STATISTICAL ANALYSIS PLAN

Study Title:	A Randomized, Double-Blinded, Placebo-Controlled, Multicenter, Phase 2 Proof-of-Concept Study to Evaluate Safety, Tolerability, and Efficacy of GS-9876 in Subjects with Active Rheumatoid Arthritis on Background Therapy with Methotrexate
Name of Test Drug:	GS-9876, Filgotinib
Study Number:	GS-US-379-1582
Protocol Version (Date):	Amendment 2 (11 July 2016)
Analysis Type:	Final Analysis
Analysis Plan Version:	1.0
Analysis Plan Date:	26 October 2017
Analysis Plan Author(s):	PPD

CONFIDENTIAL AND PROPRIETARY INFORMATION

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ACR	American College of Rheumatology
ACR 20/50/70	American College of Rheumatology 20/50/70% improvement
AE	adverse event
ALT	alanine aminotransferase
AST	aspartate aminotransferase
ATC	anatomical therapeutic chemical drug class
AUC	area under the curve
bDMARD	biological disease modifying anti-rheumatic drug
BLQ	below the limit of quantitation
BMI	body mass index
BPM	beats per minute
CDAI	clinical disease activity index
CI	confidence interval
CMH	Cochran-Mantel-Haenszel
CRF	case report form
CRP	C-reactive protein
CPK	creatinine phosphokinase
CSR	clinical study report
CTCAE	common terminology criteria for adverse events
DAS28	disease activity score for 28 joint count
ECG	electrocardiogram
eCRF	electronic case report form
FAS	full analysis set
Gilead	Gilead Sciences
HAQ-DI	health assessment questionnaire-disability index
HLGT	high-level group term
HLT	high-level term
ID	identification
IXRS	interactive web and mobile response system
JAK	janus kinase
LLOQ	lower limit of quantitation
LLT	lower-level term
LOCF	last observation carried forward
LOQ	limit of quantitation
MedDRA	medical dictionary for regulatory activities
MMRM	mixed model repeated measures
MTX	methotrexate
NRI	non-responder imputation

PhGA	physician's global assessment
PT	preferred term
PtGA	patient's global assessment
PK	pharmacokinetic
POC	proof-of-concept
PTM	placebo-to-match
Q1, Q3	first quartile, third quartile
QD	quaque die (each day)
QT	electrocardiographic interval between the beginning of the Q wave and termination of the T wave, representing the time for both ventricular depolarization and repolarization to occur
QTc	corrected QT
QTcF	corrected CT interval using the Fridericia formula
RA	rheumatoid arthritis
SAE	serious adverse event
SAP	statistical analysis plan
SD	standard deviation
SDAI	simplified disease activity index
SE	standard error
SJC	swollen joint count
SOC	system organ class
SYK	spleen tyrosine kinase
TE	treatment-emergent
TEAE	treatment-emergent adverse event
TFLs	tables, figures, and listings
TJC	tender joint count
ULN	upper limit of normal
VAS	visual analog scale
WHO	World Health Organization

PHARMACOKINETIC ABBREVIATIONS

AUC_{tau}	area under the concentration versus time curve over the dosing interval
C_{last}	last observed quantifiable concentration of the drug
C_{max}	maximum observed concentration of drug
C_{tau}	observed drug concentration at the end of the dosing interval
CL_{ss}/F	apparent oral clearance after administration of the drug: at steady state: $CL_{ss}/F = Dose/AUC_{tau}$, where "Dose" is the dose of the drug
$t_{1/2}$	estimate of the terminal elimination half-life of the drug, calculated by dividing the natural log of 2 by the terminal elimination rate constant (λ_z)
T_{last}	time (observed time point) of C_{last}
T_{max}	time (observed time point) of C_{max}
V_z/F	apparent volume of distribution of the drug
λ_z	terminal elimination rate constant, estimated by linear regression of the terminal elimination phase of the concentration of drug versus time curve

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-379-1582. This SAP is based on the study protocol amendment 2 dated 11 July 2016 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 effect of GS-9876 versus placebo for the treatment of signs and symptoms of rheumatoid arthritis (RA) in subjects with active RA as measured by change from baseline in disease activity score for 28 joint count using the C-reactive protein (CRP) (DAS28[CRP]) at Week 12

The secondary objectives of this study are as follows:

- To evaluate the safety and tolerability of GS-9876 in subjects with RA
- To explore the effect of GS-9876 on other disease-specific outcomes in RA

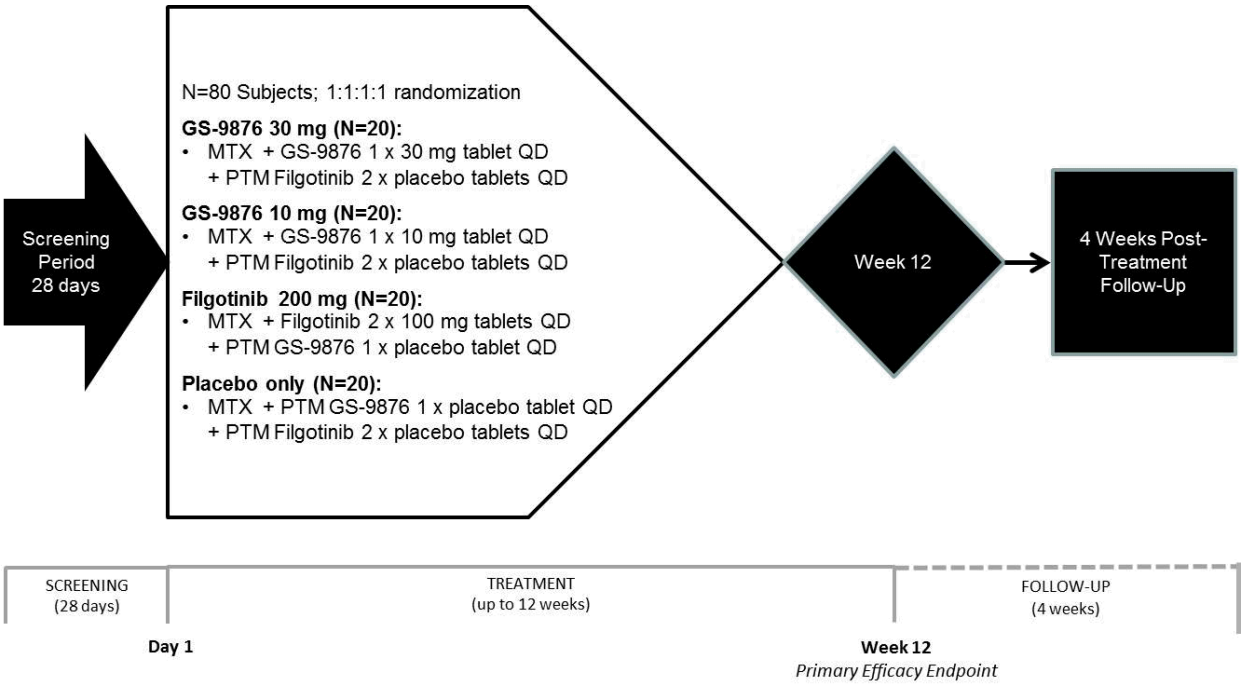
The exploratory objectives of this study are as follows:

1.2. Study Design

This is a randomized, double-blind, placebo-controlled, Phase 2 proof-of-concept (POC) study evaluating the safety, tolerability, and efficacy of GS-9876 in adult male and female subjects with active RA despite MTX therapy who had an inadequate response to MTX (either alone or in combination with biological disease modifying anti-rheumatic drug [bDMARD]). In addition, the effect of GS-9876 and filgotinib on disease and/or pathway markers relevant to RA and/or the SYK and JAK pathways will be assessed.

A total of approximately 80 subjects will be randomized in this study. The study design is shown in Figure 1-1.

Figure 1-1. Study Schema



Following completion of screening assessments, eligible subjects will be randomized in a blinded fashion in a 1:1:1:1 ratio as follows (all groups will continue to take MTX):

- **GS-9876 30 mg:** GS-9876 (1 x 30 mg tablet QD) + placebo-to-match (PTM) filgotinib (2 x placebo tablets QD) (N=20)
- **GS-9876 10 mg:** GS-9876 (1 x 10 mg tablet QD) + PTM filgotinib (2 x placebo tablets QD) (N=20)
- **Filgotinib 200 mg:** Filgotinib (2 x100 mg tablets QD) + PTM GS-9876 (1 x placebo tablets QD) (N=20)
- **Placebo only:** PTM GS-9876 (1 x placebo tablet QD) + PTM filgotinib (2 x placebo tablets QD) (N=20)

Randomization will be stratified by prior inadequate response to biologic therapy and geographic region.

Subjects will be dosed for up to 12 weeks and then followed for 4 weeks after their last dose of study drug.

1.3. Sample Size and Power

Sample sizes for GS-9876 and placebo groups are determined based on the superiority test of 1 dose of GS-9876 compared to placebo based on the change from baseline in DAS28(CRP) at Week 12. When assuming a difference of 1.2 between the 2 groups and a common standard deviation of 1.35, 20 subjects in each of the GS-9876 groups and 20 in the placebo group are required to obtain 78% power at a 2-sided 0.05-level. PPD . Therefore, the total sample size will be 80 (20 per treatment group).

2. TYPE OF ANALYSIS

2.1. Interim Analyses

Administrative interim analyses of efficacy and safety data were performed by the Gilead internal unblinded team. The purpose of the analyses was to assess efficacy of GS-9876 for further planning and development as a single agent and in combination. This unblinded team was independent of the GS-US-379-1582 study team who were not directly involved in routine study conducts. People who were unblinded were documented in the unblinded form GF-27011A, and the justification was documented in GT-27011A per SOP-BM-27011. The Study Team has remained blinded to treatment assignments throughout the trial until all subjects have completed the planned study visits and the database has been locked and unblinded.

2.2. Final Analysis

After all 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.

All statistical tests will be 2-sided and performed at the 5% significance level unless otherwise specified.

By-subject listings will be presented for all subjects in the All Randomized Analysis Set and sorted by treatment group, site/investigator identification (ID) number, 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 included in the listings, as well as age, sex at birth, race, and ethnicity.

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 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. The study drugs in this study are filgotinib, GS-9876, and PTMs for GS-9876 or filgotinib. This is the primary analysis set for efficacy analyses.

3.1.3. 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.

3.1.4. Pharmacokinetic Analysis Set

The pharmacokinetic (PK) analysis set includes all subjects in the Safety Analysis Set who have at least 1 nonmissing PK concentration data for GS-9876, filgotinib and/or their metabolites. This is the primary analysis set for general PK analyses.

3.1.5. Pharmacokinetic Substudy Analysis Set

The PK Substudy Analysis Set includes all subjects in the Safety Analysis Set who have enrolled into the PK Substudy, and have intensive PK concentration data to provide interpretable results for the specific parameters of interest for the analyte under evaluation. This is the primary analysis set for intensive PK analyses.

3.2. Subject Grouping

For analyses based on the All Randomized Analysis Set or 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.

For the PK Analysis Set, subjects will be grouped according to the actual treatment they received.

3.3. Strata and Covariates

Subjects will be randomly assigned to treatment groups via the IXRS in a 1:1:1:1 ratio using a stratified randomization schedule. Stratification will be based on the following variables:

- Geographic region (**Region 1** includes the United States; **Region 2** includes the following countries: Bulgaria, Ukraine, Czech Republic, Georgia, Hungary, Moldova, and Poland)
- Prior inadequate response to biologic therapy (Yes or No). In the eCRF data, subjects who discontinued prior RA biologic therapy before the first dose date of any study drug due to any of the following reasons are considered to have prior inadequate response to biologic therapy:
  - Primary inadequate response
  - Loss of response
  - Intolerance (Non-allergic)
  - Allergic response

Please refer to Appendix 2 for the RA biologic therapy list.

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

For efficacy endpoints, the baseline value of the efficacy variable(s) will be included as covariates in the efficacy analysis model, as specified in Section 6.

3.4. Examination of Subject Subgroups

Subgrouping of subjects based on stratification factors and Arensia sites will be explored for subgroup analyses.

The primary efficacy endpoint and the secondary endpoint of ACR20/50/70 will be examined using the following subgroups:

- Geographic region (Region 1, Region 2, as listed in Section 3.3)
- Prior inadequate response to biologic therapy (Yes, No, as defined in Section 3.3).
- Arensia sites and non-Arensia sites. The Arensia site numbers are 13632 and 13455.

If the number of subjects in a subgroup is not sufficient for modeling-based statistical analyses between treatment groups, decriptive statistics for each treatment group within the corresponding subgroup will be provided instead.

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. 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. The handling of missing data for PK is described in Section 8.

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 will be included in the data analyses.

3.6. Data Handling Conventions and Transformations

In general, age (in years) on the date of the first dose of study drug will be used for analyses and presentation in listings. If an enrolled subject was not dosed with any study drug, the randomization date will be used instead of the first dosing date of study drug. For screen failures,

the date the last informed consent was signed will be used for age calculation. If only birth year is collected on the case report form (CRF), "01 January" will be used for the unknown birth day and month for the purpose of age calculation. If only birth year and month are collected, "01" will be used for the unknown birth day.

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 for 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).

Natural logarithm transformation will be used for plasma/blood concentrations and analysis of PK parameters. 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 postbaseline time points.

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".

PK parameters that are BLQ will be imputed as one-half LOQ before log transformation or statistical model fitting.

3.7. Analysis Visit Windows

3.7.1. Definition of Study Day

The first dose date of individual study drug will be calculated separately for each study drug (ie, GS-9876, filgotinib, and PTMs) in a treatment group. Study Day 1 is defined as the first dose date of any study drug, which is the minimum of the first dose dates of individual study drugs in a treatment group.

The last dose date of individual study drug will be calculated separately for each study drug in a treatment group. The last dose date for an individual study drug will be the end date on study drug administration CRF for the record where the "study drug was permanently withdrawn" flag is "Yes". The last dose date of any study drug will be defined as the maximum of the last dose dates of individual study drugs in a treatment group.

Study Day will be calculated from the Study Day 1 and derived as follows:

- For postdose study days: Assessment Date – Study Day 1 + 1
- For days prior to the first dose: Assessment Date – Study Day 1

3.7.2. Analysis Visit Windows

Subject visits may occur on days not specified in the protocol (eg, unscheduled visits). 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.

Table 3-1 Analysis Visit Windows for Joint Count Assessment, HAQ-DI, PhGA, PtGA, and On-Treatment CRP, Safety Laboratory Data, Weight, Vital Signs, ECG

Nominal Visit	Nominal Study Day	Lower Limit	Upper Limit
Baseline	1	(none)	1
Week 2	15	2	22
Week 4	29	23	43
Week 8	57	44	71
Week 12	85	72	≥ 85

The analysis windows for joint count assessment, health assessment questionnaire - disability index (HAQ-DI) (including the patient's assessment of pain), physician's global assessment of disease activity (PhGA), patient's global assessment of disease activity (PtGA), on-treatment CRP, safety laboratory, weight, vital signs and electrocardiogram (ECG) are provided in Table 3-1. CRP, safety laboratory, weight, vital signs, and ECG data collected up to the last dose date of any study drug + 3 days are considered to be on-treatment data.

CRP, safety laboratory, weight, vital signs, and ECG data collected after last dose date of any study drug + 3 days will be considered as posttreatment data and will be summarized up to last dose date plus 30 days, and labeled as "FU-Week 4". Data obtained after last dose date plus 30 days will be excluded from the summaries, but will be included in the listings.

3.7.3. Selection of Non-Efficacy 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:
  - 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 value with the worst severity will be taken (eg, abnormal will be selected over normal for safety ECG findings), unless otherwise specified.

4. SUBJECT DISPOSITION

4.1. Subject Enrollment and Disposition

A summary of subject enrollment will be provided by treatment group for each country within each geographic region and investigator within a country. 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 stratification factor 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 IXRS 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 IXRS 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 not randomized, 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
- Full Analysis Set
- PK Analysis Set
- PK Substudy Analysis Set
- Completed study drug
- Did not complete study drug with reasons for premature discontinuation of study drug
- Continuing study (if applicable)
- 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. The denominator for the percentage calculation will be the total number of subjects in the Safety Analysis Set corresponding to that column. In addition, a flowchart will be provided to depict the subject disposition.

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

- Reasons for premature study drug or study discontinuation
- Reasons for screen failure (will be provided by screening ID number in ascending order)
- Lot number and kit 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 dose date of any study drug - first dose date of any study drug + 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, weight and vital signs assessment date that occurred during the on-treatment period will be used.

The total duration of exposure to any 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 2 (Day 15), Week 4 (Day 29), Week 8 (Day 57), and Week 12 (Day 85).

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

Adherence will be calculated separately for GS-9876/PTM (tablets) and filgotinib/PTM (tablets).

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 CRF using the following formula:

Total Number of Tablets Adminstered =
(Σ No. of Tablets Dispensed) - (Σ No. of Tablets Returned)

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 administered will be counted as zero.

4.2.2.1. On-Treatment Adherence

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:

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

Note: If calculated adherence is greater than 100%, the result will be set to 100%.

Study drug expected to be administered for GS-9876/PTM (tablets) = 1 tablet × total duration of exposure to any study drug (days).

Study drug expected to be administered for filgotinib/PTM (tablets) = 2 tablets × total duration of exposure to any 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.

Categorical displays will be presented for the number of subjects who are at least 80% adherent to their study drug regimen (ie, adherence is ≥ 80% for each study drug).

No formal statistical testing is planned.

A by-subject listing of study drug administration and drug accountability will be provided separately by treatment group, 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. 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 by deviation reason 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.

5. BASELINE CHARACTERISTICS

5.1. Demographics

Subject demographic variables 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 data will be provided for the Safety Analysis Set for the following:

- age (on the first dose date of any study drug) as a continuous variable
- sex at birth (male, female)
- race
- ethnicity (Hispanic/Latino, not Hispanic/Latino)

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

5.2. Other Baseline Characteristics

Other baseline characteristics include:

- prior inadequate response to biologic therapy and number of prior biologics used
- geographic region
- concurrent use and weekly dose of methotrexate (refer to Appendix 2 for methotrexate medication list)
- concurrent use of hydroxychloroquine (refer to Appendix 2 for hydroxychloroquine medication list)
- concurrent use of corticosteroid (refer to Appendix 3 for the definition of corticosteroid medications)
- height (cm)
- weight (kg)
- body mass index (BMI; in kg/m^2)
- baseline swollen joint count based on 66 joints (SJC66) and tender joint count based on 68 joints (TJC68)

- baseline swollen joint count based on 28 joints (SJC28) and tender joint count based on 28 joints (TJC28)
- baseline HAQ-DI
- baseline patient's pain assessment
- baseline DAS28(CRP)
- baseline patient's global assessment of disease activity
- baseline physician's global assessment of disease activity
- baseline Simplified Disease Activity Index (SDAI) and Clinical Disease Activity Index (CDAI)
- baseline CRP (mg/L)
- duration of RA (years)

duration of RA (years) = ((first dose date) – (date of initial diagnosis) + 1 day) / 365.25. If the date of initial diagnosis is incomplete, then the following rules will be applied:

  - Missing day: use the first of the month.
  - Missing month: use January.

These baseline characteristics 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 baseline characteristics will be provided for the Safety Analysis Set.

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

5.3. Medical History

Medical history collected at screening will not be coded. Medical history will be summarized by treatment groups. Numbers and percentages of subjects who experienced myocardial infarction before the age of 50 years, experienced stroke before the age of 50 years, and have diabetes (type I or II) will be provided. Reported medical condition terms of subject having any relevant medical history, including any medication allergies, will be presented by descending order of the total frequencies. Same reported terms will be counted only once in the summary. The summary will be provided for the Safety Analysis Set. No formal statistical testing is planned.

A by-subject listing of disease-specific medical history will be provided by treatment group and subject ID number (in ascending order) and medical history of abnormalities (in chronological order).

6. EFFICACY ANALYSES

6.1. General Considerations

The primary analysis set for efficacy analyses will be the FAS, defined in Section 3.1.2.

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

- If record collection date/time is known, latest record will be taken
- If time is unknown, the worst scenario will be selected

For the composite endpoints including DAS28(CRP), ACR 20/50/70, SDAI, and CDAI, we use the following steps unless otherwise specified:

- Step 1: Put each component value into a visit window
- Step 2: Within each visit window, select the component data point for analysis
- Step 3: Combine these selected data points within each visit window across the components.

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 above 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 considered as missing.
- Last observation carried forward (LOCF): For the above composite endpoints, First LOCF all components over all visits, then combine the components as mentioned above.
- Non-responder imputation (NRI): For all binary response measurements, starting from OC, all missings will be treated as non-responders.

6.1.1. Tender/Swollen Joint Counts (TJC/SJC)

TJC68 and SJC66 will be collected during the course of the study. The assessment for each joint will be from the following selections: Tender Only, Swollen Only, Tender and Swollen, Not Tender or Swollen, or Joint Non-Evaluable or Missing.

Individual joint with missing assessment will not be imputed. If at least half of the joints are assessed at a given visit, the *prorated* tender and swollen joint counts will be calculated using the following formula:

$$TJC68 = \frac{Total\ number\ of\ tender\ joints}{68 - (Number\ of\ nonevaluable\ or\ missing\ joints\ out\ of\ 68\ joints)} \times 68$$

$$SJC66 = \frac{Total\ number\ of\ swollen\ joints}{66 - (Number\ of\ nonevaluable\ or\ missing\ joints\ out\ of\ 66\ joints)} \times 66$$

If less than half of joints are assessed at a given visit, joint counts are treated as missing for the visit.

A more abbreviated assessment considering 28 joints as listed in Table 6-1 for both tenderness and swelling will also be conducted (as part of the TJC68 and SJC66 assessment), denoted as TJC28 and SJC28, respectively.

Table 6-1. Composition of the 28 Joints

Joints	Number
Shoulder Joints (Left and Right)	2
Elbow Joints (Left and Right)	2
Wrist Joints (Left and Right)	2
Metacarpophalangeal Joints I-V (Left and Right)	10
Proximal Interphalangeal Joints I-V (Left and Right)	10
Knee Joints (Left and Right)	2

If there are non-evaluable or missing joints among the 28 joints but at least half of the joints are assessed at a given visit, similar prorated tender and swollen joint counts will be calculated as follows:

$$TJC28 = \frac{Total\ number\ of\ tender\ joints}{28 - (Number\ of\ nonevaluable\ or\ missing\ joints\ out\ of\ 28\ joints)} \times 28$$

$$SJC28 = \frac{Total\ number\ of\ swollen\ joints}{28 - (Number\ of\ nonevaluable\ or\ missing\ joints\ out\ of\ 28\ joints)} \times 28$$

If less than half of the 28 joints are assessed at a given visit, TJC28 and SJC28 are treated as missing for the visit.

6.1.2. Global Assessment of Disease Activity

The Patient's Global Assessment of Disease Activity is a horizontal (0 – 100) visual analog scale (VAS) that ranges from "No Arthritis" to "Severe Arthritis."

The Physician's Global Assessment of Disease Activity is a horizontal (0 – 100) VAS that ranges from "No Disease Activity" to "Maximum Disease Activity."

6.2. Primary Efficacy Endpoints

6.2.1. Definition of the Primary Efficacy Endpoints

The primary endpoint for the study is change from baseline in DAS28(CRP) at Week 12. The DAS28(CRP) score is calculated as follows:

$$\mathrm{DAS28(CRP)} = 0.56\sqrt{\mathrm{TJC28}} + 0.28\sqrt{\mathrm{SJC28}} + 0.36\ln(\mathrm{CRP}+1) + 0.014 \times \mathrm{PtGA} + 0.96,$$

where

TJC28 = number of joints tender out of 28 joints

SJC28 = number of joints swollen out of 28 joints

CRP = CRP measurement in unit of *mg/L*

PtGA = patient's global assessment of disease activity on a 0-100 VAS

Higher DAS28(CRP) value indicates more severe disease activity.

DAS28(CRP) will be calculated only for the time points with all measurements available.

6.2.2. Statistical Hypotheses for the Primary Efficacy Endpoints

The primary analysis will consist of a superiority test at the 2-sided 0.05-level of each of the GS-9876 doses compared to placebo based on the change from baseline in DAS28(CRP) at Week 12. There is no adjustment for multiple testing for this study.

The mean change from baseline in DAS28(CRP) at Week 12 for treatment groups GS-9876 10 mg, GS-9876 30 mg and placebo group are denoted as μ_1, μ_2 and μ_P, respectively.

The primary statistical hypotheses for the superiority tests are as follows:

H_0: $\mu_1 = \mu_P$ vs H_1: $\mu_1 \neq \mu_P$

and

H_0: $\mu_2 = \mu_P$ vs H_1: $\mu_2 \neq \mu_P$

6.2.3. Analysis of the Primary Efficacy Endpoint

Change from baseline in DAS28(CRP) at Week 12 will be analyzed using a mixed-effects model for repeated measures (MMRM) approach. The model will include the fixed effects of treatment, visit, treatment by visit interaction, and baseline DAS28(CRP) value. Subjects will be included as a random effect. Missing change scores (eg, due to early withdrawal, etc.) will not be otherwise imputed using the MMRM approach. An unstructured variance-covariance matrix for the repeated measures and Kenward-Roger approximation for the denominator degrees of freedom will be used in the mixed-effects model.

From the MMRM approach, estimated differences in mean change from baseline between each of the GS-9876 doses and the placebo group will be presented with 95% confidence intervals (CIs) and p-values. PPD

Change from baseline in DAS28(CRP) will be summarized by treatment and visit using descriptive statistics (sample size, mean, SD, median, Q1, Q3, minimum, and maximum). Plots of mean ± SD of DAS28(CRP) values and changes from baseline in DAS28(CRP) over time will be presented.

6.2.4. Sensitivity Analysis of the Primary Efficacy Endpoints

Sensitivity analyses of the primary efficacy endpoint will not be performed.

6.2.5. Subgroup Analysis of the Primary Efficacy Endpoint

Subgroup analysis will be performed for the primary efficacy endpoint for the subgroups outlined in Section 3.4. The difference between treatment groups and the corresponding 95% CIs will be presented.

6.3. Secondary Efficacy Endpoints

The secondary efficacy endpoints are:

- The proportion of subjects who achieve ACR 20/50/70 at Week 12
- Change from baseline in HAQ-DI score at Week 12

6.3.1. Definition of Secondary Efficacy Endpoints

6.3.1.1. ACR 20/50/70

A subject achieves ACR20 response when this subject has

- ≥ 20% improvement from baseline in the tender joint count (TJC68), AND
- ≥ 20% improvement from baseline in the swollen joint count (SJC66), AND

- ≥ 20% improvement from baseline in at least 3 of the following 5 items:

  1) Physician's Global Assessment of Disease Activity (PhGA),

  2) Patient's Global Assessment of Disease Activity (PtGA),

  3) Patient's pain assessment,

  4) Patient's assessment of physical function (HAQ-DI) score,

  5) C-reactive protein (CRP)

Percent improvement from baseline at a post-baseline visit is calculated as follows for all 7 components mentioned above:

([baseline value – post-baseline value] / baseline value) * 100%

If the baseline value is 0 and the post-baseline value is also 0, then the percent improvement from baseline is set to 0. If the baseline value is 0 and the post-baseline value is non-zero then the percent improvement from baseline is set to missing.

In the case that some ACR20 components are missing, the ACR20 assessment will be derived based on the non-missing components. If non-missing components are not sufficient to determine ACR20 response, then the ACR20 response will be considered as missing.

ACR50 and ACR70 are similarly defined as ACR20 above, except the improvement threshold from baseline is 50% and 70%, respectively.

6.3.1.2. Health Assessment Questionnaire Disability Index (HAQ-DI)

The HAQ-DI score is defined as the average of the scores of eight functional categories (dressing and grooming, arising, eating, walking, hygiene, reach, grip, and other activities), administered by the patient. Responses in each functional category are collected as: without any difficulty; with some difficulty; with much difficulty; unable to do a task in that area; and with or without aids or devices. The HAQ-DI score ranges from 0 (no disability) to 3 (completely disabled), when 6 or more categories are non-missing. Detailed algorithm for calculating HAQ-DI score is described in Appendix 1.

The HAQ-DI also includes a separate pain assessment and the patient will be requested to mark the severity of the pain in the past week on a 0-100 VAS, with 0 indicating "no pain" and 100 indicating "severe pain."

6.3.2. Analysis Methods for Secondary Efficacy Endpoints

The ACR20 response rates at Week 12 between each of the GS-9876 dosed groups and the placebo group will be compared using a stratified Cochran-Mantel-Haenszel (CMH) chi-square test adjusting for stratification factors in randomization (ie, prior inadequate response to biologic

therapy, yes vs. no; and Geographic region, region 1 vs. region 2 as defined in Section 3.3). The difference in response rates between treatment groups and the corresponding 95% CIs will be presented. The CMH chi-square p-value for testing the superiority of each of the GS-9876 doses as compared to placebo will be provided, as well as the 2-sided 95% exact CI of the ACR20 response rate based on Clopper-Pearson method for each treatment group. Subjects who do not have sufficient measurements to establish efficacy at Week 12 will be considered as failures (ie, non-responder imputation [NRI]). Sensitivity analysis will be conducted by imputing missing components for determining ACR20 as well as early dropouts via the last observation carried forward (LOCF) method.

For ACR50 and ACR70 at Week 12, the same CMH approach with NRI as in the analysis of ACR20 described above will be adopted.

The difference in change from baseline in HAQ-DI between each of the GS-9876 dosed groups and the placebo group will be analyzed using a similar MMRM approach as described in Section 6.2.3. The model will include the fixed effects of treatment, visit, treatment by visit interaction, and baseline value. Subjects will be included as a random effect. Missing change scores (eg, due to early withdrawal, etc) will not be otherwise imputed using the MMRM approach.

PPD

PPD

Mean ACR 20/50/70 response rates with NRI will be plotted by treatment and visit. Mean ± SD over time will be plotted for each component of ACR 20/50/70.

6.4. Exploratory Efficacy Endpoints

PPD

6.4.1. Definition of Exploratory Efficacy Endpoints

PPD

[redacted]

6.4.2. Analysis Methods for Exploratory Efficacy Endpoints

PPD

[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 MedDRA v 20.0. 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 definition of an SAE specified in the study protocol. SAEs captured and stored in the clinical database will be reconciled with the SAE database from the Gilead Drug Safety and Public Health (DSPH) 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

The number and percentage of subjects who experienced 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
- All Treatment Emergent (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
- All TE SAEs
- All TE treatment-related SAEs
- All TEAEs leading to premature discontinuation of any study drug
- All TEAEs leading to premature discontinuation of study
- All TEAEs leading to death (ie, outcome of death)
- All TEAEs leading to temporary interruption of any study drug

A brief, high-level summary of AEs described above will be provided by treatment group and by the number and percentage of subjects who experienced 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 to the above summary tables, all TEAEs and TE treatment-related AEs 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
- SAEs
- Deaths
- AEs leading to death (ie, outcome of death)
- AEs leading to premature discontinuation of any study drug
- AEs leading to premature discontinuation of study
- AEs leading to temporary interruption of any study drug

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. Hemolized test results 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 visit in chronological order for hematology, serum chemistry, coagulation, serology, lipid panel, and urinalysis 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.

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 selected laboratory tests 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 3.7.3.

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 5) 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: > 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

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 with AST or ALT > 3 x ULN will be provided.

7.2.4. Shifts Relative to the Baseline Value

Shift tables will be presented by showing change in severity grade from baseline to postbaseline visits for the following laboratory tests:

- Complete blood count components
- Creatinine clearance (estimated glomerular filtration rate)
- Liver function tests
- Creatinine phosphokinase (CPK)
- CRP

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 3.7.3. No formal statistical testing is planned.

A by-subject listing of vital signs (systolic and diastolic blood pressure [mmHg], pulse [beats/min], respiration [breaths/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 Q22016.

All the analyses in this section will be performed for general prior /concomitant medications and RA-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 Anatomical Therapeutic Chemical (ATC) drug class Level 4 and 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 ATC medical class and then by preferred name in order of descending overall frequency within each ATC medical class. 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 ATC drug class Level 4 and 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 ATC medical class and then by preferred name in descending overall frequency within each ATC medical class. 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 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

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); or missing. 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 treatment group and subject ID number and visit in chronological order.

7.5.2. Corrected QT Intervals

The QT interval (measured in millisecond [msec]) is a measure of the time between the start of the Q wave and the end of the T wave in the heart's electrical cycle. The QT interval represents electrical depolarization and repolarization of the ventricles. The QT interval is affected by heart rate, and a number of methods have been proposed to correct QT for heart rate.

Corrected QT (QTc) intervals will be derived using Fridericia's correction (QTcF) as follows:

$$QTcF = \frac{QT}{\sqrt[3]{RR}}$$

where QT is measured in msec; RR = 60/Heart Rate (beats per minute [bpm]) and RR is measured in seconds

QTcF and uncorrected QT values at each visit and change from baseline at each visit will be summarized for the Safety Analysis Set by treatment group using descriptive statistics.

7.5.3. PR and QRS Intervals

The PR interval (measured in msec) is a measure of the time between the start of the P wave (the onset of atrial depolarization) and the beginning of the QRS complex (the onset of ventricular depolarization). The QRS interval measures the duration of the QRS complex.

VR, PR, RR, and QRS values at each visit and change from baseline at each visit will be summarized for the Safety Analysis Set by treatment group using descriptive statistics.

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

Plasma concentrations of GS-9876 and filgotinib and the active metabolite of filgotinib, GS-829845, will be listed and summarized using descriptive statistics (eg, sample size, arithmetic mean, geometric mean, % coefficient of variation, standard deviation, median, minimum and maximum).

8.1. PK Analyses Related to Intensive PK Sampling

Steady-state PK over a 24 hour dosing interval will be determined in subjects in the PK Substudy Analysis Set. Concentrations of GS-9876, filgotinib, and GS-829845 in plasma will be determined using validated bioanalytical assays.

8.1.1. Estimation of Pharmacokinetic Parameters

PK parameters will be estimated using Phoenix WinNonlin® software using standard noncompartmental methods. The linear/log trapezoidal rule will be used in conjunction with the appropriate noncompartmental model, with input values for dose level, dosing time, plasma concentration, and corresponding real-time values, based on drug dosing times whenever possible.

All predose sample times before time-zero will be converted to 0.

For area under the curve (AUC), samples BLQ of the bioanalytical assays occurring prior to the achievement of the first quantifiable concentration will be assigned a concentration value of 0 to prevent overestimation of the initial AUC. Samples that are BLQ at all other time points will be treated as missing data in WinNonlin. The nominal time point for a key event or dosing interval (τ) may be used to permit direct calculation of AUC over specific time intervals. The appropriateness of this approach will be assessed by the PK scientist on a profile-by-profile basis.

Pharmacokinetic parameters such as AUC_{tau}, λ_z and $t_{1/2}$ are dependent on an accurate estimation of the terminal elimination phase of drug. The appropriateness of calculating these parameters will be evaluated upon inspection of PK data on a profile-by-profile basis by the PK scientist.

8.1.2. Pharmacokinetic Parameters

PK parameters will be generated for all subjects in the PK Substudy analysis set. The analytes presented in Table 8-1 will be evaluated if data are available.

Table 8-1. Study Treatments and Associated Analytes

Treatment	Analyte
GS-9876 10 mg	GS-9876
GS-9876 30 mg	GS-9876
Filgotinib 200 mg	Filgotinib, GS-829845

The analytes and parameters presented in Table 8-2 will be used to evaluate the PK objectives of the study. The PK parameters to be estimated in this study are listed and defined in the Pharmacokinetic Abbreviations section.

Table 8-2. Pharmacokinetic Parameters for Each Analyte

Analyte	Parameters
GS-9876	AUC_{tau}, C_{tau}, C_{max}, T_{max}, C_{last}, T_{last}, λ_z, CL_{ss}/F, V_z/F, and $t_{1/2}$, if appropriate
Filgotinib	AUC_{tau}, C_{tau}, C_{max}, T_{max}, C_{last}, T_{last}, λ_z, CL_{ss}/F, V_z/F, and $t_{1/2}$, if appropriate
GS-829845	AUC_{tau}, C_{tau}, C_{max}, T_{max}, C_{last}, T_{last}, λ_z, and $t_{1/2}$, if appropriate

8.1.3. Statistical Analysis Methods

Individual subject concentration data and individual subject PK parameters for GS-9876, filgotinib, and GS-829845 will be listed and summarized using descriptive statistics by treatment. Summary statistics (number [n], mean, SD, coefficient of variation [%CV], median, min, max, Q1, and Q3) will be presented for both individual subject concentration data by time point and individual subject PK parameters by treatment. Moreover, the geometric mean, 95% CI, and the mean and SD of the natural log-transformed values will be presented for individual subject PK parameter data.

Individual concentration data listings and summaries will include all subjects with concentration data. The sample size for each time point will be based on the number of subjects with nonmissing concentration data at that time point. The number of subjects with concentration BLQ will be presented for each time point. For summary statistics, BLQ values will be treated as 0 at predose and one-half of the lower limits of quantitation (LLOQ) for postdose time points.

Individual PK parameter data listings and summaries will include all subjects for whom PK parameter(s) can be derived. The sample size for each PK parameter will be based on the number of subjects with nonmissing data for that PK parameter.

The following tables may be provided for each analyte by treatment:

- Individual subject concentration data and summary statistics
- Individual subject plasma PK parameters and summary statistics.

The following figures may be provided for each analyte by treatment:

- Individual subject concentration data versus time (on linear and semilogarithmic scales)
- Mean (± SD) concentration data versus time (on linear and semilogarithmic scales)
- Median (Q1, Q3) concentration data versus time (on linear and semilogarithmic scales)

Individual, mean, and median postdose concentration values that are ≤ LLOQ will not be displayed in the figures and remaining points connected.

The following listings will be provided:

- PK sampling details (and PK concentrations) by subject, including procedures, differences in scheduled and actual draw times, and sample age
- Individual data on determination of plasma half-life and corresponding regression correlation coefficient.

9. BIOMARKER ANALYSIS

A separate biomarker analysis plan will document methods to analyze biomarker assessments.

10. REFERENCES

Aletaha D, Nell VP, Stamm T, Uffmann M, Pflugbeil S, Machold K, et al. Acute phase reactants add little to composite disease activity indices for rheumatoid arthritis: validation of a clinical activity score. Arthritis research & therapy 2005;7 (4):R796-806.

11. SOFTWARE

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

nQuery + n Terim Version 4.0. Statistical Solutions, Cork, Ireland.

12. SAP REVISION

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

13. APPENDIX

Appendix 1. Health Assessment Questionnaire Disability Index (HAQ-DI)

The HAQ-DI score is defined as the average of the scores of eight functional categories (dressing and grooming, arising, eating, walking, hygiene, reach, grip, and other activities), usually administered by the patient. Responses in each functional category are collected as 0 (without any difficulty) to 3 (unable to do a task in that area), with or without aids or devices.

The highest score for questions in each category (range 0 to 3) determines the score for the category, unless aids or devices are required. Dependence on equipment or physical assistance increases a lower score (ie, scores of 0 or 1) to the level of 2 to more accurately represent underlying disability. The eight category scores are averaged into an overall HAQ-DI score on a scale from 0 (no disability) to 3 (completely disabled) when 6 or more categories are non-missing. If more than 2 categories are missing, the HAQ-DI score is set to missing. The HAQ-DI can be treated as a continuous measure.

The HAQ-DI score using aids (and/or devices) is computed by taking the maximum score of the questions in each category (range: [0, 3]) and whether or not aids/devices are used (0 or 1):

A = max(dressing & grooming category questions, 2*aids indicator) +
max(rising category questions, 2*aids indicator) +
max(eating category questions, 2*aids indicator) +
max(walking category questions, 2*aids indicator) +
max(hygiene category questions, 2*aids indicator) +
max(reach category questions, 2*aids indicator) +
max(grip category questions, 2*aids indicator) +
max(usual activities category questions, 2*aids indicator)

HAQ-DI = A/(total number of categories with at least 6 non-missing)

The following table shows the contribution of the 43 questions used to calculate the HAQ-DI:

HAQ-DI Category:	Category questions: At least 6 categories must have scores to compute the HAQ-DI.		HAQ-DI Category Score with Aids/Devices Calculation:
	Category Questions	Aids/Devices Indicators	
Dressing / Grooming	HAQ0101, HAQ0102 (DRESS, HAIR)	HAQ0114, HAQ0119 (DRSG, GROOM)	Using each question with a scale of 0-3, calculate the category score as the maximum of the category questions. If the Aids/Devices indicator is "No", no need to adjust the category score. If the Aids/Devices indicator is "Yes" and the category score is <2, then the category score with the Aids/Devices is set to 2. If the Aids/Devices indicator is "Yes" and the category score is ≥2, then the category score with Aids/Devices is the calculated category score without adjustment. For example: The Dressing/Grooming category score is 2 if subject answered 1 for both questions 1 and 2 and "Yes" for both question 14 and 18. In the HAQ-DI score calculation, questions on other device/aids will not be used.
Arising	HAQ0103, HAQ0104 (STAND, BED)	HAQ0116, HAQ0120 (CHAIR, ARISING)	
Eating	HAQ0105,HAQ0106, HAQ0107 (MEAT, LIFT,MILK)	HAQ0115, HAQ0121 (UTENSIL, EAT)	
Walking	HAQ0108, HAQ0109 (WALK, STEPS)	HAQ0110, HAQ0111, HAQ0112, HAQ0113, HAQ0122 (CANE, WALKER, CRUTCH, WHEEL, WALKING)	
Hygiene	HAQ0123, HAQ0124, HAQ0125 (WASH, BATH, TOILET)	HAQ0134, HAQ0135, HAQ0137, HAQ0139, HAQ0142 (RAISEAT, BATHBAR, BATHSEAT, LONGBATH, HYGIENE)	
Reach	HAQ0126, HAQ0127 (REACH, BEND)	HAQ0138, HAQ0143 (LONGRCH, REACH)	
Grip	HAQ0128, HAQ0129, HAQ0130 (OPENCAR, JAR, FAUCET)	HAQ0136, HAQ0144 (JAROPEN, GRIP)	
Activity	HAQ0131, HAQ0132, HAQ0133 (SHOP, INCAR, CHORES)	HAQ0145 (ERRAND)	

Handling Missing Data: If no more than 2 categories have missing category scores, then the HAQ-DI is the mean of the non-missing category scores. Otherwise, the HAQ-DI score is set to missing.

If any of the category questions are missing, but the aids/device indicator is non-missing, the category score can still be computed. However, if all category questions and its aids/device indicators are missing, then the category score is considered missing.

Appendix 2. RA Medication List

WHO-DDE PREFERRED TERM	Biologic	Hydroxychloroquine	Methotrexate
ABATACEPT	y		
ADALIMUMAB	y		
ANAKINRA	y		
BIMEKIZUMAB	y		
CERTOLIZUMAB	y		
CERTOLIZUMAB PEGOL	y		
ETANERCEPT	y		
GOLIMUMAB	y		
HYDROXYCHLOROQUINE		y	
HYDROXYCHLOROQUINE SULFATE		y	
IMMUNOSUPPRESSANTS	y		
INFLIXIMAB	y		
METHOTREXATE			y
METHOTREXATE SODIUM			y
MONOCLONAL ANTIBODIES	y		
NORFLOX-TZ	y		
SARILUMAB	y		
SIRUKUMAB	y		
SULFASALAZINE			
TOCILIZUMAB	y		
TUMOR NECROSIS FACTOR ALPHA	y		
INVESTIGATIONAL DRUG (note: reported term is METHOTREXATE/PLACEBO)			y

Appendix 3. Corticosteroid Medications

Corticosteriod medications were searched in all the concomitant medication database using WHO ATC drug class Level 2 term “CORTICOSTEROIDS FOR SYSTEMIC USE”.