PROTOCOL B7451037

A PHASE 2A RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, PARALLEL GROUP, MULTI-CENTER STUDY TO INVESTIGATE THE MECHANISM OF ACTION OF PF-04965842 MONOTHERAPY IN ADULT PARTICIPANTS WITH MODERATE TO SEVERE ATOPIC DERMATITIS

STATISTICAL ANALYSIS PLAN (SAP)

Version: 3

Date: 8 SEP 2021

[redacted]

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

APPENDICES

1. VERSION HISTORY

This is the first amendment of the Statistical Analysis Plan (SAP) for Study B7451037 and is based on the final protocol dated 12NOV2020.

Table 1. Summary of Changes

Version/Date	Associated Protocol Amendment	Rationale	Specific Changes
1 15 Oct 2019	Protocol Amendment 1 05 Sep 2019	N/A	N/A
2 29 Mar 2021	Protocol Amendment 3 12 Nov 2020	Due to lack of assay of sufficient quality to study immature fraction.	Section 2.1, 3.3.1, 4.3.1: Deleted "immature platelet fraction" and changed "platelet volume" to "platelet count, red blood cell count and reticulation".
		For clarification and to add metabolites.	Section 2.1: Updated PK objectives/endpoints.
		Suggestion by SciOps for clarification.	Section 5.1: Added text to require baseline and a 12-week assessment.
		To allow for flexibility in analytical device.	Section 5.2.2: Deleted text "by Illumina HiSeq 2500".
		An interim analysis has been added.	Section 7: Added text for interim analysis on the first 15 completers of the trial.
		Hypothetical estimands are not the best approach in this case.	Section 5.1.1 and 5.2.1: Deleted all references to a hypothetical estimand.
		Align with supplemental SAP.	Section 5.2.2: Remove RNA analyses since they will be part of the supplementary biomarker SAP.
3 8 Sep 2021	Protocol Amendment 3 12 Nov 2020	Clarify dose compliance derivation.	Section 6.7.3: Added detailed derivation for dose compliance.

2. INTRODUCTION

This SAP provides the detailed methodology for summary and statistical analyses of the data collected in Study B7451037. 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

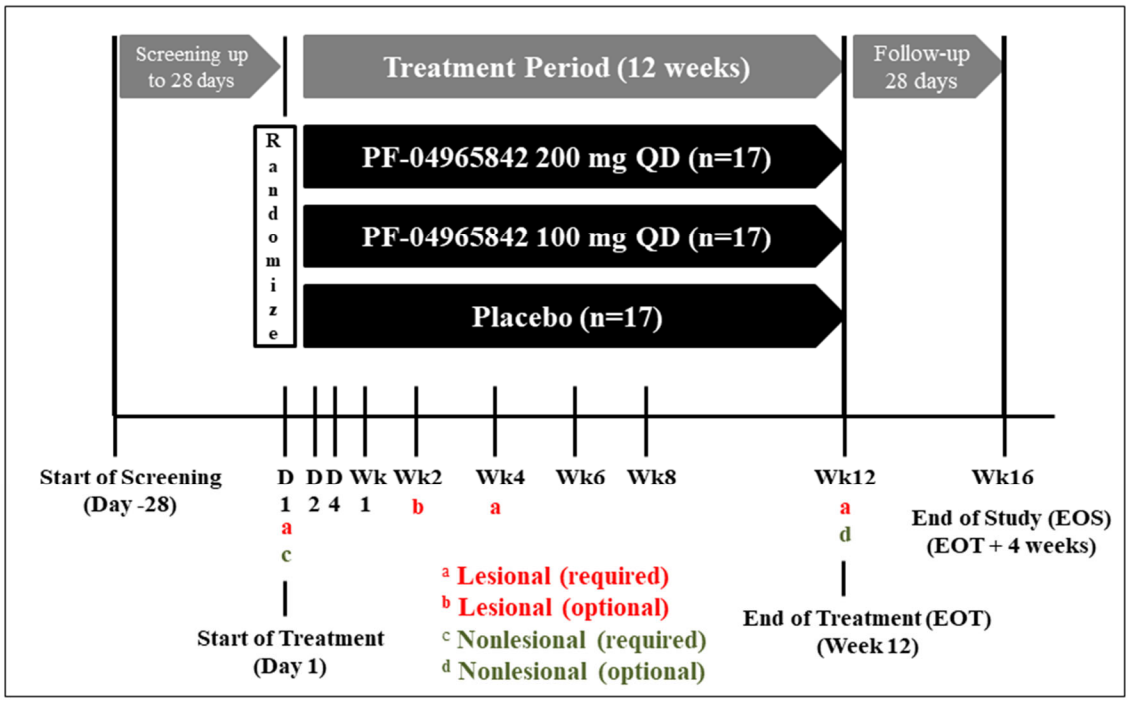
Primary Objective:	**Primary Endpoint:**
To assess the PF-04965842on lesional and non-lesional skin biomarkers of adult participants with moderate to severe atopic dermatitis (AD).	• Changes from baseline in AD biomarkers in skin, including biomarkers for general inflammation (MM12), hyperplasia (K16), Th2 immune response (CCL17, CCL18, CCL26) and Th22 immune response (S100A8, S100A9, S100A12).
Secondary Objectives:	**Secondary Endpoints:**
To evaluate the effect of PF-04965842 on gene expression (evaluated by mRNA PCR and/or gene arrays).	• Changes from baseline in gene expression (evaluated by mRNA PCR and/or gene arrays) over time in skin lesions.
To evaluate the effect of PF-04965842 on inflammatory infiltrates (T-cell and dendritic cells) using immunohistochemistry IHC) in skin biopsies.	• Changes from baseline in cellular (T-cells and dendritic cells) markers using IHC in skin biopsies at various time points.
To evaluate the effect of PF-04965842 on epidermal hyperplasia using IHC and RT PCR in skin biopsies.	• Changes from baseline in hyperplasia markers (thickness, K16, Ki67) in skin biopsies.
To assess the effect of PF-04965842 on blood biomarkers (OLINK proteomic in serum).	• Changes from baseline in blood biomarkers (OLINK proteomic microassay for inflammation and immune response in serum).
To assess changes in T-cell lymphocyte subset populations in blood using flow cytometry.	• Changes from baseline in T-cell lymphocyte subset populations in blood using flow cytometry.
To evaluate the effect of PF-04965842 on pruritus and its correlation to IHC and genetic markers in lesional skin.	• Response based on at least 4 points improvement in the severity of Peak Pruritus numerical rating scale (NRS) from baseline and change from baseline in IHC and genetic markers in lesional skin.
Efficacy Objectives:	**Efficacy Endpoints:**
To evaluate the effect of PF-04965842 on clinical efficacy outcomes.	• Response based on the Investigator's Global Assessment (IGA) score of clear (0) or almost clear (1) on a 5 point scale) and a reduction from baseline of ≥2 points at all scheduled time points. • Response based on the Eczema Area and Severity Index ≥75% improvement from baseline (EASI-75) response at all scheduled time points. • Response based on at least 4 points improvement in the severity of Peak Pruritus NRS from baseline at all scheduled time points. • Response based on a ≥50% and ≥90% improvement in the EASI total score (EASI-50 and EASI-90) at all scheduled time points.

	• Change from baseline in the percentage Body Surface Area (BSA) affected at all scheduled time points.
Safety Objectives	**Safety Endpoints**
To evaluate the safety and tolerability of PF-04965842 during 12 weeks of treatment.	• Incidence of treatment-emergent adverse events (TEAEs). • Incidence of serious AEs (SAEs). • Incidence of AEs leading to discontinuation. • Incidence of clinical abnormalities and changes from baseline in clinical laboratory values, and vital signs.
Tertiary/Exploratory Objectives	**Tertiary/Exploratory Endpoints**
CCI [redacted]	[redacted]
To assess the effect of PF-04965842 on Night Time Itch.	• Changes from baseline in the Night Time Itch Scale at all scheduled time points.
CCI [redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]

2.2. Study Design

This is a randomized, double-blind, placebo-controlled, parallel-group, Phase 2a study to investigate the mechanism of action of PF-04965842 by correlating efficacy outcomes with changes from baseline in key skin and blood biomarkers in adult participants ≥18 years of age with moderate-to-severe AD. Participants will be screened within 28 days prior to the first dose of study intervention to confirm study eligibility. A total of approximately 51 participants will be randomized in a 1:1:1 ratio to receive PF-04965842 200 mg QD, PF-04965842 100 mg QD, or matching placebo QD for 12 weeks. A study design schematic is presented in Figure 1.

Figure 1 Study Design Schematic



At the end of the 12-week study treatment, qualified participants will have the option to enter the long-term extension (LTE) Study B7451015. Participants discontinuing early from this study will undergo a 4-week off-treatment follow-up period.

Eligible participants will be randomized into 3 intervention groups in the study:

Group 1 (N=17): PF-04965842 200 mg QD taken orally for 12 weeks.

Group 2 (N=17): PF-04965842 100 mg QD taken orally for 12 weeks.

Group 3 (N=17): placebo QD taken orally for 12 weeks.

The total duration of participation in the study is up to 20 weeks, including up to 4 weeks for screening, 12 weeks study intervention, and a follow-up period of 4 weeks after study intervention (for those participants who do not enter the LTE study).

A full schedule of activities for the study is provided in Appendix 7.

Sample Size Determination

The sample size of this study is based on the objective to evaluate changes in key AD biomarkers in skin lesions, including biomarkers for general inflammation (MMP12),

hyperplasia (K16), Th2 (CCL17, CCL18, CCL26), and Th22 (S100A8, S100A9, S100A12) with PF 04965842 200 mg QD, PF 04965842 100 mg QD, and placebo.

A total sample of approximately 51 participants, with 17 participants randomized to PF-04965842 200 mg QD, 17 participants randomized to PF-04965842 100 mg QD, 17 participants randomized to matching placebo (1:1:1 randomization) is planned.

A sample size of 17 participants per treatment group would provide about 92% chance that the 95% CI for the mean fold-change (post-baseline relative to baseline) has a half-width of no more than 3.2 (which is within 64% of an assumed maximal standard deviation of 5 for the fold-change, as based on published data). Calculations are based on the assumption of a chi-square distribution for the variance so that the width of the estimated confidence interval is a random quantity.

3. ENDPOINTS AND BASELINE VARIABLES: DEFINITIONS AND CONVENTIONS

3.1. Primary Endpoints

The primary endpoints are:

- Changes from baseline in AD biomarkers in skin, including biomarkers for general inflammation (MMP12), hyperplasia (K16), Th2 immune response (CCL17, CCL18, CCL26), and Th22 immune response (S100A8, S100A9, S100A12).

3.2. Secondary Endpoints

The secondary endpoints are:

- Changes from baseline in gene expression (evaluated by mRNA PCR and/or gene arrays) over time in skin lesions;
- Changes from baseline in cellular (T cells and dendritic cells) markers using IHC in skin biopsies at various time points;
- Changes from baseline in hyperplasia markers (thickness, K16, Ki67) in skin biopsies;
- Changes from baseline in blood biomarkers (OLINK proteomic microassay for inflammation and immune response in serum);
- Changes from baseline in T cell lymphocyte subset populations in blood using flow cytometry;
- Change from baseline in IHC and genetic markers in lesional skin.

A detailed description of how the pruritus NRS score is derived is provided in Appendix 3.

3.3. Secondary Efficacy Endpoints

The secondary efficacy endpoints are:

- Response based on the Investigator's Global Assessment (IGA) score of clear (0) or almost clear (1) (on a 5 point scale) and a reduction from baseline of ≥2 points at all scheduled time points;
- Response based on the Eczema Area and Severity Index ≥75% improvement from baseline (EASI 75) response at all scheduled time points;
- Response based on at least 4 points improvement in the severity of Peak Pruritus NRS from baseline at all scheduled time points;
- Response based on a ≥50% and ≥90% improvement in the EASI total score (EASI-50 and EASI-90) at all scheduled time points;
- Change from baseline in the percentage Body Surface Area (BSA) affected at all scheduled time points.

3.3.1. Tertiary/Exploratory Endpoints

- CCI [redacted]
- Changes from baseline in Night Time Itch Scale at all scheduled time points;
- CCI [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]

3.4. Baseline Variables

In general, for all analyses, baseline will be defined based on observations collected prior to first dose. Baseline values for demographics, medical and other history, atopic dermatitis history will be based on measures collected at Visit 1/Screening visit. Study Day 1 is defined as the day the subject receives first dose of study drug. For purposes of all other analyses including analyses for change from baseline, the baseline value will be defined as measured on Day 1 (before time of first dose, if time

is available). If a value is missing on Day 1, then the last available observation before Day 1 will be used.

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. Endpoints will be assessed as:

- Incidence of treatment emergent adverse events (AEs) and serious AEs (SAEs).
- AEs leading to discontinuation.
- Serious infections, defined as any infection (viral, bacterial, and fungal) requiring hospitalization or parenteral antimicrobials or met other criteria that required the event be classified as serious;
- Incidence of clinical abnormalities and change from baseline in selected clinical laboratory values, ECG measurements, and vital signs.

The safety endpoints will be defined in accordance with Clinical Data Interchange Standards Consortium (CDISC) aligned Pfizer Standards (CaPS).

3.5.1. Adverse Events

An adverse event will be considered a Treatment-Emergent Adverse Event (TEAE) 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 the last dose plus the lag time (28 days) will be flagged as TEAEs. The algorithm will not consider any events that started prior to the first dose date.

3.5.2. Laboratory Data

Below is a list of hematology and serum chemistry test parameters.

- Hematology: hemoglobin, hematocrit, red blood cell count, red blood cell indices (MCV, MCH, MCHC, RBC morphology), reticulocyte count, platelet count, white blood cell count with differential, total neutrophils, eosinophils, monocytes, basophils, lymphocytes, coagulation panel.
- Clinical chemistry: blood urea nitrogen, creatinine, creatine phosphokinase, glucose (non-fasting, Day 1 and Week 12 fasting), sodium, potassium, chloride, calcium, total bicarbonate, aspartate aminotransferase, alanine aminotransferase, gamma-glutamyl transferase, bilirubin, alkaline phosphatase, lactate dehydrogenase, uric acid, albumin, total protein, total cholesterol, low-density lipoprotein (LDL), high-density lipoprotein (HDL), triglycerides.

Below is a list of urinalysis test parameters

- pH, glucose, protein, blood, ketones, nitrite, leukocyte esterase by dipstick;
- Microscopic examination and/or culture.

3.5.3. Vital Signs, including Height and Weight

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

Height is collected at screening and weight is collected at pre- and post-treatment.

3.5.4. Physical Examinations

Complete physical examinations consist of assessments of general appearance; skin; head, eyes, ears, nose and throat; mouth, heart; lungs; breast (optional); abdomen; external genitalia (optional); extremities; neurologic function; and lymph nodes. A targeted physical examination will include, at a minimum, assessments of the skin, heart, lungs, and abdomen and examination of body systems where there are symptom complaints by the participant.

4. ANALYSIS SETS

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

4.1. Full Analysis Set

The Full Analysis Set (FAS) is defined as all subjects who have been randomized and who have received at least one dose study of treatment. Subject is assigned to the randomized treatment group regardless of actual treatment received. Analyses for endpoints that are defined based on a threshold of change from baseline (eg, NRS4) will also require the baseline value to be equal to or greater than that threshold (eg, for NRS4, the baseline value needs to be ≥4).

4.2. Per Protocol Analysis Set

The Per Protocol Analysis Set (PPAS) is defined as a subset of FAS who had no major protocol violations. The subjects excluded from the PPAS will be determined and documented before the study is unblinded. Supportive analyses for the primary endpoint will be performed using the PPAS. This set will include subjects who:

- Were eligible for the study by way of meeting key inclusion criteria and none of the key exclusion criteria.
- Had valid and non-missing biopsies at baseline and per protocol schedule.

- Did not take a protocol-prohibited medication for the primary diagnosis prior to completion of the study dosing period (Visit 10/Week 12).
- Took the correct randomized treatment for at least 80% and at most 120% of the assigned amount until completion of the study dosing period (Visit 10/Week 12).
- Had no other major protocol violations as determined by the clinical team prior to database lock. A major protocol violation in this context is one that is likely to affect materially the efficacy responses of the patient and will be defined by the clinical team before database is locked and any analysis is performed for this study.

These items will either be assessed by programmed checks of the data or be determined by clinical review prior to unblinding of study treatment.

4.3. Safety Analysis Set

The Safety Analysis Set (SAF) will be defined as all subjects who receive at least one dose of study medication classified according to actual study treatment received. The safety analysis set is the primary population for treatment administration / compliance and safety. A randomized but not treated subject will be excluded from the safety analyses.

5. GENERAL METHODOLOGY AND CONVENTIONS

The final analysis and reporting of results will be performed after the completion of the study and the database is locked.

5.1. Hypotheses and Decision Rules

There will be no formal hypothesis testing, but the point estimates and 95% two-sided confidence intervals will be reported for the biomarker endpoints, secondary efficacy CCI endpoints. No adjustments for multiplicity will be performed. Safety endpoints will be summarized descriptively through appropriate data tabulations, descriptive statistics, categorical summaries, and graphical presentations.

5.1.1. Estimands

The primary objective of this study is to assess the effects of PF-04965842 on lesional and non-lesional skin biomarkers of adult participants with moderate-to-severe AD. A hypothetical estimand will be used for estimation. The following estimand attributes will be considered:

- Population: Participants with moderate-to-severe AD as defined by the inclusion criteria and with a valid baseline and Week 12 assessment;

- Variable: Fold-change from baseline in the biomarkers for general inflammation (MMP12), hyperplasia (K16), Th2 immune response (CCL17, CCL18, CCL26), and Th22 immune response (S100A8, S100A9, S100A12);

- Intercurrent event: All data collected will be utilized;

- Population-level summary: Mean fold-change from baseline within each treatment group.

Some secondary objectives require analyses of binary endpoints for which, a composite estimand with the following attributes will be considered:

- Population: Participants with moderate-to-severe AD as defined by the inclusion criteria and with a valid baseline;

- Variable: Binary outcome measures such as responder-based endpoints. Participants who discontinue from the study treatment for any reason are defined as non-responders after that point;

- Intercurrent event: The intercurrent event is captured through the variable definition;

- Population-level summary: Proportion of participants who are responders in each treatment group and differences in proportions of responders between each active dose and placebo.

5.2. General Methods

In general, for descriptive analyses, number and percent will be presented for binary endpoints. Number, median, Q1 and Q3 will be presented for descriptive analyses of continuous efficacy endpoints.

5.2.1. Analysis of Primary Endpoint

The population for the primary analysis will be all participants in the FAS having valid biomarker data at baseline and Week 12.

The primary endpoints in this study include mean fold-changes for the following biomarkers:

- General inflammation (MMP12);

- Hyperplasia (K16);

- Th2 (CCL17, CCL18, CCL26);

- Th22 (S100A8, S100A9, S100A12).

In each participant, expression levels from RT-PCR will be normalized to the housekeeping gene RPLP0 (validated in other studies) by negatively transforming the Ct values to –dCt, obtaining an equivalent to log2 scale expression values (reporting unit: log2 [Expression/RPLP0]). The expression values below detection will be estimated for each gene as the 20% of the minimum of unlogged values across all samples (which is equivalent to subtracting 2.321928 from the –dCt minimum value). The expression levels below detection may provide important information about some markers that significantly change with treatment (eg, the expression may decrease below the level of detection after treatment), therefore this imputation is an important step in the analysis.

RT-PCR expression data in log2 scale will be modeled using a mixed effect model with Treatment, Time Point, and Tissue Type as a fixed effect and a random intercept for each participant. This model will yield unbiased estimates and valid inferences in the presence of a missing data mechanism that is missing at random (MAR).[1] This approach introduces less bias than restricting the analysis for participants with complete observations. Contrasts will be used to estimate the log2 fold-changes for each treatment group. Fold-changes will be computed by obtaining the antilog of log2 fold-changes. Point estimates and 95% confidence intervals (CIs) will be obtained for the fold-change from baseline in the biomarker values for each treatment group based on this model.

Correlation analyses will be performed to assess if the fold-change from baseline in biomarkers correlates with change (from baseline) in the clinical efficacy endpoints such as IGA, EASI, and Peak Pruritus NRS.

Supportive analyses for the primary endpoint will also be performed using the PPAS.

While statistical tests of hypotheses will not be performed to compare the fold-change from baseline in the PF-04965842 groups versus placebo, the point estimates and 95% CIs for the two PF-04965842 dose groups will be assessed in relation to those from the placebo group. A fold-change from baseline observed in the PF-04965842 groups alone (and correlated with efficacy) but not in the placebo group is likely due to the mechanism of action of PF-04965842.

5.2.2. Analysis of Other Biomarker Endpoints

In addition to the key biomarkers considered as primary endpoints, the other biomarkers endpoints in this study include (but may not be limited to) the following:

Parameters studied using IHC will include epidermal thickness (performed on H&E sections), Ki 67+ cells, Keratin 16 (K16), CCI [redacted].

CCI [redacted]

CCI

CCI

CCI

Gene expression by TLDA RT-PCR will include the levels of the following established cytokines associated with AD: CCI

S100A7, S100A8, S100A9, S100A12,
CCI

CCL18,
CCI CCL26, CCI 18S and
RPLP0 as control genes.

CCI

The mixed-effect model approach chosen for the analysis of RT-PCR and IHC data has the advantage that estimation of the parameters in the models is unbiased and valid inferences are possible in the presence of a missing data mechanism that is MAR. This approach introduces less bias than restricting the analysis for participants with complete observations.

5.2.3. Analyses of Other Endpoints

The responses based on IGA, EASI (eg, EASI-75), and severity scale of the Peak Pruritus NRS will be summarized descriptively by visit and treatment. The difference of the improvement from baseline between each PF-04965842 dose and placebo will also be derived.

If a participant withdraws from the study, then this participant will be considered a non-responder after withdrawal for any binary valued clinical efficacy endpoint.

CCI [redacted]

CCI [redacted]

5.3. Methods to Manage Missing Data

In general, for analyses using descriptive statistics, missing values will not be explicitly imputed. In addition, for safety endpoints, missing values will not be imputed.

For the responses based on IGA, EASI (eg, EASI-75), and severity scale of the Peak Pruritus NRS, if a participant withdraws from the study, then this participant will be considered a non-responder after withdrawal for any binary valued clinical efficacy endpoint.

6. ANALYSES AND SUMMARIES

Point estimates and 95% confidence intervals will be provided for each of the PF-04965842 QD treated groups (200 mg and 100 mg) and placebo.

A summary of analyses for clinical efficacy endpoints is provided in Appendix 1. Visit windows to be used for all efficacy analyses and some relevant safety analyses are detailed in Appendix 2.

Data collected at Week 16 will be displayed in listings only, and will not be part of any analyses, unless specifically noted otherwise.

6.1. Primary Endpoint

The primary endpoints in this study include mean fold-changes for the following biomarkers:

- General inflammation (MMP12);

- Hyperplasia (K16);
- Th2 (CCL17, CCL18, CCL26);
- Th22 (S100A8, S100A9, S100A12).

6.1.1. Primary Analysis

- Summary: Fold-change from baseline at Week 12;
- Population: FAS;
- Statistical Method: Estimate and 95% CI from MMRM described in Section 5.2.1;
- Missing Data: No explicit imputation.

6.1.2. Additional/Supportive Analysis

- Summary: Fold-change from baseline at Week 12;
- Population: PPAS;
- Statistical Method: Estimate and 95% CI from MMRM described in Section 5.2.1;
- Missing Data: No explicit imputation.

6.2. Other Biomarker Endpoints

The other biomarker endpoints in this study include mean fold-changes for the following biomarkers:

- epidermal thickness (performed on H&E sections);
- Ki 67+ cells,
- Keratin 16 (K16);
- CCI [redacted]

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

6.2.1. Primary Analysis

- Summary: Fold-change from baseline at Week 12;
- Population: FAS;
- Statistical Method: Estimate and 95% CI from MMRM described in Section 5.2.2;
- Missing Data: No explicit imputation.

6.3. Efficacy Endpoints

6.3.1. NRS4 for Severity Response

- Summary: Proportion of subjects with NRS4 for severity response at Weeks 4, 8 and 12;
- Population: Subjects from the FAS with a baseline NRS score for severity ≥4;
- Statistical Method: Descriptive (n, %), 95% CI and difference from placebo. Visit windows will be used as described in Appendix 2;
- Missing Data: Missing data arising due to subject dropout are considered "non-response".

6.3.2. IGA Response

- Summary: Proportion of subjects with IGA score of clear (0) or almost clear (1) and a reduction from baseline of ≥2 points at Weeks 2, 4, 8 and 12;
- Population: Subjects from the FAS with a baseline IGA score ≥2;
- Statistical Method: Descriptive (n, %). Visit windows will be used as described in Appendix 2;
- Missing Data: Missing data arising due to subject dropout are considered "non-response".

6.3.3. EASI50, EASI75 and EASI90 Response

- Summary: Proportion of subjects with EASI response at Weeks 2, 4, 8 and 12;
- Population: FAS;
- Statistical Method: Descriptive (n, %). Visit windows will be used as described in Appendix 2;
- Missing Data: Missing data arising due to subject dropout are considered "non-response".

6.3.4. Change from Baseline in BSA%

- Summary: Change from baseline at all scheduled visits;
- Population: FAS;
- Statistical Method: Descriptive (LSmean, SD, mean difference, SE, 95% CI, median, Q1, Q3);
- Missing Data: Observed Data.

CCI

C [redacted]

6.4.2. Night Time Itch Scale

- Summary: Change from baseline in Night Time Itch Scale at all scheduled visits;
- Population: FAS;
- Statistical Method: Descriptive (mean, SD, median, Q1, Q3);
- Missing Data: Observed data.

CCI [redacted]

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

CCI [redacted]

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

6.5. PK Endpoints

PK concentration data for PF-04965842 and its metabolites will be summarized through appropriate data tabulations, descriptive statistics, and graphical presentation. A population PK model will be developed for the purpose of estimating PK parameters. Additional details of the methodology will be captured in a separate modeling plan and the results will also be reported separately.

6.6. Subset Analyses

Summary statistics for co-primary endpoints will be presented by subgroups below.

- Age (years) group; (<40, ≥40);

- Race (White, Black or African-American, Asian, Other[1]);
- Baseline disease severity (moderate, severe).

Point estimates and 95% confidence intervals will be derived from MMRM models as described in Section 5.2.1.

The primary purpose of the subgroup analyses is to check for consistency of results across subgroups, to make sure overall results are not being driven by some subset of subjects.

Graphical display (eg, forest plots) of the differences between treatment groups will be presented. There is no intention to have any specific inference within subgroups.

6.7. Baseline and Other Summaries and Analyses

6.7.1. Baseline Summaries

Demographics, medical history, primary diagnosis, history of prior AD treatments and disease characteristics will be summarized by treatment group according to CaPS.

6.7.2. Study Conduct and Subject Disposition

Subjects' evaluation, disposition and discontinuation will be summarized according to CaPS.

6.7.3. Study Treatment Exposure

Subjects were expected to take 2 tablets of 100 mg per day (on a QD schedule) during the 12 weeks treatment period. The study drug compliance is defined as:

Drug compliance (%) = 100×Actual number tablets taken/Expected number of tablets,

where

Actual number tablets taken = Total number of dispensing tablets [at Week 0 (Day 1), 4, and 8] – Total number of return tablets (at Weeks 4, 8, and 12);

Expected number of tablets = 2×Days of treatment duration (Last treatment date – First treatment date +1).

The study compliance (%) and study durations (days) will be summarized for each treatment group.

[1] For purposes of analysis, Other will comprise the categories of American Indian or Alaska Native, Hawaiian or Other Pacific Islander, Multi-Racial and any other category reported on the CRF.

6.7.4. Concomitant Medications and Non-drug Treatments

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

6.8. Safety Summaries and Analyses

Safety analyses will be based on the SAF analysis set.

All clinical AEs, SAEs, TEAEs, withdrawal due to AEs, ECGs, vital signs and safety laboratory data will be reviewed and summarized on an ongoing basis during the study to evaluate the safety of subjects.

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 CaPS. 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, ECGs and vital signs will also be summarized. Subject listings will be produced for these safety endpoints accordingly.

6.8.1. Adverse Events

The safety data will be summarized in accordance with CaPS. 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;
- Withdrawals from active treatment due to AEs;
- Serious infections, defined as any infection (viral, bacterial, and fungal) requiring hospitalization or parenteral antimicrobials.

AEs will be displayed by MedDRA system organ class (SOC).

It should be recognized that most studies are not designed to reliably demonstrate a causal relationship between the use of a pharmaceutical product and an adverse event or a group of adverse events. Except for select events in unique situations, studies do not employ formal adjudication procedures for the purpose of event classification. C
Cl
[redacted]

6.8.2. Laboratory Data

Laboratory data will be listed and summarized in accordance with the CaPS reporting standards. Summaries of subjects meeting pre-specified monitoring and discontinuation criteria will be created using methods for categorical data.

6.8.3. Vital Signs

Vital signs will be summarized at baseline, Weeks 2, 4, 8, and 12/End of Treatment visits. Height and weight will be summarized at baseline (both) and End of Treatment (weight only) visits.

6.8.4. Electrocardiogram

ECG parameters, if applicable, will be summarized at baseline

6.8.5. Physical Examination

Physical examinations will be summarized at baseline and all-available post-baseline visits.

The C-SSRS will also be collected and reported using descriptive summaries for recent history/baseline and post-baseline visits.

7. INTERIM ANALYSES

After the first 15 participants have completed the trial, there will be an interim analysis to examine platelet values. A longitudinal analysis will be conducted to determine if there is a change from baseline in platelet values. The data will remain blinded so that no one will know the participants' assigned treatment and thus the analysis will not look for any treatment differences. Correlation analyses may also be conducted to look for potential relationships between platelet values and other laboratory measures, such as TPO. There will not be any decision rules associated with this interim analysis, so there will be no changes to the conduct or continuation of the trial based on these analyses.

There will be a program-level E-DMC (see below) who will periodically review the safety data from the study.

This study will use an E-DMC. The E-DMC is independent of the study team and includes only external members. The E-DMC will be responsible for ongoing monitoring of the safety of participants 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 endpoint events and of safety data that are not endpoints, to regulatory authorities, as appropriate. Composition of the E-DMC and processes under which the E-DMC operates will be documented in the E-DMC charter.

8. REFERENCES

1. R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

intentionally blank

9. APPENDICES

Appendix 1. Summary of Clinical Efficacy Analyses

Efficacy Endpoints	Population	Analysis Method	Missing Data Imputation	Primary Analysis for Co-primary
Week 12 IGA Response	FAS	CMH	NR	Yes
	PPAS	CMH	NR	No
Week 12 EASI-75 Response	FAS	CMH	NR	Yes
	PPAS	CMH	NR	No
Weeks 2, 4, 8 and 12 NRS4 for severity Response	FAS	CMH	NR+MI	
	PPAS	CMH	NR+MI	
	FAS	CMH	NR	
Time to NRS4 for severity	FAS	Kaplan-Meier	OD	
Weeks 2, 4 and 8 IGA Response	FAS	CMH	NR	
Weeks 2, 4, 8 and 12 EASI-50 / EASI-90 / EASI-100 Response	FAS	CMH	NR	
Weeks 2, 4 and 8 EASI-75 Response	FAS	CMH	NR	
Weeks 2, 4, 8 and 12 CFBL in Total EASI Score	FAS	MMRM	OD	
Weeks 2, 4, 8 and 12 CFBL in NRS for severity	FAS	MMRM	OD	
Weeks 2, 4, 8 and 12 CFBL in %BSA	FAS	MMRM	OD	

CFBL=Change from baseline; CMH=Cochran-Mantel-Haenszel; ANCOVA=Analysis of Covariance; MMRM=Mixed-effect Model Repeated Measures; NR=Non-Responder; OD=Observed Data; MI=Multiple Imputation.

Appendix 2. Definition and Use of Visit Windows in Reporting

Visit windows will be used for efficacy variables, and for any safety data that display or summarize by study visit. For other endpoints (eg, ECG, vital signs), visit windows will be applied for summary statistics by study visits if required.

Visit Label	Target Day	Definition [Day window]
Screening		Days -28 to Day -1
Baseline	Day 1 (Day of first dose)	Last observation prior to and including day of first dose
Week 2	15	Days 2 to 22
Week 4	29	Days 23 to 36
Week 6	43	Days 37 to 50
Week 8	57	Days 51 to 71
Week 12	85	Days 72 to 99
Follow Up/End of Study		
Week 16	-	Days 100 to -

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 CaPS standards.

CCI

[redacted]

[redacted]

Appendix 4. Investigator's Global Assessment (IGA)

The Investigator's Global Assessment of atopic dermatitis is scored on a 5-point scale (0-4), reflecting a global consideration of the erythema, induration and scaling. The clinical evaluator of atopic dermatitis will perform an assessment of the overall severity of atopic dermatitis and assign an IGA score and category as described in the table below. The assessment will be a static evaluation without regard to the score at a previous visit.

IGA Score

Score	Category	Description
0	Clear	Atopic dermatitis is cleared, except for any residual discoloration (post-inflammatory hyperpigmentation and/or hypopigmentation).
1	Almost Clear	Overall, the atopic dermatitis is not entirely cleared and remaining lesions are light pink (not including post inflammatory hyperpigmentation) and/or; have barely palpable hard thickened skin and/or papules and/or; have barely perceptible lichenification; excoriation and oozing/crusting are absent.
2	Mild	Overall, the atopic dermatitis consists of lesions that are light red; with slight, but definite hard thickened skin and/or papules; with slight, but definite linear or picked scratch marks or penetrating surface injury; with slight, but definite thickened skin, fine skin markings, and lichenoid scale; oozing/crusting is absent.
3	Moderate	Overall, the atopic dermatitis consists of lesions that are red; with easily palpable moderate hard thickened skin and/or papules; with moderate linear or picked scratch marks or penetrating surface injury; with moderate thickened skin, coarse skin markings, and coarse lichenoid scale; with slight oozing/crusting.
4	Severe	Overall, the atopic dermatitis consists of lesions that are deep, dark red; with severe hard thickened skin and/or papules; with severe linear or picked scratch marks or penetrating surface injury; with severe thickened skin with very coarse skin markings and lichenoid scale; with moderate to severe oozing/crusting.

Appendix 5. 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 by the atopic dermatitis clinical evaluator 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

%BSA with Atopic Dermatitis: The number of handprints of skin afflicted with atopic dermatitis in a body region can be used to determine the extent (%) to which a body region is involved with atopic dermatitis (see table below). When measuring, the handprint unit refers to the size of each individual subject's hand with fingers in a closed position.

Handprint Determination of %BSA

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 and groin/genitals)	30	3.33%
Lower Limbs (including buttocks)	40	2.5%

Handprint refers to the hand size of each individual subject.

* The number of handprints will be for the entire body region.

EASI Area Score Criteria

Percent 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 (see table below).

EASI Body Region Weighting

Body Region	Body Region Weighting
Head and Neck	0.1
Upper Limbs	0.2
Trunk (including axillae and groin/genitals)	0.3
Lower Limbs (including buttocks)	0.4

* No adjustment for body regions excluded for assessment.

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 below.

$$EASI = 0.1Ah(Eh+Ih+Exh+Lh) + 0.2Au(Eu+Iu+ExU+Lu) + 0.3At(Et+It+Ext+Lt) + 0.4Al(El+Il+Exl+Ll)$$

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 0.1 and range from 0.0 to 72.0, with higher scores representing greater severity of atopic dermatitis.

CCI

[redacted]

[redacted]

Appendix 7. Schedule of Activities

The schedule of activities table provides an overview of the protocol visits and procedures. Refer to the STUDY PROCEDURES and ASSESSMENTS sections of the protocol for detailed information on each procedure and assessment required for compliance with the protocol.

The investigator may schedule visits (unplanned visits) in addition to those listed on the schedule of activities table, in order to conduct evaluations or assessments required to protect the well-being of the subject.

	Screening	**Intervention Period**									**Follow-up**	**Notes**
Visit Identifier	**Day -28**	**Day 1 Week 0 Base line**	**Day 2 Week 0**	**Day 4 Week 0**	**Day 8 Week 1 (by Phone)**	**Day 15 Week 2**	**Day 29 Week 4**	**Day 43 Week 6 (by Phone)**	**Day 57 Week 8**	**Day 85 Week 12 (EOT/ET)**	**Day 113 Week 16 (EOS)**	Visit 11 is 4 weeks after last dose in case of early study termination.
	Visit 1	**Visit 2**	**Visit 3**	**Visit 4**	**Visit 5**	**Visit 6**	**Visit 7**	**Visit 8**	**Visit 9**	**Visit 10**	**Visit 11**	
Visit Window (days)	**None**	**None**	**None**	**±1**	**±1**	**±1**	**±2**	**±3**	**±3**	**±3**	**±3**	
Enrollment Procedures												
Informed consent	X											
Register participant using IRT	X											See Section 6.3
Inclusion and exclusion criteria	X	X										See Section 5.1 and Section 5.2
Demographics, medical history, tobacco and alcohol history, AD disease history and prior AD treatments	X											See Section 8.2.1
Review prior/concomitant medications & treatments	X	X	X	X	X	X	X	X	X	X	X	
Provide patient emergency contact card	X											
Medical Procedures												
Complete physical examination	X									X		See Section 8.2.1
Targeted physical exam		X				X	X		X		X	
Vital signs (blood pressure, pulse rate, respiratory rate, temperature)	X	X				X	X		X	X	X	See Section 8.2.2

	Screening	Intervention Period									Follow-up	Notes
Visit Identifier	**Day -28**	**Day 1 Week 0 Base line**	**Day 2 Week 0**	**Day 4 Week 0**	**Day 8 Week 1 (by Phone)**	**Day 15 Week 2**	**Day 29 Week 4**	**Day 43 Week 6 (by Phone)**	**Day 57 Week 8**	**Day 85 Week 12 (EOT/ET)**	**Day 113 Week 16 (EOS)**	Visit 11 is 4 weeks after last dose in case of early study termination.
	Visit 1	**Visit 2**	**Visit 3**	**Visit 4**	**Visit 5**	**Visit 6**	**Visit 7**	**Visit 8**	**Visit 9**	**Visit 10**	**Visit 11**	
Weight	X									X		See Section 8.2.1
Height	X											See Section 8.2.1
Chest X-Ray	X											See Section 8.2.3
ECG (12-lead)	X											See Section 8.2.7
Laboratory Procedures												
Serum chemistry and hematology (including coagulation panel)	X	X				X	X		X	X	X	See Appendix 2. See Section 8.2.8 for guidance on abnormal lab results
Lipid panel		X								X		8-hour fast required
CCI [redacted]												[redacted]
CCI [redacted]												[redacted]
Fresh blood collection for hematopoiesis evaluation		X		X		X	X		X	X	X	See Section 8.18 Collected from participants at participating sites
CCI [redacted]												[redacted] See Section 8.5
Lymphocyte subsets		X		X		X			X	X	X	See Appendix 2
Urinalysis	X	X				X	X		X	X	X	See Appendix 2
Serum pregnancy test	X											Required for all WOCBP
Urine Pregnancy Test (conducted at study site)		X				X	X		X	X	X	

	Screening	Intervention Period									Follow-up	Notes
Visit Identifier	**Day -28**	**Day 1 Week 0 Base line**	**Day 2 Week 0**	**Day 4 Week 0**	**Day 8 Week 1 (by Phone)**	**Day 15 Week 2**	**Day 29 Week 4**	**Day 43 Week 6 (by Phone)**	**Day 57 Week 8**	**Day 85 Week 12 (EOT/ET)**	**Day 113 Week 16 (EOS)**	Visit 11 is 4 weeks after last dose in case of early study termination.
	Visit 1	**Visit 2**	**Visit 3**	**Visit 4**	**Visit 5**	**Visit 6**	**Visit 7**	**Visit 8**	**Visit 9**	**Visit 10**	**Visit 11**	
Banked Biospecimen Sample (Prep D1)		X										See Section 8.7.2 If not collected on the designated collection day, collect at the next available time point when biospecimens are being collected in conjunction with a participant visit
Banked Biospecimen Sample (Prep B1.5)		X								X		See Section 8.8.4
Banked Biospecimen Sample (Prep B2.5)		X								X		See Section 8.8.4
Blood samples collection for viral studies		X										Only analyzed on suspected viral infection/reactivation . See Section 8.2.8.2
HIV testing	X											See Appendix 2
Hepatitis B (HBsAg, HBsAb and HBcAb) and Hepatitis C (HCV Ab, HCV RNA)	X											See Section 8.2.8.1
HBV DNA (Only when HBsAg-, HBcAb+, HBsAb+ at screening)	X									X		See Section 8.2.8.1
Tuberculosis test	X											See Section 5.2 and Section 8.2.4
Study Treatment												
Randomization		X										
Drug dispensing		X					X		X			
Drug accountability						X	X		X	X		

	Screening	Intervention Period									Follow-up	Notes
Visit Identifier	**Day -28**	**Day 1 Week 0 Base line**	**Day 2 Week 0**	**Day 4 Week 0**	**Day 8 Week 1 (by Phone)**	**Day 15 Week 2**	**Day 29 Week 4**	**Day 43 Week 6 (by Phone)**	**Day 57 Week 8**	**Day 85 Week 12 (EOT/ET)**	**Day 113 Week 16 (EOS)**	Visit 11 is 4 weeks after last dose in case of early study termination.
	Visit 1	**Visit 2**	**Visit 3**	**Visit 4**	**Visit 5**	**Visit 6**	**Visit 7**	**Visit 8**	**Visit 9**	**Visit 10**	**Visit 11**	
Study intervention treatment		X----------								----------X		At Weeks 4 and 12, study intervention treatment should be taken at the study site to accommodate PK analysis. See Section 8.5
Assess eligibility for LTE Study B7451015										X		
Clinical Assessments												
Fitzpatrick Skin Type Assessment		X										See Section 8.2.6
Investigator's Global Assessment (IGA)	X	X				X	X		X	X	X	See Section 8.1.6
Eczema Area and Severity Index (EASI)	X	X				X	X		X	X	X	See Section 8.1.7
Skin Biopsies												
Target lesion(s) identification		X										See Section 8.1.2
Lesional skin punch biopsy sampling		X				X (Optional)	X			X		See Section 8.1.2
Non-lesional skin punch biopsy sampling		X								X (Optional)		See Section 8.1.2
Examination of Post-biopsy Site(s)		X	X	X		X	X		X	X	X	See Section 8.1.2
CCI [redacted]		[redacted]	[redacted]			[redacted]	[redacted]			[redacted]	[redacted]	[redacted]
CCI [redacted]												[redacted]
Photography		X					X			X		See Section 8.1.5

	Screening	Intervention Period									Follow-up	Notes
Visit Identifier	**Day -28**	**Day 1 Week 0 Base line**	**Day 2 Week 0**	**Day 4 Week 0**	**Day 8 Week 1 (by Phone)**	**Day 15 Week 2**	**Day 29 Week 4**	**Day 43 Week 6 (by Phone)**	**Day 57 Week 8**	**Day 85 Week 12 (EOT/ET)**	**Day 113 Week 16 (EOS)**	Visit 11 is 4 weeks after last dose in case of early study termination.
	Visit 1	**Visit 2**	**Visit 3**	**Visit 4**	**Visit 5**	**Visit 6**	**Visit 7**	**Visit 8**	**Visit 9**	**Visit 10**	**Visit 11**	
Patient-reported Outcomes (PRO)												
Peak Pruritus Numerical Rating Scale (NRS)	X------X	X----------X					X		X	X	X	See Section 8.1.9.1. During screening, assessed for the 7 days prior to Day 1 and also each day from Day 1 to Day 15
Night Time Itch Scale	X------X	X----------X					X		X	X	X	See Section 8.1.9.2. During screening, assessed for the 7 days prior to Day 1 and also each day from Day 1 to Day 15.
Safety												
C-SSRS	X	X							X	X	X	See Section 8.2.9.1
PHQ-8	X											See Section 8.2.9.2
Serious and non-serious adverse event monitoring	X	X	X	X	X	X	X	X	X	X	X	
Contraception check	X	X	X	X	X	X	X	X	X	X	X	Required for all females

Abbreviations: AD = atopic dermatitis; C-SSRS = Columbia Suicide Severity Rating Scale; EASI = Eczema Area and Severity Index, ECG = electrocardiogram; EOS = End of Study, EOT = End of Treatment; ET= early termination; HBsAg = hepatitis B surface antigen; HBsAb = hepatitis B surface antibody; HBcAb = hepatitis B core antibody; HBV DNA = hepatitis B viral deoxyribonucleic acid, HCV Ab = hepatitis C antibody; HCV RNA = hepatitis C viral ribonucleic acid; HIV = human immunodeficiency virus; hsCRP = high sensitivity C-reactive protein, IGA = Investigator's Global Assessment; IRT = Interactive Response System; LS = lesional; LTE = long-term extension; NL = non-lesional; NRS = Numerical Rating System; PHQ-8 = Patient Health Questionnaire - 8 Items; PK = Pharmacokinetic; WOCBP = women of childbearing potential.