Janssen Research & Development

Statistical Analysis Plan

A Phase 2a, Multicenter, Randomized, Double-blind Study Evaluating the Efficacy and Safety of Subcutaneously Administered Guselkumab and Golimumab Combination Therapy in Patients with Active Psoriatic Arthritis

Protocol CNTO1959PSA2003; Phase 2a

CNTO1959 (Guselkumab)
CNTO148 (Golimumab)

Status: Approved
Date: 10 May 2024
Prepared by: Janssen Research & Development, LLC
Document No.: EDMS-RIM-617366, 2.0

Compliance: The study described in this report was performed according to the principles of Good Clinical Practice (GCP).

TABLE OF CONTENTS

VERSION HISTORY

Table 1 - SAP Version History Summary

SAP Version	Approval Date	Change and Rationale
0.1		Not Applicable
0.2		Add EDMS number
0.3		MRI related sections with STAR KoL opinion considered
1.0	26/10/2022	
2.0		1) Updated study inclusion criteria from "1" to "1 or 2" prior anti-TNF treatment, due to protocol amendment 2) Removed ICE4 from all efficacy Estimand definition, i.e., discontinued study intervention due to study conduct affected by COVID-19. The reason for this update is because there are no such cases in study (one-month prior DBL) 3) MRI analysis plan is updated by removing Estimand and sensitivity analysis, with the plan to just perform complete case analysis. 4) Included additional endpoints in Section 5.5.1 per clinical request

1. INTRODUCTION

This statistical analysis plan (SAP) contains definitions of analysis sets, derived variables, and statistical methods for the analysis of efficacy, safety, pharmacokinetics (PK), and pharmacodynamics (PD), and immunogenicity in the CNTO1959PSA2003 study.

1.1. Objectives

Primary Objective

To evaluate the efficacy of guselkumab + golimumab combination treatment in participants with active psoriatic arthritis (PsA) and inadequate response (IR) to 1 or 2 prior anti-TNFα treatment(s) by assessing clinical response compared with guselkumab monotherapy.

Major Secondary Objective

To evaluate the efficacy of guselkumab + golimumab combination treatment in participants with active PsA and IR to 1 or 2 prior anti-TNFα treatment(s) by assessing the reduction in signs and symptoms of PsA compared with guselkumab monotherapy.

Other Secondary Objectives

- Safety: to evaluate the safety of guselkumab + golimumab combination treatment in participants with active PsA and IR to 1 or 2 prior anti-TNFα treatment(s) compared with guselkumab monotherapy
- Pharmacokinetic (PK) and Immunogenicity: to evaluate the PK and immunogenicity of guselkumab + golimumab combination treatment in participants with active PsA and IR to 1 or 2 prior anti-TNFα treatment(s) compared with guselkumab monotherapy.

1.2. Study Design

This is a Phase 2a, randomized, double-blind, active-controlled, parallel-group, multicenter, interventional, proof of concept clinical study designed to evaluate the efficacy and safety of combination therapy with guselkumab and golimumab (guselkumab CCI CCI and golimumab CCI CCI versus guselkumab monotherapy (guselkumab CCI CCI and placebo) in adults with active PsA. The target population is men or women 18 to 65 years old, with active PsA and a previous history of inadequate response (IR) to 1 or 2 anti-TNFα treatment(s), either due to primary non-response or loss of efficacy. participants must not have received prior treatment with golimumab or guselkumab.

There will be three phases in this study: screening, double-blind, and safety follow-up.

- The screening phase of up to 6 weeks will allow for sufficient time to perform screening study evaluation and determine study eligibility.
- The second phase of the study will be from week 0 to week 24, which includes the active treatment phase and the final efficacy visit. The primary endpoint will be evaluated at the final efficacy visit at week 24. This duration will provide adequate time to evaluate the efficacy and short-term safety of combination therapy in PsA
- The final phase of the study will be the safety follow-up phase and will be 16 weeks from the last administration of study intervention to the final safety visit. The safety follow-up

period allows for monitoring of participants for a period equivalent to approximately 5 times the half-life of guselkumab and golimumab.

Approximately 90 participants who have met the study inclusion and exclusion criteria are to be randomized by using dynamic central randomization stratified by below five baseline measurements:

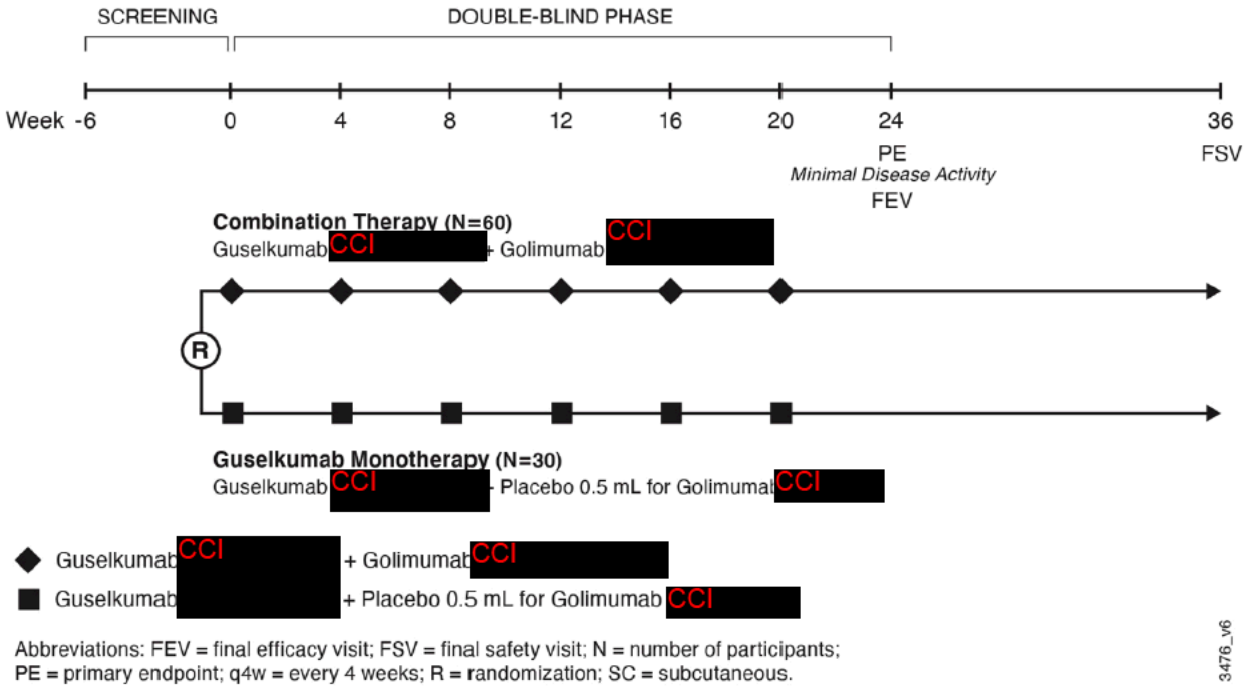
- Non-biologic DMARDs (methotrexate [MTX], sulfasalazine [SSZ], hydroxychloroquine [HCQ], leflunomide [LEF]) use (yes/no)
- Skin biopsy sub-study enrollment status (enrolled/not enrolled)
- Tender joint count (<12, ≥12 and <18, ≥18 and <27, ≥27)
- PASI (<3, ≥3 and <8.8, ≥8.8 and <20.7, ≥20.7)
- Patient's pain score: <50, ≥50 and <67, ≥67 and <79, ≥79

The randomization will be in a blinded fashion in a 2:1 ratio of 1 of the 2 treatment groups:

- Group I (n=60): Combination therapy: guselkumab CCI CCI and golimumab CCI CCI CCI at weeks CCI
- Group II (n=30): Monotherapy: guselkumab CCI CCI CCI and placebo CCI CCI to maintain the blind throughout the duration of study at weeks CCI

An overview of this clinical development program is presented in Figure 1 below:

Figure 1: Schematic Overview of Study CNTO1959PSA2003



Through the study, stable dose of concomitant NSAIDs, oral corticosteroid, and selected non-biologic DMARDs (limited to MTX, SSZ, HCQ, LEF, see Table 2) will be allowed. participants should not initiate any new treatment and/or any investigational treatment for PsA through Week 24. Participants may not be receiving more than one non-biologic DMARD from baseline through

Week 24. At baseline, participants are permitted to enter the study on a stable dose of one non-biologic DMARD up to the maximum allowed dose specified in Table 2. Participants may not be receiving MTX and oral corticosteroids same time within 2 weeks prior to the first administration of study intervention.

Table 2 - Permitted Commutant Medication for PsA and the Maximum Allowed Doses During the Study

Permitted Concomitant Medications for PsA [a]	Maximum Allowed Dose
NSAIDs and other analgesics	Marketed dose approved in the country where study is being conducted
Oral corticosteroids	Equivalent to 10 mg/day of prednisone
Methotrexate (MTX)[b]	25 mg/week
Sulfasalazine (SSZ)	3 g/day
Hydroxychloroquine (HCQ)	400 mg/day
Leflunomide (LEF)	20 mg/day

a. Permitted concomitant medications are not supplied by the sponsor.
b. It is recommended that all participants taking MTX in this study receive at least 5 mg oral folate or 5 mg folinic acid weekly. Guidelines for dose adjustment in the event of MTX toxicity are included in the Trial Center File.

Participants will be followed for adverse events (AE) and serious adverse events (SAE) up to week 36, i.e., 16 weeks following the last study intervention administration.

An interim analysis is planned to inform future clinical development when 60 participants (two third of the target population) reach Week 24. Participants, regardless of their study intervention discontinuation and/or termination status will be counted towards the 60.

Database locks (DBLs) are scheduled at Week 24 and End of Study (Week 36). The first DBL will occur when all randomized participants have either completed the Week 24 assessments or terminated study participation prior to the Week 24 visit (referred to as Week 24 DBL). The second DBL will occur when all randomized participants have either completed their final safety visit (Week 36) or have terminated study participation prior Week 36 (referred to as Final DBL).

The end of the study is defined as the last visit for the last participant. The last visit is the last safety follow-up visit for participants who complete the week 20 study intervention and the week 24 final efficacy assessment, and for participants who discontinue study intervention prior to week 20. For participants who terminate the study early, the last visit is the last attended visit. The study will remain blinded for the duration of the trial, until after the first DBL (Week 24).

An independent, external Data Monitoring Committee (DMC) will monitor the safety of the study in an unblinded fashion on a regular basis and whenever deemed necessary. Additional details related to DMC are provided in Section 5.9.

2. STATISTICAL HYPOTHESES

The primary endpoint of this study is the proportion of participants who achieve a Minimal Disease Activity (MDA) response at Week 24 (Section 5.3.1).

The hypothesis related to the primary endpoint is that: treatment with combination therapy (guselkumab CCI CC CCI and golimumab CCI CC CCI is superior to treatment with

monotherapy (guselkumab CCI CC CCI with respect to reduction of PsA signs and symptoms as measured by proportion of participants who achieve an MDA response at week 24. If the primary hypothesis achieves a statistical significance at a 1-sided α-level of 0.1, the study will be considered positive.

3. SAMPLE SIZE DETERMINATION

The sample size was determined based on the primary endpoint of proportion of participants who achieve an MDA response at Week 24. The response rate of guselkumab CCI CCI monotherapy treatment is estimated to be 20%, which is based on indirect estimation of different dose/population from 3 completed global Phase 3 studies: CNTO1959PSA3001(hereafter referred to as PSA3001), CNTO1959PSA3002 (hereafter referred to as PSA3002), and CNTO1959PSA3003 (hereafter referred to as PSA3003). With an assumed response rate of 45% in the combination therapy and 20% in the monotherapy, sample size of 90 would achieve above 80% power by using the one-sided Fisher's Exact test at significance level 0.1.

Participants will be randomized in a 2:1 ratio with 60 in combination therapy and 30 in monotherapy. Table 3 shows the power to detect a difference in the proportion of participants achieving MDA response between combination regimen and monotherapy regimen with various assumptions.

Table 3 - Statistical Power for Treatment Difference in MDA Response at Week 24

MDA Response Rate			
Guselkumab CCI CC	Guselkumab CCI CC + Golimumab CCI CC	Difference (Δ)	Power
20%	35%	15%	48%
20%	40%	20%	67%
20%	**45%**	**25%**	**82%**
25%	35%	10%	29%
25%	40%	15%	47%
25%	45%	20%	65%
25%	50%	25%	80%

4. POPULATIONS (ANALYSIS SETS) FOR ANALYSIS

Full Analysis Set

All participants who were randomized in the study. Participants will be analyzed according to the intervention they were randomized into, i.e., regardless of the intervention that participants received.

Safety Analysis Set

All participants who were randomized in the study and received at least one (complete or partial) administration of study intervention. Participants will be analyzed according to the intervention they received, i.e., regardless of the intervention that participants randomized to.

Pharmacokinetics (PK) Analysis Set

All participants who were randomized in the study and received at least one (complete) administration of study intervention and had at least one valid post-injection blood sample drawn

for PK analysis. Participants will be analyzed according to the intervention they received, i.e., regardless of the intervention that participants randomized to.

Immunogenicity Analysis Set
All participants who were randomized in the study and received at least one (complete) administration of study intervention and had appropriate samples for antibodies to guselkumab and golimumab detection. Participants will be analyzed according to the intervention they received, i.e., regardless of the intervention that participants randomized to.

Pharmacodynamic (PD) Analysis Set
All participants who were randomized in the study and received at least one (complete or partial) administration of study intervention and had at least one valid post-injection blood sample drawn for PD analysis. Participants will be analyzed according to the intervention they received, i.e., regardless of the intervention that participants randomized to.

5. STATISTICAL ANALYSES

5.1. General Considerations

5.1.1. Visit Windows

All post-baseline visits through Week 24 will have a visit window of ± 4 days. The final safety visit/Week 36 will have a visit window of 7 days (+ 7 days). Unless otherwise specified, actual scheduled visits will be used for the summaries and listings over time with no visit windows applied.

For PK analyses, if a participant has an administration outside the visit window at a visit, the concentration data collected at and after that visit will be excluded from the by-visit data analyses.

For MRI, the window for taking MRI assessments at respective scheduled visit are specified in the protocol. An analytical window of -4 weeks inclusive will be used in MRI data analyses at Week 0. An analytical window of +/-4 weeks inclusive of target Week 24 date will be used in MRI data analyses at Week 24. MRI data taken at end of study intervention will be slotted based on the analysis window to the appropriate visits to be included in the data analysis.

5.1.2. Pooling Algorithm

Data from all investigational centers/sites will be pooled for analyses.

5.2. Participant Dispositions

Screened participants and reason for screen failures will be summarized overall.

The number of participants who met Intercurrent Event (ICE) will be summarized through tabulation by intervention group, in the following categories:

- Initiated protocol-prohibited medications and therapies for PsA.
- Initiation or increased the dose of non-biologic DMARDs (MTX, SSZ, HCQ, LEF) or oral corticosteroids over baseline for PsA.

- Discontinued study intervention due to any reason

The number of participants who discontinued study intervention administration early will be summarized through tabulation by intervention group and overall, in the following categories:

- Number of participants randomized and received study intervention.
- Number of participants discontinued study intervention early and the primary reasons for the discontinuation of treatment.

The number of participants who terminated from study prematurely will be summarized through tabulation by intervention group and overall, in the following categories:

- Number of participants randomized and received study intervention.
- Number of participants ongoing in the study
- Number of participants who terminated from study prematurely and the primary reasons for terminating from study.

Per-protocol follow-up is defined as:

- Completing study scheduled visits if participants did not discontinue study intervention prematurely: Week 0, 4, 8, 12, 16, 20, 24 and 36.
- In the case that participant discontinue study intervention prematurely, participant completed Week 24 efficacy visit and the Final safety visit, which is 16 weeks from last injection.

Listings of participants will be provided for the following categories:

- Participants who discontinued study intervention early (i.e., prior Week 24)
- Participants who terminated study prematurely (prior Week 24 and, prior Week 36)

5.3. Primary Endpoint Analysis

The primary endpoint of this study is proportion of participants who achieved an MDA response at week 24. This section outlines the definitions and analyses of this primary endpoint.

5.3.1. Definition of Endpoint

MDA response is a composite measurement of PsA signs and symptoms, it is a 'state' of disease activity rather than a numerical measurement (Gossec, et al., 2018). MDA response is defined when at least 5 of the following 7 assessment criteria are met (Coates, et al., 2010):

- Tender Joint Count (68 joints) [**TJC68**] is equal or less than 1.
- Swollen Joint Count (68 joints) [**SJC66**] is equal or less than 1.
- Psoriasis Area and Severity Index [**PASI**] is equal or less than 1 OR Body Surface Area [**BSA**] involvement is equal or less than 3%.
- Patient's Global Assessment of Pain [**PAIN**] is equal or less than 15 on a 100 mm unit Visual Analog Scale (VAS)
- Patient's Global Assessment of Disease Activity (arthritis and psoriasis, VAS) [**PGADAPA**] is equal or less than 20 on a 100 mm unit VAS.
- Disability Index of the Health Assessment Questionnaire (**HAQ-DI**) is equal or less than 0.5.
- Leeds Enthesitis Index (**LEI**) is equal or less than 1.

Following is the definition of each of the forgoing disease assessment criteria (components) that are used in the determination of MDA response:

TJC68: a total number of tender joints among the 68 joints evaluated for tenderness. Each of the 68 joints will be evaluated for tenderness, categorized as tender or not tender. Joint evaluability rules specified in Appendix 2 for overwriting joint evaluation will be applied to those joints with joint injection(s)/surgical joint procedure(s). For patients with any joint not evaluable in the 68-joint set, joint count adjustment rules described in Appendix 2 will be applied in determining the ultimate count of tender joints.

SJC66: a total number of swollen joints among the 66 joints evaluated for swelling (note: the 2 hip joints are excluded from swelling assessment). Each of the 66 joints will be evaluated for swelling, categorized as swollen or not swollen. Joint evaluability rules specified in Appendix 2 for overwriting joint evaluation will be applied to those joints with joint injection(s)/surgical joint procedure(s). For participants with any joint not evaluable in the 66-joint set, joint count adjustment rules described in Appendix 2 will be applied in determining the ultimate count of swollen joints.

PASI: The PASI is a system used for assessing and grading the severity of psoriatic lesions and their response to therapy. The PASI produces a numeric score that can range from 0 to 72 with a higher score indicating a more severe disease. In the PASI system, the body is divided into 4 regions: head and neck, trunk (including axillae and groin), upper extremities, and lower extremities (including buttocks), which account for 10%, 30%, 20%, and 40% of the total BSA, respectively. Each of these 4 regions is assessed separately for erythema, induration, and scaling on a scale of 0-4. If any of the components required for computing the PASI score is missing, the PASI score will be set as missing.

BSA: The BSA is an assessment measures the total area of body affected by psoriasis. The size of the BSA with psoriatic lesions is estimated based on the number of pals of the participant, each palm (from the basis of wrist to the proximal interphalangeal joints including thumb) is equivalent to 1% BSA. The score can range between 0% and 100%. A score of less than 5% is considered as mild, a score between 5% and 10% is considered as moderate, a score between 10% and 15% is considered as severe, and last, a score that greater than 15% is considered as very severe. In the PASI data collection, percentage area affected was measured by each of the four regions, while BSA is a one measurement of the total area of body affected by psoriasis. Approximately, 1 palm of psoriasis equal to 10%, 5%, 3.3%, 2.5% area of involvement at head, upper limbs, trunk, and lower limbs area respectively. Therefore, the percentage area affected collected in PASI should be strongly linear correlated with BSA.

PAIN: VAS instrument is used to collect PAIN and is defined on a sliding scale from 0 to 100mm, where the far-left position (0mm) is representing no pain, and the far-right position (100mm) is representing the worst possible pain.

PGADAPA: VAS instrument is used to collect PGADAPA and is defined on a sliding scale from 0 to 100mm, where the far-left position (0mm) means participant feel very well, and the far-right position (100mm) means participant feel very poor.

HAQ-DI: The HAQ-DI is an evaluation of the functional status for a participant. The 20-question instrument assesses the degree of difficulty a person has in accomplishing tasks in 8 functional areas (dressing, arising, eating, walking, hygiene, reaching, gripping, and activities of daily living). Response in each functional area is scored from 0 (no difficulty) to 3 (inability to perform task). The HAQ-DI score is the sum of scores divided by the number of categories answered. The HAQ-DI score will not be computed if more than 2 of the 8 categories are not scored.

Enthesitis (LEI): Enthesitis is an important feature of psoriatic arthritis and other spondyloarthropathies. In this study, Enthesitis will be assessed by an independent joint assessor using the Leeds Enthesitis Index (LEI) (Healy & Helliwell, 2008). The LEI was developed to assess Enthesitis in participants with PsA and evaluates the presence (score of 1) or absence (score of 0) of pain by applying local pressure to three entheses on both left and right: Lateral epicondyle humerus, medial condyle femur, and Achilles tendon insertion. The Enthesitis index score is a total score of the 6 evaluated sites as listed. For participants who have an incomplete set of 6 evaluated sites, the Leeds Enthesitis Index score cannot be used.

5.3.2. Estimand

5.3.2.1. Primary Estimand (Composite Variable)

Primary Trial Objective

To assess the superiority of combination therapy (guselkumab CCI CC CCI and golimumab CCI CC CCI versus monotherapy (guselkumab CCI CC CCI on reducing disease activity in participants with active PsA and previous history or inadequate response (IR) to one or two prior anti-TNFα treatment(s) either due to primary non-response or loss of efficacy.

Estimand Scientific Question of Interest

For a participant with active psoriatic arthritis, what would be the expected effect of assigning combination therapy versus monotherapy on the likelihood of achieving minimal disease activity at Week 24, while administered together with the protocol allowed concomitant medication?

Estimand Attributes

The primary estimand, aligned with the above primary trial objective and question of interest, provide description of the treatment effect of interest that defined by the following 5 attributes:

Table 4 - Primary Estimand Attributes

Population	Participants who are ≥18 to ≤65 years of age (inclusive), with active PsA and previous history of IR to one or two prior anti-TNFα treatment(s) either due to primary non-response or loss of efficacy.
Variable	Binary response variable, where a responder is defined as a participant with an MDA response at Week 24 and not having had ICE 1-3.
Study intervention	- Combination therapy: guselkumab CCI C CC and golimumab CCI C CC - Monotherapy: guselkumab CCI C CC and placebo C CC

Population-level summary	Difference in proportion of responders (as defined in Variable) between study interventions	
ICE and the corresponding strategies	1. Initiated protocol-prohibited medications and therapies for PsA. 2. Initiation or increased the dose of non-biologic DMARDs (MTX, SSZ, HCQ, LEF) or oral corticosteroid over baseline for PsA. 3. Discontinued study intervention due to any reason	Composite Variable Strategy

This question aims to capture a treatment effect that can be translated into benefit at the participant level. Events that might be linked to unfavorable efficacy or safety are considered as unfavorable outcomes.

5.3.2.2. Supplementary Estimand (Treatment Policy)

Estimand Scientific Question of Interest

For a participant with active psoriatic arthritis, what would be the expected effect of assigning combination therapy versus monotherapy on the likelihood of achieving minimal disease activity at Week 24, while administered together with the protocol allowed concomitant medication, regardless of the initiation of protocol-prohibited medication/therapies for PsA, initiation or increased dose of concomitant medication, discontinuation of study intervention?

Table 5 - Supplementary Estimand Attributes

Population	Participants who are ≥18 to ≤65 years of age (inclusive), with active PsA and previous history of IR to one or two prior anti-TNFα treatment(s) either due to primary non-response or loss of efficacy.	
Variable	Binary response variable, where a responder is defined as a participant with an MDA response at Week 24	
Study intervention	- Combination therapy: guselkumab CCI C CC and golimumab CCI C CC - Monotherapy: guselkumab CCI C CC and placebo C CC	
Population-level summary	Difference in proportion of responders (as defined in Variable) between study interventions	
ICE and the corresponding strategies	1. Initiated protocol-prohibited medications and therapies for PsA. 2. Initiation or increased the dose of non-biologic DMARDs (MTX, SSZ, HCQ, LEF) or oral corticosteroid over baseline for PsA. 3. Discontinued study intervention due to any reason	Treatment Policy Strategy

5.3.3. Analysis Methods

The primary endpoint will be analyzed at Week 24 DBL based on the Full Analysis Set (Section 4).

Data that are not used will be different between the two estimand:

- For primary estimand, values of the outcome measure collected after ICE 1-3 are not used.
- For supplementary estimand, all available data will be used for analysis.

Missing data is assumed as missing at random (MAR), and conservatively impute the missing data at Week 24 as non-responder.

Descriptive statistics (i.e., counts and percentages) will be used to summarize the data. Graphical data displays (e.g., line plots) may also be used to summarize the data.

The treatment comparison between combination therapy and monotherapy will be tested using logistic regression on binary endpoint at Week 24. The explanatory variable might include treatment group, baseline non-biologic DMARDs use, TJC68, PASI, and PAIN, where the last

three variables will be treated as continuous variable. Odds ratio (OR), CI and p-value will be reported. Study will be considered as positive if the test is significant at 1-sided α level of 0.1.

Joint evaluability rules specified in Appendix 2 for overwriting joint evaluation will be applied to those joints with surgical procedure and/or joint injection.

5.3.4. Sensitivity Analysis 1

Here are the steps in performing the re-randomization test (EMA, 2015) (Han, et al., 2013)

1) Obtain the test statistics T^* for the superiority of Combination therapy against monotherapy in primary endpoint analysis (Section 5.3.2.1)
2) Estimate the distribution of the test statistics with repetitions by following below steps:
   a. Re-randomize the treatment assignment with the dynamic randomization by maintaining the same biased coin probability, but rather with a new seed to generate new serious of random number for treatment assignment. Based on the new treatment assignment, obtain the new test statistics T
   b. Repeat the re-randomize step for 10,000 times.
3) Derive the p-value associated with the observed test statistics T^* based on the empirical distribution in step 2. Here the p-value can be calculated as $(n + 1)/(10{,}000 + 1)$, where the n is the number of times that $T > T^*$ among the 10,000 repetitions.

5.3.5. Sensitivity Analysis 2

To evaluate the robustness of the assumption of all the missing data are non-responder, a sensitivity analysis with multiple imputation will be performed. More detail on MI method is specified in Appendix 3.

5.3.6. Subgroup Analyses

Subgroup analyses will be performed (if the number of participants within each subgroup level permit) using logistic regression model to evaluate treatment consistency in proportion of participants who achieve MDA response at Week 24 over the subgroups.

A forest plot based on logistic regression model will be produced for all subgroups. Odds ratio of the combination therapy group versus monotherapy group and the associated 90% CI from the logistic regression model will be provided for each of subgroups. The logistic regression model might include treatment group, baseline non-biologic DMARDs use, TJC68, PASI, and PAIN, where the last three variables will be treated as continuous variable.

To evaluate whether there is a statistically significant difference between subgroups, logistic regression model will be used, which will include treatment group, subgroup variable, an interaction term between treatment group and subgroup variable, baseline non-biologic DMARDs use, TJC68, PASI, and PAIN, where the last three variables will be treated as continuous variable.

For each subgroup analysis, the main estimator for the primary Estimand (composite variable) will be used.

For details on subgroup, it is defined in Section 5.7.5.

5.3.7. Summary of Analyses

Table 6 - Primary Endpoint Analysis Summary

Analysis (Estimand)	Missing data	Additional notes
Primary Estimand (Composite Variable) (Section 5.3.2.1) Supplementary Estimand (Treatment Policy) (Section 5.3.2.2) Sensitivity Analysis 1 for Primary Estimand (Composite Variable) (Section 5.3.4)	Missing data will be assumed as MAR and imputed conservatively as non-responder	- Summarized descriptively. - Odds ratio and 90% CI for treatment comparison - P-value from logistic regression
Sensitivity Analysis 2 for Primary Estimand (Composite Variable) (Section 5.3.5)	Missing data will be assumed as MAR and imputed through MI [a] with FCS regression of component scores	- Summarized descriptively. - Odds ratio and 90% CI for treatment comparison though all imputation datasets - P-value from logistic regression across multiply imputed data sets
Subgroup Analysis for Primary Estimand (Composite Variable) (Section 5.3.6)	Missing data will be assumed as MAR and imputed conservatively as non-responder	- Odds ratio and 90% CI for treatment comparison - P-value from logistic regression for the interaction of intervention group and subgroup variable - Graphical: forest plots
[a] MI method is summarized in Appendix 3		

5.4. Secondary Endpoints Analysis

5.4.1. Major Secondary Endpoints

Table 7 – Major Secondary Endpoints

1	Proportion of participants who achieve American College of Rheumatology (ACR) 50 at Week 24
2	Proportion of participants who achieve MDA at Week 16
3	Proportion of participants who achieve PASI 90/100 at Week 24 among the participants with ≥3% BSA psoriatic involvement and an Investigator's Global Assessment (IGA) score of ≥2 at baseline.
4	Proportion of participants with an IGA-psoriasis response (i.e., an IGA score of 0 or 1 AND ≥2 grade reduction from baseline) at Week 24 among the participants with ≥3% BSA psoriatic involvement and an IGA score of ≥2 at baseline
5	Change from baseline in HAQ-DI at Week 24
6	Resolution of Enthesitis (LEI) at Week 24 among the participants with Enthesitis (LEI) at baseline
7	Resolution of dactylitis at Week 24 among the participants with dactylitis at baseline
8	Change from baseline in Short Form Health Survey (SF-36) Physical Component Score (PCS) at Week 24

5.4.1.1. Definition of Endpoints

ACR 50

ACR response is a composite measurement of change in PsA signs and symptoms and is presented as the numerical measurement of improvement in multiple disease assessment criteria. An ACR50 response is defined as meeting below two criteria:

- ≥50% improvement from baseline in both TJC68 and SJC66
- ≥50% improvement from baseline in at least 3 of the following 5 assessments:
  - PAIN
  - Patient's Global Assessment of Disease Activity (arthritis, VAS) [PGADAA]
  - Physician's Global Assessment of Disease Activity (VAS) [PGAD]
  - HAQ-DI

- C-reactive protein (CRP)

Following is the definition of each of the forgoing disease assessment criteria (components) that are used in the determination of ACR response:

- TJC68: See Section 5.3.1
- SJC68: See Section 5.3.1
- PAIN: See Section 5.3.1
- PGADAA: a measure from 0 (very well) to 100(very poor) on a 100-unit VAS
- PGAD: a measure from 0 (no arthritis activity) to 100 (extremely active arthritis) on a 100-unit VAS
- HAQ-DI: See Section 5.3.1
- CRP: a lab parameter measured in mg/dL. LLOQ rule specified in Appendix 2 will be applied to values <LLOQ.

If a participants' baseline value for a component is zero (i.e., no disease activity as measured by the component), then the participant should be considered as not achieving 20% improvement from baseline for that component since there is no room for improvement.

MDA
For MDA definition, see Section 5.3.1

PASI
For PASI definition, see Section 5.3.1

BSA
For BSA definition, see Section 5.3.1

IGA
IGA documents the investigator's assessment of the participant's psoriasis at a given time point. Overall lesions are graded for induration, erythema, and scaling. The participant's psoriasis is assessed as cleared (0), minimal (1), mild (2), moderate (3), or severe (4)

An IGA-psoriasis response is defined as an IGA psoriasis score of 0 (cleared) or 1 (minimal) AND ≥2 grade reduction from baseline in the IGA psoriasis score. An IGA-psoriasis response is only applicable to participants who had an IGA psoriasis score of ≥2 (mild) at baseline.

HAQ-DI
For HAQ-DI definition, see Section 5.3.1

For change from baseline in HAQ-DI, it is a measure of the change in the functional status, where a negative change reflects an improvement, and a positive change reflects worsening. In PsA, a decrease in score of 0.35 has been determined to indicate a meaningful improvement (Mease, et al., 2016) (Genovese, et al., 2018).

Enthesitis (LEI)

For Enthesitis definition (based on LEI), see Section 5.3.1

For resolution of Enthesitis, it is defined as participants who had at least one tender enthesis at baseline and none at the at the analysis visit among the 6 sites included in the LEI measurement.

Dactylitis
Presence and severity of dactylitis will be assessed in both hands and feet using a scoring system from 0 to 3 (0: no dactylitis, 1: mild dactylitis, 2: moderate dactylitis, 3: severe dactylitis). The results for each digit are summed to produce a final dactylitis score with a range from 0 to 60.

Dactylitis count will be derived based on dactylitis score, each score will be recoded to 0 or 1 from 0 to 3, where any score >0 is recoded as 1.

For resolution of dactylitis, it is defined as participants who had a dactylitis score greater than 0 at baseline and a score of 0 at the analysis visit.

SF-36 PCS/MCS
The questionnaire of the 36-item short form health survey (SF-36) is a health-related quality of life instrument with 36 questions, developed as part of the Rand Health Insurance Experiment. Version 2 will be used. This instrument consists below 8 multi-item scales (domains):

- Limitations in physical functioning due to health problems
- Limitation in usual role activities due to physical health problems
- Bodily pain
- General mental health (psychological distress and well-being)
- Limitation in usual role activities due to personal or emotional problems
- Limitation in social functioning due to physical or mental health problems
- Vitality (energy and fatigue)
- General health perception

Each of these 8 scales (domains) is scored from 0 to 100 with higher scores indicating better health. Based on the scale scores, physical component summary (PCS) score and mental component summary (MCS) score will be derived. These summary scores are also scaled with higher scores indicating better health. A sample SF-36 assessment tool is provided in protocol Appendix 14.

For change from baseline in SF-36 PCS/MCS score, it is a measure of the change in health-related quality of life, where a positive change indicates an improvement, and a negative change indicates a worsening.

5.4.1.2. Estimand

5.4.1.2.1. Main Estimand (Composite Variable)

Table 8 – Major Secondary Endpoints Main Estimand Attributes

Study interventions: - Combination therapy: Guselkumab CCI C C and Golimumab CCI C C - Monotherapy: Guselkumab CCI C C and Placebo C C			
	Population	**Variable**	**Summary measure**

1	Same as the primary Estimand (Section 5.3.2.1 Table 4)	Binary response variable, where a responder is defined as a participant with an ACR 50 response at Week 24 and not having had ICE 1-3	Difference in proportion of responders (as defined in the Variable) between study intervention
2		Binary response variable, where a responder is defined as a participant with an MDA response at Week 16 and not having had ICE 1-3	
3	Same as the primary Estimand (Section 5.3.2.1 Table 4), but limit to subset of participants with ≥3% BSA psoriatic involvement and an IGA score of ≥2 at baseline	Binary response variable, where a responder is defined as a participant with an PASI 90 response at Week 24 and not having had ICE 1-3	
4		Binary response variable, where a responder is defined as a participant with an PASI 100 response at Week 24 and not having had ICE 1-3	
5		Binary response variable, where a responder is defined as a participant with an IGA-psoriasis response at Week 24 and not having had ICE 1-3	
6	Same as the primary Estimand (Section 5.3.2.1 Table 4), but limit to subset of participants with Enthesitis (LEI) at baseline	Binary response variable, where a responder is defined as a participant with Enthesitis (LEI) resolution at Week 24 and not having had ICE 1-3	
7	Same as the primary Estimand (Section 5.3.2.1 Table 4), but limit to subset of participants with dactylitis at baseline	Binary response variable, where a responder is defined as a participant with Dactylitis resolution at Week 24 and not having had ICE 1-3	
8	Same as the primary Estimand (Section 5.3.2.1 Table 4)	Continuous variable, which is defined as the change from baseline to Week 24 in HAQ-DI. In the case of an ICE 1-3, participant will be considered as having no change from baseline	Difference in average change (as defined in the Variable) between study intervention
9		Continuous variable, which is defined as the change from baseline to Week 24 in SF-36 PCS. In the case of an ICE 1-3, participant will be considered as having no change from baseline	
ICE and the corresponding strategies			
1. Initiated protocol-prohibited medications and therapies for PsA 2. Initiation or increased the dose of non-biologic DMARDs (MTX, SSZ, HCQ, LEF) or oral corticosteroid over baseline for PsA 3. Discontinued study intervention due to any reason			Composite Variable Strategy

5.4.1.2.2. Supplementary Estimand (Treatment Policy)

For all major secondary endpoints' supplementary estimand, it will share same population, study intervention, ICEs, variable and summary measures as the major secondary endpoints' main estimand (Section 5.4.1.2.1). The only difference is the ICE strategy, which will be Treatment Policy Strategy, i.e., a participant with this ICE will not be treated differently from participant without ICE. The occurrence of the ICE is irrelevant.

5.4.1.3. Analysis Methods

All major secondary endpoints will be analyzed at Week 24 DBL based on the FAS (Section 4).

Data that are not used will be different between the two estimand:
- For primary estimand, values of the outcome measure collected after ICE 1-3 are not used.
- For supplementary estimand, all available data will be used for analysis.

Binary endpoints

- For binary endpoints, same analysis method will be used as primary endpoint (Section 5.3.3).

Continuous endpoints

- Descriptive statistics (i.e., mean, median, SD, IQ range, minimum, and maximum) will be used to summarize continuous variables.
- Missing data is assumed as missing at random (MAR), and imputed through MI.
- The treatment comparisons between combination therapy and monotherapy will be tested using Analysis of Covariance (ANCOVA). The explanatory variable might include treatment group, endpoint baseline measurement, non-biologic DMARDs use, TJC68, PASI, and PAIN, where the last three variables will be treated as continuous variable. LS means, and LS mean difference, and associated CI will be reported.

5.4.1.4. Sensitivity Analysis

For binary endpoints, to evaluate the robustness of the assumption of all the missing data are non-responder, a sensitivity analysis with multiple imputation will be performed. More detail on MI method is specified in Appendix 3.

5.4.1.5. Summary of Analyses

Table 9 - Major Secondary Endpoints Analysis Summary

Binary endpoints			
	Main Estimand (Section 5.4.1.2.1)	Missing data will be assumed as MAR and imputed conservatively as non-responder	- Summarized descriptively. - Odds ratio, CI, and P-value from logistic regression
	Supplementary Estimand (Section 5.4.1.2.2)		
	Sensitivity Analysis for Main Estimand (Section 5.4.1.4)	Missing data will be assumed as MAR and imputed through MI [a]	- Summarized descriptively. - Pooled odds ratio, CI, and P-value from logistic regression across multiply imputed data sets
Continuous endpoints			
	Main Estimand (Section 5.4.1.2.1)	Missing data will be assumed as MAR and imputed through MI [a]	- Summarized descriptively. - Treatment comparisons including p-value from ANCOVA model.
	Supplementary Estimand (Section 5.4.1.2.2)		
[a] MI method is summarized in Appendix 3			

5.4.2. Other Secondary Endpoints

The other secondary endpoints are:

1) Frequency and type of Adverse Events (AEs), Serious AEs (SAEs), reasonably related AEs, AEs leading to discontinuation of study intervention, infections, and injection site reactions.
2) Serum concentrations of guselkumab and golimumab
3) Incidence of antibodies to guselkumab or golimumab

For continuous outcome, descriptive status include mean, standard deviation, median, quantiles, minimum and maximum will be prepared by study intervention group. For categorical outcome, event frequency counts and percentages will be prepared by study intervention group.

5.4.2.1. Adverse Events

To accomplish the first secondary endpoint, the Adverse Event analysis method are discussed below.

All Adverse Events (AE) will be analyzed at Week 36 DBL based on the Safety Analysis Set (Section 4), that is all randomized participants who have received at least one administration of study intervention will be included in analysis according to the intervention they actually receive.

The verbatim terms used in the CRF by investigators to identify adverse events will be coded using the Medical Dictionary for Regulatory Activities (MedDRA). Any AE occurring at or after the initial administration of study intervention is considered to be treatment emergent. If the event occurs on the day of the initial administration of study intervention, and either event time or time of administration are missing, then the event will be assumed to be treatment emergent. If the event date is recorded as partial or completely missing, then the event will be considered to be treatment emergent unless it is known to be prior to the first administration of study intervention based on partial onset date or resolution date. AEs that occur prior to the administration of study intervention but then increase in intensity after dosing are also considered to be treatment-emergent. All reported treatment-emergent adverse events will be included in the analysis. For each adverse event, the number and percentage of participants who experience at least 1 occurrence of the given event will be summarized by intervention group.

Treatment-emergent AEs will be tabulated by treatment group and MedDRA System Organ Class (SOC) and be sorted in descending frequency of the Preferred Term within the SOC. Summary tables will be provided for treatment-emergent adverse events:

- Treatment-emergent AE
- Treatment-emergent Serious AE
- Treatment-emergent AE leading to discontinuation of study intervention
- Treatment-emergent Infections (defined as any AE that was recorded as an infection by the investigator on the CRF)
- Treatment-emergent Serious Infections (defined as any SAE that was recorded as an infection by the investigator on the CRF)
- Injection-site reactions
- Suicidal Ideation or Behavior (defined as any AE that was recorded as an event of suicidal ideation or behavior by the investigator on the CRF)

Few selected summary tables will be also performed by subgroup of interest:

- Treatment-emergent AE
- Treatment-emergent Serious AE
- Treatment-emergent Infections
- Treatment-emergent Serious Infections

These summary tables will provide the count and percentage of participants with 1 or more of the specified AEs by treatment group and by SOC / preferred term respectively. At each level of summarization (e.g., treatment group level, SOC level, preferred term level), participants who experienced 1 or more incidences of the same event will be counted only once.

The count and percentage of participants with events of Suicidal Ideation or Behavior will be provided by treatment group and by post-baseline worst severity. Participant will be counted only once for the worst severity experienced.

Since safety should be assessed relative to study intervention exposure, all AE summary tables will include average (mean) number of study agent administrations received and average (mean) duration of follow-up in weeks ([date of last contact- reference date+1]/7). The table for Injection-site reactions will include number of injections and injections with injection-site reactions. The tables will present each SOC / preferred term by descending percentages of participants in the treatment group of 'Guselkumab CCI CCI + Golimumab CCI CCI

In addition to the summary tables, listings of participants with the following Treatment-emergent AEs will be provided:

- Treatment-emergent Serious AE
- Treatment-emergent AE leading to discontinuation of study intervention.
- Death
- Treatment-emergent Serious Infections
- Suicidal Ideation or Behavior
- Post-baseline positive or suspicious tuberculosis (TB) testing results

The data listings will be sorted by treatment group, participant identification, AE onset date and preferred term, wherever appropriate. The listings will also present the treatment that the participant is on when a specified event is reported.

5.4.2.2. Pharmacokinetics

To accomplish the second secondary endpoint, the Pharmacokinetics data analysis method are discussed below.

The Pharmacokinetic (PK) analysis will be performed at Week 24 and Week 36 DBL based on the Pharmacokinetics Analysis Set (Section 4), that is all randomized participants who have received at least one administration of golimumab and guselkumab and had at least one valid post-injection blood sample drawn. Participants will be analyzed according to the treatment groups that they actually received, i.e., regardless of the intervention that participants randomized to.

Serum samples will be analyzed to determine serum guselkumab and golimumab concentration. A concentration not quantifiable (below LLOQ) will be treated as 0 in the summary statistics and shown as the lower limit of quantification (<LLOQ) in the data listings. No imputation for missing data will be performed.

Descriptive statistics (N, mean, SD, median, range, interquartile range) will be used to summarize concentrations at each sampling time point. PK data may be displayed graphically, such as mean +/- SD PK concentrations over time by intervention group. The following analyses will be performed by treatment group as appropriate:

- Summary of serum guselkumab and golimumab concentration at each visit
- Proportion of participants without detectable serum guselkumab and golimumab concentration at each visit
- Summary of serum guselkumab and golimumab concentration at each visit by body weight quartiles

- Summary of serum guselkumab ang golimumab concentration at each visit by baseline MTX use (Yes/No)
- Summary of serum guselkumab and golimumab concentrations by baseline CRP levels
- Plot of median serum guselkumab and golimumab concentrations over time

In addition, the relationship between serum guselkumab and golimumab concentrations and safety or efficacy may be explored.

Once a participant meets one of the following dosing deviation criteria, the participant's data will be excluded from the by visit data analyses from that point onward:

- Discontinue either guselkumab or golimumab administrations.
- Skipped a guselkumab or golimumab administration.
- Received an incomplete/incorrect dose.
- Received an incorrect study agent.
- Received an additional study agent.

In addition, if a participant has an administration outside of visit window (Section 5.1.1), the concentration data collected at and after that visit will be excluded from the by-visit data analyses. Additional exclusions for incongruous PK data to be implemented based on Janssen TV-SOP-11008 Version 4.0 and TV-GDL-00362 Version 5.0. All participants and samples excluded from analysis will be documented in the Clinical Study Report.

If sufficient data are available, the population PK analysis using serum concentration-time data of guselkumab or golimumab will be performed using nonlinear mixed-effects modeling. Details may be given in a population PK analysis plan and the results of the analysis will be presented in a separate report.

5.4.2.3. Immunogenicity

To accomplish the third secondary endpoint, the Immunogenicity data analysis method are discussed below.

The Immunogenicity analysis will be performed at Week 24 and Week 36 DBL based on the Immunogenicity Analysis Set (Section 4), that is all randomized participants who have received at least one administration of guselkumab and golimumab and had appropriate samples for antibodies to guselkumab and golimumab detection. Participants will be analyzed according to the treatment groups that they received, i.e., regardless of the intervention that participants randomized to.

Serum samples will be collected to examine the formation of antibodies to guselkumab and antibodies to golimumab at the specified visits as shown in the schedule of events of the protocol. Serum samples will also be collected at the final visit from participants who terminate prematurely.

The following analysis of antibodies to guselkumab and golimumab will be performed by treatment group:

- Summary of antibodies to guselkumab or antibodies to golimumab status
- Summary of neutralizing antibodies to guselkumab or antibodies to golimumab status

- List of participants positive for antibodies to guselkumab and/or antibodies to golimumab

In addition, to explore the relationship between antibodies to guselkumab and golimumab status, and serum guselkumab and golimumab concentration, efficacy and safety, the following analysis may be performed as appropriate:

- Summary of clinical response (e.g., MDA, ACR 50, PASI) by antibody to guselkumab or antibodies to golimumab status
- Summary of injection site reaction by antibody to guselkumab or antibodies to golimumab status
- Summary of serum guselkumab and golimumab concentration by antibody to guselkumab or antibodies to golimumab status
- Plots of median and IQ through serum guselkumab and golimumab concentration over time by antibody to guselkumab or antibodies to golimumab status

5.5. Other Endpoints Analysis

5.5.1. Other Efficacy Endpoints

Table 10 – Other Efficacy Endpoints

Other efficacy endpoints	
1	Proportion of participants who achieve MDA by visit over time through Week 24
2	Proportion of participants who achieve ACR 20 and ACR 70 at Week 24
3	Proportion of participants who achieve ACR 20/50/70 response at Week 16
4	Change from baseline in PASI score by visit over time through Week 24 among participants with ≥3% BSA psoriatic involvement and an IGA score of ≥2 at baseline.
5	Change from baseline in Enthesitis score (based on Leeds Enthesitis Index [LEI]) by visit over time through Week 24 among the participants with Enthesitis (LEI) at baseline.
6	Change from baseline in Enthesitis score (based on Spondylarthritis Research Consortium of Canada [SPARCC]) by visit over time through Week 24 among the participants with Enthesitis (SPARCC) at baseline.
7	Change from baseline in dactylitis score by visit over time through Week 24 among the participants with dactylitis at baseline.
8	Change from baseline in Patient-Reported Outcomes Measurement Information System (PROMIS-29 v2.1) at Week 24
9	Proportion of participants who achieve MDA by Baseline Rheumatoid Factor and Anti-Cyclic Citrullinated Peptide (CCP) Antibodies Result at Week 16 and Week 24
10	Proportion of participants who achieve ACR 50 by Baseline Rheumatoid Factor and Anti-Cyclic Citrullinated Peptide (CCP) Antibodies Result at Week 16 and Week 24
11	Proportion of participants who achieve ACR 20/50/70 by visit over time through Week 24
12	Proportion of participants who achieve >=0.35 improvement in HAQ-DI Score from Baseline by visit through Week 24
13	Change from baseline in HAQ-DI score by visit through Week 24
14	Proportion of participants who achieve ≥4 Improvement in FACIT-Fatigue from Baseline by Visit through Week 24
15	Change from baseline in FACIT-Fatigue Score by visit through Week 24
16	Change from baseline in SF-36 PCS score by visit through Week 24
17	Change from baseline in SF-36 MCS score by visit through Week 24
18	Change from baseline in PROMIS-29 score by visit through Week 24
19	Change from baseline in DAPSA by visit through Week 24
20	Change from baseline in cDAPSA by visit through Week 24
21	Proportion of participants who achieve Resolution of Dactylitis Response by Visit through Week 24 among the participants with Dactylitis at Baseline
22	Proportion of participants who achieve Resolution of Enthesitis (LEI) Response by Visit through Week 24 among the participants with Enthesitis (LEI) at Baseline
23	Proportion of participants who achieve Resolution of Enthesitis (SPARCC) Response by Visit through Week 24 among the participants with Enthesitis (SPARCC) at Baseline

24	Proportion of participants who achieve PASI 90/100 by Visit through Week 24 among the participants with ≥3% BSA psoriatic involvement and an IGA score of ≥2 at baseline.
25	Proportion of participants who achieve PASI 100 by Visit through Week 24 among the participants with ≥3% BSA psoriatic involvement.
26	Proportion of participants who achieve Low Disease Activity or Remission Based on the Disease Activity Index for Psoriatic Arthritis (DAPSA) by Visit through Week 24

5.5.1.1. Definition of Endpoints

MDA
For PASI definition, see Section 5.3.1

ACR
For ACR definition, see Section 5.4.1.1

PASI
For PASI definition, see Section 5.3.1

BSA
For BSA definition, see Section 5.3.1

HAQ-DI
For HAQ-DI definition, see Section 5.3.1

IGA
For IGA definition, see Section 5.4.1.1

SF-36 PCS/MCS
For SF-36 PCS/MCS definition, see Section 5.4.1.1

Dactylitis
For Dactylis definition, see Section 5.4.1.1

Enthesitis (LEI)
For Enthesitis definition (based on LEI), see Section 5.3.1

Enthesitis (SPARCC)
Enthesitis not only can be assessed with LEI, but it can also be assessed by the Spondylarthritis Research Consortium of Canada (SPARCC) The SPARCC developed a measure for Enthesitis in general spondylarthritis which evaluates the presence or absence of pain by applying local pressure to the following entheses:

- Supraspinatus insertion, left and right
- Medial epicondyle humerus, left and right
- Lateral epicondyle humerus left and right
- Greater trochanter, left and right
- Quadriceps to patella, left and right
- Patellar to tibia, left and right
- Achilles tendon insertion, left and right

- Plantar fascia, left and right

Tenderness on examination is recorded as either present (1) or absent (0) for each of the 16 sites for an overall score range of 0–16.

For change from baseline in Enthesitis score (based on SPARCC), it is a measure of the change in Enthesitis, where a negative change indicates an improvement, and a positive change indicates a worsening.

DAPSA

The Disease Activity in Psoriatic Arthritis (DAPSA) score is a derived score combining SJC66 (Section 5.3.1), TJC68 (Section 5.3.1), CRP (Section 5.4.1.1, mg/dL), PAP (Section 5.3.1, on a 100-unit VAS), and PGADAA (Section 5.4.1.1, on a 100-unit VAS). The DAPSA is a continuous parameter and is defined as follow.

If any of the component required for computing the DAPSA score is missing, the DAPSA score will be set to missing. $DAPSA = SJC66 + TJC68 + CRP + PAP/10 + PGADAA/10$

Change from baseline in DAPSA measures the change in disease activity, where a negative change indicates an improvement and positive change indicates a worsening.

Participants were assigned to the following disease activity categories with DAPSA: remission is when DAPSA ≤ 4; low disease activity is when DAPSA > 4 and ≤ 14; moderate disease activity is when DAPSA > 14 and ≤ 28; and high disease activity is when DAPSA > 28.

cDAPSA

The clinical DAPSA (cDAPSA) consists of the same parameters as the conventional DAPSA but omits CRP values and thus enables immediate assessment of disease activity without a time delay while waiting for CPR results. The cDAPSA has shown a high agreement with DAPSA.

If any of the component required for computing the cDAPSA score is missing, the cDAPSA score will be set to missing. $DAPSA = SJC66 + TJC68 + PAP/10 + PGADAA/10$

Disease activity classification for cDAPSA was remission when cDAPSA ≤ 4; low disease activity when cDAPSA > 4 and ≤ 13; moderate disease activity when cDAPSA > 13 and ≤ 27; and high disease activity when cDAPSA > 27.

BASDAI

The Bath Ankylosing Spondylitis Disease Activity Index (BASDAI) was developed as a participant self-assessment for and consists of 6 questions relating to the 5 major symptoms of ankylosing spondylitis. Only participants with spondylitis with peripheral arthritis as their primary arthritic presentation of PsA will complete the BASDAI using a 10-unit VAS to indicate the degree of their symptoms over the past week on the following criteria:

- A. Fatigue on a 10-unit VAS
- B. Spinal pain on a 10-unit VAS

- C. Joint pain on a 10-unit VAS
- D. Enthesitis on a 10-unit VAS
- E. Qualitative morning stiffness on a 10-unit VAS
- F. Quantitative morning stiffness on a 10-unit VAS

The BASDAS is a continuous parameter and is defined as follow. If any of the components required for computing the BASDAI score is missing, then the BASDAI score will be set to missing. $BASDAI = 0.2 \times (A + B + C + D + 0.5 \times E + 0.5 \times F)$

Higher BASDAI score indicate greater disease severity and a score decrease of 50% or 2 points is considered clinically meaningful.

Change from baseline in BASDAI score measures the change in disease activity, where a negative change indicates an improvement, and a positive change indicates a worsening.

Note that change from baseline in BASDAI score, and ≥20%, ≥50%, ≥70%, and ≥90% improvement from baseline in BASDAI score are only applicable to participants with spondylitis and peripheral joint involvement as their primary arthritic presentation of PsA and BASDAI>0 at baseline.

PROMIS-29

Patient-Reported Outcomes Measurement Information System 29 Profile (PROMIS-29 v2.1) is a National Institutes of Health initiative to develop state-of-the-science self-report measure to assess functioning and well-being in physical, mental, and social domain of health (Hays, et al., 2018). It is a profile instrument intended for adults (ages 18+) and consists of a collection of short forms containing 4 items for each of 7 PROMIS domains (physical function, anxiety, depression, fatigue, sleep disturbance, pain interference, and ability to participate in social roles and activities) and an additional pain intensity 0-10 numeric rating scale.

The PROMIS-29 Profile is universal rather than disease-specific. They assess all domains over the past seven days except for Physical Function which has no timeframe specified. The raw score of each domain is converted into a standardized score with a mean of 50 and a standard deviation (SD) of 10 (T-Score). The standardized T-score is reported as the final score for each participant. Pain Intensity is presented as raw responses (0 -10). For PROMIS domains of Depression, Anxiety, Physical Function, Pain Interference, Fatigue, a score of 50 is the average for the United States general population with a standard deviation of 10, because testing was performed on a large sample of the general population. However, the other two domains (Ability to Participate in Social Roles and Activities and Sleep Disturbance) were not centered in a national sample. For these two domains, a score of 50 represents the average of the calibration sample which was generally more enriched for chronic illness, and a score of 50 likely represents somewhat sicker people than the general population. A higher PROMIS T -score represents more of the concept being measured. For negatively worded concepts like Anxiety, a T-score of 60 is one SD worse than average. By comparison, an Anxiety T -score of 40 is one SD better than average. However, for positively worded concepts like Physical Function, a T-score of 60 is better than average while a T-score of 40 is better.

PASDAS

The Psoriatic Arthritis Disease Activity Score (PASDAS) is a derived score combining PGAD (Section 5.4.1.1, on a 100-unit VAS), PGADAPA (Section 5.3.1, on a 100-unit VAS), SF-36 PCS, SJC66 (Section 5.3.1), TJC68 (Section 5.3.1), Enthesitis (LEI, Section 5.3.1), Dactylitis count (Section 5.4.1.1), CRP (Section 5.4.1.1, mg/dL).

The PASDAS is a continuous parameter and is defined as follow. If any of the component required for computing the PASDAS score is missing, the PASDAS score will be set to missing.

$$PASDAS = 1.5 \times [0.18 \times \sqrt{PGAD} + 0.159 \times \sqrt{PGADAA} - 0.253 \times \sqrt{SF36\ PCS} + 0.101 \times \log(SJC66 + 1) + 0.048 \times \log(TJC68 + 1) + 0.23 \times \log(LEI + 1) + 0.377 \times \log(Dactylitis\ count + 1) + 0.102 \times \log(CRP_{mg/dL} \times 10 + 1) + 2]$$

Change from baseline in PASDAS measures the change in disease activity, where a negative change indicates an improvement, and a positive change indicates a worsening.

Participants with very low or low disease activity based on the PASDAS score are those participants who have a PASDAS score less than or equal to 1.9, or greater than 1.9 and less than or equal to 3.2, respectively.

FACIT-Fatigue

The FACIT-Fatigue is a questionnaire that assesses self-reported tiredness, weakness, and difficulty conducting usual activities due to fatigue.1 The questionnaire consists of 13 questions that assess a participant's level of fatigue and tiredness over the last 7 days. Each question is graded on a 5-point scale (0 - 4); and accordingly, the total FACIT-Fatigue scores can range from 0 to 52, with lower score reflecting more fatigue and higher scores reflecting less fatigue. Note that, when at least 7 questions are answered, the score can be calculated and should be adjusted by the number of available questions.

Although not developed for PsA, the FACIT-Fatigue has been used to assess fatigue in clinical trials of participant with RA and has demonstrated sensitivity to change in these participants. In rheumatology, a change of 4 points is considered clinical meaningful and has been used as response definition in the RA population.

Change from baseline in FACIT-Fatigue scores measures the change in fatigue where a positive change indicates an improvement, and a negative change indicates a worsening.

5.5.1.2. Other Efficacy Endpoints Estimand

5.5.1.2.1. Main Estimand (Composite Variable)

Table 11 - Other Efficacy Endpoints Main Estimand Attributes

Study interventions: - Combination therapy: Guselkumab CCI C C and Golimumab CCI C C - Monotherapy: Guselkumab CCI C C and Placebo C C			
	Population	**Variable**	**Summary measure**
1	Same as the primary Estimand (Section 5.3.2.1 Table 4)	Binary response variable, where a responder is defined as a participant with an MDA response at all study scheduled visits and not having had ICE 1-3	Difference in proportion of responders (as defined in the Variable) between study intervention
2		Binary response variable, where a responder is defined as a participant with an ACR 20 response at Week 24 and not having had ICE 1-3	

3		Binary response variable, where a responder is defined as a participant with an ACR 70 response at Week 24 and not having had ICE 1-3	
4		Binary response variable, where a responder is defined as a participant with an ACR 20 response at Week 16 and not having had ICE 1-3	
5		Binary response variable, where a responder is defined as a participant with an ACR 50 response at Week 16 and not having had ICE 1-3	
6		Binary response variable, where a responder is defined as a participant with an ACR 70 response at Week 16 and not having had ICE 1-3	
7	Same as the primary Estimand (Section 5.3.2.1 Table 4), but limit to subset of participants with ≥3% BSA psoriatic involvement and an IGA score of ≥2 at baseline	Continuous variable, which is defined as the change from baseline by visit through Week 24 in PASI. In the case of an ICE 1-3, participant will be considered as having no change from baseline	Difference in average change (as defined in the Variable) between study intervention
8	Same as the primary Estimand (Section 5.3.2.1 Table 4), but limit to subset of participants with Enthesitis (LEI) at baseline	Continuous variable, which is defined as the change from baseline by visit through Week 24 in Enthesitis score (LEI). In the case of an ICE 1-3, participant will be considered as having no change from baseline	
9	Same as the primary Estimand (Section 5.3.2.1 Table 4), but limit to subset of participants with Enthesitis (SPARCC) at baseline	Continuous variable, which is defined as the change from baseline by visit through Week 24 in Enthesitis score (SPARCC). In the case of an ICE 1-3, participant will be considered as having no change from baseline	
10	Same as the primary Estimand (Section 5.3.2.1 Table 4), but limit to subset of participants with dactylitis at baseline	Continuous variable, which is defined as the change from baseline by visit through Week 24 in Dactylitis score. In the case of an ICE 1-3, participant will be considered as having no change from baseline	
11	Same as the primary Estimand (Section 5.3.2.1 Table 4)	Continuous variable, which is defined as the change from baseline at Week 24 in PROMIS-29. In the case of an ICE 1-3, participant will be considered as having no change from baseline	
12	Same as primary estimand (Section 5.3.2.1 Table 4)	Binary response variable, where a responder is defined as a participant with an ACR 20 response at all study scheduled visit and not having ICE 1-3	Difference in proportion of responders (as defined in Variable) between study interventions
13		Binary response variable, where a responder is defined as a participant with an ACR 50 response at all study scheduled visit and not having ICE 1-3	
14		Binary response variable, where a responder is defined as a participant with an ACR 70 response at all study scheduled visit and not having ICE 1-3	
15		Binary response variable, where a responder is defined as a participant with >=0.35 improvement in HAQ-DI score at all study scheduled visit and not having ICE 1-3	
16		Binary response variable, where a responder is defined as a participant with >=4 improvement in FACIT-Fatigue at Week 24 and not having ICE 1-3	
17		Continuous variable, which is defined as the change from baseline at all study scheduled visit in HAQ-DI. In the case of an ICE, participant will be considered as having no change from baseline.	Difference in average change (as defined in Variable) between study interventions
18		Continuous variable, which is defined as the change from baseline at all study scheduled visit in FACIT-Fatigue. In the case of an ICE, participant will be considered as having no change from baseline.	
19		Continuous variable, which is defined as the change from baseline at all study scheduled visit in SF-36 PCS. In the case of an ICE, participant will be considered as having no change from baseline.	
20		Continuous variable, which is defined as the change from baseline at all study scheduled visit in SF-36 MCS. In the case of an ICE, participant will be considered as having no change from baseline.	
21		Continuous variable, which is defined as the change from baseline at all study scheduled visit in PROMIS-29. In the case of an ICE, participant will be considered as having no change from baseline.	

22		Continuous variable, which is defined as the change from baseline at all study scheduled visit in DAPSA. In the case of an ICE, participant will be considered as having no change from baseline.	
23		Continuous variable, which is defined as the change from baseline at all study scheduled visit in cDAPSA. In the case of an ICE, participant will be considered as having no change from baseline.	
24	Same as primary estimand (Section 5.3.2.1 Table 4), but limit to participant with Dactylitis at Baseline	Binary response variable, where a responder is defined as a participant with an Dactylitis response at all study scheduled visit and not having ICE 1-3	Difference in proportion of responders (as defined in Variable) between study interventions
25	Same as primary estimand (Section 5.3.2.1 Table 4), but limit to participant with Enthesitis (LEI) at Baseline	Binary response variable, where a responder is defined as a participant with an Enthesitis (LEI) response at all study scheduled visit and not having ICE 1-3	
26	Same as primary estimand (Section 5.3.2.1 Table 4), but limit to participant with Enthesitis (SPARCC) at Baseline	Binary response variable, where a responder is defined as a participant with an Enthesitis (SPARCC) response at all study scheduled visit and not having ICE 1-3	
27	Same as primary estimand (Section 5.3.2.1 Table 4), but limit to participant with >=3% BSA psoriatic involvement and an IGA score of >=2 at Baseline	Binary response variable, where a responder is defined as a participant with an PASI 90 response at all study scheduled visit and not having ICE 1-3	
28	Same as primary estimand (Section 5.3.2.1 Table 4), but limit to participant with >=3% BSA psoriatic involvement and an IGA score of >=2 at Baseline	Binary response variable, where a responder is defined as a participant with an PASI 100 response at all study scheduled visit and not having ICE 1-3	
29	Same as primary estimand (Section 5.3.2.1 Table 4), but limit to participant with >=3% BSA psoriatic involvement at Baseline	Binary response variable, where a responder is defined as a participant with an PASI 100 response at all study scheduled visit and not having ICE 1-3	
30	Same as primary estimand (Section 5.3.2.1 Table 4)	Binary response variable, where a responder is defined as a participant with a low disease activity or remission based on DAPSA at all study scheduled visit and not having ICE 1-3	
ICE and the corresponding strategies			
1. Initiated protocol-prohibited medications and therapies for PsA 2. Initiation or increased the dose of non-biologic DMARDs (MTX, SSZ, HCQ, LEF) or oral corticosteroid over baseline for PsA 3. Discontinued study intervention due to any reason			Composite Variable Strategy

5.5.1.2.2. Supplementary Estimand (Treatment Policy)

For all other efficacy endpoints' supplementary estimand, it will share same population, study intervention, ICEs, variable and summary measures as the other efficacy endpoints' main estimand (Section 5.5.1.2.1). The only difference is the ICE strategy, which will be Treatment Policy Strategy, i.e., a participant with this ICE will not be treated differently from participant without ICE. The occurrence of the ICE is irrelevant.

5.5.1.3. Analysis Methods

All other efficacy endpoint analysis will follow same analysis method as major secondary endpoint analysis (Section 5.4.1.3). The difference is missing data for Supplementary Estimand will not be imputed (for both binary and continuous endpoints).

5.5.1.4. Sensitivity Analysis

To evaluate the robustness of the assumption of all the missing data are non-responder, a sensitivity analysis with multiple imputation will be performed on MDA by visit and ACR50 at Week 16 only. More detail on MI method is specified in Appendix 3.

5.5.1.5. Summary of Analyses

Table 7 - Other Efficacy Endpoints Analysis Summary

Binary endpoints			
	Main Estimand (Section 5.5.1.2.1)	Missing data will be assumed as MAR and imputed conservatively as non-responder	- Summarized descriptively. - Odds ratio, CI, and P-value from logistic regression
	Supplementary Estimand (Section 5.5.1.2.2)	Missing data will not be imputed	
	Sensitivity Analysis for Main Estimand for just MDA by visit and ACR50 at Week 16 (Section 5.5.1.4)	Missing data will be assumed as MAR and imputed through MI [a]	- Summarized descriptively. - Pooled odds ratio, CI, and P-value from logistic regression across multiply imputed data sets
Continuous endpoints			
	Main Estimand (Section 5.5.1.2.1)	Missing data will be assumed as MAR and imputed through MI [a]	- Summarized descriptively. - Treatment comparisons including p-value from ANCOVA model.
	Supplementary Estimand (Section 5.5.1.2.2)	Missing data will not be imputed	
[a] MI method is summarized in Appendix 3			

5.5.2. Other Safety Endpoints

The other safety endpoints are:

1) Suicidal ideation and behavior
2) Laboratory parameters and change from baseline in laboratory parameters (hematology and chemistry)
3) Vitals signs over time

All other safety endpoints will be analyzed at Week 36 DBL based on the Safety Analysis Set (Section 4), that is all randomized participants who have received at least one administration of study intervention will be included in analysis according to the intervention they actually receive.

For continuous outcome, descriptive status include mean, standard deviation, median, quantiles, minimum and maximum will be prepared by study intervention group. For categorical outcome, event frequency counts and percentages will be prepared by study intervention group.

5.5.2.1. Suicidal Ideation and Behavior

To address the first other safety endpoint, the analysis method is discussed below.

The electronic Columbia-Suicide Severity Rating Scale (eC-SSRS) will be used as a screening tool to prospectively evaluate suicidal ideation and behavior among study participants. The eC-SSRS measures 5 possible levels of suicidal ideation and 4 possible suicidal behaviors, as well as non-suicidal self-injurious behavior. The assessment is a fully structured, participant self-report eC-SSRS questionnaire, including standardized questions, follow-up prompts, error handling routines, and scoring conventions (Mundt 2013; Posner 2011).

Two versions of the eC-SSRS will be used in this study, the Lifetime version and the Since Last Contact version. The Lifetime version will be conducted during the screening visit and the Since Last Contact version will be conducted at all other visits. Participants will complete the eC-SSRS questionnaire using the sponsor-provided electronic device. Study site personnel will train the participants on how to use the electronic device. The eC-SSRS will be provided in the local languages in accordance with local guidelines.

Any eC-SSRS findings, which in the opinion of the investigator are new or considered to be a worsening and clinically significant, will be reported on the AE eCRF. As discussed previously (Section 5.4.2.1), suicidal ideation and behavior will be tabulated by treatment group and MedDRA System Organ Class (SOC) and be sorted in descending frequency of the Preferred Term within the SOC.

5.5.2.2. Clinical Laboratory Tests

To address the second other safety endpoint, the analysis method is discussed below.

Clinical laboratory tests will be displayed for the participants included in the Safety Analysis Set.

Applicable clinical laboratory test values are to be graded using the criteria defined in Appendix 13. All of these criteria come from to National Cancer Institute's Common Terminology Criteria for Adverse Events (NCI-CTCAE) version 5.0.

Hematology and chemistry parameters (see below) values by worst toxicity grade, post baseline, according to NCI-CTCAE will be summarized by treatment group:

- Hematology: hematocrit, hemoglobin, lymphocytes, neutrophils, platelet count, total and differential White Blood Cell (WBC) count
- Chemistry: total and direct bilirubin, Aspartate aminotransferase (AST), Alanine aminotransferase (ALT), alkaline phosphatase (ALP), albumin, total protein, calcium, phosphate, sodium, potassium, chloride, blood urea nitrogen (BUN), creatinine, glucose

Number of participants with a shift from baseline value to post baseline value with respect to normal reference range for below selected chemistry and hematology laboratory tests will be summarized. Box plot of observed value and change from baseline to week 24 by visit will be provided for the following laboratory tests:

- Hematology: Platelet count, Hemoglobin, White Blood Cell (WBC) count (neutrophils, lymphocytes)
- Chemistry: Aspartate aminotransferase (AST), Alanine aminotransferase (ALT), total and direct bilirubin, Gamma-glutamyl transferase (GGT), alkaline phosphatase (ALP)

Listings of participants with any post-baseline toxicity grade 2 or higher hematology and chemistry laboratory value will be presented.

Listings of participants with any liver function test (ALT, AST) greater than 3 times upper limit of normal post-baseline will be presented.

For by-visit data summaries, only the data from the scheduled visits will be included in the analyses. For data summaries not by visit (such as: post-baseline maximum lab toxicity grade, post-baseline worst or maximum value. etc.) and for data listings, data from both scheduled visits and unscheduled visits will be included.

5.5.2.3. Vital Signs and Physical Examination Findings

To address the third other safety endpoint, the analysis method is discussed below.

Continuous vital sign parameters including temperature, pulse/heart rate, and blood pressure (systolic and diastolic) will be summarized over time. BMI will be calculated as weight (kg)/ (height (m))^2. The weight, height measurement collected at baseline (Week 0) will be used in the calculation. Percent change from baseline will be summarized for the intervention period. Descriptive statistics (mean, standard deviation, median, minimum, and maximum) will be prepared.

Abnormality criteria (based on criteria defined below) will be applied to baseline and postbaseline values. For baseline values, increase or decrease criteria are not applied. Postbaseline values will be considered Treatment-Emergent if they meet both value and change criteria in the table below.

Incidence of treatment-emergent vital signs during intervention, as defined in Table below, will be summarized for participants who had a baseline assessment and at least 1 postbaseline assessment for that vital sign.

Table 8 - Clinically Important/Markedly Abnormal Vital Signs

Vital Sign	Markedly Abnormal Low	Markedly Abnormal High
Temperature	<35°C and with >=1°C decrease from baseline	>38°C and with >=1°C increase from baseline
Pulse/Heart rate	Absolute value ≤ 50 bpm and a decrease from baseline by ≥ 15 bpm	Absolute value ≥ 120 bpm and an increase from baseline by ≥ 15 bpm
Systolic blood pressure	Absolute value ≤ 90 mm Hg and a decrease from baseline by ≥ 20 mm Hg	Absolute value ≥ 180 mm Hg and an increase from baseline by ≥ 20 mm Hg
Diastolic blood pressure	Absolute value ≤ 50 mm Hg and a decrease from baseline by ≥ 15 mm Hg	Absolute value ≥ 105 mm Hg and an increase from baseline by ≥ 15 mm Hg

Treatment-Emergent will be reported as adverse events and are included in the analysis of adverse events (Section 5.4.2.1)

5.6. Exploratory Endpoints Analysis

5.6.1. Pharmacodynamic

To evaluate the Pharmacodynamic (PD) effects of combination therapy in participants with active PsA compared with monotherapy, one PD endpoint is planned: change from baseline in serum T helper 17 cytokines and select inflammatory cytokines and other proteins in the IL-23 and TNFα pathway by visit over time.

The PD endpoint will be analyzed at Week 36 DBL based on the Pharmacodynamic Analysis Set (Section 4), that is all randomized participants who have received at least one administration of

study intervention and had appropriate samples for antibodies to guselkumab and golimumab detection will be included in analysis according to the intervention they received, i.e., regardless of the intervention that participants randomized to.

Samples for the analysis of PD biomarkers will be collected. Serum level of T helper 17 cytokines, including but not limited to IL-17A, IL-17F, and IL-22, and select inflammatory cytokines may be measured to assess the PD effect of guselkumab and/or golimumab.

5.6.2. Magnetic Resonance Imaging

Table 14 – Magnetic Resonance Imaging (MRI) Endpoints

Psoriatic Arthritis Magnetic Resonance Imaging (PsAMRI)	
1	Change from baseline in PsAMRIS hand at Week 24 among participants with ≥1 PsAMRIS hand at baseline.
2	Change from baseline in PsAMRIS foot at Week 24 among participants with ≥1 PsAMRIS foot at baseline.
3	Change from baseline in PsAMRIS hand at Week 24
4	Change from baseline in PsAMRIS foot at Week 24
5	Change from baseline in PsAMRIS hand composite inflammatory score at Week 24
6	Change from baseline in PsAMRIS foot composite inflammatory score at Week 24
7	Change from baseline in PsAMRIS hand total inflammatory score at Week 24
8	Change from baseline in PsAMRIS foot total inflammatory score at Week 24
9	Change from baseline in PsAMRIS hand total damage score at Week 24
10	Change from baseline in PsAMRIS foot total damage score at Week 24
Heel Enthesitis Magnetic Resonance Imaging Score (HEMRIS)	
1	Change from baseline in HEMRIS at Week 24
2	Change from baseline in HEMRIS Achilles Tendon Inflammatory score at Week 24
3	Change from baseline in HEMRIS Achilles Tendon Structural score at Week 24
4	Change from baseline in HEMRIS Plantar Fascia Inflammatory score at Week 24
5	Change from baseline in HEMRIS Plantar Fascia Structural score at Week 24

5.6.2.1. Definitions of Endpoints

5.6.2.1.1. Clinical endpoints

Enthesitis (LEI)
For Enthesitis definition (based on LEI), see Section 5.3.1

Enthesitis (SPARCC)
For Enthesitis definition (based on SPARCC), see Section 5.5.1.1

5.6.2.1.2. MRI endpoints

Central reading and analysis of MRI scans will be used to evaluate drug efficacy from MRI exploratory endpoints. The efficacy assessment will involve the analysis of MRI scans using the Outcome Measures in Rheumatology Clinical Trials (OMERACT) scoring frameworks PsAMRIS (for application to assessment of the hand and forefoot), and HEMRIS (for application to assessment of the heel).

For each participant, ideally there should be two MRI readings, one from Week 0 and another one from Week 24 or earlier termination visit. In advance of the Week 0 MRI scanning visit for each study participant, determination of the right or left hand and foot to undergo longitudinal scanning will be made based on which has greater clinical manifestations of disease. Clinical study sites

will communicate this selection to their imaging center, and the laterality of target extremities to be scanned will not change over the course of the study. The reading of MRI scans will use two independent, primary central readers who will evaluate all participants enrolled in the trial. Discrepancies identified in the reading results of the two primary readers will go to confirmation reading by a third independent central reader (2+1 reading paradigm), blinded to the results of the primary reading. The images from all visits for a participant will be displayed in random time order with blinding of the chronological sequence. There will be separate assessments for PsAMRIS hand and forefoot and HEMRIS heel, so each of these assessments will have their own adjudication criterion based on a score threshold. The criteria for triggering adjudication can be found in the study imaging Charter.

For all the inflammation and structural score, in case of joint injection or joint surgical procedure, inflammation and structural score adjustment rule specified in Appendix 2 for overwriting image evaluation will be applied to those joints with surgical procedure and/or joint injection.

5.6.2.1.2.1. PsAMRIS

The PsAMRIS is a scoring system that assesses imaging-derived features typically associated with peripheral manifestations of PsA, resulting in a sum score (Boyesen, et al., 2011) (Abrar, et al., 2020) (McQueen, et al., 2009) (Ostergaard, et al., 2009). For the study, we will have PsAMRIS of the hand and forefoot for each participant, and each participant will be measured twice, one from Week 0 and another one at Week 24 or early termination visit.

For PsAMRIS assessment of the hand, both thumb and fingers of hand are considered for this endpoint, with three regions of each finger assessed: metacarpophalangeal (MCP) joint region, proximal interphalangeal (PIP) joint region, and distal interphalangeal (DIP) joint region. Acquired images will be evaluated for features of synovitis (score: 0-3), flexor tenosynovitis (score: 0-3), periarticular inflammation (score: 0 or 1), bone oedema (score: 0-3), bone erosion (score: 0-10), and bone proliferation (score: 0 or 1) in each joint region. See table below for further detail on the scoring framework.

PsAMRIS of the hand*											
MCP Joint Region		Week 0					Week 24/Early Termination				
		Finger 1	Finger 2	Finger 3	Finger 4	Finger 5	Finger 1	Finger 2	Finger 3	Finger 4	Finger 5
Synovitis (score 0-3)											
Flexor tenosynovitis (score 0-3)											
Periarticular inflammation	Volar (score 0 or 1)										
	Dorsal (score 0 or 1)										
Bone oedema	Proximal (score 0-3)										
	Distal (score 0-3)										
Bone erosion	Proximal (score 0-10)										
	Distal (score 0-10)										
Bone proliferation (score 0 or 1)											
PIP Joint Region		Week 0					Week 24/Early Termination				
		Finger 1	Finger 2	Finger 3	Finger 4	Finger 5	Finger 1	Finger 2	Finger 3	Finger 4	Finger 5
Synovitis (score 0-3)											
Flexor tenosynovitis (score 0-3)											
Periarticular inflammation	Palmar (score 0 or 1)										
	Dorsal (score 0 or 1)										
Bone oedema	Proximal (score 0-3)										
	Distal (score 0-3)										

Bone erosion	Proximal (score 0-10)										
	Distal (score 0-10)										
Bone proliferation (score 0 or 1)											
DIP Joint Region		Week 0					Week 24/Early Termination				
		Finger 1	Finger 2	Finger 3	Finger 4	Finger 5	Finger 1	Finger 2	Finger 3	Finger 4	Finger 5
Synovitis (score 0-3)											
Flexor tenosynovitis (score 0-3)											
Periarticular inflammation	Palmar (score 0 or 1)										
	Dorsal (score 0 or 1)										
Bone oedema	Proximal (score 0-3)										
	Distal (score 0-3)										
Bone erosion	Proximal (score 0-10)										
	Distal (score 0-10)										
Bone proliferation (score 0 or 1)											
* Note: any score can be NA (Not Assessable) if there is a joint replacement, amputation or other surgery, inadequate image quality											

Based on above table, one can tell that PsAMRIS of the hand:

1) The total synovitis score can range from 0 to 42
2) The total flexor tenosynovitis score can range from 0 to 42
3) The total periarticular inflammation score can range from 0 to 28
4) The total bone oedema score can range from 0 to 84
5) The total bone erosion score can range from 0 to 280
6) The total bone proliferation score can range from 0 to 14

For PsAMRIS assessment of the forefoot, five toes of the foot are considered for this endpoint, with two regions assessed: the metatarsophalangeal (MTP) joint region and interphalangeal (IP) joint region. For MTP, all five joints will be assessed, while only the IP joint region of the 1st toe will be assessed. See table below for further detail on the scoring framework.

PsAMRIS of the forefoot *											
MTP Joint Region		Week 0					Week 24/Early Termination				
		Toe 1	Toe 2	Toe 3	Toe 4	Toe 5	Toe 1	Toe 2	Toe 3	Toe 4	Toe 5
Synovitis (score 0-3)											
Flexor tenosynovitis (score 0-3)											
Periarticular inflammation	Volar (score 0 or 1)										
	Dorsal (score 0 or 1)										
Bone oedema	Proximal (M1, score 0-3)										
	Distal (M2, score 0-3)										
Bone erosion	Proximal (M1, score 0-10)										
	Distal (M2, score 0-10)										
Bone proliferation (score 0 or 1)											
PIP Joint Region		Week 0					Week 24/Early Termination				
		Toe 1	-	-	-	-	Toe 1	-	-	-	-
Synovitis (score 0-3)											
Flexor tenosynovitis (score 0-3)											
Periarticular inflammation	Palmar (score 0 or 1)										
	Dorsal (score 0 or 1)										
Bone oedema	Proximal (P1, score 0-3)										
	Distal (P2, score 0-3)										
Bone erosion	Proximal (P1, score 0-10)										
	Distal (P2, score 0-10)										
Bone proliferation (score 0 or 1)											
* Note: any score can be NA (Not Assessable) if there is a joint replacement, amputation or other surgery, inadequate image quality											

Based on above table, the six features of PsAMRIS forefoot have below score range:

1) The total synovitis score can range from 0 to 18

2) The total Flexor tenosynovitis score can range from 0 to 18
3) The total Periarticular inflammation score can range from 0 to 12
4) The total bone oedema score can range from 0 to 36
5) The total bone erosion score can range from 0 to 120
6) The total bone proliferation score can range from 0 to 6

For the PsAMRIS measurement, features are considered as either inflammatory or structural feature:

- Inflammatory features: Synovitis, Flexor tenosynovitis, Periarticular inflammation, Bone oedema
- Structural features: Bone erosion, Bone proliferation

PsAMRI Composite Inflammatory score = Bone oedema + 2* Synovitis + 2* Flexor tenosynovitis

PsAMRI Total Inflammatory score = Bone oedema + 2* Synovitis + 2* Flexor tenosynovitis + 3* Periarticular inflammation

PsAMRI Total Damage score = Bone erosion + 20* Bone proliferation

5.6.2.1.2.2. HEMRIS

HEMRIS is the first international consensus-based and validated, comprehensive MRI score system for peripheral Enthesitis in participants with spondylarthritis and was extended the use of it for ease of scoring Enthesitis in the heel region in participants with PsA as well. The scoring method will involve assessing inflammation and structural changes at the entheses in two anatomical regions: Achilles tendon and plantar fascia, and therefore, HEMRIS endpoints include:

a) Total Achilles Tendon Inflammation score is the summation of the scores from the following four features: Intra-tendon hypersignal, Peri-tendon hypersignal, Bone marrow oedema, and Retrocalcaneal bursitis, data range between 0 and 12.
b) Total Plantar Fascia Inflammation score is the summation of the scores from the following three features: Intra-fascia hypersignal, Peri-fascia hypersignal, and Bone marrow oedema, data range between 0 and 9.
c) Total Achilles Tendon Structure is the summation of the scores from the following three features: Tendon thickening, Calcaneal enthesophyte, and Calcaneal bone erosion, data range between 0 and 9.
d) Total Plantar Fascia Structure score is the summation of the scores from the following three features: Fascia thickening, Calcaneal enthesophyte, Calcaneal bone erosion, data range between 0 and 9.

The total HEMRIS score for heel will be calculated as the sum of above four scores, and therefore the score will range between 0 and 39.

For details on the HEMRIS scoring framework, see table below (Mathew, et al., 2020) (Mathew, et al., 2019):

HEMRIS for heel *			
Achilles tendon			
Inflammatory	Intra-tendon hypersignal 0. no intra-tendon hypersignal 1. minimal intra-tendon hypersignal spots (≤25% of the tendon volume) 2. moderate intra-tendon hypersignal (>25% and ≤50% of the tendon volume) 3. Severe intra-tendon hypersignal (>50% of the tendon volume)	Peri-tendon hypersignal 0. no hypersignal 1. mild focal hypersignal 2. moderate hypersignal 3. severe hypersignal	Bone marrow oedema 0. No oedema 1. 1%-33% of the bone is edematous 2. 34%-66% pf the bone is edematous 3. 67%-100% of the bone is edematous
	Retrocalcaneal bursitis 0. No hypersignal or maximal diameter of hypersignal in the sorter or two perpendicular dimensions to be <0.25cm 1. Maximal diameter of hypersignal in the shorter of two perpendicular dimensions to be ≥0.25cm to <0.5cm 2. Maximal diameter of hypersignal in the shorter of two perpendicular dimensions to be ≥05cm to <1cm 3. Maximal diameter of hypersignal in the shorter of two perpendicular dimensions to be ≥1cm		
Structural	Tendon thickening 0. None 1. Mild 2. Moderate 3. Severe	Calcaneal enthesophyte 0. None 1. Small 2. Medium-sized 3. Large	Calcaneal bone erosion 0. None 1. Small 2. Medium-sized 3. Large
Plantar fascia			
Inflammatory	Intra-fascia hypersignal 0. no intra-fascia hypersignal 1. minimal intra-fascia hypersignal spots (≤25% of the fascia volume) 2. moderate intra-fascia hypersignal (>25% and ≤50% of the fascia volume) 3. Severe intra-fascia hypersignal (>50% of the fascia volume)	Peri-fascia hypersignal 0. no hypersignal 1. mild focal hypersignal 2. moderate hypersignal 3. severe hypersignal	Bone marrow oedema 0. No oedema 1. 1%-33% of the bone is edematous 2. 34%-66% pf the bone is edematous 3. 67%-100% of the bone is edematous
Structural	Fascia thickening 0. None 1. Mild 2. Moderate 3. Severe	Calcaneal enthesophyte 0. None 1. Small 2. Medium-sized 3. Large	Calcaneal bone erosion 0. None 1. Small 2. Medium-sized 3. Large

5.6.2.2. Analysis Methods

The exploratory MRI endpoints will be analyzed based on the FAS (Section 4), that is all randomized participants will be included in analysis according to the intervention they were randomized into, regardless of the treatment participant actually received.

Status and change score were described by calculating the average of corresponding scores from the 2 primary readers for a participant without adjudication. In the case of adjudication, the average of the two closest reads will be used. In the case that there are more than one pair or two closest reads, the average of three reads will be used.

Below are the steps (in order) that will be used for preparation of MRI endpoint data for analysis:

1) In the case that participant have had joint injection or joint procedure, MRI endpoint (only PsAMRIS) will be adjusted as per Appendix 2
2) In the case of incomplete joint or un-evaluable joint, the corresponding feature score will be considered as missing. And therefore, the summary score that containing the missing feature score will be considered as missing as well.
3) In the case of missing data, the missing data will be treated as Missing Completely at Random (MCAR) and will not be imputed.

For all the MRI endpoints (change from baseline), complete case analysis will be performed. Descriptive statistics (i.e., mean, median, SD, IQ range, minimum, and maximum) will be used to summarize continuous variables. The treatment difference between combination therapy and monotherapy will be tested using Analysis of Covariance (ANCOVA). The explanatory variable might include treatment group, non-biologic DMARDs use, and endpoint baseline measurement.

Intra-reader variability will be assessed. More specifically, MRI images (at both Week 0 and Week 24) of 15 participants will be randomly selected from the first approximately 40 enrolled participants and re-read by each of the two primary readers. The scores from the re-read will be used for intra-reader correlation analysis. Inter-rater reliability will be assessed using intraclass correlation coefficients (ICC) (Koo & Li, 2016) The smallest detectable change (SDC) will be calculated for change score (Bruynesteyn, et al., 2005). ICC is interpreted as follows:

- <0 Poor Agreement
- 0-0.20 Slight Agreement
- 0.21-0.40 Fair Agreement
- 0.41-0.60 Moderate Agreement
- 0.61-0.80 Substantial Agreement

5.7. Other Analyses

5.7.1. Extent of Exposure

The number and percentage of participants who receive each study intervention will be summarized, overall and by visit. Summary will be prepared based on the Safety Analysis Set.

5.7.2. Biomarkers

Planned biomarker analyses may be deferred if emerging study data show no likelihood of providing useful scientific information or insufficient number of samples are available for analyses. Any biomarker samples received after the cutoff date will not be analyzed, and therefore, excluded from the biomarker analysis.

The biomarker analyses will be used to examine the biologic response to treatment and to identify biomarkers that are relevant to treatment, and determine if these markers can predict response to guselkumab and/or golimumab, where local regulations permit.

The biomarker analysis may include but not limited to serum T helper 17 cytokines, inflammatory marker, whole blood RNA profile, and other categories of biomarkers potentially involved in the development and the progression of PsA.

Change in biomarkers over time may be summarized by treatment group. Associations between baseline levels and changes from baseline in select markets and clinical response may be explored. Results of biomarker analyses may also be summarized in a separate technical report.

5.7.3. Pharmacokinetic/Pharmacodynamic Analyses

If data permit, the relationships between serum concentrations of guselkumab and golimumab (Section 5.4.2.2) and efficacy may be analyzed graphically. If any visual trend is observed, a suitable population PK/PD model may be developed to describe the exposure-response relationship. Details may be given in a population PK/PD analysis plan and results of the population PK/PD analysis will be presented in a separate technical report.

5.7.4. Pharmacogenomic Analyses

Genetic (DNA) analyses may be conducted only in participants who consent in the pharmacogenomic sampling. DNA sample may be used for research related to guselkumab and/or golimumab. It may also be used to develop tests/assay related to guselkumab or golimumab. Pharmacogenomic research may consist of the analysis of one or more candidate genes or of the analysis of genetic marker throughout the genome or analysis of the entire genome (as appropriate) in relation to guselkumab or golimumab or PsA clinical endpoint.

5.7.5. Definition of Subgroups

To evaluate the consistency in the primary efficacy endpoint (proportion of participants who achieve MDA at Week 24) over demographics, baseline disease characteristic, prior and baseline medication use, subgroup analyses will be performed.

Subgroup analysis will be based on the FAS (Section 4), that is all randomized participants will be included in analysis according to the intervention they were randomized into, regardless of the treatment participant actually receive.

Summary statistics for the subgroup analysis will be prepared by treatment group. Descriptive statistics for continuous data will include the number of participants included in the summary (n), mean, standard deviation, median, quantiles, minimum and maximum. For quantile, minimum, and maximum, the same number of decimal places as in the raw data will be presented when reporting. For mean, median, one more decimal place than the raw data will be reported. For standard deviation, two more decimal places than the raw data will be reported.

For categorical data, frequency count and percentages will be used for summary. For all the percentages, one decimal place will be used with two exceptions: 100% will be presented as 100% (no decimal), and 0% will be not presented.

The subgroups include, but are not limited to, the following:

1. **Demographic subgroups**
   - Gender: Male, Female
   - Age: ≥18 to <45, ≥45 to ≤65
   - Race: White, OTHER
   - Body weight: (quantiles)
   - Body mass index: normal (<25), overweight (≥25 to <30) and obese (≥30)

- Country: Denmark, France, Hungary, Italy, Poland, Russia, Spain, Ukraine, United States

2. **Baseline disease characteristics subgroups**
   - PsA duration: <1 year, ≥1 year to <3 year, ≥3 year
   - PsA subtype: Distal interphalangeal joint involvement, Polyarticular arthritis with absence of rheumatoid nodules, Asymmetric peripheral arthritis, Spondylitis with peripheral arthritis
   - Number of swollen joints: <10, ≥10 to ≤15, >15
   - Number of tender joints: <10, ≥10 to ≤15, >15
   - HAQ-DI: <1, ≥1 to ≤2, >2
   - CRP (mg/dL): <1, ≥1 to <2, ≥2
   - CRP: (quantiles)
   - Dactylitis: Yes, No
   - Enthesitis: Yes, No
   - PASI: <12, ≥12 to ≤20, >20
   - BSA: <3%, ≥3% to <10%, ≥10% to <20%, ≥20%
   - PAIN: ≤50, >50 and ≤67, >67 and ≤79, >79
   - IGA: <2, ≥2
3. **Prior and baseline medication use**
   - Number of prior anti-TNF agents: 1 or 2
   - Reason for discontinuation for prior anti-TNF: primary inadequate response, secondary inadequate response
   - Use of non-biologic DMARDs use (MTX, SSZ, HCQ, LEF): Yes, No
   - Use of MTX: Yes, No
   - Use of oral corticosteroids: Yes, No
   - Use of NSAIDs: Yes, No
   - Number of prior non-biologic treatments including DMARDs, systemic immunosuppressive drugs, and Apremilast: 0, 1, 2, ≥3

Note: some of the subgroup cut-off points may be changes if there are no or few participants within a subgroup category.

5.8. Interim Analysis

One Interim Analysis (IA) is planned for this study, and that is when approximately 60 participants of the planned participants complete the Week 24 assessment. Participants, regardless of their study intervention discontinuation and/or termination status will be counted towards the 60. The objective of the interim analysis is to determine the feasibility of advancing Phase 2b study planning activities. A separate interim analysis plan (IAP) will be prepared for this purpose.

Further details of the interim analysis will be in the Interim Analysis Charter and SAP.

5.9. Data Monitoring Committee (DMC) or Other Review Board

Data Monitoring Committee

An independent external DMC will be established to monitor data on an ongoing basis to ensure the continuing safety of the participants enrolled in this study. This committee will consist of at least one medical expert in the relevant therapeutic area and at least one statistician; committee membership responsibilities, authorities, and procedures will be documented in DMC charter. Analysis that planned for DMC will be provided in the DMC SAP.

Interim Analysis Committee

An internal unblinded Interim Analysis Committee (IAC) will be established to review IA data as specified in IAP. This committee will consist of physicians, statisticians, clinical pharmacologists, and pharmacometricians unaffiliated with the study; committee membership responsibilities, authorities, and procedures will be documented in its charter. Recommendations based on the IA results, which will be specified in the IAC SAP/Charter, will be communicated to the Sponsor Committee. Further details of the IA will be provided in the IAC SAP/Charter.

6. SUPPORTING DOCUMENTATION

6.1. Appendix 1 List of Abbreviations

AE	adverse event
ALT	alanine aminotransferase
ALP	alkaline phosphatase
ANCOVA	analysis of covariance
Anti-TNFα	Tumor Necrosis Factor alpha
AST	aspartate aminotransferase
BMI	body mass index
BSA	body surface area
BUN	blood urea nitrogen
CI	confidence interval
CRF	case report form
CTCAE	Common Terminology Criteria for Adverse Events
CRP	C-reactive protein
cLDA	Constrained longitudinal data analysis
DAPSA	Disease Activity in Psoriatic Arthritis
DMC	Data Monitoring Committee
DBL	Database Lock
ECG	electrocardiogram
eC-SSRS	Electronic Columbia-suicide Severity Rating Scale
eCRF	electronic case report form
FAS	full analysis set
FDA	Food and Drug Administration
GDPT	Patient global assessment of disease activity (arthritis)
GDPTPSA	Patient global assessment of disease activity (arthritis and psoriasis)
GDEV	Physician global assessment of disease activity
GLMM	Generalized linear mixed model
GGT	Gamma-glutamyl transferase
HEMRIS	Heel Enthesitis Magnetic Resonance Imaging score
HAQ-DI	Disability index of the health assessment questionnaire
HCQ	Hydroxychloroquine
IAC	Interim analysis committee
ICH	International Conference on Harmonization
ICC	Intra-class correlation coefficient
ICE	Incurrent event
IGA	Investigator Global assessment
IQ	interquartile
IR	Inadequate response
LLOQ	lower limit of quantification
LEF	Leflunomide
LEI	Leeds Enthesitis Index
MAR	Missing at random
MDA	Minimal Disease Activity
MedDRA	Medical Dictionary for Regulatory Activities
MTX	Methotrexate
MMRM	Mixed-effect Model for repeated measure
MedDRA	Medical Dictionary for Regulatory Activities
NSAIDs	nonsteroidal anti-inflammatory drugs
NCI-CTCAE	National cancer institute's common terminology criteria for adverse events
PASI	Psoriasis area and severity index
PAIN	Patient global assessment of pain
PD	pharmacodynamic(s)
PK	pharmacokinetic(s)
PsA	Psoriatic Arthritis
PsAMRIS	Psoriatic Arthritis Magnetic Resonance Imaging

PROMIS-29	Patient reported outcomes measurement information system 29 profile
CCI	Every 4 week
SAE	Serious adverse event
SAP	Statistical Analysis Plan
SF-36 PCS	Short form health survey (36 items) physical component score
SSZ	Sulfasalazine
SD	standard deviation
SOC	System Organ Class
SPARCC	Spondylarthritis research consortium of Canada
SJC	Swollen joint count
TJC	Tender joint count
US NCI	United States National Cancer Institute
WBC	White blood cell

6.2. Appendix 2 Rules Applied in Definition of Endpoints

1. Joint Evaluability Rules for Sign and Symptom Data

If a joint cannot be assessed, mark the joint as 'non-evaluable'. Joints should only be marked as 'non-evaluable' if it is physically impossible to assess the joint (i.e., joint amputated or joint inaccessible due to a cast or a recent wound near or at the joint, or artificial joint). Otherwise, the joint should be assessed regardless.

For participants having a joint injection(s)/surgical joint procedure(s) ***prior to the date of study entry*** (e.g., randomization), the affected joints will be analyzed according to the impact of the joint injection and/or surgical joint procedure on the evaluability of the involved joints. If a joint is considered not evaluable at baseline due to certain procedure/injection performed prior to the date of randomization, the joint will be considered not evaluable throughout the study.

Table 9 - Joint Adjustment Rule in Case of Joint Injection/Surgical Joint Procedure During the Study

Part 1: For participants undergoing joint injection or joint procedure for the treatment of PsA	
Event	**Evaluability rule**
Joint injection	- For swollen/tender joint count, the affected joint(s) will be considered as swollen/tender from the date of injection for the next 90 days
Surgical joint procedure	- For swollen/tender joint count, the affected joint(s) will be considered as swollen/tender from the date of procedure onwards
Part 2: For participants undergoing joint injection or joint procedure for the treatment of non-PsA disease	
Event	**Evaluability rule**
Joint injection	- For swollen and tender joint count, the affected joint(s) will be considered as not evaluable from the date of injection for the next 90 days
Surgical joint procedure	- For swollen and tender joint count, the affected joint(s) will be analysed according to the impact of the surgical joint procedure on the evaluability of the involved joints

Table 10 – MRI Endpoint Adjustment Rule When Participant having a Joint Injection/Surgical Procedure

Part 1: For participants undergoing joint injection or joint procedure for the treatment of PsA	
Event	**Evaluability rule**
Joint injection	- For inflammatory features, the affected joint(s) will be considered as worst inflammatory level onwards for 90 days - For structural features, the affected joint(s) will be using the observed data as it is
Surgical joint procedure	- For both inflammatory and structural features, the affected joint(s) will be considered as not evaluable from the date of procedure onwards
Part 2: For participants undergoing joint injection or joint procedure for the treatment of non-PsA disease	
Event	**Evaluability rule**
Joint injection	- For both inflammatory and structural features, the affected joint(s) will be considered as not evaluable from the date of injection for the next 90 days
Surgical joint procedure	- For both inflammatory and structural features, the affected joint(s) will be considered as not evaluable from the date of procedure onwards

2. Joint Count Adjustment Rule

For participants who have an incomplete set of evaluable joints, the joint count/score will be adjusted to the total number joints of interest (e.g., 68 joints for tenderness and 66 joints for swelling) by dividing the number of affected joints by the number of evaluable joints and multiplying by the total number joints of interest.

3. LLOQ rule

Any value < LLOQ is considered equal to half of the value of LLOQ for numerical calculations.

4. Smallest Detectable Change

Smallest Detectable Change (SDC) is the smallest change in a score that is assessed correctly based on the limits of agreement (i.e., above the measurement error). The SDC in score of interest is determined as follows: $SDC = 1.96 * \frac{sd}{\sqrt{2k}}$ (Bruynesteyn, et al., 2005), where sd is the standard deviation of the difference between the two readers in change from baseline in the score of interest and k is the number of readers.

6.3. Appendix 3 Summary of multiple imputation methods

In case of multiple imputation will be implemented, following steps should be followed:

1) Data after ICE 1-3 are not used in the imputation model, i.e., data collected after ICEs are not contributing to the MI, it will be treated as missing instead and get imputed with MI.
2) Carry out MI as specified in Table 17. For each of the multiply imputed datasets, determine the endpoint value. In the case of composite endpoint (e.g., MDA, ACR), calculate the composite endpoint value)
3) After the imputation, in the case of ICE 1-3, apply ICE rule, i.e., composite estimand.
4) Each of the imputed dataset will be analyzed using the analysis method specified (logistic for binary, ANCOVA for continuous). The result from each of the imputed dataset will then be combined to produce inferential results based on Rubin's rule.

Binary endpoint

For binary endpoint, if it is a composite binary endpoint, MI will be performed on each component that have missing data, and then the missing composite binary endpoint will be determined based on (imputed) components data. If it is not composite binary endpoint, then the imputation will be performed on the continuous score first (continuous endpoint imputation) and then the binary endpoint will be determined.

Continuous endpoint

For continuous endpoint imputation, imputation will be performed on continuous scale, and it will be restricted to only impute within its possible range of values (e.g., PASI score may only be imputed to be between 0 and 72). The missing data will be imputed using the predicted value from an imputation model using the full conditional specification regression method for any missing pattern.

Table 17 - Multiple Imputation Method Summary

Endpoint [a, b]	Imputation Variable [c]	Data [c]/Ancillary Variables
MDA	BSA	- MIdata_BSA (N=100, Seed=11235) - Ancillary variables: treatment group, baseline DMARDs use, TJC68, PAIN and PASI
MDA	PGADAPA	- MIdata_PGADAPA (N=100, Seed=12358) - Ancillary variables: treatment group, baseline DMARDs use, TJC68, PAIN and PASI
MDA/ACR (20/50/70)	TJC68	- MIdata_TJC68 (N=100, Seed=13471) - Ancillary variables: treatment group, baseline DMARDs use, PAIN and PASI
MDA/ACR (20/50/70)	SJC66	- MIdata_SJC66 (N=100, Seed=14594) - Ancillary variables: treatment group, baseline DMARDs use, TJC68, PAIN and PASI
MDA/ACR (20/50/70)	PAIN	- MIdata_PAIN (N=100, Seed=15617) - Ancillary variables: treatment group, baseline DMARDs use, TJC68 and PASI
MDA/ACR (20/50/70) /Change from baseline in HAQ-DI	HAQ-DI	- MIdata_HAQ (N=100, Seed=16730) - Ancillary variables: treatment group, baseline DMARDs use, TJC68, PAIN and PASI
MDA/Resolution of Enthesitis (LEI) and change from baseline in Enthesitis score (LEI) among participants with Enthesitis at baseline	Enthesitis (LEI)	- MIdata_LEI (N=100, Seed=17853) - Ancillary variables: treatment group, baseline DMARDs use, TJC68, PAIN and PASI
MDA/PASI90/PASI100/ Change from baseline in PASI score among participants with ≥3% BSA psoriatic involvement and an IGA score of ≥2 at baseline	PASI	- MIdata_PASI (N=100, Seed=18976) - Ancillary variables: treatment group, baseline DMARDs use, TJC68, and PAIN
ACR (20/50/70)	PGADAA	- MIdata_PGADAA (N=100, Seed=21347) - Ancillary variables: treatment group, baseline DMARDs use, TJC68, PAIN and PASI
ACR (20/50/70)	PGAD	- MIdata_PGAD (N=100, Seed=22460) - Ancillary variables: treatment group, baseline DMARDs use, TJC68, PAIN and PASI
ACR (20/50/70)	CRP	- MIdata_CRP (N=100, Seed=23583) - Ancillary variables: treatment group, baseline DMARDs use, TJC68, PAIN and PASI

Table 17 - Multiple Imputation Method Summary

Endpoint [a, b]	Imputation Variable [c]	Data [e]/Ancillary Variables
IGA-psoriasis, among participants with ≥3% BSA psoriatic involvement and an IGA score of ≥2 at baseline	IGA	- MIdata_IGA (N=100, Seed=31459) - Ancillary variables: treatment group, baseline DMARDs use, TJC68, PAIN and PASI
Resolution of Dactylitis and change from baseline in Dactylitis score among participants with Dactylitis at baseline	Dactylitis	- MIdata_DACTYLITIS (N=100, Seed=32572) - Ancillary variables: treatment group, baseline DMARDs use, TJC68, PAIN and PASI
Change from baseline in Enthesitis score (SPARCC) among participants with Enthesitis at baseline	Enthesitis (SPARCC)	- MIdata_SPARCC (N=100, Seed=33695) - Ancillary variables: treatment group, baseline DMARDs use, TJC68, PAIN and PASI
Change from baseline in SF-36 PCS	SF-36 PCS	- MIdata_PCS (N=100, Seed=34718) - Ancillary variables: treatment group, baseline DMARDs use, TJC68, PAIN and PASI
Change from baseline in SF-36 MCS	SF-36 MCS	- MIdata_MCS (N=100, Seed=35831) - Ancillary variables: treatment group, baseline DMARDs use, TJC68, PAIN and PASI
Change from baseline in FACIT-fatigue	FACIT-Fatigue	- MIdata_FACIT (N=100, Seed=36954) - Ancillary variables: treatment group, baseline DMARDs use, TJC68, PAIN and PASI
Change from baseline in PROMIS-29	PROMIS-29	- MIdata_PROMIS (N=100, Seed=37077) - Ancillary variables: treatment group, baseline DMARDs use, TJC68, PAIN and PASI

[a] For categorical endpoints (Enthesitis resolution [LEI or SPARCC], dactylitis resolution, and IGA-psoriasis response), the imputation will be performed on the continuous scale, and then rounded to the nearest integer prior to deriving the response (resolution or response).

[b] For composite binary endpoint (MDA, ACR), the imputation will be performed on each component with missing data and then response status will be determined based on imputed component result.

[c] All the imputation variables are by visit through week 24, unless specified otherwise.

[e] The starting seed for FCS regression MI is used to generate a series of imputation seeds using the algorithm: *INT((2**31-2)*RANUNI(starting seed))*[SAS code], where each imputation seed will be used for a single imputation. To account for the possibility that some imputations may fail to complete due to out-of-range issues, 100+ initial imputation seeds will be prepared, and the first 100 successful imputations will be used for analysis.

6.4. Appendix 4 Changes to Protocol-Planned Analyses

Not applicable

6.5. Appendix 5 Demographics and Baseline Characteristics

The number of participants in each analysis set will be summarized and listed by intervention group and overall. In addition, the distribution of participants by region, country, and site ID will be presented unless otherwise noted. No format statistical comparison is planned. P-values will not be provided.

Demographic variables to be summarized will include age, weight, height, BMI, sex, race, ethnicity, BMI (see table below). Baseline characteristics variables to be summarized will include, but not be limited to, baseline disease characteristics of PsA (e.g., duration of disease, PsA subtype, baseline efficacy assessment), medical history, prior exposure to non-biologic medication, and baseline medication usage for PsA. And also the summary of previous medication and therapies for Psoriasis.

In addition, balance between randomized treatment groups will also be explored for the subpopulation used for analyses below: (1) participants with baseline ≥3% body surface area (BSA) psoriatic involvement and an IGA score of ≥2 (mild) at baseline; (2) participants with Enthesitis at baseline; (2) participants with dactylitis at baseline.

Table below presents a list of the demographic variables that will be summarized by intervention group and overall, for the FAS.

Continuous Variables:	**Summary Type**
Age ([years])	Descriptive statistics (N, mean, standard deviation [SD], median and range [minimum and maximum], and IQ range).
Weight (kg)	
Height (cm)	
Body Mass Index (BMI) (kg/m^2)	
Categorical Variables	
Age ([18-25 years, 26-50 years, 51-64 years, and >=65 years])	Frequency distribution with the number and percentage of participants in each category.
Sex (male, female, unknown, undifferentiated)	
Race [a] (American Indian or Alaska Native, Asian, Black, or African American, Native Hawaiian or Other Pacific Islander, White, Not Reported, Unknown)	
Ethnicity (Hispanic or Latino, not Hispanic or Latino)	
BMI ([underweight <18.5 kg/m^2, normal 18.5-<25 kg/m^2, overweight 25-<30 kg/m^2, obese >=30 kg/m^2])	

[an] If multiple race categories are indicated, the Race is recorded as 'Multiple'

6.6. Appendix 6 Protocol Deviations

In general, the following list of major protocol deviations may have the potential to impact participants' rights, safety or well-being, or the integrity and/or result of the clinical study. Participants with major protocol deviations will be identified prior to database lock and the participants with major protocol deviations will be summarized by category.

Developed withdrawal criteria but not withdrawn
Entered but did not satisfy inclusion/exclusion criteria
Received a not allowed concomitant treatment
Received wrong treatment or incorrect dose
Other

6.7. Appendix 7 Prior and Concomitant Medications

Prior and Concomitant medications will be coded using the World Health Organization Drug Dictionary (WHO-DD). Prior medications are defined as any therapy used before the day of first dose (partial or complete) of study intervention. Concomitant medications are defined as any therapy used on or after the same day as the first dose of study intervention, including those that started before and continue on after the first dose of study intervention.

Summaries of PsA specific concomitant medications will be presented by anatomic therapeutic chemical (ATC) classification term, intervention group. The proportion of participants who receive each concomitant medication will be summarized as well as the proportion of participants who receive at least 1 concomitant medication for PsA. In addition, concomitant medications of special interest will be presented. See Appendix 13 for list of medications in each category.

Prior medications/therapies that taken for PsA will be summarized by intervention group and ATC term. The medications include non-biologic DMARDs (MTX, SSZ, LEF, HCQ, Chloroquine, Gold preparations, Penicillamine and other), immunosuppressive (Cyclosporine, Azathioprine, Tacrolimus, Mycophenolate mofetil, and other), systemic corticosteroids, NSAIDs, and Apremilast.

6.8. Appendix 8 Medical History

Number and percentage of participants who had medical histories of interest for PsA will be collected and summarized by intervention group. For the list of medical history, it can be identified on the medical history of interest form.

6.9. Appendix 9 Intervention Compliance

Compliance will be summarized descriptively for each study agent within a study intervention. Compliance to randomized intervention versus actual intervention will be presented in a summary table. Since the study has a combination intervention, therefore the 'study agent' refer to the individual components included in the combination intervention.

Compliance (%) will be calculated as (actual number of injection received/total number of injections planned) x 100.

6.10. Appendix 10 Adverse Events of Special Interest

Adverse Events of special interest are defined as any newly identified malignancy or case of active Tuberculosis occurring after the first study intervention administration. Case(s) are identified by check boxes in the eCRF.

6.11. Appendix 11 Medications of Special Interest

The medication of special interest is methotrexate.

6.12. Appendix 12 Laboratory Toxicity Grading

The grading scale use for lab assessments is based on 'Common Terminology Criteria for Adverse Events (CTCAE) v5.0'.

If a laboratory value falls within the grading as specified below but also within the local laboratory normal limits, the value is considered to be normal and will be reset to grade 0.

Pre-baseline measurements will use the same grading ranges as applied to baseline measurements. In case a test has two sets of ranges – one for baseline normal and one for baseline abnormal, the one for baseline normal will be applied for all measurements taken pre-baseline and on baseline.

Table 18 - Grading Criteria for Clinical Laboratory Tests [CTCAE Version 5.0]

Hematology Tests		**Criteria**			
Test	**Direction**	**1**	**2**	**3**	**4**
Hemoglobin (g/dL)	Increase	>0 - 2 x ULN	>2 - 4 x ULN	>4 x ULN	
Hemoglobin (g/dL)	Decrease	<LLN - 10.0	<10.0 - 8.0	<8.0	
Lymphocytes (/mm3)	Increase		>4000 - 20,000	>20,000	
Lymphocytes (/mm3)	Decrease	<LLN - 800	<800 - 500	<500 - 200	<200
Neutrophils (/mm3)	Decrease	<LLN - 1500	<1500 - 1000	<1000 - 500	<500
Platelets (/mm3)	Decrease	<LLN - 75,000	<75,000 - 50,000	<50,000 - 25,000	<25,000
Total WBC count (/mm3)	Increase			>100,000	
Total WBC count (/mm3)	Decrease	<LLN - 3000	<3000 - 2000	<2000 - 1000	<1000
Chemistry Tests		**Criteria**			
Test	**Direction**	**1**	**2**	**3**	**4**
ALT	Increase	>ULN - 3.0 x ULN if baseline was normal; 1.5 - 3.0 x baseline if baseline was abnormal	>3.0 - 5.0 x ULN if baseline was normal; >3.0 - 5.0 x baseline if baseline was abnormal	>5.0 - 20.0 x ULN if baseline was normal; >5.0 - 20.0 x baseline if baseline was abnormal	>20.0 x ULN if baseline was normal; >20.0 x baseline if baseline was abnormal
AST	Increase	>ULN - 3.0 x ULN if baseline was normal; 1.5 - 3.0 x baseline if baseline was abnormal	>3.0 - 5.0 x ULN if baseline was normal; >3.0 - 5.0 x baseline if baseline was abnormal	>5.0 - 20.0 x ULN if baseline was normal; >5.0 - 20.0 x baseline if baseline was abnormal	>20.0 x ULN if baseline was normal; >20.0 x baseline if baseline was abnormal
Albumin (g/L)	Decrease	≥30 - <LLN	≥20 - <30	<20	
Alkaline Phosphatase	Increase	>ULN - 2.5 x ULN if baseline was normal; 2.0 - 2.5 x baseline if baseline was abnormal	>2.5 - 5.0 x ULN if baseline was normal; >2.5 - 5.0 x baseline if baseline was abnormal	>5.0 - 20.0 x ULN if baseline was normal; >5.0 - 20.0 x baseline if baseline was abnormal	>20.0 x ULN if baseline was normal; >20.0 x baseline if baseline was abnormal
Bilirubin (total)	Increase	>ULN - 1.5 x ULN if baseline was normal; > 1.0 - 1.5 x baseline if baseline was abnormal	>1.5 - 3.0 x ULN if baseline was normal; >1.5 - 3.0 x baseline if baseline was abnormal	>3.0 - 10.0 x ULN if baseline was normal; >3.0 - 10.0 x baseline if baseline was abnormal	>10.0 x ULN if baseline was normal; >10.0 x baseline if baseline was abnormal
Corrected Calcium (mmol/L)	Increase	>ULN - ≤2.9	>2.9 - ≤3.1	>3.1 - ≤3.4	>3.4
Corrected Calcium (mmol/L)	Decrease	≥2.0 - <LLN	<2.0 - ≥1.75	<1.75 - ≥1.5	<1.5
Creatinine	Increase	>ULN - ≤1.5 x ULN	>1.5 - 3.0 x baseline; >1.5 - 3.0 x ULN	>3.0 x baseline; >3.0 - 6.0 x ULN	>6.0 x ULN
Glucose (mmol/L)	Decrease	<LLN - 3.0	<3.0 - 2.2	<2.2 - 1.7	<1.7
Potassium (mmol/L)	Increase	>ULN - ≤5.5	>5.5 - 6.0	>6.0 - 7.0	>7.0
Potassium (mmol/L)	Decrease	<LLN - 3.0		<3.0 - 2.5	<2.5
Sodium (mmol/L)	Increase	>ULN - 150	>150 - 155	>155 - 160	>160
Sodium (mmol/L)	Decrease	<LLN - 130	129 - 125	124 - 120	<120

7. REFERENCES

Abrar, D. et al., 2020. Introduction of a Simplified Psoriatic Arthritis Magnetic Resonance Imaging Score (sPsAMRIS): A Potential Tool for Treatment Monitoring in Peripheral Psoriatic Arthritis. *Diagnostics,* Volume 10, p. 1093.

Boyesen, P. et al., 2011. The OMERACT Psoriatic Arthritis Magnetic Resonance Imaging Score (PsAMRIS) is reliable and sensitive to change: results from an OMERACT workshop. *The Journal of Rheumatology,* 38(9), pp. 2034-8.

Bruynesteyn, K. et al., 2005. Deciding on progression of joint damage in paired films of individual participants: smallest detectable difference or change. *Annals of the Rheumatic Diseases,* 64(2), pp. 179-82.

Coates, L. C., Fransen, J. & Helliwell, P. S., 2010. Defining minimal disease activity in psoriatic arthritis: a proposed objective target for treatment. *Annals of Rheumatic Disease,* pp. 48-53.

EMA, 2015. *Guideline on adjustment for baseline covariates in clinical trials,* London: Committee for Medicinal Products for Human Use.

Genovese, M. C. et al., 2018. Safety and efficacy of ixekizumab in patients with PsA and previous inadequate response to TNF inhibitors: week 52 results from SPIRIT-P2. *Rheumatology,* p. 57.

Gladman, D. D. et al., 2007. International spondyloarthritis interobserver reliability exercise--the INSPIRE study: II. Assessment of peripheral joints, enthesitis, and dactylitis. *The Journal of Rheumatology,* 34(8), p. 1740.

Gladman, D. D., Ziouzina, O., Thavaneswaran, A. & Chandran, V., 2013. Dactylitis in psoriatic arthritis: prevalence and response to therapy in the biologic era. *The Journal of Rheumatology,* 40(8), p. 1357.

Glinatsi, D. et al., 2015. Validation of the OMERACT Psoriatic Arthritis Magnetic Resonance Imaging Score (PsAMRIS) for Hand and Foot in a Randomized Placebo-Controlled Trial. *The Journal of Rheumatology,* 42(12), pp. 2473-9.

Gossec, L. et al., 2018. Minimal Disease Activity as a Treatment Target in Psoriatic Arthritis: A Review of the Literature. *The Journal of Rheumatology,* p. 45.

Gossec, L. et al., 2018. Minimal Disease Activity as a Treatment Target in Psoriatic Arthritis: A Review of the Literature. *The Journal of Rheumatology,* pp. 45-1.

Han, B., Yu, M. & McEntegart, D., 2013. Weighted re-randomization tests for minimization with unbalanced allocation. *Pharmaceutical Statistics,* 12(5), pp. 243-253.

Hays, R. D., Spritzer, K. L., Schalet, B. D. & Cella, D., 2018. PROMIS®-29 v2.0 Profile Physical and Mental Health Summary Scores. *Quality of Life Research,* pp. 1885-1891.

Hays, R. D., Spritzer, K. L., Schalet, B. D. & Cella, D., 2018. PROMIS-29 v2.0 Profile Physical and Mental Health Summary Scores. *Quality of Life Research,* 27(7), p. 1885.

Healy, P. J. & Helliwell, P. S., 2008. Measuring Clinical Enthesitis in Psoriatic Arthritis: Assessment of Existing Measures and Developmemntn of an Instrument Specific to Psoriatic Arthritis. *Arthritis & Rheumatism,* 59(5), pp. 686-691.

Koo, T. K. & Li, M. Y., 2016. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine,* Volume 15, pp. 155-163.

Langley, R. & Ellis, C., 2004. Evaluating psoriasis with Psoriasis Area and Severity Index, Psoriasis Global Assessment, and Lattice System Physician's Global Assessment. *Journal of American Academy of Dermatology,* pp. 563-9.

Mathew, A. et al., 2019. The OMERACT MRI in Enthesitis Initiative: Definitions of Key Pathologies, Suggested MRI Sequences, and a Novel Heel Enthesitis Scoring System. *The Journal of Theumatology,* 46(9), pp. 1232-8.
Mathew, A. et al., 2020. Atlas of the OMERACT Heel Enthesitis MRI Scording System (HEMRIS). *RMD Open,* 6(e001150), p. 1136.
McQueen, F. et al., 2009. Testing an OMERACT MRI scoring system for peripheral psoriatic arthritis in cross-sectional and longitudinal settings. *The Journal of Rheumatology,* 36(8), pp. 1811-5.
Mease, J. P. et al., 2016. Minimally Important Difference of Health Assessment Questionnaire in Psoriatic Arthritis: Relating Thresholds of Improvement in Functional Ability to Patient-related Importance and Satisfaction. *The Journal of Rheumatology,* 38(11), p. 2461.
Mease, P. et al., 2016. Minimally Important Difference of Health Assessment Questionnaire in Psoriatic Arthritis: Relating Thresholds of Improvement in Functional Ability to Patient-rated Importance and Satisfaction. *The Journal of Rheynatology,* p. 2461.
Ostergaard, M. et al., 2009. The OMERACT psoriatic arthritis magnetic resonance imaging scoring system (PsAMRIS): definitions of key pathologies, suggested MRI sequences, and preliminary scoring system for PsA Hands. *The Journal of Rheumatology,* 36(8), pp. 1816-24.