 12-17, and up to 90 subjects will be adults. Following the screening period (up to 35 days prior to Baseline/Day 1), eligible subjects will be randomized at the Baseline/Day 1 visit. Randomization will be stratified by eligibility for TCS or TCI treatment as per national approved labels and eligibility provided in Section 4 of the protocol. Cohort 1 will be for subjects who are eligible for TCS therapy, and Cohort 2 will be for subjects who are not eligible for TCS therapy but eligible for TCI therapy. The investigational product will be applied BID for 28 days to the Treatable BSA identified at Baseline/Day 1. Any new AD on treatment eligible locations occurring following Baseline/Day 1 should also be treated with the study drug after consultation with the Investigator at the next visit.

The primary efficacy endpoint is the percent change from baseline in the EASI total score at Day 29.

For the efficacy comparison of crisaborole versus vehicle, subjects from both Cohort 1 and Cohort 2 are included in the analysis, adjusted for cohort effect. For the efficacy comparison of crisaborole versus TCS, only subjects from Cohort 1 are included in the analysis. For the comparison of crisaborole versus TCI, only subjects from Cohort 2 are included in the analysis.

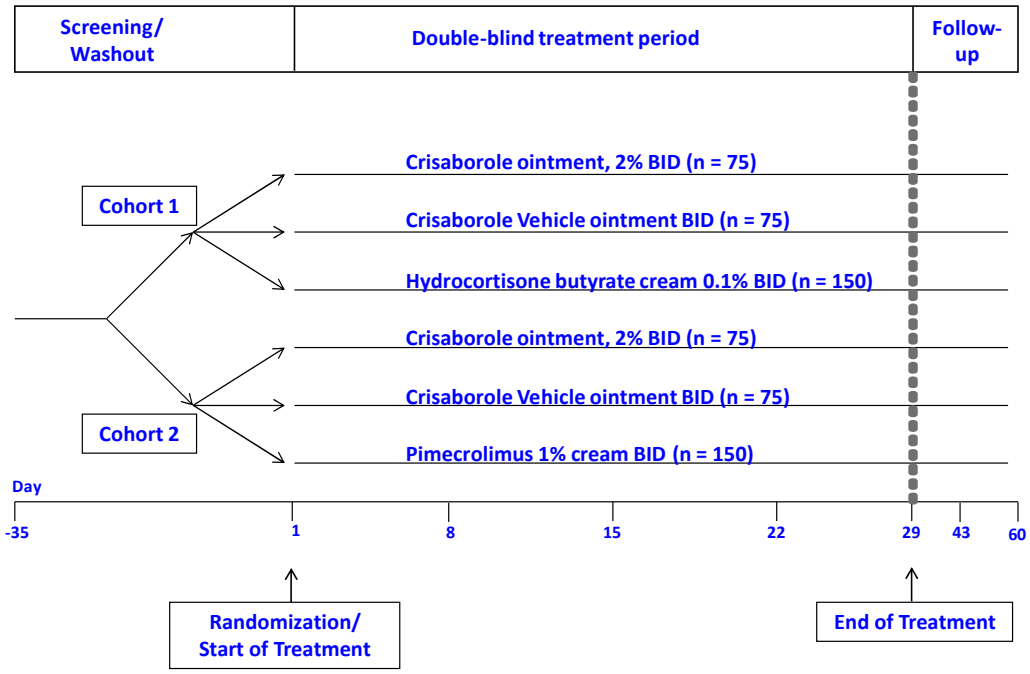
Safety and efficacy assessments will be conducted at the investigator site by a clinical assessor blinded for treatment arms.

A sub-study will be conducted at selected investigator sites to evaluate differences of changes in epidermal skin thickness as measured by Optical Coherence Tomography (OCT) between treatment groups in Cohort 1. Detailed statistical analysis for sub-study will be given in a separate SAP.

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) and Day 43 or 14 Days after last dose if subject is terminated early from treatment. A follow up telephone call will be made by site staff to the subjects/caregivers on Day 60 or at least 28 days after last dose if subject is terminated early from treatment. Subjects enrolled in the OCT sub-study will attend the clinic on Day 60 and do not require a telephone call on Day 60.

A schematic of the study design is shown in Figure 1.

Figure 1. Study Design Schematic



1

a. Treatment will be clinical assessor blinded for all treatment arms and double blinded for crisaborole ointment, 2% and vehicle treatment arms.

3. ENDPOINTS AND BASELINE VARIABLES: DEFINITIONS AND CONVENTIONS

3.1. Primary Endpoint(s)

3.1.1. Efficacy Endpoint

The primary efficacy endpoint is:

- Percent change from baseline in EASI total score at Day 29.

3.1.2. Safety Endpoints

The primary safety endpoints are:

- AEs, SAEs, local tolerability, discontinuations and clinically significant changes in vital signs and clinical laboratory parameters.

3.2. Secondary Endpoint(s)

3.2.1. Efficacy Endpoints

The secondary efficacy endpoints are:

- Percent change from baseline in EASI total score by scheduled time points except Day 29;
- Achievement of success in ISGA (score of clear (0) or almost clear (1) with at least a 2-grade improvement from baseline) by scheduled time points;
- Achievement of ISGA score of clear (0) or almost clear (1) by scheduled time points;
- Achievement of EASI75 by scheduled time points;
- Time to EASI75;
- Change from baseline in %BSA by scheduled time points.

3.2.2. Patient Reported Outcomes (PRO) Endpoints

The secondary PRO endpoints are:

(Note: The peak pruritus NRS (11-category numeric rating scale) is subject reported for 12 years and older subjects. A 5-category subject reported peak pruritus scale has been developed for subjects ≥6 and <12 years of age. The observer reported peak pruritus NRS will be completed by a caregiver for subjects <6 years old. The peak pruritus data will be analyzed separately for these three age groups):

- Change from baseline in peak pruritus NRS by scheduled time points for:
  - Subjects ≥12 years.
- Change from baseline in Patient Reported Itch Severity Scale for:
  - Subjects ≥6 and <12 years.
- Change from baseline in Observer Reported Itch Severity Scale for:
  - Subjects <6 years.
- Time to ≥2-point improvement from baseline in peak pruritus NRS for:
  - Subjects ≥12 years;
  - Subjects <6 years.
- Time to ≥3-point improvement from baseline in peak pruritus NRS for:
  - Subjects ≥12 years;
  - Subjects <6 years.
- Achievement of ≥2-point improvement from baseline in peak pruritus NRS by scheduled time points for:
  - Subjects ≥12 years;
  - Subjects <6 years.
- Achievement of ≥3-point improvement from baseline in peak pruritus NRS by scheduled time points for:
  - Subjects ≥12 years;
  - Subjects <6 years.
- Change from baseline in DLQI (for subjects ≥16 years) by schedule time points;
- Change from baseline in CDLQI (for subjects 4-15 years) by schedule time points;
- Change from baseline in DFI (completed by parent/caregiver of subjects 2-17 years) by schedule time points.

CCI

[redacted]

- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]

3.4. Baseline Variables

Demographic and baseline characteristics include:

- Age (years);
- Sex;

- Race;
- Ethnicity;
- Height (cm);
- Weight (kg);
- Body Mass Index (kg/m^2);
- Duration of disease (years);
- Prior treatment history (treatment naïve vs treatment [eg, TCS and/or TCI]);
- Geographic region;
- Country;
- ISGA;
- EASI;
- %BSA
- PROs (Peak Pruritus NRS, DLQI, CDLQI, DFI).

For multiple pre-treatment data besides peak pruritus and CCI baseline will be defined as the last evaluation taken before the first dose of investigational product. For peak pruritus and CCI the average of 7-day scores immediately prior to Day 1 (Day -6 to Day 1) will be used as the baseline.

3.5. Safety Endpoints

Safety will be assessed by the spontaneous reporting of AEs, physical examinations, 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.

3.5.1. Adverse Events

An adverse event is considered treatment-emergent adverse event (TEAE) to a given treatment if the event started during the effective duration of treatment. All events that start on or after the first dosing day and time/start time, if collected, but before end of study will be flagged as TEAEs.

Safety endpoints will be assessed by the spontaneous reporting of:

- TEAEs;
- Local tolerability AEs/SAEs;
- SAEs and AEs leading to discontinuation.

3.5.2. Laboratory Data

Below is a list of hematology and serum chemistry test parameters at Screening, Baseline/ Day 1 and Day 29/End of treatment/Early termination visit.

- Hematology: hemoglobin, hematocrit, red blood cell count, platelet count, white blood cell count (% and absolute for neutrophils, eosinophils, monocytes, basophils, lymphocytes).
- Serum chemistry: blood urea nitrogen/urea, glucose (non-fasting), 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 Examination

A detailed physical examination at Screening, Baseline/ Day 1 and Day 29 (End of treatment)/Early termination visit which will include but is 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.

4. ANALYSIS SETS

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

4.1. Full Analysis Set

The primary analysis population for efficacy data will be the Full Analysis Set (FAS) defined as all randomized subjects receiving at least one dose of investigational product. All efficacy and PRO endpoints will be analyzed based on the FAS.

For each specific endpoint, only participants who have data at ≥1 timepoint for that endpoint will be included in the analysis.

For a binary endpoint with a threshold requirement for change from baseline, only participants with a baseline value ≥ the threshold will be included in the analysis.

4.2. Safety Analysis Set

The Safety analysis set (SAF) is defined as those subjects who received at least one dose of the investigational product according to actual treatment received.

5. GENERAL METHODOLOGY AND CONVENTIONS

Final analyses will occur after database lock after Last Subject Last Visit (LSLV).

5.1. Hypotheses and Decision Rules

This study was terminated. The termination decision was made for business reasons only and was not related to any safety or efficacy concerns regarding crisaborole. Only 40% of planned participants were enrolled at termination. Due to small sample size, no hypotheses testing and decision rules. No statistical analysis and comparison will be performed. Only descriptive summary will be generated for safety data and some efficacy endpoints.

5.2. General Methods

In general, number and percent will be presented for binary and categorical variables. Number, mean, standard deviation, standard error of the mean, median, minimum, and maximum will be presented for continuous variables. In addition, graphics may be used to present the data – specific details will be outlined in the study List of Table (LOT).

Descriptive statistics will be provided for:

- Crisaborole ointment, 2% applied BID (Cohort 1);
- Crisaborole ointment, 2% applied BID (Cohort 2);
- Crisaborole ointment, 2% applied BID (combined Cohorts 1 and 2);
- Vehicle (combined Cohorts 1 and 2);
- Hydrocortisone butyrate cream, 0.1% BID (TCS, Cohort 1);
- Pimecrolimus cream, 1% BID (TCI, Cohort 2);
- Total – only for disposition, evaluation, demographic and baseline characteristics.

5.2.1. Analyses for Binary Data

Number and percentage of subjects will be summarized. Line plots of proportions may be provided.

5.2.2. Analyses for Continuous Data

Number, mean, standard deviation, standard error of the mean, median, minimum, and maximum will be presented. Line plots of means and standard errors may be provided.

5.2.3. Analyses for Categorical Data

The frequency and percentage for each category will be presented.

5.2.4. Analyses for Time to Event Data

Time to event endpoints will be summarized using the Kaplan-Meier (KM) method and estimated survival curves will be displayed graphically when appropriate. Graphs will describe the number of patients at risk over time. The median, quartiles, 95% confidence interval (CI) for median and quartiles will be estimated by the KM method.

5.3. Methods to Manage Missing Data

In general, for analyses using descriptive statistics, missing values will not be imputed. In addition, for safety endpoints, missing values will not be imputed. Other methods for handling missing values are discussed below.

5.3.1. Binary Endpoints

For binary endpoints analyzed at each scheduled visit separately, if a subject has no data for a binary endpoint at a scheduled visit, this subject will be classified as a non-responder (NR) for that endpoint at that visit.

5.3.2. Continuous Endpoints

Observed data will be used for all continuous endpoints.

5.3.3. Time to Event Endpoints

For time-to-event endpoints, subjects who complete the study without the event of interest or those who withdraw before experiencing the event of interest will have their event times right censored at the last available measurement time (or visit) used to define whether the subject experienced the associated event (ie, the event of interest is presumed to have occurred beyond this time point).

6. ANALYSES AND SUMMARIES

Due to study termination, only 40% of participants were enrolled. No statistical analysis and comparison will be performed. Only descriptive summary will be performed for selected efficacy and PRO endpoints.

Summary of selected efficacy and PRO endpoints will be based on FAS population. Section 5.2 provides treatment groups for summary.

6.1. Efficacy Endpoints

6.1.1. Percent Change from Baseline in EASI at Days 8, 15, 22, and 29

Descriptive statistics and graph will be provided for observed percent change from baseline in EASI.

6.1.2. Percent Change from Day 29 to Day 43 in EASI

Descriptive statistics will be provided for observed percent change from Day 29 to Day 43 in EASI.

6.1.3. ISGA of Success at Days 8, 15, 22, and 29

Achievement of success in the ISGA (defined as an ISGA score of clear (0) or almost clear (1) with at least a 2-grade improvement from baseline) at Days 8, 15, 22, and 29 will be summarized. Missing data will be classified as non-responder.

6.1.4. ISGA of Clear or Almost Clear at Days 8, 15, 22, and 29

Achievement of ISGA of clear (0) or almost clear (1) at Days 8, 15, 22, and 29 will be summarized. Missing data will be classified as non-responder.

6.1.5. EASI75 at Days 8, 15, 22, and 29

Achievement of EASI75 at Days 8, 15, 22, and 29 will be summarized. Missing data will be classified as non-responder. Graph depicting proportions will be provided.

6.1.6. Time to EASI75

Time to event endpoints will be summarized using the KM method and estimated survival curves will be provided for time to EASI75. The median, quartiles, 95% CI for median and quartiles be estimated by the KM method.

6.1.7. Change from Baseline in %BSA at Days 8, 15, 22, and 29

Descriptive statistics will be provided for observed change from baseline in %BSA.

6.2. PRO Endpoints

6.2.1. Change from Baseline in Weekly Average Peak Pruritus NRS/Scale

For peak pruritus NRS/Scale, weekly average score will be used in the analyses of change from baseline. Observed change from baseline in weekly average peak pruritus NRS/Scale at Days 8 (Week 1 average of Days 2-8), 15 (Week 2 average of Days 9-15), 22 (Week 3 average of Days 16-22), and 29 (Week 4 average of Days 23-29) for subjects ≥12 years will be summarized.

6.2.2. Change from Baseline in Weekly Average Patient Reported Itch Severity Scale

Due to the low enrollment, Patient Reported Itch Severity Scale will not be summarized.

6.2.3. Change from Baseline in Weekly Average Observer Reported Itch Severity Scale

Due to the low enrollment, Observer Reported Itch Severity Scale will not be summarized.

6.2.4. ≥2-Point Improvement in Weekly Average Peak Pruritus NRS

Achievement of ≥2-point improvement from baseline in weekly average peak pruritus NRS at Days 8 (Week 1 average of Days 2-8), 15 (Week 2 average of Days 9-15), 22 (Week 3 average of Days 16-22), and 29 (Week 4 average of Days 23-29) will be summarized for subjects ≥12 years. Missing data will be classified as non-responder.

6.2.5. ≥3-Point Improvement in Weekly Average Peak Pruritus NRS

Achievement of ≥3-point improvement from baseline in weekly average peak pruritus NRS at Days 8 (Week 1 average of Days 2-8), 15 (Week 2 average of Days 9-15), 22 (Week 3 average of Days 16-22), and 29 (Week 4 average of Days 23-29) will be summarized for subjects ≥12 years. Missing data will be classified as non-responder.

6.2.6. Change from Baseline in DLQI at Days 8, 15, 22, and 29

Descriptive statistics will be provided for observed change from baseline in DLQI for subjects ≥16 years.

6.2.7. Change from Baseline in CDLQI at Days 8, 15, 22, and 29

Descriptive statistics will be provided for observed change from baseline in CDLQI for subjects 4-15 years.

6.2.8. Change from Baseline in DFI at Days 8, 15, 22, and 29

Descriptive statistics will be provided for observed change from baseline in DFI completed by parent/caregiver of subjects 2-17 years.

CCI [redacted]

6.4. Subset Analyses

Summary of percentage change from baseline in EASI will be summarized by treatment and by cohort, age group, sex, race, baseline ISGA, and geographic region.

6.5. Baseline and Other Summaries and Analyses

6.5.1. Baseline Summaries

Demographics and baseline characteristics including variables defined in Section 3.4 will be summarized by treatment group according to Pfizer standards.

6.5.2. Study Conduct and Subject Disposition

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

6.5.3. Study Treatment Exposure

The exposure to study drug will be summarized by total number of applications, the total number of days of dosing, total amount of study drug used, and number and percentage of subjects who are compliant with the dosing regimen.

A subject will be considered compliant with the dosing regimen if he/she receives at least 45 but no more than 67 investigational product doses (ie, 80–120%, inclusive, of the expected number of doses) administered in accordance with the protocol.

6.5.4. Concomitant Medications and Non-drug 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

Safety analysis will be based on the SAF.

Safety data will be presented in tabular and/or graphical format and summarized descriptively, where appropriate. All safety endpoints will be listed and summarized in accordance with Pfizer Standards. Categorical outcomes (eg, AEs) will be summarized by subject counts and percentage. Continuous outcome (eg, blood pressure, pulse rate, etc) will be summarized using N, mean, median, standard deviation, etc. Change from baseline in laboratory data, and vital signs will also be summarized. Subject listings will be produced for these safety endpoints accordingly. Cohort 1 and cohort 2 will be combined for crisaborole and vehicle arms for all safety summaries and analyses.

6.6.1. Adverse Events

The safety data will be summarized in accordance with Pfizer Data Standards. All safety data will be summarized descriptively through appropriate data tabulations, descriptive statistics, categorical summaries, and graphical presentations. Safety endpoints for the study include:

- Treatment-emergent AEs and SAEs;
- Local tolerability AEs/SAEs;
- Withdrawals from treatment/study due to AEs.

6.6.2. Laboratory Data

Laboratory data will be listed and summarized in accordance with the Pfizer reporting standards.

6.6.3. Vital Signs

Vital signs will be summarized at baseline, Day 29/End of treatment/Early termination visits.

6.6.4. Physical Examination

Physical examinations will be summarized at baseline, Day 29/End of treatment/Early termination visits.

7. INTERIM ANALYSES

7.1. Data Monitoring Committee

This study uses an external data monitoring committee (E-DMC).

The E-DMC will be 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 a final decision. Pfizer will forward such decisions, which may include summaries of aggregate analyses of endpoint events and of safety data that are not endpoints, to regulatory authorities, as appropriate.

7.2. Interim Analyses and Summaries

There is no plan for an interim analysis.

8. REFERENCES

None.

9. APPENDICES

Appendix 1. Definition and Use of Visit Windows in Reporting

Visit windows will be used for efficacy variables, and for any safety data that display/summarize by study visit.

Visit Label	Target Day	Definition [Day window]
Screening		Days -35 to Day -1
Baseline	Day 1, Baseline	Day 1*
Day 8	8	Days 2 to 11
Day 15	15	Days 12 to 18
Day 22	22	Days 19 to 25
Day 29/ End of Treatment/ Early Termination	29	Days 26 to last dose +7
Day 43	43	Last dose +8 to end of study

* The baseline of Peak pruritus NRS/scale and CCI is the average of 7-day scores immediately prior to Day 1 (Day -6 to Day 1)

For the lab values, if the calculated study Day for the labelled baseline visit is not study Day 1 but falls within 28 days before the start of the study dosing, then that data should be used for the baseline instead of leaving baseline missing.

For the other values, if the calculated study day for the labelled baseline visit is not study Day 1, but falls before the start of the study dosing, then that data should be used for the baseline instead of leaving baseline missing.

If two or more visits fall into the same window, keep the one closest to the Target Day. If two visits are equaled distant from the Target Day in absolute value, the later visit should be used.

Safety analysis may follow Pfizer standards.

Appendix 2. Investigator's Static Global Assessment (ISGA)

The ISGA, a five-point global assessment of AD severity, will be assessed at times specified in the Study Procedure section of study protocol to characterize subjects' overall disease severity across all treatable AD lesions (excluding the scalp).

The ISGA will be a static evaluation without regard to the score at a previous visit. It must be completed by a clinical assessor blinded for treatment arms. 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.

ISGA Score

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/papulation with or without oozing/crusting
4	Severe	Deep or bright red erythema with severe induration/papulation and with oozing/crusting

∗ The ISGA will exclude scalp from the assessment/scoring.

Appendix 3. Eczema Area and Severity Index (EASI)

The EASI quantifies the severity of a subject's AD 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.

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 the table below.

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		

%BSA with Atopic Dermatitis: 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.

Handprint Determination of Body Region Surface Area for Subjects ≥8 Years

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

Handprint Determination of Body Region Surface Area for Subjects <8 Years

Body Region	Total Number of Handprints in Body Region*	Surface Area of Body Region Equivalent of One Handprint
Head and Neck	20	5%
Upper Limbs	20	5%
Trunk (including axillae)	30	3.33%
Lower Limbs (including buttocks)	30	3.33%

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.

Eczema Area and Severity Index (EASI) Area Score Criteria

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

Body Region Weighting: Each body region is weighted according to its approximate percentage of the whole body.

EASI Body Region Weighting

Body Region	Body Region Weighting for Subjects ≥8 Years	Body Region Weighting for Subjects <8 Years
Head and Neck	0.1	0.2
Upper Limbs	0.2	0.2
Trunk (including axillae)	0.3	0.3
Lower Limbs (including buttocks)	0.4	0.3

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 equations below.

Equation 1 (subjects ≥8 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)

Equation 2 (subjects 2-<8 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)

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 0.1 and range from 0.0 to 72.0, with higher scores representing greater severity of AD. Since the scalp will be excluded from the EASI assessment in this study, the maximum possible score will be less than 72.0.

Appendix 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 skin created problems with your **partner** or any of your **close friends** or **relatives**?	Very much 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 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, **school time?** **OR** was it **holiday time?**	**If school time**: Over the last week, how much did your skin problem affect your **school work**? **If holiday time**: How much over the last week, has your skin problem interfered with your enjoyment of the **holiday**?	Prevented school Very much Quite a lot Only a little Not at all 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 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/careers.	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/careers.	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 career and partner** or between the **main career and other children** in the family.	Very much A lot A little Not at all

10.	Over the last week, how much effect has **helping with your child's treatment** had on the main career'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**

CCI

[redacted]

CCI

[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]

[redacted]

[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]

CCI	

Appendix 8. Listing of Abbreviations

Abbreviation	Term
AD	Atopic Dermatitis
AE	Adverse Event
BID	Twice Daily
%BSA	Percent 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
EASI75	Eczema Area And Severity Index ≥75% improvement from Baseline
E-DMC	External Data Monitoring Committee
CCI [redacted]	[redacted]
FAS	Full Analysis Set
CCI [redacted]	[redacted]
ISGA	Investigator'S Static Global Assessment
KM	Kaplan-Meier
LOT	List of Table
LSLV	Last Subject Last Visit
CCI [redacted]	[redacted]
NR	Non-responder
NRS	Numeric Rating Scale
OCT	Optical Coherence Tomography
CCI [redacted]	[redacted]
PRO	Patient Reported Outcome
SAE	Serious Adverse Event
SAF	Safety Analysis Set
SAP	Statistical Analysis Plan
TCI	topical calcineurin inhibitor
TCS	topical corticosteroid
TEAE	Treatment-Emergent Adverse Event