Protocol C2501007

A PHASE 2A, MULTICENTER, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, 16-WEEK STUDY EVALUATING THE SAFETY AND EFFICACY OF PF-06650833, PF-06700841, AND PF-06826647 IN ADULTS WITH MODERATE TO SEVERE HIDRADENITIS SUPPURATIVA

Statistical Analysis Plan
(SAP)

Version: 3

Date: 03 Nov 2021

[redacted]

[redacted]

TABLE OF CONTENTS

CCI [redacted]

LIST OF TABLES

LIST OF FIGURES

APPENDICES

1. VERSION HISTORY

This Statistical Analysis Plan (SAP) for study C2501007 is based on the protocol (Amendment 1) dated 28 January 2020.

Table 1. Summary of Changes

Version/ Date	Associated Protocol Amendment	Rationale	Specific Changes
1 07Feb2020	Not Applicable	Not Applicable	Not Applicable
2 24May2020	C2501007 Protocol Amendment 1	Adjust or supplement the planned statistical analysis due to COVID-19	• Section 2.1, section 3.8, section 5.3 and section 6, details were added regarding missing data due to COVID-19. • Section 3.7 and 6.7, additional listings and safety anslysis were added for COVID-19. • Section 6.9.1, analysis for comparison between before and after TYK2 coming into study were removed.
3 03Nov2021	C2501007 Protocol Amendment 1	To be consistent with protocol	• Section 2.1, estimands were changed to E1, E2 and E3. • Sections 3.2, 3.3 and 3.5, detailed definition of endpoints were added according to protocol. • Sections 5.2.1, 6.2, 6.3 and 6.5, CIs for non-primary endpoints were updated to 90% CI, adjusted p-values were used only for primary endpoint, and was changed from Bonferroni to Hochberg correction. • Section 5.3.2, baseline value was added into the MMRM model as an independent variable. • Section 5.3.2, PROC MI was added for continuous endpoint. • Sections 5.3.5, 6.2.3, 6.2.4 and 6.2.5 multiple imputation with Bayesian model were added for NRS missing data. • Sections 6.2, 6.3 and 6.5, same types of analysis were merged. • Section 6.8, number of draing/non-draing fistula at baseline was added. • Appendix, example SAS codes were added.

2. INTRODUCTION

This statistical analysis plan (SAP) provides the detailed methodology for summary and statistical analyses of the data collected in Study C2501007. This document may modify the plans outlined in the protocol; however, any major modifications of the primary endpoint definition or its analysis will also be reflected in a protocol amendment.

2.1. Study Objectives, Endpoints, and Estimands

The objectives of the current study are to evaluate the efficacy, safety and tolerability of PF-06650833, an IL-1 receptor associated kinase 4 (IRAK4) inhibitor, PF-06700841, a dual inhibitor of human tyrosine kinase 2 (TYK2) and Janus kinase 1 (JAK1), and PF-06826647, a potent TYK2 inhibitor, in patients with moderate to severe hidradenitis suppurativa (HS). Since the pathophysiology of HS is not defined completely, it is uncertain which of these three disease targets and pathways would be of more relevance for the treatment of patients with HS. Therefore, the additional objective of this study is to compare the efficacy of PF-06650833, PF-06700841, and PF-06826647 with the goal to select one of these inhibitors for further clinical development. The detailed study objectives and corresponding endpoints are provided in the Table 2.

Table 2. Study Objectives and Endpoints

Objectives	***Endpoints***	***Estimands***
Primary:	***Primary:***	***Primary:***
To evaluate the efficacy of PF-06650833, PF-06700841, and PF-06826647 vs placebo in participants with HS as assessed by HiSCR.	*Percentage of participants with HiSCR response* at Week 16.*	*This estimand uses a composite estimand theory (ICH E9 addendum) and is intended to provide a population level estimate of the treatment effect of the IP on a binary endpoint regardless of participant compliance with the IP dosing. Population-level summary: difference between treated and placebo in proportion of participants with HiSCR response at Week 16.*
Secondary:	***Secondary:***	***Secondary:***
To evaluate the efficacy of PF-06650833, PF-06700841, and PF-06826647 vs placebo in participants with HS at Week 16 and over time.	*Percentage of participants with HiSCR response* at Weeks 1, 2, 4, 6, 8, and 12.* *Percentage of participants with a total abscess and inflammatory nodule (AN) count of 0 or 1; 0, 1, or 2 at Week 16.*	*This estimand is intended to provide a population level estimate of the treatment effect of the IP on a binary endpoint regardless of participant compliance with the IP dosing. Population-level summary: difference between treated and placebo in proportion of participants with HiSCR response at Weeks 1, 2, 4, 6, 8 and 12, respectively; difference between treated and placebo in proportion of participants with a total AN count of 0 or 1, or 0, 1 or 2, respectively at Week 16.*
	Percent change from baseline (CFB) in AN count at Weeks 1, 2, 4, 6, 8, 12 and 16.	*This estimand is intended to provide a population level estimate of the treatment effect of the IP on a continuous endpoint regardless of*

Table 2. Study Objectives and Endpoints

Objectives	***Endpoints***	***Estimands***
		participant compliance with the IP dosing. *Population-level summary: difference between treated and placebo change from baseline in AN count at Weeks 1, 2, 4, 6, 8, 12 and 16.*
To evaluate the efficacy of PF-06650833, PF-06700841, and PF-06826647 vs placebo on International Hidradenitis Suppurativa Severity Score System (IHS4).	*Absolute score and percent CFB in IHS4 at Weeks 1, 2, 4, 6, 8, 12 and 16.*	*This estimand is intended to provide a population level estimate of the treatment effect of the IP on a continuous endpoint regardless of participant compliance with the IP dosing.* *Population-level summary: difference between treated and placebo in absolute score and change from baseline on IHS4 at Weeks 1, 2, 4, 6, 8, 12 and 16.*
To evaluate the effects of PF-06650833, PF-06700841, and PF-06826647 vs placebo on HS flare.	*Proportion of participants who experience an HS flare, defined as at least a 25% increase in AN count with a minimum increase of 2 relative to Baseline, at Weeks 4, 8, 12 and 16.*	*The estimand is intended to provide a population level estimate of the treatment effect of the IP on a binary endpoint regardless of participant compliance with IP dosing.* *Population-level summary: difference between treated and placebo in proportion of participants experience flare at Weeks 4, 8, 12 and 16.*
To evaluate the efficacy of PF-06650833, PF-06700841, and PF-06826647 vs placebo on pain and pain reduction over time in participants with HS.	*Percentage of participants with ≥30% reduction and ≥1-unit reduction from baseline in PGA-Skin Pain numeric rating scale (NRS30) – at worst and on average, respectively, amongst participants with baseline NRS ≥3, at Weeks 1, 2, 4, 6, 8, 12 and 16.* *Percent CFB in NRS, at worst and on average respectively, in participants who have baseline NRS ≥3, at Weeks 1, 2, 4, 6, 8, 12 and 16.*	*This estimand is intended to provide a population level estimate of the treatment effect of the IP on a binary endpoint regardless of participant compliance with the IP dosing.* *Population-level summary: difference between treated and placebo in proportion of participants with ≥30% reduction and ≥1-unit reduction from baseline in PGA-Skin Pain NRS (NRS30), at worst and on average, respectively, amongst participants with baseline NRS ≥3, at Weeks 1, 2, 4, 6, 8, 12 and 16.*
	CFB in NRS, at worst and on average respectively, at Weeks 1, 2, 4, 6, 8, 12 and 16.	*This estimand is intended to provide a population level estimate of the treatment effect of the IP on a continuous endpoint regardless of participant compliance with the IP dosing.* *Population-level summary: difference between treated and placebo percent*

Table 2. Study Objectives and Endpoints

Objectives	*Endpoints*	*Estimands*
		change from baseline in NRS, at worst and on average respectively, in participants who have baseline NRS ≥3, at Weeks 1, 2, 4, 6, 8, 12 and 16. *Difference between treated and placebo change from baseline in NRS, at worst and on average respectively, at Weeks 1, 2, 4, 6, 8, 12 and 16.*
To evaluate the efficacy of PF-06650833, PF-06700841, and PF-06826647 vs placebo on erythema.	*Proportion of participants achieving erythema score of 1 or 0 in all affected anatomic regions among participants who have an erythema score of 2 or more in at least 1 anatomic region at baseline.*	*The estimand is intended to provide a population level estimate of the treatment effect of the IP on a binary endpoint regardless of participant compliance with IP dosing.* *Population-level summary: difference between treated and placebo in proportion of participants achieving erythema score of 1 or 0 in all affected anatomic regions among participants who have erythema score of 2 or more in at least 1 anatomic region at baseline.*
To assess the safety and tolerability of PF-06650833, PF-06700841, and PF-06826647 vs placebo in participants with HS.	*Incidence of TEAEs (AEs and SAEs), significant changes in vital signs, clinical laboratory abnormalities, and ECG.*	*There is no defined estimand for these endpoints and they will be analyzed using Pfizer data standards as applicable.*
To evaluate the effects of PF-06650833, PF-06700841, and PF-06826647 vs placebo on patient centered outcomes in participants with HS.	*Absolute score and CFB at time points specified in the SoA in HS Symptom Items and Dermatology Life Quality Index (DLQI) total score.* *Proportion of participants achieving a DLQI=0 or 1.*	*These endpoints will be analyzed descriptively and with respect to an estimand.*
To evaluate the PK of PF-06650833, PF-06700841, and PF-6826647 vs placebo in participants with HS.	*Summary of plasma concentration of PF-06700841, PF-06826647 and PF-06650833.*	*There is no defined estimand for these endpoints and they will be analyzed using Pfizer data standards as applicable.*
CCI [redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]

Table 2. Study Objectives and Endpoints

Objectives	*Endpoints*	*Estimands*
CCI		

Table 2. Study Objectives and Endpoints

Objectives	*Endpoints*	*Estimands*
CCI [redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]

**HiSCR requires:*

- *At least a 50% reduction in the total AN count relative to baseline; and*
- *No increase in abscess count; and*
- *No increase in draining fistula count.*

For all endpoints, baseline is defined as the result closest prior to dosing on Day 1.

2.1.1. Primary Estimand

The primary estimand E1 is intended to provide a population level of treated arm versus placebo arm on a binary endpoint. It includes the following 4 attributes:

- Population: Participants with moderate to severe HS, as defined by the inclusion and exclusion criteria.
- Variable: The proportion of participants achieving HiSCR response at Week 16.
- Non-pandemic related Intercurrent Events: Withdrawal and all other events leading to missing data will be treated as non-responders.
- COVID-19 pandemic related Intercurrent Events: The subjects who intermittently miss the visit due to COVID-19 will be excluded from the analysis for that visit. The subjects who discountinues treatment or withdraws from the study due to COVID-19 will be excluded from the analysis after the treatment discountinuation visit or study withdraw visit.

- Population-level summary: The mean difference in proportions of HiSCR responders between treated and placebo arms.

2.1.2. Secondary Estimands

The secondary estimand E2 is intended to provide a population level of treated arm versus placebo arm on a continuous endpoint. It includes the following 4 attributes:

- Population: Participants with moderate to severe HS, as defined by the inclusion and exclusion criteria.
- Variable: Percent CFB in AN count at Weeks 1, 2, 4, 6, 8, 12 and 16.
- Non-pandemic related Intercurrent Events: Withdrawal and all other events leading to missing data will be imputed in the active treatment arms using a jump to control method using the distribution of the placebo group (ie, missing data in active treatment participants will be imputed from the distribution of matching placebo participants). Missing data in a placebo arm will be imputed using data from the placebo arm assuming data are MAR.
- COVID-19 pandemic related Intercurrent Events: The subjects who miss visit, discountinues treatment, or withdraw from the study at or prior to Week 16 due to COVID-19 will not be excluded from the analysis. The missing data due to COVID-19 are assumed to be missing completely at random (MCAR).
- Population level summary: The mean difference in percent CFB in AN count between treated and placebo arms.

2.1.3. Additional Estimands

The estimand E3 is intended to provide a population level of treated arm versus placebo arm on time-to-event endpoint. It includes the following 4 attributes:

- Population: Participants with moderate to severe HS as defined by the inclusion and exclusion criteria.
- Viariable: Time to first flare, calculated from the first dose to the day that a flare is first observed.
- Non-pandemic related Intercurrent Events: Withdrawal subjects will be censored at the last available visit. Any other reason cause the missing value before first flare and prior the last visit will be censored.
- COVID-19 pandemic related Intercurrent Events: Subject who are observed to comply with treatment for at least 12 weeks will be included in the analysis. Missing values will be censored at the last available visit; Subjects who do not meet 12-week minimally acceptable level of drug compliance will be excluded from the analysis.

- Population-level summary: Hazard ratio of flare between treatment and placebo.

2.2. Study Design

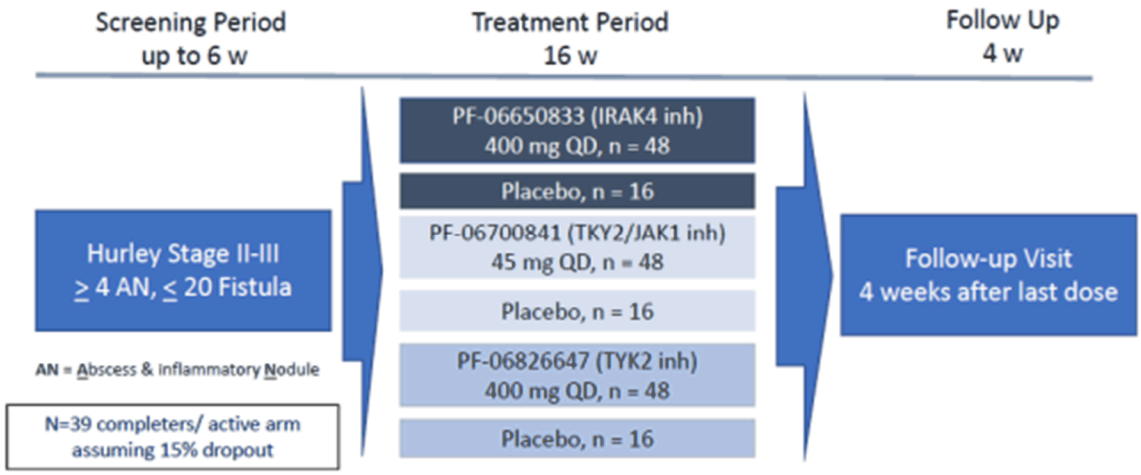
2.2.1. Design Overview

This is a Phase 2a, randomized, double-blind, placebo-controlled, parallel group, multicenter platform study with 3 kinase inhibitors (PF-06650833, PF-06700841 and PF-06826647) in participants with moderate to severe HS. The study will have a maximum duration of approximately 26 weeks. This includes an up to 6-week Screening Period, a 16-week Dosing Period and a 4-week Follow-up Period. The study will not be blinded across the PF-06650833, PF-06700841 and PF-06826647 treatment arms, but will be placebo-controlled double-blinded within each investigational product (IP) treatment arm.

Following the screening period, participants who meet eligibility criteria at the baseline visit, will be randomly assigned to receive 1 of 6 treatments. One oral dose level of each PF-06650833 (400 mg QD), PF-06700841 (45 mg QD) and PF-06826647 (400 mg QD) or matching placebo in a 3:1 ratio will be investigated. No more than 30% of enrolled participants will be inadequate anti-tumor necrosis factor (TNF) responders. Participants will be stratified according to whether they are an inadequate anti-TNF responder or not. The Figure 1 illustrates a schematic display of the study design.

Additionally, no more than 20% of enrolled participants may enter the study on a background of concomitant oral antibiotic therapy for treatment of HS; the dosing regimen (dose and frequency) must have been stable for at least 8 weeks (56 days) prior to the baseline (Day 1) visit and must remain stable throughout study participation. Participants will be stratified according to whether they are on a background of concomitant antibiotic therapy or not.

Figure 1. Study Design Schema



2.2.2. Number of Participants

The study will enroll a total of approximately 192 participants (expected to provide approximately 156 completers). The study will be conducted globally at approximately 60 study sites.

Sample size calculation is based on the primary endpoint of HiSCR response at 16 weeks. *A total of approximately 192 participants will be randomized in 3 active treatment groups (48/arm) and 3 placebo groups (16/arm) to have 39 completers for each active treatment arm, assuming a 15% dropout rate and the active treatment arm HiSCR response rate of 60% and placebo HiSCR rate of 30%. Statistical comparisons will be made between each of the active treatment arms against the 3 placebo groups pooled together. With one sided family wise error rate of 0.1 with a Bonferroni correction (0.033 after Bonferroni adjustment for 3 comparisons), this sample size will provide approximately 80% power. No statistical comparisons will be done between active treatment arms.*

3. ENDPOINTS AND BASELINE VARIABLES: DEFINITIONS AND CONVENTIONS

3.1. Primary Efficacy Endpoint

The primary efficacy endpoint is the percentage of participants with HiSCR response at Week 16. The HiSCR response is defined as:

- At least a 50% reduction in the total abscess and inflammatory nodule (AN) count relative to baseline; and
- No increase in abscess count; and
- No increase in draining fistula count.

3.2. Secondary Efficacy Endpoints

3.2.1. Binary Endpoints

Secondary binary endpoints are:

- Percentage of participants with HiSCR response at Weeks 1, 2, 4, 6, 8, and 12.
- Percentage of participants with a total AN count of 0 or 1; 0, 1, or 2 at Week 16.
- Proportion of participants who experience an HS flare at Weeks 4, 8, 12 and 16. A flare is defined as at least a 25% increase in AN count with a minimum increase of 2 relative to Baseline.

- Percentage of participants with ≥30% reduction and ≥1-unit reduction from baseline in PGA Skin Pain numeric rating scale (NRS30) – at worst and on average, respectively, amongst participants with baseline NRS ≥3, at Weeks 1, 2, 4, 6, 8, 12 and 16. The PGA of Skin Pain NRS will be used to assess the worst skin pain and the average skin pain due to HS. Ratings for the 2 items range from 0 to 10. last 7 assessments prior to the date of first dose of investigational product will be averaged and serve as the participant's baseline (See Protocol Section 8.10.2).

- Proportion of participants achieving erythema score of 1 or 0 in all affected anatomic regions among participants who have an erythema score of 2 or more in at least 1 anatomic region at baseline. The overall degree of erythema will be assessed for each anatomic region affected by HS using a four-point ordinal scale ranging between 0 (no redness), 1 (faint but discernible pink coloration), 2 (moderate red coloration), or 3 (very red or bright red coloration).

3.2.2. Continuous Endpoints

Secondary continuous endpoints are:

- Percent CFB in AN count at Weeks 1, 2, 4, 6, 8, 12 and 16.

- Percent CFB in NRS30, at worst and on average respectively, in participants who have baseline NRS ≥3, at Weeks 1, 2, 4, 6, 8, 12 and 16.

- CFB in NRS, at worst and on average respectively, at Weeks 1, 2, 4, 6, 8, 12 and 16.

- Absolute score and percent CFB in International Hidradenitis Suppurativa Severity Score System (IHS4) at Weeks 1, 2, 4, 6, 8, 12 and 16. The IHS4 score is calculated by the number of nodules, the number of abscesses, and the number of draining tunnels. IHS4 score will be calculated according to the SoA as below:

  IHS4 score = (number of nodules × 1) + (number of abscesses × 2) + {number of draining tunnels (fistulae/sinuses) × 4}

CCI [redacted]

- [redacted]

- [redacted]

[redacted]

3.4. Pharmacokinetic (PK) Endpoint

Pharmacokinetics (PK) endpoint includes the summary of plasma concentration of PF-06700841, PF-06826647 and PF-06650833.

3.5. Patient Reported Outcome Endpoints

The PGA skin pain NRS endpoints are the patient reported outcomes and key secondary endpoints which are described in the Section 3.2. Other PRO endpoints include:

- Absolute score and CFB in HS Symptom Items and Dermatology Life Quality Index (DLQI) total score at scheduled time. The HS symptoms items are 5 single items that will assess patient self-reported symptoms related to HS. The participants are asked to rate each symptom on a 0 to 10 numerical rating scale. The DLQI is a general dermatology questionnaire that consists of 10 items that assess patient health-related quality of life over the last week.

- Proportion of participants achieving a DLQI=0 or 1.

[redacted]

[redacted]

[redacted]

3.7. Safety Endpoints

Safety endpoints will be assessed by the spontaneous reporting of TEAEs (AEs and SAEs), significant changes in vital signs, clinical laboratory abnormalities, and ECG and clinical laboratory results in all subjects who receive at least one dose of the investigational product. Unscheduled clinical laboratory measurements may be obtained at any time during the study to assess any perceived safety concerns.

The COVID-19-related AEs may be summarized separately for subjects who becomes infected with the COVID-19 virus. The trial participants who missed safety assessment, discontinue treatment, or withdraw from the study due to COVID-19 will be considered as protocol deviations and will be documented. A separate listing of participants who are affected by COVID-19 study disruptions will be provided. The summary analysis maybe conducted to compare the rate of observed AEs in participants with COVID-19 in the active treatment arm to those in the control arm of the study.

3.7.1. Adverse Events

The AE endpoints will be assessed by the spontaneous reporting of:

- TEAEs (AEs and SAEs);
- Withdrawals from treatment due to AEs.

The definition and events meeting the definition of AE, SAE and TEAE is described in Appendix 2. The events will be classified as drug-related if the AE is classified as possibly, probably, or definitely related to study drug.

In additional to standard safety displays (Section 6.7.1), a 3-tier approach will be considered [redacted] to summarize AEs, SAEs and TEAEs. Under this approach, AEs are classified into 1 of 3 tiers. Different analyses will be performed for different tiers (see Section 6.7.1).

Tier 1 events: These are prespecified events of clinical importance and are maintained in a list in the product's Safety Review Plan.

Tier 2 events: These are events that are not tier 1 but are "common." A Medical Dictionary for Regulatory Activities (MedDRA) preferred term (PT) is defined as a Tier 2 event if there are at least 5% in any treatment group.

Tier 3 events: These are events that are neither Tier 1 nor Tier 2 events.

3.7.2. Laboratory Data

The safety endpoints will be assessed by the significant changes in clinical laboratory abnormalities. A list of safety laboratory test parameters is shown in the following:

Protocol-Required Safety Laboratory Assessments (to be update according to protocol amend)

Hematology	***Chemistry***	***Urinalysis***	***Other***
Hemoglobin	*BUN/urea, creatinine*	*pH*	*At screening only:*
Hematocrit	*and Cystatin C*	*Glucose (qual)*	*FSH[b]*
RBC count	*Glucose*	*Protein (qual)*	*Urine drug screening*
MCV	*Calcium*	*Blood (qual)*	*Pregnancy test (β-hCG)[c]*
MCH	*Sodium*	*Ketones*	*Hepatitis B surface antigen*
MCHC	*Potassium*	*Nitrites*	*(HBsAg)*
MPV	*Chloride*	*Leukocyte esterase*	*Hepatitis B core antibody (HBcAb)*
Platelet count	*Total*	*Urobilinogen*	*HepB reflex (HbsAb[d]), if*
WBC count	*CO_2 (bicarbonate)*	*Urine bilirubin*	*applicable*
Total neutrophils	*AST, ALT*	*Microscopy and/or*	*Hepatitis C antibody*
(Abs)	*Total bilirubin*	*urine culture[a]*	*Hepatitis C RNA, if applicable*
Eosinophils (Abs)	*Alkaline phosphatase*		*HIVUrine myoglobin[e]*
Monocytes (Abs)	*Uric acid*		
Basophils (Abs)	*Albumin*		
Lymphocytes (Abs)	*Total protein*		
	CK		

Abbreviations: Abs = absolute; ALT = alanine aminotransferase; AST = aspartate aminotransferase; β-hCG = beta-human chorionic gonadotropin; BUN = blood urea nitrogen; CO_2 = carbon dioxide; FSH = follicle-stimulating hormone; MCH = mean corpuscular hemoglobin; MCHC = mean corpuscular hemoglobin concentration; MCV = mean corpuscular volume; MPV = mean platelet volume; qual = qualitative; RBC = red blood cell; WBC = white blood cell.

a. Collect pre-dose urine sample for central laboratory urinalysis and urine microscopy at Screening and Baseline. At postdose visits: urine microscopy is indicated if the urinalysis is positive for blood, nitrite, leukocyte esterase and/or protein, or there is clinical suspicion of urinary tract infection, or decrease in renal function; urine culture is performed if urinalysis is positive for nitrite and/or leukocyte esterase, or if otherwise clinically indicated.
b. For confirmation of postmenopausal status only.
c. Local urine testing will be standard for the protocol unless serum testing is required by local regulation or institutional review board/ethics committee (IRB/EC). Serum or urine β-hCG for female participants of childbearing potential.
d. HepB reflex testing only if HBsAg negative but HBcAb positive at Screening.
e. At Screening and in case of CK >3 x ULN.

Additional laboratory results may be reported on these samples as a result of the method of analysis or the type of analyzer used by the clinical laboratory; or as derived from calculated values.

3.7.3. Vital Signs

The safety endpoints will be assessed by the significant changes in vital signs. The vital signs will be measured including temperature (Oral, Tympanic, Axillary or Temporal), pulse rate (beats/min), blood pressure (mmHg).

3.7.4. Electrocardiograms

The safety endpoints will be assessed by the significant changes in ECG. The average of the triplicate ECG measurements collected pre dose on Day 1 will serve as each participant's baseline QTc value. Summary of analysis results is described in the Section 6.7.4.

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*

3.7.5. Physical Examinations

A full physical examination will include assessments of the general appearance, skin, head, eyes, ears, nose, throat, cardiovascular, respiratory, GI, and neurological systems.

A brief physical examination will include assessments of the skin, heart, lung, abdomen, and body systems with any symptoms reported by the study participants.

3.8. Baseline Variables

The following baseline variables will be summarized for each treatment group. Details of summary analyses are described in Section 6.9.1.

In addition, the summary of baseline variables will be performed for subjects who discontinue treatment or withdraw from study caused by COVID-19 pandemic.

Demographic characteristics:

- Baseline age (<18, 18-44, 45-75, ≥75; and continuous in years);
- Gender (female vs. male);
- Race (white, black, Asian, other);
- Baseline body weight (<40, 40-100, >100 kg; and continuous in kg);
- Baseline height (cm);
- Baseline Body Mass Index (<17.5, 17.5 to <25, 25 to <30, 30 to <40, and ≥40 kg/m^2);
- Screening smoking status (never smoked, ex-smoker, smoker);
- Alcohol use (yes, no).

Baseline disease characteristics:

- Prior anti-TNF treatment (yes vs. no);
- Concomitant use of antibiotics treatment (yes vs. no);
- Hurley stages (II vs. III);
- Total AN Counts;
- Draining fistula count;
- Modified Sartorius score;
- Patient global assessment of skin pain NRS;
- [redacted]
- [redacted]
- DLQI;
- [redacted]
- [redacted]
- [redacted]

[redacted]

[redacted]

For all endpoints except height and weight, baseline is defined as the result closest prior to dosing on Day 1. The baseline for height/weight and BMI are measurements at screening visit. For ECG, the average of the triplicate ECG measurements will serve as the participant's baseline value. If the Day 1 ECG is missing, then the screening ECG will serve as baseline value screening. Also for PGA Skin Pain NRS Items, the last 7 assessments prior to the date of first dose of investigational product will be averaged and serve as the participant's baseline. The baseline data measured at Day 1 are:

- Medical history & demography, medication history, Height/Weight.
- Vital signs, ECG.
- Hurley stage, abscess count, inflammatory nodule *count*, fistula count, erythema assessments, modified Sartorius scale.
- Patient global assessment of skin pain NRS, HS symptom items, CCI [redacted] DLQI, CCI [redacted]
- CCI [redacted]

The stratification factors are in the following and will be summarized by treatment groups:

- Prior treatment with anti-TNF (yes or no),
- Concomitant use of antibiotic therapy (yes or no).

4. ANALYSIS SETS

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

Population	***Description***
Full Analysis Set (FAS)	*All participants randomized and receiving at least one dose of IP.*
Per-protocol analysis set (PPAS)	*All participants randomized and receiving at least one dose of IP, with both baseline and Week 16 primary efficacy data, and without protocol deviations that were thought to impact the efficacy evaluation during the treatment period. All protocol deviations will be reviewed and assessed by the study team prior to database release.*
Safety	*All participants randomly assigned to IP and who take at least 1 dose of IP. Participants will be analyzed according to the product they actually received.*
PK concentration set	*All enrolled participants who take at least one dose of active PF-06700841, PF-06826647 or PF-06650833 and in whom at least one concentration value is reported.*
[redacted]	CCI [redacted]

Two additional analysis sets are defined in the following for endpoints corresponding to the pain reduction and erythema score:

The analysis set for pain reduction endpoints is the subset of FAS which defined as all participants randomized and receiving at least one dose of IP and with NRS ≥3 at baseline.

The analysis set for erythema endpoint is the subset of FAS which defined as all participants randomized and receiving at least one dose of IP, and have erythema score of 2 or more in at least 1 anatomic region at baseline.

5. GENERAL METHODOLOGY AND CONVENTIONS

The final analysis will be performed after all study participants completed treatment and follow-up period, and the CSR will be issued after then.

5.1. Hypotheses and Decision Rules

The null hypothesis for primary efficacy analysis is that the percentage of participants achieving HiSCR response at Week 16 is the same for the active treatment (PF-06650833, PF-06700841 or PF-06826647) and placebo.

The Hochberg correction method is used to control the overall familywise error rate of 0.1 for multiple comparison. The treatment will be considered superior to control if the difference is statistically significant at the overall 1-sided 0.1 level. No statistical comparisons will be performed between active treatment arms.

5.2. General Methods

5.2.1. Cochran-Mantel-Haenszel (CMH) Method with Minimum Risk (MR) Weighting Strategy

The treatment difference in proportions will be computed using the Cochran Mantel Hanzel (CMH) test with Minimum Risk (MR) weighting strategy,[1] adjusting for the stratification factor of prior anti-TNF failure status and concomitant use of antibiotics status. The treatment differences in proportions will be presented. Their corresponding 2-sided 80% CI and 90% CI will be estimated using the MR weighting strategy proposed by Mehrotra and Railkar (2000).[1] In Primary efficacy analysis, the Hochberg adjusted one-sided p-values[2] for pairwise comparison between each active treatment and placebo will be presented. In secondary efficacy CCI [redacted] analysis, the p-values will not be adjusted. The summary statistics including N, frequency will be tabulated as well.

5.2.2. ANCOVA Model

For continuous data, the N, frequency, mean, standard deviation, median, minimum, and maximum will be presented. Graphics may be used to present the data – specific details will be outlined in the study List of Table (LOT). The analysis of covariance (ANCOVA) model will be implemented for statistical testing, which includes terms of treatment group, the stratification factors (if there are sufficient number of subjects in each stratum, otherwise the stratification factors will be dropped from the model), and the baseline value as the independent variable. P-values will not be adjusted in secondary and explortary analysis.

5.2.3. MMRM Model

In addition, the mixed effect model repeated measurement (MMRM) model with an unstructured covariance matrix will be used for patient reported outcome analysis, which includes the fixed effects of treatment, visit, treatment-by-visit interaction and baseline value, along with patient as a random effect. P-values will not be adjusted in secondary and explortary analysis.

5.3. Methods to Manage Missing Data

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

5.3.1. Binary Data

Withdrawal and all other events, except for COVID-19, leading to missing data will be treated as non-responders.

The subjects who intermittently miss the visit due to COVID-19 will be excluded from the analysis for that visit. The subjects who discountinues treatment or withdraws from the study due to COVID-19 will be excluded from the analysis after the treatment discountinuation visit or study withdraw visit. The intercurrent events should be considered COVID-19 pandemic-related if they occur as a result of pandemic-related factors and are not attributed

to other non-pandemic related reasons, eg, treatment discountinuation due to lack of efficacy or a toxicity. For individual instances where efficacy endpoints are not collected related to COVID-19, the reasons for failing to obtain the efficacy assessment will be documented and provided in the listing, including participant's COVID-19 infection condition, study treatment accessibility, and participant's concomitant treatment for COVID-19 in case of infection.

Participants with inadequate compliance will be handled case by case, based on the later discussion with clinician and the team.

5.3.2. Continuous Data

Withdrawal and all other events, except for COVID-19, leading to missing data will be imputed using a control-based imputation method. PROC MI will first be called and a control-based method (implemented with the MNAR (Missing Not At Random) option) will impute missing placebo observations under the assumption data are missing at random (MAR) and impute missing treatment observations assuming they are similar to corresponding placebo patients. Imputation will use the full conditional specification (FCS) method, the imputed data for the placebo arm will be combined for the analysis.

When using MMRM analysis, only observed data will be used and missing data will not be imputed.

The subjects who intermittently miss the visit due to COVID-19 will be excluded from the analysis for that visit. The subjects who discountinues treatment or withdraws from the study due to COVID-19 will be excluded from the analysis after the treatment discountinuation visit or study withdraw visit. The intercurrent events should be considered COVID-19 pandemic-related if they occur as a result of pandemic-related factors and are not attributed to other non-pandemic related reasons, eg, treatment discountinuation due to lack of efficacy or a toxicity. For individual instances where efficacy endpoints are not collected related to COVID-19, the reasons for failing to obtain the efficacy assessment will be documented and provided in the listing, including participant's COVID-19 infection condition, study treatment accessibility, and participant's concomitant treatment for COVID-19 in case of infection.

Participants with inadequate compliance will be handled case by case, based on the later discussion with clinician and the the team.

5.3.3. PK CCI [redacted] Data

In summary statistics for pharmacokinetic, concentration values below the lower limit of quantification (LLOQ) will be set to zero. CCI [redacted] The LLOQ for various PK CCI [redacted] concentrations will be noted in all tables and listings.

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. If actual sampling time is missing, the protocol-stated nominal time will be used.

5.3.4. Missing Dates

In compliance with Pfizer standards, if the day of the month is missing for any date used in a calculation, the 1st of the month will be used to replace the missing date unless the calculation results in a negative time duration (eg, date of resolution cannot be prior to date of onset; if replacing resolution date with the 1st of the month results in a negative duration, the resolution date will be set to the onset date). Pfizer standards are similarly used if both month and day are missing (January 1 unless negative time duration).

If the start date is missing for an AE, the AE is considered to be treatment emergent unless the collection date is prior to the treatment start date.

5.3.5. NRS Missing Data

Daily NRS data will be multiply imputed (MI) for any subjects with one or more missing daily scores for use in the analysis of endpoints in NRS, at worst and on average, respectively. Daily data will be imputed using a hierarchical Bayesian model with separate trends for each treatment arm modeled flexibly with a B-spline basis, a random subject effect and a first order autoregressive structure for temporal correlation. Separate models will be fitted for average and worst pain. The posterior predictive distribution (PPD) from this model will be used to impute missing pain scores and create 100 imputed datasets. The imputed daily scores will then be averaged over appropriate time periods for the analysis defined on weekly scores (see Section 3.2).

Participants with missing baseline data in NRS will be excluded from the analysis.

6. ANALYSES AND SUMMARIES

6.1. Primary Efficacy Analysis

6.1.1. Percentage of Participants Achieving HiSCR Response at Week 16.

6.1.1.1. Main Analysis

- Estimand strategy: Primary estimand E1 (Section 2.1.1).
- Analysis set: Full Analysis Set (Section 4).
- Analysis methodology: The number and percentage of participants with HiSCR response at Week 16 will be presented. The HiSCR response rate comparing treatment and placebo groups will be analyzed using the CMH test with MR weighting strategy (Section 5.2.1). Adjusted p-values for multiplicity using Hochberg correction[2] will be presented for each comparison.

- Intercurrent events and missing data: handling of missing data is described in Section 5.3.1.

- Rate (%) differences in HiSCR responders and the corresponding 2-sided 80% and 90% confidence intervals will be presented.

- The number and percent of HiSCR responders will be presented for each treatment group at Week 16.

- The number and percent of HiSCR responders will be presented for each treatment group by stratification factor at Week 16.

- A plot of number and percent of HiSCR responders for each treatment group at Week 16 will be presented. 80% and 90% CI will be calucated using Blyth-Still-Casella method.

- Plot of differences between active treatment and placebo adjusted by the stratification factor with 80% and 90% CIs will be presented.

- The number of missed or unusable assessments during COVID-19 pandemic and the reason of pandemic-related missing data will be summarized for primary endpoint.

6.1.1.2. Sensitivity Analysis

The main analysis will be repeated by using the per-protocol analysis set (PPAS) as defined in Section 4. Rate differences may also be analyzed using the method of Chan and Zhang (1999) in PROC FREQ using the PPAS and the FAS.

Sensitivity analysis may be performed excluding subject, who had used anitibiotic/analgesic during the study due to Adeverse events/topical treatment due to Adverse Events.

Handling of missing data and intercurrent event will be same as main analysis.

6.1.1.3. Supplementary Analysis

Main analysis will be repeated where both non-pandemic and pandemic-related missing response assessment are imputed as non-responder.

6.2. Secondary Efficacy Analysis

6.2.1. Percentage of Participants with HiSCR Response at Weeks 1, 2, 4, 6, 8, and 12

6.2.1.1. Main Analysis

- Estimand strategy: Primary estimand E1 (Section 2.1.1).

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

- Analysis methodology: The number and percentage of participants with HiSCR response at Weeks 1, 2, 4, 6, 8, and 12 will be presented. The HiSCR response rate comparing treatment and placebo groups at Week 8 and 12 will be analyzed using CMH test with MR weighting strategy (Section 5.2.1), but no hypothesis test will be performed to compare the treatment against the pooled placebo group at any earlier time point. P-values will not be adjusted for secondary endpoints.

- Intercurrent events and missing data: handling of missing data is described in Section 5.3.1.

- Rate (%) differences in HiSCR responders and the corresponding 2-sided 90% confidence intervals will be presented.

- The number and percent of HiSCR responders will be presented for each treatment group.

- The number and percent of HiSCR responders will be presented for each treatment group by stratification factor.

- A plot of number and percent of HiSCR responders for each treatment group will be presented. 90% CI will be calucated using Blyth-Still-Casella method.

- Plot of differences between active treatment and placebo adjusted by the stratification factor with 90% CIs over time will be presented.

6.2.1.2. Sensitivity Analysis

The main analysis will be repeated by using the per-protocol analysis set (PPAS) as defined in Section 4. Rate differences may also be analyzed using the method of Chan and Zhang (1999) in PROC FREQ using the PPAS and the FAS.

6.2.2. Percent CFB in AN Count at Weeks 1, 2, 4, 6, 8, 12 and 16

6.2.2.1. Main Analysis

- Estimand strategy: Secondary estimand E2 (Section 2.1.2).

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

- Analysis methodology: Percent CFB in AN count at Weeks 1, 2, 4, 6, 8, 12 and 16 will be presented. To compare the treatment and placebo groups at Week 8, Week 12 and Week 16, respectively, the ANCOVA model will be fitted which includes terms of treatment group, the stratification factors (if there are sufficient number of subjects in each stratum), and the baseline value as the independent variable, and with the observed (or imputed) percent CFB in AN count at Week 8, Week 12 and Week 16 as the dependent variable, but no hypothesis test will be performed to compare the treatment against the pooled placebo group at any earlier

time point. Treatment comparisons will be performed for the treatment difference in least square means. The non-adjusted p-values for treatment comparison will be presented.

- Intercurrent events and missing data: handling of missing data is described in Section 5.3.2.
- The least-square means by treatment group and visit, the treatment pairwise differences with their corresponding 90% confidence limits will be displayed.
- The N, mean, standard deviation, median, minimum, and maximum will be presented for percent CFB in AN count by treatment group and time.
- Plot of CFB in AN count by treatment group and time will be presented.
- Plot of differences between active treatment and placebo comparison with 90% CIs will be presented.

6.2.2.2. Sensitivity Analysis

- The MMRM (Section 5.2.3) will be fitted to compare the difference of active treatments with placebo participants using all data. The model will include the fixed effects of treatment, visit, treatment-by-visit interaction and baseline value, along with patient as a random effect. The p-values of comparisons between treatment groups to placebo will be calculated.
- Intercurrent events and missing data: Observed data (see Section 5.3.2).

6.2.3. Percentage of Participants with a Total AN Count of 0 or 1; 0, 1, or 2 at Week 16

Same as Section 6.1.1.

6.2.4. Percentage of Participants with ≥30% Reduction and ≥1-unit Reduction from Baseline in PGA Skin Pain Numeric Rating Scale (NRS30) – at Worst and on Average, Respectively, amongst Participants with Baseline NRS ≥3, at Weeks 1, 2, 4, 6, 8, 12 and 16

- Estimand strategy: Primary estimand E1 (Section 2.1.1).
- Analysis set: Full Analysis Set with baseline NRS ≥3 (Section 4).
- Analysis methodology: The number and percentage of participants with ≥30% reduction and ≥1-unit reduction from baseline in PGA Skin Pain NRS30 will be presented. The rate comparing treatment and placebo groups at each visit will be analyzed using CMH test with MR weighting strategy (Section 5.2.1) between each of active treatment group and placebo. The analysis will be repeated for each

imputed dataset and the results combined using Rubin's rules. P-values will not be adjusted for secondary endpoints.

- Intercurrent events and missing data: handling of missing data is described in Section 5.3.5.
- Rate (%) differences and the corresponding 2-sided 90% confidence intervals will be presented.
- The number and percent of responders will be presented for each treatment group.
- The number and percent of responders will be presented for each treatment group by stratification factors.
- A plot of number and percent of responders for each treatment group will be presented.
- Plot of differences between active treatment and placebo with 90% CIs will be presented.

6.2.5. Percent CFB in NRS30, at Worst and On Average Respectively, in Participants who have Baseline NRS ≥3, at Weeks 1, 2, 4, 6, 8, 12 and 16

- Estimand strategy: Secondary estimand E2 (Section 2.1.2).
- Analysis set: Full Analysis Set with baseline NRS ≥3 (Section 4).
- Analysis methodology: Percent CFB in NRS30, at worst and on average respectively, at Weeks 1, 2, 4, 6, 8, 12 and 16 will be presented. To compare the treatment and placebo groups at each visit, at worst and on average respectively, the ANCOVA model will be fitted which includes terms of treatment group, the stratification factors (if there are sufficient number of subjects in each stratum), and the baseline value as the independent variable, and with the percent CFB in NRS30 at each visit as the dependent variable (computed from observed and imputed daily scores as needed). Treatment comparisons will be performed for the treatment difference in least square means. The analysis will be repeated for each imputed dataset and the results combined using Rubin's rules. The non-adjusted p-values for treatment comparison will be presented.
- Intercurrent events and missing data: handling of missing data is described in Section 5.3.5.
- The least-square means by treatment group and visit, the treatment pairwise differences with their corresponding 90% confidence limits will be displayed.

- The N, mean, standard deviation, median, minimum, and maximum will be presented for percent CFB in NRS, at worst and on average respectively, by treatment group and time.

- Plot of percent CFB in NRS, at worst and on average respectively, by treatment group and time will be presented.

- Plot of differences between active treatment and placebo comparison with 90% CIs will be presented.

6.2.6. CFB in NRS, at Worst and On Average Respectively, at Weeks 1, 2, 4, 6, 8, 12 and 16

- Estimand strategy: Secondary estimand E2 (Section 2.1.2).

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

- Analysis methodology: Same as Section 6.2.5.

6.2.7. Proportion of Participants Achieving Erythema Score of 1 or 0 in All Affected Anatomic Regions among Participants who have an Erythema Score of 2 or More in at least 1 Anatomic Region at Baseline

- Estimand strategy: Primary estimand E1 (Section 2.1.1).

- Analysis set: Full Analysis Set with an erythema score of 2 or more in at least 1 anatomic region at baseline (Section 4).

- Analysis methodology: The number and percentage of participants achieving erythema response will be presented. The rate comparing treatment and placebo groups at Week 8, Week 12 and Week 16 will be analyzed using CMH test with MR weighting strategy (Section 5.2.1) between each of active treatment group and placebo. P-values will not be adjusted for secondary endpoints.

- Intercurrent events and missing data: handling of missing data is described in Section 5.3.1 and Section 2.1.1.

- Rate (%) differences and the corresponding 2-sided 90% confidence intervals will be presented.

- The number and percent of responders will be presented for each treatment group.

- The number and percent of responders will be presented for each treatment group by stratification factors.

- A plot of number and percent of responders for each treatment group will be presented.

- Plot of differences between active treatment and placebo with 90% CIs will be presented.

6.2.8. Absolute Score and Percent CFB in International Hidradenitis Suppurativa Severity Score System (IHS4) at Weeks 1, 2, 4, 6, 8, 12 and 16

- Estimand strategy: Secondary estimand E2 (Section 2.1.2).
- Analysis set: Full Analysis Set (Section 4).
- Analysis methodology: Absolute score and percent CFB in IHS4 at Weeks 1, 2, 4, 6, 8, 12 and 16 will be presented. To compare the treatment and placebo groups at Week 8, Week 12 and Week 16, the ANCOVA model will be fitted which includes terms of treatment group, the stratification factors (if there are sufficient number of subjects in each stratum), and the baseline value as the independent variable, and with the observed (or imputed) absolute score or percent CFB in IHS4 as the dependent variable. Treatment comparisons will be performed for the treatment difference in least square means. The non-adjusted p-values for treatment comparison will be presented.
- Intercurrent events and missing data: handling of missing data is described in Section 5.3.2.
- The least-square means by treatment group and visit, the treatment pairwise differences with their corresponding 90% confidence limits will be displayed.
- The N, mean, standard deviation, median, minimum, and maximum will be presented for absolute score and percent CFB in IHS4 by treatment group and time.
- Plot of absolute score and percent CFB in IHS4 by treatment group and time will be presented.
- Plot of differences between active treatment and placebo comparison with 90% CIs will be presented.

6.2.9. Proportion of Participants who Experience an HS Flare at Weeks 4, 8, 12 and 16

- Estimand strategy: Primary estimand E1 (Section 2.1.1).
- Analysis set: Full Analysis Set (Section 4).
- Analysis methodology: The number and percentage of participants who experience an HS flare at Weeks 4, 8, 12 and 16 will be presented for each treatment group by time. The rate comparing treatment and placebo groups at Week 8, Week 12 and Week 16 will be analyzed respectively, using CMH test with MR weighting strategy (Section 5.2.1) between each of active treatment group and placebo. P-values will not be adjusted for secondary endpoints.

- Intercurrent events and missing data: missing data in HS flare will no be imputed.
- Rate (%) differences and the corresponding 2-sided 90% confidence intervals will be presented.
- The number and percent of responders will be presented for each treatment group.
- The number and percent of responders will be presented for each treatment group by stratification factors.
- A plot of number and percent of responders for each treatment group will be presented.
- Plot of differences between active treatment and placebo with 90% CIs will be presented.

CCI [redacted]

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

CCI [redacted]

- [redacted]
- [redacted]

[redacted]

[redacted]

CCI

6.4. Analysis of Pharmacokinetic Endpoints

The PK analysis will be performed on the PK concentration population which defined as all enrolled participants who received at least one dose of PF-06700841, PF-06826647 or PF-06650833 and in whom at least one concentration value is reported.

- Analysis set: PK Concentration Set.
- Analysis methodology: PK concentrations will be summarized and presented with descriptive statistics by treatment group and visit. Population PK modeling may be performed with the concentration data from this study alone or combined with data from other studies. In addition, a relationship between exposures and efficacy/safety endpoints may be evaluated using population PK/PD approach. Any population analyses conducted will not be part of the clinical study report (CSR) and may be reported separately.
- Intercurrent events and missing data: the missing data will not be imputed. The subjects who miss visit due to COVID-19 will be excluded from the analysis for that visit. The subjects who discountinues treatment or withdraws from the study due to COVID-19 will be excluded from the analysis after the treatment discountinuation visit or study withdraw visit. The value below limit of quantification are discussed in Section 5.3.3.

6.5. Analysis of Patient Reported Outcome Endpoints

6.5.1. Absolute Score and CFB in HS Symptom Items and Dermatology Life Quality Index (DLQI) Total Score at Scheduled Time

- Estimand strategy: Secondary estimand E2 (Section 2.1.2).
- Analysis set: Full Analysis Set (Section 4).

- Analysis methodology: Absolute and CFB score in HS Symptom Items and Dermatology Life Quality Index (DLQI) total score for each treatment group will be presented descriptively by visit. The MMRM (Section 5.2.3) will be used to compare the difference of active treatments with placebo participants at each visit. The model will include the fixed effects of treatment, visit, treatment-by-visit interaction and baseline value, along with patient as a random effect

- Intercurrent events and missing data: Observed data.

- The N, mean, standard deviation, median, minimum, and maximum will be presented for absolute and CFB score, by treatment group and visit.

- Mean (SE) plot of absolute and CFB score will be presented by treatment group.

6.5.2. Proportion of Participants Achieving a DLQI=0 or 1

- Estimand strategy: Primary estimand E1 (Section 2.1.1).

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

- Analysis methodology: The number and percentage of participants achieving DLQI = 0 and 1 will be presented, respectively. The rate comparing each active treatment and placebo at each visit will be analyzed using CMH test with MR weighting strategy (Section 5.2.1). P-values will not be adjusted for secondary endpoints.

- Intercurrent events and missing data: handling of missing data is described in Section 5.3.1.

- Rate (%) differences and the corresponding 2-sided 90% confidence intervals will be presented.

- The number and percent of responders will be presented for each treatment group.

- The number and percent of responders will be presented for each treatment group by stratification factors.

- A plot of number and percent of responders for each treatment group will be presented.

- Plot of differences between active treatment and placebo with 90% CIs will be presented.

CCI [redacted]

CCI

CCI

CCI

CCI

[redacted]

6.7. Safety Summaries and Analyses

All safety analyses will be performed on the safety population and summarized in accordance with Pfizer Data Standards.

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 participants. The COVID-19-related AEs will be summarized according to Section 3.7.

Safety data will be summarized descriptively through appropriate data tabulations, descriptive statistics, categorical summaries, and graphical presentations, where appropriate. 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 (CFB) in laboratory data, and vital signs will also be summarized. Participant listings will be produced for the safety endpoints accordingly.

6.7.1. Adverse Events

Generally, the TEAEs (AEs and SAEs) will be summarized for each treatment group and overall total. A summary of the number and percentage of patients reporting TEAEs and number of TEAE events will be presented overall, as well as for study drug-related TEAEs. The TEAEs leading to study drug interruption and withdrawal will be summarized overall, as well as for drug-related TEAEs. The incidence of TEAEs normalized for duration of exposure will be presented.

In addition, the AEs, SAEs, TEAEs will be analyzed based on 3-tier approach (Section 3.7.1). The analyses of AEs under the 3-tier approach are considered [redacted]. The following descriptions are for the analyses of AEs under the 3-tier approach:

- Analysis Method: The risk differences and 95% confidence intervals are provided for comparing each active treatment dose versus placebo.

- There will be no adjustment for multiple comparisons or stratification factors in the analyses.

- Output: For Tier 1 and Tier 2 event outputs, footnotes will be included to provide proper interpretation of nominal p-values (Tier 1 events) and 95% confidence intervals (Tier 1 and Tier 2 events) for pairwise treatment differences in percentage of participants with events.

- For tier-3 events, observed event counts and proportions will be summarized.

- The output will be arranged in descending tier order and the descending point estimate the risk differences.

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. As such, safety analysis is generally considered as an CCI analysis and its purpose is to generate hypotheses for further investigation. The 3-tier approach facilitates this CCI analysis.

6.7.2. Laboratory Data

Unscheduled clinical labs may be obtained at any time during the study to assess any perceived safety concerns at the investigator's discretion. Laboratory data will be listed and summarized in accordance with the Pfizer reporting standards.

6.7.3. Vital Signs including Weight

Vital signs (systolic blood pressure, diastolic blood pressure pulse, pulse rates and temperature) will be measured after 5 minutes of rest as indicated in the Schedule of Activities.

The actual outcome and CFB value will be summarized at baseline and over time by treatment in accordance with the Pfizer reporting standards.

6.7.4. Electrocardiograms

The actual 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*

In addition, the number of participants with uncorrected QT values >500 msec will be summarized.

6.7.5. Physical Examination

Full and brief physical exams will be summarized separately in accordance with the Pfizer reporting standards.

6.8. Subset Analyses

The proportion of participants achieving HiSCR response at Week 16 will be summarized for the subsets below. The summary statistics will be presented.

- Gender (male vs. female),
- Race (white, black, or other),
- BMI (below or above median of observed data),
- Prior anti-TNF treatment (yes vs. no),
- Antibiotics treatment (yes vs. no),
- Hurley stages (II vs. III),
- Median disease duration (below or above median of observed data),
- AN Counts (below or above median of observed data),
- Number of draing/non-draing fistula at baseline (below or above median of observed data),
- Modified Sartorius score (below or above median of observed data),
- Surgical intervention/drainage required (yes vs. no).

6.9. Baseline and Other Summaries and Analyses

6.9.1. Baseline Summaries

Demographics and baseline disease characteristics defined in Section 3.8 will be summarized by treatment group according to Pfizer standards.

6.9.2. Study Conduct and Participant Disposition

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

6.9.3. Study Treatment Exposure

A summary of compliance and the number of doses received as well as the median total dose by visit and treatment group will be provided.

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

6.9.4. Concomitant Medications and Nondrug Treatments

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

7. INTERIM ANALYSES

An interim analysis may be performed to assess efficacy and safety. Details of the interim analysis will be described in the Interim Analysis Plan. Interim analysis results may be used for internal business decisions regarding future study planning. Before any interim analysis is initiated, the details of the objectives, decision criteria, dissemination plan, and method of maintaining the study blind as per Pfizer's SOPs will be documented and approved in an interim analysis plan. The results will only be distributed to a select list of individuals who are not part of the study team and will be involved in the internal decision-making process in order to protect the integrity of the study. This list of individuals will be provided in the interim analysis plan. The results of the interim analysis will not enable individuals directly involved in running the study (such as investigators) to identify treatment assignments for individual participants remaining in the study. There are no prospective plans to stop the study early for success as a result of any interim analysis.

8. REFERENCES

1. Mehrotra, Devan V., and Radha Railkar. "Minimum risk weights for comparing treatments in stratified binomial trials." Statistics in medicine 19.6 (2000): 811-825.

2. Yosef Hochberg. "A sharper Bonferroni procedure for multiple tests of significance." Biometrika 75.4 (1988): 800–802.

9. APPENDICES

Appendix 1. Definition and Use of Visit Windows in Reporting

Visit Label	Target Day	Start Day	End Day
Screening	NA	Day -42	Day -1
Baseline	Day 1	Day -1	Day 1
Week 1	Day 8	Day 2	Day 11
Week 2	Day 15	Day 12	Day 21
Week 4	Day 29	Day 22	Day 35
Week 6	Day 43	Day 36	Day 49
Week 8	Day 57	Day 50	Day 71
Week 12	Day 85	Day 72	Day 99
Week 16	Day 113	Day 100	Day 127
Week 20	Day 141	Day 128	Day 154

Appendix 2. Definition of AE, SAE and TEAE

AE Definition

- An AE is any untoward medical occurrence in a patient or clinical study participant, temporally associated with the use of study intervention, whether or not considered related to the study intervention.
- NOTE: An AE can therefore be any unfavorable and unintended sign (including an abnormal laboratory finding), symptom, or disease (new or exacerbated) temporally associated with the use of study intervention.

Events Meeting the AE Definition

- Any abnormal laboratory test results (hematology, clinical chemistry, or urinalysis) or other safety assessments (eg, ECG, radiological scans, vital sign measurements), including those that worsen from baseline, considered clinically significant in the medical and scientific judgment of the investigator (ie, not related to progression of underlying disease).
- Exacerbation of a chronic or intermittent preexisting condition including either an increase in frequency and/or intensity of the condition.
- New conditions detected or diagnosed after study intervention administration even though it may have been present before the start of the study.
- Signs, symptoms, or the clinical sequelae of a suspected drug-drug interaction.

- Signs, symptoms, or the clinical sequelae of a suspected overdose of either study intervention or a concomitant medication. Overdose per se will not be reported as an AE/SAE unless it is an intentional overdose taken with possible suicidal/self-harming intent. Such overdoses should be reported regardless of sequelae.

- "Lack of efficacy" or "failure of expected pharmacological action" per se will not be reported as an AE or SAE. Such instances will be captured in the efficacy assessments. However, the signs, symptoms, and/or clinical sequelae resulting from lack of efficacy will be reported as an AE or SAE if they fulfill the definition of an AE or SAE.

Events NOT Meeting the AE Definition

- Any clinically significant abnormal laboratory findings or other abnormal safety assessments which are associated with the underlying disease, unless judged by the investigator to be more severe than expected for the participant's condition.

- The disease/disorder being studied or expected progression, signs, or symptoms of the disease/disorder being studied, unless more severe than expected for the participant's condition.

- Medical or surgical procedure (eg, endoscopy, appendectomy): the condition that leads to the procedure is the AE.

- Situations in which an untoward medical occurrence did not occur (social and/or convenience admission to a hospital).

- Anticipated day-to-day fluctuations of preexisting disease(s) or condition(s) present or detected at the start of the study that do not worsen.

If an event is not an AE per definition above, then it cannot be an SAE even if serious conditions are met (ego, hospitalization for signs/symptoms of the disease under study, death due to progression of disease).

An SAE is defined as any untoward medical occurrence that, at any dose:

a. Results in death

b. Is life-threatening

The term "life-threatening" in the definition of "serious" refers to an event in which the participant was at risk of death at the time of the event. It does not refer to an event that hypothetically might have caused death if it were more severe.

c. Requires inpatient hospitalization or prolongation of existing hospitalization

In general, hospitalization signifies that the participant has been detained (usually involving at least an overnight stay) at the hospital or emergency ward for observation and/or treatment that would not have been appropriate in the physician's office or outpatient setting. Complications that occur during hospitalization are AEs. If a complication prolongs hospitalization or fulfills any other serious criteria, the event is serious. When in doubt as to whether "hospitalization" occurred or was necessary, the AE should be considered serious.

Hospitalization for elective treatment of a preexisting condition that did not worsen from baseline is not considered an AE.

d. Results in persistent disability/incapacity

- The term disability means a substantial disruption of a person's ability to conduct normal life functions.
- This definition is not intended to include experiences of relatively minor medical significance such as uncomplicated headache, nausea, vomiting, diarrhea, influenza, and accidental trauma (eg, sprained ankle) which may interfere with or prevent everyday life functions but do not constitute a substantial disruption.

e. Is a congenital anomaly/birth defect

f. Other situations:

- Medical or scientific judgment should be exercised in deciding whether SAE reporting is appropriate in other situations such as important medical events that may not be immediately life-threatening or result in death or hospitalization but may jeopardize the participant or may require medical or surgical intervention to prevent one of the other outcomes listed in the above definition. These events should usually be considered serious.
- Examples of such events include invasive or malignant cancers, intensive treatment in an emergency room or at home for allergic bronchospasm, blood dyscrasias or convulsions that do not result in hospitalization, or development of drug dependency or drug abuse.

The treatment-emergent AEs (TEAE) are considered as an adverse event that 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 will be flagged as TEAEs.

Appendix 3. SAS Code for Imputation of Continuous Endpoint

```
proc sort;
        by avisitn trtpn;
run;

proc mi data=mock_all2 seed=1007 nimpute=100 out=outimp minimum=0
maximum=&maxscore.;
        by avisitn;
        class trtpn;
        MONOTONE regpmm(aval=base /details k=5);
        mnar model(aval/modelobs=(trtpn='1') );
        var base aval;
run;
```

Appendix 4. SAS Code for CMH Test with MR Weighting Strategy Adjusted by Stratification Factors

```
PROC FREQ DATA=XXX;
  By visit;
  TABLE stratum*TREATMENT*RESP/ COMMONRISKDIFF(CL=MR Correct=NO
TEST=MR  PRINTWTS) alpha=0.1 ;
  Output out= XXXX;
RUN;
```

Appendix 5. SAS Code for Hochberg P-value

```
data analysis ;
    set rawpval ;
    p = treatment ;
    raw_p=pvalue;
proc sort ;
    by visit ;
run;
proc multtest inpvalues=analysis out=adjpval hoc;
    by visit ;
run;
```

Appendix 6. SAS Code for the Confidence Interval of a Binomial Proportion (Blyth Still Casella)

PROC BINOMIAL DATA=<DATASET> ALPHA=<value>;

BI/BS;

OU <RESPONSE VARIABLE>;

RUN;

Appendix 7. List of Abbreviations

The following is a list of abbreviations that may be used in the SAP.

Abbreviation	Term
AA	alopecia areata
Abs	absolute
ACPA+	anti-anticitrullinated protein antibody positive
ADME	absorption, distribution, metabolism, and excretion
AE	adverse event
ALC	absolute lymphocyte count
ALT	alanine aminotransferase
AN	abscess and inflammatory nodule
ANC	absolute neutrophil count
AST	aspartate aminotransferase
AUC	area under the curve
AUC_{24}	area under the curve at 24 hours
AUC_{inf}	area under the curve time zero to infinity
AV	atrioventricular
BA	bioavailability
CCI	[redacted]
BCG	Bacille Calmette Guérin
BCRP	breast cancer resistance protein
β-hCG	beta-human chorionic gonadotropin
BID	twice daily
BMI	body mass index
BP	blood pressure
bpm	beats per minute
BSEP	bile salt export pump
BUN	blood urea nitrogen
$C_{av,ss}$	average amount at steady state
CFB	change from baseline
CFR	Code of Federal Regulations
CI	confidence interval
CIOMS	Council for International Organizations of Medical Sciences
CK	creatinine kinase
CKD-EPI	Chronic Kidney Disease Epidemiology Collaboration
CL/F	apparent clearance
C_{max}	maximum observed concentration
C_{min}	minimum observed concentration
CO_2	carbon dioxide (bicarbonate)
CONSORT	Consolidated Standards of Reporting Trials
CRF	case report form
CrI	credible interval

Abbreviation	Term
CRO	contract research organization
CSR	clinical study report
C-SSRS	Columbia Suicide Severity Rating Scale
CT	clinical trial
CTCAE	Common Terminology Criteria for Adverse Events
CTMS	clinical trial management system
CV	cardiovascular
CYP	cytochrome P450
DCT	data collection tool
DILI	drug-induced liver injury
DLQI	Dermatology Life Quality Index
CCI [redacted]	[redacted]
dn	dose normalized
DNA	deoxyribonucleic acid
DU	dispensable unit
EBV	Epstein-Barr virus
EC	ethics committee
ECG	electrocardiogram
eCRF	electronic case report form
EDP	exposure during pregnancy
EFD	embryofetal developmental
eGFR	estimated glomerular filtration rate
EMA	European Medicines Agency
EPO	epogen
ePRO	electronic patient reported outcomes
CCI [redacted]	[redacted]
ET	early termination
EU	European Union
EudraCT	European Clinical Trials Database
CCI [redacted]	[redacted]
FAS	full analysis set
FIH	first in human
FSH	follicle-stimulating hormone
f_u	fraction of drug unbound in plasma
GCP	Good Clinical Practice
GFR	glomerular filtration rate
GGT	gamma-glutamyl transferase
GI	gastrointestinal
GLP	Good Laboratory Practice
HBcAb	hepatitis B core antibody
HbsAb	hepatitis B surface antibody
HbsAg	hepatitis B surface antigen

Abbreviation	Term
HCVAb	hepatitis C virus antibody
HIPAA	Health Insurance Portability and Accountability Act
HiSCR	hidradenitis suppurativa clinical response
HIV	human immunodeficiency virus
HRQoL	health-related 53uality of Life
HRT	hormone replacement therapy
HS	hidradenitis suppurativa
CCI	[redacted]
IB	investigator's brochure
IBD	inflammatory bowel disease
IC	inhibitory concentration
ICD	informed consent document
ICH	International Council for Harmonisation
ID	identification
IFN	interferon
IGRA	Interferon Gamma Release Assay
IL	interleukin
IMP	investigational medicinal product
IND	investigational new drug application
INR	international normalized ratio
IP	investigational product
IP manual	investigational product manual
IR	immediate release
IRAK4	IL -1 receptor associated kinase 4
IRB	institutional review board
IRC	internal review committee
IRT	interactive response technology
IUD	intrauterine device
IUS	intrauterine hormone-releasing system
IWR	interactive Web-based response
JAK1	Janus Kinase 1
KDIGO	Kidney Disease: Improving Global Outcomes
LASER	light amplification by stimulated emission of radiation
LBBB	left bundle branch block
LFT	liver function test
LLOQ	lower limit of quantification
MAD	multiple ascending dose
MAR	missing at random
MATE	multidrug and toxin extrusion transporter
MCH	mean corpuscular hemoglobin
MCHC	mean corpuscular hemoglobin concentration
MCS	mental component scores

Abbreviation	Term
MCV	mean corpuscular volume
MDR1	multidrug resistant protein 1
MMR	Measles Mumps Rubella
MPV	mean platelet volume
MR	Minimum Risk
MRI	magnetic resonance imaging
CCI [redacted]	[redacted]
MRP	multidrug resistance-associated protein
msec	millisecond
MTX	methotrexate
N/A	not applicable
NIMP	noninvestigational medicinal product
NOAEL	no-observed-adverse-effect level
NRS	numeric rating scale
NRS-11	Numerical Rating Scale-11
NSAIDs	nonsteroidal anti-inflammatory drugs
NTCP	sodium/taurocholate co-transporting polypeptide
OAT	organic anion transporter
OATP	organic anion transporting polypeptide
OCT	organic cation transporter
PASI	Psoriasis Area and Severity Index
PBMC	peripheral blood mononuclear cell
PCD	primary completion date
PCP	primary care physician
PCR	polymerase chain reaction
PCS	physical component score
PD	pharmacodynamic(s)
PE	physical exam
PGA	Patient's Global Assessment
CCI [redacted]	[redacted]
CCI [redacted]	[redacted]
PI	principal investigator
PK	pharmacokinetic(s)
PPAS	per protocol analysis set
PPD	Purified Protein Derivative
PRN	as needed
PRO	patient reported outcome
PsO	psoriasis
PT	prothrombin time
PVC	premature ventricular contraction/complex
QD	once daily
QID	four times a day

Abbreviation	Term
QoL	quality of life
QTc	corrected QT
QTcF	corrected QT (Fridericia method)
QW	once weekly
qual	qualitative
RA	rheumatoid arthritis
RBC	red blood cell
RNA	ribonucleic acid
SAD	single ascending dose
SAE	serious adverse event
SAP	statistical analysis plan
SCr	serum creatinine
SDAI	simplified disease activity index
SDD	sprayed dried dispersion
CCI	[redacted]
SMP	study monitoring plan
SoA	schedule of activities
SOC	system organ class
SOP	standard operating procedure
SRSD	single reference safety document
SSID	screening number
SToD	study team on demand
SUSAR	suspected unexpected serious adverse reaction
$T_{1/2}$	half-life
TB	tuberculosis
TBili	total bilirubin
CCI	[redacted]
TdP	torsade de pointes
TEAEs	treatment emergent adverse events
CCI	[redacted]
TID	three time a day
TLR	tool-like receptor
T_{max}	time to maximal concentration
TNF	tumor necrosis factor
TYK2	tyrosine kinase 2
UGT	5'-diphospho-glucuronosyltransferase
ULN	upper limit of normal
US	United States
UVA	Ultra violet A
CCI	[redacted]
VTE	venous thromboembolic events
Vz/F	volume of distribution

Abbreviation	Term
WBC	white blood cell
WOCBP	woman of childbearing potential