Protocol Number: 22104

Official Title: A Randomized, Double-blind, Placebo-controlled, Multicenter Phase 2b/3 Study to Evaluate the Efficacy and Safety of Izokibep in Subjects with Active Psoriatic Arthritis

NCT Number: NCT05623345

Document Date: 09 January 2024

Statistical Analysis Plan

A Randomized, Double-blind, Placebo-controlled, Multicenter Phase 2b/3 Study to Evaluate the Efficacy and Safety of Izokibep in Subjects with Active Psoriatic Arthritis

Statistical Analysis Plan: Final

Statistical Analysis Plan: 2.0
Statistical Analysis Plan Date: 09JAN2024

Investigational Product: Izokibep

Protocol Reference: 22104

[redacted]

Sponsor: ACELYRIN, INC.

Author:

[redacted]

Senior Principal Biostatistician

Table of Contents

List of Tables

Glossary of Abbreviations

Abbreviation	Term
ACR	American College of Rheumatology
ACR20	American College of Rheumatology 20%
ACR50	American College of Rheumatology 50%
ACR70	American College of Rheumatology 70%
ADA	Anti-Drug Antibody
AE	Adverse Event
AUC	Area under the curve
BSA	Body Surface Area
cDAPSA	Clinical Disease Activity In Psoriatic Arthritis
CRP	C-Reactive Protein
CV	Coefficient Of Variation
csDMARD	Conventional-Synthetic Disease-Modifying Anti-Rheumatic Drugs
DAPSA	Disease Activity In Psoriatic Arthritis
DAS28	Disease Activity Score In 28 Joints
DILI	Drug Induced Liver Injury
DMARD	Disease-Modifying Antirheumatic Drug
DMC	Data Monitoring Committee
ECG	Electrocardiogram
EOT	End of Treatment
eCRF	Electronic Case Report Form
FACIT-F	Functional Assessment of Chronic Illness Therapy – Fatigue
FAS	Full Analysis Set
HADS	Hospital Anxiety And Depression Scale
HAQ-DI	Health Assessment Questionnaire – Disability Index
ICF	Informed Consent Form
ICH	International Council For Harmonisation
IXRS	Interactive Response Technology
JAK	Janus Kinase
LDI	Leeds Dactylitis Index
LEI	Leeds Enthesitis Index
MCS	Mental Component Summary
MDA	Minimal Disease Activity
mNAPSI	Modified Nail Psoriasis Severity Index

Abbreviation	Term
mTSS	Modified Total Sharp Score
NCI-CTCAE	National Cancer Institute – Common Terminology Criteria for Adverse Events
NRS	Numeric Rating Scale
NSAID	Nonsteroidal Anti-Inflammatory Drug
PASDAS	Psoriatic Disease Activity Score
PASI	Psoriasis Area And Severity Index
PASI75	Psoriasis Area And Severity Index Response of 75%
PASI90	Psoriasis Area And Severity Index Response of 90%
PASI100	Psoriasis Area And Severity Index Response of 100%
PCS	Physical Component Summary
PK	Pharmacokinetic(S)
PsA	Psoriatic Arthritis
PsAID	Psoriatic Arthritis Impact of Disease
PT	Preferred Term
QW	Every Week
Q2W	Every 2 Weeks
Q4W	Every 4 Weeks
SAE	Serious Adverse Event
SC	Subcutaneous (Ly)
SF-36	Short Form-36v2 Health Survey
SoA	Schedule of Activities
SOC	System Organ Class
SPARCC	Spondyloarthritis Research Consortium of Canada
TB	Tuberculosis
TEAE	Treatment-Emergent Adverse Event
TFL	Tables/Figures/Listings
TNF	Tumor Necrosis Factor-Alpha
TNFi	Tumor Necrosis Factor-Alpha Inhibitor(S)
US	United States
VAS	Visual Analogue Scale
VLDA	Very Low Disease Activity
WPI	Widespread Pain Index

1. Source Documents

The SAP was written based on the following documentation:

Document	Date	Version
Protocol	08JAN2024	5.0
eCRF	08APR2023	2.1

2. Protocol Details

2.1. Overall Study Design

This is a multicenter, randomized, double-blind, placebo-controlled, Phase 2b/3 Study to evaluate the clinical question of the effects of the efficacy and safety of izokibep in subjects with active psoriatic arthritis (PsA) who are tumor necrosis factor-alpha inhibitor (TNFi) naïve, who have had an inadequate response or intolerance to TNFi, or for whom TNFi is contraindicated. Approximately 325 subjects with active PsA will be enrolled. The study consists of up to a 28-day (4-week) screening period, a 51-week treatment period, and a follow-up period with visits at 8 weeks and 14 weeks after the last dose of study drug to assess safety and immunogenicity.

Subjects meeting eligibility criteria will be randomized into 1 of 4 treatment groups in a 4:4:4:1 ratio as follows:

- Group 1 (n=100): placebo every week (QW) from Day 1/Week 0 to Week 15, then izokibep 160 mg QW from Week 16 to Week 51.
- Group 2 (n=100): izokibep 160 mg QW from Day 1/Week 0 to Week 51.
- Group 3 (n=100): izokibep 160 mg every 2 weeks (Q2W) from Day 1/Week 0 to Week 50, with matching placebo Q2W until Week 51 for the weeks in between izokibep doses to maintain the blind.
- Group 4 (n=25): izokibep 80 mg every 4 weeks (Q4W) from Day 1/Week 0 to Week 48, with matching placebo QW until Week 51 for the weeks in between izokibep doses to maintain the blind.

Randomization will be stratified by prior TNFi use (Yes/No) and enthesitis (Leeds Enthesitis Index [LEI] > 0/LEI = 0). Anticipated enrollment of subjects previously treated with TNFi is 20% to 30%. The number of subjects without enthesitis entering into the study (LEI = 0 at baseline) will be capped at 70%.

Subjects will complete study assessments according to the study visits outlined in the Schedule of Activities (SoA) in the protocol. The primary endpoint will be assessed at Week 16. The last dose of study drug will be administered on Week 51.

Starting at Week 24, subjects who do not meet clinical Disease Activity for PsA (cDAPSA) low disease activity (cDAPSA ≤14) at 2 consecutive visits (e.g. at Week 20 and Week 24) will be discontinued from study drug and can receive standard of care as deemed appropriate by the investigator.

An End of Treatment (EOT) visit will be conducted at Week 52. A Safety Follow-up visit will be conducted at Week 59. A pharmacokinetic/anti-drug antibody (PK/ADA) Follow-up and

End of Study Visit will be conducted at Week 65. For subjects that prematurely discontinue study drug for any reason, the EOT visit should be completed within 14 days of withdrawal, Safety follow-up visit should be completed at 8 weeks and PK/ADA follow-up visit should be completed 14 weeks after the last dose of study drug, respectively, where possible. Subjects should continue to complete study assessments as outlined in the SoA (Section 1.3 of 22104 protocol) where possible, with the exception of study drug administration and anti-drug antibody sample and PK sample collection.

The final analysis of primary and secondary endpoints will be conducted after the last subject has had the opportunity to complete Week 16 assessments or prematurely discontinues the study drug and/or study.

Before database lock, an error was discovered in dosing for subjects who were randomized to receive izokibep 160 mg every 2 weeks (Q2W) from Day 1/Week 0 to Week 50, with matching placebo Q2W and subjects who were randomized to receive izokibep 80 mg every 4 weeks (Q4W) from Day 1/Week 0 to Week 48, with matching placebo QW until Week 51. Subjects were assigned study drug in groups of four visits at a time. The order in which the subject was instructed to receive the drug was incorrect. Therefore, each subject was assigned the correct amount of study drug over each 4-week period, but possibly in the wrong sequence, resulting in subjects who were randomized to receive active drug Q2W receiving active drug on two or more consecutive weeks and/or placebo at two or more consecutive weeks. Subjects who were randomized to receive active drug Q4W similarly received active drug at two or more consecutive weeks and/or placebo at four or more consecutive weeks. As a result of this error, the following changes were made to the analysis plan:

- Testing of all primary and secondary endpoints, comparing izokibep Q2W versus placebo, will only be done after testing of all primary and secondary endpoints, comparing izokibep QW versus placebo, in the alpha-controlled testing sequence.
    - Justification: subjects randomized to receive izokibep Q2W received the correct amount of study drug, but in incorrect order. Some quantities will be as planned (total drug received, possibly area under the concentration-time curve) while others may not (C_{trough}, C_{max}, possibly area under the concentration-time curve). Therefore the group that was treated as planned will be tested first, and the group treated not as planned will be tested second.
- Exploratory analyses of the impact of incorrect dosing will be evaluated during Period 1.
    - Subjects randomized to receive izokibep Q2W will be grouped according to the largest number of consecutive doses of active drug received to understand the impact of various dosing schedules on outcome.

- Summaries of subjects dosed correctly at earlier time points (week 4 and 12 for the primary endpoint).

2.2. Study Objectives

2.2.1. Primary Objective

The primary objective of the study is to demonstrate that 1 or both regimens of izokibep (160 mg QW and 160 mg Q2W) are efficacious compared to placebo, as measured by the proportion of subjects achieving 50% improvement in ACR core set measurements (ACR50) at Week 16.

2.2.2. Secondary Objectives

The secondary objectives of the study are:

- To demonstrate that 1 or both regimens of izokibep are efficacious (160 mg QW and 160 mg Q2W) compared to placebo, as measured by:
    - Proportion of subjects achieving 90% or greater reduction in PASI score from baseline (PASI90) at Week 16 in subjects with ≥3% BSA psoriasis at baseline
    - Proportion of subjects with resolution of enthesitis at Week 16 as assessed by LEI in subpopulation that had enthesitis (LEI >0) at baseline
    - Proportion of subjects achieving minimal disease activity (MDA) at Week 16
    - Proportion of subjects achieving 20% improvement in ACR core set measurements (ACR20) at Week 16
    - Proportion of subjects achieving an improvement in PsAID of at least 3 units at Week 16 compared to baseline in subjects with PsAID ≥3 at baseline
    - Change in physical function as assessed by HAQ-DI change from baseline to Week 16
- To assess the safety and tolerability of izokibep as measured by the incidence of Treatment Emergent Adverse Events (TEAEs), events of interest, Serious Adverse Events (SAEs), and clinically significant laboratory values and vital signs
- To assess the immunogenicity of izokibep as measured by the presence of ADAs

2.2.3. Exploratory Objectives

The exploratory objectives of the study are:

- To demonstrate that 1 or both regimens of izokibep (160 mg QW and 160 mg Q2W) are efficacious compared to placebo, as measured by:
    - Proportion of subjects achieving 70% improvement in ACR core set measurements (ACR70) at Weeks 16, 24, and 52
    - Change in DAS28-CRP at Weeks 16, 24, and 52 as compared to baseline

- Change in DAPSA score at Weeks 16, 24, and 52 as compared to baseline
- Change in clinical DAPSA (cDAPSA) score at Week 16, 24, and 52, as compared to baseline
- Proportion of subjects achieving DAPSA and cDAPSA low disease activity or remission at Week 16, 24, and 52
- Proportion of subjects achieving DAPSA and cDAPSA remission at Week 16, 24, and 52
- Change in PASDAS at Weeks 16, 24, and 52 as compared to baseline
- Proportion of subjects achieving 75% or greater reduction in PASI score from baseline (PASI75) at Weeks 16, 24, and 52 in subjects with ≥3% BSA psoriasis at baseline
- Proportion of subjects achieving PASI90 at Weeks 24 and 52 in subjects with ≥3% BSA psoriasis at baseline
- Proportion of subjects achieving 100% reduction in PASI score from baseline (PASI100) at Weeks 16, 24, and 52 in subjects with ≥3% BSA psoriasis at baseline
- Proportion of subjects achieving 75% or greater reduction in PASI score from baseline (PASI75) at Weeks 16, 24, and 52 in subjects with ≥ 10% BSA psoriasis at baseline
- Proportion of subjects achieving PASI90 at Weeks 16, 24 and 52 in subjects with ≥10% BSA psoriasis at baseline
- Proportion of subjects achieving 100% reduction in PASI score from baseline (PASI100) at Weeks 16, 24, and 52 in subjects with ≥ 10% BSA psoriasis at baseline
- Proportion of subjects achieving PASI90 at weeks 16, 24, and 52 in subjects with baseline PASI score ≥10
- Proportion of subjects achieving PASI100 at weeks 16, 24, and 52 in subjects with baseline PASI score ≥10
- Proportion of subjects achieving MDA at Weeks 24, and 52
- Proportion of subjects achieving VLDA response at Weeks 16, 24, and 52
- Proportion of subjects with resolution of dactylitis at Weeks 16, 24, and 52 in subpopulation that had dactylitis at baseline (pooled izokibep doses), as assessed with Leeds Dactylitis Index (LDI)
- Change from baseline in mNAPSI at Weeks 16, 24, and 52 in subpopulation with nail psoriasis at baseline (pooled izokibep doses)
- Progression of structural damage assessed radiographically and expressed as the change in mTSS at Weeks 16 and 52, compared to baseline
- Change in SPARCC enthesitis score at Weeks 16, 24, and 52, as compared to baseline in subpopulation with enthesitis (SPARCC >0) at baseline

- Change in spinal pain NRS at Weeks 16, 24, and 52, as compared to baseline in subjects reporting spinal pain at baseline
- Change from baseline in spine pain NRS in those with sacroiliitis on pelvic x-ray at Weeks 16, 24 and 52
- Change from baseline in spine pain NRS in those with investigator assessed spinal involvement at Weeks 16, 24 and 52
- Association of investigator assessed sacroiliac involvement with sacroiliitis on x-ray
- Proportion of subjects achieving each MDA component (TJC68, SJC66, PASI, BSA, SAPVAS, SGADA, HAQ-DI, tender entheseal points from LEI) at Weeks 16, 24 and 56
- Change in quality of life as assessed by SF-36 total score, physical component summary (PCS), and mental component summary (MCS) change from baseline to Weeks 16, 24, and 52
- Fatigue as assessed by FACIT-F change from baseline to Weeks 16, 24, and 52
- Proportion of subjects achieving ACR50 at Weeks 24 and 52
- Proportion of subjects with resolution of enthesitis at Weeks 24 and 52 as assessed by LEI in subpopulation that had enthesitis (LEI >0) at baseline
- Proportion of subjects achieving an improvement in PsAID of at least 3 units at Weeks 24 and 52 compared to baseline in subjects with PsAID ≥3 at baseline
- Proportion of subjects achieving ACR20 at Weeks 24 and 52
- Mean (median) ACR-N at Week 16, where, ACR-N is defined as the largest number N such that ACR-N is met (analogous to the primary endpoint definition of ACR 50)
- Change in enthesitis at Weeks 16, 24 and 52 as assessed by change from baseline in LEI
- Time to resolution of enthesitis at weeks 4, 12 and 16 among subjects with enthesitis (LEI > 0) at baseline
- Occurrence of enthesitis (LEI > 0) at weeks 4, 12 and 16 among subjects without enthesitis (LEI = 0) at baseline
- Depression and anxiety as assessed by change in HADS score from baseline to Weeks 16 and 52
- Change in physical function as assessed by HAQ-DI change from baseline to Weeks 24 and 52
- Proportion of subjects with a decrease of at least 2 from baseline in LEI, among subjects with baseline LEI of at least 2, at weeks 4, 12 and 16
- Subject's Global Impression of Change Questionnaire at weeks 16 and 52
- Mean (median) Area under the curve (AUC) for ACR-N through week 16
- Change from baseline in LEI at Weeks 16, 24 and 52
- Time to first achievement of ACR50

- Time to first achievement of PASI90 among subjects with at least 3% BSA at baseline
- Time to first achievement of MDA
- Time to first achievement of ACR20
- Time to first achievement of PsAID improvement of at least 3 among subjects with PsAID at least 3 at baseline
- Time to first achievement of PASI100 among subjects with at least 3% BSA at baseline

- To investigate whether 1 or both regimens of izokibep (160 mg QW and 160 mg Q2W) are efficacious compared to placebo, in the subpopulation with enthesitis as baseline, as measured by:
  - Proportion of subjects achieving an improvement in PsAID of at least 3 units at Week 16 compared to baseline in subjects with PsAID ≥3 at baseline
  - Change from baseline in PsAID at Week 16
  - Proportion of subjects with a HAQ-DI decrease of at least 0.35 at Week 16 compared to baseline in subjects with HAQ-DI >0.35 at baseline (Response to HAQ-DI at Week 16)
  - Change in physical function as assessed by HAQ-DI change from baseline to Week 16
  - Proportion of subjects achieving MDA at Week 16
  - Change in DAPSA and cDAPSA at Week 16 as compared to baseline
  - Proportion of subjects achieving VLDA at Week 16
  - Change in quality of life as assessed by SF-36 total score, physical component summary (PCS), and mental component summary (MCS) change from baseline to Weeks 16
  - Proportion of subjects achieving individual MDA components at Week 16 as compared to baseline
- To investigate the relationship between ACR50 response and other factors
  - Association of baseline positive WPI ≥7 with ACR50 response
- To investigate that the 80 mg Q4W regimen is efficacious compared to placebo, as measured by the same primary, secondary and exploratory endpoints
- To compare the three regimens of izokibep for assessment of dose-response relationships, as measured by the primary and secondary endpoints
- To evaluate the pharmacokinetics of izokibep in subjects with PsA
- To evaluate the immunogenicity of izokibep in subjects with PsA

2.3. Sample Size and Power

Response rates for ACR50 are expected to be 20% for placebo and 45% for a dose regimen of 160 mg izokibep (QW or Q2W or both). With 100 subjects receiving placebo and 100 subjects receiving a dose regimen of 160 mg izokibep, there will be 97% power to show a difference in the primary endpoint of ACR50 response rate. If a 45% response rate is observed with a dose regimen of izokibep, the 2-sided 95% confidence interval for the true response rate will have a half-width of approximately 10 percentage points.

Up to 60% of subjects are expected to have enthesitis at baseline, based on the LEI. Response rates, proportion of subjects who have LEI = 0 at Week 16, are expected to be 25% for placebo and 60% for subjects receiving izokibep. With 50 subjects with enthesitis at baseline receiving placebo and 50 subjects receiving a dose regimen of izokibep, there will be 95% power to show a difference in the secondary endpoint of resolution of enthesitis. A lower rate of enthesitis at baseline will result in a lower power, but power will be at least 80% with the same assumptions on response rate, if 30 subjects with enthesitis at baseline receive placebo and 30 subjects receive a dose regimen of izokibep.

With 25 subjects receiving izokibep 80 mg Q4W, the 2-sided 95% confidence interval for the true response rate will have a half-width of up to approximately 20 percentage points.

2.4. Primary Efficacy Variable

The primary efficacy variable for the study is the ACR50 response rate, defined as the proportion of subjects who achieve a 50% improvement from baseline in ACR core set measurements. The primary timepoint for analysis is Week 16.

ACR50 criteria require that the subject experiences an improvement in the following parameters compared to baseline:

- ≥50% reduction in 68 Tender Joint Count (TJC68)
- ≥50% reduction in 66 Swollen Joint Count (SJC66)
- ≥50% reduction in 3 of 5 additional measures:
    - Subject's Assessment of Pain (SAPVAS)
    - Subject's Global Assessment of Disease Activity (SGADA)
    - Investigator's Global Assessment of Disease Activity (IGADA)
    - HAQ-DI
    - Acute phase reactant (C-reactive protein [CRP])

2.5. Secondary Efficacy Variables

The secondary efficacy variables of the study are:

- PASI90 at Week 16 in subjects with ≥3% BSA psoriasis at baseline
- Resolution of enthesitis (LEI ≥= 0) at Week 16
- MDA at Week 16
- ACR20 at Week 16
- PsAID response at Week 16 in subjects with PsAID ≥3 at baseline
- HAQ-DI change from baseline to Week 16

2.6. Exploratory Efficacy Variables

The exploratory efficacy variables of the study are:

- ACR70 at Weeks 16, 24, and 52
- Change in DAS28-CRP at Weeks 16, 24, and 52
- Change in DAPSA at Weeks 16, 24, and 52
- Change in cDAPSA at Weeks 16, 24, and 52
- DAPSA ≤14 at Weeks 16, 24, and 52
- cDAPSA ≤14 at Weeks 16, 24, and 52
- DAPSA ≤4 at Weeks 16, 24, and 52
- cDAPSA ≤4 at Weeks 16, 24, and 52
- Change in PASDAS at Weeks 16, 24, and 52
- PASI75 at Weeks 16, 24, and 52 in subjects with ≥3% BSA psoriasis at baseline
- PASI90 at Weeks 24 and 52 in subjects with ≥3% BSA psoriasis at baseline
- PASI100 at Weeks 16, 24, and 52 in subjects with ≥3% BSA psoriasis at baseline
- PASI75 at Weeks 16, 24, and 52 in subjects with ≥10% BSA psoriasis at baseline
- PASI90 at Weeks 16, 24 and 52 in subjects with ≥10% BSA psoriasis at baseline
- PASI100 at Weeks 16, 24, and 52 in subjects with ≥10% BSA psoriasis at baseline
- PASI90 at Weeks 16, 24, and 52 in subjects with baseline PASI score ≥10
- PASI100 at Weeks 16, 24, and 52 in subjects with baseline PASI score ≥10
- MDA at Weeks 24, and 52
- VLDA at Weeks 16, 24, and 52
- Resolution of dactylitis (LDI=0) at Weeks 16, 24, and 52 in subjects with LDI >0 at baseline
- Change in mNAPSI at Weeks 16, 24, and 52 in subjects with mNAPSI >0 at baseline
- Change in mTSS at Weeks 16 and 52
- Change in SPARCC enthesitis score at Weeks 16, 24, and 52 in subjects with SPARCC >0 at baseline
- Change in spinal pain NRS at Weeks 16, 24, and 52
- Change from baseline in spine pain NRS in those with sacroiliitis on pelvic x-ray at Weeks 16, 24 and 52

- Change from baseline in spine pain NRS in those with investigator assessed spinal involvement at Weeks 16, 24 and 52
- Correlation of investigator assessed spinal involvement with sacroiliitis on x-ray
- Proportion of subjects achieving MDA component (TJC68, SJC66, PASI, BSA, SAPVAS, SGADA, HAQ-DI, tender entheseal points from LEI) at Week 16
- Change in SF-36 total score, physical component summary (PCS), and mental component summary (MCS) at Weeks 16, 24, and 52
- Change in FACIT-F at Weeks 16, 24 and 52
- ACR50 at Weeks 24 and 52
- HAQ-DI change from baseline to Weeks 16, 24 and 52
- Resolution of enthesitis (LEI = 0) at Weeks 24 and 52
- PsAID response at Weeks 24 and 52 in subjects with PsAID ≥3 at baseline
- ACR20 at Weeks 24 and 52
- ACR-N at week 16
- Change from baseline in LEI at Week 16, 24, 52
- Time to resolution of enthesitis at weeks 4, 12 and 16 among subjects with enthesitis (LEI > 0) at baseline
- Occurrence of enthesitis (LEI > 0) at week 4, 12 and 16 among subjects without enthesitis (LEI = 0) at baseline
- HADS change from baseline to Weeks 16 and 52
- Decrease of at least 2 from baseline in LEI, among subjects with baseline LEI of at least 2, at weeks 4, 12 and 16
- Subject's Global Impression of Change Questionnaire at weeks 16 and 52
- AUC for ACR-N through week 16
- Time to first achievement of ACR50
- Time to first achievement of PASI90 among subjects with at least 3% BSA at baseline
- Time to first achievement of MDA
- Time to first achievement of ACR20
- Time to first achievement of PsAID improvement of at least 3 among subjects with PsAID at least 3 at baseline
- Time to first achievement of PASI100 among subjects with at least 3% BSA at baseline
- Trough plasma concentrations of izokibep at collected timepoints
- Presence of anti-drug antibodies (positive or negative)

2.7. Safety Variables

2.7.1. Adverse Events, Serious Adverse Events, and Other Safety Reporting

Adverse events will be reported by the subject (or, when appropriate, by a caregiver).

The investigator and any qualified designees are responsible for detecting, documenting, and recording events that meet the definition of an Adverse Event (AE) or SAE. The investigator is responsible for following up on all AEs that are serious, considered related to the study drug or study procedures, or that caused the subject to discontinue the study drug.

2.7.1.1. Events of Special Interest

Based on the class effects or potential risks with IL-17 inhibitors, the following events of special interest will be monitored:

- Candida infection
- Inflammatory bowel disease

In addition, based on the potential risk with an IL-17 receptor inhibitor the following events of special interest will also be monitored:

- Suicidal ideation
- Malignancies
- Major cardiovascular and cerebrovascular events (MACE): cerebrovascular accident and transient ischemic attack, non-fatal myocardial infarction or unstable angina, cardiovascular death)
- Tuberculosis (TB)
- Infections (serious, opportunistic, or fungal only)
- Cytopenias
- Systemic hypersensitivity reactions

2.7.2. Clinical Safety Laboratory Tests

See SAP Table 1 for the list of clinical laboratory tests to be performed and the SoA in the protocol for the timing and frequency.

Table 1 Laboratory Tests

Hematology **Test (SI unit)**	**Serum Chemistry** **Test (SI unit)**	**Urinalysis**

• Red Blood Cell Count (10^{12}/L)
• Hemoglobin (g/L)
• Hematocrit (%)
• White Blood Cell Count (10^9/L)
• Differential WBC (10^9/L and %)
  ○ Neutrophils
  ○ Lymphocytes
  ○ Eosinophils
  ○ Basophils
  ○ Monocytes
• Platelet count (10^9/L)
• RBC indices:
  ○ Mean corpuscular volume (MCV) (fl)
  ○ Mean corpuscular hemoglobin (MCH) (Pg)
  ○ %Reticulocytes
	• Alanine aminotransferase (ALT) (U/L)/Serum glutamic-pyruvic transaminase
	• Aspartate aminotransferase (AST) (U/L)/Serum glutamic-oxaloacetic transaminase
	• Total and direct Bilirubin (mmol/L)
	• Alkaline phosphatase (U/L)
	• Total protein (g/L)
	• Glucose (not fasting) (mmol/L)
	• Creatinine (mmol/L)
	• Blood Urea Nitrogen (mmol/L)
	• Calcium (mmol/L)
	• Potassium (mmol/L)
	• Sodium (mmol/L)
	• C-reactive protein (mg/L)
	
	• A1C (%)
	• High-density-lipoprotein (HDL) (mmol/L)
	• International normalized ratio*
	• Total cholesterol (mmol/L)
		• Specific gravity
		
		<u>Urinalysis (dipstick)</u>
		• pH
		• Glucose
		• Proteins
		• Red Blood Cell Count
		• White Blood Cell Count
		• Ketones
		
		<u>Microscopic examination</u> (if blood or protein is abnormal)

Pregnancy Testing	**Other Screening Tests**	**Other Tests**
• Highly sensitive serum human chorionic gonadotropin pregnancy test (as needed for women of childbearing potential) at screening	• Follicle-stimulating hormone and estradiol (as needed in women of nonchildbearing potential only) • Hepatitis B virus testing • Hepatitis C virus antibody • Human immunodeficiency virus testing • Tuberculosis testing • Rheumatoid factor/anti-cyclic citrullinated peptide • Fecal calprotectin	• Pharmacokinetics • Anti-drug antibody

*INR will only be measured if ALT or AST ≥ 3 x ULN as detailed in protocol Section 7.1.1. A separate blood sample for INR testing will be included in the Liver Event Monitoring Kit at the time of repeat testing for ALT or AST, as detailed in protocol Section 7.1.1.

2.7.3. Vital Signs

Vital signs will be measured in a sitting position after 5 minutes of rest and will include temperature, systolic and diastolic blood pressure, pulse (collected as "Heart rate" in the CRF), and respiratory rate.

2.7.4. Physical Examination

A complete physical examination will include, at a minimum, assessments of the dermatological, cardiovascular, respiratory, gastrointestinal, and neurological systems.

Clinically significant findings observed prior to the first dose of study drug will be listed as medical history in the eCRF and reported as AEs if observed after the first dose of the study drug.

2.7.5. Electrocardiograms

Triplicate 12-lead ECGs will be obtained to calculate the heart rate and measure PR, QRS and QT intervals.

2.7.6. Hospital Anxiety and Depression Scale

Depression and anxiety will be monitored with the Hospital Anxiety and Depression scale (HADS). The HADS questionnaire will be completed by the subject.

2.7.7. Pharmacokinetics Variable(s)

Blood samples will be collected for measurement of plasma concentrations of izokibep as specified in the SoA in the protocol.

2.8. Immunogenicity

Anti-drug antibodies to izokibep will be evaluated in serum samples collected from all subjects according to the SoA in the protocol. These samples will be tested for positive or negative by the sponsor or sponsor's designee.

Samples may be further tested for IL-17A binding domain and Albumin binding domain (positive vs negative).

3. Estimands

The ICH[1] E9 (R1) addendum on estimands[2] and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials came into effect on 30 July 2020. This section addresses the construction of estimands for the primary, secondary, and exploratory objectives. Each estimand is defined according to the following five attributes:

- The **treatment** condition of interest, and as appropriate, the alternative condition to which comparison will be made.
- The **population** of subjects targeted by the clinical question.
- The **variable** (or endpoint) to be obtained for each subject that is required to address the clinical question.
- The clinical question of interest in respect of **other intercurrent events** not covered through the precise specifications of treatment, population and variable.
- A **population-level summary** for the variable providing a basis for comparison between treatment conditions.

3.1. Estimand for the Primary Objective

The main estimand is defined through the following five attributes:

3.1.1. Treatment Condition of Interest

The primary treatment condition of interest is izokibep 160 mg QW and izokibep 160 mg Q2W. The reference treatment condition is placebo administered QW. Izokibep 160 mg QW and izokibep 160 mg Q2W will each be compared to placebo. Subjects will be analyzed according to their randomized treatment regardless of treatment compliance, use of concomitant medications or rescue medications, or any other protocol deviations.

3.1.2. Population of Subjects Targeted by the Clinical Question

The population of subjects is males and females ≥ 18 years and ≤ 75 years with active PsA (≥3 SJC66 and ≥3 TJC68 at screening and baseline visits) as further defined by the inclusion criteria and exclusion criteria. The population being studied represents a population normally seen in clinical practice. This ensures the activity of izokibep can be evaluated across a distribution of disease severity in the study.

3.1.3. Variable Obtained from Each Subject Required to Address the Clinical Question

For each subject in the study, the variable to address the clinical question is ACR50 response at Week 16.

3.1.4. Handling of Intercurrent Events to Reflect the Clinical Question of Interest

The treatment strategy approach for the estimand will be used in general, so subjects will be included using observed data at Week 16 regardless of treatment discontinuation, concomitant medication, protocol deviations, or other actions. Subjects with missing full ACR50 assessments at Week 16 will be imputed as non-responders (non-response imputation or NRI). Subjects missing one or more components of ACR50 will be included as defined in Section 5.2.1, with subjects having missing ACR50 due to missing components also imputed as non-responders.

3.1.5. Population-level Summary for Comparison between Treatment Conditions

The population level summary is the response rate, or proportion of subjects who meet ACR50 at Week 16. The null hypothesis of equal response rates will compare each dosing regimen of izokibep to the placebo group respectively. A stratified test of response rates will be implemented in the analysis. Stratified tests will use the strata of prior TNFi use (Yes/No) and enthesitis at baseline (LEI = 0 versus >0) from the randomization process. If a subject is incorrectly classified into a stratum during the randomization process, the analysis will use the correct classification, not the classification used during randomization. Within each of the strata used for randomization, the response rate for each treatment group and corresponding standard error will be calculated. The difference in response rates (risk difference) will be calculated for each stratum. The common risk difference among the strata and associated standard error will be estimated by combining the observed risk differences using Cochran-Mantel-Haenszel weighting. The estimated risk difference divided by the standard error will be used as the test statistic and a p-value calculated assuming that the test statistic follows a standard normal distribution under the null hypothesis. Analyses at earlier timepoints when data to calculate the ACR50 are collected will also be presented using the same methodology. P-values from earlier timepoints will be presented for descriptive purposes, not part of the alpha-preserving multiple testing strategy.

The treatment effect will be quantified by the response rate difference of ACR50 at Week 16 between placebo and each dosing regimen of izokibep at Week 16.

3.2. Estimands for the secondary objectives

The main estimands for each of the secondary objectives are defined through the following five attributes:

3.2.1. Treatment Condition of Interest

The primary treatment condition of interest is izokibep 160 mg QW and izokibep 160 mg Q2W. The reference treatment condition is placebo administered QW. Izokibep 160 mg QW

and izokibep 160 mg Q2W will each be compared to placebo. Subjects will be analyzed according to their randomized treatment regardless of treatment compliance, use of concomitant medications, or any other protocol deviations.

3.2.2. Population of Subjects Targeted by the Clinical Question

The population of subjects is all randomized males and females ≥ 18 years and ≤ 75 years with active PsA (≥3 SJC66 and ≥3 TJC68 at screening and baseline visits) as further defined by the inclusion and exclusion criteria.

For some of the secondary endpoints, further restrictions on inclusion in the analysis result in a smaller target population, a subset of the FAS:

- First secondary endpoint: achieving PASI90. This endpoint will be assessed in subjects with ≥3% BSA psoriasis at baseline, to allow for a calculation of PASI90.
- Second secondary endpoint: resolution of enthesitis. This endpoint will be assessed using only subjects with enthesitis at baseline (LEI >0), to ensure this represents improvement in enthesitis rather than having no enthesitis at baseline.
- Sixth secondary endpoint: proportion of subjects achieving an improvement in PsAID of at least 3 units. This endpoint will be assessed using only subjects with PsAID of at least 3 at baseline, to ensure subjects are capable of meeting endpoint, since negative values of PsAID are not possible.

As for the secondary endpoints, further restriction on inclusion in the analysis result in a smaller target population, a subset of the FAS:

- The exploratory endpoint: change from baseline in spine pain NRS. This endpoint will be assessed using only subjects with sacroiliitis at baseline.
- The exploratory endpoint: change from baseline in spine pain NRS. This endpoint will be assessed in those with investigator assessed spinal involvement.
- The exploratory endpoint: time to resolution of enthesitis. This endpoint will be assessed using only subjects with enthesitis at baseline (LEI >0), to ensure this represents improvement in enthesitis rather than having no enthesitis at baseline.
- The exploratory endpoint: occurrence of enthesitis. This endpoint will be assessed using only subjects with enthesitis at baseline (LEI=0), with occurrence defined as of enthesitis (LEI > 0) summarized at weeks 4, 8, 12 and 16.

Other secondary and exploratory endpoints will be assessed using the full analysis set. The population being studied represents a population normally seen in clinical practice. This ensures the activity of izokibep can be evaluated across a distribution of disease severity in the study.

3.2.3. Variable Obtained from Each Subject Required to Address the Clinical Question

For each subject in the study, the variables to address the clinical questions are:

- Achieving 90% or greater reduction in PASI score from baseline (PASI90) at Week 16 in subjects with ≥3% BSA psoriasis at baseline
- Resolution of enthesitis at Week 16 as assessed by LEI in subpopulation that had enthesitis (LEI > 0) at baseline
- Achieving MDA at Week 16
- Achieving 20% improvement in ACR core set measurements (ACR20) at Week 16
- Achieving an improvement in PsAID of at least 3 units at Week 16 compared to baseline in subjects with PsAID ≥ 3 at baseline
- Change in physical function as assessed by HAQ-DI change from baseline to Week 16

3.2.4. Handling of Intercurrent Events to Reflect the Clinical Question of Interest

The treatment policy strategy similar to the primary endpoint will be used to construct estimands for each secondary endpoint. Subjects will be included in the analysis regardless of treatment discontinuation, concomitant medication, protocol deviations, or other actions. Subjects who have missing data at Week 16 will be imputed as nonresponders and included in the analysis using composite strategy.

The secondary endpoint of change in physical function as assessed by HAQ-DI change from baseline from baseline to Week 16 will be analyzed as a continuous variable. The treatment policy strategy will be used to construct estimands for change in physical function of HAQ-DI from baseline to Week 16. Subjects will be included in the analysis regarding treatment discontinuation, concomitant medication, protocol deviations, or other actions.

3.2.5. Population-level Summary for Comparison between Treatment Conditions

- Proportion of subjects achieving 90% or greater reduction in PASI score from baseline (PASI90) at Week 16 in subjects with ≥3% BSA psoriasis at baseline
- Proportion of subjects with resolution of enthesitis at Week 16 as assessed by LEI in subpopulation that had enthesitis (LEI > 0) at baseline
- Proportion of subjects achieving MDA at Week 16
- Proportion of subjects achieving 20% improvement in ACR core set measurements (ACR20) at Week 16
- Proportion of subjects achieving an improvement in PsAID of at least 3 units at Week 16 compared to baseline in subjects with PsAID ≥ 3 at baseline
- Mean change in physical function as assessed by HAQ-DI change from baseline to Week 16

Pairs of treatment arms will be compared for each summary by assessing the difference in response rates or difference in means between the two arms.

4. Analysis Sets

In accordance with ICH E3 and E9[3], the following analysis sets will be used for the analyses.

4.1. All Screened Set

The All Screened Set will include every subject who has signed the informed consent form. The All Screened Set will be used for summaries of disposition and the associated listing.

4.2. Full Analysis Set

For assessing the primary and secondary efficacy objectives, all subjects randomized will be included in the analyses as the FAS. Intercurrent events (ICEs) such as missed assessments, missed or discontinued treatment, and protocol deviations, will be addressed as described in the definition of the estimands in Section 3. Subjects will be included according to randomized treatment.

4.3. Safety Analysis Set

For assessing the safety objectives, all subjects randomized who receive at least one administration of test material will be included in the summaries and analyses. In the event that a subject receives incorrect study drug, that subject will be grouped according to the treatment received more often or, if tied, the first study treatment received, and such subjects will be listed and counted in the final report.

4.4. Pharmacokinetics Analysis Set

The Pharmacokinetics analysis set for subjects will include all subjects who received at least one administration of test material and with having at least one quantifiable PK concentration.

4.5. ADA Analysis Set

The ADA analysis set for subjects will include all subjects who received at least one administration of izokibep and have both a baseline ADA measurement and at least one post-dose ADA measurement.

5. Data Handling

5.1. Time Points and Visit Windows

5.1.1. General Definitions

All assessment days will be related to the first day of the first dose of the study drug.

Day 1 is defined as the date of the first dose of the study drug. Relative days after Day 1 are calculated as (assessment date – Day 1 date) + 1. Relative days prior to Day 1 are calculated as (assessment date – Day 1 date). The day prior to Day 1 is Day -1. There is no Day 0.

The date of the first dose of study drug for each subject will be taken from the Study Drug Administration eCRF page. If the date in this eCRF page is missing, alternatively the date of enrollment/randomization will be used.

The date of the last dose of study drug for each subject will be taken from the Study Drug Administration form, with the last study drug administration date from any visit used as the date of last study drug.

5.1.2. Screening Period

For all subjects, the screening period starts when the subject signs and dates the ICF and ends when the subject is enrolled/randomized, or screen failed. The screening period is up to 28 days. Certain initial screening period procedures may be repeated during the original initial screening period. This is not considering “rescreening”.

The baseline value for a variable is therefore defined as the last non-missing value collected before randomization.

5.1.3. Treatment Period

Data collected on Day 1 will be assigned to the Treatment Period unless the time (HH:MM) of data collection and time (HH:MM) of the first dose of study drug are both recorded and the data collection time is before the time of the first dose of study drug. In this case, the assessment will be assigned to the screening period. If the time (HH:MM) of data collection is not recorded but the protocol and/or eCRF includes an instruction to the effect that all Day 1 assessments are to be performed prior to the first dose of study drug, the data collected at Day 1 will be assigned to the screening period. However, adverse events and concomitant medications starting on Day 1, will be assigned to the Treatment Period, unless the time (HH:MM) demonstrates otherwise.

The Treatment Period is defined as the period from the date/time of the first dose of study drug to Week 52 (one week after the last dose) or early termination (end of treatment visit

should be completed within 14 days of withdrawal and follow-up visits should be completed at 8 weeks and 14 weeks after the last dose of study drug, where possible).

Treatment will be divided into two periods, Period 1 and Period 2, Period 1 will start at the first dose of study drug and ends at Week 16 when Week 16 assessments are completed and before administration of study drug. Period 2 will start at the time of administration of study drug at Week 16 and end at Week 52 or early termination/end of treatment visit.

5.1.4. Visit Windows

All data will be analyzed using the nominal study visit as defined in the SoA in the protocol and eCRF.

In the event that a subject discontinues the study early, the subject will have an early termination visit which includes efficacy assessments. If the early termination visit or an unscheduled visit occurs in a timeframe that corresponds to a scheduled efficacy assessment (week 2, 4, 12, 16, 20, 24, 32, 36, 40 or 52), and if the subject does not have an actual visit at or after the corresponding scheduled efficacy assessment, the following day ranges will be used to map the early termination/unscheduled visit to a scheduled efficacy assessment for use in all efficacy summaries at that timepoint. Similarly, if an unscheduled efficacy assessment occurs during a timeframe that corresponds to a scheduled efficacy assessment, it will be windowed accordingly, provided the subject does not have an actual visit at the same nominal time.

Table 2 Definition of Time Point Windows for Efficacy Analyses

Time Point	Time Point Window [a]
Baseline	<=Day 1
Week 2	Days 12 to 21
Week 4	Days 22 to 56
Week 12	Days 57 to 98
Week 16	Days 99 to 126
Week 20	Days 127 to 154
Week 24	Days 155 to 196
Week 32	Days 197 to 238
Week 36	Days 239 to 266
Week 40	Days 267 to 322
Week 52	Days 323 or later

[a] relative to the date of first dose of study drug during the Treatment Period/date of Baseline Visit (Day 1)

5.2. Handling of Dropouts, Missing Data, and Outliers

5.2.1. Handling of Missing Efficacy Data

Tender (and swollen) joint count will be the sum of the joints where tenderness (or swelling) is marked as Present on the CRF. A missing response will be treated the same as an Absent response, in that, it will not contribute to the joint count.

Subjects with missing efficacy assessments will be imputed as non-responders (non-response imputation or NRI) for dichotomous endpoints. Subjects with a missing component of a composite endpoint will be imputed as non-responders for that component. The subject may therefore be a responder to the endpoint, if the subject met a sufficient number of criteria from other components.

Subjects missing one or more components of composite endpoints will be included, illustrated for the primary endpoint, ACR50, as follows:

Situation	ACR50 result
If both SJC66 and TJC68 have reductions ≥ 50%: - If at least 3 of the other 5 components have reductions ≥ 50%, irrespective of missing/available on the others - If at least 3 of the other 5 components have a reduction of < 50%, irrespective of missing/available on the others - All other situations	 Responder Non-responder Missing
If result is available on both of SJC66 and TJC68, and at least one has <50%: - Irrespective of results or missingness on the other 5 components	 Non-responder
If one of SJC66 and TJC68 is missing, and the other has reduction ≥ 50%: - If at least 3 of the other 5 components have reductions ≥ 50%, irrespective of missing/available on the others - If at least 3 of the other 5 components have a reduction of < 50%, irrespective of missing/available on the others - All other situations	 Missing Non-responder Missing
If one of SJC66 and TJC68 is missing, and the other has reduction <50%: - Irrespective of results or missingness on the other 5 components	 Non-responder
If both SJC66 and TJC68 missing: - If at least 3 of the other 5 components have a reduction of < 50%, irrespective of missing/available on the others - All other situations	 Non-responder Missing

For continuous efficacy endpoints, inference will be done using mixed-effects models repeated-measures[4], which is robust to missing outcomes that are missing at random or missing complete at random.

5.2.2. Handling of Missing Safety Data

In general, missing clinical laboratory data, vital signs, and ECG data will not be imputed. Adverse event imputations for missing severity or relationship are given in Section 6.7.2. Unknown or partial medication and AE date imputations are given below and are to be used only for the assessment of prior/concomitant status for medications and treatment-emergent status for AEs.

5.2.3. Handling of Missing Age and Partial and Missing Dates for Adverse Events, Prior/Concomitant Medications

Missing Age at Time of Informed Consent

Where the age at time of informed consent is missing and year of birth is collected, the following convention will be used:

The day and month will be completed as 02JUL (the 183rd day of the year). For example, Date of Birth specified as -----1980 will be completed as 02JUL1980.

Missing or Partial Adverse Event and Prior/Concomitant Medication Start/Stop Dates

Missing and/or incomplete dates for medications and AEs are imputed in a manner resulting in the earliest onset or the longest duration during the Treatment Period, whilst ensuring that the start date does not occur after the stop date. The stop date will not be imputed if the medication or AE is “Ongoing”. Technically, this will be done as follows:

For a missing/incomplete start date/time the earliest date/time of the following will be imputed:

- The later date of: the earliest possible start date/time, and the date/time of the first dose of treatment.
- The latest possible start date/time.
- The latest possible stop date/time.

For a missing/incomplete stop date/time the later date/time of the following will be imputed:

- The earlier date/time of the latest possible stop date/time and the date/time of last dose of treatment.
- The earliest possible stop date/time.

- The earliest possible start date/time.

Here, the earliest possible date/time is defined as:

- The date/time itself if available.
- The date/time of the first day of the month at 00:00hrs, if month and year are available but the day/time is missing.
- The date/time of the first day of the year at 00:00hrs, if year is available but day/time and month are missing.
- 00:00hrs on the day of informed consent, if the date/time is completely missing.

The latest possible date/time is defined as:

- The date/time itself if available.
- The date/time of the last day of the month at 23:59hrs, if month and year are available but the day/time is missing.
- The date/time of the last day of the year at 23:59hrs, if year is available but day/time and month are missing.
- 23:59hrs on the date of last known date on the study for the subject plus one year, if the date/time is completely missing.

5.2.4. Handling of Plasma Concentrations that are Below the Lower Limit of Quantification

Plasma concentrations that are below the lower limit of quantification (BLQ) will be handled as follows for descriptive statistics:

- Predose concentrations that are BLQ will be treated as zero.
- Post-dose concentrations that are BLQ will be treated as missing elsewhere.
- For pre-dose BLQ results treated as 0, geometric mean and coefficient of variation (CV) % of geometric mean will be reported as not calculated (NC).

6. Statistical Methods

6.1. General Principles

All data processing, summarization, and analyses will be performed using [redacted] SAS Environment/Version 9.4 (or later) of the SAS® statistical software package.

The default summary statistics for quantitative variables will be the number of observations (n), mean, standard deviation (SD), median, minimum (min), and maximum (max), for those subjects with data.

For qualitative variables, the number (n) and percentage (%) of subjects with non-missing data per category will be the default summary presentation, and where appropriate and present, the number of missing values as a "Missing" category. Number of subjects in the analysis population will be used as denominator for percentages calculation, unless stated otherwise in TFLs mock shell(s).

The treatment group labels and presentation order are given below:

Full treatment group label:	**Short treatment group label:**	**Treatment group ordering:**
Placebo QW Week 0 to Week 15; Izokibep 160 mg QW Week 16 to Week 51	Placebo QW - Izokibep QW	1
Izokibep 80 mg Q4W Week 0 to Week 48; Placebo QW Week 1 to Week 51 for the Weeks in Between Izokibep Doses	Izokibep SC 80 mg Q4W	2
Izokibep 160 mg Q2W Week 0 to Week 50; Placebo Q2W Week 1 to Week 51	Izokibep SC 160 mg Q2W	3
Izokibep 160 mg QW Week 0 to Week 51	Izokibep SC 160 mg QW	4

All statistical comparisons of the primary and secondary endpoints will be made using two-sided tests at the $\alpha = 0.05$ significance level (with decisions also requiring consideration of multiplicity as defined below) unless specifically stated otherwise. All null hypotheses will be of no treatment difference. All alternative hypotheses will be two-sided.

All laboratory test results will be received from the central/local laboratories, and the results will be provided in either standard internal (SI) or conventional units, or both. For the TFLs, the results will be summarized or presented in International System of Units (SI) units. Refer to Appendix of the TFLs mock shells for the SI unit corresponding to each laboratory test. Refer to Appendix of the TFL shells for the precision level in which each laboratory test is reported by the central laboratories.

Specifications for table, figures and data listing formats can be found in the TFL shells specifications for this study. Please refer to "2. General Format Guidelines" section within TFL shells for more details on presentation of results.

For comparisons of izokibep to placebo, the three dosing regimens of izokibep (QW, Q2W and Q4W) will be compared separately to placebo. Binary endpoints including the primary endpoint of ACR50 will be compared with a stratified test of risk difference. Continuous endpoints will be compared with a repeated measures mixed model.

Safety data, such as adverse events, serious adverse events, and laboratory data, will be summarized.

6.2. Subject Disposition and Data Sets Analyzed

Subject disposition will be summarized by treatment group and overall for the All Screened Set. The following information will be reported:

- Number of subjects for the following categories:
  - Screened
- Number and percentage of subjects for the following categories:
  - Randomized
  - Treated
  - Not Treated
  - Permanently discontinued study medication
  - Did not meet cDAPSA low disease activity (cDAPSA $\leq$ 14) criteria at two consecutive visits, starting at Week 24 or after
  - Discontinued study prior to the Week 16 assessment
  - Completed the Week 16 assessment
  - Completed the study
  - Ongoing in the study
  - Discontinued the Study
  - Reasons for study discontinuation
- Number and percentage of subjects included in, and excluded from, each study population together with the reasons for exclusion from the analysis set

- Number and percentage of subjects who completed/discontinued treatment, including the reasons for treatment discontinuation
- Number and percentage of subjects who met/did not meet all eligibility criteria, together with the criteria not met
- Number and percentage of subjects who failed screening prior to randomization, including the primary reason for screen failure
- Number and percentage of subjects at each region (and each country within region)
    - Regions are defined using the following countries:
        - North America (US or Canada)
        - Europe (Bulgaria, Czech Republic, Spain, Hungary, Poland, or Germany)
        - Non-US (Bulgaria, Czech Republic, Spain, Hungary, Poland, Germany, or Canada)
- Number and percentage of subjects by stratification factors. Each stratification variable, prior TNFi use (Yes/No) and enthesitis at baseline (LEI = 0 versus >0)
    - By the stratum collected in Interactive Voice/Web Response System (IXRS)
    - By the stratum according to the clinical database

A subject will be regarded as having completed the study if the status recorded on the End of Study eCRF form is Complete, or assessments at Week 65 have been completed. A subject will be considered as having discontinued the study if they have an eCRF status of premature study discontinuation. Otherwise, the subject will be considered as ongoing in the study, for summaries before final study completion.

A listing of all subjects with their treatment and study completion status, including the respective reasons for treatment discontinuation will be presented for all subjects randomized in the FAS.

A listing of all screen failed subjects with their reasons for screen failure will be presented for the All Screened Set. A separate listing of subjects who failed at least one inclusion/exclusion criteria including a text description of the criterion failed will be presented for the All Screened Set.

A listing of all randomized subjects with their randomization details, including first dose date and time, and actual treatment received will be presented for the FAS.

A listing of all randomized subjects excluded from the FAS will be presented.

6.3. Protocol Deviations

All important protocol deviations will be summarized for the FAS by treatment group and overall as described below:

- The number of unique subjects with at least one important protocol deviation as well as the number of subjects in each important protocol deviation category will be presented by default descriptive summary statistics for categorical variables.

A listing of all subjects with one or more protocol deviations will be presented for the FAS.

6.4. Demographic and Other Baseline Characteristics

6.4.1. Demographic Characteristics

Demographic characteristics will be summarized for the FAS by treatment group and overall as described below. All missing data will be presented as part of a missing category, if appropriate. Standard descriptive statistics will be presented for the continuous variables of:

- Age (years)

Total counts and percentages of subjects will be presented for the categorical variables of:

- Age group (years):
  - < 65
  - $\geq 65 - < 75$
  - ≥ 75
- Sex
- Ethnicity
- Race

Demographic characteristics will be listed for the FAS by treatment group.

6.4.2. Baseline Characteristics

Baseline characteristics will be summarized for the FAS by treatment group and overall as described below. All missing data will be presented as part of a missing category, if appropriate.

Standard descriptive statistics will be presented for the continuous variables of:

- Disease duration (years)
  - calculated as (date of randomization - date of first diagnosis of psoriatic arthritis) + 1/365.25
- Height (cm)
- Weight (kg)
- Body mass index (kg/m^2)

  - calculated as (body weight / height^2) where weight is in kg and height is in m, and presented to one decimal precision
- Body Temperature (degrees Celsius)
- Respiration Rate (breaths per minute)
- Systolic Blood Pressure (SBP [mmHg])
- Diastolic Blood Pressure (DBP [mmHg])
- Heart Rate (beats per minute [bpm])
- Swollen joint count based on 66 joints (SJC66)
- Tender joint count based on 68 joints (TJC68)
- Swollen joint count based on 28 joints (SJC28)
- Tender joint count based on 28 joints (TJC28)
- Subject's Assessment of Pain (VAS) (SAPVAS)
- Subject's Global Assessment of Disease Activity (SGADA)
- Investigator's Global Assessment of Disease Activity (IGADA)
- Health Assessment Questionnaire Disability Index (HAQ-DI) total score
- Acute phase reactant (C-reactive protein [CRP]) in mg/dL
- Psoriatic Arthritis Impact of Disease (PsAID)
- Psoriasis Area and Severity Index Score (PASI) in subjects with ≥ 3% BSA psoriasis at baseline
- Affected Body Surface Area (BSA)
- Disease Activity Score in 28 Joints (DAS28-CRP)
- Disease Activity in Psoriatic Arthritis (DAPSA)
- Clinical Disease Activity in Psoriatic Arthritis (cDAPSA)
- 36-Item Short Form Survey (total score, PCS, and MCS)
- Functional assessment of chronic illness Therapy-Fatigue (FACIT-F)
- Leeds Enthesitis Index among those with enthesitis (LEI > 0 at baseline)
- Leeds Dactylitis Index among those with dactylitis (LDI>0 at baseline)
- Psoriatic Arthritis Disease Activity Score (PASDAS)
- Subject's Assessment of Spinal Pain (NRS)
- Widespread Pain Index (WPI)
- Modified Nail Psoriasis Severity Index (mNAPSI) in those with nail psoriasis at baseline (mNAPSI>0)
- SPARCC Enthesitis Index among those with enthesitis (SPARCC >0 at baseline)
- mTSS
- Hospital Anxiety and Depression Scale (HADS)
- Number of prior biologic Disease-Modifying Antirheumatic Drugs (DMARDs)
  - Number of prior TNFi
  - Number of prior biologic DMARDs other than TNFi

- Number of prior failed biologic DMARDs

Total counts and percentages of subjects will be presented for the categorical variables of:

- Any csDMARD at Baseline group (Yes/No)
- Any csDMARD at Baseline group (MTX alone, MTX and other csDMARD, csDMARD other than MTX)
- Number of prior failed biologic DMARDs group (0, 1, 2 or more)
- Prior TNFi Use (Yes/No)
    - Failed prior TNFi (Yes/No)
- Number of prior failed biologics group (0, 1, 2 or more)
- Prior JAK Inhibitor Use
- Concurrent use of oral NSAID on the first dosing date of study drug administration
- Concurrent use of oral corticosteroids on the first dosing date of study drug administration
- Presence of enthesitis (Yes/No)
    - LEI = 0
    - LEI > 0
    - LEI = 1
    - LEI = 2
    - LEI = 1 or 2
    - LEI < 3
    - LEI ≥ 3
    - LEI < 4
    - LEI ≥ 4
    - SPARCC (SPARCC > 0 at baseline)
- Combination of enthesitis and prior TNFi exposure
    - LEI = 1 or 2 and no prior TNFi exposure
    - LEI ≥ 3 and no prior TNFi exposure
    - LEI = 1 or 2 and prior TNFi exposure
    - LEI ≥ 3 and prior TNFi exposure
- Presence of dactylitis (Yes/No)
- Investigator Question on Presence of Axial Involvement (yes/no)
- Childbearing Potential
- Alcohol (Never, Former, Current)
- Nicotine (Never, Former, Current)

Baseline characteristics will be listed for the FAS.

6.4.3. Medical History

Medical history is defined as any condition, with the exception of the study indication, that the subject may have prior to study drug administration in the study, including any chronic conditions diagnosed prior to entry into the study. See section 5.2.3 for imputation of missing or partial dates for medical history. Non-serious AEs with onset after signing ICF but before the randomization will be included as medical history.

Medical history will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) [version 24.0 (or a later version if updated during the study)] and will be presented by System Organ Class (SOC) and Preferred Term (PT) and total. The SOCs and PTs are to be sorted by Internationally Agreed order SOCs and descending PTs in the total column.

Medical history records will be summarized for the FAS by treatment group and overall as follows:

- The number and percentage of subjects with at least one medical history record will be presented.
- The number and percentage of subjects with at least one medical history record within each primary SOC and PT will be presented. The summary will be sorted using the internationally agreed order for SOC and using descending order of overall numerical counts for PT. Where terms tie, these will be sorted alphabetically.

Medical history records will be listed by-subject and within-subject by medical history start date for the FAS.

6.4.4. Prior and Concomitant Medications/Procedures

All medications will be coded using the WHO Drug Global Dictionary, Format B3 [Version March 2022 (or a later version if updated during the study)], Anatomical Therapeutic Chemical (ATC) Classification codes.

Prior medications and concomitant medications are defined as follows:

- Prior medications are those taken prior to screening with a stop date and time prior to the start of the Treatment Period.
- Concomitant medications are those with a start date and time on or before the end of the Treatment Period and a stop date on or after the start of the treatment period, including prior medications that are ongoing at the start of the treatment period.

See Section 5.2.3 for imputation of missing or partial dates for medication.

Prior and concomitant medications will be summarized for the FAS by treatment group and overall as follows:

- The number and percentage of subjects with at least one prior/concomitant medication will be presented.
- The number and percentage of subjects with at least one prior/concomitant medication within each Anatomical Group (ATC Level 1), Therapeutic Subgroup (ATC Level 2), and preferred term will be presented. The summary will be sorted using numerical counts by descending order of Anatomical Group, then descending order of Therapeutic Subgroup, then descending order of preferred term in the total column. Where groups or terms tie these will be sorted alphabetically.
- Number and percentage of subjects with each primary reason for stopping NSAID, csDMARD, TNFi or biologic medication taken for Plaque Psoriasis or PsA.

Prior medications and concomitant medications will be listed separately for the FAS. In the listings the relative start and stop day of prior/concomitant medication use will be calculated relative to the first dose date and time of study drug and will be presented for those subjects who received at least one dose of study drug. If the concomitant medication is "Ongoing" it will be indicated as such in the listing and the relative stop day will not be calculated. Only original dates will be presented in the listing even though the relative day may be based on an imputed date.

All prior and concomitant procedures will also be presented in the listing.

6.5. Measurements of Study Drug Compliance

Study Drug compliance is defined as the number of doses that were actually taken relative to the number of doses that should have been taken as per the protocol for the planned duration of treatment exposure. A partial dose in eCRF will be treated as a half dose. Study drug compliance will be summarized for period 1 and period 2 separately.

In general, the percentage of overall compliance, assessed by dose count, will be calculated as follows:

For Period 1

$$\text{Compliance (\%)} = \frac{\text{(Total doses administered from Week 0 up to before Week 16)}}{\text{Total doses expected to be administered from Week 0 to before Week 16}} \times 100\%$$

For Period 2

$$\text{Compliance (\%)} = \frac{\text{(Total doses administered from Week 16 to EOT)}}{\text{Total doses expected to be administered from Week 16 to EOT}} \times 100\%$$

For group 3 and group 4, matching placebo will be counted in the calculation of total doses administered and total doses expected. The total number of doses expected to be administered during Period 1 is therefore 16 for all groups. The total number of doses expected to be administered during Period 2 will be different after accounting for study drug withdrawal for those who meet criteria at or after week 24.

The calculated percentage compliance will be categorized as:

- < 80% compliance
- ≥ 80%

Compliance will be summarized for the Safety Analysis Set by treatment group as follows:

- Number of doses will be presented by default summary statistics. This will include both number of doses of izokibep and number of doses of placebo.
- Percent compliance will be presented by default summary statistics.
- Number and percentage of subjects within each of the compliance categories will be presented. Treatment compliance will be listed for the Safety Analysis Set. Missing data will not be imputed and only original data for those fields (for example, date fields) will be presented in the listing together with derived variables such as the calculated compliance (%) and exposure duration.

Treatment compliance will be listed for the Safety Analysis Set. Missing data will not be imputed and only original data for those fields (for example, date fields) will be presented in the listing together with derived variables such as the calculated compliance (%) and exposure duration.

6.6. Efficacy

Hypotheses tested will be adjusted to control the familywise error rate in the strong sense at $\alpha = 0.05$, 2-sided.

The statistical comparisons for the primary efficacy endpoint and the secondary endpoints, all at Week 16, will be carried out in sequential order. The primary endpoint, comparing izokibep dosed QW to placebo, will be tested first, with significance concluded if $p < 0.05$.

Testing of secondary endpoints, comparing izokibep dosed QW to placebo, will only be carried out if all prior tests, including the tests of the primary endpoint, first show significance with $p < 0.05$. If all primary and secondary endpoints, comparing izokibep doses QW to placebo, are significant, testing of izokibep doses Q2W to placebo will begin and follow the same order. As long as all prior tests are significant, testing will proceed in the following order:

- The first secondary endpoint, achieve PASI90 at Week 16 in subjects with ≥3% BSA psoriasis at baseline, comparing izokibep dosed QW to placebo.
- The second secondary endpoint, achieving resolution of enthesitis at Week 16 in subjects that had enthesitis (LEI > 0) at baseline, comparing izokibep dosed QW to placebo.
- The third secondary endpoint, achieving MDA at Week 16, comparing izokibep dosed QW to placebo.
- The fourth secondary endpoint, achieving ACR20 at Week 16, comparing izokibep dosed QW to placebo.
- The fifth secondary endpoint, achieving improvement in PsAID of at least 3 units at Week 16 compared to baseline in subjects with PsAID ≥3 at baseline, comparing izokibep dosed QW to placebo.
- The sixth secondary endpoint, change in physical function of HAQ-DI from baseline to Week 16, comparing izokibep dosed QW to placebo.
- The primary endpoint, achieving ACR50 at Week 16, comparing izokibep dosed Q2W to placebo.
- The first secondary endpoint, achieve PASI90 at Week 16 in subjects with ≥3% BSA psoriasis at baseline, comparing izokibep dosed Q2W to placebo.
- The second secondary endpoint, achieving resolution of enthesitis at Week 16 in subjects that had enthesitis (LEI > 0) at baseline, comparing izokibep dosed Q2W to placebo.
- The third secondary endpoint, achieving MDA at Week 16, comparing izokibep dosed Q2W to placebo.
- The fourth secondary endpoint, achieving ACR20 at Week 16, comparing izokibep dosed Q2W to placebo.
- The fifth secondary endpoint, achieving improvement in PsAID of at least 3 units at Week 16 compared to baseline in subjects with PsAID ≥3 at baseline, comparing izokibep dosed Q2W to placebo.
- The sixth secondary endpoint, change in physical function of HAQ-DI from baseline to Week 16, comparing izokibep dosed Q2W to placebo.

6.6.1. Primary Efficacy Analysis

The primary efficacy variable is defined as ACR50, the proportion of subjects achieving 50% improvement in ACR core set measurements.

The statistical null and alternative hypotheses to be used to assess the primary objective are:

$$H_0: \pi_{ABY} - \pi_{PBO} = 0$$

$H_A: \pi_{ABY} - \pi_{PBO} \neq 0$

where π_{ABY} and π_{PBO} are the proportion of subjects achieving ACR50 at Week 16 among subjects randomly assigned to receive izokibep and placebo, respectively. One set of hypotheses will be tested for each dose frequency of izokibep (QW and Q2W).

The treatment strategy approach for the estimand will be used in general, so subjects will be included using observed data at Week 16 regardless of treatment discontinuation, concomitant medication, protocol deviations, or other actions. Subjects with missing ACR50 assessments at Week 16 will be imputed as non-responders (non-response imputation or NRI).

Three components of the primary endpoint (subject's assessment of pain, subject's global assessment of disease activity, and investigator's global assessment of disease activity) were assessed using a visual analog scale (VAS), consisting of a horizontal line, 100 mm in length, with anchor statements at either end. The subject or investigator was asked to place a single vertical line through the VAS at the point that best corresponds to assessment. At some sites, and for some subjects, the VAS was not administered as intended. If a line was not 100 mm in length, the score used for analysis was calculated as 100 x (distance from left to mark)/(total length of horizontal line). If the subject was requested to write a number or was verbally asked for a number, that value will be used for analysis. Such protocol deviations will be noted in listings, and sensitivity analyses will be reported to assess the impact of such deviations. A spreadsheet will be created, after database lock and before unblinding, to note which subjects answered questions using a different length line, or by writing or verbally responding with a number. This spreadsheet will be signed and dated, and entered into the study files, before database lock to verify the protocol deviations and sensitivity analyses defined below.

The null hypothesis of equal response rates will compare each dosing regimen of izokibep to the placebo group. A stratified test of response rates will be used. Within each of the strata used for randomization, the response rate for each treatment group and corresponding standard error will be calculated. If a subject is incorrectly classified into a stratum during the randomization process, the analysis will use the correct classification, not the classification used during randomization. The difference in response rates (risk difference) will be calculated for each stratum. The common risk difference among the strata and associated standard error will be estimated by combining the observed risk differences using Cochran-Mantel-Haenszel weighting. The estimated risk difference divided by the standard error will be used as the test statistic and a p-value calculated assuming that the test statistic follows a standard normal distribution under the null hypothesis. A confidence interval on the difference will be calculated using analogous methods. Analyses at earlier timepoints when data to calculate the ACR50 are collected will also be presented using the same methodology. P-values from earlier timepoints will be presented for descriptive purposes, not part of the alpha-preserving multiple testing strategy.

The number and proportion of subjects achieving ACR50 in each treatment group will be reported at each scheduled assessment, along with Clopper-Pearson 95% CI without strata. The common risk difference with Cochran-Mantel-Haenszel weighting will be used to calculate a 95% CI for difference in response rates. A bar graph showing the proportion of subjects achieving ACR50 for each treatment group along with the associated p-value will be produced. ACR50 results will also be listed.

Sample SAS code with stratification factors of comparing each dosing regimen of izokibep to the combined placebo group is given as following:

```
ods output commonpdiff = ;
proc freq data= ;
  tables stratum * treatment * response / cmh commonriskdiff (test=mh cl=mh);
  weight count / zeros;
run;
```

6.6.2. Supplementary Analyses for the Primary Efficacy Analysis

Tipping point analysis: A tipping point analysis will be performed in order to evaluate the robustness of the results of the primary analysis. This analysis will use simulations to assess the impact of different response rates in subjects with missing primary endpoint data, and how different the response rates must be to change the conclusions of the study. The tipping point analysis will impute mean response rates in the two groups. Imputed response rates will be based on the response rates among subjects with observed primary endpoint data at Week 16. Sets of imputed mean response rates will include the observed response rates ($\hat{\pi}_{ABY}$, $\hat{\pi}_{PBO}$), as observed in each stratum, and rates that differ in increments of ±0.02, adjusted independently in each arm. For each of the 121 imputed mean response rates ($\hat{\pi}_{ABY}$ + a, $\hat{\pi}_{PBO}$ + b), with a and b each taking on all values from -0.10 to +0.10 in increments of 0.02, 1000 simulations will be reported. Within each simulation, a random number mechanism will be used to assign each subject who is missing the primary endpoint at Week 16 to either response or non-response. The primary stratified analysis will be reported on each simulation, and the probability that the null hypothesis of equal response rates is rejected will be reported for each set of imputed response rates. A graphical output will summarize the rejection probability for various sets of response rates. The same tipping point analysis will be performed in comparing each izokibep group to placebo.

Multiple Imputation: Multiple imputation[5] using a Markov chain Monte Carlo (MCMC) technique will be implemented to replace missing endpoint data using multiple draws from a posterior predictive distribution estimated from the placebo group only. For ACR50, assuming monotone missingness pattern, missing ACR50 at Weeks 2, 4, 12 and 16 will be imputed by first imputing the 7 components of the ACR50 criteria (SJC66, TJC68, subject's VAS assessment of pain, SGADA, IGADA, HAQ-DI, and CRP), then assigning ACR50

status from the imputed values. Predictors in the regression model for missing values at Week 2 will be the randomization stratification factors of prior TNFi use, and LEI>0/LEI=0, as well as sex (male/female), race (white/non-white), age (continuous), and BMI (continuous). For each of the 7 components, baseline value will be included as a continuous predictor. Predictors in the regression model for missing values after Week 2 will be all of these variables, plus the value of the component being predicted at all prior scheduled assessments. Missing ACR50 values in both the placebo group and three izokibep groups will be imputed using observed data from the placebo group only. Data will be processed sequentially by repeatedly calling SAS® PROC MI to impute missing outcomes. A total of 100 imputations will be imputed to form the complete data set with a pre-specified seed (6848251) to ensure reproducibility. These multiple imputed data sets are then analyzed by using the same method for the primary analysis for complete data as specified in Section 6.6.1. The final inference on treatment difference is conducted from the multiple datasets using Rubin's combining rules[6]. Other imputation, fully conditional specification method may be implemented if the multiple imputation model does not converge. Sample SAS code:

```
proc mi nimpute=100 seed=6848251;
  var acr50;
  ods select misspattern;
run;
```

VAS Supplemental Analyses: A supplemental primary analysis using the VAS data on a scale from 0 to 100 mm will be performed in order to evaluate the robustness of the results when variations in the collection of the VAS data were reported by the investigator. The impact of VAS collection on paper (single vertical line through the VAS at the point that best corresponds to their feeling of pain), VAS collection via verbal response from the subject, and VAS scales that that measured 100 mm will be examined by conducting the following analyses:

- Using only VAS data collected on a 0 to 100 scale (paper or verbal), excluding data collected on a 0 to 10 scale.
- Using only VAS data collected on a 0 to 100 horizontal line (paper only), excluding data collected verbally.
- Using only VAS data collected on a 0 to 100 scale (paper only, with line that measured 100 mm), excluding those collected with lines of a different length (0 to 10, 0 to <100mm, 0 to >100mm), excluding data collected verbally.

If there are no subjects who meet one of the protocol deviation criteria (data collected on a 0 to 10 scale, data collected verbally, or data collected using a line not 100 mm) then the corresponding subgroup analysis above will be omitted.

Q2W and Q4W Supplemental Analyses: A supplemental analysis is planned for subjects who were randomized to receive izokibep Q2W or izokibep Q4W. The analysis of subjects randomized to receive izokibep Q2W will group subjects according to the longest number of consecutive doses of izokibep or placebo that were received within the first 16 weeks of treatment of Period 1:

- Did not receive consecutive doses of izokibep or placebo
- Received consecutive doses, but no more than two consecutive doses of either izokibep or placebo
- Received at least three consecutive doses of izokibep or placebo, but not four consecutive doses.
- Received four or more consecutive doses of izokibep or placebo.

Because the number of subjects in the first group who did not receive consecutive doses of izokibep or placebo is expected to be small, the first two groups will be combined. The primary analysis will be run in each of the resulting three subsets of subjects, comparing each subset of subjects to the entire group of placebo subjects. Subjects will be classified according to doses actually received (not assigned), and included only if the subjects receive at least 8 total doses (izokibep or placebo) and have an assessment after the 8th dose (with other subjects excluded from this supplemental analysis). Informal comparisons of the resulting analyses (p-values, point estimates and treatment differences) will be used to assess whether the three subsets of subjects, with different PK characteristics, have similar responses.

All subjects who were randomized to receive Q2W and who received correct treatment at each administration until week 12 (or until discontinuation of study drug, whichever came first) will be summarized with week 12. The number and proportion of subjects achieving ACR50, as an additional supplemental analysis to understand the impact of dosing errors.

An analogous analysis of subjects who were randomized to receive izokibep Q4W will also be reported. This will classify such subjects into two subsets:

- Received no more than four consecutive doses of placebo
- Received more than four consecutive doses of placebo.

6.6.3. Secondary Efficacy Analysis

The following secondary endpoints will be analyzed as dichotomous endpoints and analogously to the primary endpoint:

- proportion of subjects achieving PASI90 at Week 16 in subjects with ≥3% BSA psoriasis at baseline
- proportion of subjects achieving resolution of enthesitis at Week 16 in subjects that had enthesitis (LEI > 0) at baseline
- proportion of subjects achieving MDA at Week 16
- proportion of subjects achieving ACR20 at Week 16

- proportion of subjects achieving improvement in PsAID of at least 3 units at Week 16 compared to baseline in subjects with PsAID ≥3 at baseline

Subjects who have missing data at Week 16 will be imputed as nonresponders and included in the analysis.

The secondary endpoint of change in physical function of HAQ-DI from baseline to Week 16 will be analyzed using the mixed-effects model for repeated measures (MMRM) that includes data at postbaseline visits up to the time point of interest. Subjects that have a baseline value and at least 1 post-baseline value are included in the analysis. The model includes the baseline physical function of HAQ-DI, treatment group, and stratification variables, visit week as fixed effects and subject being the random effect. An unstructured variance-covariance matrix will be used. Other variance-covariance matrices, such as autoregressive, and compound-symmetric will be considered in this order if the model does not converge. The Kenward-Roger method will be used to estimate the degrees of freedom. Missing change scores will not be imputed using the MMRM approach. The least squares (LS) means and 95% CIs of the difference in mean change from Baseline in HAQ-DI between each dose regimen of izokibep group and placebo group from MMRM will be provided. Sample SAS code is given as following:

```
proc mixed;
        class week subjectid treatment(ref = "placebo")
        model change_y = baseline_y treatment week prior_tnfi baseline_lei
        repeated week / subject = subjectid (treatment) TYPE = UN;
        lsmeans treatment * week /slice = week cl;
        slice treatment * week/ sliceby = week pdiff cl;
run;
```

6.6.4. Supplementary Analyses for the Secondary Efficacy Analysis

Sensitivity analyses will use analysis of covariance on change in physical function of HAQ-DI from baseline to Week 16, with stratification variables used for randomization also included as covariates, along with baseline physical function of HAQ-DI.

Sensitivity analysis for physical function of HAQ-DI will also use multiple imputations. Physical function of HAQ-DI missing as Week 16 will be replaced with the imputed value from the imputation dataset. The imputation dataset will be created from the placebo subjects who have complete data.

Multiple Imputation: Multiple imputations[7] will be implemented to replace missing endpoint data using multiple draws from a posterior predictive distribution estimated from the placebo group only. For HAQ-DI at Week 2, 4, 12, and 16, assuming monotone missingness

pattern, missing HAQ-DI will be imputed using previous HAQ-DI (except for Week 2) up to the time point of interest and randomization strata using placebo group only. Missing HAQ-DI in the placebo group and two izokibep groups will be imputed. Monotonicity in missing data will first be imputed with MCMC. The resulting data will then be processed sequentially by repeatedly calling SAS® PROC MI to impute missing outcomes. A total of 100 imputations will be imputed to form the complete data set with a pre-specified seed (36767) to ensure reproducibility. These multiple imputed data sets will then be analyzed by using the same method for the HAQ-DI endpoint analysis as specified in Section 6.6.3. The final inference on treatment difference is conducted from the multiple datasets using Rubin's combining rules[8]. Other imputation, fully conditional specification method may be implemented if the multiple imputation model does not converge.

6.6.5. Exploratory Analysis

Exploratory endpoints will be analyzed analogously to primary and secondary endpoints, with the exception of a prespecified alpha-controlling testing strategy. Exploratory endpoints of dichotomous measurements will be analyzed analogously to primary endpoints. Exploratory endpoints of continuous measurements will be analyzed analogously to the secondary endpoint of change in physical function of HAQ-DI from baseline to Week 16. P-values for exploratory endpoints will be considered descriptive and not considered conclusive. A summary table of all primary, secondary and exploratory endpoints will be presented using point estimates only, to visually assess izokibep 160 mg dosed QW, izokibep 160 mg dosed Q2W, izokibep 80 mg dosed Q4W and placebo.

All primary, secondary and exploratory endpoints defined for comparing izokibep 160 mg QW or izokibep 160 mg Q2W to placebo will also be reported for comparing izokibep 80 mg Q4W to placebo.

Dose-response will be investigated through analyses of the primary and secondary endpoints, which will use all four treatment arms (izokibep 160 mg QW, izokibep 160 mg Q2W, izokibep 80 mg Q4W and placebo) in a single analysis for each endpoint. A p-value for the global test of the hypothesis that all groups have identical response rates will be reported for each endpoint, not subject to type I error rate control or any multiple comparison procedure. P-values will also be reported for all pairwise comparisons for each endpoint. Dose-response relationships will be assessed qualitatively from these analyses. Due to the error in dose sequence in the Q2W and Q4W arms, additional investigations will include assessment of endpoints that may be achieved earlier, including ACR20 and PASI75 at week 4 (at which time all subjects received correct dose sequencing).

The following binary endpoints before or on Week 16 will be analyzed analogously to the primary endpoints of ACR50 at Week 16. The number and proportion of subjects for binary

endpoints in each treatment group will be reported without strata along with Clopper-Pearson 95% CI. Common risk difference with Cochran-Mantel-Haenszel weighting will be used to calculate a 95% CI, and p-values from the stratified test will be reported as well. The binary endpoints after week 16 will include number and proportion along with Clopper-Pearson 95% CI without strata according to the four randomized treatment groups. No p-value will be reported.

- ACR70 at Weeks 16, 24, and 52
- DAPSA ≤14 at Weeks 16, 24, and 52
- cDAPSA ≤14 at Weeks 16, 24, and 52
- DAPSA ≤4 at Weeks 16, 24, and 52
- cDAPSA ≤4 at Weeks 16, 24, and 52
- PASI75 at Weeks 16, 24, and 52
- PASI100 at Weeks 16, 24, and 52
- MDA at Weeks 24, and 52
- VLDA at Weeks 16, 24, and 52
- Resolution of dactylitis at Weeks 16, 24, and 52 among those with dactylitis at baseline
- ACR50 at Weeks 24 and 52
- PASI90 at Weeks 24 and 52 among those with ≥3% BSA psoriasis at baseline
- Resolution of enthesitis (LEI = 0) at Weeks 24 and 52 among those with LEI > 0 at baseline
- Resolution of enthesitis (SPARCC = 0) at Weeks 24 and 52 among those with SPARCC > 0 at baseline
- PsAID response at Weeks 24 and 52
- ACR20 at Weeks 24 and 52
- Occurrence of enthesitis (LEI > 0) at weeks 4, 8, 12 and 16, among subjects with no enthesitis (LEI = 0) at baseline
- Decrease of at least 2 in LEI, among subjects with baseline LEI of at least 2, at weeks 4, 8, 12 and 16.

The following continuous endpoints will be analyzed at week 16 with an MMRM model as specified in sec 6.6.3, and after week 16 with standard summary statistics for continuous endpoints. LS mean changes and associated SEs with associated 95% CI and p-values will be presented. Models through week 16 will include all data up to Week 16. The following continuous endpoints after Week 16 will be analyzed according to the four randomized treatment groups. Mean changes and associated SEs along with 95% CI according to the four randomized treatment groups will be presented without strata. No p-values will be reported.

- Change from baseline in DAS28-CRP at Weeks 16, 24, and 52
- Change from baseline in DAPSA/cDAPSA at Weeks 16, 24, and 52
- Change from baseline in PASDAS at Weeks 16, 24, and 52

- Change from baseline in mNAPSI at Weeks 16, 24, and 52 in those with nail psoriasis at baseline
- Change from baseline in mTSS at Weeks 16 and 52
- Change from baseline in SPARCC enthesitis score at Weeks 16, 24, and 52 in those with enthesitis at baseline (SPARCC>0)
- Change from baseline in spinal pain NRS at Weeks 16, 24, and 52
- Change from baseline in spine pain NRS in those with sacroiliitis on pelvic x-ray
- Change from baseline in spine pain NRS in those with investigator assessed spinal involvement
- Correlation of investigator assessed spinal involvement with sacroiliitis on x-ray
- Change from baseline in SF-36 total score, PCS, and MCS at Weeks 16, 24, and 52
- Change from baseline in FACIT-F at Weeks 16, 24, and 52
- Change from baseline in LEI enthesitis score at Weeks 16, 24, and 52 in those with enthesitis at baseline (LEI>0)
- HAQ-DI change from baseline to Weeks 24 and 52
- HADS change from baseline to Week 24 and 52

Change from baseline for the individual components of the ACR and MDA scores will be summarized at Week 16. Summary statistics (n, mean, standard deviation, median, minimum, and maximum) will be provided

- Swollen Joint Count
- Tender Joint Count
- SAPVAS
- SGADA
- IGADA
- HAQ-DI
- CRP
- PASI
- BSA
- Tender entheseal points ≤ 1 site out of 6 sites included in LEI

The following will be analyzed with a time to event model, using a log rank test with stratification by the factors used to stratify randomization. Subjects who do not achieve the endpoint at or before the week 16 visit will be censored at the time of the last assessment could be derived at or before week 16.

- Time to resolution of enthesitis (LEI = 0) among subjects with enthesitis (LEI = 0) at baseline.
- Time to first achievement of ACR50

- Time to first achievement of PASI90 among subjects with ≥3% BSA psoriasis at baseline
- Time to first achievement of MDA
- Time to first achievement of ACR20
- Time to first achievement of PsAID improvement of at least 3 among subjects with PsAID ≥3 at baseline
- Time to first achievement of PASI100 among subjects with PsAID ≥3 at baseline

WPI will be calculated at baseline. WPI is a scale of 0-19, and is a count of how many of 19 body sites the subject reports pain or tenderness in the prior 7 days. Standard summary statistics for binary data will be calculated by baseline WPI (WPI ≥7 vs. <7) for the following endpoints:

- ACR50 at week 16 (observed responder, observed non-responder, missing)
- LEI responders at week 16 (observed responder, observed non-responder, missing)PsAID at week 16 (observed responder, observed non-responder, missing)
- HAQ-DI at week 16: improvement of at least 0.35 (observed responder, observed non-responder, missing)

Area under the curve (AUC) for ACR-N will be calculated for each subject by triangulating the observed values at baseline (by definition ACR-N = 0 at baseline) and weeks 2, 4, 12 and 16. The area for consecutive visits, using nominal time from Week t and Week u, will be calculated as

$$AUC_{\text{Week t to Week u}} = ((u - t) \times (\text{ACR-N}_{\text{Week t}} + \text{ACR-N}_{\text{Week u}})/2$$

The AUC for the period through week 16 will be the sum of the individual values from week 0 through week 16. Any visits at which ACR-N cannot be calculated (including Week 0) will be imputed as ACR-N = 0. In case one or more components of ACR is missing, the smallest value N for which ACR-N is known to be achieved will be used.

Listings of all efficacy data, including derived composite values will be provided for the Full Analysis Set.

PK concentrations will be summarized by dose group and scheduled visit. Summary statistics reported will be number of samples, number of samples that are BLQ, arithmetic mean, standard deviation, coefficient of variation in percent (CV), minimum, median, maximum, geometric mean and the CV of the geometric mean in percent.

6.6.6. Subgroup Analysis

Various subgroup analyses of the efficacy data will be reported. These analyses will investigate the treatment response within specific subgroups of interest, and assess whether the treatment response is consistent across different subgroup levels. All analyses will be performed on the FAS data set. Treatment effect will be summarized by comparison (QW versus placebo; Q2W versus placebo) and presented in tables and a forest plot.

Summaries of primary and secondary efficacy endpoints will be presented by subgroup:

The following subgroups are of interest for the primary and all secondary endpoints:

- Race (Asian, White, Black or African American, other)
- Sex (male, female)
- Age (< 65, ≥ 65)
- Prior TNFi use (Yes/No)

The following subgroups are also of interest for ACR50 at Week 16 and all secondary endpoints:

- Any csDMARD at Baseline (Yes/No)
  - MTX only
  - MTX + other csDMARD
  - Non-MTX csDMARD only
  - No csDMARD at baseline
- Any prior TNFi exposure (Yes/No)
- Prior TNFi failure (Yes/No)
- Number of prior failed biologics group (0, 1, 2 or more)
- Geographic region
  - North America (US or Canada)
  - Europe (Bulgaria, Czech Republic, Spain, Hungary, Poland or Germany)
  - Non-US (Bulgaria, Czech Republic, Spain, Hungary, Poland, Germany or Canada)
- Country
  - United States
  - Canada
  - Bulgaria
  - Czech Republic
  - Poland
  - Hungary
  - Spain
  - Germany
- BMI (≤ 18, > 18 - ≤ 25, > 25 - ≤ 30, > 30 - ≤ 35, > 35)
- Smoking status (never, former, current)

- Enthesitis at baseline
  - LEI = 0
  - LEI > 0
  - LEI = 1
  - LEI = 2
  - LEI = 1 or 2
  - LEI < 3
  - LEI ≥ 3
  - LEI < 4
  - LEI ≥ 4
- Combination of enthesitis and prior TNFi exposure
  - LEI = 1 or 2 and no prior TNFi exposure
  - LEI ≥ 3 and no prior TNFi exposure
  - LEI = 1 or 2 and prior TNFi exposure
  - LEI ≥ 3 and prior TNFi exposure

The following endpoints are of interest for the subjects with enthesitis at baseline (LEI > 0) and subjects without enthesitis (LEI = 0) at baseline:

- PsAID response at Week 16
- PsAID change from baseline to Week 16
- HAQ-DI response at Week 16
- HAQ-DI change from baseline to Week 16
- Change in DAPSA and cDAPSA from baseline to Week 16
- VLDA at Week 16
- Change in SF-36 total score, physical component summary (PCS), and mental component summary (MCS) from baseline to Week 16
- MDA at Week 16 (including each component of MDA)

The following subgroups are of interest for summarizing Time to resolution of enthesitis (LEI = 0) among subjects with enthesitis (LEI = 0) at baseline.

- Sex (male, female)
- Prior exposure to TNFi (yes, no)
- LEI score at baseline
  - LEI = 1
  - LEI = 2
  - LEI = 1 or 2
  - LEI ≥3

The following subgroup is of interest for the endpoints of PASI75, PASI90 and PASI100:

- Baseline BSA 10% or greater

Response rates and difference in response rate will be reported for each subgroup. Cochran-Mantel-Haenszel weighting will be used to calculate a 95% CI for subgroup analyses for each endpoint similar to Section 6.6.1 and 6.6.3 using stratified tests. A forest plot for the difference in response rate and 2-sided 95% CI will also be presented. No p-values will be reported for the subgroup analyses.

PsAID and HAQ-DI response rates will be determined by the proportion of subjects who achieve at least a minimum clinically important difference (MCID) improvement from baseline to Week 16. MCID values are 3 units for PsAID and 0.35 units for HAQ-DI.

Similar to the subgroup analysis above, difference in response rate and 2-sided Cochran-Mantel-Haenszel weighted stratified 95% CI will be reported for subgroup analyses for each binary endpoint.

6.7. Safety

For all Safety Tables, data will be summarized as follows, unless otherwise specified:

- Period 1 will include the first 16 weeks, summarized by the randomized treatment arms (Placebo QW; Izokibep QW; Izokibep Q2W, Izokibep Q4W)
- Period 2 will include data after Week 16 visit, summarized by the randomized treatment groups (Placebo - Izokibep QW; Izokibep QW; Izokibep Q2W, Izokibep Q4W)
    - Selected AE summaries will be presented as: Period 2 will include data after Week 16 visit, summarized by Placebo - Izokibep QW; Izokibep QW (those who receive QW during any part of period 2); Izokibep Q2W (those who receive Q2W during any part of period 2); Izokibep Q4W (those who receive Q4W during any part of period 2); Izokibep QW only (those who receive only izokibep QW during period 2); Izokibep Q2W only (those who receive only izokibep Q2W during period 2); Izokibep Q4W only (those who receive only izokibep Q4W during period 2)

6.7.1. Extent of Exposure

Exposure will be reported through the planned collection of the primary endpoint and for the entire study for the Safety Analysis Set. Exposure duration will also be summarized and will be calculated as (last dose – first dose + 1). Descriptive statistics will be presented for duration of exposure.

A listing of overall treatment exposure data, including the first and last dates of treatment will be presented together with compliance for the Safety Analysis Set. Further, study treatment administration data will be listed for the Safety Analysis Set.

6.7.2. Adverse Events

All SAEs will be collected from the signing of the ICF until the 8-week follow-up visit. All AEs will be collected from the first dose of study drug until 4 weeks after the last dose of study drug. All AEs recorded on the eCRF will be coded using the latest MedDRA dictionary around the time of Database Lock and classified as either AEs or treatment–emergent AEs (TEAEs) as follows:

- TEAEs are either events with start date and time after the start of the Treatment Period and up to 4 weeks after the end of the Treatment Period, or events with start date and time prior to the start of the Treatment Period whose severity worsens on or after the start of the Treatment Period. TEAEs for period 2 are either events with start date and time after the start of administration of Week 16 treatment and up to 4 weeks after the end of the Treatment Period, or events with start date and time prior to the start of administration of Week 16 treatment whose severity worsens on or after the start of administration of Week 16 treatment. For subjects who receive treatment in period 2, TEAEs with start date after the start of period 2 will be assigned to period 2 only, not to period 1.
- Treatment-Emergent Serious AEs (TESAEs) will be defined as TEAEs which meet the criteria as Serious = "Yes".
- The relationship between a TEAE and treatment is assessed as related or not related. A treatment-related TEAE will be defined as a TEAE considered by the investigator as related to treatment or with unknown/missing relationship to treatment.
- Assessment of AE severity will be based on eCRF AEs form.
- TEAEs leading to discontinuation of treatment are defined as TEAEs where "Action Taken with Study Treatment" is indicated as "Drug Withdrawn".

In addition to the aforementioned AE types, the following TEAEs of special interest will be monitored:

- Candida infection,
- Inflammatory bowel disease,
- Suicidal ideation.
- Malignancies
- MACE (major adverse cardiovascular and cerebrovascular events: cerebrovascular accident and transient ischemic attack, non-fatal myocardial infarction or unstable angina, cardiovascular death)
- Tuberculosis (TB)
- Infections (serious, opportunistic, or fungal only)
- Cytopenias

- Systemic hypersensitivity reaction

Adverse events of special interest will be identified manually. After all data are entered into the database and before database unblinding, a list of all unique PTs will be generated. ACELYRIN pharmacovigilance and clinical development representatives will review the list and determine which are TEAEs of special interest in each category listed above. This determination will be documented in a memo, signed and dated before database lock, and stored in the study trial master file. The database will not be unblinded until the memo is finalized and signed.

Adverse events will be summarized by default descriptive summary statistics for categorical variables for the Safety Analysis Set as follows:

- An overview of TEAEs including the number and percentage of subjects with at least one of each mentioned TEAE type according to both Period 2 groupings as described in section 6.7:
    - Any TEAE
        - Leading to discontinuation of study treatment
        - Mild (Grade 1)
        - Moderate (Grade 2)
        - Severe (Grade 3)
        - Life-threatening (Grade 4)
        - Death (Grade 5)
        - Any TEAE
            - Grade 2 or higher
            - Grade 3 or higher
            - Grade 4 or higher
    - Any study treatment related TEAE
        - Leading to discontinuation of study treatment
    - Any serious TEAE
        - Leading to discontinuation of study treatment
        - Leading to death
    - Any serious study treatment related TEAE
        - Leading to discontinuation of study treatment
        - Leading to death
- The number and percentage of subjects reporting each TEAE and the count of number of events will be summarized by SOC and PT for the following types of TEAEs.
    - TEAEs
    - TEAEs of Special Interest
    - TEAEs Leading to Discontinuation of Study Treatment

  - TEAEs leading to interruption of study treatment
  - TEAEs by Maximum Severity
  - TEAEs by Relationship to Treatment
  - TEAEs by Relationship and Maximum Severity
  - Study Treatment related TEAEs
  - Study Treatment Related TEAEs Leading to Discontinuation of Study Treatment
  - Serious TEAEs
  - Serious TEAEs Leading to Discontinuation of Study Treatment
  - Serious TEAEs Leading to Death
  - Study Treatment Related Serious TEAEs
  - Study Treatment Related Serious TEAEs Leading to Discontinuation of Study Treatment
  - Study Treatment Related Serious TEAEs Leading to Death
- The number and percentage of subjects reporting each TEAE and the count of number of events will be summarized by SOC and PT for two subgroups:
  - Subjects < 65 years of age at baseline
  - Subjects ≥ 65 years of age at baseline
- The number and percentage of subjects who died will be summarized
- The number and percentage of subjects reporting each TEAE and the number of events will be summarized by PT for all TEAEs
- The number and percentage of subjects reporting each TEAE and the number of events will be summarized by PT for TEAEs of Special Interest
- The number and percentage of subjects reporting each TEAE and the count of number of events according to the 5 Period 2 groupings as described in section 6.7 will be summarized by SOC and PT for the following types of TEAEs.
  - All TEAEs
  - Serious TEAEs

In the above summaries, subjects with more than one TEAE within a particular SOC are counted only once for that SOC. Similarly, subjects with more than one TEAE within a particular PT are counted only once for that PT.

For summaries by maximum severity, subjects with multiple TEAEs within a particular SOC or PT will be counted under the category of their most severe TEAE within that SOC or PT. TEAEs with missing severity will be included in the counts as severe (Grade 3) unless the TEAE has an outcome of fatal; in that case, the TEAE with missing severity will be included in the counts as death (Grade 5).

Summaries by SOCs and PTs will be sorted by SOCs by their Internationally Agreed Order (MedDRA) and PTs within SOC by descending order of total incidence. Where PTs tie, PTs will be sorted alphabetically.

All AE data will be listed and AEs and TEAEs will be presented together. Treatment-emergence status will be flagged in the listing. The listing will present the relative start and stop day of the AE calculated relative to the first dose of treatment and will be presented for those subjects who received at least one dose of treatment. If the AE is "Ongoing" it will be indicated as such in the listing and the relative stop day will not be calculated. Only original dates will be presented in the listing even though the relative day may be based on an imputed date.

In addition, the following listings will be presented:

- Listing of Deaths
- Listing of TEAEs
- Listing of Serious TEAEs
- Listing of AEs Leading to Interruption of Study Treatment
- Listing of AEs Leading to Discontinuation of Study Treatment
- Listing of AEs of Special Interest

6.7.3. Laboratory Evaluations

Data for the following hematology, clinical chemistry, routine urinalysis, pregnancy testing, other screening tests, and other tests analytes received from the central laboratory are to be measured at the scheduled timepoints (Baseline, Week 4, 8, 12, 16, 20, 24, 32, 36, 40, 52 and 59).

In accordance with the baseline value definition in Section 5.1.2, the change from baseline will be derived as follows:

Change (unit) = (post-baseline value – baseline value)

All laboratory data will be reported in SI units. All quantitative laboratory test values at each assessed timepoints will be compared with the relevant reference range in SI units and categorized as:

- Low: Below the lower limit of the reference range.
- Normal: Within the reference range (upper and lower limits included).
- High: Above the upper limit of the reference range.

For summaries which present the worst value with respect to the reference range at the subject level, low and high are each chosen in preference to normal values. For parameters with both

low and high reference ranges, subjects who have assessments within both low and high ranges will be counted within each category for worst value summary tables.

For analysis purposes, values preceded by a "<" or a ">" sign (i.e. those below or above the limits of quantification) will be considered equal to the lower or upper limit of quantification, respectively.

Laboratory values will be assigned toxicity grades, when available, using criteria based on the National Cancer Institute – Common Terminology Criteria for Adverse Events (NCI CTCAE) version 5.0 or higher scale.

For those parameters which are graded with two toxicities such as potassium (hypokalemia/hyperkalemia), the toxicities will be summarized separately. Low direction toxicity (e.g., hypokalemia) grades at baseline and post baseline will be set to 0 when the variables are derived for summarizing high direction toxicity (e.g., hyperkalemia), and vice versa.

Laboratory data will be summarized by default descriptive summary statistics for continuous and categorical variables for the Safety Analysis Set:

- Observed values and change from baseline at each assessed timepoint for each standard continuous laboratory parameter
- Number and percentage of subjects with categorized shift (low, normal, and high) values relative to the reference range at baseline compared to each post-baseline timepoint for hematology and clinical chemistry
- Number and percentage of subjects with worst categorized (low, normal, and high) values relative to the reference range
- Number and percentage of subjects with NCI-CTCAE Grade 1, 2, 3, 4, 3/4, and any grade during post-baseline
- Number and percentage of subjects with categorized shift NCI-CTCAE toxicity values at baseline compared to worst post-baseline result

Listings of all clinical laboratory data including derived change from baseline will be provided for the Safety Analysis Set. Within each listing, laboratory values outside the normal ranges will be flagged as either high or low.

6.7.3.1. Liver toxicity review

Potential liver toxicity will be assessed by listing subjects who met potential drug induced liver injury (DILI) criteria as described in the protocol and by summarizing subjects who met Hy's Law.

Subjects who met any (one or more) of the following criteria will be included in a listing for consideration of DILI:

- ALT or AST ≥5 x ULN at any time
- ALT or AST ≥3 x ULN at all (2 or more) assessments in a span of ≥4 weeks
- ALT or AST ≥3 x ULN and total bilirubin ≥2 x ULN (>35% direct bilirubin) simultaneously at any time
- ALT or AST ≥3 x ULN and international normalized ratio (INR) ≥1.5 simultaneously at any time
- ALT or AST ≥3 x ULN at any time during the study and weekly assessments after this finding are not available for at least 4 weeks
- ALT or AST ≥3 x ULN associated with symptoms (new or worsening) believed to be related to liver injury or hypersensitivity at any time

To determine subjects who meet the last criterion, a listing of all subjects who have at least one ALT or AST ≥3 x ULN at any time, and who do not meet any of the first five criteria, will be reviewed by qualified medical professionals along with a listing of all AEs reported by such subjects. In coordination with the investigator, the medical professionals will assess whether any AE is believed to be related to liver injury or hypersensitivity and, if so, whether the elevated ALT or AST is in temporal association with the AE. Any subjects who meet these criteria will be included on a memo that will be filed in the TMF prior to database lock and unblinding, and included in the listings. For subjects who meet any of the six criteria, all liver function test results at any time during the study will be listed, sorted by subject and date. Listing will include subject, treatment assignment, ALT, AST, bilirubin and INR, along with specific criterion met.

Subjects who meet the criteria for Hy's Law will be displayed in shift tables. Hy's Law is met when ALT or AST is ≥3 x ULN, total bilirubin ≥2 x ULN and ALP is <2 x ULN, with all occurring simultaneously (at the same draw). At each assessment, subjects will be classified as meeting Hy's Law or not meeting Hy's Law. The shift tables will be presented as follows:

- Period 1
  - Baseline (meet Hy's Law / not meet Hy's Law)
  - Any time during period 1 (meet Hy's Law at least one time / not meet Hy's Law at any time)
- Period 2
  - Baseline prior to period 2 (meet Hy's Law / not meet Hy's Law)
  - Any time during period 2 (meet Hy's Law at least one time / not meet Hy's Law at any time)

A listing of all liver function test results from all subjects who meet Hy's Law at least one time during the study will be produced, sorted by subject and date. This listing will show subject, treatment assignment, period, ALT, AST, ALP, bilirubin and INR.

6.7.4. Vital Signs

The analyses described below will be conducted for the following vital signs assessments respectively:

- SBP (mmHg)
- DBP (mmHg)
- Heart Rate (bpm)
- Respiratory Rate (breaths/min)
- Body Temperature (°C)

In accordance with the baseline value definition in Section 5.1.2, the change from baseline will be derived as follows:

Change (unit) = (post-baseline value – baseline value)

The following will be summarized for the Safety Analysis Set:

- Observed values and change from baseline at each assessed timepoint for each standard vital sign parameter using default summary statistics for continuous variables.

A listing of all vital signs data including derived change from baseline will be provided for the Safety Analysis Set.

6.7.5. Electrocardiograms

The following electrocardiogram (ECG) assessments will be taken at baseline:

- An overall assessment classified as normal or abnormal
- Abnormal assessments will be further classified as clinically significant or not clinically significant

At baseline assessment, the subject will be classified as abnormal if any (one or more) of the replicate assessments are classified as abnormal. Similarly, the abnormal assessment will be classified as clinically significant if any (one or more) abnormal assessments are classified as clinically significant. The subject will be classified based on the worst finding of triplicate 12-Lead ECG. A list of ECG assessments will be presented.

6.7.6. Physical Examination

Clinically significant findings observed prior to the first dose of study drug will be entered in the CRF and reported as medical history. Clinically significant findings observed after the first dose of study drug will be reported as AEs in the CRF and reported as AEs.

Physical examination findings (normal/abnormal/clinically significant) and details of abnormalities will be listed only for each subject at each assessment.

6.7.7. Interim Analysis and Data Monitoring

An independent Data Monitoring Committee (DMC) will be established to monitor data on an ongoing basis to ensure the continuing safety of the subjects enrolled in this study.

The DMC will consist of two clinicians and a biostatistician that, collectively, has experience in the management of subjects with Psoriatic Arthritis and in the conduct and monitoring of randomized clinical trials. The DMC responsibilities, authorities, and procedures will be documented in the DMC charter. The committee will meet approximately every 6 months to review interim data. After each review, the DMC will make recommendations regarding the continuation of the study based on safety.

The DMC will review unblinded efficacy data when approximately 120 subjects have had the opportunity to complete the Week 12 visit. Data from the first 105 subjects randomly assigned to receive izokibep 160 mg QW, izokibep 160 mg Q2W or placebo will be used for these DMC summaries. The DMC may make a recommendation on further development activities for izokibep based on this data review. The DMC will not be empowered to recommend stopping this study or changing the sample size due to a demonstration of positive efficacy. The study team operationalizing the day-to-day of the study, the subjects, and the investigators will remain blinded to these interim results until after the study is completed.

No interim analyses will be conducted for the purpose of stopping the study or altering the study design.

The primary analysis of primary and secondary endpoints will occur after all subjects have had an opportunity to complete the Week 16 assessments (have completed the Week 16 assessments, or have completed a subsequent timepoint assessment, or permanently discontinued the study before Week 16). For the primary and secondary objectives, this will include all data for all subjects for the final analysis, and so it will be the final analysis. The final analysis will be conducted at the end of the study when all subjects have completed the study or permanently discontinued the study. For the primary endpoint, all available data at visits after Week 16 will be included in the summaries at these timepoints, but some

subjects will be ongoing and may have data collected after the primary analysis is reported. Data collected at subsequent timepoints may be summarized when all data at that timepoint have been collected. All data collected for the primary endpoint for all subjects through the end of the study will be reported in a subsequent analysis, but the p-values from each comparison of a regimen of izokibep to placebo in this subsequent analysis will not be used for any alpha-preserving hypothesis testing. Investigators, site personnel, and all sponsor personnel who have contact with subjects will remain blinded to individual subject treatment assignments until the final analysis.

When the primary efficacy analysis is conducted, safety data will be summarized. Summaries will include all safety data through Week 16 and all available safety data on all subjects through the end of the study. All safety data will be summarized through the end of the study when all subjects have completed all visits.

6.8. Pharmacokinetic Assessments

PK concentrations will be summarized by dose group and scheduled visit. Summary statistics will be reported with the number of samples, number of samples that are BLQ, arithmetic mean, standard deviation, coefficient of variation in percent (CV), minimum, median, maximum, geometric mean and the CV of the geometric mean in percent. For the calculation of summary statistics, predose concentrations that are below the limit of quantification (BLQ) will be treated as zero and post-dose concentrations that are BLQ will be treated as missing.

Plots of arithmetic mean time courses by dose group will be provided using linear and semi-logarithmic scale.

6.9. Anti-Drug Antibodies

Samples tested for ADAs will be reported as positive or negative based on a non-domain specific confirmatory assay. Samples that are positive, and are further tested, will be reported as positive or negative for the IL-17A binding domain, positive or negative for the albumin binding domain, and/or positive or negative as a neutralizing ADA. Any continuous values of concentration (titers) will be summarized using observed data.

Subjects will be summarized at each scheduled visit according to the confirmatory test, with number (%) positive. Positive samples on the confirmatory test, that are further tested, will be reported according to each specific binding domain, with number (%) positive for the IL-17A binding domain, number (%) positive for the albumin binding domain, and number (%) positive for either (one or both) binding domain. Samples that are further tested on a neutralizing antibody assay will be reported as number (%) positive for neutralizing activity, with percentage calculated using the number that were positive according to the screening

test as the denominator. In addition, incidence persistent ADA will be summarized for IL-17A and albumin binding domain. Persistent ADA is characterized as treatment-induced ADA detected at two or more sampling time points during the treatment irrespective of any negative samples in between (including follow-up period if any).

7. Appendices

Appendix 1: Document History

Document Version, Status, Date	Summary/Reason for Changes
Version 1.0, Final, 21JUN2023	Not applicable; First final version

Appendix 2: End Points Definitions and Scoring Method

Appendix 2.1 Health Assessment Questionnaire Disability Index (HAQ-DI)[9]

- The Disability Index

The eight categories assessed by the Disability Index are 1) dressing and grooming (2 items), 2) arising (2 items), 3) eating (3 items), 4) walking (2 items), 5) hygiene (3 items), 6) reach (2 items), 7) grip (3 items), and 8) activities (3 items).

- Scoring Conventions for the Disability Index

There are four possible responses for the Disability Index questions:

WITHOUT ANY DIFFICULTY = 0 WITH SOME DIFFICULTY = 1

WITH MUCH DIFFICULTY =2 UNABLE TO DO =3

The highest score for questions in each category (range 0 to 3) determines the score for the category. If either AIDS OR DEVICES and/or HELP FROM ANOTHER PERSON are checked for a category, the score is set to "2", unless the score is already "3" (i.e., scores of "0" or "1" are increased to "2"). For example, if the highest score for the dressing category is "1", and the patient says they use a device for dressing, the computed category score would be "2". The sum of the computed categories scores is then calculated and divided by the number of categories answered. This gives a score in the 0 to 3 range. A disability index cannot be computed if the patient does not have scores for at least six (6) categories

The following list details how each aid and device is associated with the category scores:

HAQ-DI Category	AIDS OR DEVICES	HELP FROM ANOTHER PERSON
Dressing & grooming	Devices used for Dressing (button hook, zipper pull, etc.)	Dressing and grooming

Arising	Special or built up chair	Arising
Eating	Built up or special utensils	Eating
Walking	Crutches, Cane, Wheelchair, Walker	Walking
Hygiene	Raised toilet seat, Bathtub bar, Bathtub seat, Long-handled appliances in bathroom	Hygiene
Reach	Long-handled appliances for reach	Reach
Grip	Jar opener (for jars previously opened)	Gripping and opening things
Activities		Errands and chores

Appendix 2.2 36-Item Short Form Survey (SF-36)[10]

The SF-36 questionnaire consists of eight scales yielding two summary measures: physical and mental health. The physical health measure includes four scales of physical functioning (10 items, PF), role limitations due to physical health (4 items, RP), bodily pain (2 items, BP), and general health (5 items, GH). The mental health measure is composed of vitality (4 items, VT), social functioning (2 items, SF), role limitations due to emotional problems (3 items, RE), and mental health (5 items, MH).

Two versions of the SF-36 were used in this study. SF-36v1 and SF-36v2®. A small subset of subjects received v1. For these subjects, coding will be done programmatically as follows.

- Step 1: Recoding Items

Item numbers	**Change original response category ***	**To recoded value of:**
1, 2, 20, 22, 34, 36	1	100
	2	75
	3	50
	4	25
	5	0
3, 4, 5, 6, 7, 8, 9, 10, 11, 12	1	0
	2	50
	3	100
13, 14, 15, 16, 17, 18, 19	1	0
	2	100
21, 23, 26, 27, 30	1	100
	2	80
	3	60
	4	40

	5	20
	6	0
24, 25, 28, 29, 31	1	0
	2	20
	3	40
	4	60
	5	80
	6	100
32, 33, 35	1	0
	2	25
	3	50
	4	75
	5	100

* Precoded response choices as printed in the questionnaire.

- Step 2: Averaging Items to Form Scales

Scale	**Number of items**	**After recoding per Step 1, average the following item**
Physical functioning	10	3 4 5 6 7 8 9 10 11 12
Role limitations due to physical health	4	13 14 15 16
Role limitations due to emotional problems	3	17 18 19
Energy/fatigue/vitality	4	23 27 29 31

Emotional well-being/mental health	5	24 25 26 28 30
Social functioning	2	20 32
Bodily pain	2	21 22
General health	5	1 33 34 35 36

The SF-36 questionnaire consists of eight scales yielding two summary measures: physical and mental health. The physical health measure includes four scales of physical functioning (10 items, PF), role limitations due to physical health (4 items, RP), bodily pain (2 items, BP), and general health (5 items, GH). The mental health measure is composed of vitality (4 items, VT), social functioning (2 items, SF), role limitations due to emotional problems (3 items, RE), and mental health (5 items, MH).

From the individual subscales, two component summary scores are generated for physical (PCS) and mental health (MCS). The mean of the first five subscales (PF, RP, BP, GH, and VT) produce the PCS and the mean of the last five subscales (GH, VT, SF, RE, MH) produce the MCS; the GH and VT subscales overlap between the two overall components.

Items that are left blank (missing data) are not taken into account when calculating the scale scores. Hence, scale scores represent the average for all items in the scale that the respondent answered.

Scoring of the SF-36v.2 will be done using Quality Metrics[10].

Abbreviated Item Content for the SF-36v2® Health Survey Health Domain Scales

Scale	Item	Abbreviated Item Content
Physical Functioning (PF)	3a	Vigorous activities, such as running, lifting heavy objects, or participating in strenuous sports
	3b	Moderate activities, such as moving a table, pushing a vacuum cleaner, bowling, or playing golf
	3c	Lifting or carrying groceries
	3d	Climbing several flights of stairs
	3e	Climbing one flight of stairs
	3f	Bending, kneeling, or stooping
	3g	Walking more than a mile
	3h	Walking several hundred yards

	3i	Walking one hundred yards
	3j	Bathing or dressing oneself
Role-Physical (RP)	4a	Cut down the amount of time one spent on work or other activities
	4b	Accomplished less than you would like
	4c	Limited in kind of work or other activities
	4d	Had difficulty performing work or other activities (e.g., it took extra effort)
Bodily Pain (BP)	7	Intensity of bodily pain
	8	Extent pain interfered with normal work
General Health (GH)	1	Is your health: excellent, very good, good, fair, poor
	11a	Seem to get sick a little easier than other people
	11b	As healthy as anybody I know
	11c	Expect my health to get worse
	11d	Health is excellent
Vitality (VT)	9a	Feel full of life
	9e	Have a lot of energy
	9g	Feel worn out
	9i	Feel tired
Social Functioning (SF)	6	Extent health problems interfered with normal social activities
	10	Frequency health problems interfered with social activities
Role-Emotional (RE)	5a	Cut down the amount of time spent on work or other activities
	5b	Accomplished less than you would like
	5c	Did work or other activities less carefully than usual
Mental Health (MH)	9b	Been very nervous
	9c	Felt so down in the dumps that nothing could cheer you up
	9d	Felt calm and peaceful
	9f	Felt downhearted and depressed
	9h	Been happy
Reported Health Transition (HT)	2	How health is now compared to 1 year ago
An excerpt from the User's Manual for the SF-36v2 Health Survey		

Appendix 2.3 Functional Assessment of Chronic Illness Therapy-Fatigue (FACIT-F)[11]

For any FACIT measure, items are scored on a 0 – 4 response scale with anchors ranging from "Not at all" to "Very much so". To score the FACIT-F, scores are calculated by first reversing negatively stated-items (subtracting the response from '4') and then summing the reverse scores. A total score is then derived by summing all scores.

In cases where individual items are missing, subscale scores can be prorated using the average of the other answers in the scale. This is acceptable as long as more than 50% of the items were answered in the subscale (e.g., a minimum of 7 of 13 items.).

Appendix 2.4 Disease Activity Score in 28 joints-C-reactive protein (DAS28-CRP)[12]

The DAS28-CRP considers 28 tender and swollen joint counts, general health (subject's global assessment of disease activity (SGADA) using a 100 mm visual analogue scale (VAS) with 0=best, 100=worst), plus levels of an acute phase reactant (CRP (mg/L)).

DAS28 values were calculated as follows:

$$DAS28\text{-}CRP = 0.56*\sqrt{(TJC28)} + 0.28*\sqrt{(SJC28)} + 0.014* SGADA + 0.36*\ln(CRP+1) + 0.96;$$

where TJC=tender joint count and SJC=swollen joint count;

SGADA = subject's global assessment of disease activity on a 0-100 VAS.

If one or more components are missing in the formula, then DAS28-CRP is set to be missing.

Appendix 2.5 Disease Activity index for Psoriatic Arthritis (DAPSA)[13]

DAPSA is the numerical sum of 66 swollen and 68 tender joint counts, pain, SGADA (pain and SGADA scored on a 0–10 scale) and CRP (mg/dL))

DAPSA = TJC68 + SJC66 + SGADA/10 + SAPVAS/10 + CRP;

where TJC=tender joint count and SJC=swollen joint count;

ln = natural logarithm;

SGADA = subject's global assessment of disease activity on a 0-100 VAS;

SAPVAS = subject's global assessment of pain on a 0-100 VAS;

CRP = C-reactive protein measured in mg/dL.

The cDAPSA, which was derived from the DAPSA, includes all DAPSA components but excludes CRP to obtain a fully clinical score.

A DAPSA, cDAPSA score of ≤14 represents low disease activity, and a score of ≤4 represents remission.

Appendix 2.6 Psoriatic Arthritis Disease Activity Score (PASDAS)[14]

The PASDAS is calculated using the following formula:

PASDAS = (0.18√IGADA) + (0.159√SGADA) – (0.253√SF-36PCS) + (0.101*ln(SJC66+1)) + (0.048*ln(TJC68+1)) + (0.23*ln(LEI+1)) + (0.37*ln(LDI+1)) + (0.102*ln((CRP*10) +1) +2) * 1.5

where swollen and tender joint counts can range from 0 to 66 and 68 joints, respectively; and CRP = C-reactive protein measured in mg/L. The PASDAS score ranges from 0-10 with higher scores indicating worse disease activity.

where TJC=tender joint count and SJC=swollen joint count;

ln = natural logarithm; all VAS scores are from 0-100mm;

IGADA = investigator's global assessment of disease activity;

SGADA = subject's global assessment of disease activity;

PCS = physical component summary scale of the SF-36;

LEI = Leeds enthesitis index;

LDI = Leeds dactylitis index;

CRP = C-reactive protein measured in mg/dL.

The score range of the PASDAS is 0 to 10, with worse disease activity represented by higher scores.

Appendix 2.7 Psoriatic Arthritis Impact of Disease Scoring and Calculation Rules (PsAID)[15]

The PsAID is calculated based on 9 Numerical rating scales (NRS) questions. Each NRS is assessed as a number between 0 and 10.

1. Calculation

PsAID final value = (PsAID1 (pain) NRS value (range 0-10) x 0.174) + (PsAID2 (fatigue) NRS value (range 0-10) x 0.131) + (PsAID3 (skin) NRS value (range 0-10) x 0.121) + (PsAID4 (Work and/or leisure activities) NRS value (range 0-10) x 0.110) + (PsAID5 (function) NRS value (range 0-10) x 0.107) + (PsAID6 (discomfort) NRS value (range 0-10) x 0.098) + (PsAID7 (sleep) NRS value (range 0-10) x 0.089) + (PsAID8 (coping) NRS value (range 0-10) x 0.087) + (PsAID9 (anxiety) NRS value (range 0-10) x 0.085)

Thus, the range of the final PsAID value is 0-10 where higher figures indicate worse status.

2. Missing data imputation

If one of the 9 NRS values composing the PsAID is missing, the imputation is as follows:

a. calculate the mean value of the 8 other (non-missing) NRS (range, 0-10)

b. impute this value for the missing NRS

c. Then, calculate the PsAID as explained above.

If 2 or more of the NRS are missing, the PsAID is considered as missing value (no imputation).

Appendix 2.8 Psoriasis Area and Severity Index and Affected Body Surface Area (PASI)[16]

Psoriasis Area and Severity Index is an index used to express the severity of psoriasis.

1. Intensity

A representative area of psoriasis is selected for each body region. The intensity of redness/ erythema, thickness, and scaling of the psoriasis is assessed as absent/none (0), mild (1), moderate (2), severe (3), or very severe (4).

2. Calculation for intensity

The three intensity scores are added up for each of the four body regions to give subtotals A1, A2, A3, A4.

Each subtotal is multiplied by the body surface area represented by that region.

- A1 x 0.1 gives B1
- A2 x 0.2 gives B2
- A3 x 0.3 gives B3
- A4 x 0.4 gives B4

3. Area

The percentage area affected by psoriasis is evaluated in the four regions of the body. In each region, the area is expressed as nil (0), 1-9% (score 1), 10-29% (score 2), 30-49% (score 3), 50-69% (score 4), 70-89% (score 5) or 90-100% (score 6).

- Head and neck
- Upper limbs
- Trunk
- Lower limbs

4. Calculations for area

Each of the body area scores is multiplied by the area affected.

- B1 x (0 to 6)= C1
- B2 x (0 to 6)= C2
- B3 x (0 to 6)= C3
- B4 x (0 to 6)= C4

5. Total score

The PASI score is C1 + C2 + C3 + C4.

Appendix 2.9 modified Nail Psoriasis Severity Index (mNAPSI)

mNAPSI is a scale used to evaluate the severity of nail bed psoriasis and nail matrix psoriasis by area of involvement in the nail unit.

NAPSI is used to assign a score to each nail for nail bed and nail matrix psoriasis: Nail plate is assessed for nail matrix psoriasis by the presence of any feature of nail matrix psoriasis, including nail pitting, leukonychia, red spots in the lunula, and crumbling in each quadrant of the nail. Nail bed psoriasis is assessed by the presence of any features of nail bed psoriasis, including onycholysis, oil drop (salmon patch) dyschromia, splinter hemorrhages, and nail bed hyperkeratosis in each quadrant of the nail.

The score is 0 if the findings are not present, 1 if they are present in 1 quadrant of the nail, 2 if present in 2 quadrants of a nail, 3 if present in 3 quadrants of a nail, and 4 if present in 4 quadrants of a nail. Thus, each nail has a matrix score (0-4) and a nail bed score (0-4), and the total nail score is the sum of those 2 individual scores (0-8). Sum of the total score of all involved fingernails is the total mNAPSI score.

Appendix 2.10 van der Heijde modified Total Sharp Score (mTSS)

The mTSS is a detailed scoring method evaluating erosions, joint space narrowing, (sub)luxation, ankylosis, gross osteolysis, and pencil in cup phenomena.

Hands/wrists radiographs will be obtained by sites and sent for central review. The central reviewers will use the modified total Sharp Score (mTSS) to evaluate for radiographic damage and progression.

Appendix 2.11 Minimal Disease Activity (MDA) and Very Low Disease Activity (VLDA)

A subject is classified as in MDA when 5 of the following 7 criteria are met:

- 68 tender joint count ≤ 1
- 66 swollen joint count ≤ 1
- PASI ≤ 1 or BSA ≤ 3%
- Subject's Pain Assessment (visual analogue scale [VAS]) ≤ 15 mm
- Subject's Global Activity VAS ≤ 20 mm (corresponds to Subject's Global Assessment of Disease Activity)
- HAQ-DI ≤ 0.5
- Tender entheseal points ≤ 1 site out of 6 sites included in LEI.

A subject is classified as in very low disease activity (VLDA) when all 7 criteria described above are met.

8. References

[1]**ICH.** *Statistical Principles for Clinical Trials*, Guideline E9, 1998. Available at https://database.ich.org/sites/default/files/E9_Guideline.pdf

[2]ICH. *Addendum on Estimands and Sensitivity Analysis in Clinical Trials*, Guideline E9(R1). Available at https://database.ich.org/sites/default/files/E9-R1_Step4_Guideline_2019_1203.pdf

[3]ICH. *Structure and Content of Clinical Study Reports*, Guideline E3, 1995. Available at https://database.ich.org/sites/default/files/E3_Guideline.pdf

[4]Brown H and Prescott R. *Applied Mixed Models in Medicine*. John Wiley & sons (Chichester), 1999.

[5] Fairclough D L. *Design and Analysis of Quality of Life Studies in Clinical Trials.* Chapman & Hall/CRC, 2002.

[6]Rubin, D. B. (1987). *Multiple Imputation for Nonresponse in Surveys*. New York: John Wiley & Sons.

[7]Fairclough D L. *Design and Analysis of Quality of Life Studies in Clinical Trials.* Chapman & Hall/CRC, 2002.

[8]Rubin, D. B. (1987). *Multiple Imputation for Nonresponse in Surveys*. New York: John Wiley & Sons.

[9]THE HEALTH ASSESSMENT QUESTIONNAIRE (HAQ) DISABILITY INDEX (DI) OF THE CLINICAL HEALTH ASSESSMENT QUESTIONNAIRE (VERSION 96.4): https://www.niehs.nih.gov/research/resources/assets/docs/haq_instructions_508.pdf

[10]36-Item Short Form Survey (SF-36) Scoring Instructions
QualityMetric, Inc., September 2022

[11]Scoring OF THE FACIT MEASURES
https://www.facit.org/scoring#:~:text=For%20any%20FACIT%20measure%2C%20subscale,derived%20by%20summing%20subscale%20scores

[12]Table 1 Formulae to calculate the different DAS and SDAI score
https://arthritis-research.biomedcentral.com/articles/10.1186/ar1787/tables/1

[13] DAPSA (Disease Activity in PSoriatic Arthritis) Score
https://rheuma.charite.de/fileadmin/user_upload/microsites/ohne_AZ/m_cc13/rheuma/Templates/DAPSA_ENG.pdf

[14]TREATING PSORIATIC ARTHRITIS TO TARGET: DEFINING PSORIATIC ARTHRITIS DISEASE ACTIVITY SCORE (PASDAS) THAT REFLECTS STATE OF MINIMAL DISEASE ACTIVITY (MDA)
https://www.jrheum.org/content/jrheum/early/2019/06/11/jrheum.181472.full.pdf

[15] The EULAR Psoriatic Arthritis Impact of Disease: PsAID9 for clinical trials
https://oml.eular.org/sysModules/obxOML/docs/id_5/PSAID-English.pdf

[16] PASI score
https://dermnetnz.org/topics/pasi-score