STATISTICAL ANALYSIS PLAN

Study Title:	A Phase 2, Randomized, Placebo-Controlled Trial Evaluating the Efficacy and Safety of Filgotinib in Subjects with Active Noninfectious Uveitis	
Name of Test Drug:	Filgotinib	
Study Number:	GS-US-432-4097	
Protocol Version (Date):	Amendment 4:	17 March 2020
Analysis Type:	Final Analysis	
Analysis Plan Version:	Version 1.0	
Analysis Plan Date:	06 October 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

AC	Anterior chamber
AE	adverse event
AESIs	adverse events of special interest
ALT	alanine aminotransferase
ANCOVA	analysis of covariance
AST	aspartate aminotransferase
ATC	Anatomical Therapeutic Chemical drug class
BCVA	best corrected visual acuity
BLQ	below the limit of quantitation
BMI	body mass index
CI	confidence interval
CMH	Cochran-Mantel-Haenszel
CRF	case report form
CSR	clinical study report
CTCAE	Common Terminology Criteria for Adverse Events
DMC	data monitoring committee
ECG	electrocardiogram
eCRF	electroni case report form
CCI	[redacted]
ET	early termination
FAS	full analysis set
FDA	Food and Drug Administration
Gilead	Gilead Sciences, Inc.
HLGT	high-level group term
HLT	high-level term
ID	identification
IWRS	interactive web response system
LLOQ	lower limit of quantitation
LLT	lower-level term
logMAR	logarithm of the minimal angle of resolution
LOQ	limit of quantitation
MedDRA	Medical Dictionary for Regulatory Activities
NEI/SUN	National Eye Institute/Standardization of Uveitis Nomenclature
OCT	optical coherence tomography
PD	pharmacodynamics
PK	pharmacokinetic
PT	preferred term
PTM	Placebo to match

Q1, Q3	first quartile, third quartile
SAE	serious adverse event
SAP	statistical analysis plan
SD	standard deviation
SE	standard error
CCI	[redacted]
SMQ	standardised MedDRA queries
SOC	system organ class
TEAE	treatment-emergent adverse event
TFLs	tables, figures, and listings
ULN	upper limit of normal
US	United States
CCI	[redacted]
VH	vitreous haze
WHO	World Health Organization
CCI	[redacted]

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-432-4097. This SAP is based on the study protocol dated 17 March 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 as follows:

- To evaluate the efficacy of filgotinib versus placebo for the treatment of the signs and symptoms of noninfectious uveitis as measured by the proportion of subjects failing treatment for active noninfectious uveitis by Week 24

The secondary objectives of this study are as follows:

- To evaluate the efficacy of filgotinib versus placebo for the treatment of noninfectious uveitis as measured by the time to treatment failure on or after Week 6
- To evaluate the efficacy of filgotinib versus placebo for the treatment of noninfectious uveitis as measured by the change in vitreous haze (VH) grade in each eye (National Eye Institute/Standardization of Uveitis Nomenclature [NEI/SUN] criteria), from best state achieved prior to Week 6 to Week 52/ End of Treatment (EOT) visit or Early Termination (ET) visit
- To evaluate the efficacy of filgotinib versus placebo for the treatment of noninfectious uveitis as measured by the change in anterior chamber (AC) cell grade in each eye, from best state achieved prior to Week 6 to Week 52/ EOT visit or ET visit
- To evaluate the effects of filgotinib versus placebo as measured by the change in logarithm of the minimal angle of resolution (logMAR) best corrected visual acuity (BCVA) in each eye, from best state achieved prior to Week 6 to Week 52/EOT visit or ET visit
- To evaluate the efficacy of filgotinib versus placebo for the treatment of noninfectious uveitis as measured by the change in central retinal thickness in each eye, from best state achieved prior to Week 6 to Week 52/ EOT visit or ET visit
- To evaluate the efficacy of filgotinib versus placebo for the treatment of noninfectious uveitis as measured by the time to development of macular edema in at least one eye as determined by optical coherence tomography (OCT) on or after Week 6
- To evaluate the safety and tolerability of filgotinib
- To evaluate the pharmacokinetics of filgotinib and its metabolite GS-829845

CCI [redacted]

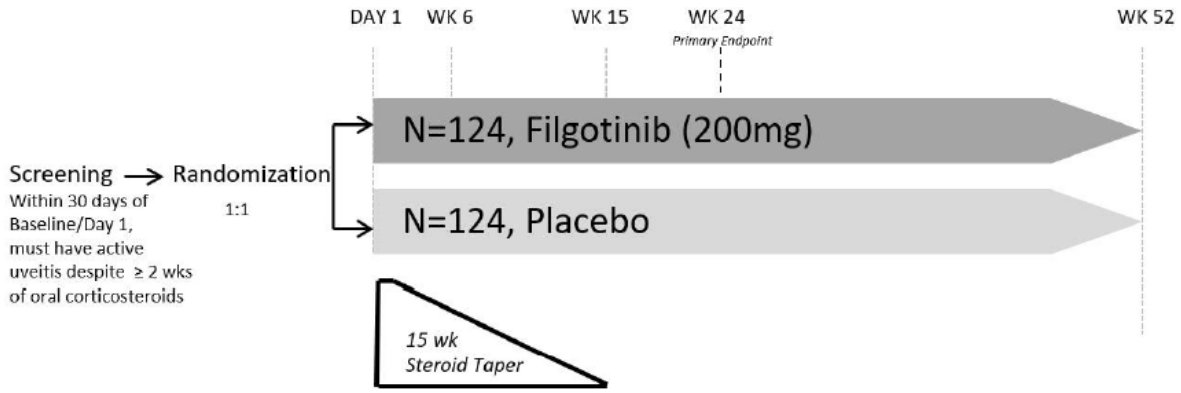
1.2. Study Design

This is a randomized, double-masked, placebo-controlled Phase 2 trial in adult subjects with active noninfectious intermediate-, posterior-, or pan-uveitis in at least one eye despite at least 2 weeks of maintenance therapy with oral prednisone at a dose of ≥ 10 mg/day to ≤ 60 mg/day (or oral corticosteroid equivalent, conversion chart is in Appendix 1). The study is designed to demonstrate the efficacy and safety of filgotinib in adult subjects requiring high dose steroids for active noninfectious intermediate-, posterior-, or pan-uveitis. Both of the subject's eyes as applicable should be evaluated for the purpose of determining eligibility based on inclusion and exclusion criteria, and for the purpose of assessing treatment failure. There will be no designated "study eye".

Subjects will be randomized in a 1:1 ratio to receive filgotinib 200 mg once daily or placebo to match (PTM) for up to 52 weeks, as shown in Figure 1.1. Subjects that meet criteria for treatment failure or discontinue for any other reason prior to the end of the study will have an EOT or ET visit (as appropriate).

It is anticipated that approximately 25% of subjects will have their uveitis attributed to sarcoidosis. Subjects will be stratified according to whether their uveitis is attributed to sarcoidosis (yes/no), baseline use of immunosuppressant(s) (yes/no), and prior use of anti-TNF therapy (yes/no). All subjects will be administered 60 mg/day of oral prednisone at Day 1/Baseline followed by a protocol-defined mandatory taper schedule in which all subjects continuing in the study will discontinue prednisone no later than Week 15. Subjects who enter the study on topical ocular corticosteroids will undergo a standardized taper schedule until they have completely discontinued topical ocular corticosteroids no later than Week 9. Depending on the topical ocular corticosteroid dose that the subject is on at the baseline visit, the number of drops per day will be decreased every week according to a predefined schedule.

Figure 1.1. Study Schema



The schedule of assessments is provided as an appendix to this analysis plan (Appendix 2). Beginning at Week 6 and at all subsequent visits, subjects will be examined for evidence of treatment failure. Treatment failure will be defined as a subject meeting at least one of the elements of failure criteria in at least 1 eye as outlined in Table 1-1.

Table 1-1. Treatment Failure Criteria

Parameter	Treatment Failure *	
	Week 6 Visit	**All Visits After Week 6**
Inflammatory chorioretinal and/or inflammatory retinal vascular lesions	New active, inflammatory lesions relative to Day 1/Baseline	New active, inflammatory lesions relative to Day 1/Baseline
Anterior Chamber Cell grade (SUN Criteria)	Inability to achieve ≤ Grade 0.5+	2-step increase relative to best state achieved **
Vitreous Haze grade (NEI/SUN Criteria)	Inability to achieve ≤ Grade 0.5+	2-step increase relative to best state achieved **
Visual Acuity Early Treatment Diabetic Retinopathy Study (ETDRS)	Worsening of BCVA by ≥ 15 letters relative to best state achieved	Worsening of BCVA by ≥ 15 letters relative to best state achieved

* To be considered a treatment failure, ≥ 1 of these criteria need to be present in at least 1 eye. "Best state" refers to the best measures recorded at all prior visits.
** A 2-step increase is represented by a change of Grade 0 to Grade 2+; or Grade 0.5+ to Grade 3+.

Subjects with evidence of treatment failure at or after Week 6 will be discontinued from study and treated at the discretion of the investigator.

1.3. Sample Size and Power

Sample size calculations are based on published data from the VISUAL I trial {Jaffe 2016}.

Assuming the observed proportion of subjects failing treatment by Week 24 is 70% in the placebo group, 107 subjects per treatment group will provide at least 85% power to detect a reduction of 20% in the proportion of subjects with treatment failure by Week 24 in the filgotinib group using a 2-sided significance level of 0.05. Given an expected attrition rate of 15%, a total of 248 subjects will need to be randomized into the study.

On 15 December 2020, Gilead decided to halt the study at an enrollment of 74 subjects despite the study not achieving complete enrollment of planned 248 subjects. The discontinuation of global development for filgotinib was a business decision and not because of any safety concerns with filgotinib or the study procedures.

2. TYPE OF PLANNED ANALYSIS

2.1. Interim Analyses

An external multidisciplinary Data Monitoring Committee (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 initial DMC safety data review meeting will be conducted after approximately 25% of enrolled subjects have either completed their Week 24 visit or discontinued from the study before Week 24. Following the initial data review meeting, further DMC reviews will be scheduled to occur after approximately 50% and 75% of enrolled subjects have either completed their Week 24 visit or discontinued from the study before Week 24.

There will be 1 planned interim analysis of efficacy and futility based upon the primary endpoint when approximately 50% of enrolled subjects have either completed their Week 24 visit or discontinued from the study before Week 24. The rejection boundary for efficacy will be based on a significance level of 0.001. A non-binding futility rule is defined as a predictive power (i.e., the probability of obtaining a statistically significant result at the final analysis given the observed interim data) being < 20%.

To control for the overall type I error rate at 0.05, an alpha of 0.001 will be spent for the filgotinib group comparison to the placebo group based on the Haybittle-Peto boundaries at the interim analysis. As a result, an adjusted 2-sided significance level of 0.0498 will be used to declare statistical significance for the final primary analysis.

If the DMC recommends early termination of the study, an internal Gilead team who are not involved with the study may perform an unmasked review of the interim data to decide the path forward. While the DMC will be asked to advise Gilead regarding future conduct of the study, including possible early study termination, Gilead retains final decision-making authority on all aspects of the study.

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

At the time of study early termination, two DMC safety data review meetings had taken place, but no interim analysis of efficacy and futility were performed.

2.2. Final Analysis

After all subjects enrolled prior to study early termination have completed the study, outstanding data queries have been resolved or adjudicated as unresolvable, and the data have been cleaned and finalized, 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 the 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.

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. Subjects included in each analysis set will be determined before the study blind is broken for analysis. The analysis set will be identified and included as a subtitle of each table, figure, and listing (TFL).

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 are 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 received at least 1 dose of study drug.

3.1.3. Evaluable Analysis Set

The Evaluable Analysis Set is the primary analysis set for efficacy analyses which includes all randomized subjects who receive at least one dose of study drug, and do not permanently discontinue from the study prior to Week 6.

3.1.4. Safety Analysis Set

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

3.2. Subject Grouping

For analyses based on the All Randomized Analysis Set, Evaluable Analysis Set, and FAS, subjects will be grouped according to the treatment to which they were randomized. For analyses based on the Safety Analysis Set, subjects will be grouped according to 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. In this case, the actual treatment received is defined as the treatment received for the entire treatment duration.

3.3. Strata and Covariates

Subjects will be randomly assigned to treatment groups via the interactive web response system (IWRS) in a 1:1 ratio using a stratified randomization schedule. Stratification will be based on the following variables:

- Uveitis is attributed to sarcoidosis (yes/no)
- Baseline use of immunosuppressant (yes/no)
- Prior use of anti-TNF therapy (yes/no)

If there are discrepancies in stratification factor values between the IWRS and the clinical database, the values recorded in the clinical database will be used for analyses.

Efficacy endpoint will be evaluated using stratification factors as covariates or stratification variables for analyses, as specified in Section 6. If there are a small number of subjects (eg, < 5 subjects) within a stratum of a stratification factor, then the stratification factor will not be included as covariates or stratification variables for efficacy analysis.

For efficacy endpoints, the baseline value of the efficacy variable(s) will be included as a covariate in the efficacy analysis model.

3.4. Examination of Subject Subgroups

Subgrouping of subjects based on the following two stratification factors will be explored for subgroup analyses.

- Uveitis is attributed to sarcoidosis (yes/no)
- Baseline use of immunosuppressant (yes/no)

3.5. Multiple Comparison

The number of enrolled subjects is much smaller than planned, due to early termination of the study. CCI [redacted] The correspondingly reported *p*-values for all the comparisons planned in this SAP will be nominal, and need to be interpreted with caution.

3.6. Missing Data and Outliers

3.6.1. 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. 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. Imputation rules adopted in the efficacy analyses are specified in Section 6.

3.6.2. Outliers

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

3.7. 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 (e.g., 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.

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 lower LOQ at the same precision level of the originally reported value will be used to calculate descriptive statistics if the datum is reported in the form of "< x" (where x is considered the lower 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 for calculate summary statistics.

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

- The lower or upper 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 lower or upper LOQ, respectively).

If methods based on the assumption that the data are normally distributed are not adequate, analyses may be performed on transformed data or nonparametric analysis methods may be used, as appropriate.

PK concentration values that are below the limit of quantitation (BLQ) will be presented as "BLQ" in the data listing.

3.8. Analysis Visit Windows

3.8.1. Definition of Study Day

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

- For postdose study days: Assessment Date – First Dosing Date + 1

- For days prior to the first dose: Assessment Date – First Dosing Date

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

3.8.2. 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.

In general, the baseline value will be the last nonmissing value on or prior to the first dose date of study drug.

The efficacy endpoints, vital signs, body weight, and safety laboratory data collected during the 'On-treatment' period will be mapped to the analysis visit window as specified below. The 'On-treatment' period is defined as from Study Day 1 to the last dose date of any study drug + 5 days. The on-treatment visit windows will be calculated from Study Day 1.

CCI

CCI

CCI

CCI

Efficacy data, vital signs, body weight, and safety laboratory collected in the post-treatment follow-up period will be summarized as a separate visit, and labeled "Follow-up Visit" and also be included in the listings. The analysis window for the post-treatment follow-up period is defined as from the last dose date of any study drug + 6 days to the last dose date of any study drug + 35 days. The analysis visit window won't be applied to data collected after the last dose + 35 days, and the data obtained will only be included in the listings with the nominal visit name displayed only.

When data are summarized by visit, the following visits will be summarized together in a separate visit labeled as "Week 52/EOT visit or ET visit " visit.

- Week 52 visit for subjects who complete study until Week 52 visit,
- First treatment failure visit for subjects who have a treatment failure,
- ET visit, or last available visit prior to study discontinuation when ET visit is not available, for subjects who prematurely discontinue from the study (early termination).

3.8.3. 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 dose date of study drug, unless otherwise specified. If multiple measurements occur on the same day, the last nonmissing value prior to the time of first dose 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, if there are multiple records with the same time or no time recorded on the same day, the value with the worst severity within the window will be selected (eg, abnormal will be selected over normal for safety ECG findings).

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 randomized subjects. If there are discrepancies in the value used for stratification assignment between the IWRS and the clinical database or other source documents, 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. 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:

- Safety Analysis Set
- Evaluable Analysis Set
- Full Analysis Set
- Completed study drug
- Did not complete study drug with reasons for premature discontinuation of study drug
- Completed prednisone taper
- Did not complete Prednisone taper with reasons for discontinuation of prednisone taper
- Completed study
- Did not complete study with reasons for premature discontinuation from the 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. In addition, the status of prednisone taper completion and reasons for discontinuation 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
- Reasons for premature prednisone taper discontinuation
- Reasons for screen failure (will be provided by screening ID number in ascending order)
- Lot number and bottle ID of assigned study drugs

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.

4.2.1. Duration of Exposure to Study Drug

Total duration of exposure to any 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 for subjects included in the final analyses. If month and year of the last dose are known, and the last study drug dosing date imputed above is different from the month collected, the last date of that month will be used. If only year of the last dose is known, and the last study drug dosing date imputed above is after the year collected, the last date of that year will be used; if the last study drug dosing date imputed above is before the year collected, the first date of that year will be used.

The total duration of exposure to study drug will be summarized using descriptive statistics and using the number (ie, cumulative counts) and percentage of subjects exposed through the following time periods: Baseline (Day 1), Week 1 (Day 8), Week 4 (Day 29), Week 6 (Day 43), Week 8 (Day 57), Week 12 (Day 85), Week 16 (Day 113), Week 20 (Day 141), Week 24 (Day 169), Week 28 (Day 197), Week 32 (Day 225), Week 36 (Day 253), Week 40 (Day 281), Week 44 (Day 309), Week 48 (Day 337), and Week 52 (Day 365).

The total duration of exposure to prednisone will also be summarized using descriptive statistics and using the number and percentage of subjects exposed through the following time periods: Baseline (Day 1), Week 1 (Day 8), Week 4 (Day 29), Week 6 (Day 43), Week 8 (Day 57), Week 12 (Day 85), and Week 15 (Day 106).

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:

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

If a bottle is dispensed and the bottle is returned empty, then the number of tablets returned will be entered as zero. If a bottle is dispensed but not returned (missing), the number of tablets returned will be counted as zero. On-treatment adherence will be calculated to assess the level of adherence to study drugs.

The level of on-treatment adherence to the study drug 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:

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

Study drug expected to be administered for filgotinib 200 mg/PTM (tablets) = 1× total duration of exposure to study drug (days).

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 separately 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 randomized in the study will be summarized. 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 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 (IPDs) by deviation category (as specified in the IPD plan) will be summarized by treatment group for the All Randomized Analysis Set. A by-subject listing will be provided for those subjects with important protocol deviation.

4.4. Assessment of COVID-19 Impact

This study was ongoing during the novel coronavirus (COVID-19) pandemic which has an impact on the study conduct. Some participants were unable to attend onsite visits due to shelter in place guidelines, site closures, or other reasons. This section describes how special situations due to COVID-19 will be handled in the analysis.

4.4.1. Study Drug or Study Discontinuation Due to COVID-19

A by-subject listing of reasons for premature study drug or study discontinuation due to COVID-19 will be provided if applicable. Similarly, a by-subject listing of reasons for prednisone discontinuation due to COVID-19 will be provided.

4.4.2. Protocol Deviations Due to COVID-19

A by-subject listing will be provided for subjects with IPDs related to COVID-19 if applicable. A separate listing will be provided for subjects with non-IPD related to COVID-19 if applicable.

4.4.3. 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 will be collected as free text in the CRF comment fields. The determination of missed or virtual visits due to COVID-19 will be performed using Natural Language Processing (NLP) to search the CRF comment fields. A detailed explanation of the algorithm is given in Appendix 3.

4.4.4. Adverse Events Due to COVID-19

Adverse events (AEs) of COVID-19 will be included in analyses of AEs if applicable, which will be determined through COVID-19 SMQ broad search. A by-subject listing of AEs of COVID-19 will be provided if applicable.

4.4.5. Overall Assessment of COVID-19 Pandemic Impact

For subjects affected by COVID-19 infection and/or pandemic while participating in the study, a listing of the following individual COVID-19 related outcome categories will be provided:

- Death due to COVID-19
- Adverse event of COVID-19, as determined by COVID-19 SMQ broad search
- Specific adverse event directly associated with the pathogen causing COVID-19, as determined by MST
- Hospitalization (using data from AE eCRF) due to adverse event of COVID-19 as defined above
- Tested positive for COVID-19
- Study drug discontinuation due to COVID-19
- Study discontinuation due to COVID-19
- Missed visits due to COVID-19
- Missed key assessments due to COVID-19 (key assessments are the assessments contributing to primary and secondary endpoints)

In addition, composite broad COVID-19 impact indicator will be derived based on the following individual categories defined above: death, adverse event, hospitalization, tested positive, study drug discontinuation, study discontinuation, missed visits, and missed key assessments.

Composite specific COVID-19 impact indicator will be derived based on death, specific adverse event, and tested positive.

5. BASELINE CHARACTERISTICS

5.1. Demographics and Baseline Characteristics

Subject demographic variables (ie, age, sex, race, and ethnicity) and other baseline characteristics (ie, body weight [in kg], height [in cm], body mass index [BMI; in kg/m^2], smoking status) will be summarized by treatment group and overall using descriptive statistics for continuous variables, and using number and percentage of subjects for categorical variables. The summary of demographic and baseline characteristics data will be provided for the Safety Analysis Set.

- Age (on the first dose date of any study drug)
- Age group (< 65 years, ⩾ 65 years)
- Age group (< 30, ⩾ 30 to < 50, ⩾ 50 years)
- Sex at birth (male, female)
- Race
- Ethnicity (Hispanic or Latino, not Hispanic or Latino)
- Weight (kg)
- Height (cm)
- Body mass index (BMI; in kg/m^2)
- Smoking status (non-smoker, former smoker, current smoker)
- Drinking status (non-drinker, former drinker, current drinker)

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

5.2. Other Baseline Disesase Characteristics

Baseline disease characteristics include:

- Uveitis attributed to sarcoidosis (yes/no)
- Baseline immunosuppressant usage (yes/no)
- Prior use of anti-TNF therapy (yes/no)

- Prior immunosuppressants (excluding corticosteriod) use (yes/no) during screening period: n (%)
  - Azathioprine
  - Methotrexate (MTX)
  - Adalimumab
  - Infliximab
  - Ciclosporin
  - Mycophenolate Mofetil
- Prior Corticosteriod use (yes/no) during screening period: n (%)
  - Oral
  - Ophthalmic
  - Intraocular
  - Other
- Concurrent immunosuppressants use (excluding corticosteriod) on the first dose date (yes/no): n (%)
  - Azathioprine
  - Methotrexate (MTX)
  - Adalimumab
- Concurrent corticosteriod use on the first dose date (yes/no): n (%)
  - Oral
  - Ophthalmic
  - Intraocular
  - Other

[redacted]

[redacted]

- Duration of current uveitis flare (month)
  - Calculated as earlier start date (from two eyes) of the current uveitis flare to the first dose date, where the current uveitis flare is the study qualifying uveitis flare identified at the screening visit
- Eye affected (left, right, both) for current uveitis flare
- Type of Uveitis (pan-uveitis, posterior uveitis, intermediate uveitis, anterior + intermediate, posterior + intermediate) for current uveitis flare
- Etiology for current uveitis flare
- History of infectious uveitis (yes, no)
- Duration of uveitis (months) from the first uveitis flare
  - Calculated as earlier start date (from two eyes) of the first uveitis flare to the first dose date
- Time since last uveitis flare (months) at Baseline
  - Calculated as from later end date (from two eyes) of the last uveitis flare to earlier start date (from two eyes) of current uveitis flare
- Number of uveitis flares in the past 12 months (1, 2, ≥ 3)
- Corticosteroid use within 14 days prior to first dose
  - Oral corticosteroid use , n (%)
- Maximum oral corticosteroid dose use within 14 days prior to first dose (mg) for subjecs who use oral corticosteroid, expressed as prednisone-equivalent dose, mean (SD)

The following baseline **ophthalmologic variables** will be summarized separately for left and right eye:

- Optical coherence tomography (OCT) evidence of macular edema (yes, no) (based on central reader assessment at baseline)
- Central retinal thickness (microns)
- Intraocular pressure (mmHg)
- Active lesions (yes, no) via Dilated Indirect Ophthalmoscopy (DIO) and lesion location
  - Chorioretinal: n (%)

  - Retinal vascular: n (%)
  - Both chorioretinal and retinal vascular: n (%)
  - Other: n (%)
- Anterior chamber cell grade (SUN criteria) (continuous)
- Anterior chamber cell grade (SUN criteria) (categorical)
- AREDS lens opacity grading
  - Nuclear opacity (grade < 1.0, 1.0, 1.5, 2.0, 2.5, 3.0, > 3.0)
  - Cortical opacity (grade < 1.0, 1.0, 1.5, 2.0, 2.5, 3.0, > 3.0)
  - Posterior subcapsular opacity (grade < 1.0, 1.0, 1.5, 2.0, 2.5, 3.0, > 3.0)
- Vitreous haze grades (NEI/SUN criteria) (continuous)
- Vitreous haze grades (NEI/SUN criteria) (categorical)
- Best Corrected Visual Acuity (BCVA)
- LogMAR BCVA, calculated as follows:

  if ≥ 20 letters were read at 4 m OR if < 20 letters were read at 4 m but the test was not done at 1 m,

  logMAR = 1.7 – 0.02 * (number of letters at 4 m + 30),

  if < 20 letters were read at 4 m and the test was also done at 1 m,

  logMAR = 1.7 – 0.02 * (number of letters read at 4 m + number of letters read at 1 m)

  if 0 letters were read at 1 m, the following logMAR values will be calculated according to Holladay:

  in case of finger counting: logMAR = (–1) * log10 (distance in meters/60)

  in case of hand movements: logMAR = 3.0

These baseline ocular characteristics will be summarized by treatment group and overall using descriptive statistics (n, mean, SD, median, Q1, Q3, minimum, and maximum) for continuous variables and using number and percentage of subjects for categorical variables. The summary of

baseline characteristics will be provided for the Safety Analysis Set. No formal statistical testing is planned.

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

5.3. Medical History

Medical history will be collected at screening for disease-specific and general conditions (ie, conditions not specific to the disease being studied).

General medical history collected at Screening will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) v24.0. Medical history will be summarized by system organ class (SOC), preferred term (PT), treatment group, and overall. Subjects who report 2 or more medical history items that are coded to the same SOC and/or PT will be counted only once by the unique coded term in the summary.

A by-subject listing of general medical history will be provided by subject ID number (in ascending order) and medical history of abnormalities (in chronological order). A by-subject listing of uveitis-specific medical history will also be provided.

6. EFFICACY ANALYSES

6.1. General Consideration

The primary analysis set for efficacy analyses will be the Evaluable Analysis Set, defined in section 3.1.3.

Due to study early terminiation, the number of enrolled subjects is much smaller than planned, all the hypothesis testings performed and *p*-values reported will be nominal, and, unless specified otherwise, all the confidence intervals will be calculated based on the nominal level of 95%. The results need to be interpreted with caution.

Missing Data Consideration

Unless specified otherwise, below are the descriptions for the imputation methods that will be used throughout the efficacy analyses:

- Observed case (OC): Missing values remain missing. For the categorical composite endpoints, in the case that some components are missing, the composite endpoint assessment will be derived based on the non-missing components. If non-missing components are not sufficient to determine final composite endpoint, then the composite endpoint will be set as missing. For example, the treatment failure status is a binary composite score derived based on four criteria. If one of the criteria (e.g., new lesion relative to baseline) is missing, the treatment failure status will be derived based on other three non-missing criteria (VH grade, ACC grade and BCVA score). However, if none of the three non-missing criteria meets the treatment failure definition, the treatment failure status will be set as missing.
- Non-responder imputation (NRI): For all binary response measurements, all subjects with missing values will be analyzed as having the clinical event (i.e., treatment failure).

6.2. Primary Efficacy Endpoint

The primary efficacy endpoint is the proportion of subjects failing treatment by Week 24.

6.2.1. Definition of the Primary Efficacy Endpoint

Subjects will be evaluated for treatment failure criteria as defined in Table 1-1.

6.2.2. Primary Analysis of the Primary Efficacy Endpoint

The primary hypothesis will consist of a superiority test of filgotinib compared to placebo based on proportion of subjects failing treatment by Week 24. The analysis will use Cochran-Mantel-Haenszel (CMH) approach adjusting for the stratification factors. The *p*-value from the CMH test for testing the superiority of filgotinib as compared to placebo will be provided. Subjects with missing values on treatment failure status will be analyzed as treatment failure using Non-Responder Imputation (NRI).

The non-stratified point estimate for treatment differences in proportions of subjects failing treatment by Week 24 between filgotinib and placebo along with its 95% confidence interval (CI) will be provided based on the normal approximation with a continuity correction. In addition, the numbers and percentages of subjects failing treatment in each treatment group with the corresponding 2-sided 95% CIs for the percentage based on a normal approximation method with a continuity correction will be presented. Appendix 4 provides sample SAS models statements for constructing the confidence interval for the proportion.

A logistic regression analysis will also be performed to assess the odds ratio between treatment groups adjusting for stratification factors. The point estimate of odds ratio and the corresponding 95% CI, as well as the *p*-value, will be presented.

6.2.3. Sensitivity Analysis of the Primary Endpoint

The following sensitivity analyses of the primary efficacy endpoint will be performed.

- The analysis specified in section 6.2.2 will be performed using Full Analysis Set. The missing data will be imputed using non-responder imputation.

- The analysis will also be performed using the observed data only. The missing data will be not be imputed, i.e. the subjects who discontinue from the study will be excluded from the analysis.

6.2.4. Subgroup Analysis of the Primary Endpoint

Subgroup analyses comparing filgotinib group to the placebo group will be performed for the primary endpoint, for the subgroups specified in Section 3.4.

The proportion of subjects failing treatment by Week 24 between filgotinib and placebo will be compared using the stratified CMH test based on the NRI method with corresponding *p*-value provided. The non-stratified point estimate and 95% CIs for the differences in proportions between filgotinib and placebo, and 95% CIs for proportion of responders in each treatment group based on a normal approximation method with a continuity correction will be presented.

When the grouping variable is identical to one of the stratification factors, this stratification factor will not be adjusted in the modeling. In the case there were not sufficient data to meet the assumptions of the analysis, i.e. < 5 subjects within a stratification stratum, a Fisher exact test will be conducted instead.

6.3. Secondary Efficacy Endpoints

The secondary efficacy endpoints include:

- Time to treatment failure on or after Week 6

- Change in VH grade in each eye (NEI/SUN criteria), from best state achieved prior to Week 6 to Week 52/EOT visit or ET visit
- Change in AC cell grade in each eye, from best sate achieved prior to Week 6 to Week 52/EOT visit or ET visit
- Change in logMAR BCVA in each eye, from best state achieved prior to Week 6 to Week 52/EOT visit or ET visit
- Log change in central retinal thickness in each eye, from best state achieved prior to Week 6 to Week 52/EOT visit or ET visit
- Time to development of macular edema in at least one eye as determined by OCT on or after Week 6

6.3.1. Definition of Secondary Efficacy Endpoints

6.3.1.1. Anterior Chamber Cell Grade

The number of Anterior Chamber (AC) cells observed within a 1 mm × 1 mm slit beam will be recorded for each eye. The reported number will be used to determine the grade according to the Standardization of Uveitis Nomenclature (SUN) criteria (Table 6-1).

Table 6-1. Anterior Chamber Cells

Grade	Cells in Field
0	< 1
0.5 +	1 – 5
1 +	6 – 15
2 +	16 – 25
3 +	26 – 50
4 +	> 50

Anterior chamber cell grades range from 0 to 4+, with higher scores indicating more cells visible in the anterior chamber and greater severity of uveitis.

6.3.1.2. Vitreous Haze Grading

Grading of vitreous haze will be based on the publication from the National Eye Institute (NEI) which has also been adapted by the SUN working group (Table 6-2).

Table 6-2. Vitreous Haze

Grade	Description
0	No evident vitreal haze
0.5 +	Slight blurring of the optic disc margin because of the haze; normal striations and reflex of the nerve fiber layer cannot be visualized
1 +	Permits a better definition of both the optic nerve head and the retinal vessels (compared to higher grades)
2 +	Permits better visualization of the retinal vessels (compared to higher grades)
3 +	Permits the observer to see the optic nerve head, but the borders are quite blurry
4 +	Optic nerve head is obscured

Vitreous haze grades range from 0 to 4+, with higher scores indicating greater severity of uveitis.

6.3.1.3. Best Corrected Visual Acuity (BCVA)

Best Corrected Visual Acuity (BCVA) is the best possible vision that an eye can achieve with the use of glasses or contact lenses. A refraction test is performed to measure the appropriate lens strength to focus light on the retina. Using the appropriate corrective lenses based on that visit's refraction, subject's BCVA is measured using an ETDRS chart. In the ETDRS system, 15 letters is equal to a change in 3 lines of visual acuity, for example, from 20/40 to 20/20 or vice versa, a halving or doubling of the visual angle; a change of 7.5 letters corresponds to a 25% decrease or 50% increase in the visual angle. Clinically, it is the single BCVA for each eye that represents the maximal visual potential. If the subject is unable to read letters on a testing chart, visual acuity is described as ranging from ability to count fingers, recognize hand movements, or light perception.

The smaller BVCA score indicates greater severity of uveitis.

6.3.1.4. Central retinal thickness

Central retinal thickness is measured by optical coherence tomography (OCT). Full retinal thickness is defined as the distance between the inner limiting membrane and the apical boundary of the retinal pigment epithelium. Central retinal thickness is defined as the thickness of the retina in the center of the foveal pit (1 mm subfield).

The larger central retinal thinkness value indicates greater severity of uveitis.

6.3.1.5. Macular edema

Macular edema is defined as central retinal thickness $\geqslant$ 300 microns if using Cirrus machine, or $\geqslant$ 315 microns if using Spectralis machine.

6.3.2. Analysis Methods for Secondary Efficacy Endpoints

The time to treatment failure and the time to OCT evidence of macular edema will be based on the first eye to meet the criteria for treatment failure or macular edema.

The change in anterior chamber cell grade in each eye from best state achieved prior to Week 6 to Week 52/EOT or ET Visit, change in vitreous haze grade in each eye from best state achieved prior to Week 6 to Week 52/EOT or ET, change in best corrected visual acuity (logarithm of the minimum angle of resolution) in each eye from best state achieved prior to Week 6 to Week 52/EOT or ET, and log change in central retinal thickness in each eye from best state achieved prior to Week 6 to Week 52/EOT or ET Visit will be analyzed using data from each eye individually.

The best state for the VH grade and AC cell grade is defined as the minimum grade prior to Week 6, and the best state for BCVA is defined as the maximum score prior to Week 6. The best state for central retinal thickness is the minimum value prior to Week 6.

The time to treatment failure on or after Week 6 will be compared between the filgotinib group and the placebo group with a stratified log-rank test. A proportional-hazards model with the treatment group as a factor and stratification factors as covariates will be fitted to estimate the hazard ratio with its 95% confidence interval. Appendix 5 provides sample SAS models statements for constructing the stratified log-rank test and the proportional-hazards model. In additional, time to treatment failure on or after Week 6 due to each individual component of treatment failure criteria will be analyzed similarly.

Treatment failures on or after Week 6 will be counted as events; subjects who complete the study and don't have events or permanently discontinue from study due to reasons other than treatment failure at any time during the study will be considered as censored observations. Time (in weeks) to the clinical event (eg, treatment failure) will be calculated as: (earliest date of clinical event - first dose date + 1) / 7. Those who are not observed to have the specified events by the end of study completion or by the time of discontinuation from study will be censored at the date of their last available assessment.

The time to macular edema on or after Week 6 will be analyzed in the same way as for time to treatment failure analysis. OCT evidence of macular edema on or after Week 6 will be counted as event; subjects who complete the study and don't have macular edema or permanently discontinue from study at any time during the study will be considered as censored observations. Similar analyses will also be performed for subjects with and without pre-existing macular edema at baseline, respectively.

For time to event endpoints, the Kaplan - Meier (KM) estimates of clinical event probability and 95% CIs for each treatment group will be presented at Week 6, 12, 24, 36, 48, and 52. The 95% CI will be calculated based on a log-log transformation of the survival function. Median, Q1 and Q3 of the time to first clinical event will also be provided for each treatment group with the corresponding 95% CIs based on the log-log transformation of the survival function. The KM curves will be plotted for treatment failure or macular edema by treatment groups, respectively.

For continuous endpoints of change in Vitreous Haze grade, change in AC cell grade, change in BCVA (logarithm of the minimum angle of resolution), and log change in central retinal thickness will be compared between treatment groups using the repeated measures Analysis of Covariance (ANCOVA) to control for clustered observations (i.e., observations from each of the subject's eyes), with treatment, subject's eyes, interaction of treatment and subject's eys, stratification factors and best state values as covariates. For log change in central retinal thickness, the analysis will additionally be adjusted for type of OCT machines. Least squares mean estimate and 95% CI of treatment difference between filgotinib and placebo in change from baseline at Week 52/EOT or ET visit and corresponding *p*-value from the ANCOVA model will be provided. Least squares mean estimate and 95% CI will also be presented for each treatment group.

For all the continuous endpoints, absolute value and change from best state will be summarized using descriptive statistics (sample size, mean, SD, median, Q1, Q3, minimum, and maximum) by treatment and by visit. For central retinal thickness, the descriptive summary will also be provided by visit for subjects with and without pre-existing macular edema at baseline, respectively. In addition, plots of mean ± SD for change from the best state by visit will be presented. For log change in central retinal thickness, plots of mean ± SD for change from Baseline will also be provided.

6.3.3. Sensitivity Analysis for Key Secondary Efficacy Endpoints

The following sensitivity analyses of the key secondary efficacy endpoints will be performed.

- For time to treatment failure on or after Week 6, subjects who permanently disconinute from study at any time during the study will be considered as having events at the time of discontinuation from study.

- For time to macular edema on or after Week 6, subjects who permanently disconinute from study at any time during the study will be considered as having events at the time of discontinuation from study.

6.3.4. Subgroup Analysis of the Key Secondary Endpoints

Subgroup analyses comparing filgotinib group to the placebo group will be performed for the key secondary endpoints, for the subgroups specified in Section 3.4.

For continus key secondary endpoint, the repeated measures ANCOVA with treatment, subject's eyes, interaction of treatment and subject's eyes, stratification factors and best state value as covariates will be performed.

For time to event endpoints, the stratified log-rank test will be used to compare the filgotinib group and the placebo group. The Kaplan-Meier estimates of clinical event probability and 95% CIs at Week 6, 12, 24, 36, 48, and 52 for each treatment group will also be presented. Median, Q1 and Q3 of the time to the first clinical event will also be provided for each treatment group.

The point estimates and 95% CIs based on proportional hazard model will also be provided in the subgroup analysis.

When the grouping variable is identical to one of the stratification factors, this stratification factor will not be adjusted in the modeling.

CCI

CCI

[redacted]

[redacted]

6.5. Changes From Protocol-Specified Efficacy Analyses

On 15 December 2020, Gilead decided to halt the study at an enrollment of 74 subjects despite the study not achieving complete enrollment of planned 248 subjects. The discontinuation of global development for filgoginib was a business decision and not because of any safety concerns with filgotinib or the study procedures.

The final analysis will be based on these 74 subjects. CCI [redacted]

[redacted]

7. SAFETY ANALYSES

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 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 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 CRF to the question of "Related to Study Treatment." Relatedness will always default to the investigator's choice, not that of the Medical Monitor. 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 SAEs that were specified in the study protocol. SAEs captured and stored in the clinical database will be reconciled with the SAE database from the Gilead Global Patient Safety Department 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 one or both of the following:

- Any AEs with an onset date on or after the study drug start date 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 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 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.

7.1.6.1. Summaries of AE incidence in Combined Severity Grade Subsets

A brief, high-level summary of the number and percentage of subjects who experienced at least 1 TEAE in the categories described below will be provided by treatment group. All deaths observed in the study will also be included in this summary.

The number and percentage of subjects who experienced at least 1 TEAE will be provided and summarized by SOC, HLT (if applicable), PT, and treatment group:

- TEAEs

For the AE categories described below, summaries will be provided by SOC, PT, and treatment group:

- TEAEs of Grade 3 or higher
- TEAEs of Grade 2 or higher
- TE treatment-related AEs
- TE treatment-related AEs of Grade 3 or higher
- TE treatment-related AEs of Grade 2 or higher
- TE SAEs
- TE treatment-related SAEs

- TEAEs leading to premature discontinuation of any study drug
- TEAEs leading to premature discontinuation of study
- TEAEs leading to temporary interruption of study drug
- TEAEs leading to death (ie, outcome of death)
- Death

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 to the above summary tables, all above AE categories will be summarized by PT only, in descending order of total frequency.

In addition, 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
- All SAEs
- All AEs leading to premature discontinuation of any study drug
- All AEs leading to premature discontinuation of study
- All AEs leading to temporary interruption of study drug
- All deaths
- All AEs leading to death (ie, outcome of death)

7.1.6.2. Summaries of AE incidence by Severity

A brief, high-level summary of the number of percentage of subjects who experienced at least 1 TEAE in the categories described below will be provided by treatment group. All deaths observed in the study will also be included in this summary.

For the AE categories described below, summaries will be provided by SOC, PT, and treatment group:

- TEAEs by maximum severity
- TE treatment-related AEs by maximum severity

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.

7.1.7. Adverse Events of Interest

Events of interest will be identified by the use of either SMQs or MedDRA search terms (MSTs). However, should additional cases not detected by the predefined search term listings be identified during the clinical review process, these cases will also be reported by respective category.

7.1.7.1. Adjudication Committee for Major Adverse Cardiovascular Events (MACE) and Thromboembolic Events

An independent adjudication committee governed by an adjudication charter will be set up to perform periorically adjudication of potential MACE as well as thromboembolic events reported during the study.

7.1.7.1.1. Major Adverse Cardiovascular Events (MACE)

MACE will be comprised of cardiovascular (CV) death, non-fatal myocardial infarction (MI) and non-fatal stroke. To identify the MACE, the following potential cases will be adjudicated:

- All deaths
- Cardiovascular events (meeting serious criteria)
- Myocardial infarction (Narrow scope SMQ)
- Hospitalization for unstable angina
- Transient ischemic attack
- Stroke
- Hospitalization for cardiac failure
- Percutaneous coronary intervention

7.1.7.1.2. Thromboembolic Events

Thromboembolism events will be comprised of arterial systemic thromboembolism events (ASTE) and venous thromboembolism events (VTE) associated with deep vein thrombosis (DVT) and/or pulmonary embolism (PE). To identify the thromboembolism events, the potential cases identified using the Embolic And Thrombotic Events SMQ search will be adjudicated.

The adjudication committee will review those potential MACE and thromboembolic events and related clinical data to determine whether a MACE or thromboembolism event has developed.

The number and percentage of subjects with positively adjudicated MACE or thromboembolic events will be summarized by treatment group using PT.

By-subject listings for all randomized subjects who have potential MACE or thromboembolic events will be provided respectively.

7.1.7.2. Other Adverse Events of Interest

In addition to general safety parameters, MACE and thromboembolic Events, safety information on other adverse events of interest (AEIs) will also be analyzed. AEIs will be identified by either laboratory results, standardized MedDRA queries (SMQs), or sponsor defined MSTs, or a combination of these methods as indicated below.

- All infections (defined as all PTs in the Infections and Infestations SOC)
- Serious infections (defined as all PTs in the Infections and Infestations SOC that are SAEs)
- Infections of special interest identified by MSTs
    a) Herpes zoster
    b) Active tuberculosis
    c) Opportunistic infections (OIs)
    d) Hepatitis B or C infections
- Malignancies (including lymphoma; excluding non-melanoma skin cancers) identified by MSTs
- Nonmelanoma skin cancer
- Gastrointestinal (GI) perforation
- Uveitis-related AEs

The number and percentage of patients with aforementioned events of interest will be provided by the PT term for each AE of special interests.

A by-subject listing for all patients having AE of interests at any time will be provided for each AE of interest.

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 and will include data collected up to the last dose of any study drug plus 30 days for subjects who have permanently discontinued study drug, or all available data at the time of the database snapshot for subjects who were ongoing at the time of an interim analysis. 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 3.7.

A by-subject listing for laboratory test results will be provided by subject ID number and visit in chronological order for hematology, serum chemistry, and urinalysis separately. Values falling outside of the relevant reference range and/or having a severity grade of 1 or higher on the CTCAE severity grade will be flagged in the data listings, as appropriate.

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 as follows:

- Baseline values
- Values at each postbaseline visit
- Change from baseline at each postbaseline visit

A baseline laboratory value will be defined as the last measurement obtained on or prior to the date/time of first dose of any 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 values for the laboratory tests specified above will be plotted using a line plot by treatment group and visit.

In the case of multiple values in an analysis window, data will be selected for analysis as described in Section 3.8.3.

7.2.2. Graded Laboratory Value

The CTCAE Version 4.03 (with one modification specified in protocol) will be used to assign 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 any study drug plus 30 days for subjects who permanently discontinued study drug, or the last available date in the database snapshot for subjects who were still on treatment at the time of an interim analysis.

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 from baseline by at least 3 toxicity grades at any postbaseline time point, up to and including the date of the last dose of any study drug plus 30 days for subjects who permanently discontinued study drug or the last available date in the database snapshot for subjects who were still on treatment at the time of an interim analysis.

If the relevant baseline laboratory value is missing, any Grade 3 or higher 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 visit.

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 lab 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 higher laboratory abnormalities and marked laboratory abnormalities will be provided separately by subject ID number and visit in chronological order. These listings will include all test results that were collected throughout the study for the lab test of interest, with all applicable severity grades or abnormal flags 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:

- Aspartate aminotransferase (AST): (a) > 3 times of the upper limit of reference range (ULN); (b) > 5 x ULN; (c) > 10 x ULN; (d) > 20 x ULN
- Alanine aminotransferase (ALT): (a) > 3 x ULN; (b) > 5 x ULN; (c) > 10 x ULN; (d) > 20 x ULN
- AST or ALT: (a) > 3 x ULN; (b) > 5 x ULN; (c) > 10 x ULN; (d) > 20 x ULN
- Total bilirubin: (a) > 1 x ULN; (b) > 2 x ULN
- Alkaline phosphatase (ALP) > 1.5 x ULN
- AST or ALT > 3 x ULN and total bilirubin: (a) > 1.5 x ULN; (b) > 2 x ULN
- AST or ALT > 3 x ULN, total bilirubin > 2 x ULN and ALP < 2 x ULN

The summary will include data from all postbaseline visits up to 30 days after the last dose of any study drug. 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 and Vital Signs

Descriptive statistics will be provided by treatment group for body weight, BMI and vital signs (systolic and diastolic blood pressures [mmHg], pulse [beats/min]) as follows:

- Baseline value
- Values at each postbaseline visit
- Change from baseline at each postbaseline visit

A baseline value will be defined as the last available value collected on or prior to the date/time of first dose of any 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 3.8.3. No formal statistical testing is planned.

A by-subject listing of vital signs (systolic and diastolic blood pressure [mmHg], pulse [beats/min], and body temperature [^{o}C]) will be provided by subject ID number and visit in chronological order. In the same manner, a by-subject listing of body weight, height, and BMI will be provided separately

7.4. Prior and Concomitant Medications

Medications collected at screening and during the study will be coded using the World Health Organization (WHO) Drug dictionary version BMAR21.

All the analyses in this section will be performed for general prior /concomitant medications and Uveitis-specific prior /concomitant medications separately, unless otherwise specified.

7.4.1. Prior Medications

Prior medications are defined as any medication taken before a subject took the first study drug.

Prior medications will be summarized 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 alphabetically by preferred name in order of 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 any 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. Use of concomitant medications will be summarized 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 alphabetically by preferred name 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 any 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 any study drug will also be considered concomitant. Medications with a stop date prior to the date of first dosing date of any study drug or a start date after the last dosing date of any 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

7.5.1. Investigator Electrocardiogram Assessment

ECGs are only collected at the screening visit. A by-subject listing of ECG assessment results at screening will be provided.

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.

8. PHARMACOKINETIC ANALYSES

Concentrations of filgotinib and its metabolite GS-829845 in plasma will be determined using validated bioanalytical assays.

Individual subject concentration data for filgotinib and GS-829845 will be listed. Individual concentration data listings will include all subjects with concentration data. PK sampling details by subject, including procedures, differences in scheduled and actual draw times if applicable, and sample age will be provided in listings.

Due to study early termination, individual concentration data summaries will not be provided; No pharmacokinetic (PK) parameters will be derived and no analyses for PK parameters will be performed.

CCI [redacted]

[redacted]

10. REFERENCES

Jaffe GJ, Dick AD, Brezin AP, Nguyen QD, Thorne JE, Kestelyn P, et al. Adalimumab in Patients with Active Noninfectious Uveitis. N Engl J Med. 2016;375 (10):932-43.

Mangione CM, Lee PP, Gutierrez PR, Spritzer K, Berry S, Hays RD, et al. Development of the 25-item National Eye Institute Visual Function Questionnaire. Arch Ophthalmol 2001;119 (7):1050-8.

11. SOFTWARE

SAS® Software Version 9.4. SAS Institute Inc., Cary, NC, USA.

12. SAP REVISION

Revision Date (DD MMM YYYY)	Section	Summary of Revision	Reason for Revision

13. APPENDICES

Appendix 1. Corticosteroids Dose Conversion Chart

The following table will be used for converting non-prednisone medications to prednisone equivalent:

*Example: Subject is taking 8 mg of Methylprednisolone orally daily. To get the equivalent dose of prednisone: 8 mg Methylprednisolone = (5*8)/4 = 10 mg prednisone.*

Corticosteroids Name	Equivalent Dose (mg) to 5 mg Prednisone
Dexamethasone	0.75 mg
Methylprednisolone	4 mg
Prednisolone	5 mg
Prednisone	5 mg

Appendix 2. Study Procedures Table

	Screening	Day 1 / Baseline	Wk 1	Wk 4	Wk 6	Wk 8	Wk 12	Wk 16	Wk 20	Wk 24	Wk 28	Wk 32	Wk 36	Wk 40	Wk 44	Wk 48	Unscheduled Visit related to Uveitis Symptoms[u]	Wk 52 / End of Treatment –or– Early Termination Visit	4 Week Post-Treatment Follow-Up Visit[t]
Visit Window	**Within -30 days**		**(±3 days)**	**(±3 days)**	**(±3 days)**	**(±3 days)**	**(±3 days)**	**(±3 days)**	**(±3 days)**	**(±3 days)**	**(±3 days)**	**(±3 days)**	**(±3 days)**	**(±3 days)**	**(±3 days)**	**(±3 days)**			**(+7 days)**
Written Informed Consent	X																		
Review Inclusion/Exclusion		X																	
Medical History	X																		
Complete/ Symptom-Driven Physical Examination [a]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Height	X																		
Vital Signs [b]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Weight	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
12-lead ECG	X																		
TB Test and Chest X-ray[c,d]	X																		
Hematology[d]	X[v]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X
Chemistry[d]	X[v]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X
Lipid Profile (Fasting)[e]		X								X								X	
ANA Test[f]		X																	
Syphilis Test[d,g]	X																		
Serology Testing[h]	X[v]																		

	Screening	Day 1 / Baseline	Wk 1	Wk 4	Wk 6	Wk 8	Wk 12	Wk 16	Wk 20	Wk 24	Wk 28	Wk 32	Wk 36	Wk 40	Wk 44	Wk 48	Unscheduled Visit related to Uveitis Symptoms[u]	Wk 52 / End of Treatment –or– Early Termination Visit	4 Week Post-Treatment Follow-Up Visit[t]
Visit Window	**Within -30 days**		**(±3 days)**	**(±3 days)**	**(±3 days)**	**(±3 days)**	**(±3 days)**	**(±3 days)**	**(±3 days)**	**(±3 days)**	**(±3 days)**	**(±3 days)**	**(±3 days)**	**(±3 days)**	**(±3 days)**	**(±3 days)**			**(+7 days)**
HBV Viral Monitoring[h]							X			X			X			X		X	
Urine drug test [i]	X[v]																		
Urinalysis[d]	X[v]	X					X		X									X	X
Urine or Serum Pregnancy Test[j]	X[v]	X		X		X	X	X	X	X	X	X	X	X	X	X		X	X
Serum FSH[k]	X[v]																		
[redacted]																			
PK[m]		X		X	X		X			X			X					X	
Best Corrected Visual Acuity Testing (BCVA)	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Slit Lamp Exam	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Tonometry	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

	Screening	Day 1 / Baseline	Wk 1	Wk 4	Wk 6	Wk 8	Wk 12	Wk 16	Wk 20	Wk 24	Wk 28	Wk 32	Wk 36	Wk 40	Wk 44	Wk 48	Unscheduled Visit related to Uveitis Symptoms[u]	Wk 52 / End of Treatment –or– Early Termination Visit	4 Week Post-Treatment Follow-Up Visit[t]
Visit Window	**Within -30 days**		**(±3 days)**	**(±3 days)**	**(±3 days)**	**(±3 days)**	**(±3 days)**	**(±3 days)**	**(±3 days)**	**(±3 days)**	**(±3 days)**	**(±3 days)**	**(±3 days)**	**(±3 days)**	**(±3 days)**	**(±3 days)**			**(+7 days)**
Dilated Indirect Ophthalmoscopy	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Fundus Color Photography (both eyes)		X								X								X	
Fluorescein Angiogram (both eyes)		X								X								X	
Optical Coherence Tomography (OCT) Scan (both eyes)	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Prior and Current Uveitis Medication	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
CCI																			
CCI																			
CCI																			
CCI																			
In Clinic Study Drug Dosing[r]		X		X	X		X												
Study Drug Dispensation[r]		X		X		X	X	X	X	X	X	X	X	X	X	X			
Prednisone Dispensation[s]		X			X	X													
Drug Accountability			X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	

	Screening	Day 1 / Baseline	Wk 1	Wk 4	Wk 6	Wk 8	Wk 12	Wk 16	Wk 20	Wk 24	Wk 28	Wk 32	Wk 36	Wk 40	Wk 44	Wk 48	Unscheduled Visit related to Uveitis Symptoms[a]	Wk 52 / End of Treatment –or– Early Termination Visit	4 Week Post-Treatment Follow-Up Visit[t]
Visit Window	Within -30 days		(±3 days)	(±3 days)	(±3 days)	(±3 days)	(±3 days)	(±3 days)	(±3 days)	(±3 days)	(±3 days)	(±3 days)	(±3 days)	(±3 days)	(±3 days)	(±3 days)			(+7 days)
Adverse Events		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Con Meds	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

a Complete PE will be performed at Screening and Week 52 or ET/EOT. Symptom-driven PE will be performed at all other visits.
b Vital signs include resting blood pressure, pulse, and temperature.
c Subjects must have the QuantiFERON® TB-Gold In-Tube test AND a chest radiograph (views as per local guidelines) taken at screening or within 3 months prior to screening (with the report or films available for investigator review). Subjects with evidence of a negative QuantiFERON® TB-Gold In-Tube test result within 3 months will not require a re-test at Screening. QuantiFERON® TB-Gold can be processed locally.
d The required screening labs can be obtained either using the sites local lab or by submitting appropriate samples to the central lab. For Chemistry panel, Uric acid to be tested at screening only per protocol Appendix 6. Refer to list provided in Laboratory assessment table (protocol Appendix 6).
e Fasting lipid on Day 1 and Week 24. Week 52/EOT or ET visit lipid can be obtained fasting or non-fasting.
f Study personnel will collect a sample for antinuclear (ANA) and reflex double stranded-DNA antibody testing at the Day 1/Baseline visit. A repeat ANA/Anti-ds-DNA would be warranted if a subject has clinical signs and symptoms suggestive of lupus. The Anti-ds-DNA antibody testing will be performed in case of positive ANA result.
g Subjects whom underwent syphilis screening (either a FTA test or a syphilis IgG test) for standard of care purposes within 3 months prior to Screening and evidence of a nonreactive FTA or syphilis IgG test result will not require a re-test at Screening.
h HBsAg and core Ab (if positive core Ab, then reflex Hep B DNA), Hepatitis C Ab, (if positive then, reflex HCV RNA), HIV 1/2 antigen/antibody combination test (4th generation). (protocol Appendix 6). The required screening labs can be obtained either using the sites local lab or by submitting appropriate samples to the central lab. Subjects with positive HBV core Ab, and negative HBV DNA are eligible per investigator judgment, but may require prophylactic treatment in accordance with HBV treatment guidelines/local standard of care and require ongoing monitoring with blood tests for HBV DNA every 3 months.
i Drug screen to be conducted at screening only per protocol Appendix 6. The required screening labs can be obtained either using the sites local lab or by submitting appropriate samples to the central lab.
j Females of childbearing potential only. Serum pregnancy test at Screening and Urine pregnancy test to be conducted every 4 weeks during the dosing period through 35 days after their last dose of study drug. The required screening labs can be obtained either using the sites local lab or by submitting appropriate samples to the central lab.
k For female subjects <54 and amenorrhea ≥12 months. The required screening labs can be obtained either using the sites local lab or by submitting appropriate samples to the central lab.
[redacted]
m Day 1/Baseline and Week 12 PK sample to be drawn at least 30 minutes post dose. Week 4 and/or 6, PK sample to be drawn within 2 hours prior to dose. Weeks 24, 36 & 52/EOT or ET PK can be drawn at any time.
[redacted]

[redacted]

r Study drug will be dispensed to subjects every 4 weeks until Week 48 or prior to the ET/EOT study visit.
s Prednisone will be dispensed to subjects at Day 1/Baseline (Dose strengths 20 mg and 10 mg), Week 6 (Dose strengths 5 mg and 2.5 mg) and Week 8 (Dose strength 1 mg).
t 4 Week Post Treatment Follow Up Visit to be completed at the end of the Randomized Double-Masked study.
u If a subject presents at a site for an Unscheduled visit with symptoms related to a uveitis flare, subjects should be evaluated for treatment failure criteria. If treatment failure criteria are met, investigators should complete EOT procedures. If a subject presents at a site for an Unscheduled visit not related to a uveitis flare, investigators should complete Unscheduled visit procedures at their discretion.
v Tests obtained for standard of care purposes within 30 days of Day 1/Baseline may be used to satisfy eligibility criteria. Refer to list provided in Laboratory assessment table (protocol Appendix 6)

Appendix 3. Data Collection of COVID-19 Data

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

Data collection

A COVID-19 supplement to the eCRF Completion Guidelines (CCG) was provided by Clinical Data Management to instruct clinical trial sites with data entry expectations pertaining to scenarios related to the COVID-19 pandemic. If a visit was missed, sites were instructed to enter "Visit missed due to COVID-19" and if a in-person visit was conducted virtually, sites were instructed to enter "Virtual visit due to COVID-19".

Determination of Missed and Virtual Visits

Natural Language Processing (NLP) will be used to search the CRF comment fields to identify instances of "COVID-19", "Virtual", or synonyms (see Table 13-1). The search terms will be maintained in a global lookup and can be modified to tune the NLP model. Any comments with COVID-19 search terms, "Missed visit" or "Virtual visit" will be assigned 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 subjects and the same visit, if one record could be categorized as "Virtual Visit", all records associated with this subject and this visit will be categorized as "Virtual Visit"

iii) Otherwise result is missing

Table 13-1. Examples Search Terms for "COVID-19" and "Virtual" Used to Identify Missed/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

Appendix 4. Sample SAS Code for Analysis of Binary Endpoints

The following model statement will be used to construct the confidence interval for the binomial proportion difference described in Section 6.3.2:

```
proc freq data = test ;
   by avisitn;
   tables TRT01PN*aval_ / riskdiff (CORRECT);
   output out = f200vp_ci RISKDIFF;
   where TRT01PN = 1 | TRT01PN = 3;
run;
```

where the confidence interval for the difference of two independent binomial proportions is constructed based on the normal approximation (ie, the Wald method) with a continuity correction, to adjust for the difference between the normal approximation and the binomial distribution, which is a discrete distribution:

$$(\hat{p}_1 - \hat{p}_2) \pm \left(\frac{1}{2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right) + z_{\alpha/2} \times \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}\right)$$

The following model statement will be used to construct the confidence interval for the binomial proportion described in the Section 6.3.2:

```
proc freq data = test ;
   by avisitn TRT01PN;
   tables aval_ / BINOMIAL (CORRECT);
run;
```

the confidence interval for the binomial proportions is constructed based on the normal approximation to the binomial distribution with continuity correction:

$$\hat{p}_1 \pm \left(\frac{1}{2}\left(\frac{1}{n_1}\right) + z_{\alpha/2} \times \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}}\right)$$

Appendix 5. Sample SAS Code for Analysis of Time to Event Endpoints

In the following SAS example codes, *aval* is the number of weeks from the first dose date to the first clinical event or to the censoring date for subjects who don't have clinical events. *Cnsr* is a binary variable with 1 representing no clinical event observed and 0 representing a clinical event observed.

```
proc lifetest data=analysis;
        time aval*cnsr(1);
        strata strat1n strat2n/group=trt diff=all test=logrank;
run;
```

Below SAS example codes are provided for the the proportional hazards model:

```
proc phreg data=analysis;
        class trt (ref = "2");
        strata strat1n strat2n;
        model aval*cnsr(1)= trt /rl alpha=0.05 ties=efron;
        hazardratio trt01pn/alpha=0.05 diff=ref;
run;
```

CCI


CCI


CCI


CCI

CCI

[redacted]

CCI

CCI

GS-US-432-4097-SAP

ELECTRONIC SIGNATURES

Signed by	Meaning of Signature	Server Date (dd-MMM-yyyy hh:mm:ss)
PPD	Biostatistics eSigned	08-Oct-2021 07:09:13
PPD	Clinical Research eSigned	08-Oct-2021 19:04:54