1.0 Title Page

Statistical Analysis Plan

Study 1311.3

ABBV-066 (Risankizumab) Versus Ustekinumab and Placebo Comparators in a Randomized Double Blind Trial for Maintenance Use in Moderate to Severe Plaque Type Psoriasis (UltIMMa-1)

Date: 12 Sep 2017

Version 2.0

2.0 Table of Contents

List of Tables

List of Figures

3.0 Introduction

This statistical analysis plan (SAP) describes the statistical analysis to be completed by the AbbVie Clinical Statistics Department for study Protocol 1311.3 dated 13 May 2016, which incorporates Amendments 1 (original Protocol: 19 October 2015).

This SAP will provide details to further elaborate statistical methods as outlined in the Protocol 1311.3 and will describe analysis conventions to guide the statistical programming work. The SAP will be signed off before the study database is locked.

Analyses will be performed using SAS® version 9.3 (SAS Institute, Inc., Cary, NC 27513) or higher using the UNIX operating system.

4.0 Study Objectives, Design and Procedures

4.1 Objectives

The objectives of this study are to assess the efficacy and safety of risankizumab compared to ustekinumab and placebo in subjects with moderate to severe chronic plaque psoriasis.

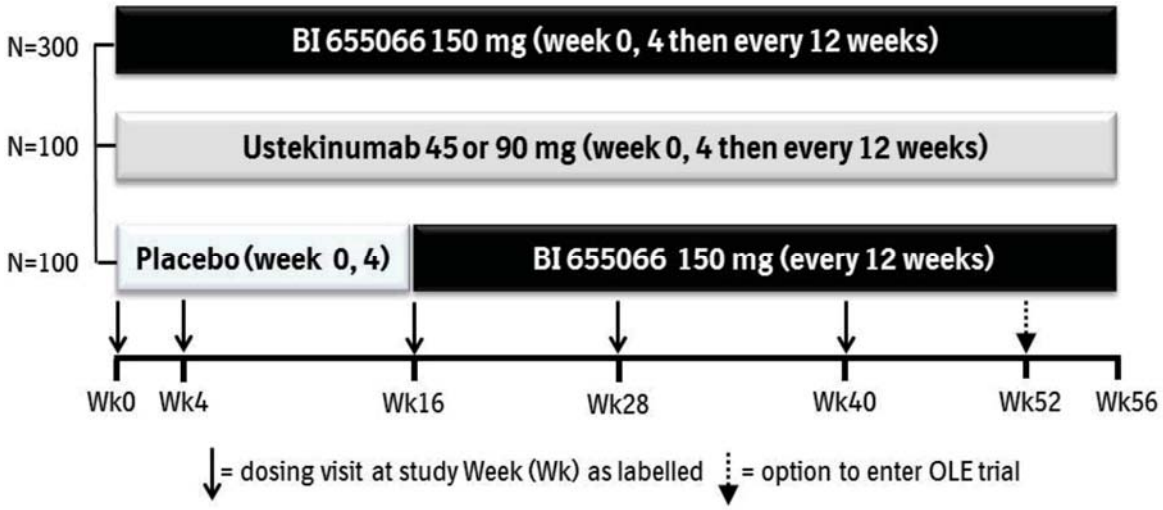
In addition, this trial will assess pharmacokinetics (PK) and the emergence of anti-drug antibodies (ADA) and their effect on efficacy and safety. Moreover, it will be explored how the use of risankizumab may influence gene and protein expression levels and disease specific protein markers. PK and ADA analyses, as well as DNA banking, exploratory biomarker, and metabolic risk factor analyses, will not be included in this report and may be analyzed in separate documents.

4.2 Design Diagram

This Phase 3 multi-national, randomized, double-blind, double dummy, placebo and active comparator controlled, parallel design study compares risankizumab with ustekinumab and placebo. In total, approximately 500 subjects with moderate to severe chronic plaque psoriasis will be randomized in this trial.

Subjects who have successfully finished screening will be eligible to participate in the 40 week treatment period and will be randomized at a ratio of 3:1:1 to one of 3 treatment arms (risankizumab, ustekinumab, and placebo) stratified by weight (≤ 100 kg versus > 100 kg) and prior exposure to TNF antagonists (0 versus ≥ 1) as shown in Figure 1.

Figure 1. Study Design Schematic



BI 655066 is risankizumab.

Subjects randomized to placebo at Week 0 will cross over to risankizumab at Week 16.

Subjects completing the study will be offered to roll over into an OLE trial, if they meet the inclusion criteria for the Open Label Extension (OLE) trial. Subjects who will not participate in the OLE trial will be followed up for adverse event (AE) assessment 16 weeks after having received their last dose.

4.3 Sample Size

The study is powered to show a benefit of risankizumab over ustekinumab in terms of achieving 90% reduction in Psoriasis Area and Severity Index (PASI 90) and achievement of static Physician Global Assessment (sPGA) clear or almost clear (0 or 1) at Week 16.

Based on the outcome from the Trials 1311.1 and 1311.2, the PASI 90 response rate at Week 16 is assumed to be at least 65% in the risankizumab arm and at most 45% in the ustekinumab arm. Using a 3:1 randomization scheme (risankizumab:ustekinumab), 300 subjects in the risankizumab arm and 100 in the ustekinumab arm will provide 94% power.

Based on the outcome from the Trials 1311.1 and 1311.2, the achievement of sPGA clear or almost clear rate at Week 16 is assumed to be at least 85% in the risankizumab arm and at most 67.5% in the ustekinumab arm. Using a 3:1 randomization scheme (risankizumab:ustekinumab 237 subjects in the risankizumab arm and 79 in the ustekinumab arm will provide 90% power. Using the PASI 90 required sample sizes, 300 subjects in the risankizumab arm and 100 subjects in the ustekinumab arm will provide at least 95% power for the sPGA endpoint.

Thus, the total sample size for this trial is 500 subjects (300 in risankizumab, 100 in ustekinumab, and 100 in placebo). Assuming a 5% response rate in the placebo group for both PASI 90 and sPGA clear or almost clear, this trial will have > 99% power for comparing the risankizumab arm to placebo on both of these endpoints. The assumptions used for sample size calculations are based on the current knowledge from the Phase I Study 1311.1, Phase II Study 1311.2, and Phase III studies of ustekinumab for moderate to severe chronic plaque psoriasis (Studies R11-1519, R13-3519, R13-3513, c03272682).

All calculations were performed using ADDPLAN Version 6.0.4, an Aptiv Solutions Company.

4.4 Interim Analysis

Unblinded data will be reviewed during the study by an independent data monitoring committee (DMC). The membership, roles and responsibilities and activities of the DMC have been defined in a written charter. The DMC will include representatives from external safety experts who are not directly involved in the study. Clinical site personnel and the study team will remain blinded to the randomized treatment assignments during

the course of the study. No decisions on efficacy are to be made for superiority or futility; therefore, no type I error adjustments for interim analysis are made. Communications from the DMC to the study team will not contain information that could potentially unblind the study personnel.

No other interim analyses are planned.

5.0 Analysis Populations and Stratification

5.1 Analysis Populations

In this SAP, Part A refers to Week 0 to Week 16, and Part B refers to Week 16 and after.

Efficacy Population:

Intent-to-Treat (ITT) Population: The ITT Population is defined as all subjects who are randomized at Week 0. ITT Population will be used for the efficacy analyses in Part A and Part B, as well as efficacy across the entire study (cross-period) for subjects who are randomized to risankizumab or ustekinumab. Subjects who are randomized to placebo in Part A and do not continue into Part B will be excluded from the analysis in Part B.

ITT Population will be analyzed by treatment group as randomized.

Per-protocol (PP) Population: To evaluate the impact of protocol deviations on the primary and ranked secondary endpoints, additional analyses will be performed on the Per-protocol Populations. The Per-protocol Population will include those who were most compliant with the protocol in ways that could impact the primary and ranked secondary endpoints.

Final results and the criteria for exclusion of subjects will be finalized prior to database lock.

The PP Population is defined as all subjects from the ITT Population who meet all the following criteria:

- Subjects must receive at least 75% of planned study drug injections (including dummy placebo injections). For the analyses of endpoints in Part A, only Part A doses (prior to Week 16) will be considered. For the analyses of endpoints at Week 52, the doses in the entire study will be considered for both risankizumab and ustekinumab arms.
- Subjects must have either a PASI or sPGA assessment post-baseline.
- Inclusion Criterion 4: Subjects must have stable moderate to severe chronic plaque psoriasis at baseline:
  - Have an BSA ≥ 10% and
  - Have a PASI ≥ 12 and
  - Have an sPGA ≥ 3.

Primary and ranked secondary endpoints pertaining to Week 12, Week 16, and Week 52 will be analyzed using PP population. PP Populations will be analyzed by treatment group as randomized.

In addition, PP analysis for a particular endpoint will also require subjects to have at least one post-baseline assessment for the corresponding endpoint. For example, proportion of subjects with PSS = 0 at Week 16 will be analyzed in the PP Populations among subjects who also have at least one post-baseline PSS assessment in Part A.

Safety Population:

The Safety Population will be used for safety analyses in both study parts. The Safety Population will be analyzed based on the actual treatment received at the randomization visit.

- The Safety Population is defined as all randomized subjects who received at least one dose of study drug in Part A.
- Subjects who discontinue study prior to Part B will be excluded from the safety analysis in Part B.

- Safety Population will be used for the safety analyses across the entire study (cross-period) for subjects who are randomized to risankizumab or ustekinumab.
- All Risankizumab Treated (ALL_RZB) Population is defined as subjects who receive at least one dose of risankizumab in the study. Key safety summaries will be provided.

Table 1. Notation for Treatment Groups

Population	Treatment Code	Definition
ITT/Safety (Part A)	RZB	Subjects who randomized to risankizumab 150 mg.
	UST	Subject who randomized to ustekinumab
	PBO	Subjects who randomized to placebo.
ITT/Safety (Part B)	RZB/RZB	Subjects who randomized to risankizumab 150 mg.
	UST/UST	Subject who randomized to ustekinumab
	PBO/RZB	Subjects who randomized to placebo and entered Part B
ITT/Safety (entire study)	RZB/RZB	Subjects who randomized to risankizumab 150 mg.
	UST/UST	Subject who randomized to ustekinumab
Safety ALL_RZB	RZB	Subjects who received at least one dose of risankizumab 150 mg.

5.2 Variables Used for Stratification of Randomization

Subjects will be randomized in blocks to double-blind treatment, stratified by weight (≤ 100 kg versus > 100 kg) and prior exposure to TNF antagonists (0 versus ≥ 1). Subjects will be randomized to risankizumab, ustekinumab, or placebo in a ratio of 3:1:1 within each level of stratification.

6.0 Analysis Conventions

Definition of Baseline

The last non-missing observation collected on or before the date of the first dose of study drug injection will be used as Baseline for summary of demographics and disease characteristics, safety and efficacy analyses, with the exception that:

- The last evaluation on or prior to the date of first dose of risankizumab will be used for the safety analyses in Part B among subjects randomized to placebo at Week 0, and the safety analyses in the ALL_RZB Population.
- For all vital sign assessments, only assessments prior to first dose time will be considered, since vital signs are to be assessed both pre- and post-dose in some visits.
- Analysis of Psoriasis Symptoms Scale will be defined in Section 10.1.5.

Entry of Part B is defined as the last non-missing observation collected on or before the date of the first dose of study drug in Part B for PBO/RZB arm and last non-missing observation collected on or before the date of Week 16 visit for RZB/RZB and UST/UST arms.

Definition of Final Observation (Applicable to Safety Analyses)

Final observation in Part A is defined as the last non-missing observation collected within 105 days following the last dose of study drug in Part A for those who do not enter Part B, and on or before the first dose date of study drug injection in Part B for subjects who entered Part B. For vital sign assessments, only assessments prior to the first dose of study drug injection in Part B will be included in the analysis for Part A.

Final observation for Part B is defined as the last non-missing observation collected within 105 days following the last dose in Part B.

Final observation for the entire study analyses is defined as the last non-missing observation collected within 105 days following the last dose of study drug.

Final observation for ALL_RZB Population is defined as the last non-missing observation collected within 105 days following the last dose of risankizumab.

Definition of Rx Days in Each Period (Days Relative to the Date of First Dose of Study Drug in Each Period)

Rx Days are calculated for each time point relative to the date of first dose of study drug. They are defined as the number of days between the day of the first dose of study drug and the specific time point. Rx Days are negative values when the time point of interest is prior to the first study drug dose day. Rx Days are positive values when the time point of interest is on or after the first study drug dose day. The day of the first dose of study drug is defined as Rx Day 1, while the day prior to the first study drug dose is defined as Rx Day –1 (there is no Rx Day 0).

Rx Days for Part A and entire study will be defined relative to the date of first dose of study drug. Rx days for Part B will be defined relative to the date of the date of the first study drug injection on or after Week 16.

Definition of Analysis Windows

All time points and corresponding time windows are defined based on Rx Days. The visit schedule with accompanying details can be found in Flow Chart and Section 6.2 of the protocol.

For efficacy analyses, local tolerability, ECG, laboratory parameters, and vital sign variables, analysis windows are constructed using the following algorithm:

- Determine the nominal Rx day for each visit (e.g., Week 4 [4 weeks after Baseline visit] equals Rx Day 29).
- In order to include all post baseline data, the first post-baseline interval starts on the first day after the first dose of study drug (Rx Day 2).
- Determine the window around a specific nominal Rx day by adding or subtracting half of the interval between adjacent visits (e.g., days between Week 4 and Week 8 is 28). The threshold between adjacent visits is determined by splitting the interval evenly between the visits. If the resulting split is between Rx days, then the threshold is determined as the midpoint between the adjacent visits. If the resulting split is on an Rx day, then the

threshold is determined as being between that Rx day and the Rx day prior to it (e.g., the split between Week 4 and Week 8 would be between Rx Days 43 and 44).

- If more than one assessment is included in a time window the assessment closest to the nominal day will be used. If there are two observations equidistant to the nominal day, the later one will be used in analyses. If more than one assessment is included on the same day, then the worst assessment on that day will be used in analyses, except those specified in Section 11.0.

The following analysis windows have been specified for the analysis of efficacy, laboratory parameters, and vital signs, etc. for each analysis period. In order to include all post baseline data, the first post baseline interval starts on the day after the date of first study drug injection in each period (Rx Day 2) for all efficacy variables.

For subjects who were randomized and not dosed date of randomization will be used for determining window definitions.

Table 2. Visit Windows for Analysis of PASI, sPGA, PSS, Vital Signs, and Local Tolerability in Part A (ITT/PP/Safety Populations)

Window Label	Target Day	Interval
Baseline	1	≤ 1[a]
Week 4	29	[2, 43]
Week 8	57	[44, 71]
Week 12	85	[72, 99]
Week 16	113	[100, 127[b]]

Rx Day calculated relative to first dose date of study drug. For subjects randomized but not dosed, Rx Day calculated relative to Randomization.

a. There is no baseline for local tolerability measurements. If time is collected in vital signs, restrict to records prior to the first dose of study drug.

b. The minimum of upper bound and the day of the first dose date in Part B (and within 105 days of last dose for safety analyses).

Table 3. Visit Windows for Analysis of Safety Laboratory Tests in Part A (Safety Population)

Window Label	Target Day	Interval
Baseline	1	≤ 1
Week 8	57	[2, 85]
Week 16	113	[86, 141*]

Rx Day calculated relative to first dose date of study drug.

* The minimum of upper bound, the day of the first dose date in Part B, and 105 days of last dose.

Table 4. Visit Windows for Analysis of ECG in Part A (Safety Population)

Window Label	Target Day	Interval
Baseline	1	≤ 1
Week 4	29	[2, 71]
Week 16	113	[72, 155*]

Rx Day calculated relative to first dose date of study drug.

* The minimum of upper bound, the first dose date in Part B, and the last dose + 105 days.

Table 5. Visit Windows for Analysis of NAPSI, PPASI, PSSI, HAQ-DI, Pain VAS, Patient Global Assessment VAS, SJC, TJC, and DAS 28 in Part A (ITT Population)

Window Label	Target Day	Interval
Baseline	1	≤ 1
Week 16	113	[2, 225*]

Rx Day calculated relative to first dose date of study drug. For subjects randomized but not dosed, Rx Day calculated relative to Randomization.

* The minimum of upper bound and the day of the first dose date in Part B.

Table 6. Visit Windows for Analysis of DLQI/EQ-5D-5L/HADS in Part A (ITT Population)

Window Label	Target Day	Interval
Baseline	1	≤ 1
Week 12	85	[2, 99]
Week 16	113	[100, 127*]

Rx Day calculated relative to first dose date of study drug. For subjects randomized but not dosed, Rx Day calculated relative to Randomization.

* The minimum of upper bound and the day of the first dose date in Part B.

Table 7. Visit Windows for Analysis of PASI, sPGA, and PSS, and Local Tolerability in Part B (ITT Population)

Window Label	Target Day	Interval
Entry of Part B	1	≤ 1
Week 22	43	[2, 64]
Week 28	85	[65, 106]
Week 34	127	[107, 148]
Week 40	169	[149, 190]
Week 46	211	[191, 232]
Week 52	253	[233, 274]

Rx Day calculated relative to the first dose date on or after Week 16 for PBO/RZB arm, and relative to the date of Week 16 visit for RZB/RZB and UST/UST arms. Local tolerability only evaluated from Week 22 to 40.

Table 8. Visit Windows for Analysis of Vital Signs in Part B (Safety Population)

Window Label	Target Day	Interval
Entry of Part B	1	≤ 1[a]
Week 22	43	[2, 64]
Week 28	85	[65, 106]
Week 34	127	[107, 148]
Week 40	169	[149, 190]
Week 46	211	[191, 232]
Week 52	253	[233, 267]
Week 56	281	[268, 295[b]]

Rx Day calculated relative to the first dose date on or after Week 16 for PBO/RZB arm, and relative to the date of Week 16 visit for RZB/RZB and UST/UST arms. Local tolerability only evaluated from Week 22 to 40.

a. If time is collected for vital sign, restrict to records prior to the first dose of study drug in Part B.

b. The minimum of upper bound and 105 days after the last dose.

Subjects who enter the Open Label Extension Study 1311.31 will not have Week 56 evaluations.

Table 9. Visit Windows for Analysis of Laboratory Tests and ECG in Part B (Safety Population)

Window Label	Target Day	Interval
Entry of Part B	1	≤ 1
Week 28	85	[2, 127]
Week 40	169	[128, 211]
Week 52	253	[212, 294[a]]

Rx Day calculated relative to the first dose date on or after Week 16 for PBO/RZB arm, and relative to the date of Week 16 visit for RZB/RZB and UST/UST arms.

a. The minimum of upper bound and 105 days after the last dose.

Table 10. Visit Windows for Analysis of NAPSI, PPASI, PSSI, DLQI, EQ-5-5L, and HADS in Part B (ITT Population)

Window Label	Target Day	Interval
Entry of Part B	1	≤ 1
Week 52	253	[2, 504]

Rx Day calculated relative to the first dose date on or after Week 16 for PBO/RZB arm, and relative to the date of Week 16 visit for RZB/RZB and UST/UST arms.

Table 11. Visit Windows for Analysis of HAQ-DI, PainVAS, Patient Global Assessment VAS, TJC, SJC, and DAS 28 in Part B (ITT Population)

Window Label	Target Day	Interval
Entry of Part B	1	≤ 1
Week 28	85	[2, 169]
Week 52	253	[170, 336]

Rx Day calculated relative to the first dose on or after Week 16.

Table 12. Visit Windows for Summary of Study Drug Injections in Part A

Window Label	Target Day	Time Window
Week 0	1	≤ 1
Week 4	29	[2, 57[a]]

Rx Day calculated relative to the first dose date of study drug.

a. Before the first dose date on or after Week 16.

Table 13. Visit Windows for Summary of Study Drug Injections in Part B

Window Label	Target Day	Time Window
Week 16	1	≤ 1
Week 28	85	[2, 127]
Week 40	169	[128, 211]

Rx Day calculated relative to the date of first dose date on or after Week 16.

Definition of Missing Data Imputation

No global imputation is taking place at the database level. Efficacy related imputations are outlined in Section 10.0. There is no imputation for missing values in the safety analyses.

Rounding of Numeric Results

Rounding will be performed for presentation of results. No rounding will be performed before or during analyses. The ROUND function of SAS will be used to round results.

When dichotomizing continuous variables from single observations, associated continuous variables will be rounded to 9 decimal points before applying the cutoff point to determine the response status (for example, percent change from baseline in PASI score will be rounded to 9 decimal places before comparing to 90%).

When dichotomizing continuous variables based on weekly rolling average (achievement of PSS score of 0), the weekly rolling average will be rounded to integers before dichotomizing using the ROUND function of SAS.

The mean and median will be rounded for presentation to 1 decimal more than the data entered into the database. The standard deviation will be rounded to 2 decimal places more than the data entered into the database. The minimum and maximum values will be presented as entered into the database.

Probabilities will be rounded to 3 decimal places before assignment of statistical significance and will be presented in rounded format. Probabilities that round to zero or are reported by SAS as zero will be presented as "< 0.001." Probabilities that round to 1 or are reported by SAS as 1 will be presented as "> 0.999."

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

7.1 Demographic and Baseline Characteristics

Demographic and baseline characteristics will be summarized for each arm and for overall of the ITT population. The number of observations, mean, standard deviation, median, minimum and maximum will be summarized for continuous variables; and treatment groups will be compared using a one-way analysis of variance (ANOVA) model with treatment as the independent factor. Categorical or discrete variables will be summarized via counts and percentages; and treatment groups will be compared using a two-sided Pearson's Chi-Square test (or an appropriate exact test if expected cell count < 5).

The following demographic and baseline parameters will be summarized.

Subject Demographics

- Sex (male, female)
- Age (years), defined as the number of years from date of birth to date of first drug
- Age categories (< 40 years, ≥ 40 – < 65 years, ≥ 65 years)
- Race (White, Black, Asian, American Indian/Alaska Native, Native Hawaiian or Other Pacific Islander, Multi Race)
- Ethnicity (Hispanic or Latino, Other)
- Body weight (kg)
- Body weight category (≤ 100 kg, > 100 kg)
- Height (cm)
- BMI (kg/cm^2)
- BMI category (< 25, ≥ 25 – < 30, ≥ 30)
- Prior exposure to TNF antagonists (0 versus ≥ 1)

General Baseline Characteristics

- PASI (Psoriasis Area and Severity Index)
- BSA (Body Surface Area)
- sPGA categories
- NAPSI (Nail Psoriasis Severity Index)
- PSSI (Psoriasis Scalp Severity Index)
- PPASI (Palmoplantar Psoriasis Area Severity Index)
- Dermatology Life Quality Index (DLQI)
- HAQ-DI (HAQ Disability Index)
- Pain VAS (Visual Analog Scale)/Patient Global Assessment (PtGA) VAS (Visual Analog Scale)
- TJC/SJC (Tender/Swollen Joints)
- DAS28 (Disease Activity)

Psoriasis and Cardiovascular History

- Psoriatic arthritis (diagnosed, suspected, no)
- Cardiovascular Diseases (myocardial infarction, angina pectoris, transient ischemic attack, stroke, deep vein thrombosis)
- Cardiovascular Risk Factors (hypertension, hyperlipidemia, diabetes mellitus, obesity)

General Use

- Smoking history (Never-smoked, Ex-smoker, Currently smokes)
- Alcohol History (Non-drinker, drinks – no interference to study participation, drinks – possible interference to study participation)

Prior Treatment:

- Psoriasis TNF antagonist and Other Biologic Usage History – by Response to Treatment
- Psoriasis Treatment – by Therapy Type (topical therapy, phototherapy, photochemotherapy, non-biologic systemic therapy, TNF antagonist, other biologic)

Physical Exam and Pregnancy Test will be presented in listing format.

7.2 Medical History

Medical history will be summarized using body systems and condition/diagnosis as captured on the eCRF. The body systems will be presented in alphabetical order and the conditions/diagnoses will be presented in alphabetical order within each body system. The number and percentage of subjects with a particular condition/diagnosis will be summarized for each treatment arm. Subjects reporting more than one condition/diagnosis within a body system will be counted only once for that body system.

7.3 Previous Treatment and Concomitant Medications

Prior and concomitant medications will be summarized by generic name. A prior medication is defined as any medication taken at the time of entry into the study and prior to the first dose of study drug. A concomitant medication is defined as any medication that started prior to the first dose of study drug and continued to be taken after the first dose of study drug or any medication that started after the first dose of study drug, but not after the last dose of study drug plus 21 days. The number and percentage of subjects who had taken medications will be summarized by generic drug name assigned by the World Health Organization (WHO) for both prior and concomitant medications.

7.4 Protocol Deviations

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 she/he did not satisfy entry criteria
- Subject who developed withdrawal criteria during the study and was not withdrawn
- Subject who received wrong treatment or incorrect dose
- Subject who received excluded or prohibited concomitant treatment

8.0 Patient Disposition

The number of subjects for each of the following categories will be summarized, for overall and for each treatment group in the ITT population:

- Number of subjects randomized
- Number of subjects treated
- Number of subjects who completed the study
- Number of subjects who discontinued the study drug
- Number of subjects who prematurely discontinued study

In addition, the reasons for premature discontinuation will be summarized with frequencies and percentages.

In addition to patient disposition, number of screening failures and reasons for screening failure will also be summarized among all screened subjects in a subject screening status table.

The summaries will be provided for Part A and Part B.

9.0 Study Drug Exposure and Compliance

Summary of study drug treatment duration and compliance will be provided for each treatment arm in each treatment period under the ITT Population.

Study drug treatment duration (days) will be summarized using the sample size, mean, standard deviation, minimum, median and maximum. In addition cumulative exposure of RZB (including total patient years) will also be summarized in the ALL_RZB Population. Study drug treatment duration will be summarized as follows:

Study Drug Exposure (in Days) in Each Period:

Part A:

- For subjects who did not continue into Part B:
  - Date of last injection in Part A – Date of first injection in Part A + 84 days.
- For subjects who continued into Part B:
  Minimum of
  - Date of first injection in Part B – Date of first injection in Part A
  - Date of last injection in Part A – Date of first injection in Part A + 84 days.

Part B:

Date of last injection in Part B – Date of first injection in Part B + 84 days

ALL_RZB:

For study drug exposure during the administration of risankizumab in the ALL_RZB Population:

- Date of last injection of risankizumab – Date of first injection of risankizumab + 84 days.

Compliance

There will be a summary of the number of subjects receiving study drug and dose at each study drug administration visit. This will be repeated on the cumulative number of doses.

When computing compliance at each study drug administration visit, the denominator will include all subjects in each analysis population who have not prematurely discontinued the study drug prior to the scheduled study drug injection. Subjects who have prematurely discontinued the study drug but have not prematurely discontinued the study are not used in the denominator.

10.0 Efficacy Analysis

10.1 General Considerations

The treatment effect will be evaluated based on a two-sided significance level of 0.05 (when rounded to three decimal places).

The ITT population will be used for the analyses of efficacy endpoints. The following treatment periods will be analyzed:

- Part A (up to Week 16)
- Part B (Week 16 and after)

Time-to-event analysis will be performed across the entire study for subjects who are randomized to risankizumab or ustekinumab in addition to the analyses in Part A.

Primary and ranked secondary variables will be analyzed for the per-protocol populations using primary approach to handle missing data as described below.

Subjects' actual weight category and prior TNF antagonist exposure will be used as the strata in the stratified analyses.

10.1.1 Analyses of Categorical Variables

For categorical variables, frequencies and percentages will be summarized. Treatment comparison will be conducted using a Cochran-Mantel-Haenszel (CMH) test with stratification factors as strata for the analysis. The CMH test will use weights proposed by Greenland & Robins, which is calculated as follows:

$\hat{\delta}_{MH} = \frac{\sum_{i=1}^{u} w_i \cdot \hat{\delta}_i}{\sum_{i=1}^{u} w_i}$, where

$\hat{\delta}_i = \frac{x_i}{n_i} - \frac{y_i}{m_i}$ denotes the risk difference in stratum $i, i = 1, \cdots, u$

$w_i = \frac{n_i \cdot m_i}{n_i + m_i}$ denotes the weight of stratum $i, i = 1, \cdots, u$

x_i denotes the number of subjects with event in treatment_1 in stratum i, $i = 1, \cdots, u$

y_i denotes the number of subjects with event in treatment_2 in stratum i, $i = 1, \cdots, u$

n_i denotes the number of subjects on treatment_1 in stratum i, $i = 1, \cdots, u$

m_i denotes the number of subjects on treatment_2 in stratum i, $i = 1, \cdots, u$

The estimated variance of $\hat{\delta}_{MH}$ is calculated as:

$$\widehat{var}\left(\hat{\delta}_{MH}\right) = \frac{\sum_{i=1}^{u} L_i}{\left(\sum_{i=1}^{u} w_i\right)^2}$$

where $L_i = \frac{x_i(n_i - x_i)\, m_i^3 + y_i(m_i - y_i)\, n_i^3}{n_i \cdot m_i \cdot (n_i + m_i)^2}$, $i = 1, \cdots, u$

Assuming a normal distribution of $\hat{\delta}_{MH}$, an approximate 95% CI is given as follows, where $z_{0.975}$ is the 97.5% quantile of the standard normal distribution:

$$CI = \left[\hat{\delta}_{MH} \pm z_{0.975} \cdot \sqrt{\widehat{var}\left(\hat{\delta}_{MH}\right)}\right]$$

Also, the approximate p-value can be calculated using the following:

$$\text{pvalue} = 2 \cdot \Pr\left[Z > \left|\frac{\hat{\delta}_{MH}}{\sqrt{\widehat{var}\left(\hat{\delta}_{MH}\right)}}\right|\right], \text{where } Z \sim N(0,1)$$

If there is a stratum for a treatment group that has 0 subjects in any cell in the contingency table, all cells from the stratum will be added by 0.1 in order to prevent dividing by 0 in the above equations, as suggested in Greenland & Robins.

10.1.2 Analyses of Continuous Variables

Change from baseline in PSS will be analyzed by stratified van Elteren test. For other continuous variables, the model based mean and standard error will be presented. The Baseline and visit means will also be presented for each treatment group for subjects who have both Baseline and post Baseline visit values. The treatment groups will be compared using ANCOVA with treatment group, Baseline value, and stratification factors in the model.

10.1.3 Analyses of Time-to-Event Variables

Time-to-event analysis will be done for all three arms for Part A, and RZB/RZB and UST/UST treatment arms for the entire study.

For the time to first achievement of Endpoint, the time to event will be calculated as:

- Time to first achievement (with observed event) = [date of first achievement] – [date of first study drug] + 1
- If a subject never attains an endpoint, then that subject's time to first achievement will be censored at the last visit where variable was measured. Censoring will be done within Part A for the analyses in Part A, and over the entire study for cross-period analyses. For example, if the first time a PASI 90 response is achieved in Part B, then PASI 90 will be censored for Part A analysis only.

Time to loss of response will be evaluated from the time of achievement of the same endpoint, and defined using the following algorithm:

- Never attains Endpoint (Failure at time 0)
- After achieving Endpoint, subjects will be a considered loss of response if they subsequently lost the response or discontinue from the study due to AE of "Worsening of disease under study."
- Subjects will be censored at their last assessment if they do not lose the endpoint or discontinue from the study due to reasons other than AE of "Worsening of disease under study." Time to failure will be calculated starting from date of first achievement of the endpoint.

Time to event will be analyzed using Kaplan-Meier estimates for each treatment group. Treatments comparisons will be performed using stratified Log-rank test.

10.1.4 Missing Data Imputations

Missing data will be imputed using the following methods for the efficacy analyses:

- Non-Responder Imputation (NRI): the NRI analysis will categorize any subject who has a missing value at a specific visit as a non-responder for that visit. The only exception is when the subject is a responder both before and after a specific visit window, and then the subject will be categorized as a responder for the visit. This before- and after-window imputation will utilize

visits across Part A and B for the risankizumab and ustekinumab arms. The NRI will be the primary approach in the analyses of categorical variables.

- Last Observation Carried Forward (LOCF): The LOCF analyses will use the completed evaluation from the previous visit for efficacy measures assessed to impute missing data at later visits. For observations in RZB/RZB, and UST/UST arms in Part B, LOCF is continued across Part A and Part B. LOCF will be applied to the observations in PBO/RZB within the analysis period. Baseline values will not be carried forward. LOCF will be the primary approach in the analyses of continuous variables, and the secondary approach in the analyses of categorical variables.
- As-Observed Cases (OC): The as-observed 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 as-observed analysis for that visit. As-observed analysis will be the secondary approach in the analysis of continuous variables.
- Multiple Imputation (MI): The MI will be used as sensitivity approach to impute missing data in primary and ranked secondary endpoints. The variables to be included in the imputation model are listed below. If MI is not applicable due to the nature of our data (e.g., MCMC algorithm does not converge), logistic regression or mixed effect model repeat measurement (MMRM) methods will be applied as sensitivity approach, whichever applicable.

The Multiple Imputation analysis will be carried out in three steps.

- Imputation of missing data. The imputation will be generated for each efficacy endpoint measurement. The variables to be included in the imputation model are: Baseline disease severity (PASI and sPGA), Baseline weight, treatment group, prior TNF exposure, and measurements at each visit from randomization (or re-randomization) up to the end of the analysis period. For each endpoint, 20 'complete' datasets will be generated using SAS PROC MI. The imputed post-baseline measurements will be rounded to the same precision as the observed data before the determination of responder status (e.g., PASI 90).

- Analysis of imputed data sets. A CMH test, stratified by stratification factors, will be used to analyze categorical endpoints in each imputed dataset. A stratified van Elteren test will be used to analyze the change from baseline in PSS endpoint in each imputed dataset.
- Synthesis of imputation and analysis results. SAS PROC MIANALYZE will be used to generate the final inferences of the risk difference between treatment groups.

Of note, subjects who discontinued due to AE of "worsening of disease under study" will be counted as non-responders in all subsequent visits in the NRI and MI analyses, and will have their last observation prior to discontinuation carried forward in the LOCF analyses. For Observations in RZB/RZB, and UST/UST arms in Part B, LOCF is continued across Part A and Part B. LOCF will not carry forward from Period A to Period B for PBO/RZB arm.

10.1.5 Analysis of Psoriasis Symptoms Scale (PSS)

The 4-item PSS was designed to measure patient-reported psoriasis symptoms. Subjects are requested to rate the severity of each of their symptoms related to pain, redness, itching, and burning, in the last 24 hours from "0 = none" to "4 = very severe," with total scores, which is the sum of the four item responses, ranging from 0 to 16 with higher scores indicating worse symptoms. If one or more of the four items are missing then the total score for that day will be set to missing.

PSS will be completed daily in Part A. PSS from Day 1 that collects all items in the past 24 hours will be used as the Baseline value. The rolling weekly average of the total PSS score is calculated starting from Day 8 through Day 127 or the day of the first injection in Part B, whatever comes first, as follows. Let $P_{m-6}, P_{m-5}, \ldots, P_{m-1}, P_m$ be the daily total PSS scores from day $m - 6$ to day m, and N_m be the number of days with non-missing pain scores from day $m - 6$ to day m, then the rolling weekly average for day m is:

$$\frac{\sum_{i=m-6}^{m} P_i}{N_m}$$

If values from 4 or more days of the 7 day period are missing, then the rolling weekly average of day m will be set to missing. If more than one assessment is included on the same day, the worst (highest) assessment on that day will be chosen as the daily score. Analysis value for a given visit will be selected from rolling averages based on analysis window conventions.

PSS will be completed by the subject during clinic visits from Week 22 through Week 52 in Part B. For those visits, scores from single clinic visits will be used as the analysis value. Change from baseline in PSS evaluated using the total score on the PSS at Week 16 will be analyzed by the van Elteren test between the risankizumab arm and placebo arm using the stratification factors of previous TNF exposure (0 versus ≥ 1) and weight (≤ 100 kg vs > 100 kg). An example of the SAS code for this is as follows:

```
proc freq data = indata;

  tables strata * treatmnt * aval1 / cmh2 scores=modridit noprint;
```

where "strata" is numbered 1 – 4 based on the row order, "treatment" is dichotomous (Arm A or Arm B), and "aval1" is the PSS change from baseline at Week 16. In the output from this code, the p-value is the "Prob" for statistic 2 – "Row Mean Scores Differ."

In addition to the PSS total score, the four individual components, pain, redness, itching, and burning, will be analyzed using the same methodology using rolling weekly averages.

10.2 Primary Efficacy Analysis

There are two co-primary endpoints to assess the efficacy of risankizumab 150 mg for the treatment of moderate to severe plaque psoriasis. These are as follows:

- Achievement of ≥ 90% reduction from baseline PASI score (PASI 90) at Week 16
- Achievement of an sPGA of clear or almost clear at Week 16

The primary null hypothesis is that risankizumab 150 mg is not different from placebo with respect to achieving PASI 90 and achieving sPGA of clear or almost clear at Week 16.

The achievement of PASI 90 at Week 16 is the first co-primary endpoint. The difference in proportion of subjects achieving PASI 90 between treatment arms will be estimated and tested using the Cochran-Mantel-Haenszel risk difference estimate stratified by the stratification factors of baseline weight (≤ 100 kg vs. > 100 kg) and prior exposure to TNF antagonists (0 vs. ≥ 1).

The achievement of a sPGA of clear or almost clear at Week 16 is the second co-primary endpoint. The analysis method for the sPGA co-primary endpoint will be identical to that of the PASI 90 co-primary endpoint detailed above.

The co-primary endpoints of PASI 90 and sPGA clear or almost clear need to be significant simultaneously. Both endpoints will be tested using a two-sided test with a type I error rate of 0.05.

Non-responder imputation will be used as the primary approach for missing values. LOCF and MI will be performed as sensitivity analyses.

10.3 Secondary Efficacy Analyses

10.3.1 Ranked Secondary Efficacy Analysis

Ranked secondary endpoints will be analyzed in the ITT and PP Populations.

The following null hypotheses will be tested in a hierarchical order using two-sided tests with a type I error of 0.05, only if the null hypothesis for the co-primary endpoints has been rejected:

1. Risankizumab is not different from placebo with respect to achieving sPGA of clear at Week 16

2. Risankizumab is not different from placebo with respect to PASI 100 response at Week 16
3. Risankizumab is not different from placebo with respect to achieving a DLQI score of 0 or 1 at Week 16
4. Risankizumab is not different from placebo with respect to achieving a PSS score of 0 at Week 16
5. Risankizumab is not different from ustekinumab with respect to PASI 90 response at Week 16
6. Risankizumab is not different from ustekinumab with respect to achieving a sPGA of clear or almost clear at Week 16
7. Risankizumab is not different from ustekinumab with respect to PASI 100 response at Week 16
8. Risankizumab is not different from ustekinumab with respect to achieving a sPGA of clear at Week 16
9. Risankizumab is not different from ustekinumab with respect to PASI 90 response at Week 52
10. Risankizumab is not different from ustekinumab with respect to PASI 100 response at Week 52
11. Risankizumab is not different from ustekinumab with respect to achieving a sPGA of clear at Week 52
12. Risankizumab is not different from ustekinumab with respect to PASI 75 response at Week 12
13. Risankizumab is not different from ustekinumab with respect to achieving a sPGA of clear or almost clear at Week 12
14. Risankizumab is not different from ustekinumab with respect to achieving a DLQI score of 0 or 1 at Week 16

15. Risankizumab is not different from placebo with respect to mean change from baseline in PSS total score at Week 16

10.3.2 Other Secondary Efficacy Analysis

Other secondary endpoints are as follows:

- Achievement of PASI 75 at Week 16
- Achievement of a sPGA of clear or almost clear at Week 52
- Achievement of PASI 75 at Week 52

10.3.3 Further Efficacy Endpoints

Further endpoints are as follows:

- Achievement of PASI 50 at all visits collected
- Achievement of PASI 75 at all visits collected
- Achievement of PASI 90 at all visits collected
- Achievement of PASI 100 at all visits collected
- Time to first achievement of PASI 50, PASI 75, PASI 90, PASI 100 and sPGA of clear or almost clear
- Time to loss of PASI 75, PASI 90, PASI 100 and sPGA of clear or almost clear response
- Change and percent change from baseline in PASI at all visits collected
- Achievement of PASI < 3 at all visits collected
- Achievement of a sPGA of clear or almost clear at all visits collected
- Achievement of a sPGA of clear at all visits collected
- Change from baseline in total score of the PSS at all visits collected
- Achievement of PSS total score of 0 at all visits collected
- Change from baseline in DLQI at all visits collected
- Achievement of a DLQI score of 0 or 1 at all visits collected

- Achievement of a reduction of 5 or more points from baseline in DLQI score at all visits collected, among subjects with baseline DLQI ≥ 5
- Change from baseline in HADS at all visits collected
- Change from baseline in HAQ-DI at all visits collected, among subjects with PsA confirmed via ClASsification of Psoriatic Arthritis (CASPAR) and had baseline value > 0
- Achievement of a reduction of 0.3 or more from baseline HAQ-DI at all visits collected, among subjects with PsA confirmed via CASPAR and had baseline value ≥ 0.3
- Change and percent change from baseline on patient Pain VAS, among subjects with PsA confirmed via CASPAR and had baseline value > 0
- Achievement of a reduction of 10 or more from baseline Pain VAS at all visits collected, among subjects with PsA confirmed via CASPAR and had baseline value ≥ 10
- Change and percent change from baseline on patient Global Assessment VAS, among subjects with PsA confirmed via CASPAR and had baseline value > 0
- Change from baseline in Swollen or Tender Joint Count (28 joints) at all visits collected, among subjects with PsA confirmed via CASPAR and had baseline value > 0
- Change from baseline in DAS28 at all visits collected, among subjects with PsA confirmed via CASPAR and had baseline value > 0
- Change and percent change from baseline in Nail Psoriasis Severity Index (NAPSI) at all visits collected
- Change and percent change from baseline in Palmoplantar Psoriasis Severity Index (PPASI) at all visits collected
- Change and percent change from baseline in Psoriasis Scalp Severity Index (PSSI) at all visits collected
- Change from baseline in EQ-5D at all visits collected
- Achievement of an increase of 0.1 or more points from baseline in EQ-5D at all visits collected

In addition, the proportion of subjects achieving PASI 90, PASI 100, sPGA of clear, and sPGA of clear or almost clear at each visit in Part B will be summarized among those who had PASI 90, PASI 100, sPGA of clear, and sPGA of clear or almost clear, respectively, at the entry of Part B.

10.4 Handling of Multiplicity

The statistical comparisons for the primary efficacy variable and the ranked secondary variables will be carried out in the hierarchical order. This means that statistically significant results (p-value ≤ 0.05) for the comparison in the higher rank (primary, then ranked secondary variables) are necessary to initiate the testing of the next comparison in the lower rank. Since a step-down procedure is used, each comparison will be tested at a significance level of 0.05 and overall alpha level of 0.05 will be preserved.

10.5 Efficacy Subgroup Analysis

To evaluate the consistency of the efficacy over demographic and other baseline characteristics, summaries and analyses will be performed for the following subgroups for the primary efficacy endpoints.

- Age (< 40 years, $\geq 40 - < 65$ years, ≥ 65 years)
- Sex (male, female)
- Race (white, non-white);
- Smoking (current, ex or never)
- BMI (normal: < 25, over weight: $\geq 25 - < 30$, obese: ≥ 30);
- Baseline PASI score (by median);
- Baseline sPGA (3, 4)
- Psoriatic arthritis (yes [diagnosed or suspected], no)
- Ps Therapy History (Phototherapy or Photochemotherapy, TNF Antagonist, Other biologics, Non-biologic systemic therapy, All biologics, Naïve to all)
- Region (US, Asia, Other)

11.0 Safety Analysis

11.1 General Considerations

Safety analyses will include adverse events, laboratory, local tolerability, ECG, and vital sign measurements. Safety summaries will be provided using the safety population as defined in Section 5.1. Pairwise comparisons of risankizumab vs ustekinumab and placebo will be performed in the Safety Population. Continuous variables will be analyzed using one-way ANOVA and categorical variables will be analyzed using Fisher's exact test. For analyses of AEs, only P values ≤ 0.100, when rounded to three digits, will be presented.

Missing safety evaluations will not be imputed.

11.2 Analysis of Adverse Events

11.2.1 Treatment-Emergent Adverse Events

Treatment-emergent adverse events (TEAEs) are defined as any event with an onset that is after the first dose of study drug and with an onset date within 105 days after the last dose of study drug in the analysis period, or prior to the first dose in the subsequent period for subjects who entered in to the subsequent period. Events where the onset date is the same as the study drug start date are assumed to be treatment-emergent, unless the adverse event start time is prior to the study drug start time. If an incomplete onset date is collected for an adverse event, the event will be assumed to be treatment-emergent unless there is other evidence that confirms that the event is not treatment-emergent (e.g., the event end date is prior to the study drug start date).

Summary tables will be presented as follows:

1. **Adverse Event Overview**

   The number and percentage of subjects experiencing treatment-emergent adverse events will be summarized for the following adverse event categories.

   - Any AE

- Any AE that was assessed as related to study drug by the investigator
- Any severe AE
- Any serious AE (SAE)
- Any SAE that was assessed as related to study drug by the investigator
- Any AE leading to discontinuation of study drug.
- Any AE leading to death.
- Any deaths
- Areas of Safety Interest

2. Adverse Events by System Organ Class and Preferred Term

TEAEs will be summarized and presented using primary Medical Dictionary for Regulatory Activities (MedDRA) system organ classes (SOCs) and preferred terms (PTs). The SOCs will be presented in alphabetical order, and the PTs will be presented in alphabetical order within each SOC.

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.

In addition, the number and percentage of adverse events with causal relationship between the events and the study drug will be summarized using the same conventions described above.

3. Adverse Events by Maximum Severity

The severity grading of AEs follows Rheumatology Common Toxicity Criteria (RCTC).

- Grade 1 – mild
- Grade 2 – moderate
- Grade 3 – severe

- Grade 4 – life threatening

Adverse events will be summarized by maximum severity. If a subject has an adverse event with unknown severity, then the subject will be counted in the severity category of "unknown," even if the subject has another occurrence of the same event with a severity present. The only exception is if the subject has another occurrence of the same adverse event with the most extreme severity (life threatening). In this case, the subject will be counted under the "life threatening" category.

4. Adverse Events by Maximum Relationship

Adverse events will be summarized by maximum relationship to study drug, as assessed by the investigator. Relationship of an AE to study drug is assessed by the investigator and collected in the CRF as 'Yes' or 'No.' If a subject has an adverse event with unknown relationship, the subject will be counted in the relationship category of "unknown," even if the subject has another occurrence of the same event with a relationship of "No." If the subject has another occurrence of the same adverse event with a relationship assessment of "Yes," the subject will be counted under the "yes" category.

A listing of all pretreatment (i.e., events start prior to the first study drug injection) serious adverse events will be provided.

The following tables are planned.

Treatment-emergent adverse events will be summarized as follows:

- Grouped by SOC and PT
- Grouped by SOC, PT and maximum relationship to study drug
- Grouped by SOC, PT and maximum severity

Treatment-emergent SAEs will be summarized as follows:

- Grouped by SOC and PT

- A by-subject listing will be provided

Pre-treatment SAEs will be summarized as follows:

- A by-subject listing will be provided

Treatment-emergent adverse events leading to death or premature discontinuation of study drug will be summarized as follows:

- Grouped by SOC and PT
- Separate listings by subject for deaths and premature terminations of study drug due to adverse events will be provided.

Treatment-emergent areas of safety interest will be summarized as follows:

- Grouped by SOC and PT
- A listing by subject will be provided.

Areas of Safety Interest

Areas of Safety Interest groupings are listed in Table 14. These events are of interest due to a higher rate in the moderate to severe psoriasis population, or of interest for all immunoglobulin products or products in general.

The final list will be based on the most updated final version of risankizumab Product Safety Statistical Analysis Plan, which is consistent to the most updated risankizumab Risk Management Plan.

Table 14. Areas of Safety Interest

ASI Grouping	Categories (ASI)	Search Criteria	Terms to Display	Include in AE Overview (Y/N)
Adjudicated CV Events	MACE	Adjudicated events	Display underlined terms defined by the following adjudicated terms: • CV Death which includes CETERM values: Fatal CV, Fatal PE, Fatal Non-Cardiac/Non-Neuro Arterial Thrombosis/Thromboembolism, Undetermined Death, Not assessable death (cardiac/neuro/thrombotic), Fatal Stroke • Myocardial infarction • Stroke	Y
	Extended MACE	Adjudicated events	Display underlined terms from MACE and underlined terms below: • Hospitalization for Unstable Angina • Coronary Revascularization Procedures	N

Table 14. Areas of Safety Interest (Continued)

ASI Grouping	Categories (ASI)	Search Criteria	Terms to Display	Include in AE Overview (Y/N)
Adjudicated CV Events (continued)	Other CV events	Adjudicated events	Display underlined terms defined by the following adjudicated terms: • Thrombotic events which includes CETERM values: Deep Vein Thrombosis, TIA, Pulmonary Embolism, Non-fatal Non-Cardiac/Non-Neurological Arterial Thrombosis/ Thromboembolism, Other Venous Thrombosis, specified (non-fatal) • Cardiac arrhythmia which includes CETERM of: Clinically Significant Arrhythmia • Congestive heart failure which includes CETERM of Heart Failure • Hypertensive emergency	N
Serious infections, TB, fungal and opportunistic infections (including herpes zoster)	Serious infections	Serious PTs of the CMQ (company MedDRA query) Infections (CMQ 80000018)	PTs	Y
	TB	Tuberculosis (including Investigations) CMQ (code 80000033)	PTs	Y
	Opportunistic infections	Opportunistic infections CMQ (code 80000073)	PTs	N

Table 14. Areas of Safety Interest (Continued)

ASI Grouping	Categories (ASI)	Search Criteria	Terms to Display	Include in AE Overview (Y/N)
Serious infections, TB, fungal and opportunistic infections (including herpes zoster) (continued)	Fungal infections	Fungal infections CMQ (code 80000063)	PTs	N
	Herpes Zoster	Herpes zoster CMQ (code 80000175)	PTs	N
Malignancies	All possible malignancies	Narrow – Malignancies (SMQ 20000090)	PTs	N
	Malignant Tumours	Narrow – Malignant tumours (SMQ 20000194)	PTs	Y
	Non-melanoma skin cancer (NMSC)	Broad – Skin malignant tumours (SMQ 20000204) excluding terms identified by the Melanoma CMQ (code 80000119)	PTs	N
	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.	PTs	Y

Table 14. Areas of Safety Interest (Continued)

ASI Grouping	Categories (ASI)	Search Criteria	Terms to Display	Include in AE Overview (Y/N)
Hypersensitivity Reaction	Hypersensitivity	Narrow – Hypersensitivity (SMQ 20000214)	PTs	Y – serious events only
	Anaphylactic Reaction	Narrow – Anaphylactic reaction (SMQ 20000021)	PTs	N
Hepatic Events	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)	PTs	N

Adverse Event per 100 Patient-Years of Exposure

AEs occurring during the entire study will be presented by event rate per 100 patient-years. These will be presented for any TEAEs, serious adverse events, Areas of Safety Interest.

AEs per 100 patient-years of exposure is defined as the number of AEs divided by the total exposure in 100 patient-years. Note that one event per preferred term per day per subject will be counted in the calculation of the number of AEs (i.e., a preferred term will not be counted twice on the same day for the same subject). See the calculation method below.

$$100 \times \frac{\text{Number of TEAEs}}{\text{Total Patient Years}}$$

where total patient years is defined as the sum of the study drug exposure (defined as date of last dose – date of first dose + 105 days (5 half-lives)) of all subjects normalized by 365.25, and rounded to one decimal place.

11.2.2 SAEs (Including Deaths) and Adverse Events Leading to Study Drug Discontinuation

Deaths and all SAEs will be presented in listing format. In addition, SAEs will be summarized by System Organ Class and MedDRA Preferred Term.

11.2.3 Safety Subgroup Analysis

The AE overview and AE by SOC and preferred term in each period will also be analyzed with respect to the actual values of stratification factors:

- Dichotomous weight (≤ 100 kg vs. > 100 kg)
- Prior exposure to TNF antagonists (0 versus ≥ 1)

11.3 Analysis of Laboratory Data

For the assessments of laboratory data, values observed more than 105 days after the last dose of study drug will be excluded.

Listing and descriptive statistics of laboratory values over time, change from baseline, and extreme abnormal value on treatment will be provided. Extreme abnormal value is the value which is most significantly away from the reference range. Frequency of subjects with transitions relative to reference range and listing of subjects with significant abnormal laboratory values will be presented as well.

11.3.1 Variables and Criteria Defining Abnormality

Clinical laboratory tests performed by central laboratory service provider are listed below.

Table 15. Clinical Laboratory Tests

Category	Test Name
Haematology	Hematocrit (Hct) Hemoglobin (Hb) Glycosylated Hbc (HbA1c) White Blood Cells/Leukocytes Platelet Count/Thrombocytes
Diff. Automatic	Neutrophils (relative and absolute count) Eosinophils (relative and absolute count) Lymphocytes (relative and absolute count)
Enzymes	AST (GOT) ALT (GPT) Alkaline Phosphatase (AP) Creatine Kinase (CPK) Gamma-Glutamyl Transferase (GGT/γ-GT) Lactic Dehydrogenase (LDH) Amylase Lipase
Electrolytes	Calcium Sodium Potassium
Substrates	Glucose BUN Uric acid Creatinine eGFR (estimated by CKD-EPI formula) Bilirubin Total Albumin C-Reactive Protein (CRP) (high sensitivity) Cholesterol, total Triglycerides LDL-Cholesterol HDL-Cholesterol
Urinalysis (dipstick)	Urine Protein Urine Glucose

11.3.2 Statistical Methods

Analysis of Continuous Laboratory Parameters

Though the protocol indicates utilizing the Rheumatology Common Toxicity Criteria (RCTC) scale for grading laboratory values, given that the NCI National Cancer Institute Common Toxicity Criteria for Adverse Event (CTCAE) scale includes a more comprehensive list of laboratory values; therefore, the lab analyses based on the NCI CTCAE scale will be presented. Change from Baseline to each scheduled visit and to the final value in continuous laboratory parameters will be summarized with the mean, standard deviation and median. The Baseline and visit/final value means will also be presented for subjects who have both the Baseline and visit/final values (see Section 6.0 for the definition of Baseline and final values). If there are multiple post-baseline measurements on the same day, the average value will be used.

Shift Tables

Shift tables for changes from Baseline according to the normal range will be provided for each hematology and clinical chemistry parameter. Shifts from Baseline to the following endpoints will be considered: minimum value, maximum value and final value. Categories of "low or normal" and "high or normal" will be included at Baseline in addition to the categories of "low," "normal," "high" and "missing." If there are multiple measurements on the same day, the last value will be used.

Potentially Clinically Important Laboratory Values

Frequencies and percentages of subjects with post Baseline lab values meeting the following criteria will be summarized. Of note, a post baseline value must be more extreme than the baseline value to be considered a potentially clinically important finding.

Table 16. Criteria for Potentially Clinically Important Chemistry Values

		Definition of Potentially Clinically Important Current (Version 4) NCI CTCAE Grade 3 or Greater	
Chemistry Variables	**Units**	**Very Low**	**Very High**
TBL	mcmol/L		> 3.0 × 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
Triglycerides	mmol/L		> 5.7
Creatinine	mcmol/L		> 3.0 × ULN
Sodium	mmol/L	< 130	> 155
Potassium	mmol/L	< 3.0	> 6.0
Calcium	mmol/L	< 1.75	> 3.1
CPK	U/L		> 5.0 × ULN
Total Cholesterol	mmol/L		> 10.34
GGT			> 5.0 × ULN
ALP			> 5.0 × ULN

Table 17. Criteria for Potentially Clinically Important Hematology Values

		Definition of Potentially Clinically Important Current (Version 4) 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

A separate listing will be provided that presents all of the subjects and values that are NCI CTCAE toxicity grade 3 or above. For each of these subjects, the whole course of the respective parameter will be listed. The NCI CTCAE grading is shown in Table 18 below:

Table 18. NCI CTCAE Grading

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 – 2.5 × ULN	> 2.5 – 5.0 × ULN	> 5.0 – 20.0 × ULN	> 20.0 × ULN
ALP increased	> ULN – 3.0 × ULN	> 3.0 – 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	> ULN – 1.5 × ULN	> 1.5 – 3.0 × ULN	> 3.0 – 6.0 × ULN	> 6.0 × ULN
CPK increased	> ULN – 2.5 × ULN	> 2.5 – 5.0 × ULN	> 5.0 – 10.0 × ULN	> 10.0 × ULN
Hemoglobin decreased	< LLN – 100.0 g/L	< 100.0 – 80.0 g/L	< 80.0	
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 there are multiple post-baseline measurements on the same day, the worst value will be used.

Liver Function Tests

Additional summaries will be presented for liver function tests including ALT or serum glutamic-pyruvic transaminase (SGPT), AST or serum glutamic-oxaloacetic transaminase (SGOT), alkaline phosphatase, and total bilirubin. Each laboratory value will be categorized as follows:

- < 1.5 × ULN
- ≥ 1.5 × ULN – < 3.0 × ULN
- ≥ 3.0 × ULN – < 5.0 × ULN
- ≥ 5.0 × ULN – < 10.0 × ULN
- ≥ 10.0 × ULN – < 20.0 × ULN

- ≥ 20.0 × ULN

Shift tables of Baseline to the maximum (relative to the normal range, i.e., the largest multiple relative to the upper limit of normal) values, and from Baseline to final value will be presented using these categories. A listing of potentially clinically important liver function laboratory values will be provided. The listing will include all subjects who met any of the following four criteria:

- ALT ≥ 3 × ULN, or
- AST ≥ 3 × ULN, or
- Alkaline phosphatase ≥ 1.5 × ULN, or
- Total bilirubin ≥ 2 × ULN.

A listing of possible Hy's Law cases, defined as those who meet all of the following conditions will be provided:

- ALT > 3 × ULN or AST > 3 × ULN
- Associated with an increase in bilirubin ≥ 2 × ULN
- Alkaline phosphatase < 2 × ULN.

All analyses will be conducted in the Safety Population. Table for potentially clinically important lab values will also be summarized in the ALL_RZB Population.

11.4 Analysis of Vital Signs and Weight

11.4.1 Variables and Criteria Defining Abnormality

The following vital sign parameters will be assessed: Systolic blood pressure [mmHg], Diastolic blood pressure [mmHg], Pulse [beats per minute], Respiratory rate [breaths per minute], Temperature [°C], Weight [kg]. The following table presents the Criteria for Potentially Clinically Important Vital Sign Findings. Of note, a post baseline value must be more extreme than the baseline value to be considered a potentially clinically important finding.

Table 19. Criteria for Potentially Clinically Important Vital Sign Findings

Vital Sign	Category	Criteria for Potential Clinically Important 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 ≥ 15 mmHg from Baseline
	High Value	≥ 105 mmHg and increase ≥ 15 mmHg from Baseline
Pulse	Low Value	≤ 50 bpm and decrease ≥ 15 bpm from Baseline
	High Value	≥ 120 bpm and increase ≥ 15 bpm from Baseline

11.4.2 Statistical Methods

Changes from Baseline to each visit and to the final value in vital sign parameters will be summarized with the mean, standard deviation and median. The Baseline and final value means will also be presented for subjects who have both the Baseline and final values (see Section 6.0 for the definition of Baseline and final values).

For baseline, if there are multiple measurements on the same day, the last measurement prior to the first dose of study drug will be used as the Baseline vital sign value. If there are multiple post-baseline measurements on the same day, average value will be used.

For systolic blood pressure, diastolic blood pressure and pulse, a listing of all subjects with any vital sign value meeting criteria for potentially clinically important values will be provided. For each of these subjects, the whole course of the respective parameter will be listed. The number and percentage of subjects who have at least one value meeting criteria for potentially clinically important values will be provided for each selected vital sign parameter.

All analyses will be conducted in the Safety Population. Table for potentially clinically important vital signs will also be summarized in the ALL_RZB Population.

11.5 Analysis of ECG Parameters

The ECG parameters will be assessed as scheduled in the study protocol.

Summary statistics for mean change from baseline for corrected QT interval (QTc) using Bazett (QTcB) and Fridericia (QTcF) corrections, aggregated, will be provided by analysis visits.

- Values for both QTcF and QTcB interval measurements will be categorized into the following: ≤ 450 ms, > 450 ms, > 480 ms, > 500 ms, or missing. For the scheduled visits, as well as baseline, the number and percentage of subjects within each category will be presented. Additionally, for each category, the number and percentage of subjects with a maximum QTcF interval falling into the category will be presented; a similar summary will be presented for QTcB intervals.
- For QTcF and QTcB intervals, the changes from baseline will be categorized into the following: < 30 ms, 30 ≤ – < 60 ms, > = 60 or missing. For the scheduled visits, the number and percentage of subjects within each category will be presented. Additionally, for each category, the number and percentage of subjects with a maximum change from baseline in the QTcF interval falling into the category will be presented; a similar summary will be presented for QTcB intervals.

Analyses will be conducted in the Safety Population.

11.6 Local Tolerability

Local tolerability at the administration site of the subcutaneous injection will be assessed by the investigator in the measurements of "swelling," "induration," "heat," "redness," "pain," or "other findings."

Proportion of subjects reporting each condition will be analyzed. The overall rate (with at least one occurrence) during Part A and B will also be presented.

Analyses will be conducted in the Safety Population in Part A and in Part B.

12.0 Pharmacokinetic Analysis

Pharmacokinetic analysis is not covered in this SAP.

13.0 Biomarkers Analysis

Biomarkers Analysis is not covered in this SAP.

14.0 Summary of Changes

14.1 Summary of Changes Between the Latest Version of Protocol and the Current SAP

- Defined analysis populations (including the Per-protocol population) in detail.
- Adjusted secondary endpoint rankings in the null hypothesis test hierarchy, based on the discussion with FDA.
- Added additional ranked secondary endpoints:
  - Achievement of sPGA of clear at Week 16, comparing to placebo.
  - Achievement of PASI 100 at Week 16, comparing to placebo.
  - Achievement of sPGA of clear at Week 16, comparing to ustekinumab.
  - Achievement of sPGA of clear at Week 52, comparing to ustekinumab.
- Pre-specified Areas of Safety Interest.

14.2 Summary of Changes Between the Previous Version and the Current Version of the SAP

Clarified that a post baseline value must be more extreme than the baseline value to be considered a potentially clinically important finding for laboratory parameters and vital signs variables.

Updated Table 19 based on Integrated Summary of Safety SAP.

Updated the end day for inclusion concomitant medications, based on Integrated Summary of Safety SAP.

15.0 Appendix

None.

16.0 References

None.

17.0 List of Tables, Figures, and Data Listings to Be Programmed

To be provided in a separate document.

Document Approval

Study 13113 - Statistical Analysis Plan Version 2 - 12Sep2017 (E3 16.1.9)

Version: 1.0 **Date:** 13-Sep-2017 11:18:08 AM **Company ID:** 09132017-00F9F683A6A755-00001-en

Signed by:	Date:	Meaning Of Signature:
[redacted]	12-Sep-2017 04:48:15 PM	Approver
	12-Sep-2017 04:50:13 PM	Author
	12-Sep-2017 06:21:46 PM	Approver
	13-Sep-2017 11:18:07 AM	Approver