Protocol B7931023

A PHASE 2B, RANDOMIZED, DOUBLE BLIND, VEHICLE-CONTROLLED, PARALLEL-GROUP, DOSE RANGING STUDY TO ASSESS EFFICACY, SAFETY, TOLERABILITY AND PHARMACOKINETICS OF PF-06700841 TOPICAL CREAM APPLIED ONCE OR TWICE DAILY FOR 12 WEEKS IN PARTICIPANTS WITH MILD TO MODERATE CHRONIC PLAQUE PSORIASIS

Statistical Analysis Plan
(SAP)

Version: 3.0

Author: PPD

Date: 04-MAY-2021

TABLE OF- CONTENTS

LIST OF TABLES 4
LIST OF FIGURES 4
APPENDICES 4
1. VERSION HISTORY 5
2. INTRODUCTION 5
2.1. Study Objectives, Endpoints, and Estimands 7
2.1.1. Primary Estimand(s) 10
2.1.2. Secondary Estimand(s) 11
2.1.3. Additional Estimand(s) 11
2.2. Study Design 12
3. ENDPOINTS AND BASELINE VARIABLES: DEFINITIONS AND CONVENTIONS 14
3.1. Primary Endpoint(s) 14
3.2. Secondary Endpoint(s) 14
CCI
3.4. Baseline Variables 16
3.5. Safety Endpoints 17
3.5.1. Adverse Events 17
3.5.2. Laboratory Data 17
3.5.3. Vital Signs, including Height and Weight 18
3.5.4. Physical Examinations 18
4. ANALYSIS SETS (POPULATIONS FOR ANALYSIS) 18
4.1. Full Analysis Set (FAS) 18
4.2. Efficacy Analysis Set (EAS) 18
4.3. Safety Analysis Set 18
CCI
4.5. Treatment Misallocations and Missing Data 19
5. GENERAL METHODOLOGY AND CONVENTIONS 20
5.1. Hypotheses and Decision Rules 20
5.2. General Methods 21
5.2.1. Analyses for Binary Endpoints 21
5.2.2. Analyses for Continuous Endpoints 21

5.2.3. Analyses for Categorical Endpoints 22
5.2.4. Analyses for Time-to-Event Endpoints 22
5.2.5. Analyses of Safety Events 22
5.2.6. Methods to Manage Missing Data 22
6. ANALYSES AND SUMMARIES 23
6.1. Primary Endpoint(s) 23
6.1.1. Change from Baseline PASI at Week 12 23
6.1.1.1. Main Analysis 23
6.1.1.2. Sensitivity/Supplementary Analyses 24
6.2. Secondary Endpoint(s) 24
6.2.1. Key Secondary Endpoint: PGA Response at Week 12 24
6.2.1.1. Main Analysis 24
6.2.1.2. Sensitivity/Supplementary Analysis 24
6.2.2. Secondary Endpoints: Continuous Data 25
6.2.3. Secondary Endpoint: Binary Data 25
6.2.3.1. Main Analysis 25
6.2.3.2. Sensitivity/Supplementary Analyses 26
CCI
6.4. Clinical Laboratory Values, ECG and Severity of Skin Tolerability 26
CCI
CCI
6.7. Subset Analyses 27
6.8. Baseline and Other Summaries and Analyses 27
6.8.1. Baseline Summaries 27
6.8.2. Study Conduct and Participant Disposition 27
6.8.3. Study Treatment Exposure 27
6.8.4. Concomitant Medications and Nondrug Treatments 27
6.9. Safety Summaries and Analyses 28
6.9.1. Adverse Events 28
6.9.2. Laboratory Data 29
6.9.3. Vital Signs 29
6.9.4. Electrocardiograms 29
6.9.5. Physical Examination 29

7. INTERIM ANALYSES 29
7.1. Introduction 29
7.2. Interim Analyses and Summaries 29
8. REFERENCES 30

LIST OF TABLES

Table 1. Summary of Changes 5
Table 2. Treatment Groups 13
CCI [redacted]

LIST OF FIGURES

Figure 1. Study Schematic 13

APPENDICES

Appendix 1. Summary of Efficacy Analyses 33
Appendix 2. Data Derivation Details 34
Appendix 2.1. Definition and Use of Visit Windows in Reporting 34
Appendix 2.2. Endpoint Derivations 34
Appendix 2.3. Definition of Protocol Deviations That Relate to Statistical Analyses/Populations 37
Appendix 3. Statistical Methodology Details 38
Appendix 4. SAS Code for Estimand 1 –Change from Baseline in PASI Score 39
Appendix 5. SAS Code for the Generalized Linear Mixed Model for Binary Longitudinal Data 41
Appendix 6. SAS Code for Estimand 2 – Risk Difference using (Chan and Zhang 1999)[5] 42
Appendix 7. SAS Code for the Confidence Interval of a Binomial Proportion (Blyth-Still-Casella) 43
Appendix 8. List of Abbreviations 44

1. VERSION HISTORY

None.

Table 1. Summary of Changes

Version/ Date	Associated Protocol Amendment	Rationale	Specific Changes
2.0/ 20 APR 2020	Amendment 1 20 April 2020	Revisions and clarification made in the statistical analysis plan.	• Section 2.2 and 5.1. Details about interim analsyis. • Section 2.1.1, 2.1.2, 6.1 and 6.2. Details regarding to COVID-19 impact are added.
3.0/ 04 MAY 2021	Amendment 1 20 April 2020	Added PGA (0/1 and ≥2 pts decrease) and PASI 75 subgroup analyses based on PGA mild and PGA moderate at the baseline.	• Section 3.3. Included the endpoints to be analysed.

2. INTRODUCTION

PF-06700841 is a dual inhibitor of human tyrosine-protein kinase 2 (TYK2) and Janus kinase 1 (JAK1) that is currently being investigated in participants with multiple indications. The oral formulation of this compound has been studied in several studies, including in patients with psoriasis. This study is the first study for the topical formulation of PF-06700841.

The most common variant of psoriasis, plaque psoriasis, is a chronic inflammatory skin disease characterized by red, scaly, raised plaques. Chronic plaque psoriasis is a common skin disorder with a worldwide prevalence of 2% and afflicts an estimated 5.8-7.5 million Americans (Parisi 2013).[25] *Although psoriasis primarily affects the skin and is not a life-threatening disease, it can profoundly impact the patient's quality of life (QoL) resulting in an impairment akin to other major diseases, such as type 2 diabetes, myocardial infarction, and arthritis* (Rapp 1999).[26]

The Janus kinase (JAK) family, including JAK1, JAK2, JAK3 and TYK2, is a group of cytoplasmic tyrosine kinases that mediate signal transduction via interactions with Type 1 and Type 2 cytokine receptors critical for immune cell function, survival, activation, and proliferation (O'Sullivan 2007) (Murray 2007).[23,22] *JAK2 homodimers are critical for the signaling of hematopoietic cytokines and hormones including erythropoietin (EPO), IL-3, granulocyte-macrophage colony-stimulating factor (GM-CSF) and prolactin.*

JAK1-dependent cytokines include IFN-alpha, IFN-gamma, IL-6, IL-21, and IL-22. IL-6 and IL-21 play a critical role in the development of Th17 cells and production of IL-17, which is a target of several efficacious biologic therapies. JAK1-selective inhibitors are expected to spare dose-limiting effects of pan-JAK inhibition by preserving JAK2 homodimer signaling and thus provide the potential for more efficacious oral treatments for several inflammatory diseases including psoriasis. The previous positive psoriasis studies with two JAK1 selective inhibitors, INCB39110 and GSK2586184, have provided clinical support for JAK1 inhibition as a novel approach to treat plaque psoriasis (Ludbrook 2016) (Bissonnette 2016).[17,3]

Increased activation of Th17 and the main effector cytokines of Th17 cells has been linked to various inflammatory diseases, including psoriasis (Martin 2013).[20] *The number of Th17 cells are elevated in psoriatic lesions, along with levels of proinflammatory cytokines, including IL-17A and IL-17F, which are expressed by Th17 cells and are likely mediators of inflammation and tissue damage* (Martin 2013) (Lowes 2008).[20,16] *Human genetic studies implicate the Th17 pathway in psoriasis, and have uncovered likely risk alleles which include genes involved in IL-23 signaling (including TYK2, which mediates IL-23 receptor signaling) and genes that function downstream of the IL-17 receptor* (Martin 2013).[20]

In addition to genetic evidence, several effective psoriasis therapies target Th17 cytokine production, suggesting a central role of Th17 and IL-17 in the disease. Secukinumab and ixekizumab selectively target IL-17A and have been shown to be effective in the treatment of psoriasis (Gordon 2013).[9] *Other therapies such as cyclosporine, phototherapy, and infliximab also inhibit the IL-17 pathway* (Langley 2014) (Haider 2008) (Zaba 2009).[15,10,31] *The monoclonal antibodies ustekinumab and guselkumab, which are also effective in treating psoriasis* (Krueger 2015) (Reich 2017) (Blauvelt 2017)[14,27,4] *disrupt IL-23 activation of Th17 cells* (Yeilding 2012).[30] *Thus, there is strong rationale for targeting the Th17 pathway in the treatment of psoriasis.*

In the skin, cytokines mediated by JAK signaling impact several cellular inflammatory functions, such as apoptosis of inflammatory T cell infiltrates and T helper cell differentiation. C-X-C motif chemokine ligand 10 (CXCL10), chemokine ligand 26 (CCL26) and matrix metallopeptidase (MMP12) have been reported to be induced by cytokines acting via the JAK class of kinases (Bao 2012) (Fenwick 2015) (Xu 2015) (Amano 2015)[2,6,29,1] *and are implicated in inflammatory and autoimmune conditions of the skin* (Suomela 2001) (Owczarek 2010) (Kamsteeg 2010) (Fujita 2011) (Ferrari 2015) (Martel 2016).[28,24,13,8,7,19] *In addition, impairment of the skin barrier protein filaggrin has also been implicated in inflammatory and autoimmune diseases of the skin* (McLean 2016)[21] *and its expression has been reported to increase upon inhibition of JAK enzymes* (Amano 2015).[1] *Following topical application of the clinical formulation to freshly excised human skin, PF-06700841 caused a dose-dependent inhibition of gene expression of pro-inflammatory molecules CXCL10, CCL26 and MMP12 (measured by changes in messenger ribonucleic acid [mRNA] in the presence and absence of PF-06700841) and a dose-dependent stimulation of the skin barrier protein, filaggrin. Thus, PF-06700841 showed pharmacological modulation in human skin by the topical application, consistent with the known activity of PF-06700841 on JAK1 and TYK2 enzymes.*

PF-06700841 is a dual TYK2/JAK1 inhibitor with good selectivity profile over other human kinases. Based on its cytokine inhibition profile, PF-06700841 is expected to target the Th17 pathway directly by inhibiting TYK2 and indirectly by inhibiting JAK1, thereby providing therapeutic benefit in the treatment of plaque psoriasis.

This statistical analysis plan (SAP) provides the detailed methodology for summary and statistical analyses of the data collected in Study B7931023 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.

Throughout this document, the investigational product (IP) is PF-067008741.

2.1. Study Objectives, Endpoints, and Estimands

Objectives	***Endpoints***	***Estimands***
Primary:	***Primary:***	***Primary:***
• *To compare the efficacy of multiple dose levels of PF-06700841 versus vehicle on change from baseline in PASI score in participants with mild to moderate plaque psoriasis.*	• *Change from baseline in PASI score at Week 12.*	• *Estimand E1: This estimand is intended to provide a population level estimate of the treatment effect of the IP alone on a continuous endpoint; without the benefit of additional prohibited medications, regardless of participants compliance with the IP dosing.* • *Population: Participants with mild to moderate plaque psoriasis as defined by the inclusion and exclusion without the benefit of additional prohibited medications regardless of compliance.* • *Intercurrent Events: A) Prohibited medication – all scores after participants receive prohibited medication will be omitted 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 will be treated similarly as in A). C) Inadequate compliance – participants data will be used as recorded.* • *Population level summary: The mean difference between treated and vehicle control arms of the change from baseline PASI score.*
Secondary:	***Secondary:***	***Secondary:***
• *To compare the efficacy of multiple dose levels of PF-06700841versus vehicle on PGA score in participants with mild to moderate plaque psoriasis.*	*Key secondary endpoint* • *Proportion of participants with PGA score clear (0) or almost clear (1) and ≥2 points improvement from baseline at Week 12.*	• *Estimand E2: This estimand is intended to provide a population level estimate of the treatment effect of the IP alone on a binary responder endpoint; without the benefit of additional prohibited medications, regardless of*

• *To compare the efficacy of multiple dose levels of PF-06700841versus vehicle on the proportion of participants with mild to moderate plaque psoriasis achieving PASI 75.*	*Proportion of participants achieving PASI 75 (75% or greater improvement from Baseline) at time points specified in the Schedule of Activities (SoA).*	*participants compliance with the IP dosing.* • *Population: Participants with mild to moderate plaque psoriasis as defined by the inclusion and exclusion criteria without the benefit of additional prohibited medication regardless of compliance.* • *Intercurrent Events: A) Prohibited medication –response will be considered negative for participants after receiving prohibited medication. B) Withdrawal and all other events leading to missing data will be treated as in A). C) Inadequate compliance – participants data will be used as recorded.* • *Population level summary: The difference in proportions between treated and vehicle response rates.*
• *To compare the efficacy of multiple dose levels of PF-06700841 versus vehicle on measures of disease and symptom severity in participants with mild to moderate plaque psoriasis.*	• *Change from baseline and percent change from baseline in PASI scores at time points specified in the SoA.* • *Absolute score and change from baseline in Peak-Pruritus Numerical Rating Scale score at time points specified in the SoA.* • *Absolute score and change from baseline Psoriasis Symptom Inventory at time points specified in the SoA.* • *Proportion of participants with PGA score clear (0) or almost clear (1) and ≥2 points improvement from baseline at time points specified in the SoA.* • *The proportion of participants who achieved a Psoriasis Symptom Inventory score of 0 (not at all) or 1 (mild) on every item at time points specified in the SoA.*	• *All other continuous secondary endpoints will be analyzed descriptively and using estimand E1 described above when appropriate.* • *All other categorical secondary endpoints will be analyzed descriptively and using estimand E2 described above when appropriate.*
• *To assess safety and tolerability of PF-06700841 in participants with mild to moderate plaque psoriasis.*	• *Incidence and severity of adverse events, serious adverse events and withdrawals due to adverse events.* • *Change from baseline in clinical laboratory values (chemistry and hematology, lipids).* • *Change from baseline in ECG parameters (heart rate, QT, QTc, PR and QRS intervals).* • *Change from baseline in vital*	• *There is no defined estimand for these endpoints and they will be analyzed using Pfizer data standards as applicable.*

	signs (blood pressure, pulse rate and temperature measurements). • *Incidence of severity grades in skin tolerability at times indicated in SoA.*	
CCI		

2.1.1. Primary Estimand(s)

The primary estimand will provide a population level estimate of the treatment effect of PF-07600841 on change from the baseline Psoriasis Area and Severity Index (PASI) under the hypothetical scenario that prohibited medication is not used regardless of dosing compliance.

- Population: Patients with mild to moderate plaque psoriasis as defined by the inclusion and exclusion criteria of the study who do not take prohibited medications and without Week 12 missing visit due to COVID-19 pandemic.

- Variable: Change from baseline PASI at Week 12.

- Intercurrent events: *A) Prohibited medication – all scores after participants receive prohibited medication will be omitted 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 will be treated similarly as in A). C) Inadequate compliance – participants data will be used as recorded.*

- Population-level summary: The mean difference between treated and vehicle control arms of the change from baseline PASI score.

2.1.2. Secondary Estimand(s)

The key secondary estimand will provide a *population level estimate of the treatment effect of the IP alone on a binary responder endpoint; without the benefit of additional prohibited medications, regardless of participants compliance with the IP dosing.*

- Population: Patients with mild to moderate plaque psoriasis as defined by the inclusion and exclusion criteria of the study who do not take prohibited medications and without Week 12 missing visit due to COVID-19 pandemic.

- Variable: Binary response of a PGA score of of clear (0) or almost clear (1) and ≥2 point improvement from baseline at Week 12.

- Intercurrent events *A) Prohibited medication –response will be considered negative for participants after receiving prohibited medication. B) Withdrawal and all other events leading to missing data will be treated as in A). C) Inadequate compliance – participants data will be used as recorded.*

- *Population-level summary: The difference in proportions between treated and vehicle response rates.*

The primary estimand strategy will be applied to the following endpoints:

- Change from baseline and percent change from baseline in PASI scores at time points specified in the SoA.

- Absolute score and change from baseline in Peak-Pruritus Numerical Rating Scale score at time points specified in the SoA.

- Absolute score and change from baseline Psoriasis Symptom Inventory at time points specified in the SoA.

The primary and key secondary estimand strategies will be applied to the following endpoints:

- The proportion of participants who achieved a Psoriasis Symptom Inventory score of 0 (not at all) or 1 (mild) on every item at time points specified in the SoA.

- Proportion of participants with PGA score clear (0) or almost clear (1) and ≥2 points improvement from baseline at time points specified in the SoA.

- The key secondary estimand will also be used for PASI 75 response at all time points specified in the Schedule of Activities.

2.1.3. Additional Estimand(s)

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

2.2. Study Design

This is a Phase 2b, randomized, double-blind, vehicle-controlled, parallel group, multicenter study in participants with mild to moderate plaque psoriasis. The duration of study participation will be approximately 22 weeks, including up to a 6 week screening period, 12 week treatment period, and approximately 4 week follow-up period in both Stage 1 and Stage 2 participants.

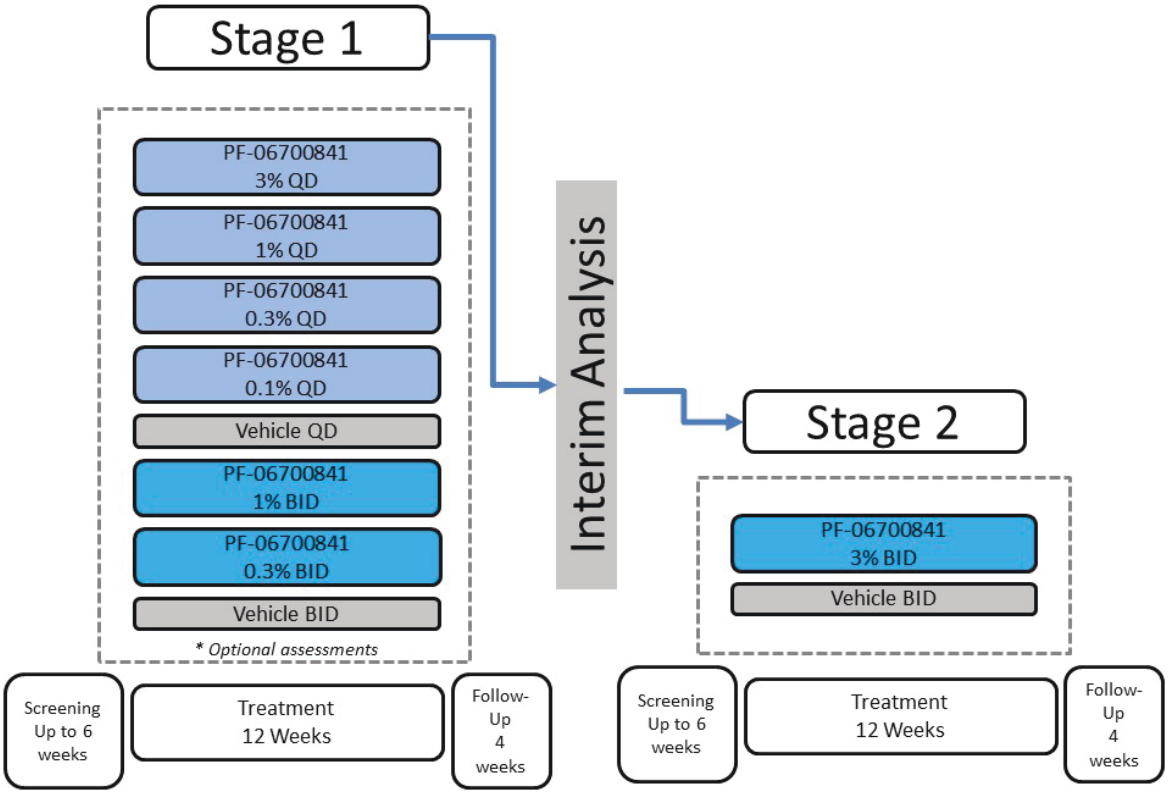
The study will be conducted in 2 stages where an additional cohort (Stage 2) of participants may be added, following completion of enrollment into Stage 1 and at the discretion of the Sponsor. Investigators, study participants, and Pfizer study team members will remain blinded to results of Stage 1.

In Stage 1, approximately 240 participants are planned to be randomized into the study, to allow for approximately 192 (20% drop out rate) evaluable participants (approximately 24 completers per arm, 8 arms). Participants will be randomized to 1 of 8 treatment groups in QD (4 active and 1 vehicle) or BID (2 active and 1 vehicle) regimen in the ratio of 1:1:1:1:1:1:1:1.

Stage 2 will be undertaken at the discretion of the Sponsor, in order to provide more extensive exploration of the dose response curve. In Stage 2, approximately 40 participants are planned to be randomized to allow for approximately 32 evaluable participants. Participants will be randomized into 1 of 2 treatment groups in BID (1 active and 1 vehicle) dosing regimen, in ratio 3:1. The eligibility criteria for participants enrolled in Stage 1 and Stage 2 are the same.

A study schematic is shown in Figure 1, treatment groups are shown in Table 2.

Figure 1. Study Schematic



**Biopsy sub-study, Sleep & Scratch Assessment, Plaque Lesion Photography*

Table 2. Treatment Groups

Stage	Treatment Group	Target Number of Participants Randomized	Approx. Number of Completers	Investigational Product
1	A	30	24	Vehicle cream QD
	B	30	24	PF-06700841 0.1% cream QD
	C	30	24	PF-06700841 0.3% cream QD
	D	30	24	PF-06700841 1.0% cream QD
	E	30	24	PF-06700841 3.0% cream QD
	F	30	24	Vehicle cream BID
	G	30	24	PF-06700841 0.3% cream BID
	H	30	24	PF-06700841 1.0% cream BID
2	I	30	24	PF-06700841 3.0% cream BID
	J	10	8	Vehicle cream BID

CCI [redacted]

CCI [redacted]

Throughout the 12-week treatment period, participants will treat all psoriatic areas identified on Day 1, regardless of clearing or improvement. Any new psoriatic areas occurring after Day 1 will also be treated with IP, provided they are treatment eligible.

Investigators, participants, and the sponsor study team (with the exception of the sponsor supply chain lead) will be blinded as to investigational drug (PF-06700841 cream vs placebo cream [vehicle]) assignment during the conduct of the trial.

An interim analysis will be conducted at the end of Stage 1. Some Pfizer team members and other key stakeholders will become unblinded to the results of Stage 1. Unblinded Pfizer team members will be replaced by blinded team members. Investigators and study participants will remain blinded to the results of Stage 1.

If a participant is withdrawn from IP treatment, the participant will proceed with the Early Termination (ET) and Follow up visits per Schedule of Activities. If a participant uses prohibited medication, guidance in Section 10.9 of the protocol should be followed.

3. ENDPOINTS AND BASELINE VARIABLES: DEFINITIONS AND CONVENTIONS

3.1. Primary Endpoint(s)

- *Change from baseline in PASI score at Week 12* (see Protocol Section 6.1.2). The baseline will be defined as the PASI total score on Day 1 predose.

3.2. Secondary Endpoint(s)

- *Proportion of participants with PGA score clear (0) or almost clear (1) and ≥2 points improvement from baseline at Week 12.*
- *Proportion of participants achieving PASI 75 (75% or greater improvement from Baseline) at time points specified in the Schedule of Activities (SoA).*
- *Change from baseline and percent change from baseline in PASI scores at time points specified in the SoA.*
- *Absolute score and change from baseline in Peak Pruritus Numerical Rating Scale score at time points specified in the SoA.*
- *Proportion of participants achieving ≥2pt improvement in Peak Pruritus Numerical Rating Scale score at time points specified in SoA.*
- *Absolute score and change from baseline Psoriasis Symptom Inventory (PSI) at time points specified in the SoA.*

- *Proportion of participants with PGA score clear (0) or almost clear (1) and ≥ 2 points improvement from baseline at time points specified in the SoA.*
- *The proportion of participants who achieved a Psoriasis Symptom Inventory total score of ≤ 8, with 0 (not at all) or 1 (mild) on every item at time points specified in the SoA.*
- *Incidence and severity of adverse events, serious adverse events and withdrawals due to adverse events.*
- *Change from baseline in clinical laboratory values (chemistry and hematology, lipids).*
- *Change from baseline in ECG parameters (heart rate, QT, QTc, PR and QRS intervals).*
- *Change from baseline in vital signs (blood pressure, pulse rate and temperature measurements).*
- *Incidence of severity grades in skin tolerability at times indicated in SoA.*

The baseline is defined as the total measure on Day 1 predose.

CCI

3.4. Baseline Variables

Baseline will be defined as the last available measurement prior to randomization on Day 1, if a Day 1 measurement is not available, the value from the Screening visit can be used. For the PRO endpoints (except PtGA) baseline will be defined as the average of all values/values recorded on Day 1. The baseline for the endpoint PtGA will be measured as the value on Day 1.

3.5. Safety Endpoints

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

- Incidence of treatment emergent adverse events.
- Incidence of SAEs and AEs leading to discontinuation.
- Incidence of clinical abnormalities and change from baseline in selected clinical laboratory values, ECG measurements, and vital signs.
- Incidence of severity grades in 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

An adverse event will be considered a Treatment-Emergent Adverse Event (TEAE) if the event started during the effective duration of treatment. All events that start on or after the first dosing day (after the first application of IP) and time/start time, if collected, but before the last dose plus the lag time will be flagged as TEAEs. The algorithm will not consider any events that started prior to the first dose date.

A 3-tier approach will be used to summarize AEs. Under this approach, AEs are classified into 1 of 3 tiers. Different analyses will be performed for different tiers (see Section 6.9.1).

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 4 in any treatment group.

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

3.5.2. Laboratory Data

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

- Hematology: hemoglobin, hematocrit, red blood cell count, reticulocyte count, platelet count, white blood cell count with differential, total neutrophils, eosinophils, monocytes, basophils, lymphocytes, coagulation panel.

- Serum chemistry: blood urea nitrogen, creatinine, creatine phosphokinase, glucose, sodium, potassium, chloride, calcium, total bicarbonate, aspartate aminotransferase, alanine aminotransferase, gamma-glutamyl transferase, bilirubin, alkaline phosphatase, lactate dehydrogenase, uric acid, albumin, total protein, total cholesterol, low-density lipoprotein (LDL), high-density lipoprotein (HDL), triglycerides.

3.5.3. Vital Signs, including Height and Weight

Vital sign measurements are oral or tympanic temperature, pulse rate, and blood pressures.

Weight is collected at pre- and post-treatment.

3.5.4. Physical Examinations

Complete physical examinations consist of assessments of general appearance; skin; head, eyes, ears, nose and throat; mouth, heart; lungs; breast (optional); abdomen; external genitalia (optional); extremities; neurologic function; and lymph nodes.

4. ANALYSIS SETS (POPULATIONS FOR ANALYSIS)

4.1. Full Analysis Set (FAS)

As specified in the protocol, the analysis of the efficacy, health outcome and biomarker endpoints will be performed for the modified intent-to-treat (mITT) population, defined as all randomized subjects who receive at least 1 dose of investigational product (PF-06700841 or placebo).

4.2. Efficacy Analysis Set (EAS)

Primary and key secondary analysis will be performed for the modified intent-to-treat (mITT) population, defined as all randomized subjects who receive at least 1 dose of investigational product (PF-06700841 or placebo) and Week 12 visits are not missing due to COVID-19 pandemic.

4.3. Safety Analysis Set

The safety analysis set (SAS) will be all subjects who receive at least 1 dose of investigational product. The safety analysis set will include the follow up period. The final safety database will include all reported safety data at the time of database release.

CCI [redacted]

4.5. Treatment Misallocations and Missing Data

If a subject was:

- Randomized but not treated: the subject will appear on the subject evaluation table as randomized but not treated; this is the extent of how much the subject will be reported;

- Treated but not randomized: the subject will be reported under the treatment they actually received for all safety analyses, but will not be included in the efficacy analyses;

- Randomized but took incorrect treatment: If a subject received the incorrect treatment for the whole duration of the study, then the subject will not be reported for any efficacy analysis, but will be summarized under the treatment they actually received for all safety analyses; if a subject received the incorrect treatment at only some dosing occasions then the subject will be reported under their randomized treatment group for both efficacy and safety analyses. If sufficient doses were incorrect and therefore deemed a major protocol deviation, the subjects may be excluded as sensitivity analysis.

For the continuous PRO variables such as pruritus NRS, DLQI, EQ-5D-5L, SF-36 etc. rules suggested by the developers of these instruments will be followed in calculating the missing values. If these rules are not enough for imputing a value, then the missing values will be handled in the same way as non-PRO variables.

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 randomly assigned to IP and who apply at least 1 dose of IP.
Efficacy analysis set (EAS)	All participants randomly assigned to IP and who apply at least 1 dose of IP and whose Week 12 visit was not missing due to COVID-19 pandemic.
CCI [redacted]	[redacted]
Safety	All participants who receive at least 1 dose of study intervention. Participants will be analyzed according to the intervention they actually received. A randomized but not treated participant will be excluded from the safety analyses.

The following patient level data descriptions are also required for defining the pre-specified analyses:

Defined Analysis Data set (at the data level) – endpoint specific	Description
Primary Estimand Continuous Endpoint Set	This set will include all patients in the EAS population. All data for a subject after the initiation of prohibited medications will be set to missing. Note the primary estimand requires multiple imputation which will be performed on this dataset, the multiple imputations themselves will not be saved in the database, however, the SAS specifications and random number seed will be.
Secondary Estimand Categorical Endpoint Set	This set will include all patients in the EAS population. All data for a subject after the initiation of prohibited medications, withdrawl of either study drug or the study itself will be set to a failure. Any additional missing data will be recorded as a failure. Subjects will have either a success or failure in the dataset for all scheduled visits.
Observed Efficacy Set	This set will include all patients in the FAS population and all observed data and includes all data recorded from the CRF pages. No data will be set to missing or modified from the original CRF record.

5. GENERAL METHODOLOGY AND CONVENTIONS

The primary analysis will be performed when all randomized subjects have either completed their 12-week study participation period or withdrawn early or should the study be stopped prematurely due to any reason, and the database is released. *An interim analysis may be performed for making internal business decisions regarding future study planning. In that case, before any interim analysis is instigated, 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 the IRC charter.*

5.1. Hypotheses and Decision Rules

Statistical inference will be made on the primary endpoint: change from baseline PASI score at Week 12 based on an analysis of the Primary Estimand Continuous Endpoint Set. The null hypothesis is that there is no difference between any arm of PF-06700841 (3% QD, 1% QD, 0.3% QD, 0.1% QD, 3% BID, 1% BID and 0.3% BID) and its corresponding vehicle arm (ie, the 3% BID, 1% BID and 0.3% BID arms will be compared with the BID vehicle, and the remaining QD arms will be compared with the QD vechicle). The alternative hypothesis is that one of the PF-06700841 arms being tested is superior to its corresponding vehicle at Week 6. The study will be considered positive, if this null hypothesis is rejected.

The Type I familywise error-rate will be controlled at the 0.05 level for these tests with a Hochberg step-up procedure (Hochberg, A sharper Bonferroni procedure for multiple significance testing 1988)[11] (Hochberg and Tamhane, Multiple Comparison Procedures 1987)[12] (Marcus, Peritz and Gabriel 1976).[18] The p-values from the 6 comparisons will be ordered from smallest to largest, p[1], … p[7]. Testing will begin with the largest p-value, p[7], if significant at the 0.05 level that comparison and all other comparisions will be declared significant. If it is not significant, p[6] will be compared to an alpha of 0.05/2, if p[6] ≤0.05/2 it and all remaining hypotheses will be declared significant. If not significant, the process will continue using 0.05/3, 0.05/4, 0.05/5, 0.05/6 and lastly 0.05/7.

At the interim analysis, all QD treatment arms will be compared against the QD vehicle arm, whereas the 0.3% BID and 0.1% BID treatment arms will be compared against the available BID vehicle arm data. In Stage 2, participants will only be randomized among 3% BID treatment arm and the BID vehicle arm. After Stage 2, pooled BID vehicle arms from two the stages will be used to compare against all three BID treatments (3% BID, 1% BID and 0.3% BID).

5.2. General Methods

5.2.1. Analyses for Binary Endpoints

Landmark (cross-sectional) analyses of key binary endpoints will calculate and test for risk differences using the method of (Chan and Zhang 1999).[5] Covariates will not be included in the primary analyses. Risk differences and 90% confidence intervals will be presented.

For all key binary endpoints such as PASI75, PASI90, PASI100 and PGA scores etc that are measured repeatedly over time, will be analyzed using generalized linear mixed effect model (GLMM) with treatment group (defined as factor variable), visit, treatment group by visit interaction and subjects as the random effect. An unstructured covariance matrix will be used to fit such model. In case, if the model fails to converge, a covariance structure such as compound symmetry or autoregressive model may be used. Bayesian information criterion (BIC) will be used to assess the goodness of fit of the models. The model with the smallest BIC will be selected for inference. P-values and inference for relative risks between treatments will be provided based on the link function of logit.

For all binary endpoints, a summary based on the mITT Observed Efficacy Set of the number of subjects in each category based on observed cases in each treatment arm at each time point will be produced and the response rate will also be plotted against time, by treatment group.

Explortatory categorical analyses that include or assess the effects of covariates may be done on an exploratory basis. Exploratory longitudinal analyses may also be performed including appropriate covariates of interest.

5.2.2. Analyses for Continuous Endpoints

Landmark (cross-sectional) analysis of key continuous endpoints will use analysis of covariance (ANCOVA). The ANCOVA model will include terms for treatment arm and baseline score of the dependent variable. Least-squares means at the mean overall baseline score will be presented along with 90% confidence intervals.

Mixed model repeated measures (MMRM) models will be used. The fixed effects of treatment, visit (Weeks 2, 4, 8 and 12), and treatment-by-visit interaction will be included. Visit will be modeled as a categorical covariate. Unstructured covariance matrix will be assumed for the model errors. Compound symmetry covariance matrix will be used if the model with unstructured covariance doesn't converge.

When modeling the change from baseline values, the variable for visit will start with the first post-baseline visit, and the actual baseline value will be included as a covariate. At each visit, estimates of least square mean (LSM) values and the LSM differences between the PF-06700841 treated groups (QD and BID) and the corresponding QD and BID vehicle group will be derived from the model. The corresponding p-values and 90% confidence intervals will also be derived from the model.

Unless stated otherwise, descriptive summary statistics for all continuous variables will be presented on mITT Observed Efficacy Set by treatment group and will include the following: n, mean, median, standard deviation, minimum and maximum.

5.2.3. Analyses for Categorical Endpoints

NA

5.2.4. Analyses for Time-to-Event Endpoints

NA

5.2.5. Analyses of Safety Events

Number and percentage of subjects with AEs over the duration of treatment will be provided for each treatment group. Tier-2 events will be analyzed using methods proposed by (Chan and Zhang 1999).[5] Exact methods will be used for Tier-1 events while an asymptotic method will be used for Tier-2 events. Risk differences (each PF-dose group compared to corresponding vehicle) and 2-sided 90% CIs will be reported.

5.2.6. Methods to Manage Missing Data

The primary analysis will use the primary estimand continuous endpoint set. For each landmark analysis (eg, Cross sectional analysis by week) missing data, except missing due to COVID-19 pandemic, will be imputed using a control based imputation method. PROC MI will first be called separately for the QD and BID arms at the visit and a control based method (implemented with the MNAR option) will impute missing vechicle observations under the assumption data are missing at random (MAR) and impute missing treatment observations assuming they are similar to corresponding vechicle patients. Imputation will use the full conditional specification (FCS) method, the imputed data for the QD and BID arms will be combined for the analysis.

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

Analysis of binary data will use the Secondary Estimand Categorical Endpoint Set. This dataset by definition has no missing data since all missing values, except missing due to COVID-19 pandemic, will have been set to a failure. Summaries will use the Observed Efficacy Set and will report results on an observed case (OC) basis.

If a subject misses a visit due to COVID-19, he/she will be excluded from the analysis for that visit; If a subject 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.

6. ANALYSES AND SUMMARIES

Analysis will be done pairwise between each of the PF-06700841 QD treated groups (0.1% cream QD, 0.3% cream QD, 1% cream QD, 3% cream QD) and the corresponding vehicle cream QD, and (0.3% cream BID, 1% cream BID) and the corresponding vehicle cream BID respectively. An analysis with point estimates of the difference and the associated 90% confidence intervals will also be done for the QD versus corresponding BID doses without testing any hypotheses.

A summary of analyses for clinical efficacy endpoints is provided in Appendix 1. Visit windows to be used for all efficacy analyses and some relevant safety analyses are detailed in Appendix 2.1.

Data collected at Week 12 will be displayed in listings only, and will not be part of any analyses, unless specifically noted otherwise.

6.1. Primary Endpoint(s)

6.1.1. Change from Baseline PASI at Week 12

6.1.1.1. Main Analysis

- Estimand strategy: Primary Estimand (Section 2.1.1). 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: EAS (Section 4) using data prepared in the description of the Primary Estimand Continuous Endpoint Set.

- Analysis methodology: Change from baseline will be analyzed using an ANCOVA with the observed (or imputed) Week 12 change from baseline PASI score as the dependent variable with treatment arm and baseline PASI score as the independent variables.

- Intercurrent events and missing data: Data after study drug discontinuation and prohibited medication will be excluded and set to missing. Missing data which will be multiply imputed using a control-based strategy as described in Section 5.2.6. Twenty (100) imputed datasets will be used in the analysis and results combined using PROC MIANALYZE.

- The least-squares (LS) means, the 90% confidence interval for the LS means, the difference between the LS means for each pair of treatment groups, and the corresponding 90% confidence interval will be presented for change from baseline in PASI score.

6.1.1.2. Sensitivity/Supplementary Analyses

Sensitivity Analyses

- To assess the impact of baseline distribution, the main analysis will be repeated with the ANCOVA model replaced by an ANOVA model (excluding baseline as a covariate).

- To assess the impact of COVID-19, the main analysis will be repeated with Full Analysis Set (FAS).

Supplementary Analyses

- An analysis assuming all data is missing at random will be perfomed. In this analysis PROC MI will use and FCS regression approach for all study arms. This analysis is does not penalize the treated arms for intercurrent events and is expected to provide an upper bound for a treatment effect estimates.

6.2. Secondary Endpoint(s)

6.2.1. Key Secondary Endpoint: PGA Response at Week 12

6.2.1.1. Main Analysis

- Estimand strategy: Secondary estimand (Section 2.1.1) use a composite endpoint strategy. 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: EAS (Section 4) using data prepared in the description of the Secondary Estimand Categorical Endpoint Set.

- Analysis methodology: Risk differences will be analyzed using the method of (Chan and Zhang 1999)[5] in PROC BINOMIAL.

- Intercurrent events and missing data: These have been accounted for in the preparation of the Secondary Estimand Categorical Endpoint Set (Section 4.5). This prepared data set has no missing values.

- Proprotions, risk differences and 90% confidence intervals will be presented.

6.2.1.2. Sensitivity/Supplementary Analysis

- To assess the impact of COVID-19, the main analysis will be repeated with Full Analysis Set (FAS).

- Estimand strategy: Primarily this estimand is intended to provide an estimate of the treatment effect in the population level assunming missing data are either MCAS (Missing Completely at Random) or MAR (Missing at Random). Applicable statistics will be calculated at Week 12.

- Population: FAS.

- Analysis methodology: GLMM and descriptive statistics.

- Missing Data: Observed Data.

6.2.2. Secondary Endpoints: Continuous Data

Analyses of all continuous endpoints such as PASI, CCI , PP-NRS, PSI, CCI [redacted] etc for their percentage change or absolute change from baseline or absolute values:

- Estimand strategy: Primarily this estimand is intended to provide an estimate of the treatment effect in the population level assunming missing data are either MCAS (Missing Completely at Random) or MAR (Missing at Random). Applicable statistics will be calculated at time points specified in the SoA.

- Population: EAS.

- Analysis methodology: MMRM and descriptive statistics.

- Missing Data: Observed Data.

6.2.3. Secondary Endpoint: Binary Data

Analyses of all binary endpoints such as CCI [redacted], PASI 75, CCI [redacted] PGA (0/1or ≥2), CCI [redacted] PSI (≤8 and 0/1), PP-NRS (≥2pt improvement), CCI [redacted] for their proportions of participants will be as follows:

6.2.3.1. Main Analysis

- Estimand strategy: Secondary estimand (Section 2.1.1) use a composite endpoint strategy. 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: EAS (Section 4) using data prepared in the description of the Secondary Estimand Categorical Endpoint Set.

- Analysis methodology: Risk differences will be analyzed using the method of (Chan and Zhang 1999)[5] in PROC BINOMIAL.

- Intercurrent events and missing data: These have been accounted for in the preparation of the Secondary Estimand Categorical Endpoint Set (Section 4.5). This prepared data set has no missing values.

- Proprotions, risk differences and 90% confidence intervals will be presented.

6.2.3.2. Sensitivity/Supplementary Analyses

- Estimand strategy: Primarily this estimand is intended to provide an estimate of the treatment effect in the population level assunming missing data are either MCAS (Missing Completely at Random) or MAR (Missing at Random). Applicable statistics will be calculated at time points specified in the SoA.

- Population: mITT.

- Analysis methodology: GLMM and descriptive statistics.

- Missing Data: Observed Data.

Using appropriate model such 4 parameter logistic or E_{max}, dose-response profile for PF-06700841 in QD and BID form will be explored for Primary and Key secondary endpoint.

CCI

[redacted]

6.4. Clinical Laboratory Values, ECG and Severity of Skin Tolerability

The following endpoints will be summarized using available data and not modeled.

Change from baseline in clinical laboratory values, Change from baseline in ECG parameters (hear rate, QT, QTc, PR and QRS intervals). Incidence of severity grades in skin tolerability at times indiciated in the SoA.

- Estimand strategy: No estimand is applicable. Summary statistics will be calculated at time points specified in the SoA.

- Population: EAS.
- Analysis methodology: Summary statistics.
- Missing Data: Observed Data.

CCI [redacted]

CCI [redacted]

6.7. Subset Analyses

There are no subset analyses are planned.

6.8. Baseline and Other Summaries and Analyses

6.8.1. Baseline Summaries

Demographic and baseline characteristics will be summarized by randomized treatment group for all randomized and treated subjects. Continuous variables will be summarized using mean and standard deviation. Categorical variables will be summarized using relative frequency. Key demographic and baseline variables to be summarized include: geographic region, age, gender, race, ethnicity, height, weight, body mass index, disease duration, baseline PASI score, baseline PGA, baseline BSA score etc.

6.8.2. Study Conduct and Participant Disposition

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

6.8.3. Study Treatment Exposure

A summary of dosing compliance by treatment group will be provided.

The exposure to study drug will be summarized by the total number of days of dosing, mean/median number of days of exposure and number and percent of subjects in exposure duration categories.

6.8.4. Concomitant Medications and Nondrug Treatments

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

6.9. Safety Summaries and Analyses

The analysis population for safety is described in Section 4.5. Safety and tolerability will be assessed by clinical review of all relevant parameters including adverse experiences (AEs) and laboratory tests. A complete list of laboratory parameters can be obtained in Section 10.2 of the protocol.

All the tables, listings and graphs for adverse events, lab parameters and vital sign and ECG will follow Pfizer standards. The binary safety endpoints including the incidences of on-treatment AEs, withdrawals due to AEs and serious AEs will be analyzed using the exact test described in Section 6.9.1. A 3-tier approach will be used to summarize AEs. Under this approach, AEs are classified into 1 of 3 tiers.

6.9.1. Adverse Events

All the tables, listings and graphs for adverse events, lab parameters and vital sign and ECG will follow Pfizer standards. The binary safety endpoints including the incidences of on-treatment AEs, withdrawals due to AEs and serious AEs will be analyzed using the exact test described in Section 6.2.3. A 3-tier approach will be used to summarize AEs. Under this approach, AEs are classified into 1 of 3 tiers.

Tier-1 events: These are pre-specified events of clinical importance and are maintained in a list in the product's Safety Review Plan. There are no Tier 1 events for this study. Tier 1 displays will not be created.

Tier-2 events: These are events that are not tier-1 but are "common". A MedDRA Preferred Term (PT) is defined as a tier-2 event if there are at least 4 subjects with an event in in any treatment group.

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

There will be no adjustment for multiple comparisons or stratification factors in the analyses unless specified. For tier-1 and tier-2 events, the proportion of AEs observed in each treatment groups will be presented along with the point estimates and associated 95% confidence intervals of the risk difference for each active treatment compared with placebo using the exact methods described in Section 6.2.3. For tier-1 events p-values may be included in the presentations. AEs will be arranged in the output sorted in descending point estimate of the risk difference within system organ class. Footnotes in the outputs will include the methods used to derive any p-values and confidence intervals as per Pfizer standards.

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 exploratory analysis and its purpose is to generate hypotheses for further investigation. The 3-tier approach facilitates this exploratory analysis.

6.9.2. Laboratory Data

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

6.9.3. Vital Signs

Vital signs will be summarized at baseline, Weeks 6, and Week 12/End of Treatment visits.

6.9.4. Electrocardiograms

ECG parameters, if applicable, will be summarized at baseline, Weeks 6, and Week 12/End of Treatment visits.

6.9.5. Physical Examination

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

7. INTERIM ANALYSES

7.1. Introduction

There will be an interim analyses for this study.

Interim analyses will be performed to assess efficacy and safety after at least 60% of the planned participants, ie, approximately 80 participants, complete their study participation through Week 12. Interim analysis results will be used for internal business decisions regarding future study planning or stopping for futility. Before the interim analysis is instigated, 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 the internal review committee (IRC) charter.

This study will not use a data monitoring committee (DMC). This study will use an internal review committee (IRC) which will be comprised of internal Pfizer experts, independent of the study team.

Members of the study team will not be part of the IRC. An unblinded sponsor statistician or a clinician who is not directly involved with the study will perform the interim analyses and provide the results only to the IRC. No randomization information for individual participants will be reported to the IRC as part of the interim analysis; however, unblinded aggregate results will be reported.

The IRC will be responsible for evaluating safety of participants in the study according to the IRC Charter. The recommendations made by the IRC to alter the conduct of the study will be forwarded to Pfizer for final decision. Pfizer will forward such decisions, which may include summaries of aggregate analyses of endpoint events and of safety data that are not endpoints, to regulatory authorities, as appropriate.

7.2. Interim Analyses and Summaries

A separate interim analysis plan might be developed, if needed.

8. REFERENCES

1. Amano, W. et al. 2015. "The Janus kinase inhibitor JTE-052 improves skin barrier function through suppressing signal transducer and activator of transcription 3 signaling." *The Journal of allergy and clinical immunology* 136: 667-677. doi:10.1016/j.jaci.2015.03.051.

2. Bao, L., Shi, V. Y. & Chan, L. S. 2012. "IL-4 regulates chemokine CCL26 in keratinocytes through the Jak1, 2/Stat6 signal transduction pathway: Implication for atopic dermatitis." *Molecular immunology* 50: 91-97. doi:10.1016/j.molimm.2011.12.008.

3. Bissonnette, R. et al. 2016. "A randomized, double-blind, placebo-controlled, dose-escalation study of the safety and efficacy of INCB039110, an oral janus kinase 1 inhibitor, in patients with stable, chronic plaque psoriasis." *The Journal of dermatological treatment* 27: 332-338. doi:10.3109/09546634.2015.1115819.

4. Blauvelt, A. et al. 2017. "Efficacy and safety of guselkumab, an anti-interleukin-23 monoclonal antibody, compared with adalimumab for the continuous treatment of patients with moderate to severe psoriasis." *Journal of the American Academy of Dermatology* 76: 405-417. doi:10.1016/j.jaad.2016.11.041.

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

6. Fenwick, P. S., Macedo, P., Kilty, I. C., Barnes, P. J. & Donnelly, L. E. 2015. "Effect of JAK Inhibitors on Release of CXCL9, CXCL10 and CXCL11 from Human Airway Epithelial Cells." *PloS one* 10: e0128757. doi:10.1371/journal.pone.0128757.

7. Ferrari, S. M. et al. 2015. "CXCL10 in psoriasis." *Advances in medical sciences* 60: 349-354. doi:10.1016/j.advms.2015.07.011.

8. Fujita, H. et al. 2011. "Lesional dendritic cells in patients with chronic atopic dermatitis and psoriasis exhibit parallel ability to activate T-cell subsets." *The Journal of allergy and clinical immunology* 128: 574-582.e571-512. doi:10.1016/j.jaci.2011.05.016.

9. Gordon, K. B. et al. 2013. "Phase 3 Trials of Ixekizumab in Moderate-to-Severe Plaque Psoriasis." *The New England journal of medicine* 375: 345-356. doi:10.1056/NEJMoa1512711.

10. Haider, A. S. et al. 2008. "Identification of cellular pathways of "type 1," Th17 T cells, and TNF- and inducible nitric oxide synthase-producing dendritic cells in autoimmune inflammation through pharmacogenomic study of cyclosporine A in psoriasis." *Journal of immunology (Baltimore, Md. : 1950)* 180: 1913-1920.

11. Hochberg, Y. 1988. ". A sharper Bonferroni procedure for multiple significance testing." *Biometrika* 75: 800-802.

12. Hochberg, Y., and AC. Tamhane. 1987. *Multiple Comparison Procedures.* New Yrok: Wiley.

13. Kamsteeg, M. et al. 2010. "Molecular diagnostics of psoriasis, atopic dermatitis, allergic contact dermatitis and irritant contact dermatitis." *The British journal of dermatology* 162: 568-578. doi:10.1111/j.1365-2133.2009.09547.x.

14. Krueger, J. G. et al. 2015. "Anti-IL-23A mAb BI 655066 for treatment of moderate-to-severe psoriasis: Safety, efficacy, pharmacokinetics, and biomarker results of a single-rising-dose, randomized, double-blind, placebo-controlled trial." *The Journal of allergy and clinical immunology* 136: 116-124.e117. doi:10.1016/j.jaci.2015.01.018.

15. Langley, R. G. et al. 2014. "Secukinumab in plaque psoriasis--results of two phase 3 trials." *The New England journal of medicine* 371: 326-338. doi:10.1056/NEJMoa1314258.

16. Lowes, M. A. et al. 2008. "Psoriasis vulgaris lesions contain discrete populations of Th1 and Th17 T cells." *J Invest Dermatol* 128: 1207-1211. doi:10.1038/sj.jid.5701213.

17. Ludbrook, V. J. et al. 2016. "Investigation of selective JAK1 inhibitor GSK2586184 for the treatment of psoriasis in a randomized placebo-controlled phase IIa study." *The British Journal of Dermatology* 174: 985-995. doi:10.1111/bjd.14399.

18. Marcus, R., E. Peritz, and KR. Gabriel. 1976. "On closed testing procedure with special reference to ordered analysis of variance." *Biometrika* 63: 655–660.

19. Martel, B. C. et al. 2016. "Distinct molecular signatures of mild extrinsic and intrinsic atopic dermatitis." *Experimental dermatology* 25: 453-459. doi:10.1111/exd.12967.

20. Martin, D. A. et al. 2013. "The emerging role of IL-17 in the pathogenesis of psoriasis: preclinical and clinical findings." *J Invest Dermatol* 133: 17-26. doi:10.1038/jid.2012.194.

21. McLean, W. H. 2016. "Filaggrin failure - from ichthyosis vulgaris to atopic eczema and beyond." *The British journal of dermatology* 175 Suppl 2: 4-7. doi:10.1111/bjd.14997.

22. Murray, P. J. 2007. "The JAK-STAT signaling pathway: input and output integration." *Journal of immunology (Baltimore, Md. : 1950)* 178: 2623-2629.

23. O'Sullivan, L. A., Liongue, C., Lewis, R. S., Stephenson, S. E. & Ward, A. C. 2007. "Cytokine receptor signaling through the Jak-Stat-Socs pathway in disease." *Molecular immunology* 44: 2497-2506. doi:10.1016/j.molimm.2006.11.025.

24. Owczarek, W. et al. 2010. "Analysis of eotaxin 1/CCL11, eotaxin 2/CCL24 and eotaxin 3/CCL26 expression in lesional and non-lesional skin of patients with atopic dermatitis." *Cytokine* 50: 181-185. doi:10.1016/j.cyto.2010.02.016.

25. Parisi, R. et al. 2013. "Global epidemiology of psoriasis: a systematic review of incidence and prevalence." *prevalence.* 133: 377-385. doi:10.1038/jid.2012.339.

26. Rapp, S. R., Feldman, S. R., Exum, M. L., Fleischer, A. B., Jr. & Reboussin, D. M. 1999. "Psoriasis causes as much disability as other major medical diseases." *Journal of the American Academy of Dermatology* 41: 401-407.

27. Reich, K. et al. 2017. "Efficacy and safety of guselkumab, an anti-interleukin-23 monoclonal antibody, compared with adalimumab for the treatment of patients with moderate to severe psoriasis with randomized withdrawal and retreatment." *Journal of the American Academy of Dermatology* 76: 418-431. doi:10.1016/j.jaad.2016.11.042.

28. Suomela, S., Kariniemi, A. L., Snellman, E. & Saarialho-Kere, U. 2001. "Metalloelastase (MMP-12) and 92-kDa gelatinase (MMP-9) as well as their inhibitors, TIMP-1 and -3, are expressed in psoriatic lesions." *Experimental dermatology* 10: 175-183.

29. Xu, M. et al. 2015. "JAK inhibition alleviates the cellular senescence-associated secretory phenotype and frailty in old age." *Proceedings of the National Academy of Sciences of the United States of America* 112: E6301-6310. doi:10.1073/pnas.1515386112.

30. Yeilding, N. et al. 2012. "Development of the IL-12/23 antagonist ustekinumab in psoriasis: past, present, and future perspectives--an update." *Annals of the New York Academy of Sciences* 1263: 1-12. doi:10.1111/j.1749-6632.2012.06670.x.

31. Zaba, L. C. et al. 2009. "Effective treatment of psoriasis with etanercept is linked to suppression of IL-17 signaling, not immediate response TNF genes." *The Journal of allergy and clinical immunology* 124: 1022-1010. doi:10.1016/j.jaci.2009.08.046.

Appendix 1. Summary of Efficacy Analyses

Endpoint	Analysis Type	Population	Data Inclusion and Rules for Handling Intercurrent Events and Missing Data	Analysis Model
Change from baseline to Week 16 in weekly mean of the 24-hour average pain score measured by the NRS	Summary	Evaluable	Separately for observed data and imputed data.	N/A
	Main analysis	Evaluable	All data collected will be included regardless of intercurrent events. Missing data will be imputed.	ANCOVA with terms XXX, XXX
	Sensitivity/supplementary analysis	Evaluable	All data collected will be included regardless of intercurrent events. Missing data will be imputed.	ANOVA
	Sensitivity/supplementary analysis	Evaluable	Only data collected are included. Missing data will not be imputed.	MMRM
XXX				

Appendix 2. Data Derivation Details

Appendix 2.1. Definition and Use of Visit Windows in Reporting

Visit No.	Visit Label	Target Day	Visit Window
1	Screening	N/A	-42 ≤ day ≤ -1
2	Baseline*	1	day = 1
3	Week 1	8	2≤day≤11
4	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≤92
10	Week 14	100	93≤day≤109
11	Week 16 FUP/ET	118	110≤day≤127
* Baseline analysis visit window may be considered as day ≤1 in some analyses (eg, those involving change from baseline). That is, in case that Day 1 observation is missing, the last observation by 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.			

Appendix 2.2. Endpoint Derivations

Psoriasis Area Severity Index (PASI)

The Psoriasis Area and Severity Index quantifies the severity of a subject's psoriasis based on both lesion severity and the percentage of body surface area affected. Lesion severity: the basic characteristics of psoriatic lesions - erythema, induration and scaling - provide a means for assessing the severity of lesions. Assessment of these three main signs is performed separately for four areas of the body: head, upper limbs, trunk, and lower limbs. Average erythema, induration and scaling are rated for each body area according to a 5-point scale: 0, no involvement; 1, slight; 2, moderate; 3, marked; 4, very marked.

Body surface area (BSA) involvement: the extent (%) to which each of the four areas of the body is affected by psoriasis is assigned a numerical score according to the following area scoring criteria: 0, no involvement; 1, >0 to 9%; 2, 10 to 29%; 3, 30 to 49%; 4, 50 to 69%; 5, 70 to 89%; 6, 90 to 100%. Details see Table 5 of the protocol.

Derivation of PASI score

In each area, the sum of the severity rating scores for erythema, induration and scaling is multiplied by the score representing the percentage of this area involved by psoriasis, multiplied by a weighting factor (head 0.1; upper limbs 0.2; trunk 0.3; lower limbs 0.4). The sum of the numbers obtained for each of the four body areas is the PASI.

PASI = 0Ah(Eh + Ih + Sh) + 0.2Au(Eu + Iu + Su) + 0.3At(Et + It + St) + 0.4Al(El + Il + Sl)

where A = area of involvement score; E = erythema; I =induration; S = scaling; h = head; u = upper limbs; t = trunk; l = lower limbs

The PASI score can vary in increments of 0.1 units from 0.0 to 72.0, with higher scores representing increasing severity of psoriasis.

PASI [redacted]/75/[redacted] response

At least [redacted]/75[redacted]% reduction in PASI relative to baseline PASI Score.

Body Surface Area (BSA)

Assessment of body surface area involved in psoriasis is performed separately for four areas of the body: head, upper limbs, trunk, and lower limbs. The percentage surface area affected by psoriasis is estimated by means of the "handprint method", where the full hand of the subject (ie, the subject's flat hand, thumb and fingers) represents approximately 1% of the total BSA2 and a set percentage of each of the four areas of the body:

- 1 handprint corresponds to approximately 10% of the head/neck;
- 1 handprint corresponds to approximately 5% of the arm/upper limbs;
- 1 handprint corresponds to approximately 3.3% of the trunk;
- 1 handprint corresponds to approximately 2.5% of the lower limbs.

[redacted]

[redacted]

Physician Global Assessment (PGA)

The Physician Global Assessment of psoriasis is scored on a 5-point scale, reflecting a global consideration of the erythema, induration and scaling across all psoriatic lesions. Average erythema, induration and scaling are rated separately over the whole body according to a 5-point severity scale, scored from 0 to 4, with appropriate morphologic descriptors. The severity rating scores are summed and the average taken - the total average is rounded to the nearest whole number score to determine the PGA. The 5-point scale for PGA is: 0, "clear"; 1, "almost clear"; 2, "mild"; 3, "moderate"; 4 "severe".

CCI [redacted]

CCI [redacted]

Peak Pruritus Numerical Rating Scale (PP-NRS)

The intensity of pruritus will be assessed by a numerical rating scale (NRS), an 11-category numeric rating scale from 0 to 10. Pruritus will be evaluated by asking participants to assign a numerical score representing the worst imaginable itch over the last 24 hours on a scale from 0 to 10, with 0 indicating no symptoms and 10 indicating the worst imaginable symptoms. This item will be administered to all participants.

Select the number that best describes your worst itching due to psoriasis over the past 24 hours. (Check one number only)

☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

0 1 2 3 4 5 6 7 8 9 10

No itching — Worst possible itching

CCI [redacted]

Night Time Itch Severity:

On a scale of 0 to 10, with 0 being "no itch" and 10 being "worst itch imaginable", how would you rate your itch at the worst moment during your most recent night's sleep?

☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

0 **1** **2** **3** **4** **5** **6** **7** **8** **9** **10**

No itch — Worst itch imaginable

Night-time itch frequency

During your most recent night's sleep, how often did your skin itch?

- ☐ Never
- ☐ Rarely
- ☐ Sometimes
- ☐ Often
- ☐ Almost Always

Appendix 2.3. Definition of Protocol Deviations That Relate to Statistical Analyses/Populations

Appendix 3. Statistical Methodology Details

Hochberg Testing Procedure

Consider testing null hypotheses H01, …H07 and let pi, i=1,…,7 denote the corresponding 1-sided p-values from the individual pairwise comparisons against the placebo arm prior to multiplicity adjustment. Furthermore, let [1], ….[7] denote the order of the p-values so that p[1] ≤p[2] ≤…. P[7]. The procedure starts with the largest p-value p[7] as follows:

1. If p[7] < α reject all null hypotheses, otherwise go to next step
2. If p[6] <α/2 reject hypotheses H0[1] through H0[6], otherwise go to next step
3. If p[k] < α/(7-k +1), reject hypotheses H0[1] through H0[k]
4. If p[1] < α/7, reject H0[1], otherwise stop and do not reject any hypotheses

   Alternatively, the unadjusted raw p-values can be read into Proc Multtest and adjusted using the HOC option.

At the interim analysis time, since there will not be any data related to 3% BID treatment group, the corresponding p-value will be set to 1 and the above procedure will be performed.

```
data pvals;
  Input Test$ Raw_P;
Datalines;
Test1 .xxxxx
Test2 .xxxxx
…..
Test6 .xxxxx
;
Proc multtest pdata=pvals hoc out=new;
run;
```

Appendix 4. SAS Code for Estimand 1 –Change from Baseline in PASI Score

```
libname b7931023  "/Volumes/app/…… /data_vai" ;


data ps;
        set b7931023.adps;
run;

data look;
        set ps;
        where param = "PASI02-Total Sum"  and viswin ne .;
        keep subjid param paramn avisit age sex trta viswin visit xviswin aval base chg pchg trtan;
run;

**example data for a single visit;
**created from systemic study ;
**ignore values, only used for illustration;

**data already has one record per subject per visit even if
**data is missing. Proc Mi needs missing values in order to impute;
data ex1;
        set look;
        where viswin = 9;
        regimen = "BID";
        if trtan  < 6 then regimen = "QD";
        if trtan = 5 then dose=0;
        if trtan = 4 then dose=0.1;
        if trtan = 3 then dose=0.3;
        if trtan = 2 then dose=1.0;
        if trtan = 1 then dose=3.0;
        if trtan = 8 then dose=0;
        if trtan = 7 then dose=0.3;
        if trtan = 6 then dose=1.0;
        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 usin 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=1023 nimpute=20 out=outimp  max=72 min=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=.1 ;
          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=.1;  **specify alpha for 90% CIs here;
     by regimen _dose;
     modeleffects estimate;
     stderr stderr;
     ods output parameterestimates=parameterestimates;
run;
```

Appendix 5. SAS Code for the Generalized Linear Mixed Model for Binary Longitudinal Data

```
PROC GLIMMIX DATA =<DATA>   METHOD=RMPL;
   CLASS SUBJID TRTPN VISIT;
   MODEL RESPONSE (EVENT = "1") = TRTPN VISIT TRTPN * VISIT / ALPHA = 0.1
DIST=BINARY LINK=LOGIT;
   RANDOM VISIT /SUBJECT = SUBJID TYPE=UN RESIDUAL;
   LSMEANS TRTPN * VISIT / ILINK COV DIFF CL;
RUN;
```

Appendix 6. SAS Code for Estimand 2 – Risk Difference using (Chan and Zhang 1999)[5]

PROC BINOMIAL DATA=<DATASET> GAMMA=0 ALPHA=<Value>;

PD/EX ONE STD;

PO <POPULATION VARIABLE>;

OU <OUTCOME VARIABLE>;

RUN;

Appendix 7. 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 8. List of Abbreviations

Abbreviation	Term
Abs	absolute
AE	adverse event
ANCOVA	analysis of covariance
ANOVA	analysis of variance
ATC	Anatomic Therapeutic Chemical
AUC	area under the curve
BA	bioavailability
BE	bioequivalence
BLQ	below the limit of quantitation
BOCF	baseline observation carried forward
BP	blood pressure
CDARS	Clinical Data Analysis and Reporting System (of US Food and Drug Administration)
CI	confidence interval
C_{max}	maximum observed concentration
CMH	Cochran-Mantel-Haenszel
CRF	case report form
CSR	clinical study report
DMC	data monitoring committee
EAC	event adjudication committee
EAS	Efficacy analysis set
ECG	electrocardiogram
E-DMC	external data monitoring committee
FAS	full analysis set
FDA	Food and Drug Administration (United States)
GCP	Good Clinical Practice
GLIMMIX	generalized linear mixed-effects model with repeated measures
GMC	geometric mean concentration
GMFR	geometric mean fold rise
GMR	geometric mean ratio
GMT	geometric mean titer
ICD	informed consent document
ICH	International Council for Harmonisation
IRC	internal review committee
IST	independent statistical team
ITT	intent-to-treat
LLOQ	lower limit of quantitation
LOCF	last observation carried forward
LOD	limit of detection
LS	least-squares
LSM	least-squares mean
MAR	missing at random

Abbreviation	Term
MedDRA	Medical Dictionary for Regulatory Activities
mITT	modified intent-to-treat
MMRM	mixed-effects model with repeated measures
MNAR	missing not at random
N/A	not applicable
NNB	number needed to benefit
NNH	number needed to harm
NNT	number needed to treat
NOAEL	no-observed-adverse-effect level
CCI	[redacted]
PP	per-protocol
PPAS	per-protocol analysis set
PRO	patient-reported outcome
PT	preferred term
QTc	corrected QT
QTcF	corrected QT (Fridericia method)
qual	qualitative
RCDC	reverse cumulative distribution curve
RR	relative risk
SAE	serious adverse event
SAP	statistical analysis plan
SD	standard deviation
SGS	Statistical Guidance Standards
SOP	standard operating procedure
SUSAR	suspected unexpected serious adverse reaction
TA	therapeutic area
ULN	upper limit of normal
WHO	World Health Organization
WHODD	World Health Organization Drug Dictionary