Statistical Analysis Plan for Study M15-994

A Phase 3b Multicenter, Randomized, Double-Blind, Parallel-Group, Placebo-Controlled Study Evaluating Safety and Efficacy of Risankizumab Compared to Placebo in Adult Subjects with Moderate to Severe Plaque Psoriasis with Palmoplantar (Non-Pustular) Involvement (PPPsO)

Date: 16 May 2023

Version 3.0

Table of Contents

List of Tables

List of Figures

List of Appendices

1.0 Introduction

This Statistical Analysis Plan (SAP) describes the statistical analyses for risankizumab Study M15-994, A Phase 3b Multicenter, Randomized, Double-Blind, Parallel-Group, Placebo-Controlled Study Evaluating Safety and Efficacy of Risankizumab Compared to Placebo in Adult Subjects with Moderate to Severe Plaque Psoriasis with Palmoplantar (Non-Pustular) Involvement (PPPsO).

Study M15-994 assesses the safety and efficacy of risankizumab (150 mg) versus placebo for the treatment of signs and symptoms of moderate to severe plaque psoriasis in patients with (non-pustular) palmoplantar involvement.

This SAP provides summaries of the planned statistical analyses for safety and efficacy endpoints, and overall type-I error control strategies.

The SAP will not be updated in case of administrative changes or amendments to the protocol unless the changes impact the analysis.

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

2.0 Study Design and Objectives

2.1 Objectives, Hypotheses and Estimands

The primary objective of this study is to assess the safety and efficacy of risankizumab (150 mg) versus placebo for the treatment of signs and symptoms of moderate to severe plaque psoriasis in patients with (non-pustular) palmoplantar involvement.

Primary Efficacy

The primary efficacy objective is to demonstrate a higher rate of Palmoplantar Investigator's Global Assessment (ppIGA) of "clear" or "almost clear" (0 or 1) with at least a 2-point reduction from Baseline at Week 16 of treatment with risankizumab

(150 mg) compared to placebo in the Intent-to-Treat (ITT) Population, which consists of all randomized subjects (Section 5.0).

The hypothesis corresponding to the primary endpoint is:

- The proportion of subjects achieving ppIGA of 0 or 1 with at least a 2-point reduction from Baseline with risankizumab is greater than that with placebo at Week 16.

The estimand corresponding to the primary endpoint is defined using composite variable strategy:

- Difference in the proportion of subjects achieving ppIGA of 0 or 1 with at least a 2-point reduction from Baseline at Week 16 without premature discontinuation of study drug due to lack of efficacy prior to Week 16, in the risankizumab group in comparison with the placebo group in the ITT population.

Secondary Efficacy

The secondary efficacy objectives are based on the ranked secondary endpoints as defined in Section 3.2. The hypotheses and estimands corresponding to these ranked secondary endpoints are summarized in Table 1.

Table 1. Hypotheses and Estimands Corresponding to the Ranked Secondary Endpoints

Hypothesis	Estimand
1. The proportion of subjects achieving PPASI 75 with risankizumab is greater than that with placebo at Week 16.	Difference in the proportion of subjects achieving PPASI 75 at Week 16 without premature discontinuation of study drug due to lack of efficacy prior to Week 16 in the risankizumab group in comparison with the placebo group in the ITT population.
2. The proportion of subjects achieving PPASI 90 with risankizumab is greater than that with placebo at Week 16.	Difference in the proportion of subjects achieving PPASI 90 at Week 16 without premature discontinuation of study drug due to lack of efficacy prior to Week 16 in the risankizumab group in comparison with the placebo group in the ITT population.
3. The proportion of subjects achieving sPGA of 0 or 1 with at least a 2-point reduction from Baseline with risankizumab is greater than that with placebo at Week 16.	Difference in proportion of subjects achieving sPGA of 0 or 1 with at least a 2-point reduction from Baseline at Week 16 without premature discontinuation of study drug due to lack of efficacy prior to Week 16 in the risankizumab group in comparison with the placebo group in the ITT population.
4. The proportion of subjects achieving PPASI 100 with risankizumab is greater than that with placebo at Week 16.	Difference in the proportion of subjects achieving PPASI 100 at Week 16 without premature discontinuation of study drug due to lack of efficacy prior to Week 16 in the risankizumab group in comparison with the placebo group in the ITT population.

ITT = intent-to-treat; PPASI = Palmoplantar Psoriasis Area and Severity Index; sPGA = static Physician's Global Assessment

2.2 Study Design Overview

This is a Phase 3b, multicenter, randomized, double-blind, parallel group, placebo-controlled study to evaluate the safety and efficacy of risankizumab (150 mg) compared to placebo in adult subjects with moderate to severe plaque psoriasis with palmoplantar involvement.

The study is comprised of a Screening period of up to 35 days, a 52-week treatment period, and a follow up phone call for safety.

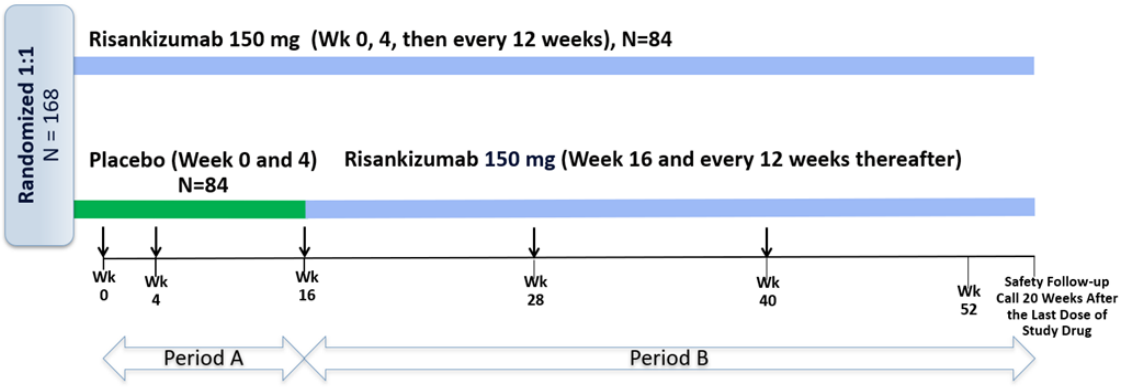
The 52-week treatment period includes Period A and Period B.

- Period A, Double-Blind Period (Baseline to Week 16): Eligible subjects will be centrally randomized at the Baseline visit in a 1:1 ratio to receive either risankizumab 150 mg as a single subcutaneous (SC) injection, or matching placebo. Study drug administration will occur at Baseline and Week 4. The final efficacy evaluation of Period A will be at Week 16.
- Period B, Open-Label Period (Week 16 to Week 52): Starting at Week 16, all subjects will receive open label risankizumab 150 mg once every 12 weeks at Weeks 16, 28, and 40. The final efficacy evaluation of Period B will take place at Week 52.

A follow-up phone call for safety will be conducted approximately 20 weeks after administration of the last dose of risankizumab (e.g., Week 60).

The schematic of the study is shown in Figure 1.

Figure 1. Study Schematic



Solid arrows indicate administration of study drug at study visits.
Wk = week.

The Primary Analysis will be performed after the last ongoing subject completing the Week 16 visit or permanently discontinuing the placebo-controlled period, and the data up to Week 16 cutoff date have been cleaned.

2.3 Treatment Assignment and Blinding

All subjects will be assigned a unique identification number by the interactive response technology (IRT) system at the screening visit. For subjects who rescreen, the screening number assigned by the IRT at the initial screening visit should be used. The IRT will assign a randomization number that will encode the subject's treatment group assignment according to the randomization schedule generated by Data and Statistical Sciences (DSS) at AbbVie.

Subjects will be centrally randomized at the Baseline visit in a 1:1 ratio to receive either SC risankizumab 150 mg or matching placebo during Period A, and all subjects will receive open label risankizumab 150 mg once every 12 weeks during Period B. Randomization will be stratified by Baseline ppIGA ("moderate" [3] versus "severe" [4]) and Baseline body surface area (BSA) affected (< 10% versus ≥ 10%).

All AbbVie personnel with direct oversight of the conduct and management of the study (with the exception of AbbVie Drug Supply Management Team) will remain blinded until the Primary Analysis at Week 16 is available. The investigator, study site personnel, and the subject will remain blinded to each subject's treatment throughout the study. To maintain the blind, the risankizumab pre-filled syringe (PFS) and placebo PFS provided for the study will be identical in appearance.

In the event of a medical emergency in which the investigator believes that knowledge of study drug treatment is required, reasonable efforts must be made to contact the AbbVie emergency contact prior to breaking the blind, as long as it does not compromise subject safety. However, if an urgent therapeutic intervention is necessary which warrants breaking the blind prior to contacting the AbbVie emergency contact, the investigator can directly access the IRT system to break the blind without AbbVie agreement. The date

and reason that the blind was broken must be recorded in the source documentation and electronic case report form (eCRF), as applicable.

2.4 Sample Size Determination

The ppIGA is an endpoint that has been used in another clinical trial in the palmoplantar psoriasis population and which is based on the modified Investigator's Global Assessment version of 2011.[1,2] Given the lack of historical ppIGA data from risankizumab trials, the assumption of response rates in the current study are based on the relevant endpoints of PPASI 75 and PPASI 90 from risankizumab PsO Studies M15-992, M15-995, and M16-008, among subjects who had Baseline PPASI of at least 8. Among these subjects, approximately 80% and 75% of subjects from the risankizumab (RZB) group achieved PPASI 75 and PPASI 90, respectively; compared to 50% and 40% in the placebo (PBO) group.

Assuming that 75% of subjects from the RZB group and 45% of subjects from the PBO group would achieve the primary endpoint of ppIGA of 0 or 1 with at least a 2-point reduction from Baseline at Week 16, the sample size of 168 subjects (84 subjects per group) will provide more than 90% power to detect the treatment difference between RZB and PBO, under a 2-sided significance level of 0.05.

3.0 Endpoints

3.1 Primary Endpoint

The primary endpoint is the achievement of ppIGA of 0 or 1 with at least a 2-point reduction from Baseline at Week 16.

3.2 Secondary Endpoint

The following ranked secondary endpoints will be tested in a hierarchical order, only if the null hypothesis for the primary endpoint has been rejected:

1. Achievement of PPASI 75 response at Week 16
2. Achievement of PPASI 90 response at Week 16
3. Achievement of sPGA of 0 or 1 with at least a 2-point reduction from Baseline at Week 16
4. Achievement of PPASI 100 response at Week 16

3.3 Additional Endpoints

All primary and ranked secondary endpoints will be analyzed at all other visits collected. In addition, the following endpoints will be analyzed at all visits collected:

- Achievement of ppIGA of "clear" (0)
- Achievement of sPGA of "clear" (0)
- Change from Baseline in PPASI
- Percent change from Baseline in PPASI
- Change from Baseline in PASI
- Percent change from Baseline in PASI among subjects with Baseline PASI ≥ 12
- Change and percentage change from Baseline in Patient Global Assessment of Skin Pain (PGA-SP), among subjects with Baseline PGA-SP NRS ≥ 4
- Change from Baseline in Psoriasis Symptom Scale (PSS)
- Achievement of PSS 0 or 1
- Change from Baseline in Dermatology Life Quality Index (DLQI)
- Achievement of DLQI 0 or 1
- Achievement of DLQI improvement (reduction) of ≥ 4 points, among subjects with Baseline DLQI ≥ 4

In addition, the achievement of ppIGA of "clear" (0), ppIGA of "clear" or "almost clear" (0 or 1) with at least a 2-point reduction from Baseline, and PPASI 75/90/100 response will also be summarized at later visits in Period B, among subjects who are randomized to risankizumab in Period A and have achieved the corresponding response status at the entry of Period B.

3.4 Safety Endpoints

The following safety evaluations will be performed throughout the study as measures of safety and tolerability:

- Adverse event (AE) monitoring
- Vital sign measurements
- Physical examinations
- Clinical laboratory testing

3.5 Pharmacokinetic Endpoints

No PK data will be collected for the purpose of PK analysis in present study.

4.0 Analysis Populations

The Intent-to-Treat (ITT) Population includes all randomized subjects. The ITT Population will be used for all efficacy analyses. Subjects who are randomized to placebo in Period A and do not continue into Period B will be excluded from the analysis in Period B.

The following populations will be used for the safety analysis:

- The Safety Population in Period A (Safety_A) is defined as all subjects who are randomized and received at least 1 dose of study drug in Period A.
- The Safety Population in Period B (Safety_B) is defined as all subjects who received at least 1 dose of study drug in Period B.

- The All Risankizumab Treated (ALL_RZB) Population is defined as subjects who received at least 1 study drug of risankizumab. This population will be used to provide a comprehensive summary of key safety variables.

5.0 Subject Disposition

A summary of subject accountability will be provided where the number of subjects in each of the following categories will be summarized by treatment groups:

- Subjects randomized in the study
- Subjects who received at least one dose of study drug in each Study Period
- Subjects who completed each Study Period
- Subjects who discontinued study drug in each Study Period (all reasons and primary reason)

For end of study participation, the number and percentage of subjects who did not complete each Study Period with associated reasons will be summarized overall and by treatment group.

6.0 Study Drug Duration and Compliance

For the Safety Populations in Period A and in Period B, the duration of treatment will be summarized by treatment groups. The duration of treatment for each period is defined as follows:

Period A: the minimum of (the last dose date in Period A + 84 days, the end of study date + 1 day, and the first dose date in Period B if not missing [and cutoff date during the Primary Analysis + 1 day]) minus the first dose date in Period A.

Period B: the minimum of (the last dose date in Period B + 84 days, and the end of study date + 1 day [and cutoff date during the Primary Analysis + 1 day]) minus the first dose date in Period B.

The duration of treatment will also be summarized among the ALL_RZB Population, which is defined as: the minimum of (the last dose date of risankizumab + 84 days, and the end of study date + 1 day [and cutoff date during the Primary Analysis + 1 day]) minus the first dose date of risankizumab.

Duration of treatment will be summarized by descriptive statistics of mean, standard deviation, median, minimum and maximum, among the number of subjects treated in each period.

Treatment compliance will be summarized by treatment groups among the Safety Populations in Period A and in Period B. The compliance will be summarized by the percentage of planned injections which are administered at each study drug administration visit. The cumulative summary of compliance in each period will also be provided. When computing the compliance at each study drug administration visit for each treatment group, the denominator will include all subjects in this treatment group who received at least one dose of study drug in the Study Period and have not prematurely discontinued the study drug prior to that scheduled visit.

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

7.1 Demographics and Baseline Characteristics

Continuous demographic variables include age, weight, height, and body mass index (BMI). Categorical demographic variables include sex, ethnicity, race, age (< 40 or ≥ 40 years), weight (≤ 100 or > 100 kg), BMI (< 25, ≥ 25 – < 30, ≥ 30 kg/m^2), tobacco user (current, former, never, unknown), and alcohol user (current, former, never, unknown).

Baseline disease characteristics include history of psoriatic arthritis (yes, no), Baseline ppIGA categories, Baseline sPGA categories, Baseline BSA categories (< 10%, ≥ 10%), Baseline PASI, Baseline PPASI, Baseline BSA, Baseline PSS, Baseline PGA-SP, and Baseline DLQI.

Demographics and baseline disease characteristics will be summarized among the ITT Population, overall and by treatment groups. Categorical variables will be summarized with the number and percentage of subjects; percentages will be calculated based on the number of non-missing observations. Continuous variables will be summarized with descriptive statistics (number of non-missing observations, mean, standard deviation, median, minimum and maximum).

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 groups. The system organ class (SOC) will be presented in the 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 corresponding category (SOC or preferred term).

Medical history will be summarized among the ITT Population.

7.3 Prior and Concomitant Medications

Prior and concomitant medications will be summarized by generic names. A prior medication is defined as any medication taken prior to the first dose date of study drug. A concomitant medication is defined as any medication that is started prior to the first dose date of study drug and continued to be taken on or after the first dose date of study drug, or any medication that is started on or after the first dose date of study drug, but not after the last dose date of study drug plus 140 days. The number and percentage of subjects taking medications will be summarized by generic drug names based on the World Health Organization (WHO) Drug Dictionary for both prior and concomitant medications.

In addition, subjects' prior biologic therapy for psoriasis will also be summarized by the reason for discontinuation.

Prior and concomitant medications will be summarized among the ITT Population.

8.0 Handling of Potential Intercurrent Events for the Primary Endpoint and Key Secondary Endpoints

The primary efficacy endpoint (defined in Section 3.1) and four ranked secondary endpoints (defined in Section 3.2), will be analyzed in the ITT Population and the following method will be used to address the potential intercurrent event:

- Subjects who discontinue study drug due to lack of efficacy prior to Week 16 will be considered as non-responders at Week 16.

9.0 Efficacy Analyses

9.1 General Considerations

All efficacy analyses will be conducted among the ITT Population. All statistical tests will be performed at a 2-sided alpha level of 0.05.

The Primary Analysis will be conducted after all continuing subjects complete Week 16 and all data pertaining to Period A are cleaned. This will be the only and final analysis for efficacy in Period A. Study sites and subjects will remain blinded to their initial treatment assignment for the duration of the entire study.

For categorical endpoints, comparisons will be made between risankizumab and placebo using the Cochran-Mantel-Haenszel (CMH) test, adjusting for the actual values of stratification factors (Baseline ppIGA and Baseline BSA categories). In case of any stratum with zero subject in either treatment group, no stratification factor will be controlled. Non-responder imputation (NRI) incorporating multiple imputation (MI) to handle missing data due to COVID-19 (NRI-C) will be the primary approach to handle missing values.

For continuous endpoints, comparisons will be made between risankizumab and placebo based on the fixed term of treatment from a Mixed-effect Model Repeat Measurement (MMRM) model including the Baseline value and observed measurements at all post-baseline visits, adjusting for the fixed effects of treatment, actual values of stratification factors, visit and treatment-by-visit interaction as covariates. The MMRM will be the primary approach to handle missing values.

"Baseline" refers to the last non-missing observation on or before the date of the first administration of study drug or the date of randomization if no study drug is administered. For variables where the assessment time is collected on the ePRO device, the baseline measurement must be prior to the time of the first administration of study drug.

9.2 Handling of Missing Data

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.

Sensitivity analyses will be performed to assess the impact of missing data and the robustness of the conclusion.

Handling of missing data for the efficacy analyses is described below.

9.2.1 Categorical Endpoints

Non-Responder Imputation incorporating multiple imputation to handle missing data due to COVID-19 (NRI-C) will be the primary approach for handling missing data in the analysis of categorical endpoints. The NRI-C will categorize any subject who does not have an evaluation during a pre-specified visit window as a non-responder for the visit. The only exceptions are: 1) when the subject is a responder both before and after the visit window in the particular Study Period, the subject will be categorized as a responder for the visit; 2) missing data due to COVID-19 infection or logistical restriction will be handled by MI. Subjects who discontinue study drug due to lack of efficacy will be counted as non-responders at later visits. Subjects whose change/percent change from Baseline cannot be calculated because of a missing Baseline will be considered as a non-responder at all post-baseline visits, unless the post-baseline value is zero. More details are provided in Appendix E.

Of note, during the Primary Analysis upon completion of Week 16, the NRI-C analysis will only be performed at all visits up to Week 16.

Multiple Imputation (MI) will be used as a sensitivity analysis for the primary endpoint at Week 16. PROC MI with the Markov Chain Monte Carlo (MCMC) statement will be first applied to generate 30 augmented datasets with monotonic missing pattern. The random seed for the MCMC will be the SAS numerical value of the first subject randomization date. PROC MI will then be used to impute 30 complete datasets using the regression method. The random seed for this imputation step will be the SAS numerical value of the last subject randomization date. The variables to be included in the imputation model are: treatment group, Baseline BSA, Baseline ppIGA category, and ppIGA measurements at each visit up to the end of the double blind Period A. The

imputed post-baseline measurements will be rounded to the same precision as the observed data to determine the responder status. Using the Cochran-Mantel-Haenszel (CMH) model adjusted by actual stratification factors (Baseline BSA category and Baseline ppIGA category), the imputed endpoints will be analyzed using each of the 30 datasets. SAS PROC MIANALYZE will be used to generate the final inferences of the risk difference between risankizumab and placebo. Note that measurements will be considered as missing after discontinue study drug due to lack of efficacy before MI. Regardless of MI imputed values, subjects with a missing ppIGA assessment after discontinuation from study due to lack of efficacy will be counted as non-responders for these visits.

9.2.2 Continuous Endpoints

Mixed-Effect Model Repeat Measurement (MMRM) will be the primary approach to handle missing data for continuous endpoints. The MMRM will be conducted using mixed model including observed measurements at all visits, using all available data even if a subject has missing data at some post-baseline visits. The mixed model includes the fixed effects of categorical variables of treatment, actual stratification factors (Baseline BSA category and Baseline ppIGA category), visit and treatment-by-visit interaction at all post-baseline visits in each Study Period, and the continuous variable of Baseline measurement as covariates. Subjects' observations after discontinued study drug due to lack of efficacy will be excluded from the model. An unstructured variance covariance matrix (UN) will be used. If the model cannot converge, an appropriate covariance structure matrix (e.g., autoregressive (1) or compound symmetry) will be used. The parameter estimations are based on the method of restrictive maximum likelihood (REML). The fixed effects will be used to report model-based means at corresponding visits.

Of note, during the Primary Analysis upon completion of Week 16, the MMRM analysis will only be performed at all visits up to Week 16.

Long-term efficacy in the Period 2 will be summarized using the observed case approach.

Observed Case (OC): The OC analysis will be used for the summaries of long-term efficacy in Period 2, which will use observed data.

9.3 Primary Efficacy Endpoint and Analyses

9.3.1 Primary Efficacy Endpoint

The primary efficacy endpoint is the achievement of ppIGA of 0 or 1 with at least a 2-point reduction from Baseline at Week 16 as defined in Section 3.1.

9.3.2 Main Analysis of Primary Efficacy Endpoint

The following null hypothesis corresponding to the primary endpoint will be tested under a two-sided significance level of 0.05:

- There is no difference between risankizumab and placebo, with respect to the proportion of subjects achieving ppIGA of 0 or 1 with at least a 2-point reduction from Baseline at Week 16.

The primary endpoint will be analyzed among the ITT Population, using the Cochran-Mantel-Haenszel (CMH) test adjusting for the actual values of stratification factors.

The attributes of the estimand corresponding to the primary efficacy endpoint are summarized in Table 2.

Table 2. Summary of the Estimand Attributes of the Primary Efficacy Endpoint

Estimand Label	Attributes of the Estimand				
	Treatment	Endpoint	Population	Handling of Intercurrent Events	Statistical Summary
ppIGA 0 or 1 at Week 16	Risankizumab 150 mg / Placebo	Achievement of ppIGA 0 or 1 with a least 2-point reduction from Baseline at Week 16	ITT	Subjects who discontinued study drug due to lack of efficacy prior to Week 16 will be considered as non-responders.	Difference in proportion of subjects achieving ppIGA 0 or 1 with a least 2-point reduction from Baseline at Week 16 between risankizumab 150 mg and placebo

Comparison will be made between risankizumab and placebo using the CMH test, adjusting for the actual values of stratification factors (Baseline ppIGA and Baseline BSA categories). NRI-C will be the primary approach to handle missing data.

9.3.3 Sensitivity and Supplementary Analyses of the Primary Efficacy Endpoint

A sensitivity analysis using MI to handle missing data will be performed on the primary endpoint, as defined in Section 9.2.1. Treatment, Baseline BSA, and the observed ppIGA assessments at all Baseline and post-baseline visits will be included in the imputation model.

9.4 Secondary Efficacy Endpoints and Analyses

9.4.1 Key Secondary Efficacy Endpoints

The ranked secondary endpoints are as defined in Section 3.2.

9.4.2 Main Analyses of Key Secondary Efficacy Endpoints

The following null hypotheses pertaining to the ranked secondary efficacy endpoints will be tested between the risankizumab and placebo groups among the ITT Population in a hierarchical order, only if the null hypothesis for the primary endpoint has been rejected:

1. There is no difference between risankizumab and placebo, with respect to the proportion of subjects achieving PPASI 75 at Week 16
2. There is no difference between risankizumab and placebo, with respect to the proportion of subjects achieving PPASI 90 at Week 16
3. There is no difference between risankizumab and placebo, with respect to the proportion of subjects achieving sPGA of "clear" or "almost clear" (0 or 1) with at least a 2-point reduction from Baseline at Week 16
4. There is no difference between risankizumab and placebo, with respect to the proportion of subjects achieving PPASI 100 at Week 16

Ranked secondary endpoints will be analyzed, using the CMH test adjusting for the actual values of stratification factors. NRI-C will be the primary approach to handle missing data.

The attributes of the estimands corresponding to the ranked secondary efficacy endpoints are summarized in Table 3.

Table 3. Summary of the Estimand Attributes of the Ranked Secondary Efficacy Endpoint

	Attributes of the Estimand				
Estimand Label	**Treatment**	**Endpoint**	**Population**	**Handling of Intercurrent Events**	**Statistical Summary**
PPASI 75 at Week 16	Risankizumab 150 mg / Placebo	Achievement of PPASI 75 at Week 16	ITT	Subjects who discontinued study drug due to lack of efficacy prior to Week 16 will be considered as non-responders.	Difference in proportion of subjects achieving PPASI 75 at Week 16 between risankizumab 150 mg and placebo
PPASI 90 at Week 16	Risankizumab 150 mg / Placebo	Achievement of PPASI 90 at Week 16	ITT	Subjects who discontinued study drug due to lack of efficacy prior to Week 16 will be considered as non-responders.	Difference in proportion of subjects achieving PPASI 90 at Week 16 between risankizumab 150 mg and placebo
sPGA 0 or 1 at Week 16	Risankizumab 150 mg / Placebo	Achievement of sPGA 0 or 1 with a least 2-point reduction from Baseline at Week 16	ITT	Subjects who discontinued study drug due to lack of efficacy prior to Week 16 will be considered as non-responders.	Difference in proportion of subjects achieving sPGA 0 or 1 with a least 2-point reduction from Baseline at Week 16 between risankizumab 150 mg and placebo
PPASI 100 at Week 16	Risankizumab 150 mg / Placebo	Achievement of PPASI 100 at Week 16	ITT	Subjects who discontinued study drug due to lack of efficacy prior to Week 16 will be considered as non-responders.	Difference in proportion of subjects achieving PPASI 100 at Week 16 between risankizumab 150 mg and placebo

9.4.3 Sensitivity and Supplementary Analyses for Key Secondary Efficacy Endpoints

Not applicable.

9.5 Additional Efficacy Analyses

Additional efficacy endpoints will be compared between the risankizumab and placebo treatment groups among the ITT Population at each visit in Period A.

Categorical endpoints will be analyzed, using the CMH test adjusting for the actual values of stratification factors. NRI-C will be the primary approach to handle missing data.

Continuous endpoints will be analyzed based on the fixed term of treatment from an MMRM model including the Baseline value and observed measurements at all post-baseline visits, adjusting for the fixed effects of treatment, actual values of stratification factors, visit and treatment-by-visit interaction as covariates.

Summary statistics will be provided for the additional efficacy endpoints at each visit in Period B by treatment groups. Categorical endpoints will be summarized by the number and proportion of subjects who achieved the endpoint, as well as the 95% confidence interval of that proportion. Missing data will be handled by NRI-C. Continuous endpoints will be summarized based on OC by descriptive statistics including the mean, standard error, and the 95% confidence interval of the mean.

9.6 Efficacy Subgroup Analyses

To evaluate the consistency of the efficacy over demographic and other baseline disease characteristics, the primary efficacy endpoint will be analyzed in selected subgroups as follows:

- Age group (< 40 years, ≥ 40)
- Sex (male, female)
- Race (white, non-white)

- Smoking (current, ex- or never)
- Body mass index (BMI) (normal: < 25, overweight: ≥ 25 to < 30, obese: ≥ 30)
- Baseline ppIGA ("moderate" [3], "severe" [4])
- Baseline BSA ($< 10\%$ versus $\geq 10\%$);
- Psoriatic arthritis (yes, no)
- Body weight (≤ 100 kg, > 100 kg)

Of note, if the BMI ≥ 30 subgroup have fewer than 10% of the total subjects, it will be combined with its adjacent subgroup.

10.0 Safety Analyses

10.1 General Considerations

Safety analyses will include adverse events, laboratory, and vital sign measurements.

Safety data will be summarized among the Safety_A and Safety_B populations by treatment groups. For the safety analysis, subjects are analyzed based on the actual treatment group, determined by the first dose of study drug that the subject received.

The overview of TEAEs, areas of safety interest (ASIs), and potentially clinically important (PCI) findings in laboratory variables and vital sign variables will be summarized among the ALL_RZB Population.

Missing safety data will not be imputed.

10.2 Adverse Events

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. When summarizing the number and percentages of subjects, subjects with

multiple occurrences of the same adverse event will be counted once, and only the maximum severity level will be presented in the severity summaries, and the worst/highest relationship level in the relationship summaries.

10.2.1 Treatment-Emergent Adverse Events

Treatment-emergent adverse events (TEAEs) for each safety period are defined as follows:

Period A: A TEAE in Period A is defined as any event with an onset date on or after the first dose date of study drug in Period A and within 140 days after the last dose date of study drug in Period A, as long as it does not exceed the first dose date in Period B.

Period B: A TEAE in Period B is defined as any event with an onset date on or after the first dose date of study drug in Period B and within the minimum of (140 days after the last dose date of study drug in Period B, and the end of study date).

The All-Risankizumab Treated Period: A TEAE is defined as any event with an onset date on or after the first dose date of risankizumab and within the minimum of (140 days after the last dose date of risankizumab, and the end of study date).

10.2.2 Adverse Event Overview

An overview of TEAEs by treatment groups in each period will be presented consisting of the number and percentage of subjects experiencing at least one event for each of the following TEAE categories:

- Any treatment-emergent AE
- Any treatment-emergent AE related to study drug according to the investigator
- Any severe treatment-emergent AE
- Any serious treatment-emergent AE
- Any treatment-emergent AE leading to discontinuation of study drug
- Any treatment-emergent AE under the Areas of Safety Interest, as defined in Appendix B

- Any treatment-emergent AE leading to death
- Any COVID-19 related TEAEs

All deaths will also be summarized:

- COVID-19 related deaths
- Deaths occurring ≤ 140 days after last dose of study drug
- Deaths occurring > 140 days after last dose of study drug.

10.2.3 Treatment-Emergent Adverse Events by SOC and/or PT

Treatment-emergent adverse events will be summarized by SOC and PT; by maximum relationship to study drug as assessed by the investigator (e.g., reasonable possibility or no reasonable possibility) and by SOC and PT; by maximum severity and by SOC and PT; and by SOC and PT listing associated subject number. When summarizing the number and percentages of subjects, subjects with multiple occurrences of the same adverse event will be counted once, and only the maximum severity level will be presented in the severity summaries, and the worst/highest relationship level in the relationship summaries.

In addition, treatment-emergent adverse events will be summarized by PT and sorted by decreasing frequency for the active group.

10.2.4 Treatment-Emergent Adverse Events per Patient-Years of Exposure

Exposure-adjusted TEAEs per 100 patient-years will be provided, where TEAEs per 100 patient-years of exposure are defined as the number of TEAEs 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 TEAEs (i.e., a preferred term will not be counted twice on

the same day for the same subject). The exposure-adjusted TEAE rate per 100 patient-years is calculated as:

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

where total patient years in each period are defined below.

Total patient years in Period A: Sum of study drug exposure in Period A, defined as the minimum of (the last dose date in Period A + 140 days, the end of study date + 1 day, and the first dose date in Period B if not missing [and the cutoff date during the Primary Analysis + 1 day]) minus the first dose date in Period A, normalized by 365.25 and rounded to one decimal place.

Total patient years in Period B: Sum of study drug exposure in Period B, defined as the minimum of (the last dose date in Period B + 140 days, and the end of study date + 1 day [and the cutoff date during the Primary Analysis + 1 day]) minus the first dose date in Period B, normalized by 365.25 and rounded to one decimal place.

Total patient years in the All-Risankizumab Treated Period: Sum of study drug of risankizumab exposure, defined as the minimum of (the last risankizumab dose date + 140 days, and the end of study date + 1 day [and the cutoff date during the Primary Analysis + 1 day]) minus the first risankizumab dose date, normalized by 365.25 and rounded to one decimal place.

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

Treatment-emergent SAEs (including deaths) and AEs leading to study drug discontinuation will be summarized by SOC and PT and in listing format.

A listing of pre-treatment SAEs with onset dates prior to the first dose of study drug will be provided.

10.2.6 Area of Safety Interest

Detailed information about the search criteria for areas of safety interest (ASIs) are provided in Appendix B.

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.

Tabular listings of selected area of safety interest will be provided.

10.3 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 derivation where SAE-triggered 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.

Mean change from Baseline to each applicable post-baseline visit will be summarized for selected laboratory variables. The following descriptive statistics will be presented by treatment groups: number of observations, Baseline mean, visit mean, mean change from Baseline and its standard error, and the 95% confidence interval of the mean change from Baseline.

In addition, laboratory parameters will be tabulated using shift tables from Baseline to minimum and maximum values in each period (Period A and Period B), categorized by the toxicity grade according to NCI CTCAE Version 4.03[3] of the laboratory used for each sample. A similar shift table will also be provided to summarize shifts from Baseline to the final post-baseline value in each period.

Laboratory abnormalities meeting CTC criteria grade 3 and 4 at least once in each period will be summarized.

Laboratory abnormalities will be evaluated based on Potentially Clinically Important (PCI) criteria (Appendix C). For each laboratory PCI criterion, the number and percentage of subjects who have a laboratory value meeting the criteria at least once in each period will be summarized. Listings will be provided to summarize subject-level laboratory data for subjects meeting PCI criteria.

In addition, the frequencies and percentages of subjects with post baseline liver specific function test values in ALT/SGPT, AST/SGOT, alkaline phosphatase, and total bilirubin that meet the following criteria will be summarized by treatment groups:

- ALT > 3.0 × ULN
- ALT > 5.0 × ULN
- ALT > 10.0 × ULN
- ALT > 20.0 × ULN
- AST > 3.0 × ULN
- AST > 5.0 × ULN
- AST > 10.0 × ULN
- AST > 20.0 × ULN
- Alkaline phosphatase > 1.5 × ULN
- Total bilirubin > 1.5 × ULN
- Total bilirubin > 2.0 × ULN
- ALT and/or AST > 3.0 × ULN and Total bilirubin > 1.5 × ULN
- ALT and/or AST > 3.0 × ULN and Total bilirubin > 2.0 × ULN
- ALT > 3 × ULN and Total bilirubin > 1.5 × ULN
- ALT > 3 × ULN and Total bilirubin > 2.0 × ULN

A listing of potentially clinically important (PCI) liver function laboratory values will include all subjects who met any of the following four criteria:

- ALT > 3.0 × ULN, or
- AST > 3.0 × ULN, or

- ALP > 1.5 × ULN, or
- Total bilirubin > 1.5 × ULN.

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

- ALT of > 3.0 × ULN or AST of > 3.0 × ULN,
- Total bilirubin ≥ 2.0 × ULN

10.4 Analysis of Vital Signs

Vital sign measurements of systolic and diastolic blood pressure will be summarized.

Change from baseline to each applicable post-baseline visit will be summarized for each vital sign variable. The following descriptive statistics will be presented by treatment groups: number of observations, Baseline mean, visit mean, mean change from Baseline and its standard error, and the 95% confidence interval of the mean change from Baseline.

Vital sign variables will be evaluated based on potentially clinically important (PCI) criteria (Appendix D). For each vital sign PCI criterion, the number and percentage of subjects who have a vital sign value meeting the criteria at least once in each period will be summarized. Listings will be provided to summarize subject-level vital sign data for subjects meeting PCI criteria.

10.5 Safety Subgroup Analyses

No subgroup for safety analyses.

10.6 Other Safety Analyses

No other safety analyses.

11.0 Other Analyses

No other analyses.

12.0 Interim Analyses

12.1 Data Monitoring Committee

There is no data monitoring committee (DMC) planned in this study.

13.0 Overall Type-I Error Control

The Primary Analysis will be the only and final analysis for the primary and all ranked secondary efficacy endpoints.

Overall type-I error will be controlled by testing the primary efficacy endpoint, followed by the ranked secondary efficacy endpoints, in a hierarchical order as described in Section 9.3 and Section 9.4.

14.0 Version History

Table 4. SAP Version History Summary

Version	Date	Summary
1.0	07 April 2021	Original version
2.0	11 July 2022	Clarified the time point of conducting Primary Analysis in Section 2.2. Clarified in Section 9.5 that categorical endpoints in Period B will be summarized by NRI-C, and continuous endpoints in Period B will be summarized by OC. Updated summary criteria in Section 10.3 based on PSSAP 5.0. Corrected a format issue in Table C-2.
3.0	16 May 2023	Due to MedDRA version update, CMQ codes in Appendix B. Safety Topics of Interest were updated accordingly.

15.0 References

1. Gottlieb A, Sullivan J, van Doorn M, et al. Secukinumab shows significant efficacy in palmoplantar psoriasis: Results from GESTURE, a randomized controlled trial. J Am Acad Dermatol. 2017;76(1):70-80.

2. Langley RGB, Feldman SR, Nyirady J, et al. The 5-point Investigator's Global Assessment (IGA) Scale: A modified tool for evaluating plaque psoriasis severity in clinical trials. J Dermatolog Treat. 2015;26(1):23-31.
3. Common Terminology Criteria for Adverse Events (CTCAE) v4.03 (2010). Available from: https://evs.nci.nih.gov/ftp1/CTCAE/CTCAE_4.03/CTCAE_4.03_2010-06-14_QuickReference_5x7.pdf.

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.
- Subject took prohibited concomitant medication.

Appendix B. Safety Topics of Interest

Area of safety interest (ASI) will be identified using the following search criteria:

Area of Safety Interest	Search Criteria	
MACE	Adjudicated terms will be identified as described in PSSAP Table 4 using CECAT and CETERM from the CE SDTM dataset.	
Extended MACE	Adjudicated terms will be identified as described in PSSAP Table 3 (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 **10000002**)	
Opportunistic Infections excluding tuberculosis and herpes zoster	Opportunistic infection excluding tuberculosis and herpes zoster CMQ (code **10000105**)	
Injection Site Reactions	Narrow	Injection site reaction CMQ (code **10000091**)
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 **10000100**)
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 hypersensitivity reactions	Narrow	Serious AEs in the Hypersensitivity (SMQ 20000214)
Adjudicated Anaphylactic Reaction*	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)

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

Appendix C. Potentially Clinically Important Criteria for Safety Endpoints

The criteria for Potentially Clinically Important (PCI) laboratory findings are described in Table C-1 and Table C-2.

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

		Definition of Potentially Clinically Important: NCI CTCAE (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

Note: A post-baseline grade must be more extreme than the baseline grade to be considered a potentially clinically important finding.

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

		Definition of Potentially Clinically Important: NCI CTCAE (Version 4) Grade 3 or Greater	
Chemistry Variables	**Units**	**Very Low**	**Very High**
Total Bilirubin	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
Triglycerides	mmol/L		> 5.7
Creatinine	mcmol/L		> 3.0 × ULN or > 3.0 × baseline
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

Note: A post-baseline grade must be more extreme than the baseline grade to be considered a potentially clinically important finding.

Appendix D. Criteria for Potentially Clinically Important Vital Sign Values

Vital Signs Variables	Criterion	Definition of Potentially Clinically Important
Systolic Blood Pressure (mmHg)	Low	Value ≤ 90 mmHg and decrease ≥ 20 mmHg from Baseline
	High	Value ≥ 160 mmHg and increase ≥ 20 mmHg from Baseline
Diastolic Blood Pressure (mmHg)	Low	Value ≤ 50 mmHg and decrease ≥ 10 mmHg from Baseline
	High	Value ≥ 100 mmHg and increase ≥ 10 mmHg from Baseline

Appendix E. Non-Responder Imputation Incorporating Multiple Imputation to Handle Missing Data Due to COVID-19 Pandemic for Dichotomized Outcome Variables

1.0 Overview

1.1 Background and Justification for Missing at Random (MAR) Assumption

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. In some cases, sensitivity analyses may be performed to assess the impact of missing data and the robustness of the conclusion.

1.2 FDA Guidance

FDA provided two guidance documents[1,2] in March 2020 and June 2020 on the efficacy collection and possible changes in the statistical analysis plan:

- "With respect to efficacy assessments, FDA recommends consultation with the appropriate review division regarding protocol modifications for the collection of efficacy endpoints, such as use of virtual assessments, delays in assessments, and alternative collection of research-specific specimens, if feasible. For individual instances where efficacy endpoints are not collected, the reasons for failing to obtain the efficacy assessment should be documented (e.g.,

identifying the specific limitation imposed by COVID-19 leading to the inability to perform the protocol-specified assessment)."

- "If changes in the protocol will lead to amending data management and/or statistical analysis plans, the sponsor should consider doing so in consultation with the applicable FDA review division. Prior to locking the database, sponsors should address in the statistical analysis plan how protocol deviations related to COVID-19 will be handled for the prespecified analyses."

1.3 EMA Guidance

EMA provided guidance[3] in March 2020:

- "At this point in time it is not possible to give general applicable advice on how the different aspects related to the pandemic should be handled, as implications on clinical trials are expected to be manifold. Impact on the data collection, analysis and interpretation of results for each trial will need a thorough case-by-case assessment."
- "As a general principle, there are strong scientific reasons to conduct trials as planned and implement changes only when there is a convincing scientific reason that it improves interpretability of results."

1.4 Missing Data Handling for Missing Due to COVID-19 for Dichotomized Variables

In this document, a missing data handling method is proposed to handle missing data due to COVID-19 infection or logistical restrictions related to the COVID-19 pandemic under the general MAR framework. In particular, we explain using multiple imputation (MI) to handle missing data due to COVID-19 in dichotomized variables in conjunction with non-responder imputation (NRI) for missing data due to other reasons.

2.0 Non-responder Imputation Incorporating Multiple Imputation (NRI)

2.1 Overall Description of the Method

For a dichotomized outcome variable with missing data, the NRI will categorize any subject who does not have evaluation during a pre-specified visit window as a non-responder for the visit, with two exceptions:

- If the subject is a responder both before and after the pre-specified visit window in the particular Study Period, the subject will be categorized as a responder for the visit.
- If the reason for missing (e.g., missed visits, incomplete visit, out-of-schedule visits, or discontinuations of study drug) is due to COVID-19, the information will be captured in the database and the subject's response status will be imputed using multiple imputation.

Subjects whose change/percent change from Baseline cannot be calculated because of a missing Baseline will be considered as a non-responder at all post-baseline visits, unless the post-baseline value is zero.

Non-responder imputation incorporating multiple imputation (NRI) for missing due to COVID-19 will be implemented as follows.

2.2 Multiple Imputation (MI) and MAR Assumption

When a dichotomized variable is derived from a continuous scale, for example, PPASI 75 (at least a 75% reduction in PPASI relative to Baseline), the multiple imputation will be applied to the original scale, PPASI (ranges from 0 – 72) assuming multivariate normal distribution. Then the dichotomized variable will be derived from the imputed value.

The MI procedure assumes that the data are missing at random (MAR). That is, for an outcome variable Y, the probability that an observation is missing depends only on the

observed values of other variables, not on the unobserved values of the outcome variable Y. Statistical inference from the MI procedure is valid under the MAR assumption.

2.3 Imputation Algorithm

It is reasonable to assume the missing values of the longitudinal data for an outcome variable (e.g., PPASI, the original scale of PPASI 75, at each post-baseline visit) follows a monotone missing pattern. In practice, the missing data of the outcome variable might have an arbitrary (non-monotone) missing data pattern. An extra step may be added accordingly, to augment data into a monotone missing pattern.

For the outcome variable (e.g., PPASI at each visit), K 'complete' datasets can be generated in two steps: augmentation step and imputation step. K, the number of repetitions, is determined below.

Augmentation Step

For datasets with non-monotone missing data pattern, the augmentation step will first impute enough values to augment the data into a monotonic missing pattern:

Markov Chain Monte Carlo (MCMC) will be applied to augment the data using PROC MI with the MCMC IMPUTE=monotone statement, assuming a multivariate normal distribution. The augmented data will be used in the subsequent imputation step to generate 'complete' datasets. Covariates included in the model are treatment group, Baseline BSA, Baseline ppIGA category, Baseline (if other than Baseline ppIGA), and all post-baseline visits of the outcome variable up to the end of the Study Period. Of note, categorical variables are included using the form of dummy variables.

Repeat the imputation process K=30 times using the procedure described above to form K=30 monotone missing datasets, where K is determined as described in "Repetition of Imputations (K)."

Imputation Step

For missing data with monotone missing patterns, the choice of multiple imputation using a parametric regression model that assumes multivariate normality is appropriate.

The imputation step is described below:

- The imputation model for the missing data is a regression model, which controls for treatment group, Baseline BSA, Baseline ppIGA category, Baseline (if other than Baseline ppIGA), and all post-baseline visits of the outcome variable up to the end of the Study Period. The covariates included in the model and the order of these variables are consistent with the augmentation step.
- For each monotone missing dataset, using SAS PROC MI with MONOTONE REG model statement, the outcome variable at each post-baseline visit with missing values will be imputed sequentially with covariates constructed from their corresponding sets of preceding variables.

A 'complete' dataset with imputed values for the missing data is generated after the augmentation and imputation steps are completed.

Repetition of Imputations (K)

Repetition of imputations, K, must be determined in advance. When estimating the overall variance of multiple imputation, the additional sampling variance is the between-imputation variance divided by K. This value represents the sampling error associated with the overall or average coefficient estimates. It is used as a correction factor for using a specific number of imputations. The more imputations (K) are conducted, the more precise the parameter estimates will be. For example, with a 1% power falloff tolerance in multiple imputation, as compared to an infinite number of imputations, multiple imputation requires 20 repetitions of imputation for 30% missing information and 40 repetitions for 50% missing information (Graham, Olchowski, and Gilreath 2007[4]). In the usual clinical settings expecting less than 30% missing information, K=30 repetitions

are deemed sufficient. When missingness exceeds 30%, depending on the power falloff tolerance level, number of repetitions may need to be increased. Recent research[4] suggested that the number of repetitions (K) should be at least equal to the percentage of missing (White et al., 2011[6]).

2.4 Derivation of Response Status and Non-Responder Imputation

For each 'complete' dataset, the imputed post-baseline values will be rounded to the same precision as the observed data. Response status (e.g., PPASI 75 at each visit) will be determined accordingly.

The imputed response status for missing due to reasons other than COVID-19 will be overridden by non-responder imputation (Section 2.1) to ensure that multiple imputation is only applied to missing due to COVID-19:

- Using NRI approach, all missing due to reasons other than COVID-19 will be categorized as non-responders. In addition, subjects whose change/percent change from Baseline cannot be calculated because of a missing Baseline will be considered as a non-responder at all post-baseline visits, unless the post-baseline value is zero.
- The only exception is that a subject will be categorized as a responder for the visit if the subject is a responder both before and after an SAP-specified visit window in the particular Study Period.

2.5 Analysis

The statistical analysis will use the Cochran-Mantel-Haenszel (CMH) test adjusted by the actual stratification factors (Baseline BSA category and Baseline ppIGA category).

2.5.1 Analysis of Each Dataset

For each of the K 'complete' datasets, the CMH test will be used to estimate the treatment difference versus placebo and the corresponding standard error.

2.5.2 Synthesis of Results for Statistical Inference

The results from the K 'complete' datasets will be synthesized using the SAS procedure PROC MIANALYZE, following Rubin's formula (Rubin, 1987[5]), to derive the MI estimator of the treatment difference for the final inferences.

Rubin's formula

We fit the analysis model to the kth 'complete' dataset, denoting the estimate of the treatment difference q by $\tilde{\theta}_k$ from the kth 'complete' dataset, and denoting the corresponding estimate of the variance as V_k.

The MI estimator of q (point estimator obtained from PROC MIANALIZE), $\tilde{\theta}_{MI}$, is the average of the K individual estimators:

$$\tilde{\theta}_{MI} = \frac{1}{K}\sum_{k=1}^{K}\tilde{\theta}_k.$$

The estimated variance of $\tilde{\theta}_{MI}$, is a combination of the between- and within-imputation variability as follows:

$$V_{MI} = W + (1+\frac{1}{K})B$$

Where $W = \frac{1}{K}\sum_{k=1}^{K}V_k$ is the within-imputation variability and $B = \frac{1}{K-1}\sum_{k=1}^{K}(\tilde{\theta}_k - \tilde{\theta}_{MI})^2$ is the between-imputation variance.

It has been shown[5] that the statistic

$$T = \frac{\tilde{\theta}_{MI} - \theta}{\sqrt{V_{MI}}}$$

has an approximate t_v distribution where v=(K-1)[(1+W/B)]^2. Statistical inference, including hypothesis testing and confidence intervals for the treatment effect, will be based on this T-statistic.

3.0 Sample SAS Code

```
/**********************/
/*IMPUTATION ALGORITHM*/
/**********************/
/*NOTE: THIS APPROACH REQUIRES NO MISSING IN CATEGORICAL COVARIATES AND
REQUIRES AT LEAST ONE OBSERVATION IN BASELIBE OR ONE OF THE POST-
BASELINE VISIT*/

/*PRE-AUGMENTATION - CREATE DUMMY FOR CATEGORICAL VARIABLES*/
/*****************************************/
DATA PPASI_2; SET PPASI;
  /*THE MCMC STATMENT BELOW ASSUMES MULTI-VARIATE NORMAL*/
  IF TRT01PN=1 THEN TRT1=1; ELSE TRT1=0;
  /*BASELINE BSA and BASELINE PPIGA CATEGORY*/
  IF BSAGR1N = 1 THEN REG1 = 1 ;ELSE REG1 = 0;
  IF PPIGAGR1N = 1 THEN REG2 = 1 ;ELSE REG2 = 0;

RUN;

/*AUGMENTATION STEP -- TO HAVE 30 MONOTONE MISSING DATASETS*/
PROC MI DATA= PPASI_2 OUT= PPASI_MONO NIMPUTE=30 SEED= 21423 /*RANDOM
SEED PRE-DEFINED*/
  ROUND=. . . 0.1 0.1 0.1 /*VALUE ROUND TO 1ST DECIMAL*/
  MIN=. . . 0 0 0  /*MINIMUM VALUE OF PPASI IS 0*/
  MAX=. . . 72 72 72   /*MAXIMUM VALUE OF PPASI IS 72*/
MCMC IMPUTE=MONOTONE ;
/*NOTE: CATEGORICAL VARIABLES SUCH AS TRT1 TRT2 ARE DUMMY, CREATED
ABOVE*/
/*NOTE: ALL OTHER NON-DUMMIED COVARIATES MUST BE CONTINUOUS*/
/*SUPPOSE STRATAN (NUMERIC VARIABLE FOR STRATA) HAS ONLY 2 LEVELS, NO
NEED TO CREATE DUMMY*/
VAR TRT1 BSABL REG2 BASE WK4 WK16;
/*CAUTION TO USE THE "BY" STATEMENT IN MCMC: */
/*MVN MODEL IS FITTED WITHIN EACH 'BY' GROUP, INSTEAD OF ACROSS ALL
GROUPS*/
RUN;

/*IMPUTATION STEP - DETERMINE IMPUTATION DISTRIBUTION AND RANDOMLY
IMPUTE MISSING VALUE TO GENERATE 'COMPLETE' DATASETS*/
/************************************************/
```

```
PROC MI DATA= PPASI_MONO OUT= PPASI_FULL NIMPUTE=1 SEED= 21931  /*RANDOM
SEED PRE-DEFINED*/
  ROUND= =. . . 0.1 0.1 0.1 /*VALUE ROUND TO 1ST DECIMAL*/
  MIN=. . .  0 0 0  /*MINIMUM VALUE OF PPASI IS 0*/
  MAX=. . .  72 72 72 /*MAXIMUM VALUE OF PPASI IS 72*/
  MINMAXITER=1000;
  /*CLASS CATEGORICAL VARIABLES TRT1 AND REG2*/
  CLASS TRT1 REG2;
  VAR TRT1 BSABL REG2 BASE WK4 WK16;
  MONOTONE REG (WK4 WK16);  /* IMPUTED SEQUENTIALLY, FROM WK 4 TO 16,
WITH COVARIATES CONSTRUCTED FROM THE CORRESPONDING PRECEDING VARIABLES*/
  BY _IMPUTATION_;                    /*FOR EACH OF THE 30 MONOTONE
MISSING DATASETS, IMPUTE A 'COMPLETE' DATASET*/
RUN;


/*DETERMINE DICHOTOMOUS RESPONSE STATUS, PPASI 75 AT WEEK 16*/
DATA ALL; SET PPASI_FULL;
  IF 0<=WK16<=0.25*BASE THEN PPASI75_16=1;
    ELSE PPASI75_16=0;
RUN;


/*********************************************************************
*/
/*         DATA HANDLING STEPS TO MERGE COVID-19 STATUS OMITTED
*/
/*            PLACE TO ADD DATA HANDLING AND MERGING STEPS
*/
/*********************************************************************
*/

/*FOR MI, SKIP THE FOLLOWING CODE, PROCEED TO THE CODE AFTER ANALYSIS
MODEL *//*OVERRIDE MISSING VALUES NOT DUE TO COVID-19 WITH TRADITIONAL
NRI*/
DATA ALLF; SET ALL;
  /*COVID19_XX='Y' IF MISSING AT WEEK XX IS DUE TO COVID-19; IF NOT,
OVERRIDE WITH TRADITIONAL NRI*/
  /*VARIABLE PPASI75NRI_XX: TRADITIONAL NRI DATA AT WEEK XX, WHICH
COVERS THE SPECIAL HANDLING SUCH AS THE BEFORE-AND-AFTER EXCEPTION IN
THE PARTICULAR STUDY PERIOD*/
  IF COVID19_16 NE 'Y' THEN PPASI75_16= PPASI75NRI_16;
RUN;
PROC SORT DATA=ALLF; BY _IMPUTATION_ SUBJID; RUN;

/****************/
/*ANALYSIS MODEL*/
/****************/
```

```
/*KEY CODE: ANALYZING EACH 'COMPLETE' DATASET*/
/*********************************/
/*INDIVIDUAL-LEVEL DATA --> # OF RESPONDERS & # OF SUBJECTS, TO BE READ-
IN TO PROC STDRATE*/
PROC FREQ DATA=ALL;
  BY _IMPUTATION_;
  TABLES TRT01PN*STRATAN* PPASI75_16/LIST NOCUM NOPRINT OUT=COUNT_TABLE;
  /*WEEK 16 RESULTS AS AN EXAMPLE*/
RUN;
DATA COUNT_TABLE; SET COUNT_TABLE;
  DROP PERCENT;
RUN;
PROC TRANSPOSE DATA=COUNT_TABLE OUT=FREQ_TABLE PREFIX=RESP;
ID PPASI75_16;
BY _IMPUTATION_ TRT01PN STRATAN;
VAR COUNT;
RUN;
DATA FREQ_TABLE1; SET FREQ_TABLE;
  CASE=RESP1;
  SIZE=SUM(RESP0, RESP1);
  KEEP _IMPUTATION_ TRT01PN STRATAN CASE SIZE;
RUN;

/*RE-ORDER TO SET 1 (PLACEBO) AS THE REFERENCE GROUP*/
DATA FREQ_TABLE2; SET FREQ_TABLE1;
  IF TRT01PN=2 THEN TRT01PN=0;
RUN;

/*CALCULATE THE COMMON RISK DIFF FOR EACH COMPLETE DATASET*/
PROC STDRATE DATA=FREQ_TABLE2
  METHOD=MH STAT=RISK EFFECT=DIFF;
  BY _IMPUTATION_;
  POPULATION GROUP=TRT01PN EVENT=CASE TOTAL=SIZE;
  STRATA STRATAN / ORDER=DATA STATS(CL=NONE) EFFECT;
  ODS OUTPUT EFFECT=EFFECT;
RUN;


/*COMBINING RESULTS USING PROC MIANALYZE*/
/*************************************/
PROC MIANALYZE DATA=EFFECT;
  ODS OUTPUT PARAMETERESTIMATES=RISK_DIFF_MH;
  MODELEFFECTS RiskDiff;
  STDERR StdErr;
RUN;
```

4.0 Reference

1. FDA Guidance on Conduct of Clinical Trials of Medical Products during COVID-19 Pandemic - Guidance for Industry, Investigators, and Institutional Review Boards. FDA. 2020.
2. Statistical Considerations for Clinical Trials During the COVID-19 Public Health Emergency - Guidance for Industry. FDA. 2020.
3. Points to consider on implications of Coronavirus disease (COVID-19) on methodological aspects of ongoing clinical trials. EMA. 2020.
4. Graham JW, Olchowski AE, Gilreath TD. How many imputations are really needed? Some practical clarifications of multiple imputation theory. Prev Sci 2007;8(3):206-13.
5. Rubin DB, Schenker N. Interval estimation from multiply-imputed data: a case study using agriculture industry codes. J Am Stat Assoc. 1987;81:366-74.
6. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. Stat Med. 2011;30(4):377-99.