Statistical Analysis Plan

NCT Number: NCT05153148

Title: A Phase 2b, Randomized, Multi-Center, Double-Blind, Placebo-Controlled, Multiple Dose Study to Evaluate the Efficacy, Safety, and Tolerability of NDI-034858 in Subjects With Active Psoriatic Arthritis

Study Number: 4858-202

Document Version and Date: Version 2.0, 26 June 2023

Certain information within this document has been redacted (ie, specific content is masked irreversibly from view) to protect either personally identifiable information or company confidential information.

Parexel International

Nimbus Lakshmi, Inc.

4858-202

A PHASE 2B, RANDOMIZED, MULTI-CENTER, DOUBLE-BLIND, PLACEBO-CONTROLLED, MULTIPLE DOSE STUDY TO EVALUATE THE EFFICACY, SAFETY, AND TOLERABILITY OF NDI-034858 IN SUBJECTS WITH ACTIVE PSORIATIC ARTHRITIS

Statistical Analysis Plan

Version v2.0

Parexel Project Number 264773

SPONSOR SIGNATURE PAGE

Approved by:

[redacted]

Takeda Pharmaceuticals USA, Inc.

[redacted]

Takeda Pharmaceuticals USA, Inc.

[redacted]

Takeda Pharmaceuticals USA, Inc.

SPONSOR SIGNATURE PAGE

Parexel SIGNATURE PAGE

Signature(s) below confirm that the Statistical Analysis Plan was developed in accordance with SOP-GDO-WW-019 and that it is approved for release.

This document has been approved and signed electronically by the following:

Signatory	
Author	[redacted] Project Role: Bios[redacted]

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

REVISION HISTORY

Version No.	Effective Date	Summary of Change(s)
1.0	14-Jul-2022	New document
2.0	01-Jul-2023	• Updated the SAP per Study Protocol Version 4.0 (May 05, 2023) • Section 4.2: Clarifications and editorial changes • Section 4.3: Clarifications, editorial changes, update to the analysis window • Section 4.5: Clarifications, editorial changes, adding specificity to the criteria for the per protocol set • Section 4.6 to 4.11: Clarifications and editorial changes • Section 4.12: Clarifications, editorial changes, and updates throughout the section, with significant updates to the following: o Added the estimand framework in Section 4.12.1.2. o Updated the efficacy analysis and missing data handling for secondary binary and continuous endpoints in Section 4.12.1.3 o Clarified derivation of primary endpoint in Section 4.12.2, added subsections for supplementary analysis and sensitivity analysis o Clarified the derivation rules for TJC, SJC, HAQ-DI, dactylitis in Section 4.12.3 • Section 4.13: Clarifications, editorial changes, update and adding analyses under clinical laboratory evaluation and consideration for COVID • Section 4.14: Clarifications and editorial changes • Section 4.15: Clarifications, editorial changes, updates to the derivations for selected exploratory endpoints • Section 4.17: Providing the changes in planned analysis from protocol • Section 5: Updating reference • Section 6: Clarifications, editorial changes, and incorporating updates

LIST OF ABBREVIATIONS

Abbreviation / Acronym	Definition / Expansion
ACR20/50/70	American College of Rheumatology 20/50/70
AE	adverse event
AESI	adverse event of special interest
AIC	Akaike's Information Criterion
AR	autoregressive
BMI	body mass index
BSA	body surface area
CI	confidence interval
COVID-19	Coronavirus Disease 2019
CRP	C-reactive protein
CS	compound symmetry
CTCAE	Common Terminology Criteria for Adverse Events
DAPSA	Disease Activity Index for Psoriatic Arthritis
[redacted]	[redacted]
DILI	drug-induced liver injury
DMARD	Disease-Modifying Anti-Rheumatic Drug
EAIRs	exposure-adjusted incident rates
ECG	electrocardiogram
eCRF	electronic Case Report Form
EOS	end of study
EOT	end of treatment
ESR	erythrocyte sedimentation rate
ET	early termination
[redacted]	[redacted]
FAS	Full Analysis Set
FSH	follicle-stimulating hormone

Abbreviation / Acronym	Definition / Expansion
GA	Global Assessment
geoCV	geometric coefficient of variation
geoMean	geometric mean
HAQ-DI	Health Assessment Questionnaire-Disability Index
HBsAg	hepatitis B surface antigen
HBV	hepatitis B virus
HCT	hematocrit
HCV	hepatitis C virus
HDL	high-density lipoprotein
Hgb	hemoglobin
HIV	human immunodeficiency virus
hsCRP	high sensitivity C-reactive protein
ICE	intercurrent event
ICH	International Conference on Harmonisation
IP	Investigational product
ITT	Intent-to-Treat
LDL	low-density lipoprotein
LEI	Leed's Enthesitis Index
LLOQ	lower limit of quantification
MACE	major adverse cardiac event
MAR	missing-at-random
MCS	Mental Component Summary
MDA	Minimal Disease Activity
MedDRA	Medical Dictionary for Regulatory Activities
MH	Mantel-Haenzel
MMRM	mixed model for repeated measures
NA	not available

Abbreviation / Acronym	Definition / Expansion
NRI	non-response imputation
[redacted]	[redacted]
PASI	Psoriasis Area and Severity Index
PASI-75	75% improvement from baseline in Psoriasis Area and Severity Index
PCS	Physical Component Summary
PD	pharmacodynamic
PhGA	Physician Global Assessment
PGA-PsA	Patient Global Assessment of Psoriatic Arthritis
PGAAP	Patient Global Assessment of Psoriatic Arthritis Pain
PhGA-PsA	Physician Global Assessment of Psoriatic Arthritis
PK	pharmacokinetic
PKS	Pharmacokinetic Set
PP	Per Protocol
PPS	Per Protocol Set
PsA	psoriatic arthritis
[redacted]	[redacted]
PSO	psoriasis
PT	preferred term
Q1	25th percentile
Q3	75th percentile
QD	once daily
SAP	statistical analysis plan
SAE	serious adverse event
SD	standard deviation
SE	standard error
[redacted]	[redacted]

Abbreviation / Acronym	Definition / Expansion
SJC	Swollen Joint Count
SMQ	Standardized MedDRA query
SOC	system organ class
SS	Safety Analysis Set
TEAE	treatment-emergent adverse event
TEAESI	treatment-emergent adverse event of special interest
TESAE	treatment-emergent serious adverse event
TJC	Tender Joint Count
TNF	tumor necrosis factor
TOEP	Toeplitz
UN	unstructured
VAS	visual analog scale
WHO-DD	World Health Organization - Drug Dictionary
WOCBP	Women of Childbearing Potential

1.0 INTRODUCTION

NDI-034858 (also known as TAK-279) is a small molecule allosteric inhibitor which leads to inhibition of the tyrosine kinase 2, therefore, has the potential to add value to the treatment algorithm of Psoriatic Arthritis (PsA), particularly considering the lack of highly efficacious oral agents. This study is a Phase 2b, randomized, multi-center, double-blind, placebo-controlled, multiple dose study designed to evaluate the efficacy, safety, and tolerability of NDI-034858 in subjects with active psoriatic arthritis.

On February 08, 2023, clinical study sponsor Nimbus Lakshmi, Inc became a directly owned subsidiary of Takeda Pharmaceuticals, U.S.A., Inc.

The Statistical Analysis Plan (SAP) details the statistical methodology to be used in analyzing study data and outlines the statistical programming specifications for the tables, listings, and figures. It describes the variables, analysis populations, anticipated data manipulations and other details of the analyses not provided in the protocol.

The analyses described in this SAP are based upon the following study document:

- Study Protocol, Version 4.0 (May 05, 2023)

The structure and content of the SAP are based upon the International Conference on Harmonisation (ICH) E3 – Guideline for Industry Structure and Content of Clinical Study Reports.

[redacted]

2.0 STUDY OBJECTIVES

2.1 Primary Objective(s)

- To assess the efficacy of NDI-034858 orally administered QD at [redacted] for 12 weeks on the rheumatological signs, symptoms, and function in subjects with active PsA

2.2 Secondary Objective(s)

- To assess additional evaluations of efficacy of NDI-034858 orally administered QD at [redacted] for 12 weeks in subjects with active PsA
- To assess the safety and tolerability of NDI-034858 orally administered QD at [redacted] for 12 weeks in subjects with active PsA
- To evaluate the plasma concentration of NDI-034858 orally administered QD at [redacted] in subjects with active PsA

2.3 Exploratory Objective(s)

- [redacted]
- [redacted]

3.0 INVESTIGATIONAL PLAN

3.1 Overall Study Design and Plan

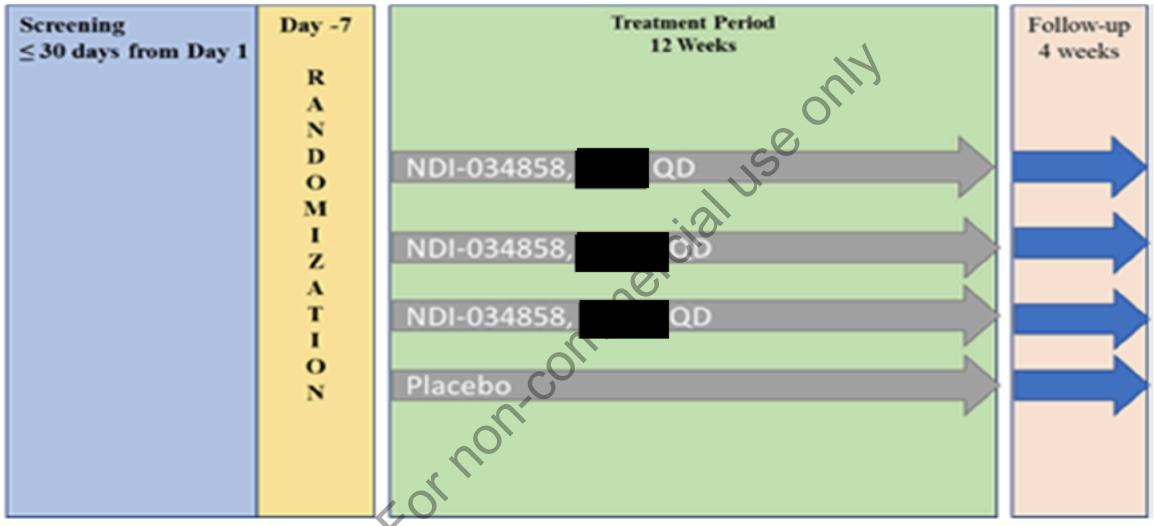
This is a Phase 2b, randomized, multi-center, double-blind, placebo-controlled, multiple-dose study designed to evaluate the efficacy, safety, and tolerability of NDI-034858 in subjects with active PsA.

Approximately 260 male and female subjects, aged 18 and higher, with active PsA will be randomized in this study.

This study consists of three periods (Figure 3-1):

- Screening Period: ≤30 days
- Treatment Period: 12 weeks
- Follow-up Period: 4 weeks

Figure 3-1: Study Diagram



QD = once daily

Note: Study drug should be taken at the same time each day whenever possible.

During a screening period of no longer than 30 days, subjects will be randomized (on Day -7) to receive either one of three doses of NDI 034858 ([redacted]), or placebo on Day 1. The goal is to have approximately 65 subjects randomized per treatment group (1:1:1:1 ratio). During the treatment period, NDI 034858 ([redacted]) or placebo will be orally administered QD for 12 weeks. The 12-week treatment period will be followed by a 4-week safety follow-up period.

For scheduled study visits, subjects will come to the study site on 8 occasions: screening, Day 1, and Weeks 1, 2, 4, 8, 12 (end of treatment [EOT] / early termination visit [ET]), and 16 (end of study [EOS]).

Primary efficacy will be assessed using the American College of Rheumatology (ACR) 20 composite measure, derived from improvements of the number of tender and swollen joint count (TJC and SJC), patient global assessment of PsA pain (PGAAP) visual analog scale (VAS), patient global PsA assessment (PGA-PsA) VAS, physician global PsA assessment VAS

(PhGA-PsA), Health Assessment Questionnaire-Disability Index (HAQ-DI), and high sensitivity C-reactive protein (hsCRP).

Safety will be assessed by collecting adverse events (AEs), recording vital signs, performing physical examinations, and evaluating clinical laboratory and ECGs results.

Blood samples will be collected as follows to measure plasma concentrations of NDI-034858:

- Predose and 1 hour (± 5 min) post-dose on Day 1
- Predose, 1 hour (± 5 min) post-dose, and 4 hours (± 10 min) post-dose at Week 4
- Predose at Week 8
- Anytime at Week 12

[redacted]

No interim analysis is planned in this study.

3.2 Schedule of Assessments

For the schedule of assessments, please reference Table 1 of protocol Version 4.0 (May 05, 2023)

3.3 Endpoints

The efficacy endpoints listed below in this section specify only the Week 12 timepoint.

3.3.1 Primary Efficacy Endpoint and Estimand

3.3.1.1 Primary Efficacy Endpoint

The primary efficacy endpoint is defined as proportion of subjects achieving at least ACR20 response at Week 12. The primary clinical question of interest is to determine the effectiveness of different doses of NDI-034858 vs placebo on ACR20 response at week 12 in subjects with active PsA.

3.3.1.2 Primary Estimand

The estimand for the primary clinical question of interest is defined in Section 4.12.1.2.

3.3.2 Secondary Efficacy Endpoints

- Proportion of subjects achieving at least an ACR50 response at Week 12
- Proportion of subjects achieving at least an ACR70 response at Week 12
- Change from baseline (Day 1) in TJC at Week 12
- Change from baseline (Day 1) in SJC at Week 12
- Change from baseline (Day 1) in PGA-PsA at Week 12
- Change from baseline (Day 1) in PGAAP at Week 12
- Change from baseline (Day 1) in PhGA-PsA at Week 12
- Change from baseline (Day 1) in HAQ-DI score at Week 12
- Change from baseline (Day 1) in dactylitis count at Week 12, among subjects who have dactylitis at Day 1

- Change from baseline (Day 1) in Leed's Enthesitis Index (LEI) at Week 12, among subjects who have enthesitis at Day 1
- Proportion of subjects with Minimal Disease Activity (MDA) at Week 12
- Change from baseline (Day 1) in Disease Activity Index for Psoriatic Arthritis (DAPSA) at Week 12
- Proportion of subjects achieving 75% improvement from baseline (Day 1) in Psoriasis Area Sensitivity Index [(PASI)-75] at Week 12 among subjects with ≥ 3% body surface area (BSA) psoriatic involvement at Day 1
- Proportion of subjects achieving a Physician Global Assessment (PhGA) of Psoriasis (PsO) of 0 or 1 and at least a 2-point improvement from baseline (Day 1) at Week 12

3.3.2.1 Secondary Endpoints – Safety

- Incidence of treatment-emergent AEs (TEAEs), treatment-emergent serious adverse events (TESAEs), and Incidence of treatment emergent AEs of special interest (TEAESIs)
- Assessment of clinically relevant changes in vital signs, clinical laboratory parameters, and proportion of subjects with clinically relevant abnormal electrocardiograms (ECGs), and physical examinations

3.3.2.2 Secondary Endpoints – Pharmacokinetics

- Measurement of plasma concentrations of NDI-034858 in subjects receiving [redacted] of NDI-034858

3.3.3 Exploratory Endpoints

[redacted]

4.0 STATISTICAL METHODS

4.1 Data Quality Assurance

All tables, figures and data listings to be included in the report will be independently checked for consistency, integrity and in accordance with standard Parexel procedures.

4.2 Definitions

Baseline value

Baseline values of laboratory and pharmacokinetic (PK) data is defined as the last non-missing value (including scheduled or unscheduled measurements) prior to the first dose of study drug or randomization if no study drug was given. Both day and time of assessment will be considered when defining baseline for laboratory and PK data.

For all other endpoints, the baseline value is defined as the last non-missing value (including scheduled or unscheduled measurements) prior to or on the same day of the first dose of study drug. If multiple measurements occur at the same time, then the average of these measurements (for continuous data) or the worst among the measurements (for categorical data) should be considered as baseline.

Change from Baseline

Change from baseline is defined as post-baseline value minus baseline value.

Percent Change from Baseline

Percent change from baseline is defined as:

$$\text{Percent change from Baseline} = 100 \times \frac{\text{Postbaseline value} - \text{Baseline value}}{\text{Baseline value}}\%$$

Study Day 1

Study day 1 is defined as the date of first dose for treated subjects or the date of randomization for untreated randomized subjects.

Study Day

Study day of any event will be calculated relative to the date of first dose for treated subjects or the date of randomization for untreated randomized subjects.

If the event is prior to study day 1, then study day will be negative (prefixed with '-') and is calculated as:

Date of Event – Date of Day 1

If event is on or after study day 1, then study day is calculated as:

Date of Event – Date of Day 1 + 1

Treatment-Emergent Adverse Events (TEAEs)

TEAEs are defined as any adverse events (AEs) occurring (onset date/time) at the time of or after study drug dosing until study end (inclusive of Week 16).

Partial dates will be imputed as described in Section 4.3.2.1. In the event a start or end date is partial, the imputed dates will be used to determine treatment-emergence.

4.3 General Presentation Considerations

Continuous data will be summarized in terms of the mean, standard deviation (SD), median, minimum, maximum, 25th and 75th percentiles (Q1 and Q3), and number of observations. The minimum and maximum will be reported to the same number of decimal places as the raw data recorded in the database. The mean (arithmetic and geometric), median, Q1, and Q3 will be reported to one more decimal place, and SD (and the standard error [SE], if applicable) will be reported to two more decimal places, than the raw data recorded in the database. The geometric coefficient of variation (geoCV) will be reported as percentage to 1 decimal place. In general, the maximum number of decimal places reported shall be four for any summary statistic.

If the number of decimal places reported in the original raw data is varied, then use either the maximum raw number of reported decimal places or 3, whichever is the lowest, as a guide for the descriptive statistics.

For NDI-034858 plasma concentration, the geometric mean (geoMean) and corresponding coefficient of variation will be presented. The geoCV is calculated as below:

$$geoCV = \sqrt{e^{SD_{ln}^2} - 1}$$

Where SD_{ln}^2 is the variance of the log-transformed data.

If no subjects have data at a given time point, for example, then only n=0 will be presented. However, if 0<n<3, the n, minimum and maximum will be presented only. If n=3, n, mean, median, minimum, and maximum will be presented only. The other descriptive statistics will be left blank.

Categorical data will be summarized in terms of the number of subjects providing data at the relevant time point (n), frequency counts and percent. Unless otherwise stated, the denominator for the percentage calculations will be based on the number of subjects in the respective analysis set.

Percentages will be presented to one decimal place except 100%. 100% will be presented as 100% instead of 100.0%.Percentages will not be presented for zero counts. If sample sizes are small, the data displays will show the percentages, but any textual report will describe frequencies only.

When specified, a missing category shall be included only for categorical variables with missing data. The missing category will be omitted if there were no missing values for that variable.

Changes from baseline in categorical data will be summarized using shift tables where appropriate.

P-values greater than or equal to 0.001, in general, will be presented to three decimal places. P-values below 0.001 will be presented as "<0.001". P-values greater than 0.999 will be presented as ">0.999".

Confidence intervals (CI) will be presented to one more decimal place than the raw data. Unless specified otherwise, CIs will be calculated using the normal approximation. If there are less than 5 events, the Clopper Pearson calculation will be used.

In the event that a subject was re-randomized (ie, subject [redacted] was initially randomized but never dosed and randomized again with a new randomization number and dosed), the new randomization number and the data associated with the re-randomization number will be used for statistical analyses.

4.3.1 Analysis Window

Subjects do not always adhere strictly to the visit timing stated in the protocol. The designation of all visits, including unscheduled visits and early termination visit (for patients with premature study termination) will be assigned to the relevant analysis visit per the schedule of activities based on the following table by using the study days relevant to first dose of study drug. Accordingly, the study is divided into continuous and mutually exclusive analysis windows.

Table 4-1: Analysis Visit Windows

Analysis Visit (Visit Number)	Target Day	Parameters and Allowed Window (in days)				
		Safety and efficacy Parameters*	PASI/ PhGA-PsO/ BSA	[redacted]	ECG	Fasting Lipids
Baseline	1	≤1	≤1		≤1	≤1
Week 1 (Visit 3)	8	[2, 11]			[2, 11]	
Week 2 (Visit 4)	15	[12, 22]	[2, 22]		[12, 22]	
Week 4 (Visit 5)	29	[23, 43]	[23, 43]		[23, 43]	
Week 8 (Visit 6)	57	[44, 71]	[44, 71]		[44, 71]	
Week 12 (Visit 7)	85	[72, 99]	[72, 99]		[72, 99]	[72, 99]
Week 16 (Visit 8)	113	[100, EOS]	[100, EOS]			

EOS=End of Study, PhGA = Physician Global Assessment.

* Safety endpoints include Vital Signs, Clinical Laboratory Tests, ESR, Physical Exam (Complete/Targeted); efficacy endpoints include TJC68, SJC66, Dactylitis, PGA-PsA, PGAAP, PhGA-PsA, HAQ-DI, LEI, MDA, and hsCRP.

For repeated assessments collected within baseline window, see Section 4.2, Definitions.

For repeated assessments collected during post-baseline visits, if more than one assessment is available within the same analysis visit, the assessment with closest study day to the target study day will be reported for the analysis window. If two observations exist with the same distance to the target study day, then the last observation will be used. Analysis visit will be used for table summaries.

4.3.2 Handling of Missing Data

Every effort will be undertaken to avoid any missing or incomplete data, and missing data will not be imputed unless specified. If variables are imputed, the analysis dataset will contain a new variable with imputed value and the original variable will contain the original missing value.

Missing data handling related to efficacy endpoints is described in Section 4.12.1.3.

4.3.2.1 Partial dates and times

Inevitable partial dates will be imputed for start/end dates of AEs or prior/concomitant medications.

Partial Start Dates for AE and prior/concomitant medications

- If only day is missing, then it will be set to:
  - First day of the month that the event occurred, if month and year of the event start date is different than the month and year of the first dose date.

  - The day of first dose date, if month and year of the event start date is the same as month and year of the first dose date and month and year of the event end date is different.
  - The day of first dose date or day of event end date, whichever is earliest, if the month and year of the event start and month and year of the first dose date and month and year of the event end date is the same.
- If only month is missing, then it will be set to the earliest of the following:
  - January, as long as this date is after the first dose date
  - Month of the first dose date, if this date is the same day and year that the event occurred
- If the day and month are both missing, then it will be set to following:
  - Month and day of the first dose date, if the year of first dose is same as the year of the start date and the (imputed) event stop date is after the date of first dose
  - January 1 of the year, if the year of first dose date in not same as the year of the start date or the date of the (imputed) event end is before the date of first dose
- If dates are completely missing:
  - Use the date of first dose of study drug if the stop date is fully missing prior to any imputation per the rules below or not prior to the date of first dose of study drug
  - if the stop date is prior to the date of the first dose of study drug, then set the start date to the 1st of January of the year of the event stop date
- If time is missing, time will be imputed as:
  - 00:00 h if start date is not the same as the date of first dose of study drug
  - The time of first dose of study drug if the start date is the same as the date of first dose of study drug

<u>Partial End Dates for AE and prior/concomitant medications</u>

- If only day is missing, then it will be set to the earliest of the last follow-up date, last day of the month, or the date of death.
- If only month is missing, then it will be set to the earliest of the last follow-up date, December, or the date of death.
- If year is missing or event is ongoing, the end date will not be imputed.
- If both month and day are missing, then use December 31st of that year.
- If dates are completely missing, then use discharge date or data cut-off date. Discharge date refers to the date of the end of study visit for completed subjects or the date of discontinuation for subjects that were withdrawn. For any AEs with known start date after the date of discontinuation, the date of last contact will be used as the discharge date. For subjects still ongoing in the study at the time of the data cut-off, and for whom no discharge date is available, the date of the data cut-off will be used instead of the discharge date.
- If time is missing and (imputed) end date is not missing, time will be imputed as 00:00 h

If the imputed event end date is before the corresponding event start date, then the event end date will be set to the event start date.

4.3.2.2 Partial Date of First Diagnosis of PsA

Partial dates of the first diagnosis of PsA will be imputed as following:

- If only the month and year are specified and the month and year of first dose are not the same as the month and year of the start date, then the 1st of the month will be used.

- If only the month and year are specified and the month and year of first dose is the same as the month and year of the start date, then the date of first dosing will be used.
- If only the year is specified, and the year of first dose is not the same as the year of the start date, then January 1st of the year of the start date will be used.
- If only the year is specified, and the year of first dose is the same as the year of the start date, then then start date will be set as missing and no imputation will be applied
- If start date is completely unknown, then then start date will be set as missing and no imputation will be applied

4.3.2.3 Missing Data for AEs

In addition to the missing data handling rules for partial event dates, other missing AE data will be imputed for inclusion in the summary tables.

Any missing intensity will be queried for completion. If the intensity is still missing in the final data and the outcome of AE is death, then it will be deemed as Grade 5, and if the life-threatening is marked, then it will be deemed as Grade 4. Otherwise, the intensity still missing in the final data will be deemed as Grade 3.

Any missing causality will also be queried for completion, and any events that still have missing causality in the final data will be deemed as 'Related'.

4.4 Software

All report outputs will be produced using SAS® version 9.4 or a later version in a secure and validated environment.

4.5 Analysis Sets

For purposes of analysis, the following analysis sets are defined:

Population (Analysis Set)	Description
All Subjects Screened	All subjects who signed the informed consent
Intent-to-Treat (ITT) Analysis Set	All randomized subjects. Subjects will be included into the analysis as randomized.
Full Analysis Set (FAS)	All randomized subjects who receive at least one dose of study drug. Subjects will be included into the analysis as randomized.
Per-Protocol Set (PPS)	All FAS subjects who are compliant with study treatment, and without any major protocol deviations that could impact the primary endpoint.
Safety Analysis Set (SS)	All randomized subjects who receive at least one dose of study drug. Subjects will be included into the analysis based on actual treatment received, regardless of the treatment randomized.
PK Analysis Set (PKS)	All subjects in the Safety Analysis Set with at least one evaluable post-dose PK assessment.

PK = pharmacokinetics

All subjects who signed the informed consent will be used to report disposition of subjects screened. A summary will be presented for the reasons for screen failure overall and for subjects

who were rescreened and failed again. The ITT set will be used to summarize the disposition and discontinuation reasons for randomized subjects. The FAS will be used for all efficacy and baseline analyses. The PPS will be determined prior to database lock and the ITT Set and PPS will be used in the context of sensitivity analyses of the primary endpoint.

Protocol deviation and analysis set outputs will be produced. A Data Review Meeting will be arranged prior to the database lock/unblinding to identify analysis sets or data to be included/excluded from analysis including a review/confirmation of subjects with ICEs. Decisions made regarding the exclusion of subjects and/or subject data from analyses will be made prior to unblinding and will be documented in a data review report that will be approved by the Sponsor prior to the analysis.

The PPS includes all subjects in the FAS who met the following criteria:

1. Subjects who had completed 12 weeks of treatment with valid efficacy endpoint assessment at baseline and Week 12
2. Treatment compliance ≥80%
3. Without major protocol deviation that could impact the primary endpoint including but not limited to the following:
    a. Without violations of relevant inclusion/exclusion criteria
    b. Without receiving prohibited concomitant medications that could impact the efficacy assessments before Week 12

All safety analyses will be based on the SS.

All PK tabulations and statistical analyses will be based on the PKS.

The number and percentage of subjects in each analysis set by treatment group and overall will be summarized on ITT Set. By-subject listing of analysis population assignment will also be provided.

4.6 Study Subjects

4.6.1 Disposition of Subjects

A clear accounting of the disposition of all subjects who enter the study will be provided, from screening to study completion.

All subjects screened will be used to summarize below subject disposition:

- Number and reasons for screen failure (collected on Screening Disposition eCRF).
- Number of subjects screened by region, site, and overall, including first subject in date, last subject out date will be provided. Additionally, number of subjects by treatment group in each analysis set (ITT, FAS, PPS, SS, PKS) will also be provided.

In addition, the following subject disposition will be summarized on ITT and FAS by treatment group and overall:

- Number and percentage of subjects randomized
- Number and percentage of subjects who received at least one dose of study drug
- Number and percentage of subjects randomized but not treated who did not receive any study drug after randomization (only in ITT version).
- Number of subjects who completed study drug

- Number of subjects who completed the study but prematurely discontinued the study drug
- Number and percentage of subjects who prematurely discontinued from study drug and primary reason for discontinuation
- Number of subjects who completed the study
- Number and percentage of subjects who prematurely discontinued from study and primary reason for discontinuation

By-subject listings of disposition details will also be provided for all subjects screened, including the date of informed consent, the date and time of first and last dosing, and the date of premature termination and primary reason (if applicable).

4.6.2 Protocol Deviations

Protocol deviations (PDs) will be classified as "major" or "minor" on an ongoing basis by the clinical study team and sponsor.

Major PDs are defined as those deviations from the protocol likely to have an impact on the subject's rights, safety, well-being, and/or the validity of the data for analysis. Minor PDs include all deviations from the protocol excluding the major PDs.

The number and percentage of subjects with at least one major PD, and major PD categories, as well as at least one minor PD will be summarized by treatment and overall on the SS. By-subject listing of all PDs will also be provided.

4.7 Demographic and Other Baseline Characteristics

Demographics and other baseline characteristics as described below will be summarized on FAS by treatment groups and overall. Additional demographic or baseline data may be added to the summary tables.

- Age (years)
- Age (<65 years, ≥65 years)
- Age at initial disease diagnosis (<18, 18-39, ≥40 years)
- Sex (Male, Female)
- Race (American Indian or Alaska Native, Asian, Japanese descent, Black or African American, Native Hawaiian or Other Pacific Islander, White, Unknown, Not Available)
- Race (White, Non-White)
- Ethnicity
- Weight (kg)
- Height (cm)
- Body Mass Index (kg/m^2)
- Body Mass Index (kg/m^2) (Normal [<25 kg/m^2], Overweight [≥25 to <30 kg/m^2], Obese [≥30 kg/m^2])
- Psoriatic arthritis subtype (collected in Diagnosis eCRF)
- Duration of disease
- Duration of disease (<Median, ≥Median)
- Region (USA/Germany, Eastern Europe) [according to randomization stratum]
- Region (USA/Germany, Eastern Europe) [according to eCRF]

- TJC
- SJC
- SJC (<median, ≥median)
- PhGA-PsA
- PGA-PsA
- PGAAP
- PhGA-PsO
- Dactylitis score
- Dactylitis (total score ≥1)
- LEI
- Enthesitis (LEI ≥1)
- DAPSA
- BSA psoriatic involvement ≥3%
- BSA psoriatic involvement
- Baseline PASI (Overall, <3% BSA, ≥3% BSA)
- HAQ-DI score
- [redacted]
- hsCRP
- [redacted]
- Prior treatment with biologics (Yes, No, if yes, number tried [1, 2, ≥3])
- Prior treatment with biologics or non-traditional DMARDs (Yes, No) [according to randomization stratum]
- Prior treatment with biologics or non-traditional DMARDs (Yes, No) [according to eCRF]
- Prior treatment with traditional/conventional DMARDs (Yes, No)
- Number of traditional/conventional DMARDS tried in subjects with prior treatment with traditional/conventional DMARDs (1, 2, ≥3)
- Prior anti-TNF use (Yes, No)

Prior anti-TNF use is determined by medications with the ATC4 code L04AB collected in Prior Biologic Therapy for Psoriatic Arthritis eCRF

- Prior treatment with anti-TNF, number tried (1, 2, ≥3)
- Prior anti-TNF failure due to lack of efficacy (Yes, No)
- Baseline use of sulfasalazine (under sulfasalazine at the baseline visit) (Yes, No)
- In subjects taking sulfasalazine, daily dose of sulfasalazine at NDI-034858 initiation (Section 4.9)
- Baseline use of methotrexate (under methotrexate at the baseline visit) (Yes, No)
- In subjects taking methotrexate, weekly dose of methotrexate at NDI-034858 initiation (Section 4.9)
- Baseline use of steroid (Yes, No)

- In subjects taking steroids, daily dose at NDI-034858 initiation (Section 4.9): daily dose of steroid will be calculated by averaging the daily dose of all types of steroids the subjects are taking after converting the dose to prednisone equivalent (see Appendix 6)
- Baseline use of NSAIDs (under NSAID at the baseline visit) (Yes, No)
- For baseline use of medications listed above, below calculation rules for weekly dose will be applied:
- If Frequency = Daily, then weekly dose = Dose × 7
- If Frequency = Twice Daily, then weekly dose = Dose × 7 × 2
- If Frequency = Three Times Daily, then weekly dose = Dose × 7 × 3
- If Frequency = Four Times Daily, then weekly dose = Dose × 7 × 4
- If Frequency = Every other Day, then weekly dose = Dose × 7 × 0.5
- If Frequency = Once Weekly, then weekly dose = Dose
- If Frequency = Twice Weekly, then weekly dose = Dose × 0.5
- If Frequency = Monthly, then weekly dose = Dose × 0.25
- If Frequency = Unknown, then weekly dose = missing
- If Frequency = Other, then weekly dose will be calculated according to the detailed information collected (e.g., number of tablets taken for one week). If detailed information is not available, then weekly dose will be set as missing.
- If Frequency = As Necessary, then weekly dose = missing

For daily dose calculation, apply the above calculation rules and divide by 7.

By-subject listing of demographic and other baseline characteristics as well as subgroups will also be provided for FAS.

4.8 Medical and Surgical History

Medical and surgical history at screening will be collected under Medical History form of the eCRF and will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) Version 25.0 or later. Medical and surgical history will be summarized on the SS by treatment and overall by system organ class (SOC) and preferred term (PT). By-subject listing of medical and surgical history will also be provided.

4.9 Prior and Concomitant Medication

Prior medication is defined as any medications that started and stopped on or before the first dose of study drug. Concomitant medication is defined as any medications that started before the first dose of study drug and stops after the first dose of study drug, as well as any medications that started on or after the first dose of study drug. Baseline concomitant medication at study drug initiation is defined as any medications being taken on the day or ongoing at the day of first study drug initiation.

Medication start and stop dates will be compared to the date of first dose of study drug to allow medications to be classified as either Prior or Concomitant. Medications starting after the study completion/withdrawal date will not be classified or summarized.

Incomplete start and end date of medications will be imputed as described in Section 4.3.2.1.

Medications will be coded using the World Health Organization Drug (WHO Drug) Global March 2021 B3 or higher. Prior and concomitant medications will be summarized by Pharmacological

Subgroup (level 3), and preferred drug name on the SS by treatment and overall. Subjects with more than one medication in a given ATC level and preferred term will be counted only once.

By-subject listings of both prior and concomitant medications will also be provided for SS. A separate summary table and listing will also be provided for:

- Prior non-traditional DMARDs therapy for PsA collected in Prior Non-Traditional DMARDs Therapy for Psoriatic Arthritis eCRF
- Prior and concomitant traditional/conventional DMARDs therapy as marked on the Prior and Concomitant Medications eCRF
- Prior biologic therapy for PsA collected in Prior Biologic Therapy for Psoriatic Arthritis eCRF

4.10 Concomitant Procedures

A by-subject listing of concomitant procedures will be provided on the SS.

4.11 Treatment Compliance

Treatment compliance (%) will be estimated as:

$$\frac{\text{Total Amount of Study Drug Administered [redacted]}}{\text{Total Amount of Study Drug expected [redacted]}} \times 100\%$$

The study drug administered per blister packet will be obtained by subtracting the number of returned [redacted] from the number of dispensed [redacted] Both the number of dispensed and returned [redacted] should come from the same blister packet (i.e. with the same kit ID number). Dispensed packets without return information will not be included in the calculation. The total amount of study drug administered in [redacted] will be the sum of study drug administered from all packets with complete dispensing and return information.

The total amount of study drug expected in [redacted] is defined as (treatment duration ×2) (mg). The treatment duration is calculated as (date of last dose – date of day 1+1). The expected dose for all subjects is two [redacted] daily, thus the total amount of study drug expected will be (duration ×2) [redacted]

Examples for calculating treatment compliance can be found in Appendix 1.

Significant noncompliance is defined as missing 3 consecutive doses (ie, 3 days in a row) or as missing a total of 10 doses.

A summary of the treatment compliance and number and percentage of subjects with significant noncompliance (along with reasons such as COVID-19, other reasons) will be provided on the SS by treatment. A by-subject listing of treatment compliance will also be provided.

4.12 Efficacy Evaluation

4.12.1 Analysis and Data Conventions

Unless otherwise specified, all statistical tests will be 2-sided and performed at a significance level of 0.05. When applicable, 95% CIs will be presented.

Selected efficacy endpoints will be calculated according to the algorithm specified in the respective section using response of individual component rather than results calculated by Medidata.

A by-subject listing for selected efficacy endpoints will be provided.

4.12.1.1 Adjustments for Covariates

The binary efficacy analyses will be adjusted for the following randomization stratification factors:

1. Prior treatment with biologics or non-traditional DMARDs [Yes, No]
2. Region [USA/Germany (USA, Germany), Eastern Europe (Czech Republic, Poland)]

All continuous efficacy endpoints will be adjusted for baseline score and the randomization stratification factors (prior treatment with biologics or non-traditional DMARDs [Yes, No], and region [USA/Germany, Eastern Europe]).

Following the intent to treat principle, subjects will be analyzed with the strata to which they were randomized. A sensitivity analysis will be performed on the primary endpoint using the actual stratification factors re-derived based on the data collected on the eCRF. A summary of mis-stratifications will be provided.

4.12.1.2 Estimands

Objectives	**Estimands**
[Primary] To assess the efficacy of NDI-034858 (also known as TAK-279) orally administered QD at [redacted] for 12 weeks on the rheumatological signs, symptoms, and function in subjects with active PsA	The **primary estimand** will be based on a mixture of "composite strategy" and "treatment policy strategy" to assess the treatment effect of TAK-279 compared to placebo in subjects with active PsA. **Population**: The target population consists of subjects with active PsA as defined by the protocol inclusion/exclusion criteria. The analysis population for the primary endpoint is FAS which includes all randomized subjects who receive at least one dose of study drug. **Endpoint of interest**: primary efficacy endpoint, assessed as the proportion of subjects achieving at least an ACR20 response at Week 12. **Treatment of interest**: TAK-279 [redacted] QD, TAK-279 [redacted] QD, or TAK-279 [redacted] QD versus placebo.

Objectives	Estimands
	Strategy for addressing ICEs: The following events will be considered ICEs: • ICE1: Premature discontinuation of the study drug due to the lack of efficacy, AE, or any other reason prior to Week 12. • ICE2: Death due to any cause prior to Week 12. • ICE3: Use of a prohibited medication as PsA treatment or of another treatment expected to have an effect on PsA clinical outcomes prior to the assessment of ACR20 at Week 12. • ICE4: Use of other prohibited medications (not PsA treatment or other treatments not expected to have a clinical effect on PsA clinical outcomes) prior to the assessment of ACR20 at Week 12. The "composite strategy" will be implemented to handle ICE1, ICE2, and ICE3. Subjects with ICE1, ICE2, or ICE3 will be considered as non-responders (ie, treatment failure). The "treatment policy strategy" will be implemented to handle ICE4. Subjects with ICE4 will be encouraged to be followed for efficacy and safety until EOS unless they have withdrawn consent for study participation for other reasons. The observed ACR20 will be considered regardless of the occurrence of the ICE following intent-to-treat principle as this event reflects clinical practice. The subjects who have nevertheless missing endpoint data at Week 12 will be imputed as non-responder (ie, treatment failure). If there are subjects with multiple ICEs applying different strategies, the ICEs will be handled in the above order. The subjects who have otherwise missing endpoint data at Week 12 will be imputed as non-responders (ie, treatment failure). Population-level summary: Difference in the proportion of subjects achieving ACR20 response at Week 12 between each TAK-279 treatment group and the placebo group.
	The **supplementary estimand** will be based on "treatment policy strategy" to assess the treatment effect of TAK-279 compared to placebo in subjects with active PsA. Population, Endpoint of interest, and Treatments are the same as for the primary estimand of the primary efficacy endpoint. Strategy for addressing ICEs. The following events will be considered ICEs: • ICE1: Premature discontinuation of the study drug due to the lack of efficacy, AE, or any other reason prior to Week 12. • ICE2: Death due to any cause prior to Week 12.

Objectives	Estimands
	• ICE3: Use of a prohibited medication as PsA treatment or of another treatment expected to have an effect on PsA clinical outcomes prior to the assessment of ACR20 at Week 12. • ICE4: Use of other prohibited medications (not PsA treatment or other treatments not expected to have a clinical effect on PsA clinical outcomes) prior to the assessment of ACR20 at Week 12. The "treatment policy strategy" will be implemented to handle ICE1 to ICE4. Subjects with ICEs will be encouraged to be followed for efficacy and safety until EOS unless they have withdrawn consent for study participation for other reasons. The observed ACR20 will be considered regardless of the occurrence of the ICEs following intent-to-treat principle as this event reflects clinical practice. The subjects who have nevertheless missing endpoint data at Week 12 will be imputed as non- responder (ie, treatment failure). Population-level summary: Difference in the proportion of subjects achieving ACR20 response at Week 12 between each TAK-279 treatment group and the placebo group.
[Secondary] To assess additional evaluations of efficacy of NDI-034858 orally administered QD at [redacted] for 12 weeks in subjects with active PsA	The **secondary estimands** for binary secondary efficacy endpoints listed below will be based on a mixture of "composite strategy" and "treatment policy strategy" the same as to the primary estimand for primary efficacy endpoint to assess the treatment effect of TAK-279 compared to placebo in subjects with active PsA. Population: The target population consists of subjects with active PsA as defined by the protocol inclusion/exclusion criteria. The analysis population for the primary endpoint is FAS which includes all randomized subjects who received at least one dose of study drug. Endpoint of interest: the following binary secondary efficacy endpoints at Week 12. • ACR50 response at Week 12 • ACR70 response at Week 12 • MDA at Week 12 • PASI-75 at Week 12 among subjects with ≥3% BSA psoriatic involvement at Day 1 • Physician Global Assessment of Psoriasis of 0 or 1 and at least a 2 -point improvement from baseline at Week 12 Treatments, ICEs, ICE handling strategies, and population-level summaries are the same as for the primary estimand of primary efficacy endpoint.
[Secondary] To assess additional evaluations of efficacy of NDI-034858 orally administered QD at [redacted] for 12 weeks in subjects with active PsA	The **secondary estimands** for continuous secondary efficacy endpoints will be based on "treatment policy strategy" to assess the treatment effect of TAK-279 compared to placebo in subjects with active PsA. Population: The target population consists of subjects with active PsA as defined by the protocol inclusion/exclusion criteria. The

Objectives	Estimands
	analysis population for the primary endpoint is FAS which includes all randomized subjects who receive at least one dose of study drug. Endpoints of interest: the following continuous secondary efficacy endpoints at Week 12. • Change from baseline (Day 1) in tender joint count at Week 12 • Change from baseline (Day 1) in swollen joint count at Week 12 • Change from baseline (Day 1) in Patient Global Assessment of Psoriatic Arthritis at Week 12 • Change from baseline (Day 1) in Patient Global Assessment of Psoriatic Arthritis pain at Week 12 • Change from baseline (Day 1) in Physician Global Assessment of Psoriatic Arthritis at Week 12 • Change from baseline (Day 1) in HAQ-DI score at Week 12 • Change from baseline (Day 1) in dactylitis count at Week 12, among subjects who have dactylitis at Day 1 • Change from baseline (Day 1) in LEI at Week 12, among subjects who have enthesitis at Day 1 • Change from baseline (Day 1) in Disease Activity Index for DAPSA at Week 12 Treatment of interest: TAK-279 [redacted] QD, TAK-279 [redacted] QD, or TAK-279 [redacted] QD versus placebo. Strategy for addressing ICEs. The following events will be considered ICEs: • ICE1: Premature discontinuation of the study drug due to the lack of efficacy, AE, or any other reason prior to Week 12. • ICE2: Death due to any cause prior to Week 12. • ICE3: Use of a prohibited medication as PsA treatment or of another treatment expected to have an effect on PsA clinical outcomes prior to the assessment of ACR20 at Week 12. • ICE4: Use of other prohibited medications (not PsA treatment or other treatments not expected to have a clinical effect on PsA clinical outcomes) prior to the assessment of ACR20 at Week 12. The "treatment policy strategy" will be implemented to handle ICE1 to ICE4. Subjects with ICEs will be encouraged to be followed for efficacy and safety until EOS unless they have withdrawn consent for study participation for other reasons. The observed data will be considered regardless of the occurrence of the ICEs following intent-to-treat principle as this event reflects clinical practice. Missing endpoint data despite effort to retain subjects after treatment discontinuation will be handled under the missing at random (MAR) assumption.

Objectives	Estimands
	Population-level summary: Treatment difference in LS means at Week 12 between each TAK-279 treatment group and the placebo group.
ACR20: American College of Rheumatology 20% response; AE: adverse events; BSA: body surface area; Dudy APSA: Disease Activity Index for Psoriatic Arthritis; EOS: end of study; FAS: full analysis set; HAQ-DI: Health Activities Questionnaire – Disability Index; ICEs: intercurrent events; LEI: Leed's Enthesitis Index; LS: least squares; MAR: missing at random; MDA: Minimal Disease Activity; PsA: psoriatic arthritis; [redacted]; QD: once daily; SAP: statistical analysis plan.	

4.12.1.3 Handling of Missing Data

The following approaches will be used to handle missing data for the efficacy endpoints taken into consideration the respective estimand framework and strategy to handle ICEs as specified in Section 4.12.1.2. If there are subjects with multiple ICEs applying the same strategies, the earliest onset of ICE will be used as the starting point for imputation.

Non-Responder Imputation (NRI)

The NRI approach will categorize any subject who has a missing data to determine the binary endpoint status at a specific visit as non-responder (eg, treatment failure) for that visit.

MMRM

For continuous efficacy endpoints, the efficacy data will be analyzed as observed via MMRM assuming MAR.

Hybrid Imputation Method

For primary efficacy endpoint, ACR20 at Week 12, a supportive analysis will be performed to evaluate the robustness of missing data handling through a hybrid imputation method under different missing data mechanisms (Chen et al, 2021). The subjects who discontinued study drug prematurely due to lack of efficacy, due to AE or death, and subjects who received a prohibited medication as PsA treatment prior to Week 12 will be imputed as non-responders via NRI under MNAR; subjects who discontinued study drug prematurely due to other reason and subjects who have otherwise missing endpoint data to determine the status of ACR20 at Week 12 will be imputed via MI using a fully conditional specification (FCS) method under MAR.

For MI, each component of ACR20 will firstly be multiply imputed before deriving ACR20 to minimize the information loss. For example, missing data for TJC will be multiply imputed first before combining with other component data to derive ACR20 response. The MI will be conducted by treatment group via FCS regression method. The missing baseline data, if any, will be imputed using randomization stratification factors. The missing data from each subsequent postbaseline visit will be imputed using all previous visits and the randomization stratification factors in a stepwise fashion. A hundred (100) imputation datasets will be generated. Once complete datasets for each component of ACR20 are generated, ACR20 will be derived as specified in Section 4.12.2.

Combine the imputed ACR20 via NRI under MNAR and via MI under MAR to obtain a complete dataset containing all subjects in the analysis population. The comparison of each dose group to placebo using the Mantel-Haenszel (MH) test specified in Section 4.12.2.1 will be performed on the 100 datasets. An estimate for the difference in proportions and corresponding standard error

will be computed for each of these datasets. The results from these analyses will be combined into a single estimate of the difference in predicted proportion of response and a 2-sided 95% CI via Rubin's rules (ie, SAS® PROC MIANALYZE).

4.12.1.4 Multiple Comparisons/Multiplicity

A fixed-sequence hierarchical testing approach will be applied to primary endpoint in the following pre-specified order:

- [redacted] vs placebo comparison,
- [redacted] vs placebo comparison,
- [redacted] vs placebo comparison.

If one dose vs. placebo is tested and found not to be significant (i.e., $p \geq 0.05$), all subsequent comparisons of lower doses vs. placebo could not be claimed if found significant. In that case, comparison will be performed but p-values will be considered as nominal.

No multiplicity adjustment is planned for other efficacy endpoints in this study and nominal p-value may be presented as appropriate.

4.12.1.5 Interim Analyses

No interim analyses are planned.

4.12.1.6 Examination of Subgroups

The treatment effect for the primary efficacy variable will be examined for the following subgroups:

1. Age (<65, ≥65 years)
2. Gender (Male, Female)
3. Race (White, Non-white)
4. BMI (Normal [<25 kg/m^2], Overweight [≥25 to <30 kg/m^2], Obese [≥30 kg/m^2])
5. Prior treatment with biologics as recorded on the eCRF (Yes, No)
6. Region (USA/Germany, Eastern Europe) [as recorded on the IRT]
7. Number of swollen joints (<median, ≥median)
8. Duration of disease (<median, ≥median)
9. [redacted]
10. Baseline presence of enthesitis (Yes, No)
11. Baseline presence of dactylitis (Yes, No)
12. Baseline use of methotrexate (Yes, No)
13. Age at disease onset (<18, 18-39, ≥40 years)
14. Prior anti-TNF use (Yes, No)

The subgroup analysis for primary efficacy endpoints at Week 12 will be analyzed using the primary estimand framework as described for the primary efficacy analysis in Section 4.12.1.2 (applying ICEs and missing data imputation). A forest plot will be presented as well.

4.12.2 Primary Efficacy Analysis

4.12.2.1 Derivation of Primary Efficacy Endpoint

The primary efficacy endpoint is a binary variable, where a subject will be classified as responder if they achieve ACR20 at Week 12. The ACR20 is a composite measure that needs to satisfy the following criteria:

1. Improvement of 20% in the number of tender joints (TJC68);

And

2. Improvement of 20% in the number of swollen joints (SJC66);

And

3. A 20% improvement in at least three of the following five criteria:
   - o PGA-PsA
   - o PGAAP
   - o PhGA-PsA
   - o HAQ-DI
   - o hsCRP

Given the limitation of Medidata eCOA system, three components of ACR20 (ie, TJC68, SJC66, HAQ-DI) will be re-derived based on the source data.

While computing ACR20, missingness in individual component will affect the derivation differently. The following algorithm should be followed:

a. If TJC68 ≥20% and SJC66 ≥20% and at least 3 of 5 components in criteria 3) has improvement ≥20%, then ACR20 will be considered as responder;
b. If TJC68, SJC66, or at least 3 of the 5 components in criteria 3) has improvement <20%, then ACR20 will be considered as non-responder;
c. For all other cases, the ACR20 will be missing since insufficient data for determination.

The following table illustrates examples for ACR response determination.

Scenario	TJC68	SJC66	Component 1	Component 2	Component 3	Component 4	Component 5	ACR20
a)	≥20%	≥20%	≥20%*	≥20%*	≥20%*			Responder
b)	≥20%	<20%	Regardless status of the components					Non-Responder
	<20%	≥20%	Regardless status of the components					Non-Responder
	<20%	<20%	Regardless status of the components					Non-Responder
	.	<20%	Regardless status of the components					Non-Responder
	<20%	.	Regardless status of the components					Non-Responder
	≥20%	≥20%	<20%*	<20%*	<20%*			Non-Responder
	.	.	<20%*	<20%*	<20%*			Non-Responder
c)	≥20%	.	Fewer than 3/5 components with <20%					Missing
	.	≥20%	Fewer than 3/5 components with <20%					Missing

Scenario	TJC68	SJC66	Component 1	2	3	4	5	ACR20
	≥20%	≥20%	At least 3 of the 5 components in criteria 3) missing					Missing
	≥20%	≥20%	≥20%	≥20%	<20%	<20%	.	Missing
	.	.	≥20%	≥20%	≥20%	≥20%	≥20%	Missing

*Components 1, 2, and 3 are used as an example in the table. The "at least 3 components" can be any combination and can be more than 3 components.
Note: scenarios in the table above are provided as examples and may not be all-inclusive of all possibilities.

For any of the ACR20 components, if the baseline value equals 0 and the post-baseline value is greater than or equal to 0, then the percent improvement of that component will be treated as 0 for the purpose of ACR response calculations.

4.12.2.2 Main Analysis Approach

Comparison of the primary endpoint will be made between each dose group and the placebo group using a two-sided MH test of the difference in two proportions stratified by the randomization stratification factors (prior treatment with biologics or non-traditional DMARDs [Yes, No] and region [USA/Germany, Eastern Europe]). The estimated MH risk difference will be summarized along with the two-sided 95% CI using MH stratum weights (Mantel, 1959) and the Sato variance estimator (Sato, 1989).

The count and proportion of subjects with ACR20 response following primary estimand will be provided by study visit and treatment arm. A line plot showing the proportion of subjects with ACR20 response over time within each treatment group will be provided.

4.12.2.3 Supplementary Analysis

The primary efficacy endpoint will be analyzed following the supplementary estimand. The observed ACR20 will be analyzed regardless of the occurrence of the ICEs. The data that is nevertheless missing will be imputed as non-response via NRI.

Comparison of the primary endpoint will be made between each dose group and the placebo group using the same method as used for the main analysis approach for the primary endpoint.

The count and proportion of subjects with ACR20 response following supplementary estimand will be provided by study visit and treatment arm. A line plot showing the proportion of subjects with ACR20 response over time within each treatment group will be provided.

4.12.2.4 Sensitivity Analysis

The following analyses will be performed for the primary efficacy endpoint to assess the robustness of the primary endpoint result. These analyses will be conducted using the FAS unless otherwise specified.

Type	Sensitivity Analysis
1	Repeat the primary endpoint analysis using ITT and PPS.
2	In the event that there is discrepancy between original randomization stratification factors assigned at the time of randomization in IRT system and the data collected in EDC, the stratification factors will be re-derived based on the EDC data subsequent to randomization. The primary efficacy analysis as described in Section 4.12.2.2 will be repeated on the FAS using the MH test adjusting for the re-derived stratification factors.
3	Apply a hybrid imputation approach under different missing data mechanism to evaluate the impact of missing data handling in the primary endpoint analysis via NRI under MNAR and MI under MAR (see Section 4.12.1.3).
4	To evaluate the robustness of missing data handling, additional sensitivity analysis may be performed, such as tipping point analysis.
5	Sensitivity analyses of the ACR20 response will be conducted using the TJC and SJC score calculated by the alternative algorithms via a proportional extrapolation method (see Section 4.12.3.2.3) using FAS.

In addition, a multiple comparison procedure modelling (MCP-Mod) analysis may be conducted to support the dose selection for future trial/s.

4.12.3 Secondary Efficacy Analysis

Secondary endpoints will be analyzed using the FAS unless otherwise specified.

4.12.3.1 Analysis Approach for Secondary Efficacy Endpoints

Secondary binary efficacy endpoints

The secondary binary efficacy endpoints at Week 12 will be analyzed using the secondary estimand framework as described in Section 4.12.1.2 (applying ICEs and missing data imputation).

The count and proportion of subjects achieving binary efficacy endpoints will be provided by study visit and treatment arm. For binary endpoints, a line plot showing the proportion of subjects achieving response by treatment group over time will be provided.

Comparison of each binary efficacy endpoint will be made between each dose group and the placebo group using the same method as described for the primary efficacy endpoint. The estimated MH risk difference will be summarized along with the two-sided 95% CI using MH stratum weights (Mantel, 1959) and the Sato variance estimator (Sato, 1989). Nominal p-values will be presented.

Secondary continuous efficacy endpoints

For longitudinal continuous secondary efficacy endpoints, the changes from baseline will be analyzed as observed regardless of ICEs (ie, using treatment policy strategy) using the mixed model for repeated measures (MMRM) where the change from baseline will be the dependent variable, treatment group, visit, treatment-by-visit interaction and the randomization stratification factors (prior treatment with biologics or non-traditional DMARDs [Yes, No], and region [USA/Germany, Eastern Europe]) will be the fixed effects, and baseline score and will be included as a covariate. The parameter estimations will be based on the assumption of data being missing at random (MAR) and using the method of restrictive maximum likelihood (REML). The Kenward-Roger approximation (Kenward MG, 1997) will be used to estimate the denominator degrees of freedom. An unstructured (UN) covariance matrix will be used to model the within-subject correlation. If this model fails to converge, then covariance matrix in the order of compound symmetry (CS), autoregressive (1) [AR(1)], and Toeplitz (TOEP) will be used.

The treatment effect measured by the continuous secondary efficacy endpoints at Week 12 between each dose group and the placebo group will be estimated based on the least square (LS) mean treatment difference at Week 12 obtained from the MMRM. The LS mean, the LS mean treatment difference, the associated 95% CI, and nominal p-value for each timepoint obtained from the MMRM will also be provided. A line plot showing the least square mean (SD) over time within each treatment group from MMRM will be provided.

4.12.3.2 Derivation of Secondary Endpoints

4.12.3.2.1 Proportion of subjects achieving ACR50 Response at Week 12

The ACR50 will be derived using the same derivation rules as for ACR20 except that the improvement of each of the seven components will be based on a 50% improvement instead of 20%.

4.12.3.2.2 Proportion of subjects achieving ACR70 response at Week 12

The ACR70 will be derived using the same derivation rules as for ACR20 except that the improvement of each of the seven components will be based on a 70% improvement instead of 20%.

4.12.3.2.3 Change from baseline in tender joint count (TJC) at Week 12 and change from baseline in swollen joint count (SJC) at Week 12

The tender joint count in 68 joints (TJC68) and swollen joint count in 66 joints (SJC66) are instruments measuring the peripheral arthritis activity. A total of 68 and 66 joints are assessed for tenderness and swollenness, respectively (note: the hips are assessed for tenderness but not for swelling).

The following joints are included in the assessment:

- Temporomandibular, sternoclavicular, acromioclavicular, glenohumeral, elbow, wrist, hip (TJC only), knee, ankle, tarsus;
- Metacarpophalangeal, finger proximal interphalangeal, finger distal interphalangeal, metatarsophalangeal, toe proximal interphalangeal

TJC68 is scored as a sum of the tender joints and SJC66 is scored as a sum of the swollen joints. When counting the number of joints with tenderness or swelling, the following scoring rules will be used:

- Response as "Present" will be given a score of 1,
- Response as "Absent", "Not Done", or "Not Applicable", or joints with missing response will be given a score of 0.
- Any response other than "Present", "Absent", "Not Done", or "Not Applicable" is considered as invalid response (eg, "Present, Not Done", "Present, Not Applicable", "Absent, Not Done", "Absent, Not applicable") and will be given with a score of 0.

The sum of the scores of all 68 or 66 joints will be used as the TJC68 or SJC66 scores, respectively, for all analyses involving tender/swollen joint counts. Injected joints will be counted as both tender and swollen from the date of injection up to the end of the study. If a subject had received any joint injection since last visit but without joint location specified, then 1 will be added to TJC68 and SJC66, respectively.

Note, given the limitation of the Medidata eCOA system, instead of using scores calculated by Medidata, TJC68 and SJC66 scores used in all analyses will be re-derived based on the assessment of individual joints. In the event these scores are discrepant, the differences will be documented in the CSR narrative section.

Acknowledging the data limitation, two sets of sensitivity analyses will be conducted using TJC68 and SJC66 estimated by proportional extrapolation method, which has the advantage of providing meaningful imputation of the joints that are replaced, not assessed, or without valid responses. The sum of the scores of all joints will be divided by the number of joints with valid non-missing response, i.e., "Present" or "Absent". This proportion will then be used to impute joint counts for the joints with missing score.

$$TJC68 = 68 \times \frac{The\ Total\ Number\ of\ Joints\ with\ Tenderness\ (\textit{"Present" or 1})}{The\ Total\ Number\ of\ Joints\ with\ non\text{-}missing\ Response}$$

$$SJC66 = 66 \times \frac{The\ Total\ Number\ of\ Joints\ with\ Swelling\ (\textit{"Present" or 1})}{The\ Total\ Number\ of\ Joints\ with\ non\text{-}missing\ Response}$$

Injected joints will be counted as both tender and swollen from the date of injection to the end of study. If a subject had received any joint injection since last visit but without joint location specified, 1 will be added to TJC68 and SJC66, respectively. When more than 50% of the joints have missing score, the TJC68 and SJC66 will be set as missing.

The two sets of sensitivity analyses follow the same proportional extrapolation approach but differ in the way of handling invalid response. Separate scoring rules are described as follows:

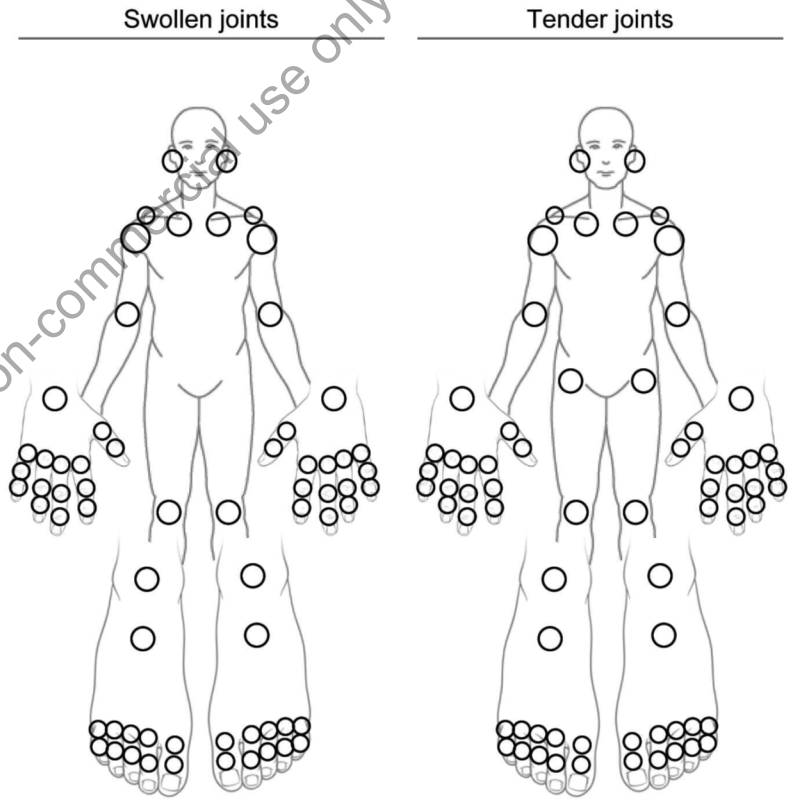
Scoring rules for sensitivity analysis #1

- Response as "Present" will be given a score of 1.
- Response as "Absent" will be given a score of 0.
- Response as "Not Done", "Not Applicable", or missing will be set as missing.
- Invalid response (eg, "Present, Not Done", "Present, Not Applicable", "Absent, Not Done", "Absent, Not applicable") will be set as missing.

Scoring rules for sensitivity analysis #2

- Response as "Present" will be given a score of 1
- Response as "Absent" will be given a score of 0
- Response as "Not Done", "Not Applicable", or missing will be set as missing.
- Response containing "Present" (eg, "Present, Not Done" or "Present, Not Applicable") will be given a score of 1
- Response containing "Absent" (eg, "Absent, Not Done", "Absent, Not applicable") will be given a score of 0

	Swollen joints	Tender joints
Temporomandibular joint	_____(0-2)	_____ (0-2)
Sternoclavicular joints	_____ (0-2)	_____ (0-2)
Acromioclavicular joints	_____ (0-2)	_____ (0-2)
Glenohumeral(s)	_____ (0-2)	_____ (0-2)
Elbow(s)	_____ (0-2)	_____ (0-2)
Wrist(s)	_____ (0-2)	_____ (0-2)
Metacarpal phalangeal joints	_____ (0-10)	_____ (0-10)
Finger Proximal interphalangeal joints	_____ (0-10)	_____ (0-10)
Finger Distal interphalangeal joints	_____ (0-8)	_____ (0-8)
Hip(s)	NA	_____ (0-2)
Knee(s)	_____ (0-2)	_____ (0-2)
Ankle(s)	_____ (0-2)	_____ (0-2)
Tarsus/Midfoot(feet)	_____ (0-2)	_____ (0-2)
Metatarsal phalangeal joints	_____ (0-10)	_____ (0-10)
Toe PIP(s)	_____ (0-10)	_____ (0-10)
Total joint counts	_____ (0-66)	_____ (0-68)



4.12.3.2.4 Change from baseline in Patient Global Assessment of Psoriatic Arthritis (PGA-PsA) at Week 12 and change from baseline in Patient Global Assessment of Psoriatic Arthritis Pain (PGAAP) at Week 12

In this study, PGA-PsA and PGAAP are scored using a Visual Analogue Scale (VAS) which is anchored on an unnumbered 100-mm horizontal line. The PGA-PsA VAS is anchored at the ends,

with 0 being "very good, no symptoms" and 100 being "very poor, severe symptoms". The PGAAP VAS is anchored at the ends, with 0 is 'no pain' and 100 is 'most severe pain'.

4.12.3.2.5 Change from baseline in Physician Global Assessment of Psoriatic Arthritis (PhGA-PsA) at Week 12

The investigator or qualified sub-investigator will assess the subjects' overall disease status, taking into account signs, symptoms, and function, of all components of joint and skin which is affected at the time of the visit and will rate this overall status using a VAS scale where 0 is 'very good, asymptomatic, and no limitation of normal activities' and 100 is 'very poor, very severe symptoms which are intolerable, and inability to carry out all normal activities'.

4.12.3.2.6 Change from baseline in Health Activities Questionnaire – Disability Index score at Week 12

The HAQ-DI is comprised of eight domains: dressing, arising, eating, walking, hygiene, reach, grip, and activities. Each domain contains two or three questions scorings from 0 (without any difficulty) to 3 (unable to do). The worst score within each domain will be used as the domain score (i.e., if the score for one question is 1 and 2 for the other, then the worst score for the domain is 2). When an aide or device is used or help is required from another individual, the minimum score for that relevant domain is 2.

The HAQ-DI score is calculated by dividing the sum of the domain scores by the number of non-missing domains. If all questions within a given domain are unanswered, no score will be provided for that domain (this rule applies even if aids and devices are non-missing). If more than 2 domains are missing, score will be considered as missing. The HAQ-DI score ranges from 0 to 3.0. A higher score indicates worse function and greater disability.

Given the limitation of the Medidata eCOA system, HAQ-DI score will be re-derived based on source data for analysis purpose. In the event the re-derived score and Medidata eCOA system calculated score are discrepant, the differences will be documented in the CSR narratives.

4.12.3.2.7 Change from baseline in Dactylitis count at Week 12, among subjects who have dactylitis at Baseline

Dactylitis is defined as tenderness observed in any of the individual digit in the hands and feet. In the CRF, each individual digit is assessed for tenderness and four values are assigned: 0="No tenderness", 1="Tender", 2="Tender and wince", and 3="Tender and withdraw". Digit with values 1, 2, and 3 will be considered as having tenderness. Dactylitis is present if the total number of digits with tenderness >=1, or absent if the total count is zero.

The dactylitis score counts the total number of single digits in the hands and feet with tenderness. A higher score indicates worse symptom of tenderness in the hands and feet. If a total score is ≥1, a subject is defined to have dactylitis. If a total score is 0, a subject is defined as not having dactylitis.

Given the limitation of the Medidata eCOA system, dactylitis score will be re-derived based on the source data for analysis purpose. In the event, the re-derived dactylitis score and Medidata eCOA system calculated score are discrepant, the differences will be documented in the CSR narratives.

Analysis on dactylitis will be performed for all subjects in the FAS who have dactylitis at baseline.

4.12.3.2.8 Change from baseline in Leed's Enthesitis Index (LEI) at Week 12, among subjects who have enthesitis at Baseline

The LEI is comprised of review of six bilateral sites: Achilles tendon insertions (left and right), medial femoral condyles (left and right), and lateral epicondyles of the humerus (left and right). Tenderness at each site is quantified on a dichotomous basis: 0 means nontender and 1 means tender.

When assessments of 4 sites are missing, the LEI total score will be set to missing. If assessments of at least 3 sites are available, the total LEI score will be computed as a weighted score:

$$LEI = 6 \times \frac{\sum_{i=1}^{n} score}{\sum_{i=1}^{n} AI}$$

n refers to the number of total sites assessed and AI refers to the site with non-missing score.

The LEI total score ranges from 0 to 6, and a higher score indicates worse symptom of tenderness in the six bilateral sites. If LEI total score ≥ 1, a subject is defined to have enthesitis.

Analysis on LEI will be performed for all subjects in the FAS with baseline LEI score ≥1.

4.12.3.2.9 Proportion of subjects with Minimal Disease Activity (MDA) at Week 12

MDA is measured as whether a subject meet ≥ 5 of the 7 criteria (yes/no). The 7 criteria include:

1. TJC ≤1
2. SJC ≤1
3. PASI score ≤1 or BSA ≤3%
4. PGAAP, VAS score ≤15
5. PGA-PsA, VAS score ≤20
6. HAQ-DI ≤0.5
7. Tender enthesial points (LEI ≤1)

If any subject meets ≥5 of the 7 criteria, MDA is considered as "Yes", whereas if any subject fails ≥3 of the 7 criteria, MDA is considered as "No". MDA will be set to "Missing" for any other scenarios.

4.12.3.2.10 Change from baseline in Disease Activity Index for Psoriatic Arthritis (DAPSA) at Week 12

The DAPSA score is a composite score calculated using TJC68, SJC66, PGA-PsA, PGAAP, and hsCRP level.

The formula to calculate DAPSA composite score is as follows:

DAPSA = SJC66 + TJC68 + PGA-PsA/10 + PGAAP/10 + hsCRP

Where hs-CRP is expressed in mg/dL.

If any individual component score is missing, the DAPSA score will be set to missing.

DAPSA values will also be categorized in the following disease activity categories:

- DAPSA ≤4 for Remission
- DAPSA from >4 to ≤14 for Low Disease Activity
- DAPSA from >14 to ≤28 for Medium Disease Activity
- DAPSA >28 for High Disease Activity

The count and percentage of DAPSA in each category will be summarized by visit, and treatment group.

4.12.3.2.11 Proportion of subjects achieving 75% improvement from baseline in psoriasis area severity index [(PASI)-75] among subjects with ≥ 3% BSA psoriatic involvement at baseline

The PASI is a composite score ranging from 0 to 72 and takes into account the degree of erythema, induration/infiltration, and desquamation (each scored from 0 to 4 separately) for each of four body regions, with adjustments for the percentage of BSA involved within each body region and for the proportion of the body region to the whole body.

PASI is a tool to measure the severity and extent of psoriasis involvement. To calculate PASI, the human body is divided into four regions: head and neck, upper extremities, trunk, and lower extremities, which account for 10%, 20%, 30%, and 40% of the total BSA, respectively. Within each region, the severity and the extent of the disease are evaluated separately. The severity is assessed using the intensity of erythema (redness), induration (thickness), and desquamation (scaling of the psoriasis). The intensity is rated on a scale of 0 to 4, with 0=no symptoms, 1=slight, 2=moderate, 3=marked, and 4=very marked. The extent of the disease is measured by the percentage of the area affected by psoriasis within each region, and is expressed as 0-6, with 0 = 0%, 1 = 1%-9%, 2 = 10%-29%, 3 = 30%-49%, 4 = 50%-69%, 5 = 70%-89%, and 6 = 90-100%.

The following formula is used to calculate the PASI score:

$$PASI = 10\% \times (I_e + I_i + I_d) \times A_h + 20\% \times (I_e + I_i + I_d) \times A_u + 30\% \times (I_e + I_i + I_d) \times A_t + 40\% \times (I_e + I_i + I_d) \times A_l$$

Where I_e = intensity of erythema, I_i = intensity of induration, I_d = intensity of desquamation, A_h = percentage of psoriasis involvement in head and neck area, A_u = percentage of psoriasis involvement in upper extremities, A_t = percentage of psoriasis involvement in trunk area, A_l = percentage of psoriasis involvement in lower extremities.

4.12.3.2.12 Proportion of subjects achieving a Physician Global Assessment of Psoriasis of 0 or 1 and at least a 2-point improvement from baseline at Week 12

The PhGA-PsO is measured using a 0 to 4 scale with a 0 meaning clear or a 4 meaning severe.

PhGA-PsO responder is defined as subjects 1) having PhGA-PsO score of 0 or 1 at any given visit; and 2) having at least 2-point improvement from baseline. The proportion of subjects achieving PhGA-PsO response at Week 12 will be calculated and analyzed for subjects with a score of at least 2 at baseline.

4.13 Safety Evaluation

4.13.1 Extent of Exposure

Duration of study drug exposure, cumulative dose (mg), and average daily dose (mg/day) will be summarized on SS by treatment and overall.

Duration of study drug exposure (days) will be calculated as: Date of Last Dose – Date of First Dose + 1

Duration of study drug exposure including safety follow-up used for exposure adjusted incidence rate (EAIR) calculation will be calculated as: Date of EOS – Date of First Dose + 1

The cumulative dose taken by each subject during the treatment period will be determined using data collected in Study Drug Accountability eCRF and is defined as: Cumulative dose (mg) = [Total Amount Dispensed [redacted] - Total Amount Returned [redacted]] × dose level (mg) / 2

Note: calculation of cumulative dose is only based on blister packets with complete data, i.e. both the number of [redacted] dispensed and returned for the same packet should not be missing.

Average daily dose will be calculated as below: Average daily dose= cumulative dose/Duration of study drug exposure.

In addition, the number and percentage of subjects will be summarized by the following study drug exposure duration (days) categories:

- 1 day to ≤4 weeks (1 to ≤28 days)
- 4 weeks to ≤8 weeks (>28 to ≤56 days)
- 8 weeks to ≤12 weeks (>56 to ≤84 days)
- >12 weeks (>84 days)

A by-patient listing of study drug exposure (including the start and end dates of study drug, duration of study drug exposure, cumulative dose, and average daily dose) will be provided on the SS. In addition, data collected from investigational product (IP) Administration-on site, IP Administration-at Home, and Patient Diary will be listed as well.

4.13.2 Adverse Events

Adverse events will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) Version 26.0 or higher. Intensity and causality of AE will be evaluated by the investigator. Intensity of AE will be graded according to the Common Terminology Criteria for Adverse Events (CTCAE), version 5.0. Partial start and end date of the AE will be imputed as described in Section 4.3.2.

Subjects with multiple incidences of events for a given PT and SOC will be counted only once. Similarly, if a subject experienced multiple incidents of events for a given PT and SOC, the worst intensity will be used in the summaries presenting intensity, and the worst causality to study drug will be used in the summaries presenting causality, respectively.

AE summaries will be sorted in decreasing frequency by SOC, and PT within SOC, in the overall SS, and then alphabetically for SOC, and PT within SOC if there are any ties in frequency. Analysis and reporting of AEs will be based on TEAEs. A TEAE is defined as an AE occurring (onset

date/time) at the time of or after dosing on Day 1. An overall summary of TEAEs will include the number and percentage of subjects with:

- TEAEs
- Drug-related TEAEs
- TEAEs leading to discontinuation of study drug
- TEAEs leading to study discontinuation
- TESAEs
- Drug-related TESAEs
- TESAEs leading to discontinuation of study drug
- TESAEs leading to study discontinuation
- TEAEs by intensity (CTCAE grades 1, 2, 3, 4, 5)
- TEAEs leading to death
- TEAESIs
- Drug-related TEAESIs
- TEAESIs leading to discontinuation of study drug
- TEAESIs leading to study discontinuation

Summaries of the following categories of TEAEs will be presented by treatment group, SOC, and PT:

- All TEAEs
- Drug-related TEAEs
- TEAEs leading to discontinuation of study drug
- TEAEs leading to study discontinuation
- TEAEs Grade 3 or higher
- TESAEs
- Drug-related TESAEs
- TESAEs leading to discontinuation of study drug
- TESAEs leading to study discontinuation
- TEAEs by intensity (CTCAE grades 1, 2, 3, 4, 5)
- TEAEs leading to death

In order to assess the impact of AEs on geriatric subjects, tables for all TEAEs and Drug-related TEAEs will also be presented by age group (<65 years old and ≥65 years old), with associated listing for TEAEs for subjects ≥65 years old. Other age groups could be evaluated based on age distribution.

Summaries of the following events will be presented by treatment group and overall:

- Potential Drug-induced Liver Injury (DILI) and Hy's Law Summary
- TEAEs by PT (descending order)
- Drug-related TEAEs by PT (descending order)
- Most frequent (≥5%) TEAEs by PT (descending order)
- Most frequent (≥5%) drug-related TEAEs by PT (descending order)
- Exposure Adjusted Incidence Rate (EAIRs) for TEAEs by SOC and PT

Number and percentage of subjects who meet one or more of the following potential DILI criteria at any visit (including unscheduled visit) will be summarized by treatment group and overall:

1. Hy's Law: ALT or AST of >3 times ULN (or ≥3 times the baseline value if greater than the ULN at baseline) AND total Bilirubin >2 times ULN (or >2 times the baseline value if greater than ULN at baseline, OR international normalized ratio >1.5) AND ALP <2 times the ULN.

   In order to meet the above criteria, a subject must experience the elevation in bilirubin and ALT or AST at the same visit. For example, a subject who experiences a ≥2 × ULN elevation of bilirubin at one visit and a ≥3 × ULN elevation in ALT (or AST) at a subsequent visit has not fulfilled the potential Hy's law criteria.

2. ALT or AST ≥8 times the ULN, or ≥8 times the baseline value, if greater than the ULN at baseline, regardless of other parameters

   Note: ALT or AST ≥8 times the baseline value alone does not qualify the event to be classified as DILI. The baseline ALT/AST must be >ULN at the same time.

Baseline ALT or AST	Postbaseline ALT or AST	ALT or AST ≥8 × ULN, or ≥8 × baseline value, if greater than the ULN at baseline, regardless of other parameters
Normal (≤1 ULN)	≥8 × ULN	Yes
Normal (≤1 ULN)	<8 × Baseline	No
Abnormal (>1 ULN)	≥8 × Baseline	Yes
Abnormal (>1 ULN)	≥8 × ULN but <8× Baseline	No
Abnormal (>1 ULN)	<8 × ULN	No

3. Either ALT or AST ≥3 times the ULN, or ≥3 times the baseline value, if greater than the ULN at baseline and the appearance of fatigue, nausea, vomiting, right upper quadrant pain or tenderness, fever, rash, and/or eosinophilia (>5%)

   Note: to be classified as DILI, both ALT/AST elevation and symptoms must be present at the same time. ALT/AST ≥3 times the baseline value alone does not qualify the event to be ALT/AST elevation. The baseline ALT/AST must be > ULN at the same time.

Baseline ALT or AST	Post-baseline ALT or AST	Selected AE Present	ALT or AST ≥3 times the ULN, or ≥ 3 times the baseline value, if greater than the ULN at baseline, and with the appearance of fatigue, nausea, vomiting, right upper quadrant pain or tenderness, fever, rash, and/or eosinophilia (> 5%)
Normal (≤1 ULN)	≥3 × ULN	Yes	Yes
Normal (≤1 ULN)	≥3 × ULN	No	No
Normal (≤1 ULN)	<3 × ULN	-	No
Abnormal (>1 ULN)	≥3× Baseline	Yes	Yes
Abnormal (>1 ULN)	≥3 × Baseline	No	No
Abnormal (>1 ULN)	≥3 × ULN but <3 × Baseline	-	No
Abnormal (>1 ULN)	<3 × ULN	-	No

4. Total bilirubin or direct bilirubin >3 times the ULN, or >3 times the baseline value, if greater than the ULN at baseline.

   Note: Total Bili or direction Bili >3 times the baseline value alone does not qualify the event to be classified as DILI. The baseline total BILI/direct BILI must be > ULN at the same time. In addition, postbaseline total bilirubin can only be compared with baseline total bilirubin and postbaseline direct bilirubin can only be compared with baseline direct bilirubin.

Baseline Total Bilirubin or Direct Bilirubin	Post-baseline Total Bilirubin or Direct Bilirubin	Total bilirubin or direct bilirubin >3 times the ULN, or >3 times the baseline value
Normal (≤1 ULN)	>3 × ULN	Yes
Normal (≤1 ULN)	≤3 × ULN	No
Abnormal (>1 ULN)	>3 × Baseline	Yes
Abnormal (>1 ULN)	>3 × ULN but ≤3 × Baseline	No
Abnormal (>1 ULN)	≤3 × ULN	No

Exposure-adjusted incident rate (EAIR) is calculated as:

$$EAIR = \frac{\text{The Number of Adverse Events}}{\text{Total Exposure Time}} \times 100\%$$

Where the total exposure time (person-year) sums the duration of study drug exposure (including safety follow up) from all subjects included in the SS. The total of the duration of study drug exposure in days will be converted into years by dividing it with 365.25. The duration of study drug exposure including safety follow-up = Date of EOS – Date of First Dose + 1.

By-subject listings of corresponding TEAE summary will also be provided including additional information such as action taken with study drug, other action taken, outcome, and pattern, etc.

4.13.3 Adverse Events of Special Interest

The following are AESIs from laboratory data:

- CTCAE Grade ≥2 cytopenia (anemia, leukopenia, neutropenia, lymphopenia, or thrombocytopenia; clinically significant or not) defined by the following ranges:
  - Hemoglobin <10.0 g/dL (<100.0 g/L).
  - White blood cell count $<3.0 \times 10^9$/L ($<3000/mm^3$).
  - Absolute neutrophil count of $<1.5 \times 10^9$/L ($<1500/mm^3$).
  - Absolute lymphocyte count of $<0.8 \times 10^9$/L ($<800/mm^3$).
  - Platelet count $<75 \times 10^9$/L ($<75{,}000/mm^3$).
- CTCAE Grade ≥3 elevation of CPK (clinically significant or not) defined as CPK >5 × ULN

The following AEs are considered AESIs from adverse events data. These AESIs will initially be identified using search criteria below and will be adjudicated by Sponsor prior to the database lock.

- Major adverse cardiovascular events (MACE), defined as cardiovascular death, non-fatal myocardial infarction, and non-fatal stroke
  MACE will be identified using the following rules:
  - Standardized MedDRA queries (SMQs): "myocardial infarction", "Haemorrhagic central nervous system vascular conditions", or "ischaemic central nervous system vascular condition"
  - Death
- Thromboembolic events, defined as pulmonary embolism, deep vein thrombosis, and other venous and arterial thromboembolic events (e.g., non-cardiac, non-neurologic, fatal and nonfatal)
  Thromboembolic events will be identified using the following 3 broad SMQs: "Embolic and thrombotic events, arterial", "Embolic and thrombotic events, venous", "Embolic and thrombotic events, vessel type unspecified and mixed arterial and venous".
- Gastrointestinal perforation
  Gastrointestinal perforation will be identified using the PT "gastrointestinal perforation".
- Malignancies
  - All malignancies
  - Non-melanoma skin cancer (NMSC)
  - Malignancy excluding NMSC
  - Lymphoma

  Malignancies will be identified using SOC "Neoplasms benign, malignant and unspecified (incl. cysts and polyps).
- Infections
  - All infections
  - Serious infection
  - Opportunistic infection, excluding tuberculosis and herpes zoster
  - Herpes zoster
  - Active tuberculosis

Infections will be identified using the SOC "Infections and Infestations".

- AEs of abnormal liver function tests
  AEs of abnormal liver function tests will be identified using the high level group term "Hepatobiliary investigations".
- AEs of renal dysfunction
  AEs of renal dysfunction will be identified using the SOC "renal and urinary disorders".

Summaries of the TEAESIs will be presented by treatment group, SOC, and PT:

- TEAESIs [overall and by specific categories defined above]
- Drug-related TEAESIs
- TEAESIs leading to discontinuation of study drug
- TEAESIs leading to study discontinuation

By-subject listings of corresponding TEAESI summaries will also be provided including additional information such as action taken with study drug, other action taken, outcome, and pattern, etc.

4.13.4 Electrocardiogram

12-lead ECG interpretation will be summarized by treatment group.

A by-subject listing of 12-lead ECG will also be provided.

4.13.5 Vital Signs

Descriptive summaries of actual values and changes from baseline at each visit will be summarized for temperature, systolic blood pressure, diastolic blood pressure, heart rate, and respiratory rate by treatment and overall. Clinically significant vital signs measurements meeting the criteria specified in Appendix 4 will also be summarized by treatment group.

A by-subject listing of vital sign parameters will also be provided.

4.13.6 Clinical Laboratory Evaluation

The following laboratory tests will be summarized by treatment and overall using standard international units:

Table 4-2: Clinical Laboratory Parameters

Laboratory Evaluation	Tests Included
Hematology	aPTT, HCT, Hgb, INR, MCH, MCHC, MCV, MPV, PLT, PT, RBC, WBC, and differentials (neutrophils, lymphocytes, monocytes, eosinophils, and basophils relative and absolute)
Biochemistry	Albumin, alkaline phosphatase, ALT, AST, calcium, CPK, creatinine (enzymatic), GGT, glucose random, potassium, sodium, total bilirubin, direct bilirubin if total bilirubin > ULN, urea (BUN), uric acid
Fasting lipids	Cholesterol (total, LDL, and HDL), triglycerides (all fasting)
Urinalysis	Color, clarity, pH, specific gravity, bilirubin, glucose, ketones, leukocytes, nitrite, blood, protein, urobilinogen Microscopic analysis (as required)
Urine pregnancy test (conducted at the investigator site)	For WOCBP (at each visit, except screening)
Specialty evaluations	hsCRP, ESR, Rheumatoid factor, anti-CCP
Laboratory tests required at screening only	FSH levels for female subjects who have had a cessation of menses for at least 12 months prior to the screening visit without an alternative medical cause β-hCG for WOCBP Tuberculosis test (QuantiFERON-TB Gold) • Note: A T-Spot.TB test (TBT) may be used as an alternative to the QuantiFERON-TB Gold test as per Exclusion Criterion 37. If the TBT is used, the QuantiFERON-TB Gold test should not be performed. Serology (HBV [HBsAg, anti-HBc], HCV[a], HIV)

ALT = alanine aminotransferase; anti-HBc = antibody to hepatitis B core antigen = aPTT = activated partial thromboplastin time; AST = aspartate aminotransferase; β-hCG = β-human chorionic gonadotropin; BUN = blood urea nitrogen; CPK = creatine phosphokinase; FSH = follicle-stimulating hormone; GGT = gamma-glutamyltransferase; HBsAg = hepatitis B surface antigens; HBV = hepatitis B virus; HCT = hematocrit; HCV = hepatitis C virus; HDL = high-density lipoprotein; Hgb = hemoglobin; HIV = human immunodeficiency virus; INR = international normalized ratio; LDH = lactate dehydrogenase; MCH = mean corpuscular hemoglobin; MCHC = mean corpuscular hemoglobin concentration; MCV = mean corpuscular volume; MPV = mean platelet volume; PCR = polymerase chain reaction, PLT = platelets; PT = prothrombin time; RBC = red blood cell (count); ULN = upper limit of normal; WBC = white blood cell (count); WOCBP = women of childbearing potential

[a]Samples testing positive for HCV antibodies will require PCR qualitative testing for HCV RNA. Any HCV RNA PCR result that meets or exceeds detection sensitivity is exclusionary.

Hematology, chemistry, and urine protein will be summarized using the SS. Continuous test results and changes from baseline will be summarized at each visit by treatment and overall using descriptive statistics. Categorical test results will be summarized at each visit by treatment and

overall using frequencies and percentages. If a subject has both low and high out-of-normal post-baseline values, each low and high will be counted separately.

All laboratory summaries will be based on observed values. For the purpose of summarizing the data, laboratory values below the lower limit of quantification (LLOQ) will be set to the midpoint between 0 and the LLOQ. If laboratory values are expressed as above a fixed value (ie, ">xx.xx"), then the numeric portion of the value will be used (ie, xx.xx).

The following summaries will be produced by treatment and overall on the SS:

- A summary of the absolute values and changes from Baseline values in each laboratory variable (Hematology, Chemistry, Urine) by visit
- Shift from baseline to each post-baseline visit for each laboratory parameter (Hematology, Chemistry, Urine Protein)
- Shift from baseline to highest/lowest post-baseline range for select laboratory parameters (Hematology, Chemistry, Urinalysis). Highest post-baseline will be considered for the following laboratory variables: hematology (aPTT, INR, hematocrit, hemoglobin, platelets, white blood cell count, neutrophil count, lymphocyte count), chemistry (ALP, AST, ALT, bilirubin, CPK, creatinine, sodium, potassium, total cholesterol, LDL, triglycerides), and urinalysis (urine protein, leukocytes, erythrocytes). Lowest post-baseline will be considered for the following laboratory parameters: hematology (hematocrit, hemoglobin, platelets, white blood cell count, neutrophil count, lymphocyte count) and chemistry (albumin, sodium, potassium)
- A summary of the number and percentage of subjects experiencing treatment-emergent markedly abnormal values by laboratory variable (Hematology, Chemistry and Urine protein). Markedly abnormal values are defined as those with a severity of Grade 3 and above based on the CTCAE criteria. Definitions of markedly abnormal values are given in Appendix 3.
- A summary of the number and percentage of subjects by CTCAE grade (version 5.0) based on minimum/maximum post-Baseline value by laboratory variable (Hematology, Chemistry, Urine protein)
- A summary of the number and percentage of subjects with CTCAE grade 2 or higher and at least 1 grade increase from baseline for selected lab parameters .

The selected lab abnormalities (neutropenia, lymphopenia, anemia, thrombocytopenia, white blood cell decreased, activated partial thromboplastin time prolonged, CPK elevation, ALT elevation, AST elevation, hypoalbuminemia, cholesterol elevation, triglyceride elevation, hypoglycemia, creatinine elevation, presence of urine protein, INR elevation, bilirubin elevation) and their associated CTCAE grade are described in Appendix 2.

By-subject listings of laboratory parameters (Hematology, Chemistry, Urinalysis) will be provided by study time point including changes from baseline. The abnormal values will be flagged with 'L' for values below the lower limit of the clinical reference range and 'H' for values above the upper limit of the clinical reference range and included in the listings.

A line plot for mean change from baseline of hematology labs, fasting lipids, hsCRP, and ESR will be presented by treatment group for the SS. A line plot for mean values over time will also be presented by treatment group for the following parameters: for lymphocytes, neutrophils, platelets, hemoglobin, total cholesterol, triglycerides, ALT, AST, and CPK.

The results of pregnancy, serology, and TB will be listed for each subject.

4.13.7 Physical Examination

A complete physical examination will be conducted at Screening and Week 16 (End of Study) visit, and includes:

- General appearance
- Dermatological (except plaque PSO)
- Head, eyes, ears, nose, and throat (HEENT)
- Respiratory
- Cardiovascular
- Abdominal
- Neurological
- Musculoskeletal
- Lymphatic

A targeted physical examination will be performed at Weeks 1, 2, 4, 8, 12:

- General appearance
- Dermatological (except plaque PSO)
- Respiratory
- Cardiovascular
- Abdominal

Abnormal findings of physical examinations will be summarized at each visit using frequencies and percentages by treatment and overall. A by-subject listing of all physical examination abnormalities including investigator determined clinical significance will also be provided.

4.13.8 Data and Safety Monitoring Board

No Data Safety Monitoring Board is planned.

4.13.9 Other Analysis Consideration

Depending on the prevalence of coronavirus disease (COVID) infections and illness in regions where the study is conducted, additional analysis may be performed to evaluate the impact of COVID on the safety of all participating subjects, including but not limited to the following:

- COVID related discontinuation
- COVID related AEs, including preferred term that contains “COVID-19”, “SARS-CoV-2”, “SARS-CoV-2” and “Coronavirus”.
- All SAEs in COVID infected subjects.
- All protocol deviations related to COVID.
- Data listing of all subjects affected by COVID-19 related study disruption, including subject ID, site ID, and description of how individual’s participation was altered.
- Data listing of all subjects taking COVID-19 vaccine during the course of study, if data permits.

4.14 Pharmacokinetics

Descriptive statistics of the NDI-034858 plasma concentration data will be summarized based on nominal timepoints per dose level and will be presented in a table using the PKS. NDI-034858 plasma concentration data, including actual sampling time, will be listed per subject by visit and time points using the PKS.

When reporting individual data in listings the following rules will apply:

- Missing data should be reported as NV (no value).
- Concentrations below the limit of quantification (BLQ) should be reported as BLQ.
- Concentrations should be listed to the same number of significant figures supplied by the bioanalytical laboratory.

When summarizing the data in tables the following rules will apply:

- To calculate descriptive statistics, BLQ values should be set to zero and missing values should be excluded.
- A minimum of 3 values are required to calculate summary statistics. If only 2 values are available, then these should be presented as the minimum and maximum with other descriptive statistics reported as missing ("-").
- If no subjects have data, only n=0 will be presented. The other descriptive statistics will be left blank.
- Descriptive statistics will be reported to the same level of precision as the individual data for the minimum and maximum, and to 1 additional decimal place (with a maximum of 3 significant digits) than raw data recorded in the database for the mean (arithmetic and geometric) and median. SD will be reported to 2 additional decimal places (with a maximum of 4 significant digits) than raw data. Geometric CV will be reported as a percentage to 1 decimal place.

4.15 Exploratory Analyses

[redacted]

4.15.1 Analysis Approach for Exploratory Endpoints

[redacted]

4.15.2 Derivation of Exploratory Endpoints

4.15.2.1 [redacted]

[redacted]

4.15.2.2 [redacted]

4.15.2.3 [redacted]

4.15.2.4 [redacted]

[redacted]

[redacted]

4.15.2.5 [redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

4.15.2.6 [redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

4.15.2.7 [redacted]

4.15.2.8 [redacted]

4.15.2.9 [redacted]

4.16 Determination of Sample Size

A sample size of 65 per treatment group across all sites will have 83% power using a two-sided test for the difference of two proportions, assuming a type I error rate of 0.05, a 55% response rate for each NDI-034858 dose group as a proportion of ACR20 response, and a placebo ACR20 response rate of 30%. A total number of 260 subjects (65 per treatment group) are planned to be randomized into the study with a 1:1:1:1 allocation.

The sample size was calculated in nQuery 8.7 using a two-sample Z-Test (Chi-Square Test) Pooled for difference of proportions.

4.17 Changes in the Conduct of the Study or Planned Analysis

Changes to the planned analysis in study protocol, Version 4.0 (May 05, 2023) included the following:

- The description text of PPS in Section 4.5 is updated from that of the protocol, specifically, the defining clause "as defined in Section 6.3.2 of the study protocol" is removed. The criteria of treatment compliance for PPS will be based on treatment compliance as defined in Section 4.11. The intention of the clause "Significant noncompliance is defined as missing 3 consecutive doses (ie, 3 days in a row) or missing a total of 10 doses" in Section 6.3.2 of the study protocol was to counsel patients and sites to maximize adherence on study.
- The adjustment for multiplicity was updated from protocol Section 9.1.1 by adding the fixed-sequence testing procedure (See Section 4.12.1.4).
- The last secondary endpoint listed under protocol Section 3.4.2 "Proportion of subjects achieving a Physician Global Assessment of Psoriasis of 0 or 1 and at least a 2-point improvement from baseline (Day 1) at Week 12" is clarified in Section 4.12.3.2.12 as The proportion of subjects achieving PhGA-PsO response at Week 12 will be calculated and analyzed for subjects with a score of at least 2 at baseline.

[redacted]

[redacted]

5.0 REFERENCES

1. Addendum on Estimands and Sensitivity Analysis in Clinical Trials to the Guideline on Statistical Principles for Clinical Trials E9(R1) Final version (Step 4), Adopted on 20 November 2019. (https://database.ich.org/sites/default/files/E9-R1_Step4_Guideline_2019_1203.pdf)

2. Eek D, Ivanescu C, Corredoira L, Meyers O, Cella D. Content validity and psychometric evaluation of the Functional Assessment of Chronic Illness Therapy-Fatigue scale in patients with chronic lymphocytic leukemia. J Patient Rep Outcomes. 2021, 5: 27.

[redacted]

4. Fransen J, Antoni C, Mease PJ, Uter W, Kavanaugh A, Kalden JR, et al. Performance of response criteria for assessing peripheral arthritis in patients with psoriatic arthritis: analysis of data from randomised controlled trials of two tumour necrosis factor inhibitors. Annals of the rheumatic diseases. 2006, 65 (10): 1373-8.

5. Helliwell PS, FitzGerald O, Fransen J, Gladman DD, Kreuger GG, Callis-Duffin K, et al. The development of candidate composite disease activity and responder indices for psoriatic arthritis (GRACE project). Annals of the rheumatic diseases. 2013, 72 (6): 986-91.

6. Kenward MG, Roger JH. Small sample inference for fixed effects from restricted maximum likelihood. Biometrics. 1997; 53(3): 983-97.

[redacted]

8. Mantel N, Haenszel W. Statistical Aspects of Analysis of Data from Retrospective Studies of Disease. Journal of the National Cancer Institute. 1959, 22: 719-48.

9. Perruccio AV, Got M, Li S, Ye Y, Gladman DD, Chandran V. Treating Psoriatic Arthritis to Target: Defining the Psoriatic Arthritis Disease Activity Score That Reflects a State of Minimal Disease Activity. The Journal of rheumatology. 2020, 47 (3): 362-8.

10. Ratitch B, Goel N, Mallinckrodt C, Bell J, Bartlett JW, Molenberghs G, et al. Defining Efficacy Estimands in Clinical Trials: Examples Illustrating ICH E9(R1) Guidelines. Therapeutic innovation & regulatory science. 2020, 54 (2): 370-84.

11. Sato, T. On the Variance Estimator of the Mantel-Haenszel Risk Difference. Biometrics. 1989, 45: 1323-24. Letter to the editor.

12. SAS® Version 9.2 of the SAS System for Personal Computers. Copyright ©2002-2003. SAS Institute Inc. SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc., Cary, NC, USA.

13. Chen J, Hunter S, Kisfalvi K, Lirio RA. A hybrid approach of handling missing data under different missing data mechanisms: VISIBLE 1 and VARSITY trials for ulcerative colitis. Contemporary Clinical Trials. 2021, 100:106226. https://doi.org/10.1016/j.cct.2020.106226.

6.0 APPENDICES

6.1 Appendix 1: Example for Calculating Treatment Compliance

For example, patient [redacted] in Table 4-2 completed week 4 visit. Even though eight blister packets were dispensed, only four had return record available. The total amount of study drug administered is (18 - 4) × 4 = 56. The total amount of study drug expected is (Date of last dose – Date of Day 1 + 1) × 2 = ([redacted] - [redacted] + 1) = 56. The compliance is 56/56 × 100% = 100%.

Patient [redacted] completed week 12 visit. A total of 12 packets were dispensed and return amount was missing for two packets. Therefore the total amount of study drug administered is (18 - 4) × 10 = 140. The total amount of study drug expected is (Date of last dose – Date of Day 1 + 1) × 2 = ([redacted] - [redacted] + 1) × 2 = 85 × 2 = 170. The compliance is 140/170 × 100% = 82.4%.

Table 4-2

Subject ID	Visit	Package ID Dispensed	# Dispensed	Package ID Returned	# Returned	Date of Visit	Date of Last Dose
[redacted]	Day 1	[redacted]	18		NA	[redacted]	NA
	Week 1		18	[redacted]	4		[redacted]
	Week 2		18		4		
	Week 2		18		4		
	Week 4		18		4		
	Week 4		18				
	Week 4		18				
	Week 4		18				
[redacted]	Day 1	[redacted]	18		NA	[redacted]	NA
	Week 1		18	[redacted]	4		[redacted]
	Week 2		18		4		
	Week 2		18		4		
	Week 4		18		4		
	Week 4		18		.		
	Week 4		18		4		

	Week 4	[redacted]	18	[redacted]	4	[redacted]	[redacted]
	Week 8		18		4		
	Week 8		18		4		
	Week 8		18		4		
	Week 8		18		.		
	Week 12		NA		4		

6.2 Appendix 2: CTCAE Criteria for Clinical Laboratory Parameters

CTCAE v.5 Term	Grade 1	Grade 2	Grade 3	Grade 4
Anemia	Hemoglobin (Hgb) <LLN - 10.0 g/dL; <LLN - 6.2 mmol/L; <LLN - 100 g/L	Hgb <10.0 - 8.0 g/dL; <6.2 - 4.9 mmol/L; <100 - 80 g/L	Hgb <8.0 g/dL; <4.9 mmol/L; <80 g/L; transfusion indicated	Life-threatening consequences; urgent intervention indicated
Neutrophil count decreased	<LLN - 1500/mm^3; <LLN - 1.5 × 10e9/L	<1500 - 1000/mm^3; <1.5 - 1.0 × 10e9/L	<1000 - 500/mm^3; <1.0 - 0.5 × 10e9/L	<500/mm^3; <0.5 × 10e9/L
Lymphocyte count decreased	<LLN - 800/mm^3; <LLN - 0.8 × 10e9/L	<800 - 500/mm^3; <0.8 - 0.5 × 10e9/L	<500 - 200/mm^3; <0.5 - 0.2 × 10e9/L	<200/mm^3; <0.2 × 10e9/L
Platelet count decreased (thrombocytopenia)	<LLN - 75,000/mm^3; <LLN - 75.0 × 10e9/L	<75,000 - 50,000/mm^3; <75.0 - 50.0 × 10e9/L	<50,000 - 25,000/mm^3; <50.0 - 25.0 × 10e9/L	<25,000/mm^3; <25.0 × 10e9/L
White blood cell decreased	<LLN - 3000/mm^3; <LLN - 3.0 × 10e9 /L	<3000 - 2000/mm^3; <3.0 - 2.0 × 10e9 /L	<2000 - 1000/mm^3; <2.0 - 1.0× 10e9 /L	<1000/mm^3; <1.0 × 10e9 /L
INR increased	>1.2 - 1.5; >1 - 1.5 × baseline if on anticoagulation; monitoring only indicated	>1.5 - 2.5; >1.5 - 2.5 baseline if on anticoagulation; dose adjustment indicated	>2.5; >2.5 × baseline if on anticoagulation; bleeding	(-)

CTCAE v.5 Term	Grade 1	Grade 2	Grade 3	Grade 4
Activated partial thromboplastin time prolonged	>ULN - 1.5 × ULN	>1.5 - 2.5 × ULN	>2.5 × ULN; bleeding	
ALT increased (Alanine aminotransferase increased)	>ULN - 3.0 × ULN if baseline was normal; 1.5 - 3.0 × baseline if baseline was abnormal	>3.0 - 5.0 × ULN if baseline was normal; >3.0 - 5.0 × baseline if baseline was abnormal	>5.0 - 20.0 × ULN if baseline was normal; >5.0 - 20.0 × baseline if baseline was abnormal	>20.0 × ULN if baseline was normal; >20.0 × baseline if baseline was abnormal
AST increased (Aspartate aminotransferase increased)	>ULN - 3.0 × ULN if baseline was normal; 1.5 - 3.0 × baseline if baseline was abnormal	>3.0 - 5.0 × ULN if baseline was normal; >3.0 - 5.0 × baseline if baseline was abnormal	>5.0 - 20.0 × ULN if baseline was normal; >5.0 - 20.0 × baseline if baseline was abnormal	>20.0 × ULN if baseline was normal; >20.0 × baseline if baseline was abnormal
Bilirubin (Blood bilirubin increased)	>ULN - 1.5 × ULN if baseline was normal; >1.0 - 1.5 × baseline if baseline was abnormal	>1.5 - 3.0 × ULN if baseline was normal; >1.5 - 3.0 × baseline if baseline was abnormal	>3.0 - 10.0 × ULN if baseline was normal; >3.0 - 10.0 × baseline if baseline was abnormal	>10.0 × ULN if baseline was normal; >10.0 × baseline if baseline was abnormal
Hypoalbuminemia	<LLN - 3 g/dL; <LLN - 30 g/L	<3 - 2 g/dL; <30 - 20 g/L	<2 g/dL; <20 g/L	Life-threatening consequences; urgent intervention indicated
CPK increased	>ULN - 2.5 × ULN	>2.5 × ULN - 5 × ULN	>5 × ULN - 10 × ULN	>10 × ULN

CTCAE v.5 Term	Grade 1	Grade 2	Grade 3	Grade 4
Cholesterol high	>ULN - 300 mg/dL; >ULN - 7.75 mmol/L	>300 - 400 mg/dL; >7.75 - 10.34 mmol/L	>400 - 500 mg/dL; >10.34 - 12.92 mmol/L	>500 mg/dL; >12.92 mmol/L
Hypertriglyceridemia	150 mg/dL - 300 mg/dL; 1.71 mmol/L - 3.42 mmol/L	>300 mg/dL - 500 mg/dL; >3.42 mmol/L - 5.7 mmol/L	>500 mg/dL - 1000 mg/dL; >5.7 mmol/L - 11.4 mmol/L	>1000 mg/dL; >11.4 mmol/L; life-threatening consequences
Hypoglycemia	<LLN - 55 mg/dL; <LLN - 3.0 mmol/L	<55 - 40 mg/dL; <3.0 - 2.2 mmol/L	<40 - 30 mg/dL; <2.2 - 1.7 mmol/L	<30 mg/dL; <1.7 mmol/L; life-threatening consequences; seizures
Creatinine increased	>ULN - 1.5 × ULN	>1.5 - 3.0 × baseline; >1.5 - 3.0 × ULN	>3.0 × baseline; >3.0 - 6.0 × ULN	>6.0 × ULN
Proteinuria	1+ proteinuria; urinary protein ≥ULN - <1.0 g/24 hrs	Adult: 2+ and 3+ proteinuria; urinary protein 1.0 - <3.5 g/24 hrs	Adult: Urinary protein >3.5 g/24 hrs 4+ proteinuria	(-)

6.3 Appendix 3: Criteria for markedly abnormal laboratory parameters

Parameter	Unit (conventional)	Unit (Standard)	Marked Abnormality Criteria
Hematology			
aPTT	Sec	sec	>2.5 × ULN
Hematocrit	L/L	L/L	Low: <0.8 × LLN High: >1.2 × ULN
Hemoglobin	g/dL	g/L	Low: < 8.0 g/dL; <80g/L High: >40 above ULN (g/L)
INR			2.5 × ULN; >2.5 times above baseline if on anticoagulation
Prothrombin time			Low: <0.8 × LLN High: >1.2 × ULN
Lymphocytes Absolute	10^9/L	10^9/L	Low: <0.5 × 10^9/L High: >20 × 10^9/L
Neutrophils Absolute	10^9/L	10^9/L	<1.0 × 10^9/L
Platelets	10^9/L	10^9/L	<50.0 × 10^9/L
RBC/ Erythrocytes	10^{12}/L	10^{12}/L	<3.5 × 10^{12}/L
WBC (Leukocytes)	10^9/L	10^9/L	Low: <2.0 × 10^9/L High : >100 × 10^9/L
Other			
Biochemistry			
AST	U/L	U/L	>5.0 × ULN if Baseline value is normal; > 5.0 × Baseline value if Baseline is abnormal
ALT	U/L	U/L	>5.0 × ULN if Baseline value is normal; > 5.0 × Baseline value if Baseline is abnormal
ALP	U/L	U/L	>5.0 × ULN if Baseline value is normal; > 5.0 × Baseline value if Baseline is abnormal
Albumin	g/dL	g/L	<2g/dL; <20g/L
Bilirubin (total)	mg/dL	umol/L	>3.0 × ULN if Baseline value is normal > 3.0 × Baseline value if Baseline is abnormal

Parameter	Unit (conventional)	Unit (Standard)	Marked Abnormality Criteria
Calcium	mg/dL	mmol/L	Low: Corrected serum calcium of <7.0 mg/dL; <1.75 mmol/L High: Corrected serum calcium of >12.5 mg/dL; >3.1 mmol/L
Total cholesterol	mg/dl	mmol/L	>10.34 mmol/L;
Creatinine	mg/dL	umol/L	>3.0 × ULN or >3.0 × Baseline
CPK	U/L	U/L	>5 × ULN
GGT	U/L	U/L	>5.0 × ULN if Baseline value is normal; > 5.0 × Baseline value if Baseline is abnormal
Glucose[b]	mg/dL	mmol/L	Low: <40mg/dL; < 2.2 mmol/L High: >250 mg/dL; >13.9 mmol/L
Potassium	mEql/L	mmol/L	Low: <3.0 mmol/L High: > 6.0 mmol/L
Sodium	mEql/L	mmol/L	Low: <125 mmol/L High: >155 mmol/L
Urea nitrogen	mg/dl	mmol/L	>30 mg/dL ; >10.7 mmol/L
Uric acid	mg/dl	mmol/L	>ULN
Urinalysis			
Urine protein			4+ proteinuria

ALT = alanine aminotransferase; aPTT = activated partial thromboplastin time; AST = aspartate aminotransferase; β-hCG = β-human chorionic gonadotropin; BUN = blood urea nitrogen; CPK = creatine phosphokinase; hsCRP = high sensitivity C-reactive Protein; ESR = erythrocyte sedimentation rate; GGT = gamma-glutamyltransferase; HDL = high-density lipoprotein; INR = international normalized ratio; LDH = lactate dehydrogenase; LLN= lower limit of normal ; MCH = mean corpuscular hemoglobin; MCHC = mean corpuscular hemoglobin concentration; MCV = mean corpuscular volume; MPV = mean platelet volume; PLT = platelets; RBC = red blood cell (count); ULN = upper limit of normal; WBC = white blood cell (count).

[a] Includes CRP and High Sensitivity (HS) CRP. Reference for marked abnormality criteria: Nehring, S.M.; Goyal, A.; Patel, B.C. (2020). StatPearls Publishing, web link: https://www.ncbi.nlm.nih.gov/books/NBK441843

[b] Glucose high criterion defined by Grade 3 and higher events according to CTCAE, Version 4.03, June 14, 2010.

6.4 Appendix 4: Abnormality Criteria for Vital Signs

Parameter	Abnormality Criteria
Pulse rate (beats/min)	Low: ≤50 and a decrease from Baseline ≥15 High: ≥120 and an increase from Baseline ≥15
Systolic Blood Pressure (mmHg)	Low: ≤90 and a decrease from Baseline ≥20 High: ≥160 and an increase from Baseline ≥20
Diastolic Blood Pressure (mmHg)	Low: ≤50 and a decrease from Baseline ≥15 High: ≥105 and an increase from Baseline ≥15
Temperature	>101°F (38.3°C)
Body Weight	Low: ≥10% decrease from Baseline High: ≥10% increase from Baseline

6.5 Appendix 5: Statistical Methodology

Sample SAS Code for Mantel-Haenszel stratum weights and the Sato variance estimator:

```
proc freq data = data;
        tables strata*trt01p*resp / commonriskdiff (test=mh cl=mh column=2);
run;
```

where strata variable has 4 levels (from the two stratification factors: prior treatment with biologics or non-traditional DMARDs [yes/no], and REGION [USA/Germany, Eastern Europe]), trt01p is planned treatment group, and resp is a responder outcome variable.

Sample SAS Code for MMRM:

```
proc mixed data = data;
        class trt01p avisit usubjid dmards region;
        model CHG = trt01p dmards region base avisit trt01p*avisit/ddfm=kr;
        repeated avisit / subject= usubjid type=UN;
        lsmeans trt01p*avisit / diff=all cl alpha=0.05;
run;
```


Sample SAS Code for Multiple Imputation of binary endpoint under MAR:

1) Generation of imputed data using proc mi

```
proc mi data = data out = outmi seed=25381 nimpute=100;
        by trt01p; *trt01p includes all 4 treatment groups including placebo;
        class dmards region;
        var dmards region resp_wk1 resp_wk2 resp_wk4 resp_wk8 resp_wk12;
        fcs nbiter=100
        fcs reg(resp_wk1= dmards region); *trt01p includes all 4 treatment groups including placebo;
        fcs reg(resp_wk2= dmards region resp_wk1);
        fcs reg(resp_wk4= dmards region resp_wk1 resp_wk2);
        fcs reg(resp_wk8= dmards region resp_wk1 resp_wk2 resp_wk4);
        fcs reg(resp_wk12=dmards region resp_wk1 resp_wk2 resp_wk4 resp_wk8 / details);
run;
```

2) Conduct analysis for each treatment group vs placebo using proc freq using each imputed dataset

```
proc freq data = outmi;
by _imputation_;
tables strata*trt01p*resp / commonriskdiff (test=mh column=2);
ods output CommonPdiff= commonpdiff;
run;
```

3) Pooling the results from the imputed datasets using proc mianalyze:

```
proc mianalyze data= commonpdiff;
modeleffects riskdifference;
stderr;
run;
```

6.6 Appendix 6: Corticosteroid Dose Equivalents

Equivalent Dose	Steroid
1.2 mg	Betamethasone (long-acting)
1.5 mg	Dexamethasone (long-acting)
8 mg	Methylprednisolone (intermediate-acting)
8 mg	Triamcinolone (intermediate-acting)
10 mg	Prednisone (intermediate-acting)
10 mg	Prednisolone (intermediate-acting)
40 mg	Hydrocortisone (short-acting)
50 mg	Cortisone (short-acting)

https://emedicine.medscape.com/article/2172042-overview

6.7 Appendix 7: Algorithm for prohibited medications used for intercurrent events

Prohibited medications will be considered in 2 separate ICEs in the primary analysis. The method for identifying these medications will consider the following criteria and may consider clinical review.

Immunosuppressants (ATC3): any new concomitant mediation OR any increase in dose from prior medications (programmatic + clinical review)

This category will be coded on a case by case basis given the lack on specific ATC.

- Biologics: any new initiation of these would be treatment failure from the time of initiation.
- Non-traditional DMARDs: any new initiation of these would be treatment failure from the time of initiation.
- Traditional MTX and SSZ: dose increase from (baseline) prior medication dose level would classify for treatment failure.
- Other traditional DMARDs (including hydroxychloroquine): dose increase from (baseline) prior medication dose level would classify for treatment failure.

Corticosteroids for systemic use, plain (ATC3):

Treatment failure is considered:

- If ROA oral: increase from stable baseline dose if prior medication (10 mg or less of prednisone equivalent) or add-on of new CS start within 14 days of Week 12 efficacy assessment (clinical review)
- If ROA IV/IM: treatment failure from the time of initiation
- If intraarticular corticosteroids: considered in the derivation rule for TJC/SJC (Section 4.12.3.2.3)

Opioids (ATC3): Treatment failure is considered if new continuous treatment started within 7 days of Week 12 Assessment (programmatic and clinical review)

Other analgesics and antipyretics (ATC3): Treatment failure is considered if any increase in dose compared to baseline OR add-on of daily continuous use for 7 days if not present at baseline within 7 days of Week 12 assessment OR one new dose given within 24 hours of Week 12 assessment (programmatic and clinical review)

Anti-inflammatory and antirheumatic products, non-steroids (ATC3): Treatment failure is considered if any increase in dose compared to baseline (if prior) or add-on of daily continuous administration for 7 consecutive days within 7 days of Week 12 assessment (programmatic and clinical review)

Topical products for joint and muscular pain (ATC3): Treatment failure is considered if daily continuous administration for 7 days if not present at baseline within 7 days of Week 12 assessment (programmatic and clinical review)