STATISTICAL ANALYSIS PLAN

Study Title:	A Phase 2, Double-Blind, Randomized, Placebo-Controlled Study Evaluating the Efficacy and Safety of Filgotinib in the Treatment of Perianal Fistulizing Crohn's Disease
Name of Test Drug:	Filgotinib
Study Number:	GS-US-419-4016
Protocol Version (Date):	Amendment 1: 06 January 2017 Amendment 1.1 EU: 28 February 2017 Amendment 2: 26 June 2017 Amendment 3: 11 October 2018 Amendment 4: 11 April 2019 Amendment 5: 04 February 2020
Analysis Type:	Final Analysis
Analysis Plan Version:	2.0
Analysis Plan Date:	01 March 2021
Analysis Plan Author(s):	PPD

CONFIDENTIAL AND PROPRIETARY INFORMATION

TABLE OF CONTENTS

LIST OF IN-TEXT TABLES

LIST OF IN-TEXT FIGURES

LIST OF ABBREVIATIONS

6-MP	6-mercaptopurine
AE	adverse event
AEI	adverse events of interest
ALT	alanine aminotransferase
ANCOVA	analysis of covariance
AST	aspartate aminotransferase
ASTE	arterial systemic thromboembolism
BLQ	below the limit of quantitation
BMI	body mass index
CCG	eCRF Completion Guidelines
CD	Crohn's disease
CDAI	Crohn's Disease Activity Index
CI	confidence interval
COVID-19	coronavirus disease 2019
CRF	case report form(s)
CSR	clinical study report
CTCAE	Common Terminology Criteria for Adverse Events
CV	cardiovascular
CVEAC	cardiovascular safety endpoint adjudication committee
DMC	data monitoring committee
ECG	electrocardiogram
eCRF	electronic case report form(s)
CCI	[redacted]
ET	early termination
FAS	Full Analysis Set
Gilead	Gilead Sciences, Inc.
HBV	Hepatitis B Virus
HCV	Hepatitis C Virus
HDL	high-density lipoprotein
HIV	Human Immunodeficiency Virus
HLT	high-level term
HLGT	high-level group term
HRQoL	health-related quality of life
CCI	[redacted]
IBD	inflammatory bowel disease
CCI	[redacted]
ID	identification
IM	intramuscularly

IV	intravenously
IWRS	interactive web response system
LDL	low-density lipoprotein
LLOQ	lower limit of quantification
LLT	lower-level term
LOCF	last observation carried forward
LOQ	limit of quantification
LTE	long-term extension
MACE	major adverse cardiovascular events
MRI	Magnetic Resonance Imaging
MedDRA	Medical Dictionary for Regulatory Activities
MI	myocardial infarction
MST	MedDRA search term
MTX	methotrexate
NLP	Natural Language Processing
NRI	non-responder imputation
CCI	[redacted]
O&P	ova and parasites test
OIs	opportunistic infections
PBO	Placebo
PDAI	Perianal Disease Activity Index
CCI	[redacted]
PO	orally
PR	rectally
CCI	[redacted]
PT	preferred term
PTM	placebo to match
PTx	Post-Treatment
Q1, Q3	first quartile, third quartile
SAE	serious adverse event
SAP	statistical analysis plan
SE	standard error
SD	standard deviation
SES-CD	simple endoscopic score for Crohn's disease
CCI	[redacted]
SMQ	Standardized MedDRA query
SOC	system organ class
TB	Tuberculosis
TEAE	treatment-emergent adverse event

TFLs	tables, figures, and listings
TNFα	tumor necrosis factor-alpha
ULN	upper limit of normal
US	United States
vfPBMC	viably frozen peripheral blood mononuclear cell
V-Day	visit day
VTE	venous thromboembolism
WHO	World Health Organization
CCI	[redacted]

DEFINITION OF TERMS

Simple Fistula (for stratification purposes)	Single, unbranched submucosal or intersphincteric fistula tract, no extensions, no inflammatory mass/collections, single external opening
Complex Fistula(e) (for stratification purposes)	Multiple simple fistulae **OR** Single branched (multiple external openings arising from one fistula tract), trans-, extra-, or suprasphincteric fistulae tracts, possible extensions, and/or focal to small collections
Perianal Fistula Closure	No fistula drainage despite gentle external compression
Clinical Fistula Remission	Perianal fistula closure of all external openings that were draining at baseline
Clinical Fistula Response	Reduction of ≥ 1 from baseline in the number of draining external perianal fistula openings that were present at baseline
Combined Fistula Remission	Perianal fistula closure of all external openings that were draining at baseline, and absence of fluid collections > 1cm on MRI pelvis
Combined Fistula Response	Reduction of ≥ 1 from baseline in the number of draining external perianal fistula openings that were present at baseline, and absence of fluid collections > 1cm on MRI pelvis
Proctitis SES-CD Score	Sum of ulcer size and ulcerated surface SES-CD endoscopy subscores for the rectum and anal canal, assessed by centrally read flexible sigmoidoscopy
Anal Canal Segmental SES-CD score	Sum of ulcer size and ulcerated surface SES-CD endoscopy subscores for the anal canal segment only, assessed by centrally read flexible sigmoidoscopy
Moderately to Severely Active Proctitis	Proctitis SES-CD Score > 2
Proctitis Remission	A Proctitis SES-CD Score of 0
Disease Worsening	An increase in CDAI of ≥ 100 points from the Week 10 value with CDAI ≥ 220 at 2 consecutive visits
Non-responder (luminal disease)	Subject who had a baseline CDAI score ≥ 220 and never achieves a ≥ 70 point CDAI reduction from baseline at any point up to and including Week 10 OR had a baseline CDAI score < 220 and have an increase in CDAI of ≥ 100 points from baseline, with CDAI ≥ 220 at Week 10
Non-Responder (Perianal Fistulizing Crohn's Disease (CD))	Subjects who meet the following criteria for PDAI symptom subscores: "Discharge" subscore of > 1 and ≥ 1 point increase from baseline at Week 6 and Week 10 OR "Pain/restriction of activities" subscore of > 1 and ≥ 1 point increase from baseline at Week 6 and Week 10
CCI [redacted]	[redacted]
CCI [redacted]	[redacted]
CCI [redacted]	[redacted]

CCI [redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
CCI [redacted]	[redacted]
4016 Inflammation Score	Sum of the following subscores for all primary fistula tracts and extensions (if present) identified at baseline: Hyperintensity of T2-weighted images Subscore Hyperintensity of T1-weighted Images Subscores Dominant Feature of Primary Tracts and Extensions Subscore Inflammatory Mass Subscore Presence of Proctitis Subscore
4016 Anatomy Score	Sum of the following subscores for all primary fistula tracts and extensions (if present) identified at baseline: Number of primary fistula tracks CCI [redacted] identified at baseline Location Subscore of all primary fistula tracts identified at baseline Extension Subscore for all extensions (if present) identified at baseline
Fistula Anatomy Subscore	Sum of the following subscores for each individual primary fistula tract (and extensions if present) identified at baseline: Location Subscore Extension Subscore
Fistula Inflammation Subscore	Sum of the following subscores for each individual primary fistula tract and extensions (if present) identified at baseline: Hyperintensity of T2-weighted Images Subscores Hyperintensity of T1-weighted Images Subscores Dominant Feature of Tract Subscores Inflammatory Mass Subscore
MRI Fistula Healing	A score of zero at Week 24 for Fistula Inflammation Subscore, for an individual primary fistula tract (and extensions if present) with Fistula Inflammation Subscore > 0 identified at baseline
MRI Fistula Response	≥ 50% decrease from baseline in Fistula Inflammation Subscore at Week 24 for an individual fistula tract (and extensions if present) with Fistula Inflammation Subscore > 0 identified at baseline
MRI Remission	All fistula tracts (and extensions if present) that were identified at baseline achieve MRI fistula healing at Week 24
MRI Response	All fistula tracts (and extensions if present) that were identified at baseline achieve MRI fistula response at Week 24

1. INTRODUCTION

This statistical analysis plan (SAP) describes the statistical analysis methods and data presentations to be used in tables, figures, and listings (TFLs) in the clinical study report (CSR) for Study GS-US-419-4016.This SAP is based on Protocol Amendment 5 dated 04 February 2020 and the electronic case report form (eCRF). The SAP will be finalized before database finalization. Any changes made after the finalization of the SAP will be documented in the CSR.

1.1. Study Objectives

The primary objective of this study is:

- To evaluate the efficacy of filgotinib as compared to placebo in establishing combined fistula response at Week 24

The secondary objectives of this study are:

- To evaluate the efficacy of filgotinib as compared to placebo in establishing combined fistula remission at Week 24
- To assess the time to clinical fistula response
- To assess the time to clinical fistula remission
- To evaluate the efficacy of filgotinib as compared to placebo in establishing proctitis remission at Week 24, in subjects that had moderately to severely active proctitis at baseline
- To evaluate the safety and tolerability of filgotinib

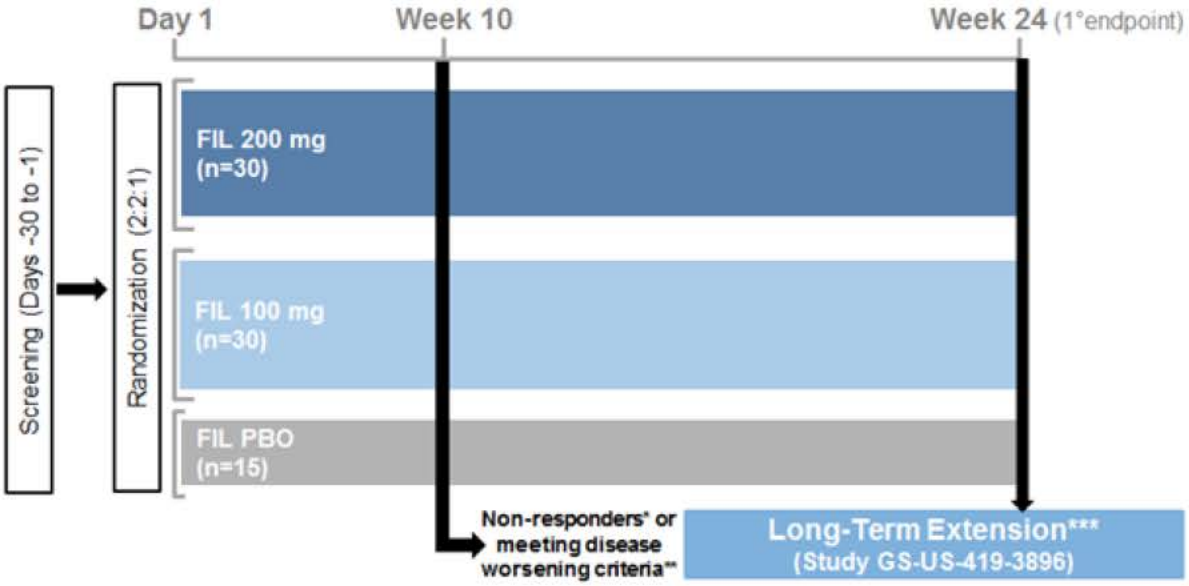
CCI

[redacted]

1.2. Study Design

This study is a Phase 2, double-blind, randomized, multicenter study evaluating the efficacy and safety of filgotinib versus placebo in the treatment of perianal fistulizing CD. A schematic of this study is provided in Figure 1-1.

Figure 1-1. Study Design Schematic



FIL = filgotinib; mg = milligram; PBO = placebo.

* Non-responders (luminal disease) are defined as subjects who either had a baseline CDAI score ≥ 220 and never achieve a ≥ 70 point CDAI reduction from baseline at any point up to and including Week 10 OR had a baseline CDAI score < 220 and have an increase in CDAI of ≥ 100 points from baseline, with CDAI ≥ 220 at Week 10
Non-responders (perianal fistulizing CD) are defined as subjects who meet the following criteria for PDAI symptom subscores: "Discharge" subscore of > 1 and ≥ 1point increase from baseline at Week 6 and Week 10 OR "Pain/restriction of activities" subscore of > 1 and ≥ 1point increase from baseline at Week 6 and Week 10

** Disease worsening is defined as an increase in CDAI of ≥ 100 points from the Week 10 value with CDAI score ≥ 220 points at 2 consecutive visits

*** Subjects who are non-responders, meeting disease worsening criteria, or complete the study at Week 24, will have the option to enter a separate LTE study, if eligible.

This study includes:

- Screening (Day -30 to -1)
- Randomization
- Blinded Treatment Period (Day 1 to Week 24)
  - Efficacy assessment: At Week 24, physical examination, flexible sigmoidoscopy to assess presence of proctitis, and MRI to assess fistula response
  - Subjects who complete all procedures per protocol, including the MRI at Week 24, may be offered the option to continue into a separate Long Term Extension (LTE) study (GS-US-419-3896), if deemed appropriate by the investigator
- Subjects who are non-responders at Week 10 (see Section 3.6 OR Section 3.7 of study protocol) or who meet disease worsening criteria (see Section 3.8 of study protocol) after Week 10 will have the option to enter the LTE study, if eligible.
- Post-treatment safety assessments:
  - All subjects completing this study will be offered the option to continue study drug in a blinded fashion in the LTE study, if eligible
  - Subjects who are eligible and opt to participate in the LTE study can continue into the LTE study without post-treatment safety assessments
  - Subjects who opt out of the LTE study will return to clinic 30 days after the last dose of study drug for post-treatment safety assessments

Based on protocol eligibility criteria, subjects will be randomized into 1 of 3 treatment groups in a blinded fashion in a 2:2:1 ratio as follows:

Treatment group 1 (n = 30): filgotinib 200 mg and PTM filgotinib 100 mg, once daily

Treatment group 2 (n = 30): filgotinib 100 mg and PTM filgotinib 200 mg, once daily

Treatment group 3 (n = 15): PTM filgotinib 200 mg and PTM filgotinib 100 mg, once daily

Note: US males who are not dual refractory will be randomized in a 2:1 ratio to either filgotinib 100 mg or matching placebo. For the purposes of this protocol, US males who have failed at least one prior TNF antagonist AND either previously failed vedolizumab induction therapy or are concomitantly taking vedolizumab at Day 1 and present with actively draining perianal fistulae despite continued vedolizumab therapy will be considered dual refractory subjects. This assessment is based on recently published data from a subgroup analysis of the GEMINI-2 study, showing a beneficial effect of vedolizumab treatment for subjects that were randomized with actively draining perianal fistulae at baseline {Feagan 2018}.

For subjects enrolled before protocol amendment 4, treatment assignments will be stratified according to the following factors:

- Anatomy of draining perianal fistulae at screening determined by MRI (simple *versus* complex)
- Prior exposure to TNFα antagonist (Yes or No)
- Presence of moderately to severely active proctitis (Yes or No)

For subjects enrolled after protocol amendment 4, treatment assignments will be stratified according to the following factors:

- Anatomy of draining perianal fistulae at screening determined by MRI (simple *versus* complex)
- Receiving vedolizumab therapy concomitantly at Day 1 (Yes or No)
- Presence of moderately to severely active proctitis (Yes or No)

1.3. Sample Size and Power

A total of approximately 75 subjects are planned to be randomized (2:2:1 ratio): 30 subjects to each of the filgotinib dose groups and 15 subjects to the placebo group. This sample size is considered adequate to assess the safety, tolerability, and efficacy of filgotinib in a descriptive manner.

Table 1-1 provides the exact 90% CI using the binomial distribution for a given combined fistula response rate ranging from 20% to 80%:

Table 1-1. Combined Fistula Response Rates (90% CI)

Number Randomized	Number of Responders	Response rate	Lower Limit of 90% CI	Upper Limit of 90% CI
30	6	20%	9%	36%
30	9	30%	17%	47%
30	12	40%	25%	57%
30	15	50%	34%	66%
30	18	60%	43%	75%
30	21	70%	53%	83%
30	24	80%	64%	91%

CI = confidence interval

2. TYPE OF PLANNED ANALYSIS

2.1. Data Monitoring Committee (DMC) Analyses

The DMC's role and responsibilities and the scope of analysis to be provided to the DMC are described in a mutually agreed upon charter, which defines the DMC membership, meeting logistics, and meeting frequency.

2.1.1. Safety Analyses

An external multidisciplinary DMC will review the progress of the study and perform interim reviews of the safety data in order to protect subject welfare and preserve study integrity. To ensure the best interests of the participants, the DMC will recommend to the sponsor if the nature, frequency, and severity of adverse effects associated with the study treatment warrant the early termination of the study, the continuation of the study, or the continuation of the study with modifications.

The DMC will meet twice to evaluate all available safety data of the study. The initial meeting will occur after approximately 20% of the planned total number of subjects complete their Week 10 visit. Following this, the next meeting will occur after approximately 50% of the subjects reach Week 10. Additionally, the DMC members may request an unscheduled review of the study data based on a concern for subject safety. Additional meetings may be triggered by safety findings, predetermined or otherwise.

2.1.2. Futility Analysis

After a minimum of 35 subjects (estimated to be 7 from the placebo group and 14 from each filgotinib treatment group based on the randomization ratio) are enrolled and complete the Week 10 visit or discontinue from the study, an interim futility analysis will be conducted to evaluate efficacy improvement in fistulizing disease. The cumulative safety analysis and summary statistics to evaluate perianal fistula closure of external openings, and CCI [redacted] will be generated and provided to the DMC. The DMC will make a recommendation to either continue the study without modification or recommend that the study be halted due to lack of efficacy, and will be based on the totality of data available at the time of the interim analysis.

2.2. Final Analysis

After all enrolled subjects have completed the study, outstanding data queries have been resolved or adjudicated as unresolvable, and the data have been cleaned and finalized, the study blind will be broken and the final analysis of the data will be performed.

3. GENERAL CONSIDERATIONS FOR DATA ANALYSES

Analysis results will be presented using descriptive statistics. For categorical variables, the number and percentage of subjects in each category will be presented; for continuous variables, the number of subjects (n), mean, standard deviation (SD) or standard error (SE), median, first quartile (Q1), third quartile (Q3), minimum, and maximum will be presented.

By-subject listings will be presented for all subjects in the All Randomized Analysis Set and sorted by subject ID number, visit date, and time (if applicable). Data collected on log forms, such as AEs, will be presented in chronological order within subject. The treatment group to which subjects were randomized will be used in the listings. Age, sex at birth, race, and ethnicity will be included in the listings, as space permits.

3.1. Analysis Sets

Analysis sets define the subjects to be included in an analysis. Analysis sets and their definitions are provided in this section. The analysis set will be identified and included as a subtitle of each table, figure, and listing.

For each analysis set, the number and percentage of subjects eligible for inclusion, as well as the number and percentage of subjects who were excluded and the reasons for their exclusion, will be summarized by treatment group.

A listing of reasons for exclusion from analysis sets will be provided by subject.

3.1.1. All Randomized Analysis Set

All Randomized Analysis Set includes all subjects who were randomized in the study. This is the primary analysis set for by-subject listings.

3.1.2. Full Analysis Set

The Full Analysis Set (FAS) includes all randomized subjects who took at least 1 dose of study drug. This is the primary analysis set for efficacy analyses.

3.1.3. Safety Analysis Set

The Safety Analysis Set includes all subjects who took at least 1 dose of study drug. This is the primary analysis set for safety analyses.

CCI [redacted]

[redacted]

CCI [redacted]

[redacted]

3.2. Subject Grouping

For analyses based on the All Randomized Analysis Set and the FAS, subjects will be grouped according to the treatment to which they were randomized. For analyses based on the Safety Analysis Set and Biomarker Analysis Set, subjects will be grouped according to the actual treatment received. The actual treatment received will differ from the randomized treatment only when their actual treatment differs from randomized treatment for the entire treatment duration.

CCI [redacted]

3.3. Strata and Covariates

Subjects will be randomly assigned to treatment groups via the interactive voice or web response system (IWRS) using a stratified randomization schedule. Details of the randomization ratio and stratification variables are provided in Section 1.2.

If there are discrepancies in stratification factor values between the IWRS and the clinical database (eCRF data), the values recorded in the clinical database will be used for analyses. For derivation of concomitant use of vedolizumab at Day 1, the start date of such medication should be before or on the same date of the first dose of study drug and with either the stop date of such medication being on or after the first dose of study drug or with "ongoing" status.

3.4. Examination of Subject Subgroups

There are no prespecified subject subgroups for efficacy and safety analyses.

3.5. Missing Data and Outliers

3.5.1. Missing Data

In general, missing data will not be imputed unless methods for handling missing data are specified. Exceptions are presented in this document.

For missing last dosing date of study drug, imputation rules are described in Section 4.2.1. Imputation and calculation rules for missing clinical fistula assessment and MRI data are described in Section 6.2.2. The handling of missing or incomplete dates for AE onset is described in Section 7.1.5.2, and for prior and concomitant medications in Section 7.4.

Values for missing safety laboratory data will not be imputed. However, a missing baseline result will be replaced with a screening result, if available. If no pretreatment laboratory value is

available, the baseline value will be assumed to be normal (ie, no grade) for the summary of graded laboratory abnormalities. If safety laboratory results for a subject are missing for any reason at a time point, the subject will be excluded from the calculation of summary statistics for that time point.

Values for missing vital signs data will not be imputed. However, a missing baseline result will be replaced with a screening result, if available.

3.5.2. Outliers

Outliers will be identified during the data management and data analysis process, but no sensitivity analyses will be conducted. All data, including outliers, will be included in the data analysis, unless otherwise specified.

3.6. Data Handling Conventions and Transformations

The following conventions will be used for the imputation of date of birth when it is partially missing or not collected:

- If only month and year of birth is collected, then "15" will be imputed as the day of birth
- If only year of birth is collected, then "01 July" will be imputed as the day and month of birth
- If year of birth is missing, then date of birth will not be imputed

In general, age collected at Day 1 (in years) will be used for analyses and presented in listings. If age at Day 1 is not available for a subject, then age derived based on date of birth and the Day 1 visit date will be used instead. If an enrolled subject was not dosed with any study drug, the randomization date will be used instead of the Day 1 visit date. For screen failures, the date the first informed consent was signed will be used for the age derivation. Age required for longitudinal and temporal calculations and analyses (eg, estimates of creatinine clearance, age at date of AE) will be based on age derived from date of birth and the date of the measurement or event, unless otherwise specified.

Duration of perianal fistulizing CD in years is the number of years between the diagnosis date of perianal fistulizing CD and date of first dose of study drug. The partial diagnosis date of perianal fistulizing CD (if any) will be imputed for calculation as follows:

- If day and month are missing but year is available, then the imputed day and month will be 01 Jan.
- If day is missing but the month and year are available, then the imputed day will be the first day of the month.
- Partial date will not be imputed if the year is missing.

Non-PK data that are continuous in nature but are less than the lower limit of quantitation (LOQ) or above the upper LOQ will be imputed as follows:

- A value that is 1 unit less than the LOQ will be used to calculate descriptive statistics if the datum is reported in the form of "< x" (where x is considered the LOQ). For example, if the values are reported as < 50 and < 5.0, values of 49 and 4.9, respectively, will be used to calculate summary statistics. An exception to this rule is any value reported as < 1 or < 0.1, etc. For values reported as < 1 or < 0.1, a value of 0.9 or 0.09, respectively, will be used to calculate summary statistics.

- A value that is 1 unit above the LOQ will be used to calculate descriptive statistics if the datum is reported in the form of "> x" (where x is considered the LOQ). Values with decimal points will follow the same logic as above.

- The LOQ will be used to calculate descriptive statistics if the datum is reported in the form of "≤ x" or "≥ x" (where x is considered the LOQ).

Plasma concentration values that are below the limit of quantitation (BLQ) will be presented as "BLQ" in the concentration data listing. Values that are BLQ will be treated as 0 at predose time points, and one-half the value of the LOQ at postdose time points for summary purposes.

The following conventions will be used for the presentation of summary and order statistics:

- If at least 1 subject has a concentration value of BLQ for the time point, the minimum value will be displayed as "BLQ."

- If more than 25% of the subjects have a concentration data value of BLQ for a given time point, the minimum and Q1 values will be displayed as "BLQ."

- If more than 50% of the subjects have a concentration data value of BLQ for a given time point, the minimum, Q1, and median values will be displayed as "BLQ."

- If more than 75% of the subjects have a concentration data value of BLQ for a given time point, the minimum, Q1, median, and Q3 values will be displayed as "BLQ."

- If all subjects have concentration data values of BLQ for a given time point, all order statistics (minimum, Q1, median, Q3, and maximum) will be displayed as "BLQ."

3.7. Analysis Visit Windows

3.7.1. Definitions

The First Dosing Date is defined as the date when subjects take the first dose of study drug, as recorded in the Study Drug Administration eCRF.

The Last Dosing Date is defined as the date when subjects take the last dose of study drug as recorded in the Study Drug Administration eCRF.

Study Day will be calculated from the first dosing date of study drug and derived as follows:

- For days on or after first dosing date: Assessment Date – First Dosing Date + 1
- For days prior to the first dosing date: Assessment Date – First Dosing Date

Therefore, Study Day 1 is the day of the first dose of study drug administration.

Baseline is defined as the last available observation prior to or on the first dosing date, unless specified otherwise.

3.7.2. Analysis Visit Windows

Subject visits might not occur on protocol-specified days. Therefore, for the purpose of analysis, observations will be assigned to analysis windows.

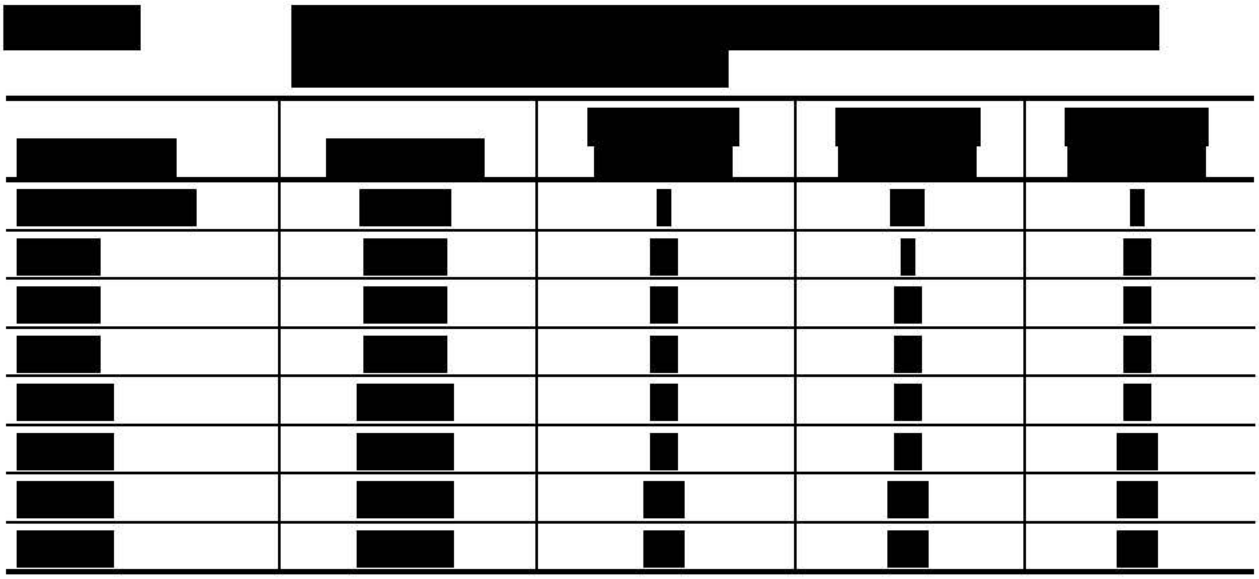
CCI [redacted]

CCI [redacted]

[redacted]

The analysis windows for weight, vital signs, ECG, safety laboratory parameters (hematology, chemistry, fasting lipid profile, serum immunoglobulin), and key inflammatory biomarkers (including hs-CRP, fecal calprotectin, and fecal lactoferrin) are provided in Table 3-4.

CCI



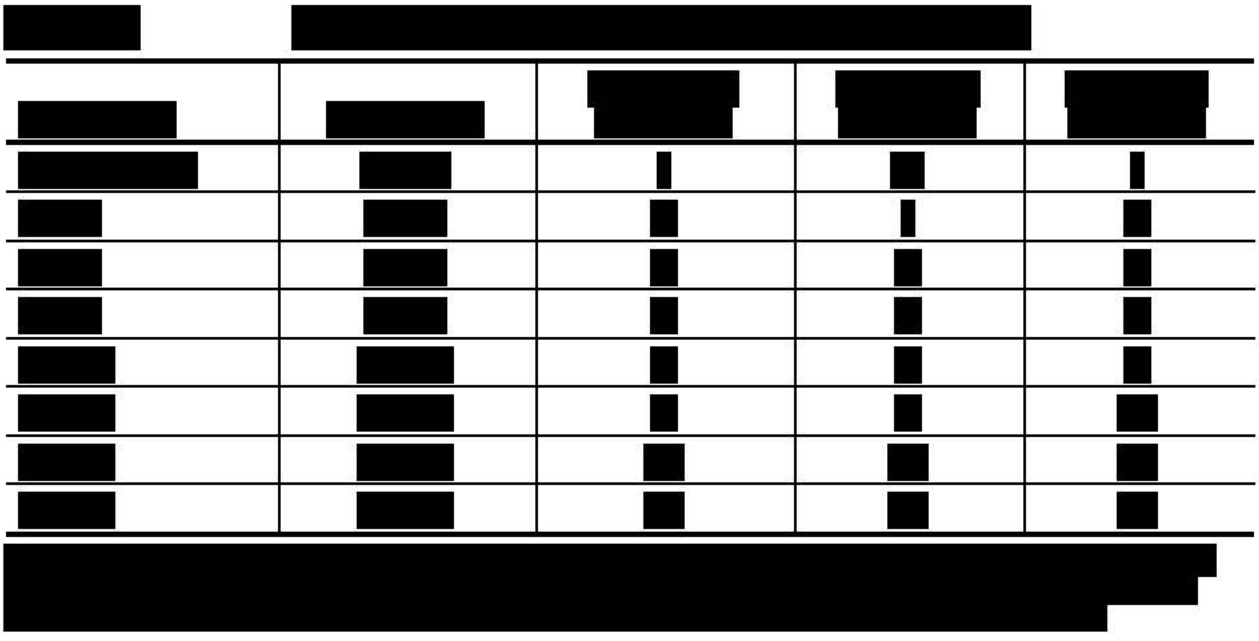


Table 3-4. Analysis Visit Windows for: Weight, Vital Signs, ECG, Hematology, Chemistry, Fasting Lipid Profile, Serum Immunoglobulin, CCI [redacted]

Nominal Visit	Analysis Visit	Nominal Day (Study Day)	Lower Limit (Study Day)	Upper Limit (Study Day)
Screening/Day 1	Baseline	1	(none)	1
Week 2	Week 2	15	2	22
Week 4	Week 4	29	23	36
Week 6	Week 6	43	37	57
Week 10	Week 10	71	58	85
Week 14	Week 14	99	86	113
Week 18	Week 18	127	114	141
Week 24	Week 24	169	142	≥ 169

Note: ECG is collected postbaseline at Week 10 and the corresponding analysis windows will be applied; fasting lipid profile, fecal lactoferrin, and fecal calprotectin are collected postbaseline at Weeks 10 and 24 and the corresponding analysis windows will be applied; serum immunoglobulin is collected postbaseline at Weeks 4, 10, and 24 and the corresponding analysis windows will be applied.

An unscheduled visit prior to the first dosing of study drug may be included in the calculation of the baseline value, if applicable.

3.7.3. Selection of Data in the Event of Multiple Records in an Analysis Visit Window

Depending on the statistical analysis method, single values may be required for each analysis window. For example, change from baseline by visit usually requires a single value, whereas a time-to-event analysis would not require 1 value per analysis window.

If multiple, valid, nonmissing, continuous measurements exist in an analysis window, records will be chosen based on the following rules if a single value is needed:

- In general, the baseline value will be the last nonmissing value on or prior to the first dosing date of study drug, unless otherwise specified. If multiple measurements occur on the same day, the last nonmissing value prior to the first dosing date of study drug will be considered as the baseline value. If these multiple measurements occur at the same time or the time is not available, the average of these measurements (for continuous data) will be considered the baseline value.
- For postbaseline visits:
  - The record closest to the nominal day for that visit will be selected.
  - If there are 2 records that are equidistant from the nominal day, the later record will be selected.
  - If there is more than 1 record on the selected day, the average will be taken, unless otherwise specified.

If multiple, valid, nonmissing, categorical measurements exist in an analysis window, records will be chosen based on the following rules if a single value is needed:

- For baseline, the last available record on or prior to the date of the first dose of study drug will be selected. If there are multiple records with the same time or no time recorded on the same day, the value with the lowest severity will be selected (eg, normal will be selected over abnormal for safety ECG findings).

- For postbaseline visits:

  — The record closest to the nominal day for that visit will be selected.

  — If there are 2 records that are equidistant from the nominal day, the later record will be selected.

- If there is more than 1 record on the selected day, the worst severity will be taken (eg, abnormal will be selected over normal for safety ECG findings), unless otherwise specified.

3.8. Assessment of Coronavirus Disease 2019 Impact

This study was ongoing during the coronavirus disease 2019 (COVID-19) pandemic.

The assessment of the impact of COVID-19 will be provided in the corresponding sections (eg, Section 4) throughout this SAP. Study drug and study discontinuation due to COVID-19 are described in Section 4.1, study drug interruptions due to COVID-19 are described in Section 4.2, important and non-important protocol deviations due to COVID-19 are described in Section 4.3, and missed and virtual visits due to COVID-19 are described in Section 4.4.

Hematocrit collected from local laboratory and weight measured at home due to COVID-19 impact will be used in CDAI calculations only, and will not be included in the safety laboratory summary (Sections 7.2 and 7.3, respectively). These data will be included in listings, with a flag to indicate which records were collected at a local laboratory or at home, respectively.

4. SUBJECT DISPOSITION

4.1. Subject Enrollment and Disposition

A summary of subject enrollment will be provided by treatment group for each country, investigator within a country, and overall. The summary will present the number and percentage of subjects enrolled. For each column, the denominator for the percentage calculation will be the total number of subjects analyzed for that column.

A similar enrollment table will be provided by randomization stratum. The denominator for the percentage of subjects in the stratum will be the total number of enrolled subjects. If there are discrepancies in the value used for stratification assignment between the IWRS and the clinical database, the value collected in the clinical database will be used for the summary. A listing of subjects with discrepancies in the value used for stratification assignment between the IWRS and the clinical database at the time of data finalization will be provided.

The randomization schedule used for the study will be provided as an appendix to the CSR.

A summary of subject disposition will be provided by treatment group and overall. This summary will present the number of subjects screened, the number of subjects who met all eligibility criteria but were not randomized with reasons subjects were not randomized, the number of subjects randomized, and the number of subjects in each of the categories listed below:

- All Randomized Analysis Set
- Full Analysis Set
- Safety Analysis Set
- [redacted]
- [redacted]
- Completed study drug
- Continuing study drug (for analysis other than final analysis)
- Did not complete study drug with reasons for premature discontinuation of study drug
- Completed study
- Continuing study (for analysis other than final analysis)
- Did not complete the study with reasons for premature discontinuation of study

For the status of study drug and study completion and reasons for premature discontinuation, the number and percentage of subjects in each category will be provided. The denominator for the percentage calculation will be the total number of subjects in the Safety Analysis Set corresponding to that column.

The following by-subject listings will be provided by subject identification (ID) number in ascending order to support the above summary tables:

- Reasons for premature study drug or study discontinuation (A separate listing of reasons for premature study drug or study discontinuation due to COVID-19 will be created.)
- Reasons for screen failure (will be provided by screening ID number in ascending order)
- Lot number and kit ID of assigned study medication

4.2. Extent of Study Drug Exposure and Adherence

Extent of exposure to study drug will be examined by assessing the total duration of exposure to study drug and the level of adherence to the study drug specified in the protocol. Summaries of extent of study drug exposure and adherence will be provided by treatment group for the Safety Analysis Set.

A by-subject listing of study drug administration will be provided for subjects with study drug interruption due to COVID-19.

4.2.1. Duration of Exposure to Study Drug

Total duration of exposure to study drug will be defined as last dosing date minus first dosing date plus 1, regardless of any temporary interruptions in study drug administration, and will be expressed in weeks using up to 1 decimal place (eg, 4.5 weeks). If the last study drug dosing date is missing, the latest date among the study drug end date, clinical visit date, laboratory sample collection date, and vital signs assessment date that occurred during the on-treatment period will be used.

The total duration of exposure to study drug will be summarized using descriptive statistics (n, mean, SD, median, Q1, Q3, minimum, and maximum) and using the number (ie, cumulative counts) and percentage of subjects exposed through the following time periods: 1 day, 2 weeks, 4 weeks, 6 weeks, 10 weeks, 14 weeks, 18 weeks, 20 weeks and 24 weeks. Summaries will be provided by treatment group for the Safety Analysis Set.

No formal statistical testing is planned.

4.2.2. Adherence to Study Drug

The total number of tablets administered will be summarized using descriptive statistics.

The presumed total number of tablets administered to a subject will be determined by the data collected on the drug accountability eCRF using the following formula. If the bottle was not returned, it is assumed that the subject took all the study drug tablets from the dispensed bottle. The number of tablets returned will be imputed as zero for the given bottle for study drug adherence calculation purpose.

$$\text{Total Number of Tablets Administered} = \left(\sum \text{No. of Tablets Dispensed}\right) - \left(\sum \text{No. of Tablets Returned}\right)$$

4.2.2.1. On-Treatment Adherence

The level of on-treatment adherence to the study drug regimen will be determined by the total amount of study drug administered relative to the total amount of study drug expected to be administered during a subject's actual on-treatment period based on the study drug regimen.

The level of on-treatment adherence will be expressed as a percentage using the following formula:

$$\text{On-Treatment Adherence (\%)} = \left(\frac{\text{Total Amount of Study Drug Administered}}{\text{Study Drug Expected to be Administered on Treatment}}\right) \text{x } 100$$

If the calculated on-treatment adherence is > 100%, it will be set to 100%. Descriptive statistics for the level of on-treatment adherence with the number and percentage of subjects belonging to adherence categories (eg, < 80%, ≥ 80% to < 90%, ≥ 90%) will be provided by treatment group for the Safety Analysis Set.

No formal statistical testing is planned.

A by-subject listing of study drug administration and drug accountability will be provided by subject ID number (in ascending order) and visit (in chronological order).

4.3. Protocol Deviations

Subjects who did not meet the eligibility criteria for study entry, but enrolled in the study will be summarized regardless of whether they were exempted by the sponsor or not. The summary will present the number and percentage of subjects who did not meet at least 1 eligibility criterion and the number of subjects who did not meet specific criteria by treatment group and overall based on the All Randomized Analysis Set. A by-subject listing will be provided for those subjects who did not meet at least 1 eligibility (inclusion or exclusion) criterion. The listing will present the eligibility criterion (or criteria if more than 1 deviation) that subjects did not meet and related comments, if collected.

Protocol deviations occurring after subjects entered the study are documented during routine monitoring. The number and percentage of subjects with important protocol deviations by deviation category (eg, eligibility criteria, informed consent) will be summarized by treatment group and overall for the All Randomized Analysis Set. A by-subject listing will be provided for those subjects with any important protocol deviation.

A by-subject listing will be provided for subjects with important protocol deviations related to COVID-19. A separate listing will be provided for subjects with non-important protocol deviations related to COVID-19.

4.4. Missed and Virtual Visits due to COVID-19

A by-subject listing of subjects with missed or virtual visits due to COVID-19 will be provided by subject ID number in ascending order.

Information regarding missed or virtual visits due to COVID-19 was collected as free text in the CRF comment fields. The determination of missing or virtual visits due to COVID-19 was done using Natural Language Processing (NLP) to search the CRF comment fields. A detailed explanation of the algorithm is provided in Appendix 12.

5. BASELINE CHARACTERISTICS

5.1. Demographics and Other Baseline Characteristics

Subject demographic variables and other baseline characteristics will be summarized by treatment group and overall using descriptive statistics for continuous variables, and using the number and percentage of subjects for categorical variables. The summary of demographic data will be provided for the Safety Analysis Set.

- Age (years, on the date of first dose of the study drug)
- Age group (< 65 years, ≥ 65 years)
- Sex at birth (female, male)
- Race
- Ethnicity (Hispanic or Latino, Not Hispanic or Latino)
- Geographic region (United States [US], non-US)
- Weight (in kg)
- Height (in cm)
- Body mass index (BMI; in kg/m^2)
- Smoking status (former, current, never)

A by-subject demographic listing, including the informed consent date, will be provided by subject ID number in ascending order.

5.2. Baseline Disease Characteristics

Stratification factors (listed in Section 3.3) and other baseline disease characteristics listed below will be summarized for the Safety Analysis Set by the same method as demographic tables, using descriptive statistics for continuous variables and using number and percentage of subjects for categorical variables.

- Duration of perianal fistulizing CD (years) from date of diagnosis to first dosing date of study drug
- Duration of perianal fistulizing CD group (< 1 year, ≥ 1 to < 3 years, ≥ 3 to < 7 years, ≥ 7 years)

- [redacted]
- [redacted]
- [redacted]
- Location of CD: ileal, colonic, ileocolonic and isolated upper disease
- [redacted]
- Active Luminal Disease (CDAI ≥ 150)
- Moderately Active Luminal Disease (CDAI ≥ 220)
- Baseline Proctitis SES-CD score (sum of size of ulcers and ulcerated surface for the rectum and anal canal segments)
- Baseline Anal Canal SES-CD score (sum of size of ulcers and ulcerated surface for the anal canal segment)
- Presence of Moderately to Severely Active Proctitis (yes, no)
- [redacted]
- Number of Draining Perianal Fistula at Baseline
- Anatomy of draining perianal fistulae at screening determined by MRI (simple vs complex)
- History of complications of perianal CD
  - Anal stenosis (yes, no)
  - Perirectal / perianal abscess (yes, no)
- Perianal Fistulizing CD treatment history
  - Prior use of antibiotics (yes, no)
  - Prior use of immunomodulators (yes, no)
  - Prior use of TNFα antagonist as listed in Appendix 2 (yes, no); and for subjects with prior use:
    - Number of TNFα antagonists used
      - 1
      - 2
      - ≥ 3

    - Worst outcome of prior use of TNFα antagonist
        - Treatment failure
        - Intolerance, including both allergic and non-allergic intolerance
        - Other

- Other CD treatment history
    - Prior use of systemic corticosteroids (yes, no)
    - Number of prior exposure to biologic agent for CD as listed in Appendix 3 (0, 1, 2, ≥ 3)
    - Prior use of vedolizumab (yes, no); and for subjects with prior use:
        - Worst outcome of prior use of vedolizumab
            - Treatment failure
            - Intolerance, including both allergic and non-allergic intolerance
            - Other
    - Prior use of ustekinumab (yes, no) and for subjects with prior use:
        - Worst outcome of prior use of ustekinumab
            - Treatment failure
            - Intolerance, including both allergic and non-allergic intolerance
            - Other

    Note: The worst outcome of a prior treatment is treatment failure, followed by intolerance, and then other outcomes.

- Concomitant CD treatment
    - Antibiotic treatment for fistulizing CD at Baseline (yes, no)
    - Concomitant use of vedolizumab at Day 1 (yes, no)

- Concomitant use of systematically absorbed corticosteroids and immunomodulators
  - Systematically absorbed corticosteroids only
  - Immunomodulator only
  - Both systematically absorbed corticosteroids and immunomodulator
  - Neither
- Prednisone equivalent dose for subjects who were on systemically absorbed corticosteroids at baseline (mg/day)
- Prednisone equivalent dose for subjects who were on systemically absorbed corticosteroids at baseline (> 0 to 10 mg/day, > 10 to 20 mg/day, > 20 mg/day)
- 5-aminosalicylates (yes, no)

A by-subject listing of the above baseline characteristics will be provided by subject ID number in ascending order.

5.3. Medical History

Medical history (disease-specific and general conditions) and IBD family history data will be collected at screening and presented in data listings.

General medical history data will be coded using the current version of Medical Dictionary for Regulatory Activities (MedDRA).

6. EFFICACY ANALYSES

6.1. General Considerations

The efficacy analysis will be conducted on the FAS, defined in Section 3.1.2, unless otherwise specified. No formal hypothesis testing will be performed.

The definitions of selected efficacy endpoints are provided in Table 6-1. Appendix 10 and Appendix 11 include detailed definition of study treatment failure and the corresponding data handling rules for efficacy analysis.

Table 6-1. Definitions of Selected Efficacy Endpoints

Type	Endpoint	Definition
Primary	Combined fistula response at Week 24	Reduction of ≥ 1 from baseline in the number of draining external perianal fistula openings that were present at baseline, and absence of fluid collections > 1cm on MRI pelvis at Week 24, among subjects with at least 1 draining external perianal fistula opening at baseline
Secondary	Combined fistula remission at Week 24	Perianal fistula closure of all external openings that were draining at baseline, and absence of fluid collections > 1cm on MRI pelvis at Week 24, among subjects with at least 1 draining external perianal fistula opening at baseline
Secondary	Time to clinical fistula response	The time interval in days from date of first dosing of study drug to the first observation (during scheduled or unscheduled clinical visits) when ≥ 1 of the draining external perianal fistula openings that were present at baseline achieves perianal fistula closure, among subjects with at least 1 draining external perianal fistula opening at baseline. Subjects not known to have a clinical fistula response will have their clinical fistula response time censored at the last time that lack of clinical fistula response was documented.

Type	Endpoint	Definition
Secondary	Time to clinical fistula remission	The time interval in days from date of first dosing of study drug to the first observation (during schedule or unscheduled clinical visits) of perianal fistula closure of all external openings that were draining at baseline, among subjects with at least 1 draining external perianal fistula opening at baseline. Subjects not known to have a clinical fistula remission will have their clinical fistula remission time censored at the last time that lack of clinical fistula remission was documented.
Secondary	Proctitis remission at Week 24, in subjects that had moderately to severely active proctitis at baseline	A Proctitis SES-CD Score (sum of ulcer size and ulcerated surface SES-CD endoscopy subscores for the rectum and anal canal) of 0 assessed by centrally read flexible sigmoidoscopy at Week 24, in subjects that had moderately to severely active proctitis at baseline
[redacted]	[redacted]	CCI [redacted]
[redacted]	[redacted]	CCI [redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]

Summary of Dichotomous Efficacy Endpoints

Numbers and percentages of subjects achieving each of the dichotomous efficacy endpoints defined in Table 6-1, and numbers and percentages of subjects not achieving those endpoints for the following reasons will be summarized by treatment, in a hierarchy order with the first reason being the highest.

1) Subjects who do not meet the endpoint based on observed data

2) Subjects who do not meet the endpoint due to study treatment failure

CCI [redacted]

CCI [redacted]

5) Subjects who do not meet the endpoint due to study drug discontinuation led by protocol-specified disease worsening

6) Subjects who do not have sufficient measurements to determine the endpoint due to study drug discontinuation for other reasons

7) Subjects who do not have sufficient measurements to determine the endpoint due to other reasons

Non-responder Imputation

For analysis of binary efficacy endpoints defined in Table 6-1, subjects who do not have sufficient measurements due to any reason (including treatment failures and CDAI or PDAI non-responders at Week 10) to determine the endpoint will be considered not meeting the endpoint (ie, non-responder imputation [NRI]).

CCI [redacted]

CCI [redacted]

CCI [redacted]

6.1.2. Calculation of Segmental SES-CD Scores

Proctitis SES-CD score (which includes an overall assessment of both rectum and anal canal segments) and Anal Canal Segmental SES-CD score will be calculated at Screening (used as baseline) and Week 24 visit (or early termination, if applicable). Only size of ulcers and ulcerated surface variables will be included for the calculation of these scores (see Appendix 6).

6.2. Primary Efficacy Endpoint

6.2.1. Definition of the Primary Efficacy Endpoint

The primary efficacy endpoint is the combined fistula response at Week 24. Please refer to Table 6-1 in Section 6.1 for the endpoint definition.

6.2.2. Analysis of the Primary Efficacy Endpoint

The number and proportion of subjects establishing combined fistula response at Week 24 for each treatment group will be summarized with corresponding 90% exact CI by Clopper-Pearson method. The difference in proportions between each filgotinib dose group and the placebo group will be presented along with the associated 90% CI. No formal hypothesis testing will be performed.

For subjects who completed the Week 24 visit, the following missing value imputation will be used before deriving the proportion of subjects establishing combined fistula response at Week 24:

- If the clinical fistula assessment data is missing at Week 24, then the number of draining external perianal fistula openings at Week 24 will be imputed by Last Observation Carried Forward (LOCF) method using the number of draining external perianal fistulae openings from the last available scheduled or unscheduled visit, among draining external perianal fistulae openings at baseline only.

- If the MRI data is missing at Week 24, then absence of fluid collection > 1 cm will be imputed based on whether there was an absence of fluid collection > 1 cm at baseline (Yes, No), which will be derived from the size of fluid collections at baseline from the Total Fistula Score CRFs.

6.3. Secondary Efficacy Endpoints

6.3.1. Definition of Secondary Efficacy Endpoints

The secondary efficacy endpoints include the following:

- Proportion of subjects establishing combined fistula remission at Week 24
- Time to clinical fistula response
- Time to clinical fistula remission
- Proportion of subjects achieving proctitis remission at Week 24, in subjects that had moderately to severely active proctitis at baseline

Please refer to Table 6-1 for the definitions of the endpoints.

6.3.2. Analysis of the Secondary Efficacy Endpoints

For binary endpoints, the same statistical analysis method for analyzing the primary efficacy endpoint will be utilized.

If the proctitis SES-CD score is missing at Week 24, no imputation will be done and the subject will be considered a non-responder when deriving the proportion of subjects achieving proctitis remission at Week 24.

For time-to-event secondary efficacy endpoints, the summary statistics (eg, median, Q1, Q3) from the Kaplan-Meier method will be provided for each treatment group. Plot of Kaplan-Meier curves will also be provided. The Cox proportional hazards model will be used to estimate the hazard ratio with the corresponding 90% CI. Subjects not known to have a clinical fistula response (or remission) will have their clinical fistula response (or remission) time censored at the last perianal fistula assessment date.

No formal hypothesis testing will be performed.

CCI [redacted]

CCI [redacted]

CCI [redacted]

[redacted]

CCI [redacted]

[redacted]

[redacted]

CCI [redacted]

[redacted]

[redacted]

CCI [redacted]

[redacted]

[redacted]

CCI [redacted]

[redacted]

CCI

[redacted]

CCI

[redacted]

[redacted]

CCI [redacted]

CCI

7. SAFETY ANALYSES

Unless otherwise specified, summaries of safety data will be provided for the Safety Analysis Set and will include data collected up to the last dose of study drug plus 30 days.

7.1. Adverse Events and Deaths

7.1.1. Adverse Event Dictionary

Clinical and laboratory adverse events (AEs) will be coded using the current version of Medical Dictionary for Regulatory Activities (MedDRA). System organ class (SOC), high-level group term (HLGT), high-level term (HLT), preferred term (PT), and lower-level term (LLT) will be provided in the AE dataset.

7.1.2. Adverse Event Severity

Adverse events are graded by the investigator as Grade 1, 2, 3, 4, or 5 according to toxicity criteria specified in the protocol. The severity grade of events for which the investigator did not record severity will be categorized as "missing" for tabular summaries and data listings.The missing category will be listed last in the summary presentation.

7.1.3. Relationship of Adverse Events to Study Drug

Related AEs are those for which the investigator selected "Related" on the AE eCRF to the question of "Related to Study Treatment." Relatedness will always reflect the investigator assessment of causality rather than the Sponsor's. Events for which the investigator did not record relationship to study drug will be considered related to study drug for summary purposes. However, by-subject data listings will show the relationship as missing.

7.1.4. Serious Adverse Events

Serious adverse events (SAEs) will be identified and captured as SAEs if AEs met the definitions of SAE specified in the study protocol. SAEs captured and stored in the clinical database will be reconciled with the SAEs captured in the Gilead safety database before data finalization.

7.1.5. Treatment-Emergent Adverse Events

7.1.5.1. Definition of Treatment-Emergent Adverse Events

Treatment-emergent adverse events (TEAEs) are defined as 1 or both of the following:

- Any AEs with an onset date on or after the study drug start date (Day 1) and no later than 30 days after permanent discontinuation of study drug
- Any AEs leading to premature discontinuation of study drug.

7.1.5.2. Incomplete Dates

If the onset date of the AE is incomplete and the AE stop date is not prior to the first dosing date of study drug, then the month and year (or year alone if month is not recorded) of onset determine whether an AE is treatment emergent. The event is considered treatment emergent if both of the following 2 criteria are met:

- The AE onset date is the same as or after the month and year (or year) of the first dosing date of study drug, and
- The AE onset date is the same as or before the month and year (or year) of the date corresponding to 30 days after the date of the last dose of study drug

An AE with completely missing onset and stop dates, or with the onset date missing and a stop date later than the first dosing date of study drug, will be considered to be treatment emergent. In addition, an AE with the onset date missing and an incomplete stop date with the same or later month and year (or year alone if month is not recorded) as the first dosing date of study drug will be considered treatment emergent.

7.1.6. Summaries of Adverse Events and Deaths

Treatment-emergent AEs will be summarized based on the Safety Analysis Set. No formal statistical testing is planned.

7.1.6.1. Summaries of AE Incidence

The number and percentage of subjects who reported at least 1 TEAE will be provided and summarized by SOC, HLT, PT, and treatment group. For other AEs described below, summaries will be provided by SOC, PT, and treatment group:

- TEAEs of Grade 3 or higher (by maximum severity)
- TEAEs of Grade 2 or higher (by maximum severity)
- All TE treatment-related AEs
- TE treatment-related AEs of Grade 3 or higher (by maximum severity)
- TE treatment-related AEs of Grade 2 or higher (by maximum severity)
- All TE SAEs
- All TE treatment-related SAEs
- All TEAEs leading to premature discontinuation of study drug

- All TEAEs leading to premature discontinuation of study
- All TE SAEs leading to death (ie, outcome of death)
- All TEAEs leading to temporary interruption of study drug

A brief, high-level summary of AEs described above will be provided by treatment group presenting the number and percentage of subjects who reported the above AEs. All deaths observed in the study will be also included in this summary.

Multiple events will be counted only once per subject in each summary. Adverse events will be summarized and listed first in alphabetic order of SOC and HLT within each SOC (if applicable), and then by PT in descending order of total frequency within each SOC. For summaries by severity grade, the most severe grade will be used for those AEs that occurred more than once in an individual subject during the study.

In addition, the following tables will be generated and summarized by PT only, in descending order of total frequency:

- TEAEs
- TEAEs of Grade 3 or higher
- TEAEs of Grade 2 or higher
- TE SAEs
- TE treatment-related AEs
- TE treatment-related SAEs
- TEAEs leading to premature discontinuation of study drug

Data listings will be provided for the following:

- All AEs, indicating whether the event is treatment emergent
- All AEs of Grade 3 or higher
- All AEs of Grade 2 or higher
- SAEs
- Deaths
- All AEs leading to death (ie, outcome of death)

- AEs leading to premature discontinuation of study drug
- AEs leading to premature discontinuation of study
- AEs leading to temporary interruption of study drug

7.1.7. Adverse Events of Interest

Adverse events of interest (AEI) include infections, malignancies (excluding nonmelanoma skin cancers), nonmelanoma skin cancers, gastrointestinal perforations, major adverse cardiovascular events (MACE), and thromboembolic events. Summaries of the following treatment-emergent AEIs will be produced to enhance the analysis of safety data.

- Events of infections, presented in the following subcategories:
  - AEs of infections, utilizing all AEs in the MedDRA Infections and Infestations SOC
  - AEs of serious infections, using all AEs in the MedDRA Infections and Infestations SOC that are classified as SAEs
  - AEs of herpes zoster, utilizing a MedDRA search term (MST) list developed by Gilead
  - AEs of opportunistic infections (OIs), utilizing a narrow scope Standardised MedDRA Query (SMQ)
- AEs of malignancies, excluding nonmelanoma skin cancers, utilizing an MST list developed by Gilead
- AEs of nonmelanoma skin cancers, utilizing an MST list developed by Gilead
- AEs of gastrointestinal perforations, utilizing an MST list developed by Gilead
- AEs of MACE, utilizing a positively adjudicated event list, presented in the following subcategories (Section 7.1.7.1):
  - Cardiovascular (CV) death
  - Non-fatal myocardial infarction (MI)
  - Non-fatal stroke
- AEs of arterial systemic thromboembolism (ASTE), utilizing a positively adjudicated event list (Section 7.1.7.1)
- AEs of venous thromboembolism (VTE), utilizing a positively adjudicated event list (Section 7.1.7.1)

The number and percentage of subjects with a reported event will be summarized for each treatment group by PT for each AEI category. Data listings for AEIs will also be provided.

The number and percentage of subjects with positively adjudicated TE MACE, TE ASTE and TE VTE will be summarized by treatment group using the adjudicated category, if applicable.

A by-subject listing of subjects with potential events for adjudication (MACE, ASTE, and VTE) and their respective adjudication results will be provided.

A by-subject listing of thromboembolic history and risk factors will be provided for subjects with potential events for adjudication (MACE, ASTE, and VTE).

7.1.7.1. Cardiovascular Safety Endpoint Adjudication Committee

An independent cardiovascular safety endpoint adjudication committee (CVEAC) will be formed to periodically review and adjudicate all potential MACE and thromboembolic events in a blinded manner. To identify potential MACE and thromboembolic events, the following AEs will be sent for adjudication. Refer to the Cardiovascular Event Adjudication Committee Charter for more details.

- All AEs leading to death
- CV events (meeting seriousness criteria), utilizing an MST list developed by Gilead
- MI, utilizing a narrow scope SMQ
- Unstable angina (meeting hospitalization criteria), utilizing an MST list developed by Gilead
- Transient ischemic attack, utilizing an MST list developed by Gilead
- Stroke, utilizing an MST list developed by Gilead
- Cardiac failure (meeting hospitalization criteria), utilizing an MST list developed by Gilead
- Percutaneous coronary intervention, utilizing an MST list developed by Gilead
- Embolic and thrombotic events, utilizing a narrow scope SMQ

The CVEAC will review the above AEs, and related clinical data to adjudicate whether the criteria for MACE (CV death, MI, and/or stroke), ASTE, and VTE have been met for each AE.

7.2. Laboratory Evaluations

Laboratory data collected during the study will be analyzed and summarized using both quantitative and qualitative methods. Summaries of laboratory data will be provided for the Safety Analysis Set for data collected up to the last dose of study drug plus 30 days for subjects who have permanently discontinued study drug.

The analysis will be based on values reported in conventional units. When values are below the LOQ, they will be listed as such, and the closest imputed value will be used for the purpose of calculating summary statistics as specified in Section 3.6. Test results from hemolyzed samples will not be included in the analysis, but they will be listed in by-subject laboratory listings.

A by-subject listing for laboratory test results will be provided by subject ID number and time point in chronological order for hematology, serum chemistry, urinalysis, lipid profile, and serum immunoglobulin separately. Values falling outside of the relevant reference range and/or having a severity grade of 1 or higher on the Common Terminology Criteria for Adverse Events (CTCAE) severity grade will be flagged in the data listings, as appropriate.

Hematocrit collected from local laboratories due to COVID-19 impact will not be included in the safety laboratory summary, but they will be included in the listing with a flag to indicate which records were collected at a local laboratory.

No formal statistical testing is planned.

7.2.1. Summaries of Numeric Laboratory Results

Descriptive statistics will be provided by treatment group for each laboratory test specified in the study protocol within hematology and chemistry panels, and also laboratory tests from lipids panel (under fasting status) including total cholesterol, low-density lipoprotein (LDL), high-density lipoprotein (HDL), triglycerides, non-HDL cholesterol (total cholesterol minus HDL cholesterol), LDL/HDL ratio, and IgA, IgM, IgG, and total Ig, as follows:

- Baseline values
- Values at each postbaseline time point
- Change from baseline at each postbaseline time point

A baseline laboratory value will be defined as the last measurement obtained on or prior to the date of first dose of study drug. Change from baseline to a postbaseline visit will be defined as the visit value minus the baseline value. The mean, median, Q1, Q3, minimum, and maximum values will be displayed to the reported number of digits; SD values will be displayed to the reported number of digits plus 1.

Median (Q1, Q3) of the observed change from baseline values for aspartate aminotransferase (AST), alanine aminotransferase (ALT), total bilirubin, alkaline phosphatase, serum creatinine, creatinine clearance, creatine phosphokinase, white blood cell count, absolute neutrophils, absolute lymphocytes, hemoglobin, platelets, fasting total cholesterol, fasting LDL, fasting HDL, total Ig, IgG, IgA, and IgM, will be plotted using a line plot by treatment group and analysis visit.

In the case of multiple values in an analysis window, data will be selected for analysis as described in Section 3.7.3.

7.2.2. Graded Laboratory Values

Modified CTCAE Version 4.03 (see Appendix 8 of the protocol) will be used for assigning toxicity grades (0 to 4) to laboratory results for analysis. Grade 0 includes all values that do not meet the criteria for an abnormality of at least Grade 1. For laboratory tests with criteria for both increased and decreased levels, analyses for each direction (ie, increased, decreased) will be presented separately.

7.2.2.1. Treatment-Emergent Laboratory Abnormalities

Treatment-emergent laboratory abnormalities are defined as values that increase at least 1 toxicity grade from baseline at any postbaseline time point, up to and including the date of last dose of study drug plus 30 days for subjects who permanently discontinued study drug.

If the relevant baseline laboratory value is missing, any abnormality of at least Grade 1 observed within the time frame specified above will be considered treatment emergent.

7.2.2.2. Treatment-Emergent Marked Laboratory Abnormalities

Treatment-emergent marked laboratory abnormalities are defined as values that increase at least 3 toxicity grades from baseline at any postbaseline time point, up to and including the date of the last dose of study drug plus 30 days for subjects who permanently discontinued study drug.

If the relevant baseline laboratory value is missing, any Grade 3 or 4 values observed within the timeframe specified above will be considered treatment-emergent marked abnormalities.

7.2.2.3. Summaries of Laboratory Abnormalities

Laboratory data that are categorical will be summarized using the number and percentage of subjects in the study with the given response at baseline and each scheduled postbaseline time point.

The following summaries (number and percentage of subjects) for treatment-emergent laboratory abnormalities will be provided by lab test and treatment group; subjects will be categorized according to the most severe postbaseline abnormality grade for a given laboratory test:

- Graded laboratory abnormalities
- Grade 3 or higher laboratory abnormalities
- Marked laboratory abnormalities

For all summaries of laboratory abnormalities, the denominator is the number of subjects with nonmissing postbaseline values up to 30 days after last dosing date.

A by-subject listing of treatment-emergent Grade 3 or 4 laboratory abnormalities will be provided by subject ID number and time point in chronological order. This listing will include all test results that were collected throughout the study for the lab test of interest, with all applicable severity grades displayed.

7.2.3. Liver-related Laboratory Evaluations

Liver-related abnormalities after initial study drug dosing will be examined and summarized using the number and percentage of subjects who were reported to have the following laboratory test values for postbaseline measurements:

- AST: (a) > 3 times of upper limit of the normal range (ULN); (b) > 5 x ULN; (c) > 10 x ULN; (d) > 20 x ULN
- ALT: (a) > 3 x ULN; (b) > 5 x ULN; (c) > 10 x ULN; (d) > 20 x ULN
- AST or ALT > 3 x ULN and total bilirubin > 2 x ULN

For individual laboratory tests, subjects will be counted once based on the most severe postbaseline values. For both the composite endpoint of AST or ALT and total bilirubin, subjects will be counted once when the criteria are met at the same postbaseline visit date. The denominator is the number of subjects in the Safety Analysis Set who have nonmissing postbaseline values of all relevant tests at the same postbaseline visit date. A listing of subjects who met at least 1 of the above criteria will be provided.

7.3. Body Weight, Height and Vital Signs

Descriptive statistics will be provided by treatment group for body weight, BMI and vital signs (resting blood pressure [systolic blood pressure and diastolic blood pressure], respiratory rate, pulse, and temperature) as follows:

- Baseline value
- Values at each postbaseline time point
- Change from baseline at each postbaseline time point

A baseline value will be defined as the last available value collected on or prior to the date/time of first dose of study drug. Change from baseline to a postbaseline visit will be defined as the postbaseline value minus the baseline value. Body weight and vital signs measured at unscheduled visits will be included for the baseline value selection.

In the case of multiple values in an analysis window, data will be selected for analysis as described in Section 3.7.3. No formal statistical testing is planned.

A by-subject listing of vital signs will be provided by subject ID number and time point in chronological order. Body weight, height, and BMI will be included in the vital signs listing, if space permits. If not, they will be provided separately.

Weight measured by the subjects at home from virtual visits due to COVID-19 impact will not be included in the vital signs summary, but they will be included in the listing with a flag to indicate which records were collected at home.

7.4. Prior and Concomitant Medications

Medications reported at screening and during the study will be coded using the current version of the World Health Organization (WHO) Drug dictionary.

7.4.1. Prior Medications

Prior medications are defined as any medications taken before a subject took the first dose of study drug at baseline.

General prior and CD-specific prior medication will be summarized separately by preferred name using the number and percentage of subjects for each treatment group and overall. A subject reporting the same medication more than once will be counted only once when calculating the number and percentage of subjects who received that medication. The summary will be ordered by preferred term in descending overall frequency. For drugs with the same frequency, sorting will be done alphabetically.

For the purposes of analysis, any medication with a start date prior to the first dosing date of study drug will be included in the prior medication summary regardless of when the stop date is. If a partial start date is entered the medication will be considered prior unless the month and year (if day is missing) or year (if day and month are missing) of the start date are after the first dosing date. Medications with a completely missing start date will be included in the prior medication summary, unless otherwise specified.

Summaries will be based on the Safety Analysis Set. No formal statistical testing is planned.

7.4.2. Concomitant Medications

Concomitant medications are defined as medications taken while a subject took study drug. General and CD-specific concomitant medications will be summarized, separately, by preferred name using the number and percentage of subjects for each treatment group. A subject reporting the same medication more than once will be counted only once when calculating the number and percentage of subjects who received that medication. The summary will be ordered by preferred term in descending overall frequency. For drugs with the same frequency, sorting will be done alphabetically.

For the purposes of analysis, any medication with a start date prior to or on the first dosing date of study drug and continued to take after the first dosing date, or started after the first dosing date but prior to or on the last dosing date of study drug will be considered concomitant medications. Medications started and stopped on the same day as the first dosing date or the last dosing date of study drug will also be considered concomitant. Medications with a stop date prior to the date of first dosing date of study drug or a start date after the last dosing date of study drug will be excluded from the concomitant medication summary. If a partial stop date is entered, any medication with the month and year (if day is missing) or year (if day and month are missing) prior to the date of first study drug administration will be excluded from the concomitant medication summary. If a partial start date is entered, any medication with the month and year (if

day is missing) or year (if day and month are missing) after the study drug stop date will be excluded from the concomitant medication summary. Medications with completely missing start and stop dates will be included in the concomitant medication summary, unless otherwise specified. Summaries will be based on the Safety Analysis Set. No formal statistical testing is planned.

All prior and concomitant medications (other than per-protocol study drugs) will be provided in a by-subject listing sorted by subject ID number and administration date in chronological order.

7.5. Electrocardiogram Results

A shift table of the investigators' assessment of ECG results at each visit compared with baseline values will be presented by treatment group using the following categories: normal; abnormal, not clinically significant; abnormal, clinically significant. The number and percentage of subjects in each cross-classification group of the shift table will be presented. Subjects with a missing value at baseline or postbaseline will not be included in the denominator for percentage calculation. No formal statistical testing is planned.

A by-subject listing for ECG assessment results will be provided by subject ID number and time point in chronological order.

7.6. Other Safety Measures

A data listing will be provided for subjects who become pregnant during the study.

7.7. Changes from Protocol-Specified Safety Analyses

There are no deviations from the protocol-specified safety analyses.

[redacted]

CCI

[redacted]

CCI [redacted]

10. REFERENCES

Daperno M, D'Haens G, Van Assche G, Baert F, Bulois P, Maunoury V, et al. Development and validation of a new, simplified endoscopic activity score for Crohn's disease: the SES-CD. Gastrointestinal endoscopy 2004;60 (4):505-12.

Feagan BG, Schwartz D, Danese S, Rubin DT, Lissoos TW, Xu J, et al. Efficacy of Vedolizumab in Fistulising Crohn's Disease: Exploratory Analyses of Data from GEMINI 2. Journal of Crohn's & colitis 2018;12 (5):621-6.

Irvine EJ. Usual therapy improves perianal Crohn's disease as measured by a new disease activity index. McMaster IBD Study Group. J Clin Gastroenterol 1995;20 (1):27-32.

Khanna R, Zou G, D'Haens G, Feagan BG, Sandborn WJ, Vandervoort MK, et al. A retrospective analysis: the development of patient reported outcome measures for the assessment of Crohn's disease activity. Aliment Pharmacol Ther 2015;41 (1):77-86.

Sandborn WJ, Feagan BG, Hanauer SB, Lochs H, Lofberg R, Modigliani R, et al. A review of activity indices and efficacy endpoints for clinical trials of medical therapy in adults with Crohn's disease. Gastroenterology 2002;122 (2):512-30.

Van Assche G, D'Haens G, Noman M, Vermeire S, Hiele M, Asnong K, et al. Randomized, double-blind comparison of 4 mg/kg versus 2 mg/kg intravenous cyclosporine in severe ulcerative colitis. Gastroenterology 2003;125 (4):1025-31.

11. SOFTWARE

SAS® Software Version 9.4 (SAS Institute Inc., Cary, NC, USA) is to be used for all programming of tables, listings, and figures.

12. SAP REVISION

Revision Date (DD MMM YYYY)	Section	Summary of Revision	Reason for Revision
Original Version (05 NOV 2019)			
Version 2 (01 March 2021)	3.5.1	Added missing data imputation rule for MRI data, laboratory data and vital signs data	To clarify missing data imputation rule for analysis
	3.7.2	Updated analysis visit window length for Week 18 and Week 24	To map data to analysis visits based on more clinical relavent rationale
	3.8, 4, 7.2, 7.3	Added assessment of COVID-19 impact	To better understand the COVID-19 impact on subject disposition and missing data
	5	Modified the list of variables to be summarized in baseline characteristic tables	To provide better characterization of the investigated patient populations
	[redacted]	CCI [redacted]	[redacted]
	7.1.7	Added adverse event of interest for MACE, ASTE, and VTE based on positively adjudicated events by the cardiovascular safety endpoint adjudication committee	To better address safety concerns
	7.2.4	Removed shift table summaries	The shift table summaries are less useful to detect safety signal given small sample size in this study.The shift table summaries will be included in integrated safety analysis if needed
	Appendix 3	Added a list of biologics for CD treatment	To clarify biologics considered as general CD treatment
	Appendix	Deleted the list of preferred terms for each category of adverse events of interest (Version 1 Appendix 11 to 16)	To reduce the burden to keep track of the list of preferred terms due to MedDRA up-versioning. The lists will be provided in the CSR.
	Global	Administrative and editorial changes have been made throughout the SAP, where appropriate	Imporve clarity and consistency

13. APPENDICES

Appendix 1. Schedule of Assessments

Period	Screening	Treatment									Follow-Up	
Visit	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**PTx[a]**	**ET**
Week		**0**	**2**	**4**	**6**	**10**	**14**	**18**	**20**	**24**		
Study Day	**-30 to -1**	**1**	**15**	**29**	**43**	**71**	**99**	**127**	**141**	**169**		
Visit Window (±)			**±3**	**±3**	**±3**	**±2**	**±3**	**±3**	**±5**	**±5**	**±3**	
Written Informed Consent	X											
Medical History & Demographics	X											
Crohn's Disease & Treatment History	X											
12-lead ECG	X					X						X[b]
Review of Inclusion/ Exclusion Criteria	X	X										
Physical Exam (Complete including perianal)[c]	X											
Physical Exam (symptom-based) and perianal assessment[c]		X	X	X	X	X	X	X		X	X	X
Vital Signs	X	X	X	X	X	X	X	X		X	X	X
Weight	X	X	X	X	X	X	X	X		X	X	X
Height	X											
Adverse Events	X	X	X	X	X	X	X	X		X	X	X
Concomitant Medications	X	X	X	X	X	X	X	X		X	X	X

Period	Screening	Treatment									Follow-Up	
Visit	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**PTx[a]**	**ET**
Week		**0**	**2**	**4**	**6**	**10**	**14**	**18**	**20**	**24**		
Study Day	**-30 to -1**	**1**	**15**	**29**	**43**	**71**	**99**	**127**	**141**	**169**		
Visit Window (±)			**±3**	**±3**	**±3**	**±2**	**±3**	**±3**	**±5**	**±5**	**±3**	
Randomization		X										
Study Drug Dispensing		X		X		X	X	X				
CCI												
eDiary instruction & review[e]	X	X	X	X	X	X	X	X		X		
Flexible Sigmoidoscopy[o]	X									X		X[r]
MRI[f]	X									X		X[s]
Stool for *C. diff* toxin, pathogenic *E. coli*, *Salmonella*, *Shigella*, *Campylobacter* and *Yersinia* testing	X											
Stool O&P	X											
CCI												
CCI												

Period	Screening	Treatment									Follow-Up	
Visit	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**PTx[a]**	**ET**
Week		**0**	**2**	**4**	**6**	**10**	**14**	**18**	**20**	**24**		
Study Day	**-30 to -1**	**1**	**15**	**29**	**43**	**71**	**99**	**127**	**141**	**169**		
Visit Window (±)			**±3**	**±3**	**±3**	**±2**	**±3**	**±3**	**±5**	**±5**	**±3**	
Urine drug screen[g]	X											
Urinalysis[q]	X									X		
Pregnancy Test[h]	X	X		X	X	X	X	X	X	X	X	X
TB screening[i]	X											
Chest x-ray[q]	X											
HBV, HCV, HIV screening[j]	X											
Hematology	X	X	X	X	X	X	X	X		X	X	X
Chemistry	X	X	X	X	X	X	X	X		X	X	X
Fasting Lipids[p]		X				X				X		
CRP	X	X	X	X	X	X	X	X		X	X	X
Blood TCR/BCR repertoire sample[k]		X		X		X				X		
CCI												
CCI												
Blood transcriptome sample		X		X		X				X		
CCI												
Serum immunoglobulin		X		X		X				X	X	X

Period	Screening	Treatment									Follow-Up	
Visit	1	2	3	4	5	6	7	8	9	10	PTx[a]	ET
Week		0	2	4	6	10	14	18	20	24		
Study Day	-30 to -1	1	15	29	43	71	99	127	141	169		
Visit Window (±)			±3	±3	±3	±2	±3	±3	±5	±5	±3	
CCI												
CCI												
CCI												

a The Post-Treatment (PTx) visit should occur 30 days after the last dose of study drug. Only subjects who roll over into the LTE study will not complete PTx assessments.
b For subjects who terminate prior to Week 10
c A complete physical examination (PE) including; vital signs, body weight, height, and perianal assessment (including completion of the Perianal Fistula Assessment Worksheet) will be performed at screening. A symptom-directed PE and perianal assessment(including completion of the Perianal Fistula Assessment Worksheet) should be performed at all other time points.
CCI
e Subjects should begin filling out the eDiary the day of their initial screening visit and continue to fill it out throughout the remainder of the study.
f Refer to Section 6.6 of study protocol for details on the MRI procedure
g Positive cocaine test disqualifies subject; positive amphetamines, barbiturates, benzodiazepines, and opioids require medical monitor review
h All females meeting the childbearing potential criteria must have a serum pregnancy testing at screening and a urine pregnancy test must be completed every 4 weeks at a minimum. If any pregnancy test is positive, study drug must be immediately interrupted and the subject should come to the site for serum pregnancy test in clinic.
i Proof of no active or untreated latent TB at screening. Subjects who are diagnosed with latent TB at screening must initiate an adequate course of prophylaxis as per local standard of care, for a minimum of 4 weeks prior to randomization. Subject may initiate study drug dosing only after consultation with the Study Medical Monitor.
j An HIV-1/HIV-2 antibody test, a HCV antibody test, a HBV surface antigen test, a HBV surface antibody test, and a HBV core antibody test will be performed on all subjects; subjects with any positivity will be excluded unless the Hepatitis B surface antibody is the sole positive result (indicating immunization)
k TCR: T-cell receptor; BCR: B-cell receptor
CCI
CCI
o Refer to Section 6.5 of study protocol for details on the flexible sigmoidoscopy procedure
p Subjects should fast (no food or drinks, except water) for at least 8 hours prior to blood sample collection at Day 1, Week 10, and Week 24
q Chest x-ray (views as per local guidelines) taken at screening or within the 3 months prior to screening (with the report or films available for investigator review) without evidence of active or latent TB infection
r Subjects meeting non-response or disease worsening criteria should undergo a flexible sigmoidoscopy assessment prior to exiting the study.
s Subjects meeting non-response or disease worsening criteria should undergo a pelvic MRI prior to exiting the study.

Appendix 2. List of Biologics for Perianal Fistulizing CD Treatment

The drug names of biologics considered as treatment for perianal fistulizing CD are listed below.

No.	Drug Class	Drug Name
1	TNFα antagonist	Adalimumab
2	TNFα antagonist	Certolizumab/Certolizumab Pegol
3	TNFα antagonist	Infliximab
4	TNFα antagonist	TNFα antagonist biosimilar (to adalimumab, Certolizumab/Certolizumab Pegol, or infliximab)

Appendix 3. List of Biologics for CD Treatment

The drug names of biologics considered as CD treatment are listed below.

No.	Drug Class	Drug Name
1	TNFα antagonist	Adalimumab
2	TNFα antagonist	Certolizumab/Certolizumab Pegol
3	TNFα antagonist	Infliximab
4	TNFα antagonist	TNFα antagonist biosimilar (to adalimumab, Certolizumab/Certolizumab Pegol, or infliximab)
5	Integrin antagonist	Natalizumab
6	Integrin antagonist	Vedolizumab
7	Interleukin antagonist	Ustekinumab

CCI

CCI [redacted]

[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]

[redacted]

CCI

CCI

CCI

[redacted]

CCI [redacted]

[redacted]

CCI [redacted]

[redacted]

CCI [redacted]

Appendix 5. Calculation of Standard Body Weight

The following table will be used to determine the standard body weight for each subject. Height from the eCRF will be used and is assumed to be measured without shoes.

Women			
Height in *cm* (in shoes)*	**Standard Weight in *kg***	**Height in *cm* (in shoes)***	**Standard Weight in *kg***
127	41.2	163	60.2
128	41.7	164	60.7
129	42.3	165	61.3
130	42.8	166	61.9
131	43.3	167	62.4
132	43.8	168	62.9
133	44.4	169	63.4
134	44.9	170	63.9
135	45.4	171	64.5
136	45.9	172	65.0
137	46.4	173	65.5
138	47.0	174	66.0
139	47.5	175	66.6
140	48.0	176	67.2
141	48.6	177	67.7
142	49.1	178	68.3
143	49.6	179	68.8
144	50.2	180	69.3
145	50.7	181	69.8
146	51.2	182	70.3
147	51.8	183	70.8
148	52.3	184	71.3
149	52.8	185	71.8
150	53.1	186	72.3
151	54.1	187	72.8
152	54.5	188	73.3
153	55.0	189	73.8
154	55.4	190	74.3
155	55.9	191	74.9
156	56.4	192	75.4
157	57.0	193	76.0
158	57.5	194	76.5
159	58.1	195	77.0
160	58.6	196	77.6
161	59.1	197	78.1
162	59.6	198	78.6

* add 2.0 cm if shoeless. Please round height to a whole number and select the appropriate standard weight (e.g., height 157.4 cm will be rounded to 157 cm; 157.5 cm will be rounded to 158 cm.)

Men			
Height in *cm* (in shoes)*	**Standard Weight in *kg***	**Height in *cm* (in shoes)***	**Standard Weight in *kg***
142	54.4	179	71.9
143	54.9	180	72.4
144	55.4	181	73.0
145	55.8	182	73.6
146	56.3	183	74.3
147	56.8	184	74.8
148	57.2	185	75.5
149	57.7	186	76.2
150	58.2	187	76.9
151	58.6	188	77.6
152	59.1	189	78.2
153	59.6	190	78.8
154	60.0	191	79.6
155	60.5	192	80.4
156	61.0	193	81.2
157	61.4	194	82.1
158	61.9	195	82.9
159	62.2	196	83.8
160	62.6	197	84.7
161	62.9	198	85.4
162	63.3	199	86.1
163	63.7	200	86.7
164	64.1	201	87.4
165	64.6	202	88.0
166	65.0	203	88.8
167	65.5	204	89.4
168	66.0	205	90.1
169	66.6	206	90.7
170	67.1	207	91.4
171	67.6	208	92.1
172	68.1	209	92.7
173	68.7	210	93.4
174	69.2	211	94.1
175	69.7	212	94.8
176	70.3	213	95.5
177	70.8	214	96.1
178	71.3		

* add 2.0 cm if shoeless. Please round height to a whole number and select the appropriate standard weight (e.g., height 157.4 cm will be rounded to 157 cm; 157.5 cm will be rounded to 158 cm.)

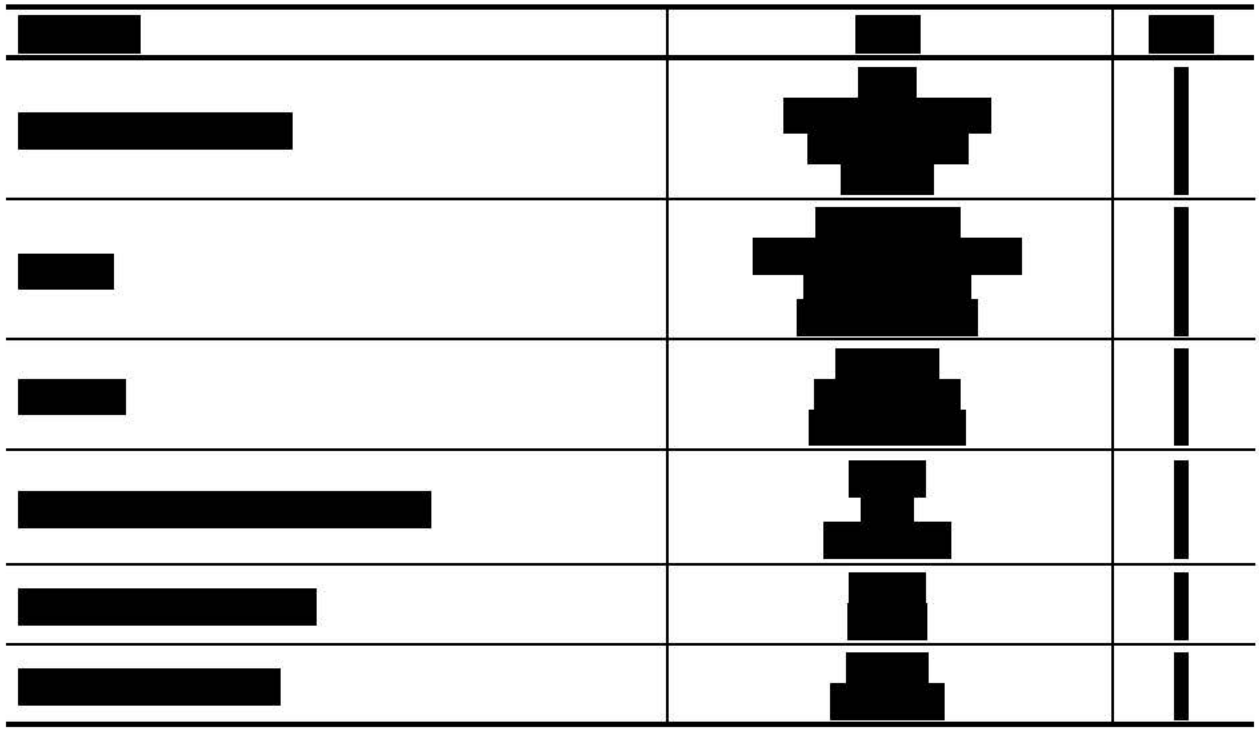
Appendix 6. Simple Endoscopic Score for Crohn's Disease (SES-CD)

Variables	Score			
	0	**1**	**2**	**3**
Size of ulcers (cm)	None	Aphthous ulcers (diameter 0.1-0.5)	Large ulcers (diameter >0.5-2)	Very large ulcers (diameter > 2)
Ulcerated Surface (%)	None	< 10	10-30	> 30
Affected surface (%)	Unaffected segment	< 50	50-75	> 75
Presence of narrowings	None	Single, can be passed	Multiple, can be passed	Cannot be passed

Total SES-CD: sum of the values of the 4 variables for the 5 bowel segments. Values are given to each variable and for every examined bowel segment (eg, rectum, left colon, transverse colon, right colon, and ileum); Total SES-CD will not be calculated for Study GS-US-419-4016.
{Daperno 2004}
For Study GS-US-419-4016, the Proctitis SES-CD score (which includes an overall assessment of both rectum and anal canal) and Anal Canal Segmental SES-CD score will **only** score **size of ulcers** and **ulcerated surface**. The Proctitis SES-CD score will be used to provide an endoscopic measure of proctitis. CCI [redacted]

CCI [redacted]

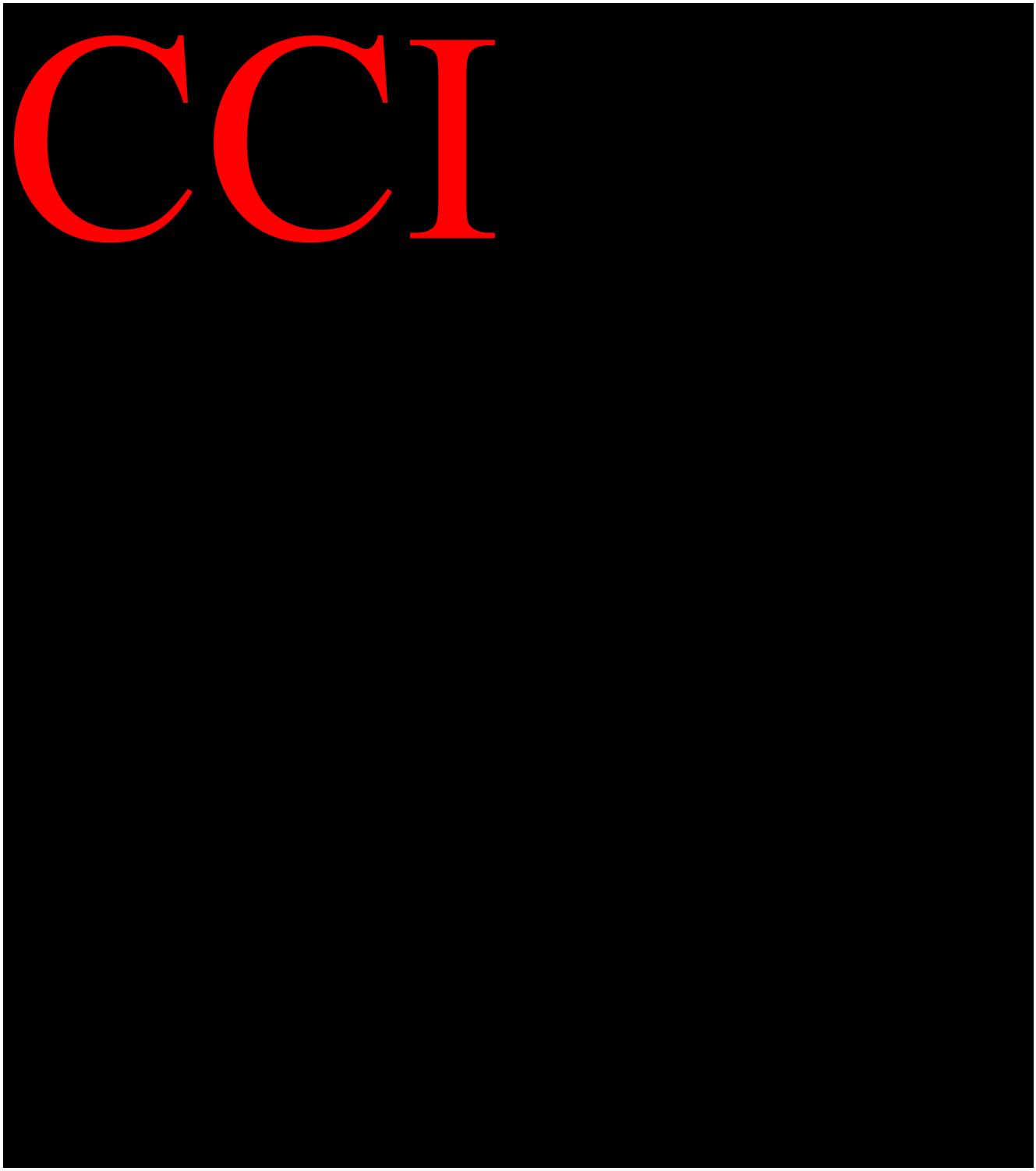
[redacted]



[redacted]

CCI

CCI


CCI

Appendix 10. Study Treatment Failure Rules for Luminal Disease Activity Related Efficacy Endpoints

Study treatment failure rules described in this appendix apply to treatment failure that occurs during the course of the study, and will be applied to all the following efficacy endpoints related to luminal disease activity:

- [redacted]
- [redacted]
- [redacted]
- Proportion of subjects achieving proctitis remission at Week 24, in subjects that had moderately to severely active proctitis at baseline

All related efficacy data will be censored (set to missing) after treatment failure criteria are met, regardless of the observed data. Subjects who do not have sufficient measurements after censoring to determine the dichotomized endpoint(s) will be considered non-responder for corresponding endpoint (s). Study treatment failure rules override other data handling rules.

Subjects who have any of the following events will be considered a study treatment failure after the earliest event through the end of the study, regardless of the actual efficacy data collected.

I. Potentially effective corticosteroid use

Potentially effective corticosteroids, for the purpose of this SAP, are corticosteroids that may impact the disease under study. **Potentially effective corticosteroids** include the following corticosteroids when administered for the indication of CD:

1) Commencement of:

   Any steroid administered intramuscularly (IM), intravenously (IV), orally (PO), or rectally (PR) at any dose for 7 or more continuous days

   - This rule applies regardless if a change in drug, dose, or route (from IM to IV to PO or PR or vice versa) occurs within the seven continuous days; and it includes oral steroids with intended local actions (eg, budesonide)

2) Escalation of concomitant steroid dose above the baseline dose for 7 or more continuous days. The baseline steroid dose is defined as the dose at Day 1 of the study. Prednisone equivalent dose will be used to determine escalation of concomitant systemic steroid dose even if there is a change in drug or route. For steroids with predominant local effect (eg, budesonide or any rectally administered steroid), this rule will apply to scenarios where the post-baseline dose is above the baseline dose via the same drug and the same route for 7 or more continuous days.

Potentially effective steroid use only includes use of steroids administered via routes that are IM, PO, IV, or PR. The below steroids will **not** be considered potentially effective steroids regardless of indication for use:

A) ocular steroids (ie, eye drops)

B) topical steroids (eg, cutaneously applied solely to the skin, or topically applied to the nasal mucosa)

C) inhaled steroids (eg, inhalational fluticasone for asthma)

D) intra-articular steroids (steroids administered directly into a joint)

E) neuraxial steroids (steroids administered into the epidural or spinal space)

II. Potentially effective immunomodulator use

Commencement of a different class of oral, IM, SC, or IV immunomodulator drugs (where the subject was not previously taking concomitant immunomodulators of the same class on Day 1), including but not limited to 6-MP, azathioprine, MTX, 6-thioguanine, and prohibited immunomodulators including but not limited to cyclosporine, leflunomide, tacrolimus, thalidomide regardless of dose, for 7 continuous days. The use of an immunomodulator will be considered potentially effective when it is administered for the indication of CD.

III. Potentially effective biologic agent use

Commencement of a biologic agent including but not limited to TNFα antagonists, IL-12/23 antagonists, and vedolizumab (or similar agents), regardless of indication for use and duration (Appendix 3).

Appendix 11. Study Treatment Failure Rules for Perianal Fistulizing CD Activity Related Efficacy Endpoints

Study treatment failure rules described in this appendix apply to treatment failure that occurs during the course of the study, and will be applied to all the following efficacy endpoints related to perianal fistulizing CD activity:

- Proportion of subjects establishing combined fistula response at Week 24
- Proportion of subjects establishing combined fistula remission at Week 24
- Time to clinical fistula response
- Time to clinical fistula remission

[redacted]

All related efficacy data will be censored (set to missing) after treatment failure criteria are met, regardless of the observed data. Subjects who do not have sufficient measurements after censoring to determine the dichotomized endpoint(s) will be considered non-responder for corresponding endpoint (s). Study treatment failure rules override other data handling rules.

Subjects who have any of the following events will be considered a study treatment failure after the earliest event through the end of the study, regardless of the actual efficacy data collected.

I. Potentially effective antibiotics use

Potentially effective antibiotics, for the purpose of this SAP, include metronidazole and ciprofloxacin, which may impact the subject's perianal fistulizing CD when administered as following for the indication of perianal fistulizing CD or CD:

- Commencement of a new antibiotics use after the first dosing date of study drug for 7 or more days;
- Escalation of the dose of antibiotics above the baseline dose for 7 or more days.

II. Potentially effective immunomodulator use

Commencement of a different class of oral, IM, SC, or IV immunomodulator drugs (where the subject was not previously taking concomitant immunomodulators of the same class on Day 1), including but not limited to 6-MP, azathioprine, MTX, 6-thioguanine, and prohibited immunomodulators including but not limited to cyclosporine, leflunomide, tacrolimus, thalidomide regardless of dose, for 7 continuous days. The use of an immunomodulator will be considered potentially effective when it is administered for the indication of perianal fistulizing CD or CD.

III. Potentially effective biologic agent use

Commencement of a biologic agent including TNFα antagonists, regardless of indication for use and duration (Appendix 2).

Appendix 12. Determining Missing and Virtual Visits due to COVID-19

This appendix describes the clinical trial site collection of COVID-19 data pertaining to missed/virtual visits and the data processing algorithm used to determine which visits were missing and which visits were virtual.

Data collection

A COVID-19 supplement to the eCRF Completion Guidelines was provided by data management to instruct clinical trial sites with respect to data entry expectations pertaining to scenarios related to the COVID-19 pandemic. If a visit was missed, sites should enter "Visit missed due to COVID-19." If an in-person visit was conducted virtually, sites should enter "Virtual visit due to COVID-19."

Determination of Missed and Virtual visits

Natural Language Processing (NLP) was used to search the CRF comment fields to identify instances of "COVID-19" (or synonyms, see Table 13-6) and "Virtual" (or synonyms, see Table 13-6). The search terms are maintained in a global lookup table and can be modified to tune the NLP model. For any comments with COVID-19 search terms, assign "Missed Visit" or "Virtual Visit" as follows:

i) If COVID-19 terms are identified through NLP and the visit date is missing, then result is "Missed Visit"

ii) If COVID-19 and Virtual terms are identified through NLP for a visit, then result is "Virtual Visit". When there are multiple records for the same subject and the same visit, NLP will be based on multiple records to ensure 1 unique category per subject per visit

iii) Otherwise result is missing

Table 13-6. Examples of Search Terms for "COVID-19" and "Virtual" Used to Identify Missed and Virtual Visits

Search terms for "COVID-19"	Search terms for "Virtual"
COVID19	VIRTUAL
CORONA	TELEMED
CORONAVIRUS	TELEHEALTH
PANDEMIC	TELEPHONE
OUTBREAK	REMOTE
CRISIS	TELEMEDICINE
LOCKDOWN	TELECONSULTATION
QUARANTINE	TELEPHONICALLY
SHELTER	PHONE
	HOME VISIT
	ZOOM
	SKYPE

GS-US-419-4016 SAP v2.0

ELECTRONIC SIGNATURES

Signed by	Meaning of Signature	Server Date (dd-MMM-yyyy hh:mm:ss)
PPD	Biostatistics eSigned	01-Mar-2021 17:42:14
PPD	Clinical Research eSigned	01-Mar-2021 20:21:43