Protocol B7581002

A PHASE 2A, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED STUDY TO EVALUATE THE EFFICACY, SAFETY, TOLERABILITY AND PHARMACOKINETICS OF PF-06687234 AS ADD-ON THERAPY TO INFLIXIMAB IN ACTIVE ULCERATIVE COLITIS SUBJECTS WHO ARE NOT IN REMISSION (BUILD UC)

Statistical Analysis Plan (SAP)

Author: PPD

Version: 1.0

Date: May 02, 2018

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

APPENDICES

1. VERSION HISTORY

This Statistical Analysis Plan (SAP) for study B7581002 is based on the protocol dated 06MAR2018.

Table 1. Summary of Major Changes in SAP Amendments

SAP Version	Change	Rationale
1	Not Applicable	Not Applicable

2. INTRODUCTION

Inflammatory bowel disease (IBD) is a chronic inflammatory condition of the gastrointestinal tract that affects five million people worldwide with considerable variability in incidence and prevalence by region. The prevalence of IBD is highest in the second and third decade of life with another peak between 60 to 70 years of age.[1] Ulcerative colitis (UC), one of two major phenotypes; the other is Crohn's disease (CD), is characterized by continuous superficial mucosal inflammation that is localized to the colon and rectum. UC involves a relapsing and remitting clinical course characterized by bloody diarrhea (with or without mucus), urgency, tenesmus, abdominal pain and weight loss, but can also present with variable extra intestinal manifestations in the eye, skin, and joint compartments.[2] There is also an increased risk of colorectal cancer in UC patients compared to the general population. Given the chronic nature of the condition, patients experience poor health-related quality of life, and the direct and indirect costs of clinical management represent a significant economic burden to society.

The goal of current treatments is to achieve resolution of signs and symptoms of active disease (clinical remission) in the short term and to decrease the frequency of subsequent disease flares in the long term. In the setting of moderate to severe UC, the treatment algorithm includes 5-aminosalicylates, corticosteroids, immunomodulators (primarily azathioprine and 6-mercaptopurine), and monoclonal antibodies against tumor necrosis factor (TNF)-α or α4β7 integrin. Anti-TNF-α therapies (infliximab, adalimumab, and golimumab) are the most commonly used biologic agents to treat UC. However, some patients do not achieve clinical remission with anti-TNF-α therapy, and up to 50% will develop loss of response over time.[3] In these patients, a dose escalation strategy with therapeutic drug monitoring (TDM) may be used to ensure adequate serum drug levels needed to achieve efficacy. Another strategy is to explore the potential for additive benefit with the addition of another immunosuppressant as seen with the combination of infliximab with azathioprine in UC.[15] However, such combinations targeting immune suppression are not without increased risk of serious infections and malignancies. Consequently, safer and more efficacious approaches that combine other mechanisms of immune modulation are needed to improve therapeutic outcomes of existing therapies, offer additional options, and delay the need for surgical intervention, particularly in those with moderate to severe disease inadequately managed or refractory to standard treatments.

This SAP provides the detailed methodology for summary and statistical analyses of the data collected in study B7581002. 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

Primary Efficacy Objective

- *To evaluate the efficacy of PF-06687234 in induction of clinical remission in subjects with UC and a partial response to anti-TNF-α.*

Primary Safety Objective:

- *To evaluate the safety and tolerability of PF-06687234 in subjects with UC and a partial response to anti-TNF-α.*

Secondary Objectives

- *To evaluate the efficacy of PF-06687234 in induction of endoscopic improvement in subjects with UC and partial response to anti-TNF-α.*
- *To evaluate histological improvement in subjects with UC and partial response to anti-TNF-α.*
- *To evaluate the efficacy of PF-06687234 in induction of clinical response in subjects with UC and a partial response to anti-TNF-α.*
- *To describe the PK of PF-06687234 in subjects with UC.*
- *To evaluate the immunogenicity of PF-06687234 in subjects with UC.*

Exploratory Objective(s)

- *To collect banked biospecimens and additional non-banked samples (eg, intestinal biopsies, stool for microbiome analysis, serum and plasma for analysis of proteins, whole blood for RNA analysis and epigenetics and/or cytometry) for exploratory research, unless prohibited by local regulations or ethics committee decision.*
- *To evaluate disease pathway and related biomarkers (ie, hsCRP and fecal calprotectin).*
- *To describe the full PK profile in a subgroup of subjects.*
- *To evaluate tissue concentrations of PF-06687234 in biopsy samples.*

2.2. Study Design

2.2.1. Study Overview

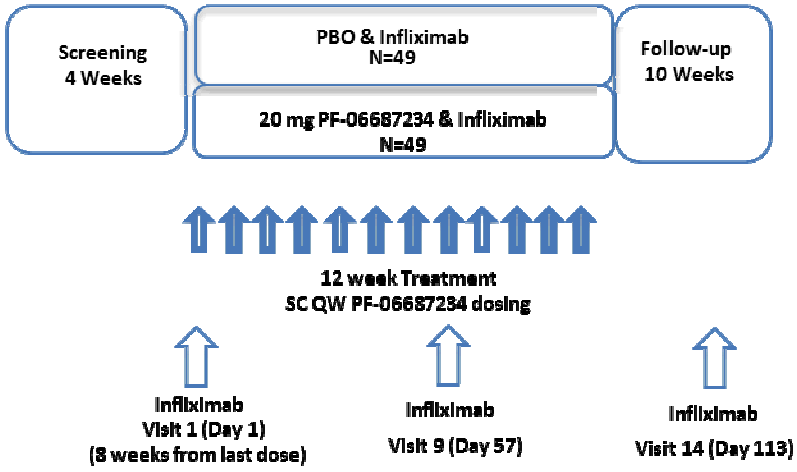
This is a Phase 2A, randomized, double-blind, placebo-controlled, parallel group, multicenter study in subjects with active UC and a non-remission (partial) response to anti-TNF-α therapy. Each subject will be randomly assigned to 1 of 2 treatment arms (1 active; 1 placebo) with approximately 98 subjects in total (49 subjects per arm) enrolled for the study to achieve a total of 78 evaluable subjects.

Using the Mayo scoring definitions in Appendix 2 and Appendix 3 of the protocol*, clinical remission is defined as those subjects with endoscopic subscore of 0 or 1, stool frequency subscore of 0 or 1 and rectal bleeding subscore of 0.*

Eligible subjects for enrollment will meet criteria for non-remission during screening, despite at least 14 weeks of infliximab therapy for active UC.

Women of childbearing potential (WOCBP) will be eligible for this study provided these women use two methods of contraception, as outlined in Section 4.4.1 of the protocol.

Figure 1. Study Design Schematic



2.2.2. Additional Pharmacokinetic Sampling

For most subjects participating in Study B7581002 a sparse PK sampling scheme will be followed. As the PK of PF-06687234 in UC subjects has not been characterized, to better understand the PK of PF-06687234, a more extensive PK sampling scheme may be used in a subgroup of approximately 12 subjects who provide consent from any investigational site. For specific details refer to Section 7.5.2 of the protocol regarding the additional PK sampling.

2.3. Duration of Subject Participation

The duration of participation for eligible subjects will be approximately 26 weeks. This will include a 4-week screening period, a 12-week treatment period, a 10-week follow-up period which will include a telephone contact conducted 6-weeks after the onsite follow-up visit. During the treatment period subjects will visit the clinic every week (±2 days) for SC administration of 20 mg of the investigator product (IP) or placebo for a total of 12 visits. In addition, all subjects enrolled in the study will be provided with their infliximab therapy (5-10 mg/kg) during the 12-week treatment period and at Week 16 visit. Subjects will be administered their infliximab therapy by IV administration at the clinic on Day 1, Week 8 visit (±2 days) and at the final onsite visit at Week 16.

Subjects will be required to return to the clinic for one visit 4 weeks after the last visit during the treatment period to evaluate for safety which will constitute the in person follow-up period. Subjects will also be administered their infliximab therapy at this visit. There will be a telephone contact conducted 6 weeks after this visit to confirm any SAEs have been reported to Pfizer post the final infliximab infusion at Week 16 and to provide results of Week 16 immunogenicity testing.

Subjects with no detectable HAFA(Human Anti-Fusion Antibodies) *or Nab* (Neutralizing Antibodies) *against IL-10 portion of PF-06687234 by the Week 16 visit will require no further follow-up beyond the Week 22 telephone contact. However, subjects with detectable HAFA or NAb against the IL-10 portion of PF-06687234 at Week 16 will be informed of their immunogenicity status at the Week 22 telephone contact. These subjects will be required to return for an onsite visit at Week 28 for repeat immunogenicity testing. If these subjects have detectable HAFA or NAb against the IL-10 portion of PF-06687234 from the immunogenicity sample collection at Week 28, they will require one additional follow-up collection 3 months after Week 28 at approximately Week 40.*

2.4. Approximate Duration of Study

Study enrollment is estimated to be completed in approximately 19 months. The completion of the study is estimated to occur in approximately 26 months.

3. ENDPOINTS AND BASELINE VARIABLES: DEFINITIONS AND CONVENTIONS

3.1. Primary Efficacy Endpoint

Proportion of subjects in clinical remission at Week 12 (as defined by a modified Mayo Score with an endoscopic subscore ≤1, stool frequency subscore ≤1 and rectal bleeding subscore = 0).[a,b]

a. *The assessment of clinical remission will utilize the current definition of the Mayo endoscopic subscore, which allows for assessment of mild friability in the subscore of 1.*

b. *The Central Readers for endoscopy Mayo subscores will also employ an assessment as to the presence or absence of any friability (including mild). In this assessment the presence of <u>any</u> friability (including mild) will be scored as 2.*

3.2. Primary Safety Endpoint

- *Incidence and severity of adverse events, serious adverse events and withdrawals due to adverse events, ECGs, vital signs and safety laboratory tests.*

3.3. Secondary Endpoints

- *Proportion of subjects with endoscopic improvement at Week 12 (defined as decrease of ≥ 1 point in Mayo endoscopy score or an absolute endoscopy score of ≤1.)*[b]
- *Mean change from baseline at Week 12 in Geboes histology score.*
- *Proportion of subjects with a clinical response at Week 12 defined with a decrease from baseline of at least 3 points in total Mayo score with at least 30% change, accompanied by at least one point decrease or absolute score of 0 or 1 in rectal bleeding subscore.*
- *Proportion of subjects with change from baseline in partial Mayo Score of ≤2 with no individual subscore >1 at Weeks 2, 4, 6, 8, 12.*
- *Plasma concentrations of PF 06687234.*
- *Incidence of the development of HAFAs and Nabs against PF 06687234.*

3.4. Exploratory Endpoints

- *Collection of banked biospecimens and additional non banked samples (eg, intestinal biopsies, stool for microbiome analysis, serum and plasma for analysis of proteins, whole blood for RNA analysis and epigenetics and/or cytometry) unless prohibited by local regulations or ethics committee decision. Additional information on collection and potential use is provided in the Banked Biospecimens section.*
- *Change from baseline in fecal calprotectin at Weeks 4, 8, 11.*
- *Change from baseline in hsCRP at Weeks 4,8,11.*
- *AUCτCmax, Tmax, t1/2, CL/F, Vz/F as data permit.*
- *PF 06687234 tissue concentration in inflamed and non inflamed biopsies.*

3.5. Baseline Variables

The covariates added into the analyses will be the baseline score (eg, in Longitudinal data analysis).

Covariates will be summarized for the mITT (FAS) and PP population, split by treatment group. Continuous baseline covariates will be summarized by: n, mean, median, standard deviation, min and max. Binary and factor covariates will be summarized by percents and counts.

3.6. Safety Endpoints

3.6.1. Adverse Events

- Safety and tolerability of PF-06687234 dose levels versus placebo: Adverse Events (AE), Serious Adverse Events (SAE), treatment-emergent adverse events (TEAE), withdrawal due to adverse events. Any adverse events that start or increase in severity following the start of treatment and occur within the follow-up period after the end of treatment will be considered as TEAE.

- Potential cases of drug-induced liver injury (ALT and AST elevation).

- Laboratory tests as specified in the protocol.

- Vital signs (blood pressure, heart rate, respirations and oral or tympanic temperature) will be measured at every study visit.

- ECG, single 12-lead ECGs will be obtained on all subjects at specified visits.

- Other safety endpoints including hospitalization, physical examination, weight, and concomitant medications, data will be collected.

A 3-tier approach will be used to summarize clinical AEs. Under this approach, AEs are classified into 1 of 3 tiers. Different analyses will be performed for different tiers (See Section 8.2.3 of the protocol).

Tier-1 events: no Tier-1 AE has been identified/specified.

Tier-2 events: A MedDRA PT is defined as a tier-2 event if it has at least 4 frequencies in any treatment group.

Tier-3 events: All other events are defined as tier-3 events.

3.6.2. Vitals and ECGs

Baseline for vitals and ECG will be defined as the last pre-dose value before the first dosing.

If not supplied, QTcF will be derived using Fridericia's heart rate correction formula: QTcF = QT / (RR)^(1/3), where RR = 60/HR (if RR is not provided).

note: if QTcB is collected, then it should be listed only.

3.6.3. Laboratory Data

Using the current Pfizer data standards, the last pre-dose value before the first dosing is used as the baseline for all laboratory parameters.

4. ANALYSIS SETS

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

4.1. Full Analysis Set

The Full Analysis Set is defined as all subjects randomized and who have received at least one dose of randomized treatment, which will also be called modified ITT (mITT) population. All efficacy, PD, and biomarker analyses will be primarily performed with the mITT population.

4.2. Per Protocol Analysis Set

The Per Protocol (PP) population will be a subset of the mITT population, which excludes those subjects who are identified as key protocol violators from mITT population.

4.3. Safety Analysis Set

The safety population includes all enrolled subjects who received at least 1 dose of investigational product. Subjects who are enrolled and treated but not randomized for various reasons will also be included in the safety population. All patients will be analyzed according to the treatment which they actually received. All safety analyses will be primarily based on the safety population.

4.3.1. Pharmacokinetic (PK) Population

The PK population is defined as all subjects who received at least 1 dose of investigational product and have data on at least one PK concentration. The PK population will be used for concentration-time data listings for the clinical study report.

4.3.2. Immunogenicity assessment population

The immunogenicity assessment population for PF-06687234 is defined as all enrolled subjects who received at least one dose of PF-06687234 with at least one post-treatment HAFA determination.

4.4. Treatment Misallocations

If a subject was:

- Randomized but not treated, the subject will appear on the subject disposition table and listing as randomized but not treated; They will be excluded from all (efficacy, safety, PK, PD, biomarker, health outcome and immunogenicity) analyses as they have not received at least one dose of study medication.

- Treated but not randomized, then they will be excluded from the efficacy analyses (both mITT and PP) since randomized treatment is missing, but will be reported under the treatment they actually received for safety analyses. They may also be included in the exploratory biomarker analyses.

- Randomized but took incorrect treatment: if a subject received the incorrect treatment for the whole duration of the Week 0-12 period then they will be reported under the randomized treatment group for all efficacy analyses (including baseline tables), but will be reported 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 all efficacy and safety analyses, but the subject will be excluded from the per-protocol population and may be excluded form the exploratory biomarker analyses if sufficient doses were incorrect and therefore deemed as a major protocol deviation.

4.5. Protocol Deviations

Different protocol deviation may have a different impact on the analyses. A full list of protocol deviations will be compiled and reviewed by the Clinician and the statistician to identify major and minor deviations prior to database release. Those major protocol deviations will be noted by declaring the subject to be a key protocol violator. Per the SAP, a key protocol violator will be excluded from the per-protocol analysis set.

4.5.1. Deviations Assessed Prior to Randomization

During the screening phase (prior to randomization), the investigator will assess and document subject eligibility against the inclusion and exclusion criteria as set out in sections 4.1 and 4.2 of the protocol. Any subject that does not meet the inclusion and exclusion criteria should not be randomized.

4.5.2. Deviations Assessed Post-randomization

Any significant deviation from the protocol occurring after randomzation will be reviewed prior to database release and a decision will be made regarding whether a subject is identified as a key protocol violator who should be excluded from the per-protocol analysis set.

5. GENERAL METHODOLOGY AND CONVENTIONS

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

An interim analysis for futility may be performed with 40% completers.

All subjects, investigators, and site personnel will continue to be blinded to randomized study treatments throughout the period of the study. The database will be officially released after last subject last visit occurs. The final analysis will be then conducted and the CSR will be issued. The decision rules for the final analyses are also described in the next section.

5.1. Hypotheses and Decision Rules

Active treatmentwill be considered superior to placebo with respect to response in primary efficacy endpoint, proportion of subjects in clinical remission at Week 12, if the difference is statistically significant at the two-sided 0.05 level.

Sample Size

The sample size calculation was based on the primary efficacy endpoint, proportion of subjects in clinical remission at Week 12.

It is anticipated that 22% of placebo subjects and 52% of subjects assigned to receive PF-06687234 will achieve a clinical remission rate at Week 12. A sample size of 39 subjects in each treatment group (78 total) is required to detect a treatment advantage of 30% in clinical remission with at least 80% power and a Type I error at 0.05 in a two-sided test. Assuming a 20% dropout rate, the target sample size for recruitment is 49 per group.

No investigational study has to date been completed in a UC population receiving infliximab therapy with mild to moderate levels of residual active disease. Based on available literature (Jairath et al, 2016),[3] and internal meta-analyses of updated clinical data available in the current literature, the placebo response rate of clinical remission in moderate to severe UC patients is assumed to be 0% to 15%. In mild to moderate disease, it is expected that the placebo response rate would be even higher. After careful consideration of the overall context of use and in consultation with external key inflammatory bowel disease (IBD) disease experts, a placebo rate of 22% and treatment rate of 52% are assumed for this design, which will guide continued clinical development for UC.

Treatment failure approach will be used for missing value imputation from drop out subjects, ie, subjects who have missing value for any reasons will be considered as treatment failures.

If the recruitment rate is very slow, an interim analysis for futility will be done with 40% completers. Since the interim analysis is for futility only, type I error rate will not be affected. A conditional power less than 10% will be used for futility.

5.2. General Methods

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

5.2.1. Analyses for Binary Data

All binary data comparing active treatment group and placebo group will be analyzed using the unconditional exact Chan and Zhang (1999) method.[1] Rate (%) differences and the corresponding 2-sided 95% confidence intervals will be computed using this methods.

Confidence interval of a single proportion will be computed using Blyth-Still-Casella method.

For binary endpoints that are measured repeatedly over time, will be analyzed using generalized linear mixed effect model with treatment group, 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. The SAS code to be used to compute the confidence intervals of relative risks, and difference of relative risks adjusting any covariates is given in the appendix.

5.2.2. Analyses for Longitudinal Continuous Data

Descriptive statistics (n, mean, median, standard deviation, minimum and maximum) for continuous endpoints will be summarized and presented by dose cohort (and by each planned measurement time point if applicable). Geometric Mean may be provided as necessary.

Mixed Effect Model Repeat Measurement (MMRM): The fixed effects of treatment, visit, and treatment-by-visit interaction will be included, along with patient as a random effect. Unstructured covariance matrix will be assumed. The baseline infliximab concentration and ADA status as covariate may be added to the model as a part of sensitivity analsysis.

When modeling the change from baseline values, the variable of 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 mean values and the mean differences between the active treated group and the placebo group will be derived from the model. The corresponding p-values, standard errors and 95% confidence intervals will also be derived from the model.

5.2.3. Analyses for Non-longitudinal Continuous Data

Analysis of Covariance (ANCOVA): The non-longitudinal continuous data will be analyzed by ANCOVA with treatment as the factor. When modelling change from baseline values, the actual baseline value will be included as a covariate. Active dose group will be contrasted versus placebo.

5.2.4. Analyses for Categorical Data

None.

5.2.5. Analyses for Time to Event Data

None.

5.3. Methods to Manage Missing Data

Observed data will be used for descriptive statistics, ie, missing data will not be imputed. Unless there is an explicit instruction, missing values will be used for lower limits of detection and quantitation.

No explicit missing value imputation will be used for longtitudinal models (such as Linear Mixed model or Generalized Linear Mixed Model) as under the missing at random (MAR) or missing completely at random (MCAR) assumptions missing values are already incorporated into those models.

No missing values will be imputed for analyses of exploratory biomarkers.

5.3.1. Binary Endpoint

For the binary response endpoints, subjects with missing values will be handled by treatment failure approach (TFA) ie, setting any missing values to be non-responsive (0).

5.3.2. Continuous Endpoints

For non-patient reported outcome variables, the missing continuous data will be used as is (Observed Case – OC). The Baseline Observation Carried Forward (BOCF) approach may be used to handle the monotone missing data or no post baseline measurements. Using BOCF, the baseline values will be carried forward for the visits after which subjects prematurely discontinued assigned treatment.

6. ANALYSES AND SUMMARIES

6.1. Primary Efficacy Endpoint: Clinical Remission

6.1.1. Primary Analysis

- Endpoint: Proportions of subjects achieving clinical remission at Week 12.
- Analysis population**:** FAS.
- Analysis methodology**:** Exact Chan and Zhang method.
- Missing Data: Missing data will be handled by TFA.

6.1.2. Sensitivity Analysis

- Endpoint: Proportions of subjects achieving clinical remission at Week 12.
- Analysis population**:** Observed Cases (OC) and Per-Protocol (PP).
- Analysis methodology**:** Confidence interval of a proportion will be computed using Blyth-Still-Casella method, and the difference in proportions will be computed using Exact Chan and Zhang method.
- Missing Data: Missing data will be handled by TFA

Reporting results:

- Raw data: The sample size, percentage, and corresponding 95% confidence intervals for each treatment arm, difference of proportions, corresponding confidence intervals and 2-sided pvalue will be presented.

Figures

- Vertical barchart of proportions and coresponding 95% confidence interval of clinical remission patients at Week 12 for treatment and placebo arms.

6.1.3. Primary Safety Endpoint:

6.1.3.1. Primary Analysis

- Endpoint: Incidence and severity of adverse events, serious adverse events and withdrawals due to adverse events, ECGs, vital signs and safety laboratory tests.
- Analysis population**:** Safety Analysis Set.
- Analysis methodology**:** Will be analyzed in accordance with Pfizer Data Standards.
- Missing Data: Observed Cases (OC).

Reporting results:

- All safety data will be summarized descriptively through appropriate data tabulations, descriptive statistics, categorical summaries, and graphical presentations.

6.2. Secondary Endpoint(s)

For all Secondary Endpoints will be analyzed based on FAS.

6.2.1. Binary Endpoint(s)

Following binary secondary endpoints will be analyzed similar to the primary efficacy endpoint:

- *Proportion of subjects with endoscopic improvement at Week 12 (defined as decrease of ≥1 point in Mayo endoscopy score or an absolute endoscopy score of ≤1).*[b]
- *Proportion of subjects with a clinical response at Week 12 defined with a decrease from baseline of at least 3 points in total Mayo score with at least 30% change, accompanied by at least one point decrease or absolute score of 0 or 1 in rectal bleeding subscore.*
- *Proportion of subjects with change from baseline in partial Mayo Score of ≤2 with no individual subscore >1 at Weeks 2, 4, 6, 8, 12.*
- *Incidence of the development of HAFAs and Nabs against PF-06687234.*

For binary secondary endpoints that are measured repeatedly over time will be analyzed using generalized linear mixed model (specified in Section 5.2.1).

6.2.2. Continous Endpoint(s)

The following secondary endpoints will be analyzed per the following:

- *Mean change from baseline at Week 12 in Geboes histology score.*

  - Analysis time points**:** Week 12.

  - Analysis population (method of imputation for missing data)**:** FAS (BOCF)
    Analysis methodology**:** Change from baseline will be analyzed using the ANCOVA model (specified in Section 5.2.3).

- *Plasma concentrations of PF-06687234.*

  - Analysis time points**:** Week 12.

    - Analysis population (method of imputation for missing data)**:** PK
      Analysis methodology**:** Change from baseline will be analyzed using the MMRM method (see Section 5.2.2) with no imputations for any missing data (OC), ie, subjects with partial data will be included in MMRM analysis.

  Reporting results:

  - Raw data: The sample size, mean, standard deviation, median, minimum and maximum at baseline and post-baseline visits will be presented for each treatment arm.

  - Change from baseline: The sample size, mean, standard deviation, median, minimum and maximum will be presented for each treatment arm. The LS means, 95% confidence interval for the LS means, difference between the LS means for each pair of treatment groups and the corresponding 95% confidence interval will be presented.

  Figures

  Vertical bar chart of placebo adjusted LS means and 95% confidence interval at Week 12.

6.3. Other Endpoint(s): Exploratory Endpoint(s)

- *Collection of banked biospecimens and additional non-banked samples (eg, intestinal biopsies, stool for microbiome analysis, serum and plasma for analysis of proteins, whole blood for RNA analysis and epigenetics and/or cytometry) unless prohibited by local regulations or ethics committee decision. Additional information on collection and potential use is provided in the Banked Biospecimens section and in the exploratory analysis plan (EAP).*

- *Change from baseline in fecal calprotectin at Weeks 4, 8, 11.*

- *Change from baseline in hsCRP at Weeks 4,8,11.*
- *AUCτC_{max}, T_{max}, $t_{1/2}$, CL/F,and Vz/F as data permit.*
- *PF-06687234 tissue concentration in inflamed and non-inflamed biopsies.*

All other continuous and binary data will be similar to the primary and secondary endpoints. The biomarker data, such as hsCRP data, fecal calprotectin data, the mean change from baseline and the corresponding confidence interval will be summarized by treatment group. Exploratory biomarkers will be summarized by dose group and may be graphically displayed; appropriate analysis may be performed as needed. The analysis population will be based on a mITT analysis set and will not account for missing values. Some of the exploratory endpoint analyses such as analyses related to fecal calprotectin and hsCRP endpoints may be included in the CSR.

6.3.1. Pharmacokinetic Analysis

Actual PK sampling times will be used in the derivation of PK parameters.

If a PK parameter cannot be derived from a subject's concentration data, the parameter will be coded as NC (ie, not calculated). (Note that NC values will not be generated beyond the day that a subject discontinues).

In summary tables, statistics will be calculated by setting NC values to missing; and statistics will be presented for a particular dose with ≥3 evaluable measurements. If an individual subject has a known biased estimate of a PK parameter (due for example to an unexpected event), this will be footnoted in summary tables and will not be included in the calculation of summary statistics or statistical analyses.

The following PK parameters will be calculated for PF-06687234 (as data permits) from the concentration-time values using standard noncompartmental methods from patients with additional PK sampling: *AUCτ, C_{max}, T_{max}, $t_{1/2}$, CL/F and Vz/F.*

To assess the pharmacokinetics of *PF-06687234*, PK parameters will be listed and summarized for subjects in the PK analysis set and will include the set of summary statistics as specified in the table below:

Table 2. PK Parameters to be Summarized Descriptively

Parameter	Summary Statistics
AUC_{τ}, C_{max}, CL/F and V_Z/F	N, arithmetic mean, median, cv%, standard deviation, minimum, maximum, geometric mean and geometric cv%.
T_{max}	N, median, minimum, maximum.
$t_{1/2}$	N, arithmetic mean, median, cv%, standard deviation, minimum, maximum.

There will be 1 summary table presenting all PK parameters. This will include data from all treatments. The treatment subheading will include the dose information and any other pertinent information related to the intense PK population .

Median trough concentration plots will be generated for all subjects. For subjects with additional PK samples, median concentration time plots (linear and log scale) on Visit 1 and Visit 12 will be generated.

Data permitting, concentration of PF-06687234 in the biopsy tissues will be summarized and presented with summary statistics.

A population PK model may be developed for the purpose of estimating population PK parameters. Any population PK model developed to characterize the PK data will be reported separately.

Data permitting, the relationship between exposure and clinical responses (efficacy and safety) during 12 weeks of treatment in subjects with mild to moderate UC may be explored using either observed or modeled exposures.

Additionally, infliximab serum concentrations by treatment group may be summarized and presented.

6.3.2. Immunogenicity Analysis

Overall incidence of development of HAFA and of NAb will be reported along with relationships of incidence with respect to time. Both continuous endpoints and categorical endpoints (ie, positive and negative) will be reported for the HAFA and NAb assays by time points samples would be collected. Data permitting, the impact of HAFA and NAb on PK, PD, safety and efficacy profiles may be explored. Data permitting, median concentration time profile in HAFA/NAb positive subjects will be compared with that in HAFA/NAb ngative subjects.

Additionally, ADA against infliximab may be summarized and reported if data permits.

6.4. Subset Analyses

Proportion of subjects achieving clinical remission at Week 12 may be summarized for the subsets below (tables will be generated if we have sufficient number of subjects in each stratum):

- Anti-TNF experience (yes or no);
- Steroid use at baseline (yes or no);
- Gender: Female/Male;
- Disease duration;
- Fecal Calprotectin at baseline: <=500, >500 ug/g and <=250, >250 ug/g.

6.5. Baseline and Other Summaries and Analyses

6.5.1. Baseline Summaries

Demographics and medical history including variables defined in Section 3.3 will be summarized by treatment group according to Pfizer standards.

6.5.2. Study Conduct and Subject Disposition

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

6.5.3. Study Treatment Exposure

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, and number and percentage of subjects who are compliant with the dosing regimen.

6.5.4. Concomitant Medications and Non-Drug Treatments

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

6.6. Safety Summaries and Analyses

Safety analysis will be based on the SAF analysis set.

Safety and tolerability of PF-06687234 dose levels versus placebo: Adverse Events (AE), Serious Adverse Events (SAE), treatment-emergent adverse events (TEAE), withdrawal due to adverse events. All clinical AEs, SAEs, treatment-emergent signs and symptoms (TEAEs), withdrawal due to AEs, ECGs, vital signs and safety laboratory data will be reviewed and summarized on an ongoing basis during the study to evaluate the safety of subjects. Any adverse events that start or increase in severity following the start of treatment and occur within the follow-up period after the end of treatment will be considered as TEAE.

- Potential cases of drug-induced liver injury (ALT and AST elevation).
- Laboratory tests as specified in the protocol.
- Vital signs (blood pressure, heart rate, respiration and oral or tympanic temperature) will be measured at every study visit.
- ECG, single 12-lead ECGs will be obtained on all subjects at specified visits.

Other safety endpoints including hospitalization, physical examination, weight, concomitant medications.

A 3-tier approach will be used to summarize clinical AEs. Under this approach, AEs are classified into 1 of 3 tiers. However, no Tier-1 AE has been identified for this study. Different analyses will be performed for different tiers (See Section 8.2.3 of the protocol).

The MedDRA preferred term, treatment in B7581002 study, n (%) for eachMedDRA preferred term per arm, risk difference, 95% confidence interval and p-values for the risk difference will be provided. Graphical format may be presented aswell. Presented in descending p-value order. Please note that this analysis is exploratory in nature. The -pvalues will be presented for descriptive purpose only and no definitive conclusion should be made based on the p-values. Tier-2 events: A MedDRA PT is defined as a tier-2 event if it occurs in at least 4 subjects in any treatment group.

Tier-3 events: All other events are defined as tier-3 events.

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.

6.6.1. Laboratory Data

Using the current Pfizer data standards, the last pre-dose value before the first dosing is used as the baseline for all laboratory parameters.

6.6.2. Vital Signs

Baseline for vitals and ECG will be defined as the last pre-dose value before the first dosing.

If not supplied, QTcF will be derived using Fridericia's heart rate correction formula: QTcF = QT / (RR)^(1/3), where RR = 60/HR (if RR is not provided).

note: if QTcB is collected, then it should be listed only.

6.6.3. Electrocardiogram

ECG parameters, if applicable, will be summarized at:

- Baseline, Week 12, and Early Termination of induction period

6.6.4. Physical Examination

Physical examinations will be summarized at:

- Baseline, Week 12, and Early Termination of induction period.

7. INTERIM ANALYSES

7.1. Introduction

If the recruitment rate is very slow, an interim analysis for futility will be done with 40% completers. Since the interim analysis is for futility only, type I error rate will not be affected. A conditional power less than 10% will be used for futility. An Internal Review Committee (IRC) is established to oversee this Pfizer-sponsored trial. The primary rationale for establishing the committee is to make certain that appropriate safeguards are in place to help ensure the safety of subjects and/or to maintain scientific rigor and study integrity for these studies with planned interim analyses where treatment arms may be compared or a single treatment arm may be assessed with respect to the accumulating safety while the trial is on-going. The IRC is an independent oversight committee comprising 3 qualified and experienced Sponsor experts in their respective fields, but who are external to and independent of the study and the study team. The IRC consists of a chairperson and at least 2 additional members, including at least one with medical qualifications and at least one other who is a statistician. All members of the IRC will be unblinded voting members, and must be qualified and experienced in reviewing and interpreting clinical study data. A separate operating guidelines for this IRC with all the necessary details are given in the IRC Charter.

7.2. Interim Analyses and Summaries

The decision at the end of the study is being at least 90% confident that active treatment will be superior to placebo. The study will be stopped for futility for the active treatment arm if the conditional power with respect to the final analysis at the end of study is <10%. The details of conditional power computation will be given in the IRC. Due to the nature of the analyses the overall Type 1 error for the study is maintained. If the futility condition is met, the study will be stopped at the interim point. The interim analysis results will be used to facilitate internal decision-making. The results will only be distributed to a select list of individuals 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 IRC charter. 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 subjects still in the study. There are no prospective plans to stop the study early for success as a result of the interim analyses.

During the interim analysis, some members of the study team may be unblinded and replaced with blinded colleagues. The subjects, investigators, and individuals from the sponsor (or designee) who interact with the investigators and monitor safety will continue to be blinded to individual study treatments throughout the study.

8. REFERENCES

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

2. Dmitrienko, A., Molenberghs, G., Chuang- Stein, C., Offen, W. (2005). Analysis of Clinical Trials Using SAS: A Practical Guide. Cary, NC: SAS Institute Inc.

3. Jairath V, Zou G, Parker CE, et al. Systematic review with meta-analysis: placebo rates in induction and maintenance trials of ulcerative colitis. J Crohns Colitis 2016; 10(5):607.

9. APPENDICES

Appendix 1. Summary of Efficacy Analyses

Efficacy analyses comparing active and placebo arms will be based on the mITT and PP population.

Efficacy Endpoints	Analysis Method	Missing Data
Proportion of subjects achieving Clinical Remission at Week 12	Chan and Zhang	TFA
Proportion of subjects achieving Clinical Remission at Week 12	Chan and Zhang	OC
Proportion of subjects with endoscopic improvement at Week 12	Chan and Zhang	OC
Proportion of subjects with endoscopic improvement at Week 12	Chan and Zhang	TFA
Mean change from baseline at Week 12 in Geboes histology score.	ANCOVA	BOCF
Proportion of subjects with a clinical response at Week 12	Chan and Zhang	TFA
Proportion of subjects with a clinical response at Week 12	Chan and Zhang	OC
Proportion of subjects with change from baseline in partial Mayo Score of ≤2 with no individual subscore >1 at Weeks 2, 4, 6, 8, 12.	MMRM	OC
Plasma concentrations of PF-06687234 over time	MMRM	OC
Incidence of the development of HAFAs and Nabs against PF-06687234.	GLMM	OC
Change from baseline in serum hsCRP levels over time	MMRM	OC
Change from baseline in fecal calprotectin over time	MMRM	OC

Appendix 2. Definition and Use of Visit Windows in Reporting

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

Visit Label	Target Day	Definition [Day window]
Screening		Days -28 to Day 0
Induction Period Weeks 0-12		
Week 0	Day 1, Baseline	Day 1
Week 2	15	Days 2 to 21
Week 4	29	Days 22 to 42
Week 8	57	Days 43 to 70
Week 12	85	Days 71 to 98

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

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

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

Safety analysis will follow Pfizer standards.

Appendix 3. SAS Program Details

Appendix 3.1. Estimate and Confidence Interval of a binomial proportion (Blyth-Still-Casella)

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

Appendix 3.2. Estimate and Confidence Interval for Risk Difference

Appendix 3.2.1. SAS code for the Confidence Interval for Risk Difference using Chan and Zhang (1999)

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

Appendix 3.2.2. SAS Code for GLIMMIX Analysis for Binary Data Analysis with Multiple Timepoints

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

Appendix 3.2.3. SAS Code for LMMs

```
PROC MIXED DATA=_DATA;
        CLASS TREATMENT SUBJID TIME;
        MODEL EFFICACYCHANGE =TREATMENT BASELINE TIME
        TIME*TREATMENT/DDFM=KR SOLUTION;
        REPEATED TIME/ TYPE=UN SUBJECT=SUBJID;
RUN;
```

Appendix 3.2.4. SAS Code for ANCOVA Models

```
PROC MIXED DATA=_DATA;
        CLASS TREATMENT;
MODEL EFFICACYCHANGE = TREATMENT BASELINE /DDFM=KR SOLUTION;
RUN;
```