Statistical Analysis Plan for Study M15-998

A Phase 3, Randomized, Double-Blind Study Comparing Risankizumab to Placebo in Subjects with Active Psoriatic Arthritis Including Those Who Have a History of Inadequate Response or Intolerance to Biologic Therapy(ies)

Date: 16 November 2020

Version 6.0

Table of Contents

1.0 Introduction 6
2.0 Study Background 6
2.1 Objectives 6
2.2 Study Design 7
2.3 Treatment Assignment and Blinding 9
2.4 Sample Size Determination 9
3.0 Endpoints 9
3.1 Primary Endpoint 9
3.2 Secondary Endpoints 9
3.3 Additional Efficacy Endpoints 10
3.4 Safety Endpoints 13
3.5 Additional Endpoints 13
4.0 Analysis Populations 13
5.0 Subject Disposition 14
6.0 Study Drug Duration and Compliance 14
7.0 Demographics, Baseline Characteristics, Medical History, and Prior/Concomitant Medications 15
7.1 Demographics and Baseline Characteristics 15
7.2 Medical History 19
7.3 Prior and Concomitant Medications 19
8.0 Efficacy Analyses 20
8.1 General Considerations 20
8.2 Handling of Missing Data and Intercurrent Events 21
8.3 Primary Efficacy Endpoint and Analyses 24
8.3.1 Primary Efficacy Endpoint 24
8.3.2 Handling of Missing Data and Intercurrent Event for the Primary Efficacy Endpoint 25
8.3.3 Primary Efficacy Analysis 25
8.3.4 Additional Analyses of the Primary Efficacy Endpoint 26
8.4 Secondary Efficacy Analyses 26
8.4.1 Secondary Efficacy Analyses 26

8.4.2 Supportive Secondary Efficacy Analyses 27
8.5 Additional Efficacy Analyses 28
8.6 Efficacy Subgroup Analyses 29
9.0 Safety Analyses 30
9.1 General Considerations 30
9.2 Adverse Events 32
9.2.1 Analysis of Adverse Events for Period 1 33
9.2.1.1 Adverse Event Overview 33
9.2.1.2 Adverse Events by System Organ Class and Preferred Term 34
9.2.1.3 Serious Adverse Events (Including Deaths) and Adverse Events Leading to Study Drug Discontinuation 35
9.2.1.4 Frequent (>= 1%) Adverse Events by Preferred Term in Decreasing Frequency 35
9.2.1.5 Adverse Event by Subgroup 36
9.2.2 Analysis of Adverse Events for Long-Term 36
9.2.2.1 Overview of Adverse Events Rates per 100 Patient-Years of Study Drug Exposure 36
9.2.2.2 Adverse Events Rates per 100 Patient-Years of Study Drug Exposure by SOC and PT 36
9.2.2.3 Serious Adverse Events (Including Deaths) and Adverse Events Leading to Study Drug Discontinuation 36
9.2.2.4 Areas of Safety Interest Rates per 100 Patient-Years of Study Drug Exposure 37
9.3 Analysis of Laboratory Data 37
9.3.1 Variables and Units 38
9.3.2 Analysis of Laboratory Data for Period 1 40
9.3.2.1 Assessment of Mean Change from Baseline in Clinical Laboratory Variables 40
9.3.2.2 Assessment of Shift from Baseline in Clinical Laboratory Variables 40
9.3.2.3 Assessment of Potentially Clinically Significant Laboratory Values 41
9.3.2.4 Assessment of Liver Function test Elevations 41
9.3.3 Analysis of Long-Term Laboratory Data 43

9.3.3.1 Assessment of Mean Change from Baseline in Clinical Laboratory Variables....43
9.3.3.2 Assessment of Potentially Clinically Significant Laboratory Values....44
9.3.3.3 Assessment of Liver Elevations....44
9.4 Analysis of Vital Signs....44
9.4.1 Analysis of Vital Sign for Period 1....45
9.4.2 Analysis of Long-Term Vital Sign....45
10.0 Other Analyses....45
11.0 Interim Analysis....45
11.1 Data Monitoring Committee....46
12.0 Overall Type-I Error Control....46
13.0 Version History....48
14.0 References....50

List of Tables

Table 1. Subgroups for Efficacy Analysis....30
Table 2. List of Laboratory Variables....39
Table 3. SAP Version History Summary....48

List of Figures

Figure 1. Study Schematic....8
Figure 2. An Example of eDISH Plot....43
Figure 3. Graphical Testing Procedure....47

List of Appendices

Appendix A. Protocol Deviations....52
Appendix B. ASI for Risankizumab with SMQs/CMQs/PTs Searches....53
Appendix C. Potentially Clinically Significant Criteria for Safety Endpoints....54
Appendix D. Definition of Estimand for Primary and Secondary Endpoints....56

Appendix E. NRI-C Procedure 58
Appendix F. Tipping Point Analysis 62
Appendix G. CTCAE Grades 65
Appendix H. Geographic Region 66

1.0 Introduction

This statistical analysis plan (SAP) describes the statistical analysis for risankizumab Study M15-998 "A Phase 3, Randomized, Double-Blind Study Comparing Risankizumab to Placebo in Subjects with Active Psoriatic Arthritis Including Those Who Have a History of Inadequate Response or Intolerance to Biologic Therapy(ies)." It provides details to further elaborate statistical methods as outlined in the protocol and describes analysis conventions to guide the statistical programming work.

Pharmacokinetics/pharmacodynamics, pharmacogenetic, and selected biomarkers will be analyzed separately and are not included in this SAP.

Unless noted otherwise, all analyses will be performed using SAS version 9.4 or later (SAS Institute Inc., Cary, NC 27513) under the UNIX operating system.

2.0 Study Background

2.1 Objectives

Primary Objective

Period 1 Double-blind

To compare the efficacy of risankizumab 150 mg versus placebo for the treatment of signs and symptoms of PsA in the study population.

Secondary Objectives

Period 1 Double-blind

To compare the safety and tolerability of risankizumab 150 mg versus placebo in the study population.

Period 2 Open-label

To evaluate the long-term safety, tolerability and efficacy of risankizumab 150 mg in subjects who have completed Period 1.

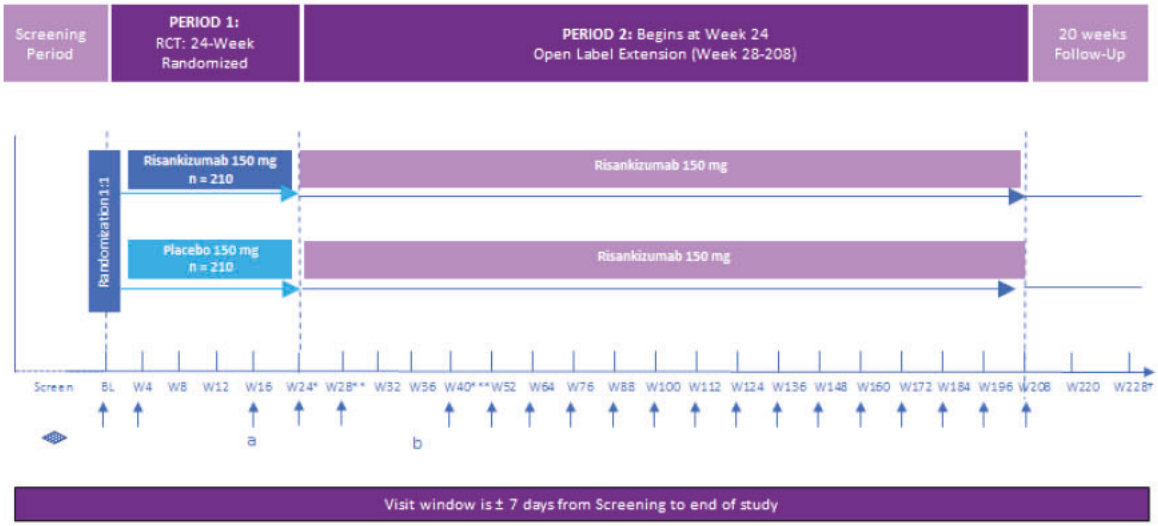
2.2 Study Design

This is a Phase 3, global, multi-center study that will evaluate moderately to severely active PsA patients. The subject population will consist of no more than 50% with an inadequate response (lack of efficacy after a minimum 12 weeks duration of therapy) or intolerance to 1 or 2 biologic therapies (Bio-IR). The remaining study population will be subjects who have an inadequate response (lack of efficacy after a minimum 12 weeks duration of therapy) or intolerance to at least 1 conventional synthetic disease modifying anti-rheumatic drugs (csDMARDs).

The study consists of a Screening Period (approximately 35 days), Period 1, Period 2 and a 20-week Follow-up Period. Period 1 is a 24-week randomized, double-blind, placebo-controlled, parallel-group period. Period 2 starts at Week 24. To maintain the blind to the original treatment allocation, treatment at Week 24 will be blinded: subjects randomized to placebo will receive blinded risankizumab 150 mg, and subjects randomized to risankizumab will receive blinded placebo. At Week 28 and for the remaining dosing visits (up to Week 208), all subjects will receive open-label risankizumab 150 mg every 12 weeks. Subjects will remain blinded to the original randomization allocation for the duration of study.

A schematic of the study is shown in Figure 1.

Figure 1. Study Schematic



LEGEND:

BL = Baseline; RCT = randomized clinical trial; W = Week

* At Week 24, subjects randomized to placebo in Period 1 will receive a blinded dose of risankizumab. Subjects randomized to risankizumab treatment in Period 1 will receive a blinded dose of placebo.

** At Week 28, subjects randomized to placebo in Period 1 will receive a 2nd dose of risankizumab. Subjects randomized to risankizumab in Period 1 will receive risankizumab (scheduled dose).

*** From Week 40 and up to Week 208 Visits, doses occur q12w.

† Follow up phone call.

Dosing.

Bilateral radiographs of hands and feet if required for CASPAR criteria.

a. At Week 16, subjects classified as non-responders (defined as not achieving at least a 20% improvement in either or both tender joint count [TJC] and swollen joint count [SJC] at both Week 12 and Week 16) compared to Baseline will add or modify rescue concomitant medications/therapy as described in Section 5.4 of the study protocol. Rescue therapy qualification occurs only at Week 16 Visit.

b. Starting at Week 36, subjects classified as non-responders will be discontinued from study drug. See Section 5.5 of the study protocol for details.

The Primary Analysis will be conducted after all ongoing subjects have completed Week 24 and an interim database lock (Week-24 Database Lock) has occurred, to support an interim CSR for the initial regulatory submission. An additional interim database lock will be conducted to provide further efficacy and safety exposure for integrated summaries of efficacy and safety supporting the initial regulatory submission; no interim

CSR will be generated based on the second interim lock. After then, additional interim analyses may be conducted as deemed appropriate. The final analysis will be performed after all ongoing subjects complete the study.

2.3 Treatment Assignment and Blinding

Subjects will be randomized to risankizumab 150 mg or placebo in a 1:1 ratio.

Randomization will be stratified by current use of csDMARD (0 vs ≥ 1), number of prior biologic therapies (0 vs ≥ 1), and extent of psoriasis (≥ 3% BSA or < 3% BSA).

2.4 Sample Size Determination

Approximately 420 subjects will be randomized to risankizumab 150 mg or placebo in a ratio of 1:1 (210 subject/treatment group). A sample size of 210 in each group will have approximately 90% power to detect a difference in HAQ-DI mean change from baseline of 0.24 assuming that the common standard deviation is 0.72 using a two group t-test with a 0.05 two-sided significance level and accounting for a 10% dropout rate. This sample size also ensures that analyses will have at least a 90% power to detect a 20% treatment difference, with the placebo response rate seen in the previous Phase 2b study of 35.7%, in ACR20 at Week 24 using a two-sided test at a 0.05 significance level and accounting for a 10% dropout rate.

3.0 Endpoints

3.1 Primary Endpoint

The primary endpoint is ACR20 response at Week 24.

3.2 Secondary Endpoints

The following is a list of ranked secondary endpoints with multiplicity adjustment:

1. Change from Baseline in Health Assessment Questionnaire-Disability Index (HAQ-DI) score at Week 24;

2. Psoriasis Area Severity Index (PASI) 90 response at Week 24 (for subjects with BSA ≥ 3% at Baseline);
3. ACR20 response at Week 16;
4. Minimal Disease Activity (MDA) response at Week 24;
5. Change from baseline in 36-Item Short Form Health Survey (SF-36) Physical Component Summary (PCS) score at Week 24;
6. Change from baseline in Functional Assessment of Chronic Illness Therapy-Fatigue (FACIT-Fatigue) score at Week 24.

Other secondary endpoints are:

1. ACR50 response at Week 24;
2. ACR70 response at Week 24;
3. Resolution of enthesitis defined as Leeds Enthesitis Index (LEI = 0) at Week 24 (for subjects with baseline presence of enthesitis (LEI > 0));
4. Resolution of dactylitis defined as Leeds Dactylitis Index (LDI = 0) at Week 24 (for subjects with baseline presence of dactylitis (LDI > 0)).

3.3 Additional Efficacy Endpoints

The primary and secondary efficacy endpoints are listed in Section 3.1 and Section 3.2, respectively. Additional efficacy endpoints at all scheduled visits (when measurements are collected) are listed below.

- ACR20/50/70 response;
- Change from Baseline in individual components of ACR response:
  - Change from Baseline in Tender Joint Count (TJC) (0 - 68);
  - Change from Baseline in Swollen Joint Count (SJC) (0 - 66);

  - Change from Baseline in Physician Global Assessment of Disease Activity (PGA) (mm on a 100-mm horizontal visual analogue scale [VAS]);
  - Change from Baseline in Patient's Global Assessment of Disease Activity (PtGA) (mm on a 100-mm horizontal visual analogue scale [VAS]);
  - Change from Baseline in Patient's Assessment of Pain (mm on a 100-mm horizontal visual analogue scale [VAS]);
  - Change from Baseline in Health Assessment Questionnaire - Disability Index (HAQ-DI) score;
  - Change from Baseline in High-Sensitivity C Reactive Protein (hs-CRP);
- MDA response;
- Change from Baseline in LDI (for subjects with Baseline LDI > 0);
- Change from Baseline in dactylitis count (for subjects with Baseline LDI > 0);
- Resolution of dactylitis (LDI = 0) (for subjects with Baseline LDI > 0);
- Change from Baseline in LEI (for subjects with Baseline LEI > 0);
- Resolution of enthesitis (LEI = 0) (for subjects with Baseline LEI > 0);
- Change from Baseline in Spondyloarthritis Research Consortium of Canada (SPARCC) Enthesitis Index (for subjects with baseline SPARCC Enthesitis Index > 0);
- Resolution of enthesitis sites included in the SPARCC Enthesitis Index (for subjects with Baseline SPARCC Enthesitis Index > 0);
- Change from Baseline in total enthesitis count (for subjects with baseline LEI > 0 and Baseline SPARCC Enthesitis Index > 0);
- Resolution of enthesitis in both LEI and SPARCC (defined as LEI = 0 and SPARCC = 0) (for subjects with Baseline LEI > 0 and Baseline SPARCC Enthesitis Index > 0);
- PASI 75/90/100 response (for subjects with BSA ≥ 3% at Baseline);
- Achievement of both PASI 90 and ACR50 (for subjects with BSA ≥ 3% at Baseline);
- Change from Baseline in BSA-PsO (for subjects with BSA ≥ 3% at Baseline);
- Modified Psoriatic Arthritis Response Criteria (PsARC);

- Change from Baseline in Disease Activity Score 28 using high sensitivity C Reactive Protein (DAS28-hsCRP);
- Change from Baseline in PsA Disease Activity Score (PASDAS);
- Change from Baseline in Disease Activity in Psoriatic Arthritis (DAPSA) score;
- Change from Baseline in FACIT-Fatigue score;
- Change from Baseline in SF-36 physical component summary, mental component summary and the 8 sub-domain scores;
- Change from Baseline in EuroQol-5D-5L (EQ-5D-5L) index and VAS scores;
- Change from Baseline in Work Productivity and Activity Impairment (WPAI) measures;
- Health resource utilization (HRU) since last study visit;
- Change from Baseline in Bath Ankylosing Spondylitis Disease Activity Index (BASDAI) score (for subjects with spondylitis at Baseline, for subjects with spondylitis at Baseline confirmed by radiography or MRI);
- Change from Baseline in modified BASDAI score (for subjects with spondylitis at Baseline, for subjects with spondylitis at Baseline confirmed by radiography or MRI);
- Change from Baseline in BASDAI – Spinal Pain (for subjects with spondylitis at Baseline, for subjects with spondylitis at Baseline confirmed by radiography or MRI);
- BASDAI50 response (for subjects with spondylitis at Baseline, for subjects with spondylitis at Baseline confirmed by radiography or MRI);
- Change from Baseline in morning stiffness score (mean of BASDAI Questions 5 and 6) (for subjects with spondylitis at Baseline, for subjects with spondylitis at Baseline confirmed by radiography or MRI);
- Change from Baseline in Ankylosing Spondylitis Disease Activity Score (ASDAS) (for subjects with spondylitis at Baseline and for subjects with spondylitis at Baseline confirmed by radiography or MRI);
- ASDAS Inactive Disease (for subjects with spondylitis at Baseline and for subjects with spondylitis at Baseline confirmed by radiography or MRI);

- ASDAS Major Improvement (for subjects with spondylitis at Baseline and for subjects with spondylitis at Baseline confirmed by radiography or MRI);
- ASDAS Clinically Important Improvement (for subjects with spondylitis at Baseline and for subjects with spondylitis at Baseline confirmed by radiography or MRI);
- Achievement of a clinically meaningful improvement in HAQ-DI defined as ΔHAQ-DI ≤ -0.35 (for subjects with HAQ-DI ≥ 0.35 at Baseline).

3.4 Safety Endpoints

Safety evaluations include adverse event (AE) monitoring, physical examinations, vital sign measurements, and clinical laboratory testing (hematology and chemistry) as a measure of safety and tolerability for the entire study duration.

3.5 Additional Endpoints

No additional endpoints will be analyzed in the SAP. Pharmacological endpoints will be analyzed separately and is not included in this SAP.

4.0 Analysis Populations

The Full Analysis Set (FAS) includes all randomized subjects who received at least one dose of study drug. The FAS will be used for all efficacy and baseline analyses. Subjects will be included in the analysis according to the treatment groups that they are randomized to.

The Per-Protocol Analysis Set consists of a subset of FAS subjects who did not have any major protocol deviations that are determined to have a potential impact on the primary efficacy endpoint up to Week 24 in Period 1 of the study. Additional analysis of the primary efficacy endpoint will be conducted on the Per-Protocol Analysis Set.

The final criteria and the exclusion of subjects for the per-protocol analysis will be identified prior to the Primary Analysis database lock.

The Safety Analysis Set consists of all subjects who received at least one dose of study drug. Subjects will be included in the analysis according to the study drug that they actually received. The Safety Analysis Set will be used for all safety analysis.

5.0 Subject Disposition

The total number of subjects who were randomized and who were treated will be summarized.

A summary of subject accountability will be provided where the number of subjects in each of the following categories will be summarized for each randomized treatment group:

- Subjects randomized in the study;
- Subjects included in key analysis populations (FAS, Per-Protocol Analysis Set, Safety Analysis Set);
- Subjects who took at least one dose of study drug;
- Subjects who completed Period 1 study participation;
- Subjects who entered Period 2;
- Subjects who completed overall study (Period 1 and Period 2) participation;
- Subjects who prematurely discontinued study drug (all reasons and primary reason).

For end of study participation, the number and percentage of subjects who completed the protocol defined follow-up period (or did not with associated reasons) will be summarized overall and by treatment group.

6.0 Study Drug Duration and Compliance

For the Safety Analysis Set, duration of treatment will be summarized for risankizumab 150 mg and placebo for Period 1, and for risankizumab 150 mg for long-term. Duration of treatment is defined for each subject as last dose date – first dose date + 84 days.

Duration of treatment will be summarized using the number of subjects treated, mean, standard deviation, median, minimum and maximum.

Treatment compliance will be summarized by treatment group at Week 24 for the FAS population. Treatment compliance is defined as the number of injections administered during the subject's participation up to Week 24 divided by the number of injections planned during the subject's participation in the treatment phase up to Week 24. Compliance will be calculated for each subject and summarized with the mean, median, standard deviation, minimum, and maximum.

7.0 Demographics, Baseline Characteristics, Medical History, and Prior/Concomitant Medications

Demographic, baseline or disease characteristics, medical history, and prior and concomitant medications will be summarized for the FAS by treatment group and overall. Categorical variables will be summarized with the number and percentage of subjects. Continuous variables will be summarized with descriptive statistics (number of non-missing observations, mean and standard deviation, median, minimum and maximum).

7.1 Demographics and Baseline Characteristics

The following demographic and baseline characteristics will be summarized.

Demographic Characteristics

- Sex (male, female)
- Age (years)
- Age Categories (< 65 years, ≥ 65 years, ≥ 65 and < 75 years, ≥ 75 years)
- Race (White, Black or African American, American Indian or Alaska Native, Native Hawaiian or Other Pacific Islander, Asian, Other)
- Geographic Region (North America, South/Central America, Western Europe, Eastern Europe, Asia, Other)
- Ethnicity (Hispanic or Latino, Not Hispanic or Latino)

- Weight (kg)
- Weight Categories (< 100 kg, ≥ 100 kg)
- Height (cm)
- Body Mass Index (BMI) (kg/m^2)
- Body Mass Index (BMI) Category (kg/m^2) (< 25, ≥ 25 and < 30, ≥ 30)

PsA Medical History and Prior and Concomitant Medications

- Duration of PsA in years
- Duration of PsA in years categories (≤ 5, > 5 and ≤ 10, > 10 years)
- Number of prior csDMARDs (0, 1, 2, ≥ 3)
- Number of prior biologics (0, ≥ 1)
- Number of prior failed biologics (0, 1, ≥ 2)
- Prior exposure to TNF antagonists (yes, no)
- Nonsteroidal anti-inflammatory drug (NSAID) use at Baseline (yes, no)
- Oral Corticosteroid use at Baseline (yes, no)
- Concomitant MTX use at Baseline (yes, no)
- Concomitant csDMARD at Baseline

  Any csDMARD
  - Any MTX
    - MTX alone
    - MTX and other csDMARD
  - csDMARD other than MTX
    - Any sulfasalazine, without MTX
    - Any leflunomide, without MTX
    - Any apremilast, without MTX

  None

Baseline Disease Characteristics

Categorical:

- Minimal Disease Activity (MDA) for PsA (yes, no)
- Presence of Dactylitis (LDI > 0) (yes, no)
- Presence of Enthesitis based on LEI (LEI > 0) (yes, no)
- Presence of Enthesitis based on SPARCC (SPARCC Enthesitis Index > 0) (yes, no)
- Body Surface Area (BSA) of psoriatic plaques (≥ 3%, < 3%)
- Presence of spondylitis (yes, no)
- Anti-cyclic citrullinated peptide (Anti-CCP) status (Positive (≥ 17 U/mL), Negative (< 17 U/mL))
- Rheumatoid Factor (RF) status (Positive (≥ 15 IU/mL), Negative (< 15 IU/mL))
- Categories for hsCRP (< 3.0, ≥ 3.0 mg/L)

Continuous:

- Psoriasis Area and Severity Index (PASI) (for subjects with BSA ≥ 3% at Baseline)
- Disease Activity Score in 28 joints – hsCRP (DAS28(CRP))
- Total dactylitis count (for subjects with presence of dactylitis (LDI > 0))
- LDI (for subjects with presence of dactylitis (LDI > 0))
- Total enthesitis count (for subjects with presence of enthesitis (total enthesitis count > 0))
- LEI (for subjects with LEI > 0)
- Classification Criteria for PsA (CASPAR classification criteria total score)
- BSA of psoriatic plaques (BSA-PsO) (for subjects with BSA ≥ 3% at Baseline and for all subjects)
- SPARCC Enthesitis Index (for subjects with SPARCC Enthesitis Index > 0)
- Tender joint count (TJC68) defined as the number of tender joints out of 68 assessed joints

- Swollen joint count (SJC66) defined as the number of swollen joints out of 66 assessed joints
- Tender joint count (TJC28) defined as the number of tender joints out of 28 assessed joints used for DAS28 calculation
- Swollen joint count (SJC28) defined as the number of swollen joints out of 28 assessed joints used for DAS28 calculation
- Physician's global assessment of disease activity (mm on a 100-mm horizontal visual analogue scale [VAS])
- Patient's assessment of pain (mm on a 100-mm horizontal visual analogue scale (VAS))
- Patient's global assessment of disease activity (mm on a 100-mm horizontal VAS)
- High sensitivity C-reactive protein (hsCRP) (mg/L)
- Health Assessment Questionnaire Disability Index (HAQ - DI) score
- Functional Assessment of Chronic Illness Therapy-Fatigue (FACIT-Fatigue) score
- 36-Item Short Form Health Survey (SF-36) Version 2: physical component summary, mental component summary and the 8 sub-domain scores
- EurQol-5d-5L (EQ-5D-5L) index and VAS score
- Work Productivity and Activity Impairment (WPAI) measures
- Bath Ankylosing Spondylitis Activity Index (BASDAI) score (for subjects with spondylitis at Baseline, for subjects with spondylitis at Baseline confirmed by radiography or MRI)
- Modified BASDAI score (for subjects with spondylitis at Baseline, for subjects with spondylitis at Baseline confirmed by radiography or MRI)
- BASDAI – Spinal Pain (for subjects with spondylitis at Baseline, for subjects with spondylitis at Baseline confirmed by radiography or MRI)
- ASDAS score (for subjects with spondylitis at Baseline, for subjects with spondylitis at Baseline confirmed by radiography or MRI)

Tobacco/Nicotine and Alcohol Use

- Tobacco/Nicotine Use (current, former, never, unknown)
- Alcohol Use (current, former, never, unknown)

7.2 Medical History

Medical history data will be coded using the Medical Dictionary for Regulatory Activities (MedDRA). The actual version of the MedDRA coding dictionary will be noted in the statistical tables and clinical study report. The number and percentage of subjects in each medical history category (by MedDRA system organ class and preferred term) will be summarized overall and by treatment group. The system organ class (SOC) will be presented in alphabetical order, and the preferred terms will be presented in alphabetical order within each SOC. Subjects reporting more than one condition/diagnosis will be counted only once in each row (SOC or preferred term).

7.3 Prior and Concomitant Medications

Prior and concomitant medications will be summarized by generic name. A prior medication is defined as any medication taken prior to the date of the first dose of study drug (risankizumab or placebo). A concomitant medication is defined as any medication that is started prior to the date of the first dose of study drug and continues to be taken after the first dose of study drug or any medication that is started on or after the date of the first dose of study drug and not after the date of the last dose of study drug + 140 days. The number and percentage of subjects taking medications will be summarized by generic drug name based on the World Health Organization (WHO) Drug Dictionary for both prior and concomitant medications. In addition, the number and percentage of subjects taking rescue medication will be summarized by ATC classification level 2, generic drug name and route of administration. A listing of subjects taking rescue medication with generic drug name, dosage and route of administration will also be provided.

8.0 Efficacy Analyses

8.1 General Considerations

All efficacy analyses will be conducted in the FAS Population. In addition, Per-protocol analysis for primary endpoint will be performed. All tests will be 2-sided at an alpha level of 0.05.

The Primary Analysis will be conducted after all ongoing subjects have completed Week 24 and the Week-24 Database Lock has occurred. Efficacy and safety analyses will be performed for Period 1 and long-term efficacy and safety will be summarized through the data cutoff date. An additional interim database lock will be conducted to provide further efficacy and safety exposure for integrated summaries of efficacy and safety supporting the initial regulatory submission; no study-specific analysis will be performed based on this database lock.

Unless otherwise specified, categorical variables will be analyzed using Cochran-Mantel-Haenszel (CMH) test, adjusting for the stratification factors. Continuous variables will be analyzed using Mixed-Effect Model Repeat Measurement (MMRM) method.

Unless otherwise specified, any subject who is randomized based on a wrong stratum will be analyzed according to the actual stratum the subject belongs to.

"Baseline" refers to the last non-missing observation before the first administration of study drug or before randomization if no study drug is given.

Efficacy analysis for Period 1

The efficacy analysis for Period 1 will be performed after all ongoing subjects have completed Period 1 and an interim database lock has occurred. This is the only and final efficacy analysis in Period 1.

Long-term efficacy analysis

Long-term efficacy will be summarized based on data in Period 1 and Period 2 up to data cutoff date for the Week-24 Database Lock. After the initial regulatory submission, additional interim analyses will be performed as deemed appropriate, for example, to address requests from regulatory agencies. The final analysis will be performed after all ongoing subjects complete the study.

The long-term efficacy analysis will be performed on As Observed data (defined in Section 8.2) by randomized treatment group sequence as described below:

1. Placebo → risankizumab 150 mg
2. risankizumab 150 mg → risankizumab 150 mg

There will be no statistical testing for long-term efficacy analysis; only descriptive statistics and 95% confidence intervals will be provided.

8.2 Handling of Missing Data and Intercurrent Events

Missing data could occur due to various reasons, including missing visits/assessments, early withdrawal from the study, or missing due to COVID-19 infection or logistic restriction.

The COVID-19 pandemic is interfering with the conduct of many ongoing trials, with potential impacts on treatment duration and the collection, analysis, and the interpretation of clinical trial data. Some protocol-specified visits in the clinical trials may be impacted due to COVID-19 infection or logistical restrictions during the pandemic. For example, some scheduled visits may be missed due to self-quarantine or local government restrictions on travel; some visits may also be delayed or canceled due to healthcare resource constraints during the pandemic. Impacted visits due to COVID-19 will be recorded in the database. The probability of having missed visits and missing data due to COVID-19 infection or logistical restrictions related to the COVID-19 pandemic can be

reasonably assumed to be unrelated to the unobserved values. Therefore, for the purpose of statistical analysis, it is reasonable to assume that these missing data are missing at random (MAR) and the statistical models that require MAR assumption are appropriate. The intent is to provide reliable estimates of the treatment effects targeted in the protocol under the scenario without the impact of COVID-19 pandemic. Sensitivity analyses will be performed to assess the impact of missing data and the robustness of the conclusion. Number of subjects with missing values due to COVID-19 will be presented.

Intercurrent events includes initiation of rescue medication and initiation of concomitant medications for PsA use that could meaningfully impact efficacy assessment. Missing data and intercurrent events will be handled using the following methods for the efficacy analysis.

Non-Responder Imputation incorporating Multiple Imputation (MI) to handle missing data due to COVID-19 (NRI-C)

The NRI-C data handling will be used for the primary estimand (refer to Section 8.3.1) for the binary endpoints. It will handle data for binary variables as follows.

a. Missing data due to COVID-19 infection or logistical restriction related to COVID-19 pandemic will be handled by Multiple Imputation.
b. Subject who does not have evaluation during a specific visit window due to reasons other than COVID-19 infection or logistical restriction related to COVID-19 pandemic will be handled by NRI for that visit. NRI will consider a subject with missing evaluation as a non-responder with the exception for composite binary endpoints including ACR20, ACR50, ACR70, MDA and modified PsARC for which the missing components will be imputed with last observation carry forward to derive composite score before imputing missing evaluations as a non-responder.
c. Subjects will be considered as non-responders after initiation of rescue medication or initiation of concomitant medications for PsA that could

meaningfully impact efficacy assessment; these medications will be identified prior to database lock and unblinding.

For composite binary endpoints including ACR20, ACR50, ACR70, MDA and modified PsARC, the missing binary values due to COVID-19 infection or logistical restriction will be imputed via MI with the logistic regression option as outlined in Appendix F. For other binary endpoints which are dichotomized from a continuous scale, the MI will be applied to the original continuous scale and the dichotomized endpoint will be derived from the imputed value for missing due to COVID-19 infection or logistical restriction.

Mixed-Effect Model Repeated Measures (MMRM)

The repeated measures analysis will be conducted using a mixed model including observed measurements at all visits. For the MMRM analysis, data collected after initiation of rescue medication or initiation of concomitant medications for PsA use that could meaningfully impact efficacy assessment will be excluded. The mixed model includes the categorical fixed effects of treatment, visit and treatment-by-visit interaction, stratification factors, and the continuous fixed covariate of baseline measurement. An unstructured variance covariance matrix will be used. Parameter estimation is based on the assumption of data being missing at random and using the method of restricted maximum likelihood (REML). The MMRM approach is appropriate in handling missing data due to COVID-19 infection or logistical restriction given the validity of the missing at random assumption. MMRM will be used for the primary estimand of continuous variables (refer to Section 8.4.1).

As Observed (AO)

The AO analysis will not impute values for missing evaluations, and thus a subject who does not have an evaluation on a scheduled visit will be excluded from the AO analysis for that visit. Regardless of premature discontinuation of study drug, initiation of concomitant medications for PsA use that could meaningfully impact efficacy assessment,

or initiation of rescue medication, all observed data will be used in the analysis. The AO analysis will be used to facilitate the supplementary analysis.

AO with Imputation

The AO with imputation will use all observed data regardless of premature discontinuation of study drug, initiation of concomitant medications for PsA use that could meaningfully impact efficacy assessment, or initiation of rescue medication and will impute missing as non-responders. It will be used for the supplementary analysis for binary variables.

Multiple Imputation (MI)

As a sensitivity analysis, MI will be utilized for both hypothetical and treatment policy estimand (Appendix D). MI analysis will be performed based on AO data. The MI analysis will impute missing data multiple times under appropriate random variation and thus generate multiple imputed "pseudo-complete" datasets. SAS PROC MI will be used to generate 30 datasets using the fully conditional specification (FCS) method. Specifically, treatment group is included in the FCS imputation model to enable stratified sampling. Additionally, the imputation model includes demographics variables, baseline disease characteristics, stratification factors, as well as longitudinal response observed at any other visits. Analysis using CMH for binary endpoints and analysis of covariance (ANCOVA) for continuous endpoints will be performed on each of the multiple imputed datasets. Subsequently SAS PROC MIANALYZE will then be used to aggregate the results for the final statistical inference using Rubin's method.[1]

8.3 Primary Efficacy Endpoint and Analyses

8.3.1 Primary Efficacy Endpoint

The primary efficacy endpoint is ACR20 response at Week 24. The attributes of primary estimand based on composite estimand framework are outlined in Appendix D.

8.3.2 Handling of Missing Data and Intercurrent Event for the Primary Efficacy Endpoint

For the primary estimand, NRI-C as defined in Section 8.2 will be used.

Subjects who initiated concomitant medications for PsA that could meaningfully impact efficacy assessment or were rescued prior to Week 24 or missed ACR20 response at Week 24 will be treated as non-responders. The number and percentage of non-responders for ACR20 will be summarized into one of the five categories for which the episode occurs first:

1. Subjects rescued before Week 24
2. Subjects who initiated concomitant medications for PsA that could meaningfully impact efficacy assessment by Week 24
3. Subjects with ACR20 measurements observed at Week 24 but did not meet ACR20 response criteria
4. Subjects with missing Week 24 ACR20 measurements due to reasons other than COVID-19 infection or logistical restriction related to COVID-19 pandemic
5. Subjects with missing Week 24 ACR20 measurements due to COVID-19 infection or logistical restriction related to COVID-19 pandemic

8.3.3 Primary Efficacy Analysis

Analysis of the primary endpoint will be conducted on the FAS based on randomized treatment groups (risankizumab 150 mg versus placebo) using NRI-C for missing data handling. Point estimate and 95% CI using normal approximation will be provided for the response rate for each randomized treatment group. Comparison of the primary endpoint between the risankizumab and placebo groups will be made using the CMH adjusting for stratification factors. Point estimate, 95% CI and nominal p-value for the treatment difference will be presented. Refer to Appendix E for the details of NRI-C procedure and respective statistical inference.

A sensitivity analysis will be conducted on the Per Protocol Analysis Set using the CMH model and NRI-C missing data handling.

8.3.4 Additional Analyses of the Primary Efficacy Endpoint

For the primary efficacy endpoint, a supplementary analysis will be based on treatment policy estimand framework using the same CMH method using AO data. The attributes of the corresponding estimand are outline in Appendix D.

For the treatment policy estimand, additional supplementary analysis will be conducted using AO with imputation. The same CMH method will be conducted.

Additional supplementary analysis using MI as outlined in Section 8.2 based on AO data, i.e., all observed data regardless of study drug adherence or initiation of rescue or concomitant medications for PsA that could meaningfully impact efficacy assessment, will also be conducted to handle missing ACR20 responses. To assess the deviation from missing at random (MAR) assumption, tipping point analysis will be conducted as a sensitivity analysis for this supplementary analysis. Details of the analysis are outlined in Appendix F.

8.4 Secondary Efficacy Analyses

8.4.1 Secondary Efficacy Analyses

The secondary endpoints are defined in Section 3.2. Ranked secondary endpoint will be tested in rank order, according to procedures in Section 12.0.

For all binary secondary endpoints, frequencies and percentages will be reported for each treatment group. The primary estimand and analysis method are the same as that for the primary efficacy endpoint as defined in Section 8.3.4. NRI-C missing data handling will be used to analyze the primary estimand (refer to Appendix D).

For the secondary endpoints applicable only for a sub-population, the estimand will be constructed similarly but based on the sub-population, as defined in the below table.

Endpoint	Sub-population
Psoriasis Area Severity Index (PASI) 90 response at Week 24	Subjects with BSA ≥ 3% at baseline are included in the analysis
Resolution of enthesitis (LEI = 0) at Week 24	Subjects with baseline presence of enthesitis (LEI > 0) are included in the analysis
Resolution of dactylitis (LDI = 0) at Week 24	Subjects with baseline presence of dactylitis (LDI > 0) are included in the analysis

For continuous secondary efficacy variables, primary analysis will be based on the hypothetical estimand. The attributes of the corresponding estimand are outline in Appendix D. The comparison is risankizumab 150 mg vs placebo for patients randomized and treated with at least one dose of study drug. Statistical inference will be conducted using the MMRM model and the associated data handling as described in Section 8.2. The least square (LS) mean and 95% CI for each randomized treatment group and the LS mean treatment difference, the associated 95% CI and the p-value between risankizumab 150 mg group and the placebo group will be provided.

8.4.2 Supportive Secondary Efficacy Analyses

For all binary secondary endpoints, the same CMH analyses as the supplementary analyses will be conducted using AO data and AO with imputation. The corresponding supplementary estimand is the same as defined in Section 8.3.4 except for the definition of the efficacy measurement (refer to Appendix D).

For continuous secondary endpoints, supplementary analyses on the FAS will be performed based on AO data using the treatment policy estimand framework (refer to Appendix D). The statistical inference will be conducted using ANCOVA model including treatment as the fixed factor and the corresponding baseline value and the stratification factor as the covariates. The LS mean and 95% CI will be reported for each randomized treatment group; the LS mean treatment difference and associated 95% CI and p-value will be reported comparing risankizumab 150 mg with the placebo group.

For HAQ-DI at Week 24, additional supplementary analysis will be conducted using MI as outlined in Section 8.2 based on AO data under MAR assumption. To assess deviations from MAR, the tipping point analysis will also be conducted using MI as additional sensitivity analyses. Details of the MI and tipping point analysis are outlined in Appendix F.

8.5 Additional Efficacy Analyses

Additional endpoints defined in Section 3.3 will be analyzed in at scheduled time points other than those specified for the primary and key secondary variables.

Additional Efficacy Analyses for Period 1

Additional endpoints will be analyzed for each randomized treatment group for all visits.

For binary endpoints, point estimate and 95% CI using normal approximation will be provided for the response rate for each randomized treatment group. Point estimate, 95% CI and p-value will be provided for the treatment comparison between risankizumab 150 mg and the placebo group based on the CMH test adjusting for stratification factors. Only the nominal p-value will be provided. NRI-C for missing data handling will be used. AO analysis will be used as supplementary analysis.

For continuous endpoints, the LS mean and 95% CI will be reported for each randomized treatment group. The LS mean treatment difference and associated 95% CI and p-values between risankizumab 150 mg and the placebo group will be provided using MMRM model. Only the nominal p-value will be provided. Data observed after initiation of rescue or concomitant medications for PsA that could meaningfully impact efficacy assessment will be excluded. Supplementary analysis will be conducted based on ANCOVA model using AO data, with treatment and stratification factors as the fixed factors and the corresponding baseline value as the covariates. Nominal p-values will be provided.

For HRU related endpoints, number and percentage of subjects with non-study health visits will be provided by visit and over the double-blind period up to Week 24. In addition, cumulative number of utilizations per time (e.g., subject-year) up to Week 24 will be provided for each category.

Long-Term Efficacy Analyses

Assessments to evaluate long-term efficacy will be analyzed for all efficacy measurements at scheduled visits.

Descriptive statistics will be provided for each randomized treatment sequence as defined in Section 8.1. These include the number of observations, mean, standard deviation, 95% CI, median, minimum, and maximum for continuous endpoints, and frequencies and percentages with 95% CI using normal approximation for binary endpoints. Plot for each randomized treatment group sequence over time will be provided for primary and multiplicity-controlled secondary endpoints and selected additional endpoints. These efficacy analyses will be based on AO analysis.

A similar analysis for HRU endpoints will be provided at Week 52.

8.6 Efficacy Subgroup Analyses

The primary efficacy endpoint will be examined in the subgroups listed in Table 1 below. Treatment difference between risankizumab 150 mg and the placebo group in each category of a subgroup will be presented with point estimate and 95% confidence interval using CMH test adjusting for stratification factors. No p-value will be provided for subgroup analysis. Forest plots presenting the treatment effect of all categories of subgroups will be provided.

Table 1. Subgroups for Efficacy Analysis

Subgroup Factor	Categories
Age	< 65 years, ≥ 65 years, ≥ 65 - < 75 years, ≥ 75 years
Sex	Male vs Female
BMI	< 25, ≥ 25 and < 30, ≥ 30 kg/m^2
Race	White vs Non-white
Geographic Region	North America, South/Central America, Western Europe, Eastern Europe, Asia, Other
Number of prior csDMARDs	≤ 1 vs > 1
Number of prior biologic therapies	0 vs ≥ 1
Number of prior anti-TNFs	0 vs ≥ 1
hsCRP at Baseline	< 3 vs ≥ 3 mg/L
Extent of psoriasis at Baseline	≥ 3% BSA vs < 3% BSA
Duration of PsA	≤ 5, > 5 and ≤ 10, > 10 years

Concomitant csDMARD at baseline	Any csDMARD
	• Any MTX
	  o MTX alone
	  o MTX and other csDMARD
	• csDMARD other than MTX
	None

In addition, subgroup analysis will be conducted for the secondary endpoints by number of prior biologics therapies (0 vs ≥ 1). Similar analysis will be provided.

9.0 Safety Analyses

9.1 General Considerations

Safety data will be summarized for the safety population. For the safety analysis, subjects are assigned to a treatment group based on the treatment actually received, regardless of the treatment randomized. A subject's actual treatment will be determined by the most frequent dose regimen received.

There are two sets of planned safety analysis: safety analysis for Period 1, and long-term safety analysis.

Safety Analysis for Period 1

Standard safety analysis by the actual treatment groups of risankizumab 150 mg and placebo will be performed on safety data up to Week 24. No protocol-defined treatment switching will occur prior to these time points.

The standard safety analyses will include reporting of adverse events (AEs), laboratory, and vital signs measurements. Frequency tables of subjects with treatment-emergent adverse events (TEAEs) by system organ class (SOC) and by preferred term (PT) as in the Medical Dictionary for Regulatory Activities (MedDRA) dictionary will be provided by treatment group. All continuous laboratory parameters and vital signs variables at each visit will also be summarized by actual treatment group. Frequency tables of subjects meeting criteria for potentially clinically significant vital sign values and for potentially clinically significant laboratory values will be provided by treatment group.

Long-Term Safety Analysis

Long-term safety analysis will be performed on all cumulative safety data in Period 1 and Period 2 up to data cutoff date for any interim lock as deemed appropriate and final lock. Long-term safety analyses that account for protocol-defined treatment switching include reporting of AE rate adjusted by cumulative exposure, descriptive summary in laboratory parameters and vital sign variables, and rate of potentially clinically significant laboratory and vital signs values.

To adjust for potentially different follow-up time between treatment groups, exposure-adjusted event rate (EAER) will be provided for long term safety analysis. The EAER will be presented for Any Risankizumab group by actual treatment received at the time of AE. The Any Risankizumab group includes all subjects who received risankisumab 150 mg, including those who started on risankizumab 150 mg at randomization and who switched from placebo to risankizumab 150 mg.

For EAER calculation, the numerator will be the total number of AEs reported for the event (i.e., a subject can contribute more than one event to the numerator) and the denominator will be the total exposure time among subjects under the treatment group. The number of AEs reported (numerator), the total number of years of study drug exposure (denominator, calculated as total number of days exposed to study drug (i.e., last dose date – first dose date + 140 days) for all treated subjects divided by 365.25), and the exposure-adjusted AE event rate per 100 patient-years calculated as ([numerator (number of AEs)/denominator]) • 100 will be presented for each treatment group.

All continuous laboratory parameters and vital signs variables at each visit will be summarized by actual treatment group sequence defined as follows.

1. Placebo → Risankizumab 150 mg
2. Risankizumab 150 mg → Risankizumab 150 mg

Frequency tables and listings of subjects meeting criteria for potentially clinically significant vital sign values and for potentially clinically significant laboratory values will be provided for Any Risankizumab group by actual treatment received at the time of event. Missing safety data will not be imputed.

9.2 Adverse Events

A treatment-emergent Adverse Event (TEAE) is defined as an adverse event with an onset date that is after the first dose of study drug, and no more than 140 days after the last dose of study drug.

Adverse events (AEs) will be summarized and presented using primary MedDRA System Organ Classes (SOCs) and preferred terms (PTs) according to the version of the MedDRA coding dictionary used for the study at the time of database lock. The actual version of the MedDRA coding dictionary used will be noted in the AE tables and in the clinical study report. Specific adverse events will be counted once for each subject for calculating percentages, unless stated otherwise. In addition, if the same adverse event occurs

multiple times within a subject, the highest severity and level of relationship to investigational product will be reported.

9.2.1 Analysis of Adverse Events for Period 1

9.2.1.1 Adverse Event Overview

An overview of AEs will be presented consisting of the number and percentage of subjects experiencing at least one event for each of the following AE categories:

- Any TEAEs
- Any COVID-19 related TEAEs
- Any TEAE related to study drug according to the investigator
- Any serious TEAE
- Any severe TEAE
- Any TEAE leading to discontinuation of study drug
- Any TEAE leading to death
- All deaths
    - COVID-19 related deaths
    - Deaths occurring ≤ 140 days after last dose of study drug
    - Deaths occurring > 140 days after last dose of study drug.

Additional AEs may be considered for tabulation/summary based on recommendations from Clinical and Safety as deemed appropriate.

The number and percentage of subjects experiencing at least one event of treatment-emergent AEs will be summarized for each treatment group. The point estimate and 95% CI (using normal approximation and separate group variance) will be provided for the treatment difference in AE percentage between risankizumab group and the placebo group. An overview of the AE of safety interest (ASI) will be provided similarly and the categories of ASI is defined in Section 9.2.1.4. In addition, an overview of AEs per 100 patient-years of study exposure will be presented for the AE categories defined

above. The number of TEAEs reported, the total number of years of study drug exposure, and the TEAE rate per 100 patient-years will be presented.

COVID-19 cases will be identified using relevant terms for COVID-19 search available in MedDRA version 23.0 or higher. All COVID-19 related AEs and COVID-19 related deaths will be summarized in the TEAE Overview table.

9.2.1.2 Adverse Events by System Organ Class and Preferred Term

Treatment-emergent adverse events will be summarized by SOC and PT, the following summaries of adverse events will be generated:

- All TEAEs
- Serious TEAEs
- Severe TEAEs
- TEAEs related to study drug according to the investigator
- TEAEs leading to discontinuation of study drug
- TEAE leading to death
- COVID-19 related TEAE

Subjects reporting more than one adverse event for a given MedDRA preferred term will be counted only once for that term (most severe incident for the severity tables and most related incident for the relationship tables). Subjects reporting more than one type of adverse event within a SOC will be counted only once for that SOC. Subjects reporting more than one type of adverse event will be counted only once in the overall total.

TEAEs will also be summarized by relationship to risankizumab 150 mg and placebo as assessed by the investigator as well as by maximum severity, by treatment groups.

The SOCs will be presented in alphabetical order, and the PTs will be presented in alphabetical order within each SOC.

9.2.1.3 Serious Adverse Events (Including Deaths) and Adverse Events Leading to Study Drug Discontinuation

All serious adverse events (SAEs) including deaths and adverse events leading to discontinuation of study drug will be listed in tables, besides summary by SOC and PT covered in Section 9.2.1.2.

9.2.1.4 Frequent (>= 1%) Adverse Events by Preferred Term in Decreasing Frequency

TEAEs occurring for at least 1% of the subjects in any of the treatment arms will be summarized by MedDRA PT in decreasing frequency of Risankizumab treatment group.

Areas of Safety Interest

The Adverse Events of Safety Interest (ASI) categories will be summarized by SOC and PT and will be based on standardized or company MedDRA queries (SMQs or CMQs). ASI are categorized in Appendix B. Adjudicated cardiovascular events will be summarized and presented by treatment group using the CAC adjudicated categories. Serious opportunistic infection excluding tuberculosis and herpes zoster, serious hypersensitivity, serious hepatic events and serious injection site reaction will also be summarized by PT.

In addition, an overview of ASI per 100 patient-years of study exposure will be presented. The number of TEAEs reported, the total number of years of study drug exposure, and the TEAE rate per 100 patient-years will be presented.

Similarly, overview tables for serious ASI in number and percentage as well as in event rate per 100 patient-years will be provided.

Additional AEs may be considered for tabulation/summary based on recommendations from Clinical and Safety teams, as deemed appropriate.

9.2.1.5 Adverse Event by Subgroup

Overview of TEAE and ASI will be assessed by number of prior biologics (0 vs ≥ 1). The number and percentage will be provided for each treatment group. Treatment difference in AE percentage between risankizumab 150 mg and the placebo group will be presented with point estimate and 95% confidence interval.

9.2.2 Analysis of Adverse Events for Long-Term

Long-term adverse event rates will be analyzed using event rates adjusted by cumulative exposure and will be based on the actual treatment received at the time of AE occurrence for the Any risankizumab 150 mg group. This includes AEs occurred under risankizumab 150 mg exposure from subjects starting on risankizumab 150 mg and subjects switching from placebo to risankizumab 150 mg.

9.2.2.1 Overview of Adverse Events Rates per 100 Patient-Years of Study Drug Exposure

An overview of TEAE per 100 patient-years of study exposure will be presented for the AE categories defined in Section 9.2.1.1.

An overview of ASI per 100 patient-years of study exposure will be presented similarly.

9.2.2.2 Adverse Events Rates per 100 Patient-Years of Study Drug Exposure by SOC and PT

The TEAE rate per 100 patient-years of exposure will be calculated overall, for each SOC and PT, for the same AEs defined in Section 9.2.1 and reported for Any risankizumab 150 mg group.

9.2.2.3 Serious Adverse Events (Including Deaths) and Adverse Events Leading to Study Drug Discontinuation

SAEs (including deaths) and AEs leading to study drug discontinuation will be listed in tables, besides summary by SOC and PT covered in Section 9.2.2.2.

9.2.2.4 Areas of Safety Interest Rates per 100 Patient-Years of Study Drug Exposure

The ASI categories will be summarized and presented for the Any risankizamub 150 mg group using SOC and MedDRA PT. The ASI categories will be identified per Standard MedDRA Queries (SMQs)/Company MedDRA Queries (CMQs) in Appendix B. Serious opportunistic infection excluding tuberculosis and herpes zoster, serious hypersensitivity, serious hepatic events and serious injection site reaction will also be summarized by PT for the Any risankizamub 150 mg group.

The ASI rate per 100 patient-years of exposure will be calculated overall, for each SOC and each PT, for each of the category listed in Appendix B.

Similarly, overview table for serious ASI rate per 100 patient-years of exposure will be summarized.

9.3 Analysis of Laboratory Data

Data collected from central and local laboratories, including additional laboratory testing due to an SAE, will be used in all analyses, except for Baseline where SAE-related laboratory assessments on or before the first dose of study drug will be excluded. The clinical laboratory tests defined in the protocol operations manual (e.g., hematology and clinical chemistry) will be summarized.

Each laboratory variable will be summarized for all time points (starting with Baseline) with the number of non-missing observations, mean and standard deviation, median, minimum and maximum. Mean change from baseline to each applicable post-baseline visit will be summarized for selected laboratory variables to compare the treatment difference between risankizumab 150 mg and placebo in Period 1. The number of observations, baseline mean, visit mean as well as the change from baseline mean, standard error, and 95% confidence interval within each treatment group and between treatment groups (risankizumab 150 mg vs. placebo) will be presented.

Changes in laboratory parameters will be tabulated using shift tables by CTCAE criteria v4.03 (Appendix G).[4] A shift table from baseline to the worse value (based on CTCAE criteria v 4.03) during treatment will be created. A similar shift table will be provided to summarize shifts from baseline to the final post-baseline value.

Laboratory abnormalities meeting CTCAE criteria grade 3 and higher will be summarized.

Laboratory abnormalities will be evaluated based on Potentially Clinically Significant (PCS) criteria (Appendix C). For each laboratory PCS criterion, the number and percentage of subjects who have a laboratory value meeting the criteria will be summarized. Listings will be provided to summarize subject-level laboratory data for subjects meeting PCS criteria. A listing of possible Hy's Law cases will be provided.

9.3.1 Variables and Units

All laboratory parameters to be collected in this study are listed below. Laboratory parameters will be reported using the standard international (SI) units.

Table 2. List of Laboratory Variables

Laboratory Variables
Hematology
White Blood Cell (WBC) Count
Red Blood Cell (RBC) Count
Hemoglobin
Hematocrit
Platelets count
Neutrophils
Basophils
Eosinophils
Lymphocytes
Monocytes
Chemistry
Total Bilirubin
Alkaline Phosphatase (ALP)
Aspartate aminotransferase (AST)
Alanine aminotransferase (ALT)
Total Protein
Albumin
Glucose
Triglycerides
Blood Urea Nitrogen (BUN)
Creatinine
Sodium
Potassium
Calcium
Chloride
Bicarbonate
Gamma Glutamyl Transferase (GGT)
Cholesterol
LDL cholesterol
HDL cholesterol

9.3.2 Analysis of Laboratory Data for Period 1

The laboratory data will be summarized by the actual treatment groups: risankizumab 150 mg and placebo group.

9.3.2.1 Assessment of Mean Change from Baseline in Clinical Laboratory Variables

Analyses of key continuous hematology and chemistry variables which are measured longitudinally will be performed by visits and by treatment group. For each parameter at each visit, the following summary statistics will be presented for each treatment group: sample size, mean, standard deviation, minimum, median and maximum.

An ANOVA model with treatment as a factor will be used to compare change from baseline between different treatment groups for selected laboratory parameters. Mean difference from placebo and associated 95% CIs will be presented. The analysis applies to the following laboratory parameters of clinical interest

- Hematology parameters: WBC, neutrophils, lymphocytes, RBC, hematocrit, hemoglobin, platelets count.
- Chemistry parameters: total bilirubin, GGT, ALP, SGOT/AST, SGPT/ALT, glucose, creatinine, potassium, sodium, calcium.

9.3.2.2 Assessment of Shift from Baseline in Clinical Laboratory Variables

The baseline and post-baseline laboratory observations will be categorized as Grade 0, Grade 1, Grade 2, Grade 3, and Grade 4 and shifts from baseline grade to worst on-therapy grade will be summarized. A similar shift table will be provided to summarize shifts from baseline to the final post-baseline value. Toxicity grading scale is based on CTCAE version 4.03. Shift tables from baseline according to the grades will be provided for laboratory variables including SGPT/ALT, SGOT/AST, GGT, ALP, total bilirubin, creatinine, hemoglobin, neutrophils count, WBC, and lymphocyte count.

Note that the minimum/maximum category is used, rather than the category of the minimum/maximum value. The two may be different due to variation in the reference range.

No statistical tests will be performed for this analysis.

9.3.2.3 Assessment of Potentially Clinically Significant Laboratory Values

Laboratory abnormalities will be evaluated based on PCS criteria (Appendix C) (defined as a lab value ≥ grade 3 NCI CTCAE v4.03 severity). For each laboratory PCS criterion, the number and percentage of subjects who have at least one post-baseline on-treatment value meeting the criteria will be summarized by actual treatment group. A listing of all observations collected will be generated for subjects who had at least one post-baseline observation meeting pre-defined criteria for potentially clinically significant values. A post baseline value must be more extreme than the baseline value to be considered a potentially clinically significant finding.

9.3.2.4 Assessment of Liver Function test Elevations

According to FDA's Guidance for Industry "Drug-Induced Liver Injury: Premarketing clinical evaluation" (July 2009), when aminotransferase abnormalities indicating hepatocellular injury are accompanied by evidence of impaired hepatic function (bilirubin elevation > 2 × ULN), in the absence of evidence of biliary obstruction (i.e., significant elevation of ALP) or some other explanation of the injury (e.g., viral hepatitis, alcohol hepatitis), the combined finding (i.e., Hy's Law cases) represents a signal of a potential for the drug to cause severe DILI.

For the purpose of assessing for potential Hy's law cases, the frequencies and percentages of subjects with post baseline liver specific function test values that meet the following criteria of potential clinical interest will be summarized by treatment groups:

- ALT ≥ 3 × ULN
- ALT ≥ 5 × ULN

- ALT ≥ 10 × ULN
- ALT ≥ 20 × ULN
- AST ≥ 3 × ULN
- AST ≥ 5 × ULN
- AST ≥ 10 × ULN
- AST ≥ 20 × ULN
- TBL ≥ 2 × ULN
- Alkaline phosphatase ≥ 1.5 × ULN
- ALT and/or AST ≥ 3 × ULN and concurrent TBL ≥ 1.5 × ULN
- ALT and/or AST ≥ 3 × ULN and concurrent TBL ≥ 2 × ULN

A listing of potentially clinically significant liver function laboratory values will also be provided.

In addition, the Evaluation of Drug-Induced Serious Hepatotoxicity (eDISH) plot [3] will be generated to evaluate the liver safety profile of risankizumab. eDISH plot is a log/log display of correlation between peak total bilirubin (TBIL) vs. ALT, both in multiples of ULN, with horizontal and vertical lines indicating Hy's law thresholds, i.e., ALT = 3 × ULN and total bilirubin = 2 × ULN. The eDISH plot makes immediately evident subjects potentially matching Hy's law laboratory criteria, all located in the upper right quadrant of the graph. Data points in the lower right quadrant, i.e., exceeding 3 × ULN for ALT, but being below 2 × ULN for total bilirubin, suggest an increased risk for liver injury as well, if incidence is differing between active treatment and control groups, however, not to the same extent and with less specificity as compared to Hy's law.

Figure 2. An Example of eDISH Plot

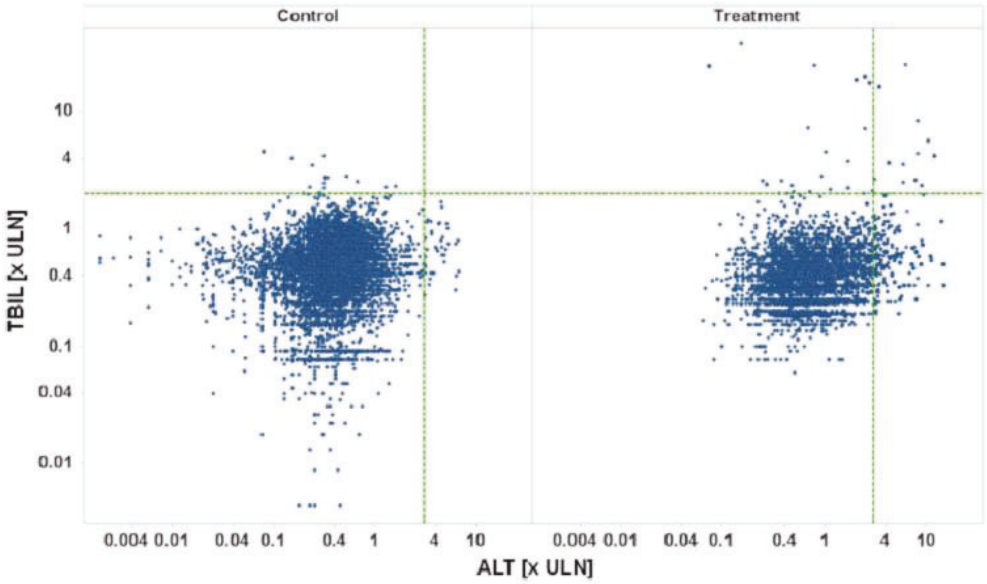


Fig. 1 eDISH plot, TBIL [× ULN] vs. ALT [× ULN] on a log/log scale, treatment by panel, pooled active versus control. *ULN* upper limit of normal, *ALT* alanine aminotransferase, *TBIL* total bilirubin

Note: Merz 2014.[3]

9.3.3 Analysis of Long-Term Laboratory Data

9.3.3.1 Assessment of Mean Change from Baseline in Clinical Laboratory Variables

Analyses of specified continuous hematology and chemistry variables which are measured longitudinally will be performed by visits and by actual treatment group sequences as described in Section 9.1. For each parameter, the following summary statistics will be presented for each treatment group sequence: sample size, mean, standard deviation, minimum, median and maximum.

9.3.3.2 Assessment of Potentially Clinically Significant Laboratory Values

Long-term laboratory data will also be summarized based on the number and percentage of subjects meeting the criteria for potentially clinically significant (PCS) laboratory values for the Any Risankizumab 150 mg group.

The baseline value will be determined by the last non-missing measurement recorded on or before the date of the first dose of risankizumab 150 mg. For a subject who started on placebo and switched to risankizumab 150 mg at Week 24, lab values under risankizamab 150 mg exposure would be evaluated against the baseline value defined as above.

A listing of all subjects with any laboratory determination meeting CTCAE criteria of Grade 3 or higher will be provided by Grade. For each of these subjects, the whole course of the respective parameter will be listed.

9.3.3.3 Assessment of Liver Elevations

The frequencies and percentages of subjects with post-baseline liver-specific function test values that meet the criteria of potential clinical interest defined in Section 9.3.2.4 will be summarized for Any risankizumab 150 mg group. A listing of potentially clinically significant liver function laboratory values will also be provided. For each of these subjects, the whole course of the respective parameter will be listed.

In addition, the eDISH plot as defined in Section 9.3.2.4 will be provided for Any risankizumab group.

9.4 Analysis of Vital Signs

Vital sign measurements of systolic blood pressure, diastolic blood pressure, pulse rate, and weight will be summarized.

Vital sign variables will be evaluated based on potentially clinically significant (PCS) criteria (Appendix C).

9.4.1 Analysis of Vital Sign for Period 1

Analyses of continuous vital sign variables which are measured longitudinally will be performed by visits and by the treatment groups of risankizumab 150 mg and placebo group. For each analysis, the following summary statistics will be presented for each treatment group: sample size, mean, standard deviation, minimum, median and maximum.

The number and percentage of subjects meeting the criteria for PCS vital sign values will be summarized by actual treatment group. Listings will be provided to summarize subject-level vital sign data for subjects meeting PCS criteria. For each of these subjects, the whole course of the respective parameter will be listed.

9.4.2 Analysis of Long-Term Vital Sign

Analyses of continuous vital signs variables which are measured longitudinally will be performed by visits and by actual treatment group sequences as described in Section 9.1.

Long-Term vital sign will also be summarized based on the number and percentage of subjects meeting the criteria for PCS vital sign values for the Any risankizumab 150 mg group. A listing of all subjects with any vital sign values meeting the criteria for PCS vital signs will also be provided.

10.0 Other Analyses

No other analyses are planned.

11.0 Interim Analysis

There will be no formal interim efficacy analysis planned before the Primary Analysis for this study. After the Primary Analysis, additional interim database locks may be conducted, as deemed appropriate, to provide updated summaries of long-term efficacy and safety.

11.1 Data Monitoring Committee

An external independent data monitoring committee (DMC) will review unblinded safety data at regular intervals during the conduct of the study. The primary responsibility of the DMC will be to protect the safety of the subjects participating in this study. When needed, high-level unblinded efficacy data may be requested by the DMC and be reviewed so that the DMC can assess benefit: risk of any emerging safety differences.

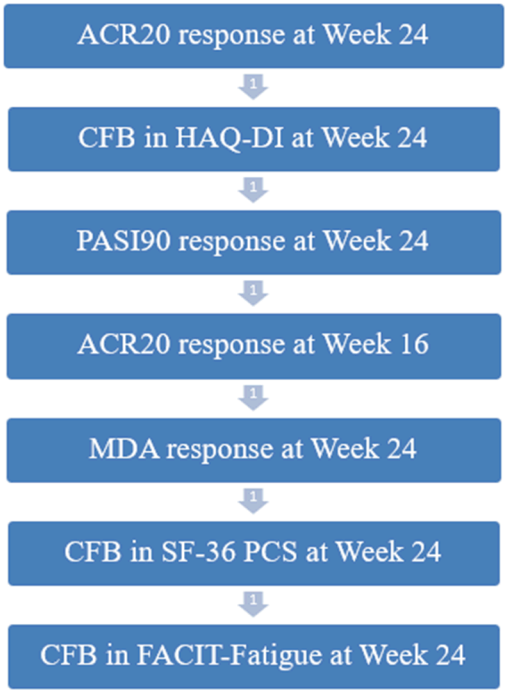
A separate DMC charter describes the roles and responsibilities of the DMC members, frequency of data reviews, relevant data to be assessed, and general operations.

12.0 Overall Type-I Error Control

A multiple testing procedure will be used to provide strong control of the type 1 error rate at alpha = 0.05 (2-sided) across analyses comparing risankizumab versus placebo with respect to the primary and the ranked secondary endpoints for Period 1.[2] Specifically, the testing will utilize a fixed sequence of hypothesis testing for the primary endpoint followed by the ranked secondary endpoints in the order as specified in Section 3.2. The test starts with the primary endpoint using two-sided $\alpha = 0.05$; testing can be conducted for a lower ranked endpoint only if the previous endpoints in the sequence meet the requirement of statistical significance.

The graph for the testing procedure is provided in Figure 3. In the graph, the arrows specify the α transfer paths. Once an endpoint is rejected (i.e., deemed statistically significant) at its assigned significance level, its significance level will be transferred to the subsequent endpoint following the arrow.

Figure 3. Graphical Testing Procedure



ACR 20 = 20% improvement in American College of Rheumatology criteria; CFB = change from baseline; DMARD-IR = disease-modifying antirheumatic drugs-intolerant or inadequate responder; FACIT Fatigue = Functional Assessment of Chronic Illness Therapy-Fatigue; HAQ-DI = Health Assessment Questionnaire-Disability Index; MDA = minimal disease activity; PASI 90 = 90% improvement in Psoriasis Area and Severity Index in subjects with BSA ≥ 3% at baseline; SF-36 PCS = Short Form-36 Physical Component Summary

13.0 Version History

Table 3. SAP Version History Summary

Version	Date	Summary
1.0	04 Aug 2017	Original version
2.0	23 Aug 2018	The following changes have occurred between the first version and the second version in order to reflect changes in the protocol and regulatory feedback. • An update to the study design. • The addition of the Analysis Population and Important Subgroups section • The addition of the Demographics and Baseline Characteristics section • The addition of subgroups and modification from 2 to 3 categories for Baseline BMI. • The addition of BASDAI 50 response to Section 7.0. • An increase in the sample size to 420 subjects • Change in the order of the ranked secondary endpoints • Addition of several non-ranked secondary endpoints • Changes in the interim analysis section to reflect unblinded analyses • Changes in the missing data section to reflect changes in the planned analysis and sensitivity analyses • Changes in the safety analysis section in order to detail short- and long-term safety analyses. • Various editorial changes to enhance readability.
3.0	21 Feb 2019	The following changes have occurred between the second version and the third version in order to reflect changes in the protocol and regulatory feedback. • Correction of typographical errors • Updated Section 13.0 (Safety Analyses) to reflect that Period 1 analyses will be performed at Week 24 and Week 52.

Table 3. SAP Version History Summary (Continued)

Version	Date	Summary
4.0	19 Mar 2020	In addition to changes to adhere to a new SAP formatting standard the following additional sections have been included between the third version and the fourth version in order to add additional information. • Subject Disposition • Study Drug Duration and Compliance • Medical History • Prior and Concomitant Medications The following changes have occurred in order to reflect changes in the protocol and regulatory guideline. • Updated secondary endpoints including replacing change from baseline in LDI/LEI with resolution of LDI/LEI as ranked secondary endpoints as well as adding ACR50/70 as other secondary endpoints. • Added estimand language for primary and secondary endpoints. • Provided analysis details for efficacy analysis in double-blind period and long-term efficacy analysis, including but not limit to, supplementary analysis for primary and secondary endpoints (e.g., MI, tipping point analysis), sensitivity analysis of primary endpoint based on Per Protocol analysis set. • Revised multiple testing procedure by removing Hochberg. • Revised and provided analysis details for safety analysis including but not limit to ASI, lab mean change summary, PCI grading.

Table 3. SAP Version History Summary (Continued)

Version	Date	Summary
5.0	14 Sep 2020	The following changes have occurred • Updated missing data handling due to COVID-19 infection or logistic restrictions. • Updated intercurrent event definition. • Added ACR20 at Week 16 as ranked secondary endpoint. • Moved resolution of LDI and resolution of LEI from ranked secondary endpoints to other secondary endpoints. • Added additional efficacy endpoints related to BASDAI and ASDAS. • Updated break-down summary of non-responders for ACR20 at Week 24.
6.0	16 Nov 2020	The following changes have occurred • Added additional database lock to support regulatory submission. • Added rescue medication summary. • Removed AE by relationship /severity in E/100PY in long-term. • Removed fungal infections and added serious anaphylactic reaction in ASI. • Updated laboratory parameters of clinical interest for the treatment comparison. • Updated liver elevations criteria according to FDA's Guidance for Industry "Drug-Induced Liver Injury (July 2009)."

14.0 References

1. Rubin DB. Multiple Imputation for Nonresponse in Surveys. New York: John Wiley & Sons; 1987.
2. Bretz F, Maurer W, Brannath W. A graphical approach to sequentially rejective multiple test procedures. Stat Med. 2009;28(4):586-604.
3. Merz M, Lee KR, Kullak-Ublick GA, et al. Methodology to assess clinical liver safety data. Drug Saf. 2014;37 (Suppl 1):S33-45.

4. National Cancer Institute. CTCAE v. 4.0. Available from: http://ctep.cancer.gov/protocolDevelopment/electronic_applications/ctc.htm#ctc. Accessed on: August 10, 2017.
5. Carpenter JR, Kenward MG. Multiple Imputation and Its Application. New York: John Wiley & Sons; 2013.
6. Greenland S, Rothman KJ. Introduction to Stratified Analysis. In: Rothman KJ, Greenland S, an Lash TL, editors. Modern Epidemiology. 3rd ed. Philadelphia: Lippincott Williams & Wilkins; 2008. p. 258-82.

Appendix A. Protocol Deviations

The number and percentage of subjects who reported at least one of the following protocol deviation categories will be provided.

- Subject entered into the study even though s/he did not satisfy entry criteria.
- Subject developed withdrawal criteria during the study and was not withdrawn.
- Subject received wrong treatment or incorrect dose of study drug.
- Subject received prohibited concomitant medication.

Appendix B. ASI for Risankizumab with SMQs/CMQs/PTs Searches

ASI	Search Criteria	
MACE	Adjudicated terms will be identified using CECAT and CETERM from the CE SDTM dataset.	
Extended MACE	Adjudicated terms will be identified (for MACE +) using CECAT and CETERM from the CE SDTM dataset.	
Serious Infections	Serious AEs in the Infections and Infestations SOC	
Active Tuberculosis	Active Tuberculosis CMQ (code 80000188)	
Opportunistic Infection excluding tuberculosis and herpes zoster	Opportunistic Infection excluding tuberculosis and herpes zoster CMQ (code 80000189)	
Herpes Zoster	Herpes Zoster CMQ (code 80000175)	
Malignant Tumours	Narrow	Malignant tumours (SMQ 20000194)
Non-melanoma Skin Cancer (NMSC)	Broad	Skin malignant tumours (SMQ 20000204) excluding terms identified by the Melanoma CMQ (code 80000119)
Malignant Tumours excluding NMSC	'Malignant Tumours excluding NMSC' is identified by the 'Malignant Tumours' search excluding terms identified by the 'Non-melanoma skin cancer (NMSC)' search.	
Hypersensitivity	Narrow	Hypersensitivity (SMQ 20000214)
Serious Anaphylactic Reactions	Narrow Serious AEs in the Anaphylactic Reaction SMQ (SMQ 20000021)	
Adjudicated Anaphylactic Reactions*	Adjudicated terms will be identified using SDTM data (e.g., CE and PR domains).	
Hepatic Events	Broad	Hepatic failure, fibrosis and cirrhosis and other liver damage-related conditions (SMQ 20000013)
	Broad	Hepatitis, non-infectious (SMQ 20000010)
	Broad	Cholestasis and jaundice of hepatic origin (SMQ 20000009)
	Broad	Liver related investigations, signs and symptoms (SMQ 20000008)
	Narrow	Liver-related coagulation and bleeding disturbances (SMQ 20000015)
Injection Site Reactions	Injection Site Reaction CMQ (code 80000019)	

* Events will be identified for adjudication by Anaphylactic Reaction SMQ Broad search as specified in the RISA AAC Charter.

Appendix C. Potentially Clinically Significant Criteria for Safety Endpoints

The criteria for Potentially Clinically Significant (PCS) laboratory and vital sign findings are described in Table C-1, Table C-2, and Table C-3.

Table C-1. Criteria for Potentially Clinically Significant Hematology Values

		Definition of Potentially Clinically Significant Current (Version 4) CTCAE Grade 3 or Greater
Hematology Variables	**Units**	**Very Low**
Hemoglobin	g/dL	< 8.0
Platelets count	10^9/L	< 50.0
WBC count	10^9/L	< 2.0
Neutrophils	10^9/L	< 1.0
Lymphocytes	10^9/L	< 0.5

Table C-2. Criteria for Potentially Clinically Significant Chemistry Values

		Definition of Potentially Clinically Significant Current (Version 4) NCI CTCAE Grade 3 or Greater	
Chemistry Variables	**Units**	**Very Low**	**Very High**
TBL	mcmol/L		> 3.0 × ULN
ALP	U/L		> 5.0 × ULN
SGOT/AST	U/L		> 5.0 × ULN
SGPT/ALT	U/L		> 5.0 × ULN
Albumin	g/L	< 20	
Glucose	mmol/L	< 2.2	> 13.9
Creatinine	mcmol/L		> 3.0 × ULN (> 3.0 × BL)
Sodium	mmol/L	< 130	> 155
Potassium	mmol/L	< 3.0	> 6.0
Calcium	mmol/L	< 1.75	> 3.1
GGT			> 5.0 × ULN

Table C-3. Criteria for Potentially Clinically Significant Vital Sign Values

Vital Sign	Category	Criteria for Potential Clinically Significant Vital Signs
Systolic blood pressure	Low	Value ≤ 90 mmHg and decrease ≥ 20 mmHg from Baseline
	High	Value ≥ 160 mmHg and increase ≥ 20 mmHg from Baseline
Diastolic blood pressure	Low	Value ≤ 50 mmHg and decrease ≥ 10 mmHg from Baseline
	High	Value ≥ 100 mmHg and increase ≥ 10 mmHg from Baseline

Appendix D. Definition of Estimand for Primary and Secondary Endpoints

Estimand	Attributes of Estimand				
	Treatments	**Population**	**Variable**	**Intercurrent Events (IE)**	**Population Level Summary**
Composite estimand for binary endpoints[a] (primary)	Risakizumab 150 mg vs placebo	Subjects who were randomized and received at least one dose of study drug.	ACR20/50/70 response/PASI 90 response/MDA response/resolution of LEI/resolution of LDI at Week 24, ACR20 at Week 16 (initiation of concomitant medications for PsA use that could meaningfully impact efficacy assessment or rescue as non-responder)	The IE is captured through the variable definition	Difference in response proportions between treatment conditions
Treatment policy estimand for binary endpoints[a] (supplementary)	Risakizumab 150 mg vs placebo (regardless of adherence to study drug or use of concomitant medications for PsA use that could meaningfully impact efficacy assessment or rescue medication)	Subjects who were randomized and received at least one dose of study drug	ACR20/50/70 response/PASI 90 response/MDA response/resolution of LEI/resolution of LDI at Week 24, ACR20 at Week 16	The IE is captured through the treatment definition	Difference in response proportions between treatment conditions

Estimand	Attributes of Estimand				
	Treatments	**Population**	**Variable**	**Intercurrent Events (IE)**	**Population Level Summary**
Hypothetical estimand for continuous endpoints[b] (primary)	Risakizumab 150 mg vs placebo	Subjects who randomized and received at least one dose of study drug	Change from baseline in HAQ-DI/SF-36 (PCS)/FACIT-fatigue at Week 24	Had subjects not initiated concomitant medications for PsA use that could meaningfully impact efficacy assessment or not rescued	Difference in variable means between treatment conditions
Treatment policy estimand for continuous endpoints[b] (supplementary)	Risakizumab 150 mg vs placebo (regardless of adherence to study drug or use of concomitant medications for PsA use that could meaningfully impact efficacy assessment or rescue medication)	Subjects who randomized and received at least one dose of study drug	Change from baseline in HAQ-DI/SF-36 (PCS)/FACIT-Fatigue at Week 24	The IE is captured through the treatment definition	Difference in variable means between treatment conditions

a. Binary endpoints include ACR20/50/70, PASI 90, MDA, resolution of LEI and resolution of LDI.

b. Continuous endpoints include HAQ-DI, SF-36 (PCS), FACIT-fatigue.

Appendix E. NRI-C Procedure

Composite Binary Endpoints

Step 1:

Missing binary values are first imputed via MI under MAR assumption using the FCS method with the logistic regression option. The FCS method can handle any arbitrary missing data pattern, without monotone missing requirement. Treatment is included in the FCS imputation model to enable stratified sampling conditional on treatment groups. The imputation model includes gender, race (white vs. non-white), ethnicity, age, baseline BMI, geographic regions, duration of PsA, baseline values of the components of the composite binary endpoint, stratification factors, intercurrent event status, as well as longitudinal response observed at all other visits. The number of imputed datasets is 30 for each PROC MI. Sample codes are as follows:

proc mi data=data1 seed=&seed nimpute=30 out=mar_imp;

```
class &predcate &postout;
fcs logistic;
by treatment;
var &predcate &predcont &postout;
run;
```

Note: &predcate (categorical covariates): gender, race (white vs. non-white), ethnicity, geographic regions, concomitant csDMARD at baseline (yes vs no), prior biologic use, extent of psoriasis, and intercurrent event status. &predcont (continuous covariates): age, baseline BMI, duration of PsA, baseline components. &postout: Binary endpoint at week 4, 8, 12, 16, 24

Step 2:

For each imputed dataset, the imputed response status for missing due to reasons other than COVID-19 is overridden by NRI to ensure that multiple imputation is only applied to missing due to COVID-19. Subjects will be considered as non-responders after initiation of rescue medication or initiation of concomitant medications for PsA use that can

meaningfully impact efficacy assessment (including COVID-19 impacted visits of which response status is imputed by MI in step 1).

Step 3:

For each of the 'complete' datasets, the CMH test is performed adjusting for the stratification factors to estimate the treatment difference of risankizumab versus placebo and the corresponding standard error. This is performed with the SAS procedure PROC STDRATE using the Mantel-Haenszel method (Greenland and Rothman 2008, p. 271).[6] Sample SAS codes are as follows:

```
proc stdrate data=data2 method=mh stat=risk effect=diff;
population group=TRTP event=N_RESPONDERS total=N_TOTAL;
strata STRATA / order=data stats effect;
run;
```

Note: The input dataset is summary level data including the following variables: treatment group (denoted by TRTP), stratification factors (denoted by STRATA), number of responders per study and treatment group (denoted by N RESPONDERS) and total number of subjects per study and treatment group (denoted by N_TOTAL).

The results from the 30 'complete' datasets will be synthesized using the SAS procedure PROC MIANALYZE, following Rubin's formula (Rubin, 1987)[1], to derive the MI estimator of the treatment difference for the final statistical inferences. Sample SAS codes are as follows:

```
proc mianalyze data=data3;
modeleffects RiskDiff;
stderr StdErr;
ods output parameterestimates=risk_diff_mh;
run;
```

Binary Endpoints Dichotomized from Continuous Scale

Step 1:

When a dichotomized variable is derived from a continuous scale, the MI is applied to the original continuous scale. Missing values are imputed via MI with the FCS method.

Treatment is included in the FCS imputation model to enable stratified sampling conditional on treatment groups. The imputation model includes gender, race (white vs. non-white), ethnicity, age, baseline BMI, geographic regions, duration of PsA, baseline value of the endpoint of interest (in continuous scale), stratification factors, intercurrent event status, as well as longitudinal response (in continuous scale) observed at all other visits. The number of imputed datasets is 30 for each PROC MI. Sample codes are as follows:

```
proc mi data=data1 seed=&seed nimpute=30 out=mar_imp;
class &predcate;
fcs;
by treatment;
var &predcate &predcont  &postout;
run;
```

Note: &predcate (categorical covariates): gender, race (white vs. non-white), ethnicity, geographic regions, current use of csDMARD, prior biologic use, extent of psoriasis, and intercurrent event status. &predcont (continuous covariates): age, baseline BMI, duration of PsA, baseline value. &postout: continuous scale of the binary endpoint at week 4, 8, 12, 16, 24

The dichotomized variable is then derived from the imputed value for each imputed dataset.

Step 2:

For each imputed dataset, the imputed response status for missing due to reasons other than COVID-19 is overridden by NRI to ensure that multiple imputation is only applied to missing due to COVID-19. Subjects will be considered as non-responders after initiation of rescue medication or initiation of concomitant medications for PsA use that can meaningfully impact efficacy assessment, including COVID-19 impacted visits of which response status is imputed by MI in step 1.

Step 3:

For each of the 'complete' datasets, the CMH test is performed adjusting the stratification factors to estimate the treatment difference of risankizumab versus placebo and the

corresponding standard error. The results from the 30 'complete' datasets will be synthesized using the SAS procedure PROC MIANALYZE, following Rubin's formula (Rubin, 1987)[1], to derive the MI estimator of the treatment difference for the final statistical inferences. Sample codes follow the same as provided above for composite binary endpoints (step 3).

Appendix F. Tipping Point Analysis

Tipping Point Analysis for ACR20

To assess the robustness of the primary analysis using NRI-C data handling, tipping point analysis is conducted on ACR20 at Week 24. The analysis is conducted on the FAS using AO data, i.e., all observed data regardless of study drug adherence or rescue.

The tipping point analyses are two-dimensional, i.e., assumptions about the missing outcomes on the risankizumab treatment group and the placebo group can vary independently. Missing binary values are first imputed via MI under MAR assumption using the FCS method with the logistic regression option. Two sets of shift parameters are applied to the imputed values by adding two MNAR statements to the same PROC MI, which allows the missing ACR20 response rate to systematically vary from 0% to 100% in both risankizumab and placebo, respectively. This is accomplished by modifying the predicted probabilities for the responses through shifting the log odds (shift parameters are log odds ratios)[5], then directly sampling the missing ACR20 response from the Bernoulli distribution with the modified probabilities. Sample codes are as follows:

```
proc mi data=data2 seed=&seed nimpute=30 out=mnar_imp;
class &predcate &postout;
fcs reg(BLj= &predcate &predcont, excluding BL);
fcs logistic(week4=&predcate &predcont);
fcs logistic(week8=&predcate &predcont week4);
fcs logistic(week12=&predcate &predcont week4 week8);
fcs logistic(week16=&predcate &predcont week4 week8 week12);
fcs logistic(week24=&predcate &predcont week4 week8 week12 week16);
by treatment;
var &predcate &predcont &postout;
mnar adjust(wk24(event="1") /shift=&sj1 adjustobs=(&treatment="PBO"));
mnar adjust(wk24(event="1")/ shift=&sj2 adjustobs=(&treatment="RIS"));
run;
```

Note: &predcate (categorical covariates): gender, race (white vs. non-white), ethnicity, geographic regions, concomitant csDMARD at baseline (yes vs no), prior biologic use, extent of psoriasis, and intercurrent event status. &predcont (continuous covariates): age, baseline BMI, duration of PsA, 7 baseline ACR components. &postout: ACR20 binary values at week 4, week 8, week 12, week 16, week 24. BL_j is the jth

baseline ACR components in the order of TJC, SJC, Pain, PGA, PhGA, HAQ-DI, CRP.

For each pair of shift parameters, the same CMH method used for the primary analysis will be performed on each of the multiple imputed datasets to obtain the results for each comparison of the risankizumab treatment group versus the placebo group adjusted by stratification factors. PROC MIANALYZE will then be used to aggregate the results for the final statistical inference using Rubin's method.

If one pair of shift parameters are found to just reverse the study conclusion, in terms of p-value larger than 0.05, then the shift parameters are identified as the tipping point. The results for a grid of shift parameter combinations are provided in tabular format.

Tipping Point Analysis for HAQ-DI

To assess the impact of potential departures from the MAR assumption, tipping point analyses are conducted as a sensitivity check for change from baseline in HAQ-DI at Week 24.

The tipping point analyses are two-dimensional, i.e., assumptions about the missing outcomes on the risankizumab treatment group and the placebo group can vary independently. Missing values are first imputed via MI under MAR assumption using AO data, and then a shift parameter is applied to the imputed values (a different shift parameter may be specified for each treatment group). This is implemented by PROC MI using the MNAR statement. The imputation model uses FCS method, thus does not rely on a monotone missing pattern. It includes the same variables as for binary endpoint and is also stratified by treatment group. The number of imputed datasets is 30 for each PROC MI.

In cases where the shifted values are smaller than the minimum or larger than maximum value of the endpoint, (i.e., out of range), the minimum or maximum value of the endpoint is used in further analysis steps. For each pair of shift parameters, the SAS procedure PROC MIXED is used for ANCOVA model which includes the fixed effects of treatment,

stratification factors and the continuous fixed covariate of baseline measurement on each of the imputed datasets to obtain the results for each risankizumab treatment group versus the placebo group comparison. These results will be aggregated using Rubin's method to get p-values.

If a pair of shift parameters are found to just reverse the study conclusion, in terms of p-value larger than 0.05, then the shift parameters are identified as the tipping point. The results for a grid of shift parameter combinations are provided in tabular format.

Appendix G. CTCAE Grades

Test	Grade 1	Grade 2	Grade 3	Grade 4
SGPT/ALT increased	> ULN – 3.0 × ULN	> 3.0 – 5.0 × ULN	> 5.0 – 20.0 × ULN	> 20.0 × ULN
SGOT/AST increased	> ULN – 3.0 × ULN	> 3.0 – 5.0 × ULN	> 5.0 – 20.0 × ULN	> 20.0 × ULN
GGT increased	> ULN – 2.0 × ULN	> 2.5 – 5.0 × ULN	> 5.0 – 20.0 × ULN	> 20.0 × ULN
ALP increased	> ULN – 2.5 × ULN	> 2.5 – 5.0 × ULN	> 5.0 – 20.0 × ULN	> 20.0 × ULN
TBL increased	> ULN – 1.5 × ULN	> 1.5 – 3.0 × ULN	> 3.0 – 10.0 × ULN	> 10.0 × ULN
Creatinine† increased	>1 - 1.5 x baseline (BL); > ULN – 1.5 × ULN	>1.5 - 3.0 x BL; > 1.5 – 3.0 × ULN	>3.0 x BL; > 3.0 – 6.0 × ULN	> 6.0 × ULN
Hemoglobin decreased	< LLN – 100.0 g/L	< 100.0 – 80.0 g/L	< 80.0 – 65.0 g/L	< 65.0 g/L
Neutrophil count decreased	< LLN – 1.5 × 10^9/L	< 1.5 – 1.0 × 10^9/L	< 1.0 – 0.5 × 10^9/L	< 0.5 × 10^9/L
WBC decreased	< LLN – 3.0 × 10^9/L	< 3.0 – 2.0 × 10^9/L	< 2.0 – 1.0 × 10^9/L	< 1.0 × 10^9/L
Lymphocyte count decreased	< LLN – 0.8 × 10^9/L	< 0.8 – 0.5 × 10^9/L	< 0.5 – 0.2 × 10^9/L	< 0.2 × 10^9/L

† If the calculation based on BL results in a different grade than the calculation based on ULN, use the higher grade.

Appendix H. Geographic Region

The following table lists the countries considered for each geographic region.

Geographic Region	Countries
Asia	Israel
Eastern Europe	Estonia, Hungary, Poland
South/Central America	Argentina, Brazil
North America	Canada, United States, Puerto Rico
Other	South Africa, Australia, New Zealand
Western Europe	Belgium, France, Germany, Greece, Italy, Portugal, Spain, United Kingdom, Denmark, Sweden