Protocol C3941005

A PHASE 2B, RANDOMIZED, DOUBLE BLIND, VEHICLE CONTROLLED, PARALLEL GROUP STUDY TO ASSESS THE EFFICACY, SAFETY, TOLERABILITY AND PHARMACOKINETICS OF MULTIPLE DOSE LEVELS OF PF-07038124 OINTMENT FOR 12 WEEKS IN PARTICIPANTS 12 YEARS AND OLDER AND WITH MILD-TO-MODERATE ATOPIC DERMATITIS OR MILD-TO-SEVERE PLAQUE PSORIASIS

Statistical Analysis Plan
(SAP)

Version: **3.0**

Date: **18Aug2023**

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. VERSION HISTORY

This SAP for Study C3941005 is based on the protocol amendment dated 02Nov2022.

Table 1. Summary of Major Changes in SAP Amendments

Version/ Date	Associated Protocol Amendment	Change	Rationale
1/19May2022	Original 21Apr2022	Not Applicable	Not Applicable
2/7Jul2023	Protocol amendment 1 02Nov2022	Section 1 updated protocol version	Protocol amendment
		Section 4. Added handling rules for multiple enrollment	A participant with multiple enrollment was identified
		Section 6.1.2: Clarified intended weight used for stratified analysis is minimum risk weight instead of Mantel-Haenszel weight.	Minimum risk weights are designed to improve precision and reduce bias and can minimize the power loss that can occur when constant odds ratio is not met.
		Section 6.5: Deleted subset analysis for adolescent	There were only few adolescent enrolled for the study.

3	Protocol amendment 1 02Nov2022	Section 4: added statement about excluding evaluation at visits by non-qualified raters	A rater unqualified for PGA/PASI evaluated 10 PsO participants at several visits and such data need to be excluded.

2. INTRODUCTION

This statistical analysis plan (SAP) provides the detailed methodology for summary and statistical analyses of the data collected in Study C3941005.

2.1. Modifications to the Analysis Plan Described in the Protocol

NA

2.2. Study Objectives, Endpoints, and Estimands

	ATOPIC DERMATITIS		
Type	**Objectives**	**Endpoints**	**Estimands**
	Primary:	**Primary:**	**Primary:**
Efficacy	• To compare the efficacy of 2 dose levels of PF-07038124 ointment versus vehicle, using IGA success as the endpoint in participants with mild or moderate AD.	• Proportion of participants achieving IGA score of clear (0) or almost clear (1) (on a 5-point scale) and a reduction from baseline of ≥2 points at Week 12.	• Estimand E1: the difference in proportions of the binary endpoint between IP treated (0.01% or 0.03% PF-07038124 ointment) and vehicle in participants with mild or moderate AD without the benefit of additional prohibited medications and regardless of treatment compliance.

	Secondary:	**Secondary:**	**Secondary:**
Efficacy	• To compare the efficacy of 2 dose levels of PF-07038124 ointment versus vehicle using EASI-75 as an endpoint in participants with mild or moderate AD.	• Proportion of participants achieving EASI-75 (75% improvement from baseline) at all study visit time points specified in the SoA.	• Estimand E1 described above.
Efficacy	• To compare the efficacy of 2 dose levels of PF-07038124 ointment versus vehicle, using IGA success as the endpoint in participants with mild or moderate AD.	• Proportion of participants achieving IGA score of clear (0) or almost clear (1) (on a 5-point scale) and a reduction from baseline of ≥2 points at all study visit time points specified in the SoA, except at Week 12.	• Estimand E1 described above.
Efficacy	• To compare the efficacy of 2 dose levels of PF-07038124 ointment versus vehicle, using percent CFB in EASI as the endpoint in participants with mild or moderate AD.	• Percent CFB in EASI total score at all study visit time points specified in the SoA.	• Estimand E2: the difference in mean of the continuous endpoint between IP treated (0.01% or 0.03% PF-07038124 ointment) and vehicle in participants with mild or moderate AD without the benefit of additional prohibited medications and regardless of treatment compliance.
Efficacy	• To compare the efficacy of 2 dose levels of PF-07038124 ointment versus vehicle, using IGA clear or almost clear as the endpoint in participants with mild or moderate AD.	• Proportion of participants achieving IGA score of clear (0) or almost clear (1) at all study visit time points specified in the SoA.	• Estimand E1 described above.

Efficacy	• To compare the efficacy of 2 dose levels of PF-07038124 ointment versus vehicle, using measures of PP-NRS PRO, in participants with mild or moderate AD.	• Proportion of participants having ≥4 points of reduction from baseline in weekly averages of PP-NRS at study visit time points specified in the SoA.	• Estimand E1 described above.
Efficacy	• To compare the efficacy of 2 dose levels of PF-07038124 ointment versus vehicle, using BSA as the endpoint in participants with mild or moderate AD.	• Percent CFB in affected BSA at all study visit time points specified in the SoA.	• Estimand E2 described above.
Safety	• To characterize the safety and tolerability of 2 dose levels of PF-07038124 ointment versus vehicle in participants with mild or moderate AD.	• Incidence of treatment emergent AEs and SAEs, clinically significant changes in vital signs, ECG, and laboratory tests. • Incidence of increase in local tolerability severity grades at times indicated in the SoA.	• Not Applicable.
	Tertiary/Exploratory:	**Tertiary/Exploratory:**	**Tertiary/Exploratory:**
CCI [redacted]			

CCI

Biomarker	**Tape Stripping Sub-study** • To evaluate pharmacological effects of applying multiple doses of PF-07038124 ointment versus vehicle ointment using tape stripping method that is used to collect upper part of epidermis.	CCI
Efficacy	**Photography Sub-study** • To compare the efficacy of two dose levels of PF-07038124 ointment versus vehicle using CCI • To document effects of applying multiple doses of PF-07038124 ointment versus vehicle ointment using photographic images.	

	PLAQUE PSORIASIS		
Type	**Objectives**	**Endpoints**	**Estimands**
	Primary:	**Primary:**	**Primary:**
Efficacy	• To compare the efficacy of multiple dose levels of PF-07038124 ointment versus vehicle, using PGA success as the endpoint in participants with mild-to-severe PsO.	• Proportion of participants achieving PGA score of clear (0) or almost clear (1) (on a 5-point scale) and a reduction from baseline of ≥2 points at Week 12.	Estimand E1: the difference in proportions of the binary endpoint between IP treated (0.01%, 0.03%, or 0.06% PF-07038124 ointment) and vehicle in participants with mild-to-severe PsO without the benefit of additional prohibited medications and regardless of treatment compliance.
	Secondary:	**Secondary:**	**Secondary:**
Efficacy	• To compare the efficacy of multiple dose levels of PF-07038124 ointment versus vehicle using PASI-75 as an endpoint, in participants with mild-to-severe PsO.	• Proportion of participants achieving PASI-75 (75% improvement from baseline) at times indicated in the SoA.	• Estimand E1 described above.

Efficacy	• To compare the efficacy of multiple dose levels of PF-07038124 ointment versus vehicle, using PGA success as endpoint in participants with mild-to-severe PsO.	• Proportion of participants achieving PGA score of clear (0) or almost clear (1) (on a 5-point scale) and a reduction from baseline of ≥2 points at times indicated in the SoA, except at Week 12.	• Estimand E1 described above.
Efficacy	• To compare the efficacy of multiple dose levels of PF-07038124 ointment versus vehicle, using CFB PASI as the endpoint in participants with mild-to-severe PsO.	• Change from baseline in PASI at times indicated in the SoA.	Estimand E2: the difference in mean of the continuous endpoint between IP treated (0.01%, 0.03%, or 0.06% PF-07038124 ointment) and vehicle in participants with mild-to-severe PsO without the benefit of additional prohibited medications and regardless of treatment compliance.
Efficacy	• To compare the efficacy of multiple dose levels of PF-07038124 ointment versus vehicle, using PGA clear or almost clear as the endpoint in participants with mild-to-severe PsO.	• Proportion of participants achieving PGA score of clear (0) or almost clear (1) at all study visit time points specified in the SoA.	• Estimand E1 described above.
Efficacy	• To compare the efficacy of multiple dose levels of PF-07038124 ointment versus vehicle, using measures of PP-NRS PRO, in adult (18-75 years old) participants with mild-to-severe PsO.	• Proportion of adult (18-75 years old) participants having ≥4 points of reduction from baseline in weekly averages of PP-NRS at study visit time points specified in the SoA.	• Estimand E1 described above.

Efficacy	• To compare the efficacy of multiple dose levels of PF-07038124 ointment versus vehicle, using BSA as the endpoint in participants with mild-to-severe PsO.	• Percent CFB in affected BSA at all study visit time points specified in the SoA.	• Estimand E2 described above.
Safety	• To characterize the safety and tolerability of multiple dose levels of PF-07038124 ointment versus vehicle in participants with mild-to-severe PsO.	• Incidence of treatment emergent AEs and SAEs, clinically significant changes in vital signs, ECG, and laboratory tests. • Incidence of increase in local tolerability severity grades at times indicated in the SoA.	• Not Applicable.
	Tertiary/Exploratory:	**Tertiary/Exploratory:**	**Tertiary/Exploratory:**
CCI [redacted]			

CCI

Efficacy	**Tape Stripping Sub-study** • To evaluate pharmacological effects of applying multiple doses of PF-07038124 ointment versus vehicle ointment using tape stripping method that is used to collect upper part of epidermis.
Efficacy	**Photography Sub-study** • To document effects of applying multiple doses of PF-07038124 ointment versus vehicle ointment using photographic images.

CCI

2.2.1. Primary Estimand

The estimand (E1) is defined by the following attributes:

Atopic Dermatitis:

Treatment condition: received treatment: PF-07038124 ointment 0.01%, 0.03% or vehicle, without the benefit of additional prohibited medications and regardless of treatment compliance.

Population: Participants with mild or moderate AD as defined by the inclusion and exclusion criteria.

Variables: Proportion of participants achieving IGA score of clear (0) or almost clear (1) and a reduction from baseline of ≥2 points at Week 12 and other binary endpoints.

Psoriasis:

Treatment condition: received treatment: PF-07038124 ointment 0.01%, 0.03%, 0.06% or vehicle, without the benefit of additional prohibited medications and regardless of treatment compliance.

Population: Participants with mild, moderate, or severe PsO as defined by the inclusion and exclusion criteria.

Variables: Proportion of participants with PGA score clear (0) or almost clear (1) and a reduction from baseline of ≥2 points at Week 12 and other binary endpoints.

Inter-current Events: A) Prohibited medication – response will be considered as failure for time points after start date of the event for participants who receive prohibited medication post-randomization. B) Withdrawal and all other events, except COVID-19 pandemic, leading to missing data will be treated similarly assuming that participants no longer receive benefit from the study intervention and hence will be considered as failure. C) COVID-19 pandemic: if a participant misses a visit due to COVID-19, he/she will be excluded from the analysis for that visit; If a participant discontinues treatment or withdraws from the study due to COVID-19, he/she will be excluded from the analyses after the treatment discontinuation visit or study withdrawal visit, respectively. D) Inadequate compliance – participants data will be used as recorded.

Population level summary: The difference in success rates between each dose level and vehicle.

2.2.2. Secondary Estimand

All binary secondary endpoints estimands are similar to the primary endpoints.

The continuous endpoint estimand (E2) is defined by the following attributes:

Atopic Dermatitis:

Treatment condition: received treatment: PF-07038124 ointment 0.01%, 0.03% or vehicle, without the benefit of additional prohibited medications and regardless of treatment compliance.

Population: Participants with mild or moderate AD as defined by inclusion and exclusion criteria.

Variables: Percent change from baseline in EASI total score and percent CFB in affected BSA at all study visit time points specified in SoA.

Psoriasis:

Treatment condition: received treatment: PF-07038124 ointment 0.01%, 0.03%, 0.06% or vehicle, without the benefit of additional prohibited medications and regardless of treatment compliance.

Population: Participants with mild, moderate, or severe PsO as defined by the inclusion and exclusion criteria.

Variables: Change from baseline in PASI total score and Percent CFB in affected BSA at all study visit time points specified in the SoA.

Intercurrent events: A) Prohibited medication – all scores measured at time points after start date of the event for participants who receive prohibited medication post randomization will

be excluded from the analysis and treated as missing scores. Missing scores will be imputed based on the assumption that participants do not benefit from the IP treatment.
B) Withdrawal and all other events leading to missing data, except COVID-19 pandemic, will be treated similarly assuming participants have efficacy values similar to control participants. C) COVID-19 pandemic: if a participant misses a visit due to COVID-19, he/she will be excluded from the analysis for that visit; If a participant discontinues treatment or withdraws from the study due to COVID-19, he/she will be excluded from the analyses after the treatment discontinuation visit or study withdrawal visit, respectively. D) Inadequate compliance – participants data will be used as recorded.

Population-level summary: The mean difference between each dose level and vehicle.

2.2.3. Additional Estimand(s)

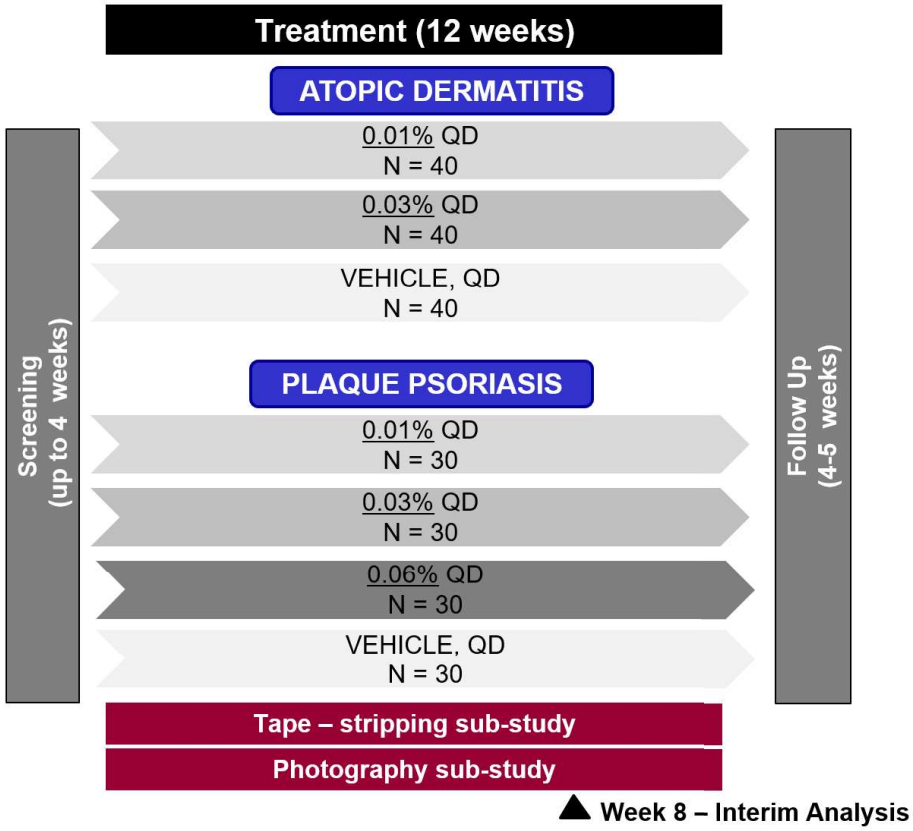
There is no defined estimand for other endpoints, and they will be analyzed using Pfizer data standards as applicable.

2.3. Study Design

This is a phase 2b, randomized, double blind, vehicle controlled, parallel group, multicenter study to assess the efficacy, safety and tolerability, and PK of multiple doses of PF-07038124 ointment versus vehicle in treatment of participants, 12 years and older, and with AD or PsO.

In this study, a total of approximately 240 participants with AD or PsO will be randomly assigned to study intervention (PF-07308124 or vehicle).

Figure 1 Study Design Schema



3. ENDPOINTS AND BASELINE VARIABLES: DEFINITIONS AND CONVENTIONS

3.1. Primary Endpoint(s)

- Proportion of participants achieving IGA score of clear (0) or almost clear (1) (on a 5-point scale) and a reduction from baseline of ≥2 points at Week 12.

- Proportion of participants achieving PGA score of clear (0) or almost clear (1) (on a 5-point scale) and a reduction from baseline of ≥2 points at Week 12.

3.2. Secondary Endpoint(s)

For AD:

- Proportion of participants achieving EASI-75 (75% improvement from baseline) at all study visit time points specified in the SoA.
- Proportion of participants achieving IGA score of clear (0) or almost clear (1) (on a 5-point scale) and a reduction from baseline of ≥2 points at all study visit time points specified in the SoA, except at Week 12.
- Percent CFB in EASI total score at all study visit time points specified in the SoA.
- Proportion of participants achieving IGA score of clear (0) or almost clear (1) at all study visit time points specified in the SoA.
- Proportion of participants having ≥4 points of reduction from baseline in weekly averages of PP-NRS at study visit time points specified in the SoA.
- Percent CFB in affected BSA at all study visit time points specified in the SoA.
- Incidence of treatment emergent AEs and SAEs, clinically significant changes in vital signs, ECG, and laboratory tests.
- Incidence of increase in local tolerability severity grades at times indicated in the SoA.

For Psoriasis:

- Proportion of participants achieving PASI-75 (75% improvement from baseline) at times indicated in the SoA.

- Proportion of participants achieving PGA score of clear (0) or almost clear (1) (on a 5-point scale) and a reduction from baseline of ≥2 points at times indicated in the SoA, except at Week 12.
- Change from baseline in PASI at times indicated in the SoA.
- Proportion of participants achieving PGA score of clear (0) or almost clear (1) at all study visit time points specified in the SoA.
- Proportion of **adult** participants (18-75 years old) having ≥4 points of reduction from baseline in weekly averages of PP-NRS at study visit time points specified in the SoA.
- Percent CFB in affected BSA at all study visit time points specified in the SoA.
- Incidence of treatment emergent AEs and SAEs, clinically significant changes in vital signs, ECG, and laboratory tests.
- Incidence of increase in local tolerability severity grades at times indicated in the SoA.

3.3. Tertiary/Exploratory Endpoint(s)

For AD:

CCI

[redacted]

CCI

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 and height will be based on measures collected at Visit 1/Screening visit. Study Day 1 is defined as the day the participant 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 the last measurement before first dose on Day 1. If a value is missing on Day 1, then the last available observation before Day 1 will be

used. For the PP-NRS score, baseline will be defined as the average of all values recorded between Day -7 and Day -1.

3.5. Safety Endpoints

Safety will be assessed by reporting of AEs, ECG, vital signs, local skin tolerability, and clinical laboratory results in all participants 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.
- Incidence of SAEs and AEs leading to discontinuation.
- Incidence of clinical abnormalities in clinical laboratory values, ECG measurements, and vital signs.
- Change from baseline in ECG measurements and vital signs.
- Incidence of skin tolerability

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

3.5.1. Adverse Events

All adverse events that start on or after the first dosing date and time, if collected, but before the last dose plus the lag time (35 days) will be flagged as TEAEs. The algorithm will not consider any events that started prior to the first dose date. Missing AE start date and time components will be imputed conservatively as the earliest date and time possible.

3.5.2. Laboratory Data

The laboratory tests will be performed at time points identified in the Schedule of Activities. Unscheduled clinical labs may be obtained at any time during the study to assess any perceived safety concerns at the investigator's discretion.

3.5.3. Vital Sign

Vital signs (blood pressure, pulse, and temperature) will be measured after a minimum of 5 minutes rest as indicated in the SoA.

3.5.4. Electrocardiograms

Standard 12-lead ECGs should be collected at times specified in the SoA.

3.5.5. Local Skin Tolerability

The investigator or designee will assess tolerability at the site of study intervention application, focusing on the non-lesion skin, immediately post application of the study intervention as indicated in the Schedule of Activities.

4. ANALYSIS SETS (POPULATIONS FOR ANALYSIS)

Data for all participants will be assessed to determine if participants meet the criteria for inclusion in each analysis population prior to releasing the database and classifications will be documented per standard operating procedures. For participants with multiple enrollment, only the participant identifier and data under initial enrollment will be included and summarized. Data collected for other enrollments will be included only in listings and narratives. For IGA, PGA, EASI and PASI related endpoints that require qualified raters, evaluations at visits by non-qualified raters will be excluded.

Participant Analysis Set	Description
ITT Analysis Set	All participants randomly assigned to study intervention. Participants will be analyzed according to the randomized intervention.
Safety Analysis Set	All participants randomly assigned to study intervention who took at least 1 dose of study intervention. A randomized but not treated participant will be excluded from the safety analyses. Participants will be analyzed according to the intervention received.

Participant Analysis Set	Description
CCI	[redacted]

5. GENERAL METHODOLOGY AND CONVENTIONS

The primary analysis will be performed when all randomized participants have either completed their 12-week study participation period or withdrawn early, or the study stop prematurely due to any reason, and the database is released.

5.1. Hypotheses and Decision Rules

Statistical inference will be made on the primary endpoints: IGA/PGA success rate at Week 12. The null hypothesis is that there is no difference between any PF-07038124 doses being tested and vehicle. The alternative hypothesis is that at least one PF-07038124 dose being tested is superior to vehicle with a higher IGA/PGA success rate. The study will be considered positive for each indication when the corresponding null hypothesis is rejected. For each indication, the overall family wise Type I error rate will be controlled at 1-sided 0.025 level using the Hochberg step up procedure.

A sufficient number of AD participants will be screened to achieve approximately 120 participants randomly assigned to study intervention. The sample size calculation is based on the primary efficacy estimand IGA success at Week 12. Assuming 15% IGA success rate for the vehicle arm, a sample size of 40 per arm will have 89% power to detect a treatment difference of 35% under a 1-sided type 1 error 0.0125 for each of the 2 dose levels comparison to vehicle by the unconditional exact test, with an overall Type 1 error 0.025 1-sided. In addition, this sample size will provide 96% power to detect a 42% difference in EASI-75 between an active arm (62% EASI-75) and vehicle (20% EASI-75).

A sufficient number of PsO participants will be screened to achieve approximately 120 participants randomly assigned to study intervention. The sample size calculation is based on the primary efficacy estimand PGA success at Week 12. Assuming 8% PGA success rate for the vehicle arm, a sample size of 30 per arm will have 92% power to detect a treatment difference of 42% under a 1-sided type 1 error 0.008 for each of the three dose levels comparison to vehicle by the unconditional exact test, with an overall Type 1 error 0.025 1-sided. In addition, this sample size will provide 96% power to detect a 44% difference in PASI-75 between an active arm (50% PASI-75) and vehicle (6% PASI-75).

5.2. General Methods

5.2.1. Analyses for Binary Endpoints

For landmark (cross-sectional) analyses of binary endpoints, Blyth-Still-Casella method will be used for estimates and confidence interval for each treatment arm, and Chan and Zhang's exact confidence interval (CI) method[1] will be used for differences between treatment arms.

For all binary endpoints, a summary of number and percentage of participants in each category by treatment arm and time will be produced and the response rate will also be plotted against time by treatment.

5.2.2. Analyses for Continuous Endpoints

Landmark (cross-sectional) analyses of key continuous endpoints will be performed by analysis of covariance (ANCOVA). The ANCOVA model will include treatment arm as fixed effect and baseline dependent variable as a covariate. Estimates of least square mean (LSM) values and the LSM differences between each PF-07038124 dose and vehicle will be derived. The corresponding p-values and 95% confidence intervals will also be presented.

Mixed model repeated measures (MMRM) models may be used when there are multiple time points. The fixed effects of treatment, visit, and treatment by visit interaction will be included. Baseline value will be included as a covariate. Unstructured covariance matrix will be assumed for the model errors when possible. Compound symmetry covariance matrix

will be used if the model with unstructured covariance doesn't converge. At each visit, estimates of least square mean (LSM) values and the LSM differences between each PF-07038124 dose and vehicle will be derived. The corresponding p-values and 95% confidence intervals will also be presented.

Unless stated otherwise, descriptive summary statistics of absolute values, change and percent from baseline for all continuous variables will be presented by treatment and time. These statistics will include n, mean, median, standard deviation, minimum and maximum.

5.2.3. Analyses for Categorical Endpoints

A summary of number and percentage of participants in each category will be presented by treatment and time point.

5.2.4. Analyses for Time-to-Event Endpoints

NA.

5.3. Methods to Manage Missing Data

5.3.1. Efficacy Data

Analysis of binary data including primary endpoints will use the primary estimand (E1). All missing values, except missing due to COVID-19 pandemic, will be set to a failure (Section 6.1.1).

Continuous secondary endpoint analysis will use the secondary estimand (E2). For each landmark analysis (e.g., cross sectional analysis by week) missing data, except missing due to COVID-19 pandemic, will be imputed using a control-based imputation method (Section 6.2.2).

Summaries of tertiary/exploratory continuous data will use the observed data only and no additional considerations are needed.

If a participant misses a visit due to COVID-19, he/she will be excluded from the analysis for that visit; If a participant discontinues treatment or withdraws from the study due to

COVID-19, he/she will be excluded from the analyses after the treatment discontinuation visit or study withdrawal visit, respectively.

5.3.2. Pharmacokinetic Concentrations and Biomarker Data

- **Concentrations outside the limit of quantification.**

  In summary statistics for pharmacokinetic and biomarker data, assayed values below the lower limit of quantification (LLOQ) will be set to zero. Other imputations (e.g., ½ LLOQ) may also be considered in other analyses (e.g., Pop-PK and PK/PD analyses), if deemed appropriate. In listings, values below LLOQ will be reported as "<LLOQ" where LLOQ will be replaced with the numerical value for the lower limit of quantification. The LLOQ for various PK and biomarker concentrations will be noted in all tables and listings.

- **Missing concentrations.**

  If a concentration value is not collected or cannot be analyzed due to sample quality issues, it will be considered as missing data and will not be imputed.

- **Missing actual sampling time.**

  If actual sampling time (date or hour) value is missing, the protocol-stated nominal time will be used.

6. ANALYSES AND SUMMARIES

6.1. Primary Endpoint(s)

6.1.1. Main Analysis

- Estimand strategy: Primary Estimand (Section 2.2.1). This estimand is intended to provide an estimate of the treatment effect in the absence of prohibited medication use in a population that may or may not be fully compliant with the treatment.

- Analysis set: ITT Analysis Set (Section 4).

- Analysis methodology: A landmark analysis of the IGA/PGA success endpoints at Week 12 will be performed using the ITT analysis set. The proportions responding and the corresponding risk difference comparing active treatment arm to vehicle group will be analyzed using the unconditional exact method; the risk differences and the corresponding 2 sided unconditional exact 95% confidence intervals will be computed using Chan and Zhang (1999)[1] method. The overall family wise Type I error rate will be controlled at the 1 sided 0.025 level using the Hochberg step up procedure.

- Intercurrent events and missing data: A) Prohibited medication – response will be considered as failure for time points after start date of the event for participants who receive prohibited medication post randomization. B) Withdrawal and all other events, except COVID-19 pandemic, leading to missing data will be treated similarly assuming that participants no longer receive benefit from the study intervention and hence will be considered as failure. C) COVID-19 pandemic: if a participant misses a visit due to COVID-19, he/she will be excluded from the analysis for that visit; If a participant discontinues treatment or withdraws from the study due to COVID-19, he/she will be excluded from the analyses after the treatment discontinuation visit or study withdrawal visit, respectively. D) Inadequate compliance – participants data will be used as recorded.

- Proportions, risk differences and 95% confidence intervals will be presented.

6.1.2. Sensitivity/Supplementary Analyses

Sensitivity Analyses:

If there are enough data in each of the disease severity groups, Cochran test stratified by baseline disease severity with the minimum risk weights proposed by Mehrotra and Railkar (2000) may be considered for IGA/PGA success at week 12.

In addition, a generalized linear mix model may be used for the repeated measures of IGA/PGA success at all time points. The fixed effects of treatment, visit, treatment by visit interaction be included in the model. If there are enough data in each of the disease severity

group, baseline disease severity may be included as additional fixed effect. Unstructured covariance will be used when possible. At each visit, estimates of mean odds ratio between the active treated group and the placebo will be presented. The corresponding p-values and 95% confidence intervals will also be derived from the model.

Supplementary Analyses:

For PGA success data with three active dose levels, a Bayesian Emax model may be used to characterize the dose response relationship (Appendix 2).

6.2. Secondary Endpoint(s)

6.2.1. Binary Secondary Endpoints

IGA/PGA success at other timepoints, EASI/ PASI-75, PP-NRS success (≥4 points of reduction from baseline), IGA/PGA of clear or almost clear will be analyzed the same way as described in Section 6.1.1. PP-NRS success rate will be summarized for all participants in ITT Analysis Set in AD, and adult participants (18–75 years old) only in ITT Analysis Set for PsO.

6.2.2. Continuous Secondary Endpoints

Continuous secondary endpoints include percent CFB in EASI total score, CFB in PASI total score, and percent CFB in affected BSA.

- Estimand strategy: Secondary estimand (Section 2.2.2). This estimand is intended to provide an estimate of the treatment effect in the absence of prohibited medication use in a population that may or may not be fully compliant with the treatment.
- Analysis set: ITT Analysis Set (Section 4).
- Analysis methodology: A landmark analysis of these continuous endpoints at Week 12 will be performed by ANCOVA with baseline as a covariate. Missing data, except missing due to COVID-19 pandemic, will be imputed using a control-based imputation method. PROC MI will first be called to impute missing vehicle observations under the assumption data are missing at random (MAR) and impute

missing active treatment observations assuming they are similar to corresponding vehicle arm. The imputation will utilize ANCOVA model including baseline as a covariate fitted using only vehicle arm data. The analysis will combine the results from the multiple imputations using Rubin's rule's as implemented in SAS PROC MIANALYZE. MMRM models may also be used on observed data at all time points as sensitivity analyses. The fixed effects of treatment, visit, and treatment by visit interaction will be included. Baseline value will be included as a covariate. No adjustments for multiplicity will be made.

- Intercurrent events and missing data: A) Prohibited medication: all scores measured at time points after start date of the event for participants who receive prohibited medication post randomization will be excluded from the analysis and treated as missing scores. Missing scores will be imputed. B) Withdrawal and all other events leading to missing data, except COVID-19 pandemic: all missing scores will be imputed. C) COVID-19 pandemic: if a participant misses a visit due to COVID-19, he/she will be excluded from the analysis for that visit; If a participant discontinues treatment or withdraws from the study due to COVID-19, he/she will be excluded from the analyses after the treatment discontinuation visit or study withdrawal visit, respectively. D) Inadequate compliance – participants data will be used as recorded.

- The least-squares (LS) means, the 95% confidence interval for the LS means, the difference between the LS means for each pair of treatment groups, and the corresponding 95% confidence interval will be presented.

6.3. Tertiary/Exploratory Endpoints

6.3.1. Binary Tertiary/Exploratory Endpoints

CCI [redacted]

6.3.2. Continuous Tertiary/Exploratory Endpoints

CCI [redacted]

CCI [redacted]

[redacted]

6.3.3. PRO Endpoints

CCI [redacted] PP-NRS (for AD and **adult** PsO aged 18-75), CCI [redacted]

6.4. Other Endpoint(s)

6.4.1. PK Endpoints

PF-07038124 concentration data will be summarized through appropriate data tabulations, descriptive statistics, and graphical presentation. A population PK model may be developed for the purpose of estimating PK parameters. In addition, a relationship between doses/exposures and efficacy/safety endpoints may be evaluated using population PK/pharmacodynamic (PD) approach. Details of the methodology will be captured in a separate modeling plan and the results will also be reported separately.

6.5. Subset Analyses

Proportion of IGA/PGA success at Week 12 may be summarized for the subsets below if we have sufficient number of participants in each stratum:

- Baseline Disease Severity: mild/moderate for AD and mild/moderate/severe for PsO.

6.6. Baseline and Other Summaries and Analyses

6.6.1. Baseline Summaries

Demographic and baseline characteristics will be summarized by treatment for the safety analysis set. Key demographic and baseline variables include age, gender, race, ethnicity,

height, weight, body mass index, disease duration, baseline affected BSA (%), EASI and IGA score for AD participants, PASI and PGA for PsO participants.

6.6.2. Study Conduct and Participant Disposition

Participants' evaluation for analysis populations, disposition and discontinuation will be summarized according to CaPS.

6.6.3. Study Treatment Exposure

The exposure to study drug measured by number of days with dosing will be summarized. The number and percent of participants in exposure duration categories (<1 week to >12 weeks with 1 week increments) will be presented. A summary of dosing compliance, defined as % days with dosing out of the duration between first and last dosing, will be provided by treatment. Number of percent of participants with treatment non-compliance (compliance < 80% or more than 120% of study intervention application) will also be summarized. Ointment application rate, defined as amount of ointment applied per dose(mg) divided by treated BSA (cm**), will also be summarized by treatment and time. The treated BSA at subsequent visits should be equal to or greater than the value at Day 1 (Protocol section 8.9), and different from affected BSA as efficacy endpoint.

6.6.4. Concomitant Medications and Nondrug Treatments

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

6.7. Safety Summaries and Analyses

The analysis population for safety is described in Section 4. Safety and tolerability will be assessed by clinical review of all relevant parameters including adverse experiences (AEs), ECG, vital signs, local skin tolerability and laboratory tests. A complete list of laboratory parameters can be obtained in Appendix 2 of the protocol.

All the tables, listings and graphs for adverse events, lab parameters and vital sign and ECG will follow CaPS.

6.7.1. Adverse Events

AE endpoints including below will be listed and summarized in accordance with the CaPS:

- Treatment-emergent AEs , SAEs, treatment area AEs and treatment related AEs.
- Discontinuation of treatment due to AEs and treatment area AEs.

Table display will be by descending order of frequency in all treatment arms.

6.7.2. Laboratory Data

Laboratory data will be listed and summarized in accordance with the CaPS. . Categorical summaries of participants meeting pre-specified abnormality criteria will be created.

6.7.3. Vital Signs

The absolute value, and change from baseline will be summarized by treatment and time. Categorical summary of participants meeting pre-specified clinical criteria will be provided by treatment and time.

6.7.4. Electrocardiograms

The absolute value and changes from baseline for the ECG parameters QT interval, heart rate, QTc interval, PR interval, and QRS complex will be summarized by treatment and time.

The number (%) of participants with maximum post-dose QTc values and maximum increases from baseline in the following categories will be tabulated by treatment:

Safety QTc Assessment

Degree of Prolongation	Mild (msec)	Moderate (msec)	Severe (msec)
Absolute value	>450-480	>480-500	>500
Increase from baseline		30-60	>60

6.7.5. Local Skin Tolerability

Categorical summary tables will be provided by treatment and time. In addition, a summary of number and percent of participants by worst local skin tolerability grade during the whole study period will be created.

7. INTERIM ANALYSES

7.1. Introduction

An interim analysis may be performed to assess efficacy and safety when approximately 60% of the planned participants, i.e., approximately 25 participants per arm in AD and 18 participants per arm in PsO have completed or had the chance to complete Week 8 visit. Interim analysis results may be used for decisions regarding stopping for futility and internal program development. The detailed decision boundary will be detailed in the IRC charter. Additional interim analysis may be performed and will be detailed in the IRC charter.

7.2. Interim Analyses and Summaries

At interim, analyses specified in Section 6.1.1 will be performed for IGA/PGA success and EASI/PASI-75 at week 8. In addition, safety data including primarily adverse event, tolerability and PK data will be analyzed.

Interim predictive probability of success (PoS) in primary endpoints based on the interim data will be calculated assuming a non-informative uniform prior for IGA/PGA success rate. The trial may stop for futility (non-binding) if the interim PoS <20%.

Interim Predictive Probability of Success

Let p_1 and p_0 be the probability of success for active arm and vehicle arm, both following a Beta(1,1) prior. Let s_{1i} and s_{0i} be the number of successes observed out of n_{1i} and n_{0i} participants at stage i for the active and vehicle arm.

For a given criteria at the end of stage 2: Posterior Prob($p_1 - p_0 >$d| $s_{12}, s_{02}\, s_{11}, s_{01}$) >c, where d is a target value and c is a probability cut-off to be detailed in IRC charter, since

$p_1|s_{11}, s_{12} \sim Beta(s_{11} + s_{12} + 1, n_{11} - s_{11} + n_{12} - s_{12} + 1)$

$p_0|\ s_{01,} s_{02} \sim Beta(s_{01} + s_{02} + 1, n_{01} - s_{01} + n_{02} - s_{02} + 1).$

we can calculate interim (stage 1) predictive probability of success of such criteria as

Predictive Prob {Posterior Prob ($p_1 - p_0 >$d$|\ s_{12}, s_{02}\ s_{11}, s_{01}) > c|\ s_{11}, s_{01}\}$

using below predictive distributions property:

$s_{12}|s_{11} \sim$ Binomial$(n_{12}, p_1|s_{11})$, $and\ s_{02}|s_{01} \sim$ Binomial$(n_{02}, p_0|s_{01})$, where

$p_1|s_{11} \sim Beta(s_{11} + 1, n_{11} - s_{11} + 1)$ and $p_0|\ s_{01} \sim Beta(s_{01} + 1, n_{01} - s_{01} + 1).$

Decision will primarily be based on interim data at week 8 and corresponding success criteria, as it is expected to have limited week 12 data at interim.

8. REFERENCES

1. Chan ISF and Zhang Z. Test based exact confidence intervals for the difference of two binomial proportions. *Biometrics*, 1999, **55**:1201–1209.

2. Mehrotra, D. V., & Railkar, R. (2000). Minimum risk weights for comparing treatments in stratified binomial trials. Statistics in medicine, 19(6), 811-825

9. APPENDICES

Appendix 1. Summary of Efficacy Analyses

Efficacy Endpoint	Analysis Type	Population	Data Inclusion and Rules for Handling Intercurrent Events and Missing Data	Analysis Model
Week 12 IGA / PGA Response	Primary	ITT	NRI	Chan & Zhang
	Sensitivity	ITT	NRI	Minimum Risk
	Supportive	ITT	NRI	Generalized Linear Mixed Model
	Supportive	ITT	NRI	Bayesian Emax
Week 12 EASI/PASI-75 Response	Secondary	ITT	NRI	Chan & Zhang
Week 12 PCFB in EASI / CFB in PASI	Secondary	ITT	MI	ANCOVA
PCFB in EASI / CFB in PASI	Sensitivity	ITT	OC	MMRM
PP-NRS Rate / IGA of clear or almost clear / / PGA of clear or almost clear	Secondary	ITT	NRI	Chan & Zhang
CCI [redacted]	Exploratory	ITT	NRI	Chan & Zhang

PCFB= Percent change from baseline; CFB= Change from baseline; ANCOVA = Analysis of Covariance; NRI = Non-Responder Imputation; OC = Observed Cases.

Appendix 2. Dose Response Analysis for PsO

The 4-parameter Emax model will be fit using the latest version (currently v2.3.5) of the clinDR package. The below model will be applied the PGA success rate p at week **12**.

$$\log\frac{p}{1-p} = E_0 + E_{max}\frac{d^{\lambda}}{d^{\lambda} + {ED_{50}}^{\lambda}}$$

Weakly informative Priors Setup:

A weakly-informative prior distribution $\boldsymbol{E_0}$**~t(mean= -2.44, scale = 4, df=5)** gives a 31% probability $p_0 < 1\%$ and 7% probability $p_0 > 99\%$, and is robust enough to cover the historical vehicle data range (4% ~15%) for PGA success rate at week 6~12. The mean is based on meta-analysis results PGA success rate with a mean of 8% and 95% CI (4% ,15%) at week 6~12, including the C3941002 study result with observed success rate of 1/17 (5.9%) and 95% CI (0.3%, 27.8%) at week 6.

We will specify the prior for the treatment effect E_T at a target dose (0.06% the maximum dose), $E_T = E_{max}\frac{{d_T}^{\lambda}}{{d_T}^{\lambda} + {ED_{50}}^{\lambda}}$. A prior distribution of $\boldsymbol{E_T}$**~t (mean= 1.23, scale =2, df=5)** will be used. C3941002 study at week **6** with observed PGA success rate of 3/17 (17.6%) for the low dose 0.01% and 1/17 (5.9%) for vehicle, results in an odds ratio of 3.43 with 95% CI (0.23, 191), or **log** odds ratio of 1.23 with 95% CI (-1.5,5.25). Such prior gives a 5% probability of log odds ratio>5.25 (equivalent of a change in success rate from 8% to 94%), and a 39% probability of log odds ratio>1.82 (equivalent of a change in success rate from 8% to 35%, the observed success rate at week 12 assuming observed treatment effect is linear in time), and 12% probability of log odds ratio <-1.5 (equivalent of a change in success rate from 8% to 2%).

ED50: log(ED50/P50)~t(mean=0, scale= 1.73, df=5), Log(lambda)~t(mean=0, scale=0.425, df=5) with a correlation of -0.45 between log(lambda) and log(ED50) are the default prior in clinDR package, under which the probability of ED50/P50>32.6 is 5%, and the probability of lambda>2.35 is 5%. P50 =0.02% is empirically estimated as mean of the first two lowest dose.

The default MCMC burn-in and number of samples will be utilized, which will include 3 chains to assess convergence. Model diagnostics will be examined including trace and auto-correlations plots. If these raise concerns over model convergence, additional burn-ins, samples and thinning will be attempted to improve convergence. Non-monotone dose response relationship will be checked before model fitting.

An independent statistician will conduct QC of the analysis. The outputs of the analysis will be provided as .csv files to the programming team for inclusion with the tables and figures in the reporting.

Appendix 3. 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.

Visit No.	Visit Label	Target Day	Visit Window
	Screening	N/A	-28 ≤ day ≤-1
1	Baseline*	1	day = 1
2	Week 1	8	2≤day≤11
3	Week 2	15	12≤day≤22
5	Week 4	29	23≤day≤36
6	Week 6	43	37≤day≤50
7	Week 8	57	51≤day≤64
8	Week 10	71	65≤day≤78
9	Week 12	85	79≤day≤99
10	FUP/EOS	N/A	day >99
* Baseline analysis visit window may be considered as day≤1 in some analyses (e.g., those involving change from baseline). That is, in case that Day 1 observation is missing, the last observation before the first dosing date may be considered as the baseline. The baseline measurements for demography, height, pre-study medical history and medications will be collected at the "Screening" visit.			

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

Appendix 4. Sample SAS Code

Sample SAS Code for Estimand 1 – Risk Difference using Chan and Zhang (1999)

```
PROC BINOMIAL DATA=<DATASET> GAMMA=0 ALPHA=<Value> OUT=<>;
     PD/EX ONE STD;
     PO <POPULATION VARIABLE>;
     OU <OUTCOME VARIABLE>;
RUN;
```

Sample SAS Code for the Confidence Interval of a Binomial Proportion (Blyth-Still-Casella)

```
PROC BINOMIAL DATA=<DATASET> ALPHA=<value> OUT=<>;
     BI/BS;
     OU <RESPONSE VARIABLE>;
RUN;
```

Sample SAS Code for the Confidence Interval of two Binomial Proportion (Minimum Risk)

```
PROC FREQ DATA= <DATASET>;
     TABLES <STRATA>*<TRT>*EVENT / COMMONRISKDIFF(CL=MR
     TEST=MR COLUMN=2) ) RELRISK ALPHA=<VALUE>;
     ods output CommonPdiff=<> CommonPdiffTests=<>
RUN;
```

Sample SAS Code for Generalized Linear Mixed Model

```
PROC GLIMMIX DATA=<DATASET>;
     CLASS AVISIT <TRT> <STRATA> USUBJID;
     MODEL CRIT1FL(EVENT="Y")=<STRATA> <TRT>|AVISIT
     /ALPHA=0.05 DIST=BINARY DDFM = KR;
     RANDOM  _RESIDUAL_/SUBJECT = USUBJID TYPE=UN;
     LSMEANS <TRT>*AVISIT/ILINK COV DIFF CL ALPHA=<VALUE>;
RUN;
```

Sample SAS Code for Estimand 2 ANCOVA

```
libname c3941005  "/Volumes/app/…… /data_vai" ;
data ps;
        set c3941005.adps;
run;
data look;
        set ps;
        where param = "EASI02-Total Sum"  and viswin ne .;
        keep subjid param paramn avisit age sex trta viswin visit xviswin aval base chg pchg
trtan;
run;
data ex1;
        set look;
        where upcase(avisit)='WEEK 12';
       regimen = "QD";
       if trtan = 1 then dose=0;
       if trtan = 2 then dose=0.01;
       if trtan = 3 then dose=0.03;
       if trtan = 4 then dose=0.06;
        keep subjid dose regimen chg aval base;
run;
proc means data=ex1;
        class regimen dose;
        var chg aval;
run;
proc sort data=ex1 out=ex2;
   *data must be sorted by regimen before use in proc mi;
        by regimen dose;
run;
*imputing aval=observed so range of endpoint 0-72 can be included in;
*mi procedure;
proc mi data=ex2 seed=1022 nimpute=1000 out=outimp  max=. 72 72 min=. 0 0;
        by regimen;
```

```
            class dose;
            monotone reg(aval= base/details);
            mnar model( aval/modelobs = (dose="0"));
            var base  aval;
run;
proc univariate data=outimp;
var aval;
histogram ;
run;
data outimp1;
            set outimp;
            chg = aval - base;**calculate chg from baseline
run;
proc sort data = outimp1 out = outimp2;
   by _imputation_ regimen dose subjid;
run;

proc mixed data=outimp2;
            by _imputation_;
            class dose regimen;
            model chg = base dose*regimen;
            lsmeans dose*regimen/ diff alpha=<Value>;
            ods output diffs=diffs lsmeans=lsmeans;
run;

data diffsout;
            set diffs;
            **only keep within regimen contrasts vs placebo;
            where regimen = _regimen and dose = 0;
run;
**now use mianalyze on lsmean differences;
**First sort by group, regimen and _dose (dose = 0 for all groups);
proc sort data=diffsout out=diffsout1;
```

```
        by regimen _dose _imputation_;
run;
**now mianalyze by regimen and _dose;
**NB mianalyze only uses estimates and standard errors not CI limits etc.;
proc mianalyze data=diffsout1  alpha=<Value>;
    by regimen _dose;
    modeleffects estimate;
    stderr stderr;
    ods output parameterestimates=parameterestimates;
run;
```

Sample SAS Code for Estimand 2 MMRM

```
proc mixed data=_dsnin order=internal alpha=<Value>;
   by PARAMN;
   class SUBJID AVISITN TRTPN;
   model PCHG = BASE TRTPN AVISITN  TRTPN*AVISITN /solution ddfm=kr;
repeated AVISITN/ participant =SUBJID type=un r rcorr;
/**try UN first, if not converged, will use CS***/
   lsmeans TRTPN*AVISITN / diff cl alpha=<Value>;
   ods output lsmeans=lsmeans diffs=diffs;
run;
/*For one-sided P-value calculation, the LOWER tail side should be used. */;
```