STATISTICAL ANALYSIS PLAN

A Phase 2, Multicenter, Randomized, Double-Blind, Placebo-Controlled, Parallel Dose Study to Evaluate the Efficacy and Safety of Oral SKI-O-703 in Patients With Active Rheumatoid Arthritis Despite Treatment With Conventional Therapies

STUDY DRUG:	**SKI-O-703**
PROTOCOL NUMBER:	**OSCO-P2201**
Version of SAP:	**Final Version 1.0**
Issued Date:	**14 OCT 2020**

Prepared by:

[redacted]

On behalf of Oscotec Inc.

Korea Bio-Park, Building A, 9th Floor
700 Daewangpangyo-ro, Bundang-gu
Seongnam-si, Gyeonggi-do 13488
The Republic of Korea

TABLE OF CONTENTS

List of Abbreviations

Abbreviation	Definition
ACR	American College of Rheumatology
AE	adverse event
ANCOVA	analysis of covariance
ATC	Anatomical Therapeutic Chemical
AUC	area under the concentration versus time curve
$AUC_{0\text{-}tau}$	area under the plasma concentration-time curve within a dosing interval
AUEC	area under the effect versus time curve
$AUEC_{0\text{-}tau}$	area under the effect versus time curve within a dosing interval
BID	twice daily
BLQ	below the limit of quantification
BMI	body mass index
CFB	change from baseline
CI	confidence interval
C_{max}	maximum observed concentration
CRF	case report form
csDMARDs	conventional synthetic disease-modifying antirheumatic drugs
CSR	clinical study report
CTCAE	Common Terminology Criteria for Adverse Events
C_{trough}	trough concentration observed at steady state
CV	coefficient of variation
DAS	disease activity score
ECG	electrocardiogram
E_{max}	maximum effect
EOS	end-of-study
EOT	end-of-treatment
HAQ-DI	health assessment questionnaire-disability index
hsCRP	high sensitivity C-reactive protein
ICF	informed consent form
ITT	intent-to-treat
IxRS	interactive voice or web response system
LOCF	last observation carried forward
LOF	lack of fit
LS	least-square
MCP	metacarpophalangeal
MedDRA	Medical Dictionary for Regulatory Activities
MAR	missing at random

Abbreviation	Definition
MCMC	Markov Chain Monte Carlo
MI	multiple imputation
mITT	modified intent-to-treat
MNAR	missing not at random
MMRM	Mixed model repeated measure
Npts	number of data points
PD	pharmacodynamic(s)
PIP	proximal interphalangeal
PK	pharmacokinetic(s)
PP	per protocol
PT	preferred term
PtGA	subject's global assessment for disease activity
RA	rheumatoid arthritis
R_{AUC}	accumulation ratio
R_{met}	metabolite ratio
Rsq (r^2)	coefficient of determination
SAE	serious adverse event
SAP	statistical analysis plan
SD	standard deviation
SE	standard error
SJC	swollen joint count
SOC	system organ class
$t_{1/2}$	apparent terminal elimination half-life
$t_{1/2,eff}$	effective half-life
TEAE	treatment-emergent adverse event
TE_{max}	Time to achieve maximum effect
TJC	tender joint count
T_{max}	time to reach maximum observed concentration
TNFα	tumor necrosis factor alpha
VAS	visual analog scale
WHODD	World Health Organization Drug Dictionary

1. Introduction

This Statistical Analysis Plan (SAP) describes the analyses and data presentations for Oscotec's protocol OSCO-P2201 "A Phase 2, Multicenter, Randomized, Double-Blind, Placebo-Controlled, Parallel Dose Study to Evaluate the Efficacy and Safety of Oral SKI-O-703 in Patients With Active Rheumatoid Arthritis Despite Treatment With Conventional Therapies" which was originally issued on 20 Dec 2017, and subsequently amended on 19 Oct 2020 (Version 5.0; Amendment 4.0). It contains definitions of analysis populations, derived variables and statistical methods for the analysis of efficacy, safety, pharmacokinetics (PK) and pharmacodynamics (PD).

There is only one final analysis for this study and no interim analyses are planned. Throughout this SAP, the treatment groups will be referred to as the 100 mg twice daily (BID) group, the 200 mg BID group, the 400 mg BID group and the placebo group. The purpose of the SAP is to ensure the credibility of the study findings by pre-specifying the statistical approaches to any data analysis prior to database lock. This SAP will be finalized and signed prior to the final database lock.

All statistical analyses detailed in this SAP will be conducted using SAS® 9.3 or later (SAS Institute Inc., Cary, North Carolina).

2. Objectives

2.1. Primary Objective

The primary objective of this study is to evaluate the efficacy of select (100 mg BID, 200 mg BID, and 400 mg BID) doses of SKI-O-703 compared with placebo in subjects with active RA who have had an inadequate response to conventional synthetic disease modifying antirheumatic drugs (csDMARDs) or anti-TNFα biologic agents.

2.2. Secondary Objective

The secondary objectives of this study are as follows:

- To evaluate the efficacy on other clinical endpoints of select (100 mg BID, 200 mg BID, and 400 mg BID) doses of SKI-O-703 compared with placebo in subjects with active RA who have had an inadequate response to csDMARDs or anti-TNFα biologic agents.
- To evaluate the safety and tolerability of select (100 mg BID, 200 mg BID, and 400 mg BID) doses of SKI-O-703 compared with placebo in subjects with active RA who have had an inadequate response to csDMARDs or anti-TNFα biologic agents.
- To investigate the PK profile of SKI-O-592 (the free base of SKI-O-703) and its metabolites (M2 and M4) in subjects with active RA who have had an inadequate response to csDMARDs or anti-TNFα biologic agents.
- To evaluate the effects of SKI-O-703 on exploratory PD biomarkers (i.e., the percentage of activated gp53/CD63+ basophils in peripheral blood) in subjects with active RA who have had an inadequate response to csDMARDs or anti-TNFα biologic agents.

2.3. Exploratory Objective

The exploratory objective of this study is as follows:

- To evaluate the PK/PD relationship between plasma concentrations of SKI-O-592, M2, and M4 and the percent change in activated gp53/CD63+ basophils in peripheral blood in subjects with active RA who have had an inadequate response to csDMARDs or previous anti-TNFα biologic agents.

3. Investigational Plan

3.1. Overall Study Design and Plan

This is a randomized, double-blind, multicenter, placebo-controlled, parallel dose study to evaluate the efficacy, safety, tolerability, PK, and PD of select (100 mg BID, 200 mg BID, and 400 mg BID) doses of SKI-O-703 in subjects with RA who have had an inadequate response to csDMARDs or anti-TNFα biologic agents.

Approximately 148 subjects are planned to participate in 4 cohorts (37 subjects each). Subjects will be randomly assigned using a 1:1:1:1 ratio to receive one of the three doses of SKI-O-703 (100 mg BID, 200 mg BID, or 400 mg BID) or placebo. Dosing will be twice daily for 12 weeks.

The duration of study participation for a subject is up to approximately 20 weeks and the study consists of 3 periods (screening period, 12-week treatment period and 4-week follow-up period). A schematic diagram of the overall study design is presented in Figure 3-1.

Figure 3-1 Study Design Schematic

Screening Period (up to 28 days)	12-Week Treatment Period (Double-Blind) 1:1:1:1							Follow-up (EOT+14 days)	EOS (EOT+28 days)
	100 mg BID 200 mg BID 400 mg BID Placebo BID								
Week	0 [a]	2	4	6	8	10	12	14	16
Visit	1	2	3		4		5 (EOT)	6	7

Abbreviation: BID, twice daily; EOT, end-of-Treatment; EOS, end-of-Study
[a] randomization

Screening Period

Screening assessments or procedures will be conducted up to 28 days prior to the start of study treatment. After signing and dating the informed consent form (ICF), all the potential subjects will be assessed as per the eligibility criteria at screening. After completing all screening assessments, subjects who meet all the inclusion and none of the exclusion criteria will proceed to the treatment period.

12-week Treatment Period

On Day 1, the subject will complete the Patient's Global Assessment of Disease Activity, the Patient's Assessment of Pain and the Health Assessment Questionnaire before site personnel perform any clinical assessments. After that, subject's eligibility will be confirmed and baseline safety evaluations will be performed (12-lead ECG, physical examination, vital sign measurements, weight, urine pregnancy test, and safety laboratory assessment). Following

predose study procedures, they will be randomized to receive either SKI-O-703 or placebo. On Day 1, subjects will be administered his/her first dose of study medication no later than 30 minutes after food with approximately 8 ounces (240 mL) of water by the site personnel.

A subset of subjects from each cohort who provide additional consent will be subjected to additional PK and PD assessments according to the Schedule of Events (Appendix 15.1). Due to PD assay limitations, PK and PD sampling will only be performed at selected sites that are deemed suitable to support PK/PD assessment (ie, which are within an acceptable shipping distance from a PD laboratory). Every effort will be made to achieve a minimum of 24 subjects (6 subjects per treatment group) in the PK/PD subgroup. It is important to note that there is no upper limit to the number of subjects that will be enrolled to the PK/PD subgroup in any region (as many subjects will be assigned to the PK/PD subgroup as is practical), and the number of subjects does not need to identical between treatment groups.

The second dose will be administered approximately 12 hours (±2 hours) after the first dose. The second dose and all subsequent doses, will be self-administered by the subject. The subjects will be instructed to take study medication no later than 30 minutes after food. Subjects will return to the clinic for routine safety and disease state review as outlined in the Schedule of Events, with a ±3 days window for all visits after randomization.

During the visits after Day 1, subject will undergo the procedures at the timepoints as specified in the Schedule of Events. The subject should complete the Patient's Global Assessment of Disease Activity, the Patient's Assessment of Pain and the Health Assessment Questionnaire before site personnel perform any clinical assessments. These should also be completed before any interaction with site personnel has occurred to avoid biasing the subject's response.

4-week Follow-up Period

All subjects who complete the treatment period, as well as subjects who discontinue the treatment at any time for any reason, are to enter the 4-week Follow-up period after the EOT visit.

3.2. Study Endpoints

3.2.1. Primary Endpoint

- Mean change from baseline in disease activity score in 28 joints (DAS28) using high sensitivity C-reactive protein (hsCRP) (based on 3 individual components including TJC, SJC and hsCRP) at Week 12

3.2.2. Secondary Endpoints

The secondary efficacy endpoints will be evaluated at Week 2, 4, 8 and 12 unless otherwise stated.

- Percentage of subjects who would achieve American College of Rheumatology (ACR) response (ACR20, ACR50 and ACR70) over time
- Change from baseline in DAS28-hsCRP score (based on 3 individual components including TJC, SJC and hsCRP)

- Change from baseline in the tender/painful and swollen joint count (28)
- Change from baseline in the physician global assessment of disease activity by visual analog scale (VAS)
- Change from baseline in the subject global assessment of disease activity by VAS
- Change from baseline in the subject assessment of arthritis pain by VAS
- Change from baseline in the health assessment questionnaire - disability index (HAQ-DI)
- Change from baseline in median hsCRP values at each schedule visit
- Safety and tolerability of SKI-O-703 compared to placebo including laboratory tests, infections, ECGs, vital signs, incidence of AEs, withdrawals due to AEs and serious adverse events (SAEs).
- PK parameters of SKI-O-592 and its metabolites (M2 and M4) from a subset of subjects per cohort on Day 1 and at Week 12, including: C_{max}, T_{max}, $AUC_{0\text{-}tau}$, $t_{1/2,eff}$, R_{met}, and R_{AUC}, as applicable
- PD parameters for the change from baseline in the percentage of activated gp53/CD63+ basophils in peripheral blood on Day 1 and at Week 12, including: E_{max}, TE_{max}, $AUEC_{0\text{-}tau}$, as applicable

3.2.3. Exploratory Endpoints

- An assessment of PK/PD relationship between plasma concentrations of SKI-O-592, M2, and M4 and change from baseline in the percentage of activated gp53/CD63+ basophils in peripheral blood on Day 1 and at Week 12

4. General Statistical Considerations

4.1. Reporting Convention

- Summary statistics will consist of the number and percentage of subjects in each category for discrete variables, and the sample size, mean, median, standard deviation (SD), minimum and maximum for continuous variables.
- All mean, median values and confidence intervals will be formatted to one more decimal place than the measured value. Standard deviations and standard errors will be formatted to two more decimal places than the measured value. Minimum and maximum values will be presented to the same number of decimal places as the measured value. The maximum decimal places will be set to 4.
- Frequency summary for categorical variables includes number and percentage. All percentages will be rounded to one decimal place. Number and percentage values will be presented as xx (xx.x). If the percentage is 100, no decimal is required.
- P-values will be formatted to 3 decimal places; if a p-value is less than 0.001, it will be displayed as <0.001.
- All analysis and summary tables will have the population sample size for each treatment group in the column header if appropriate. All the subject data listings will be presented by treatment group and subject identifier where applicable.
- All programmatically calculated values (e.g. treatment compliance) will retain all decimal places at the Analysis Data Model (ADaM) level. For reporting purpose in listings and summary tables, up to 1 decimal place will be retained for raw values, minimum and maximum. Other summary statistics will then follow the same incremental precision rules specified above
- The day of the first dose of any study drug will be defined as Study Day 1. Study Day are calculated relative to Study Day 1, as the difference between the date of interest and Study Day 1, and plus 1 day additionally if date of interest is on or after Study Day 1.

For PK and PD analysis, the reporting convention is specified in the corresponding Section 10 and Section 11. The above convention is applied to all efficacy and safety analyses.

4.2. Baseline Definition and Visit Window

Baseline value is defined as the last non-missing value on or before the date of the first dose of study drug. For subjects who were not treated, the baseline will be the assessment value taken on the screening visit, on or before the randomization day.

There is no separated visit folder designed in EDC for the treatment early terminated visit and the visit folder ‘Visit 5 EOT’ will serve as both the visit at Week 12 and the early terminated visit. Therefore, the assessments captured under visit folder ‘Visit 5 EOT’ will be summarized under visit Week 12 only when the subjects completed the treatment period; otherwise, ‘Visit 5 EOT’ will be treated as the treatment early terminated visit. There is no other visit mapping

performed and all the by-visit summaries will be based on the nominal visit as collected on the eCRF page except the visit handling of 'Visit 5 EOT'.

4.3. Sample Size

Approximately 148 subjects will be enrolled in 4 cohorts (3 cohorts of SKI-O-703 and 1 cohort of placebo). Each cohort will consist of at least 37 subjects.

The study is powered to detect a difference of 0.8 in mean change from baseline to Week 12 in DAS28-hsCRP score between one of the groups treated with SKI-O-703 and the group treated with placebo. At least 37 subjects on a course of SKI-O-703 treatment are required to complete the study to show a change from baseline to Week 12 of 0.80 and higher as statistically significant with a power of 0.80 at the significance level of 0.05, using a 1-way analysis of variance and assuming a standard deviation of 1.25, compared with in the placebo group. Sample size determinations were based on the 4-component DAS28 and are likely to be similar with the 3-component DAS28 with fewer variables.

Given the 1:1:1:1 ratio of SKI-O-703 100 mg BID, 200 mg BID, 400 mg BID to placebo allocation, at least 37 subjects in each treatment group are required to complete the study, therefore at least 148 subjects need to be enrolled and randomized in the study. Assuming a drop-out of rate of 10%, up to 20 additional subjects may be enrolled to achieve the target number of subjects required to complete the study.

4.4. Randomization, Stratification, and Blinding

Subjects will be randomized using a permuted block randomization in parallel in a 1:1:1:1 ratio (37 subjects per group) to receive SKI-O-703 100 mg BID, 200 mg BID, 400 mg BID, or matching placebo. An interactive voice or web response system (IxRS) will be used to administer the randomization schedule. The randomization schedule will be stratified by prior anti-TNFα biologic agents use (Yes or No), number of previous csDMARDs treatments (0-2, ≥3), and specific geographic region (APAC, EMEA and NA).

This study is double-blind study. Neither the subjects nor the investigator will be aware of the treatment assignment for the subjects in each cohort. The study drugs will be similar in number and appearance for each cohort to maintain the study blind. Blinding will be maintained throughout the study until the final database lock.

4.5. Analysis Sets

The following analysis sets will be used:

Safety set: The Safety set is defined as all subjects who receive at least one dose of study drug (SKI-O-703 or placebo). Subjects will be analyzed as treated. This population will be used for summaries of safety data.

Intent-to-Treat (ITT) set: The ITT set is defined as all randomized subjects who receive at least one dose of study drug (SKI-O-703 or placebo). Subjects will be analyzed as randomized. This population will be used for all efficacy analyses and summaries of demographics and baseline characteristics data.

Modified Intention-to-Treat (mITT) set: The mITT set is defined as all the subjects who receive at least one dose of study drug (SKI-O-703 or placebo) and have at least 1 postbaseline assessment. Subjects will be analyzed as randomized. This population will be used for efficacy summaries and will include imputed data for subjects who withdraw from the study prior to the Week 12 visit.

Per Protocol (PP) set: The PP set is defined as all the subjects in ITT set who don't have any major protocol deviation(s) which will be flagged in the protocol deviation list and be compliant with the treatment as defined in Section 7.2.2. All the efficacy analyses will be repeated on PP set. Subjects will be analyzed as randomized.

Pharmacokinetic (PK) set: The PK set will consist of all subjects who receive SKI-O-703 and have at least 1 measurable plasma concentration. Subjects who have partial data and/or major protocol deviations that may impact PK, and subjects who experience emesis, will be evaluated on a case-by-case basis and may be excluded from the PK population. This population will be used for summaries of PK data.

Pharmacodynamic (PD) set: The PD set will consist of all subjects who receive at least 1 dose of study drug (SKI-O-703 or placebo) and have at least 1 evaluable postdose PD value. Subjects who have partial data and/or major protocol deviations that may impact PD will be evaluated on a case by case basis and may be excluded from the PD population. This population will be used for summaries of PD data

Pharmacokinetic (PK)/pharmacodynamic (PD) set: The PK/PD set will consist of all subjects included in both the PK and PD populations. This population will be used for the PK/PD analyses.

4.6. Imputation of Incomplete Medication Date

For the purpose of inclusion in prior and/or concomitant medication tables, incomplete medication start and stop dates will be imputed as follows:

Missing start dates (where UK, UKN and UNKN indicate unknown or missing day, month and year respectively):

- UK-MMM-YYYY:
    - If the month and year are different from the month and year of the first dose of study drug, assume 01-MMM-YYYY.
    - If the month and year are the same as the first dose of study drug month and year and the end date (after any imputation) is on or after the first dose of study drug, then assume the date of the first dose of study drug.
    - If the month and year are the same as the first dose of study drug month and year and the end date (after any imputation) is prior to the first dose of study drug, then assume the end date for the start date;
- DD-UKN-YYYY/UK-UKN-YYYY:
    - If the year is different from the year of the first dose of study drug, assume 01-JAN-YYYY of the collected year.

- If the year is the same as the first dose of study drug year and the end date (after any imputation) is on or after the first dose of study drug, then assume the date of the first dose of study drug.
- If the year is the same as the first dose of study drug and the end date (after any imputation) is prior to the first dose of study drug, then assume the end date for the start date.

Missing stop dates (where UK, UKN and UNKN indicate unknown or missing day, month and year respectively):

- UK-MMM-YYYY: Assume the last day of the month;
- DD-UKN-YYYY/UK-UKN-YYYY: Assume 31-DEC-YYYY.

If start date is completely missing and end date is not prior to the first dose, then the medication will be classified as concomitant. If the start date is completely missing and the end date is prior to the first dose of study drug, then the medication will be classified as prior. If the end date is completely missing, then the medication will be classified as concomitant. Medications for which both the start and end dates are completely missing will be classified as prior and concomitant.

5. Subject Disposition

5.1. Disposition

The number of subjects screened, the number and percentage of subjects randomized and not randomized among all subjects screened, and the violated eligibility criteria will be summarized. The above percentages will be calculated using the number of subjects screened as the denominator.

The number and percentage of subjects included in the ITT set, mITT set, PP set, and Safety set will be summarized by treatment group. The percentages will be based on the number of subjects randomized.

The number and percentage of subjects who entered, completed and discontinued the treatment period, and the number and percentage of primary reasons for discontinuation of treatment will be provided by treatment group and overall. The primary reasons for discontinuation will be summarized as collected on the disposition CRF page. Repeat the summary for study completion/discontinuation.

The number and percentage of subjects by region, study site and treatment group will be tabulated based on ITT set.

Listings for subject eligibility/analysis set and listings for subject treatment will be provided.

5.2. Protocol Deviations

The protocol deviations/violations will be identified and assessed by clinical research physician or designee following institution standard operational procedures. Any major protocol deviation/violation will be assessed prior to the final database lock, and to be considered for exclusion from Per-protocol Set.

The protocol deviations/violations will be summarized by significance and treatment group and overall based on the ITT set. If there are any deviations caused by COVID-19 pandemic, a separated summary for these specific deviations will be provided as well.

A listing of subjects with protocol deviations/violations will be provided with detail protocol deviation category and description presented.

6. Demographics and Baseline Characteristics

The demographics and baseline characteristics will be summarized by treatment group and overall based on the ITT set. No inferential statistics will be presented. Individual subject listings will be provided to support the summary tables.

6.1. Demographics

The following characteristics will be summarized:

- Age (years)
- Age category (<65, ≥65 to <75, ≥75)
- Sex (Male, Female)
- Race (African American, American Indian or Alaska Native, Asian, Native Hawaiian or other Pacific Islander, White, Other)
- Geographic region (APAC, EMEA, NA)
- Height (cm) and Weight (kg)
- Body mass index (BMI; kg/m^2), calculated as BMI (kg/m^2) = baseline weight (in kilograms) / height^2 (in meters).
- Time since RA onset (in years), calculated as date of (informed consent – date of RA diagnosis + 1) / 365.25. If either month or month and day of RA diagnosis date is missing, use the January 1. If day is missing, use the first day of the month. Imputation will not be done for completely missing date.

6.2. Baseline Disease Characteristics

Variables indicated below will be summarized descriptively:

- Number of previous csDMARDs treatment
- Previous csDMARDs treatment category (0-2, ≥3)
- Number of previous anti-TNF biologic agents use
- Previous anti-TNF biologic agents use (Yes, No)
- Baseline DAS28-hsCRP (based on 3 individual components including TJC, SJC and hsCRP)
- Baseline tender joint count based on 28 joints (TJC28)
- Baseline swollen joint count based on 28 joints (SJC28)
- Baseline HAQ-DI total score
- Baseline physician global assessment of disease activity (VAS)
- Baseline subject global assessment of disease activity (VAS)
- Baseline subject assessment of arthritis pain (VAS)

- Baseline hsCRP (mg/L)

The number of previous csDMARDs treatment is calculated as the number of unique medications that were captured in the eCRF page Prior and Ongoing RA Treatments and limited to the following standardized medication list: Methotrexate, Sulfasalazine, Hydroxychloroquine, Minocycline, Leflunomide, Hydroxychloroquine sulfate, Methotrexate sodium.

The number of previous anti-TNF biologic agents is calculated as the number of unique medications that were captured in the eCRF page Prior and Ongoing RA Treatments and limited to the following standardized medication list: Adalimumab, Etanercept, Infliximab, Certolizumab Pegol, Golimumab.

6.3. Alcohol, Tobacco and Drug Usage

The following characteristics will be summarized:

- Smoking history (never a smoker or user of chewing tobacco or vaping devices, prior smoker or user, current smoker of user)
- Alcohol history (never a drinker, prior drinker, current drinker)

6.4. Medical History

Medical history will be coded according to the Medical Drug Regulatory Activities (MedDRA) version 21.0 or higher. The number and percentage of subjects who experience any medical history will be presented by treatment group and overall for each system organ class (SOC) and preferred term (PT), with SOCs and PTs within each SOC presented in a descending order of overall subject incidence (then the frequency within treatment groups from high dose to low dose and placebo group, if applicable).

7. Treatments and Medications

7.1. Prior and Concomitant Medications

Medications reported on the Prior and Concomitant Medications CRF will be coded using the World Health Organization Drug Dictionary (WHODD) Enhanced B3 March 2018 version and grouped into relevant categories based on the Anatomical Therapeutic Chemical (ATC) coding scheme.

A prior medication will be any medication stopped prior to the first dose of study drug.

A concomitant medication will be any medication that is either initiated before the first dose of study drug, or initiated on/after the date of the first dose of study drug, and continued during the study treatment period.

Frequency table will be provided by treatment group and overall for prior medications based on the Safety set. The number and percentage of subjects who have prior will be presented by treatment group and overall for each therapeutic drug class (ATC level 2) and generic drug names, with therapeutic drug class and generic drug name within each therapeutic drug class in a decreasing order of overall frequency (then the frequency within treatment groups from high dose to low dose and placebo, if applicable). Similar frequency tables will be provided by treatment group for permitted and prohibited concomitant medications based on the Safety set.

The list of prohibited concomitant medications should be reviewed and finalized by sponsor designee prior to the final database lock.

A listing of prior and concomitant medications by subject will be provided, with an indicator to identify the prior and concomitant medications.

7.2. Prior and Ongoing RA Medications

Medications reported on the Prior and Ongoing RA Treatments CRF page will be coded using the WHODD Enhanced B3 March 2018 version and grouped into relevant categories based on the ATC coding scheme.

Frequency table will be provided by treatment group and in total for prior and Ongoing RA medications using the Safety set. The number and percentage of subjects with prior and ongoing RA medications will be presented for each ATC level 2 term and generic drug names.

All the prior and ongoing RA medications will be listed.

7.3. Study Treatments

7.3.1. Extent of Exposure

Treatment duration and total dose will be summarized by treatment group based on the Safety set.

Treatment duration (in days) is calculated as the date of last dose of study drug – the date of the first dose of study drug + 1. Summary statistics of treatment duration, as well as frequency

summary of treatment duration categories (e.g. <2 weeks, ≥2 to <4 weeks, ≥4 to <8 weeks, ≥8 weeks) will be provided.

Total dose is defined as the total number of capsules taken over the treatment period, which is calculated as total number of capsules dispensed minus the total number of capsules returned. Summary statistics of total dose will be provided.

7.3.2. Treatment Compliance

Treatment compliance (%) will be calculated for each subject as 100 times total dose divided by the intended total dose that should be taken over the treatment period. The intended total dose is calculated as 8 times the treatment duration. Compliance will be summarized using descriptive statistics by treatment group. The acceptable range for compliance is defined as between 80% - 120%. The number and percentage of subjects will be summarized by category < 80%, ≥80% to ≤120%, and >120%.

A subject data listing of study drug accountability will be provided, with the capsules dispensed, capsules returned, capsules taken, intended total dose and treatment compliance presented.

8. Efficacy Analysis

The primary analyses of all efficacy endpoints will be based on the ITT set, and then repeated using the PP set. The analyses of primary endpoint and secondary efficacy endpoints ACR20, ACR50 and ACR70 will also be provided based on the mITT set.

Each efficacy endpoint will be descriptively summarized by treatment group over time. In the summary for a post-baseline time point, subjects who have both values at baseline and at the time point will be included. Listings of all the efficacy endpoints will be provided by subjects.

Statistical comparisons will be performed between each of active treatment groups and the placebo group. All the statistical tests will be 2-sided and performed at a significance level of 0.05, leading to 95% CIs (2-sided). There will be no adjustment for multiplicity.

8.1. Primary Outcome and Methodology

Efficacy will be assessed by the evaluation of the mean change from baseline in DAS28-hsCRP, ACR response, count of tender/painful and swollen joints, global assessment of disease activity, subject assessment of arthritis pain, and HAQ-DI at the specified time points.

8.1.1. DAS28-hsCRP

The DAS28 is a measure of disease activity in 28 joints and the core set of variables (individual components) for DAS28 for this study include:

- Number of tender joints with a total of 28 joints assessed for tenderness (TJC28)
- Number of swollen joints with a total of 28 joints assessed for swelling (SJC28)
- Subject global assessment for disease activity using VAS (PtGA)
- hsCRP (mg/L)

See details of the individual components in Section 8.1.3, Section 8.1.5 and Section 1.7.

Unless otherwise specified, the DAS28-hsCRP mentioned in this document is based on 3 individual components (TJC28, SJC28 and hsCRP) and calculated using the following equation:

$$\text{DAS28-hsCRP(3)} = [0.56\sqrt{TJC28} + 0.28\sqrt{SJC28} + 0.36\ln(hsCRP + 1)] \times 1.10 + 1.15,$$

where hsCRP must be in a unit of mg/L. High DAS28-hsCRP value indicates more severe disease activity.

Regarding 4-component DAS28-hsCRP analyzed in the sensitivity analysis, it is calculated using the following equation, consisting of 4 individual components (TJC28, SJC28, hsCRP and subject global assessment of disease activity):

$$\text{DAS28-hsCRP(4)} = 0.56\sqrt{TJC28} + 0.28\sqrt{SJC28} + 0.36\ln(hsCRP + 1) + 0.014 \times PtGA + 0.96,$$

where hsCRP must be in a unit of mg/L and PtGA is in a scale of 0 - 100 mm where 0 represents very well and 100 represents not well.

8.1.2. ACR responses (ACR20, ACR50, and ACR70)

The ACR responses are presented as the minimal numerical percent improvement from the baseline in multiple disease assessment criteria. ACRx (x can be 20, 50, 70) response is defined as ≥x% improvement from baseline in tender joint count (68 joint counts; TJC68), ≥x% improvement in swollen joint count (66 joint counts; SJC66), and ≥x% improvement in at least 3 of the following 5 items:

- Subject assessment of arthritis pain using VAS
- Subject global assessment of disease activity using VAS
- Physician global assessment of disease activity using VAS
- Subject's assessment of physical function as measured by the HAQ-DI
- hsCRP

The percent improvement from the baseline will be calculated as (Baseline – Change from baseline) / Baseline × 100. See details of individual component in each corresponding section below.

Full detail of algorithm to calculate the ACR response and the rules of non-responder imputation (NRI) can be found in protocol Appendix 13.4.

8.1.3. Tender/Swollen Joints Count

Each of the 68 joints will be evaluated for tenderness and 66 joints will be evaluated for swelling (hips are excluded for swelling) prior to randomization at Visit 1 and at each subsequent visit as indicated in the Schedule of Events. Each individual joint is to be examined and assessed as tender or not tender and as swollen or not swollen.

Individual joint with missing assessment will not be imputed. If a joint was replaced prior to the first dose of study drug, the joint will be considered as non-evaluable at baseline and all subsequent post-baseline visits. If the joint replacement was occurred after first dose, the joint will be counted as both tender joint and swollen joint at all subsequent visits after joint replacement. The number of tender and swollen joints will be calculated by summing all joints. If there are non-evaluable or missing joints among the 68/66 joints at a given visit, the prorated tender and swollen joints count (TJC68 and SJC66) will be calculated using the following formula. If less than half of the joints are assessed at that visit, the TJC68 and SJC66 will be treated as missing.

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

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

As for the scores considering 28 joints for both tenderness and swelling, denoted as TJC28 and SJC28 respectively, the following subset will be assessed (both right and left side):

- Shoulder

- Elbow
- Wrist
- Metacarpophalangeal (MCP) I, II, III, IV, and V
- Thumb interphalangeal
- Proximal interphalangeal (PIP) II, III, IV, and V
- Knee

If at least half of the joints are assessed at a given visit, the similar prorated tender and swollen joints count (TJC28 and SJC28) will be calculated as follows:

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

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

If less than half of the joints are assessed at that visit, the TJC28 and SJC28 will be treated as missing.

8.1.4. Subject Assessment of Arthritis Pain (VAS)

Per protocol version 3.0 (dated 08Feb2019), the subject is asked to measure his/her arthritis pain by marking a vertical tick on a 100mm horizontal line, where left end (0 mm) represents extreme pain and the right end (100 mm) represents no pain. Starting from protocol version 4.0 (dated 03Apr2020), the subject will be asked to measure his/her arthritis pain by marking a vertical tick on a 100mm horizontal line, where the left end (0 mm) represents no pain and the right end (100 mm) represents extreme pain. The visit and visit date when the subject started or switched to the correct scale described in protocol version 4.0 will be captured in eCRF page Corrected VAS Pain Scale.

Due to the VAS scale was reversed from the standard way they are presented in the clinical trial during protocol version 3.0 (dated 08Feb2019), scoring of the VAS scale was found to be variable across the sites. The clinical team will determine on a subject by subject basis how the subject assessment of arthritis pain was scored and identify subjects who completed consistently the subject assessment of arthritis pain throughout the study to be included in the relevant sensitivity analyses. The detail of subjects who will be included for the relevant sensitivity analyses will be provided in a separate document and will be reviewed and finalized by sponsor prior to the final database lock.

The arthritis pain scores to be used in the analysis will be the length from the anchor of no pain to the vertical tick marked by the subject, which is the number reported on eCRF using the correct scale and the value of 100 minus the number reported using the reversed scale.

8.1.5. Physician and Subject Global Assessment of Disease Activity (VAS)

Physician and subject global assessment of disease activity is measured by both the physician and the subject indicating the subject's current disease status by marking a vertical tick on a 100mm horizontal line, where the left end (0 mm) represents not well and the right end (100 mm) represents very well.

Due to the VAS scale was reversed from the standard way they are presented in the clinical trial during protocol version 3.0 (dated 08Feb2019), scoring of the VAS scale was found to be variable across the sites. The clinical team will determine on a subject by subject basis how the physician and subject global assessment of disease activity was scored and identify subjects whose physician and subject global assessment of disease activity was completed consistently thought the study to be included in the relevant sensitivity analyses. The detail of subjects who will be included for the relevant sensitivity analyses will be provided in a separate document and will be reviewed and finalized by sponsor prior to the final database lock.

The physician and subject global assessment score to be used in the analysis will be the length from the anchor of very well to the vertical tick marked by the physician and subject, namely the value of 100 minus the number reported on eCRF page.

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

The HAQ-DI is a subject-reported questionnaire that is commonly used to measure the disease associated disability. It consists of 8 sections which are dressing or grooming, arising, eating, walking, hygiene, reach, grip, and activities.

The disability section of the questionnaire scores the subject's self-perception on the degree of difficulty (0 = without any difficulty, 1 = with some difficulty, 2 = with much difficulty, and 3 = unable to do). The reported use of special aids or devices and/or the need for assistance of another person to perform these activities is also assessed. For each section, the score will be the worst score within that section. In addition, if an aide or device is used or if help is required from another individual, the minimum score for that section is 2. If the section score is 2 or more then no modification is made. The scores for each of the functional sections will be averaged to calculate the functional disability index. The disability index will be considered as missing when the subject does not have scores for at least 6 sections.

Full detail of algorithm to score HAQ-DI can be found in protocol Appendix 13.10.

8.1.7. hsCRP

hsCRP will be evaluated in the central laboratory. If due to COVID-19 pandemic, the central laboratory analysis cannot be performed (like the samples cannot be shipped to the central laboratory) and the samples are analyzed locally instead, then the local laboratory results will be included in the primary analyses/summaries.

8.2. Primary Efficacy Endpoint

The primary efficacy endpoint is the mean change from baseline in DAS28-hsCRP at Week 12, which is based on three individual components including TJC28, SJC28 and hsCRP (see calculation in Section 8.1.1)

8.2.1. Primary Analysis

The primary analysis will be conducted on observed cases, which means DAS28-hsCRP at Week 12 will be calculated only when all the three individual components (TJC28, SJC28 and hsCRP) are assessed at this visit; otherwise, DAS28-hsCRP at Week 12 will be considered as missing.

The change from baseline in DAS28-hsCRP at visit Week 12 will be analyzed using an analysis of covariance (ANCOVA) model. The model will include treatment groups as a factor and DAS28-hsCRP at baseline as a covariate. Within-group least-square (LS) means and the associated standard errors (SE), the treatment differences in LS means and the associated SEs, the 2-sided 95% CIs, the 2-sided p-values for treatment comparisons between each of active treatment group and placebo group, and the 2-sided p-values for treatment comparisons among active treatment groups (high dose vs. low dose) will be estimated and provided. Descriptive statistics of baseline, observed value and change from baseline at Week 12 will be summarized by treatment group.

The primary efficacy analyses will be performed based on the ITT set, and then repeated using the PP set.

8.2.2. Sensitivity Analysis: Using LOCF

The primary analysis will be repeated based on the mITT set, and the missing value of DAS28-hsCRP at Week 12 will be imputed using the LOCF method. Before calculating DAS28-hsCRP at Week 12, the component-level imputation using LOCF at Week 12 will be performed for each individual component (TJC28,SJC28 and hsCRP).

8.2.3. Sensitivity Analysis: Using MI Under MAR Framework

Repeat the primary analysis on the ITT set and before calculating the DAS28-hsCRP, the missing value of each individual component (TJC28,SJC28 and hsCRP) will be imputed by multiple imputation (MI) method for all treatment groups as described below, under the missing at random (MAR) framework.

The SAS procedure MI will be used to impute the missing value for each individual component at the scheduled post-baseline analysis visits (up to Week 12) to create 25 complete data sets. The missing data patterns will be checked by treatment and at scheduled visits. If there are non-monotone missing patterns, two separate imputation procedures as described below will be used to complete the imputation process; otherwise, skip the first step. For components SJC28 and TJC28, the imputed values will be rounded to integers. The range of imputed value should be from minimum value 0 to maximum value 28 for SJC28 and TJC28. The minimum value for hsCRP should be 0.

In the first step, the Markov Chain Monte Carlo (MCMC) method will be used to impute missing values by treatment to create 25 imputed data sets with monotone missing patterns. The seed will be set to 7032101 and a single chain will be used to procedure imputations.

In the second step, the monotone regression method will be used to impute the remaining missing values for the 25 data sets with monotone missing patterns. The MONOTONE REG statement will be used with the seed 2101703. The missing values at each visit will be imputed based on treatment, stratification factors (prior anti-TNFα biologic agents use, number of

previous csDMARDs treatments, geographic region) and values at baseline and previous visits. (If the imputation cannot be implemented due to lack of sufficient subjects in a specific stratum, remove the factor from the imputation model.)

After the completion of two steps imputations, calculate the DAS28-hsCRP based on the values of individual components in the 25 complete datasets and the same statistical analysis method described in the primary analysis will be conducted for each of the 25 completed datasets. The SAS procedure MIANALYZE will be used to combine the results for the statistical inferences.

8.2.4. Sensitivity Analysis: Excluding Local hsCRP Values

The primary analysis will also be repeated based on the ITT set using the hsCRP value analyzed from central laboratory. The hsCRP results from local laboratory (if any) will be considered as missing. Then all missing data up to Week 12 will be imputed by MI method under:

- MAR framework.

  Under this framework, the same MI method will be applied as outlined above.

- Missing not at random (MNAR) framework.

  Two imputation steps will be performed. First step is to create 25 datasets with monotone missing patterns which is the same as described above under MAR framework. The second step is control-based pattern imputation using monotone regression method with seed 2010703, in which the missing data in all active treatment groups and placebo group will be imputed based on placebo group.

The same statistical analysis method described in the primary analysis will be conducted for each of the 25 completed datasets, and the SAS procedure MIANALYZE will be used to combine the results for the statistical inferences.

8.2.5. Sensitivity Analysis: Using 4-component DAS28-hsCRP

Repeat the statistical analysis (described in Section 8.2.1) on the ITT set using 4-component DAS28-hsCRP as calculated in Section 8.1.1. Only subjects who completed the subject global assessment of disease activity consistently will be included in this sensitivity analysis.

8.3. Secondary Efficacy Endpoints

8.3.1. Primary Analysis

The primary analysis will be conducted on observed cases without any imputation. All the secondary efficacy endpoints will be summarized and analyzed by treatment group at each schedule visit (Week 2, 4, 8, and 12) based on the ITT set and then repeated on the PP set.

The DAS28-related secondary endpoint is based on 3 individual components including TJC28, SJC28 and hsCRP (see calculation in Section 8.1.1). The analyses on ACR responses (ACR20, ACR50 and ACR70) will be restricted to subjects who completed subject global assessment of disease activity consistently, as well as physician global assessment of disease activity and subject assessment of arthritis pain, using the same approach as the sensitivity analysis on 4-component DAS28-hsCRP to identify the subjects to be included in the analysis. Similar to the

analyses on ACR responses, the analyses on the individual component using VAS will be restricted to subjects who completed the corresponding individual VAS assessment.

Frequency counts and percentages will be provided for ACR responses (ACR20, ACR50, and ACR70) at each schedule visit. The Chi-square tests will be used to compare the ACR response rate between each of active treatment group and placebo group, and among active treatment groups at each schedule visit. The p-values from the Chi-square tests will be presented. In case of few responders where Chi-square test may not be appropriate or accurate, Fisher's exact test will be performed instead. Besides, 95% CI of binomial proportion of ACR response within each treatment group and 95% CI of the difference in proportion of subjects with ACR response for pairwise treatment comparison will be estimated using the Wald method.

The hsCRP will be summarized descriptively by treatment group over schedule visits. No inferential statistics will be provided.

For all other secondary efficacy endpoints, descriptive statistics of baseline, observed value and change from baseline at each schedule visit will be provided by treatment group. Treatment comparisons will be conducted using a mixed model repeated measure (MMRM) analysis, including all the scheduled visits up through Week 12. The model will be fitted with treatment group, visit, treatment by visit interaction and baseline as fixed effects, and visit within subject as a repeated effect, using an unstructured covariance to model the between- and within-subject errors. If the model fails to converge with unstructured covariance matrix, use the heterogeneous Toeplitz covariance structure, followed by the heterogeneous autoregressive covariance, followed by the compound symmetry covariance structure. The inferential statistics from this model, including within-group LS means and the associated SE, treatment differences in LS means and the associated SEs, the 2-sided 95% CIs, the 2-sided p-values for treatment comparisons between each of active treatment group and placebo, and the 2-sided p-values for treatment comparisons among different active treatment groups (high dose vs. low dose) will be provided.

Listings for ACR responses, DAS28-hsCRP assessments, number of tender and swollen joints, subject assessment of pain using VAS, subject and physician global assessment of disease activity and HAQ-DI will be provided.

Plots of LS mean change from baseline in continuous primary and secondary efficacy endpoints by visit will be provided as well. Proportion of subjects with ACR response will be plotted by visit.

8.3.2. Sensitivity Analysis

A sensitivity analysis based on 4-component DAS28-hsCRP (as calculated in Section 8.1.1) will be performed using ITT set restricted to subjects who completed consistently the subject global assessment of disease activity.

The ACR responses will also be analyzed based on the mITT set, where NRI method (refer to protocol Appendix 13.4) will be used to impute missing value. The same statistical analysis method will be used as described in Section 8.3.1. The analysis will be restricted to subjects who completed subject global assessment of disease activity consistently, as well as physician global assessment of disease activity and subject assessment of arthritis pain.

8.4. Other Efficacy Endpoints

Based on the DAS28-hsCRP (based on 3 components including TJC28, SJC28 and hsCRP), we will evaluate if the subject achieves the following disease activity category at each schedule visit:

- DAS28-hsCRP <= 3.2
- DAS28-hsCRP < 2.6

The frequency counts and percentages of above DAS28 binary endpoints will be provided by treatment group at each schedule visit based on the ITT set and the PP set. The Chi-square tests will be used to compare the response rate between each of active treatment groups and placebo group, and among different active treatment groups. Besides, 95% CI of binomial proportion of ACR response within each treatment group and 95% CI of the difference in proportion of subjects with ACR response for pairwise treatment comparison will be estimated using the Wald method.

The analyses will also be analyzed based on the mITT set, where NRI method will be used to impute missing value. If at least 1 of the 3 components (TJC28, SJC28 and hsCRP) is non-missing at the given visit and the subject is still enrolled in the study, then use LOCF method to fill in any missing values for the 3 components. Then calculate the DAS28-hsCRP and evaluate the DAS28 binary endpoints as defined above. If all the individual components are missing at the given visit or if prior to the visit the subject is considered discontinued from the treatment, assign the subject as a non-responder.

8.5. Subgroup Analysis

The focus of subgroup analysis will be limited to the baseline characteristics as listed below and the subgroup analysis will be conducted based on the ITT set for all the secondary efficacy endpoints specified in Section 3.2.2. For VAS-related secondary endpoints (ACR responses and individual VAS assessments), the subgroup analysis will be restricted to subjects who completed the corresponding related VAS assessment consistently, same approach as described in the Section 8.3.1 to identify the subjects to be included in the analyses.

- Number of previous csDMARDs treatments (0–2 csDMARDs, ≥3 csDMARDs)
- Previous anti-TNFα biologic agents use (Yes or No)
- Specific geographic region (APAC, EMEA, NA)

For ACR response endpoints, the number and percentage of responder and 95% Wald CIs will be provided by visit for each subgroup. The p-value of treatment comparison will be estimated by Chi-square tests within each subgroup. If all the subjects within a subgroup are responders or non-responders at a visit, only frequency counts and percentages will be provided for that subgroup and visit.

For other efficacy endpoints, the descriptive statistics of baseline, absolute value and change from baseline will be summarized by treatment group, scheduled visit and subgroup. By using the MMRM analysis as described in Section 8.3.1, for each subgroup, the inferential statistics for treatment effect estimation and treatment comparisons will be provided by visit. In addition, subgroup comparison within treatment group will be provided as well by using a MMRM

analysis, which will be fitted with treatment group, subgroup, visit, treatment by visit interaction, treatment by subgroup interaction and baseline as fixed effects, and visit within subject as a repeated effect.

If too few subjects fall in the subgroup, the analysis within the subgroup may not be performed and only descriptive summary will be presented.

9. Safety Analysis

All the Safety analyses will be based on the Safety set.

9.1. Adverse Events

All the Adverse Events (AE) will be coded using the MedDRA version 21.0 or higher. The severity of an AE will be graded 1 to 5 using the following criteria by the investigator.

Grade 1	No interference with activity
Grade 2	Some interference with activity not requiring medical intervention
Grade 3	Prevents daily activity and requires medical intervention
Grade 4	Emergency room visit or hospitalization
Grade 5	Fatal

For any AE that is proven to start or worsen on or after the day of first administration of study drug will be categorized as a treatment-emergent AE (TEAE). An SAE is defined as any event that results in death, is immediately life-threatening, requires inpatient hospital or prolongation of existing hospitalization, results in persistent or significant disability/incapacity, or is a congenital anomaly/birth defect. A Treatment-related TEAE is defined as TEAE which is considered to be related to the study drug and reported as “Possible”, “Probable” or “Definite” on the CRF. Any TEAE with a missing relationship will be classified as a treatment-related TEAE.

For the following AE categories, number and percentage of subjects who experience any TEAE, and the number of event will be summarized by treatment group for each system organ class (SOC) and preferred term (PT), with SOCs and PTs within each SOC presented in a descending order of subject incidence within treatment groups from high dose to low dose and placebo group.

- All TEAEs
- Treatment-related TEAEs
- All SAEs
- Treatment-related SAEs
- Grade 3 or 4 TEAEs
- Grade 3 or 4 treatment-related TEAEs
- TEAEs leading to treatment modification or discontinuation
- TEAEs leading to death

An overview summary of the above AE categories will be provided by treatment group, showing the number and percentage of subjects who experience specific events and the number of events.

Summary table of All TEAEs by maximum toxicity grade will also be provided by treatment group. If subject reports multiple occurrences of a specific AE, the subject will be counted only once on the maximum grade.

Separated listings will be prepared for All TEAEs, SAEs, Grade 3 or 4 TEAEs, TEAEs leading to treatment modification or discontinuation and TEAEs leading to death. The verbatim term, PT and SOC, as well as event dates and other attributes that were collected on CRF will be included in the listing.

9.2. Adverse Events of Special Interest

Adverse events of special interest as listed below will be summarized by treatment group.

The AE of special interest:

- Neutropenia (or Neutrophil count decreased)
- Hepatotoxicity

The number and percentage of subject who experience any TEAE of special interest and the number of events will be presented for each preferred term in a descending order of frequency within treatment groups from high dose to low dose and placebo group.

A listing of AE of special interest will be provided with verbatim term, PT and SOC presented, as well as other relevant event attributes collected on CRF.

9.3. Clinical Laboratory Evaluations

The laboratory values will be graded using National Cancer Institute's Common Terminology Criteria for Adverse Events (CTCAE) version 5.0. For hematologic and chemistry laboratory values that fall outside of the grade criteria of CTCAE version 5.0, the grade of 0 will be assigned.

Summary statistics of baseline value, observed values and change from baseline at each schedule visit in the continuous hematology and clinical chemistry parameters will be provided by treatment group. Numeric values reported as less than (<) or greater than (>) a laboratory-defined limit will be included with the "<" or ">" signs ignored. Frequency summaries of shifts from baseline grade to the post-baseline grade at each schedule visit, and to the worst post-baseline grade during 12-week treatment period will be provided in the laboratory tests with CTCAE gradings. In addition, frequency summaries of shifts from baseline category to the post-baseline category in terms of normal/abnormal will be provided over time in other laboratory parameters.

Subject data listings of all laboratory data sorted by subject identifier, date of collection and laboratory parameter will be provided based on the safety set, with abnormality indicated.

9.4. Vital Sign Measurements

Vital sign measures will include sitting systolic and diastolic blood pressures (mmHg), pulse rate (beats/ minute), respiratory rate (breaths/ minute) and oral temperature (ºC).

Summary statistics of baseline, observed value and change from baseline will be provided by treatment group over time. Frequency summaries of shift from baseline to post-baseline visits by normal indicator (in terms of low, normal and high) will be provided for pulse rate, respiratory rate and blood pressure. The normal ranges are defined as: 60 to 100 beats/minute for pulse rate, 12 to 20 breaths/minute for respiratory rate, 36.5 to 37.5 ºC for oral temperature, 90 to 140 mmHg for systolic blood pressure, and 60 to 90 mmHg for diastolic blood pressure. The low

abnormality is defined as any value lower than the low limit of normal range, and the high abnormality is defined as any value higher than the high limit of normal range.

A subject data listing of all vital signs data will be presented by subject identifier, date of collection and vital sign parameter, with abnormality indicated.

9.5. Physical Examination

Physical examination is only performed at screening, baseline, end-of-treatment and end-of-study. A subject data listing of physical examination will also be presented with examination visit, study day, system organ class examined and the examination results.

9.6. Electrocardiogram

Single 12-lead ECGs will be obtained, after the subject has been in the supine position for at least 5 minutes, prior to randomization at Day 1 and at subsequent visit as indicated in the schedule of events. ECG parameters include RR interval, PR interval, QRS width, QT interval, and corrected QT interval from Fridericia's formula (QTcF). In addition, the12-lead ECG results will be interpreted by the investigator and recorded as: normal; abnormal not clinically significant; or abnormal clinically significant.

Summary statistics of observed values and changes from baseline in ECG parameters will be provided over time. Frequency summaries of shifts from baseline interpretation results to post-baseline interpretation results at each visit will be provided. The number and percentage of subjects who have the following QTcF interval prolongation will be tabulated by treatment group over time. The descriptive summary of QTcF interval for the subjects with each specific interval prolongation will also be presented.

- QTcF > 450 msec
- QTcF > 480 msec
- QTcF > 500 msec

A subject data listing of ECG measurements will be provided by subject identifier, date of collection and ECG parameters.

10. Pharmacokinetic Analysis

10.1. Pharmacokinetic Data Handling

Plasma concentrations that are below the limit of quantification (BLQ) will be treated as zero for calculation of descriptive statistics and all PK analyses. Missing concentrations will be excluded from all calculations.

All concentrations and PK parameters will be rounded to 3 significant figures for reporting purposes. The only exceptions will be actual sampling times and T_{max} (which will be rounded to 2 decimal places) and the number of data points (which will be reported as integers). Summary statistics will be rounded to the same precision as individual data, with the exception of n (which will be reported as an integer). Estimates and confidence intervals from all statistical comparisons will be reported to 2 decimal places, and associated p-values will be reported to 4 decimal places.

10.2. Plasma Pharmacokinetic Concentrations

Serial blood samples for PK assessment of SKI-O-592, M2, and M4 will be collected from a subset of patients at the following time points:

- Day 1: 0 hour (predose), and at 15 and 30 minutes (±5 minutes), 1, 2, 4, 8 and 12 hours (±10 minutes) postdose (12 hour sample should be collected prior to the evening dose).
- Trough: predose samples will be collected pre-morning on Weeks 2 (Day 14), 4 (Day 28), and 8 (Day 56).
- Week 12 (Day 84): 0 hour (predose), and at 15 and 30 minutes (±5 minutes), 1, 2, 4, 8 and 12 hours (±10 minutes) after administration.

Individual plasma concentrations of SKI-O-592, M2, and M4 will be presented in data listings and presented graphically versus actual time by treatment and analyte on both linear and semi-logarithmic scales.

Plasma concentrations of SKI-O-592, M2, and M4 will be summarized using descriptive statistics (n, mean, SD, CV, minimum, median, and maximum) by time point for each treatment and day separately. Mean plasma concentrations will be plotted by treatment for each analyte versus nominal time on both linear and semi-logarithmic scales.

10.3. Plasma Pharmacokinetic Parameters

The following plasma PK parameters will, where possible, be estimated on Days 1 and 84 for SKI-O-592, M2, and M4 using actual sampling times via non-compartment analysis using Phoenix WinNonlin version 6.4 or higher:

C_{trough} Trough concentration observed at steady state (Day 84 only).

C_{max} Maximum observed concentration.

T_{max} Time of the maximum observed concentration.

$AUC_{0\text{-}tau}$ Area under the concentration versus time curve within a dosing interval (τ), calculated using the linear trapezoidal rule.

R_{met} Metabolite ratio, calculated as [($AUC_{0\text{-}tau}$ (metabolite)/ ($AUC_{0\text{-}tau}$ (parent)].

R_{AUC} Accumulation ratio based, calculated as [$AUC_{0\text{-}tau}$ (Day 84)/ $AUC_{0\text{-}tau}$ (Day 1)].

$t_{1/2,eff}$ Effective half-life, calculated as [ln(2)/((-ln(1-(1/R_{AUC}))/τ))].

Actual PK blood sampling times will be used for the assessment of all PK parameters. Plasma PK parameters for SKI-O-592, M2, and M4 will be presented in data listings and summarized separately using the following descriptive statistics for each treatment and analyte (n, geometric mean, geometric CV, mean, SD, CV, minimum, median, and maximum). For T_{max}, only n, median, minimum and maximum will be reported.

Note: due to the limited duration of sampling after each dose (i.e. 12 hours), parameters associated with the estimation of a terminal elimination phase (e.g., λ_z, $t_{1/2}$, $AUC_{0\text{-}\infty}$, etc.) will not be presented.

10.4. Pharmacokinetic Statistical Analyses

10.4.1. Assessment of Dose-proportionality

A non-linear power model will be used to assess dose-proportionality for SKI-O-592, M2 and M4 on Days 1 and 84 separately. The proportional relationship between the expected value of each parameter [E(y)] and dose is written as a power function:

$E(y) = a * dose^b$ [Equation 1]

where 'a' is a constant, 'b' is the proportionality constant and 'y' is the parameter of interest ($AUC_{0\text{-}tau}$ and C_{max}). The exponent, b, will be estimated by performing a linear regression of ln(parameter) on ln(dose). The exponent, b, is the estimated slope of the resulting regression line, since taking logs of Equation 1 gives the linear relationship:

$ln(y) = ln(a) + b * ln(dose)$ [Equation 2]

The relationship is dose proportional when b = 1.

To test for lack of fit of the power model, a quadratic term [$ln(dose)^2$] will be included in a separate model to test for a systematic departure from linearity; a p-value < 0.05 would indicate significant lack of fit of the power model.

The assumption of a linear relationship between ln-transformed parameter and ln-dose will be tested by including a quadratic term $[ln(dose)]^2$ in the model to test for systematic departures from linearity as a test of lack of fit (LOF) of the power model.

Scatter plots of $AUC_{0\text{-}tau}$ and C_{max} versus dose (on a log-log scale) will be presented by day for all analytes.

10.4.2. Assessment of Accumulation

The accumulation ratio will be estimated by making comparisons of $AUC_{0\text{-}tau}$ and C_{max} on Day 84 with corresponding estimates on Day 1 for each treatment separately; a mixed-effects model will be fitted to log-transformed $AUC_{0\text{-}tau}$ and C_{max} including day (Day 1 or 84) as a fixed effect and subject as a random effect. This model will be used to estimate the accumulation ratio (Day 84 / Day 1) with 95% CIs for each treatment. The geometric least squares means and ratio with 95% CIs will be tabulated by treatment and parameter.

11. Pharmacodynamic Analysis

11.1. Pharmacodynamic Data Handling

Subjects with baseline CD63+ values below the limit of detection will be defined as low responders and excluded from the PD set and all summaries of PD data. Baseline will be defined as the percentage of activated CD63+ cells at predose samples on Day 1.

All PD data will be rounded to 3 significant figures for reporting purposes. The only exceptions will be actual sampling times and TE_{max} (which will be rounded to 2 decimal places). Summary statistics will be rounded to the same precision as individual data, with the exception of n (which will be reported as an integer).

11.2. Pharmacodynamic Data

Serial blood samples for assessment of the percentage of activated CD63+ basophils will be collected in the same subset of patients for whom PK samples will be collected at the following time points:

- Day 1: 0 hour (predose), and at 30 minutes, 1, 2, 4, 8 and 12 hours postdose (12 hour sample should be collected prior to the evening dose).
- Trough: predose samples will be collected pre-morning on Day 14 (Week 2), Day 28 (Week 4), and Day 56 (Week 8).
- Day 84 (Week 12): 0 hour (predose), and at 30 minutes, 1, 2, 4, 8 and 12 hours (±10 minutes) hours after administration.

Absolute and percentage change from baseline (CFB) in activated CD63+ basophils will be listed by treatment, day and sampling time. Individual change from baseline in CD63+ data will be plotted for each treatment versus time on linear scales using actual times.

Absolute and percentage change from baseline in CD63+ basophils will be summarized by time point for each treatment and day using the following descriptive statistics: n, mean, SD, CV, median, minimum and maximum. Mean percentage change from baseline in CD63+ basophils will be plotted versus time by treatment and day using nominal times on linear scales.

11.3. Pharmacodynamic Parameters

The following PD parameters will be estimated on Day 1 and 84 using change from baseline in activated CD63+ basophils in peripheral blood using actual sampling times via non-compartment analysis using Phoenix WinNonlin version 6.4 or higher:

E_{max} Maximum effect

TE_{max} Time to achieve maximum effect

$AUEC_{0\text{-}tau}$ Area under the effect versus time curve within a dosing interval

Actual PK blood sampling times will be used for the assessment of all PD parameters, and negative baseline-adjusted values will be included in the analysis.

PD parameters will be presented in data listings and summarized separately using the following descriptive statistics for each treatment and analyte (n, mean, SD, CV, minimum, median, and maximum). For TE_{max}, only n, median, minimum and maximum will be reported.

12. Pharmacodynamic/Pharmacokinetic Analysis

The PK/PD relationship between SKI-O-592, M2, and M4 and CFB in percent of activated CD63+ basophils will be assessed. Scatter plots of individual plasma concentrations of SKI-O-592, M2, and M4 will be presented versus individual change from baseline in percentage of activated CD63+ basophils.

In addition, the relationship between individual concentrations of SKI-O-592, M2, and M4 and individual CFB in percent of activated CD63+ basophils will be modelled using an E_{max} model. If the data warrant, the model will be a nonlinear mixed effect model with inter-individual variability considered for the parameters in the model. If convergence is not obtained with the E_{max} model, the E_{max} parameter may be fixed (assuming 100% reduction is possible) or a linear model will be considered. The parameter estimates for the model will be presented.

13. Interim Analysis

No interim analysis is planned in this study.

14. Changes in the Planned Analysis

No changes.

15. Appendices

15.1. Schedule of Events