Statistical Analysis Plan for Study M20-326

Plaque Psoriasis: A Phase 4 Multicenter, Randomized, Open-label, Efficacy Assessor-blinded-Study of Risankizumab Compared to Apremilast for the Treatment of Adult Subjects with Moderate Plaque Psoriasis who are Candidates for Systemic Therapy

Date: 26 February 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 M20-326, Plaque Psoriasis: A Phase 4 Multicenter, Randomized, Open-label, Efficacy Assessor-blinded-Study of Risankizumab Compared to Apremilast for the Treatment of Adult Subjects with Moderate Plaque Psoriasis who are Candidates for Systemic Therapy.

Study M20-326 examines the efficacy and safety of risankizumab in adult subjects with moderate plaque psoriasis who are candidates for systemic therapy.

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 objective of the study is to evaluate the efficacy and safety of risankizumab (RZB) versus apremilast (APR) for the treatment of adult subjects with moderate plaque psoriasis (PsO) who are candidates for systemic therapy.

In addition, the study aims to evaluate the efficacy and safety of switching to RZB versus continuing on APR for subjects with moderate plaque PsO who do not achieve PASI 75 (defined as ≥ 75% reduction from Baseline in Psoriasis Area and Severity Index [PASI]) after 16 weeks of treatment with APR.

Primary Efficacy Objective:

The primary efficacy objective in Period A is to demonstrate a higher rate of a) PASI 90 (defined as ≥ 90% reduction from Baseline in PASI) and b) static Physicians Global

Assessment (sPGA) 0 or 1 with at least 2-grade improvement from Baseline after 16 weeks of treatment with RZB when compared to APR based on Intent to Treat (ITT) Population in Period A (ITT_A), which consists of all randomized subjects.

The hypotheses corresponding to the co-primary efficacy objective and endpoints in Period A are:

- The proportion of subjects achieving PASI 90 in the RZB group is greater than that in the APR group at Week 16, among the ITT_A Population.
- The proportion of subjects achieving sPGA 0 or 1 with at least 2-grade improvement from Baseline in the RZB group is greater than that in the APR group at Week 16, among the ITT_A Population.

The estimands corresponding to the co-primary efficacy objectives in Period A are defined as follows:

- Difference in the percentage of subjects achieving PASI 90 at Week 16 without receiving alternative PsO treatment regardless of premature discontinuation of study drug, in the RZB group in comparison to the APR group among the ITT_A population, and
- Difference in the percentage of subjects achieving sPGA of 0 or 1 with at least two grades of reduction from Baseline at Week 16 without receiving alternative PsO treatment regardless of premature discontinuation of study drug, in the RZB group in comparison to the APR group among the ITT_A population.

The primary efficacy objective in Period B is to demonstrate a higher rate of PASI 90 after switching the treatment at Week 16 from APR to RZB for 36 weeks (APR/RZB) when compared to continuing with APR (APR/APR) based on the Intent to Treat Population in Period B for Week 16 APR-PASI75NRs (ITT_B_NR), which consists of subjects randomized to APR at Baseline, failed to achieve PASI 75 at Week 16 and are re-randomized.

The hypothesis corresponding to the primary objective and endpoint in Period B is:

- The proportion of APR-PASI75NR subjects achieving PASI 90 in the APR/RZB group is greater than that in the APR/APR group at Week 52, among the ITT_B_NR Population.

The estimand corresponding to the primary efficacy objective in Period B is defined as follows:

- Difference in the percentage of subjects achieving PASI 90 at Week 52 without satisfying the rescue criteria or receiving the rescue medication, and without receiving alternative PsO treatment regardless of premature discontinuation of study drug, in the APR/RZB group in comparison to the APR/APR group among the ITT_B_NR population

Key Secondary Efficacy Objective:

The key secondary efficacy objective in Period A is to demonstrate higher efficacy of treatment with RZB when compared to APR with respect to the ranked and only secondary endpoint in Period A among the ITT_A Population, as specified in Section 3.2 below.

The hypothesis corresponding to the ranked and only secondary efficacy objective and endpoint in Period A is:

- The proportion of subjects achieving PASI 75 in the RZB group is greater than that in the APR group at Week 16, among the ITT_A Population.

The estimand corresponding to the ranked and only secondary efficacy objective in Period A is defined as follows:

- Difference in the percentage of subjects achieving PASI 75 from Baseline at Week 16 without receiving alternative PsO treatment, regardless of premature discontinuation of study drug, in the RZB group in comparison to the APR group among the ITT_A population.

The key secondary efficacy objectives in Period B are to demonstrate higher efficacy of treatment with APR/RZB when compared to APR/APR among the ITT_B_NR Population, with respect to the key secondary endpoints in a ranked order in Period B, as specified in Section 3.2 below.

The hypotheses corresponding to the ranked secondary efficacy objective and endpoints in Period B are:

- The proportion of APR-PASI75NR subjects achieving PASI 75 in the APR/RZB group is greater than that in the APR/APR group at Week 52, among the ITT_B_NR Population.
- The proportion of APR-PASI75NR subjects achieving sPGA 0 or 1 with at least 2-grade improvement from Baseline in the APR/RZB group is greater than that in the APR/APR group at Week 52, among the ITT_B_NR Population.

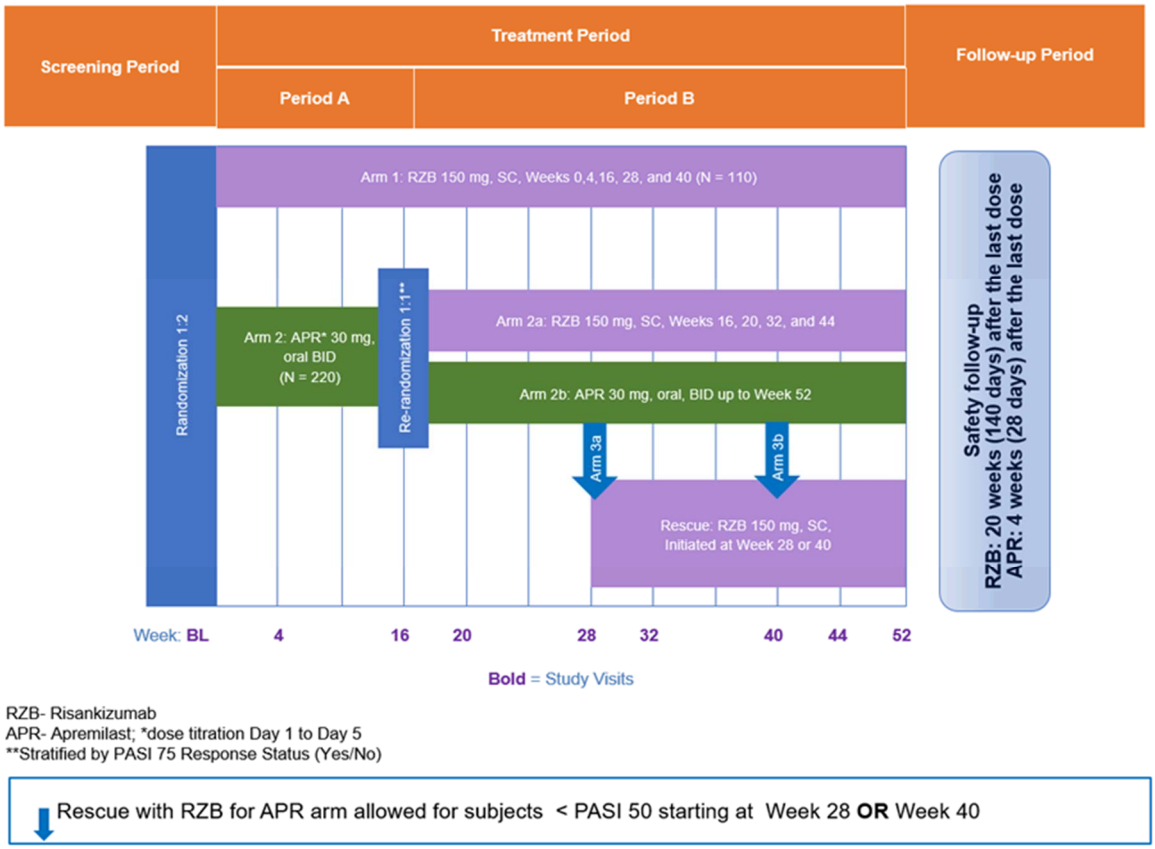
The estimands corresponding to the ranked secondary efficacy objectives in Period B are defined as follows:

- Difference in the percentage of subjects achieving PASI 75 at Week 52 without satisfying the rescue criteria or receiving the rescue medication, and without receiving alternative PsO treatment regardless of premature discontinuation of study drug, in the APR/RZB group in comparison to the APR/APR group among the ITT_B_NR population, and
- Difference in the percentage of subjects achieving sPGA of 0 or 1 with at least two grades of reduction from Baseline at Week 52 without satisfying the rescue criteria or receiving the rescue medication, and without receiving alternative PsO treatment regardless of premature discontinuation of study drug in the APR/RZB group in comparison to the APR/APR group among the ITT_B_NR population.

2.2 Study Design Overview

The schematic of the study is shown in Figure 1.

Figure 1. Study Schematic



BID = twice daily; PASI = Psoriasis Area Severity Index; SC = subcutaneous

2.3 Treatment Assignment and Blinding

All subjects will be assigned a unique identification number by the IRT 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.

In Period A (Baseline to Week 16), subjects will be centrally randomized to risankizumab (Arm 1) or apremilast (Arm 2) in a 1:2 ratio at Baseline. Randomization will be stratified

by Baseline body weight (≤ 100 kg, > 100 kg) and prior exposure to any systemic and/or biologic treatment for Po (0, ≥ 1).

In Period B (Week 16 to Week 52), subjects initially randomized to risankizumab (Arm 1) will continue to receive risankizumab. Subjects initially randomized to apremilast (Arm 2) will be centrally re-randomized at the Week 16 visit in a 1:1 ratio to receive either risankizumab (Arm 2a) or apremilast (Arm 2b). Re-randomization will be stratified by PASI 75 response (responder, non-responder) at Week 16. Rescue risankizumab will be offered to subjects who are re-randomized to apremilast (Arm 2b) and are PASI 50 non-responders at Week 28 (Arm 3a) or Week 40 (Arm 3b).

This is an open-label study; however, the efficacy assessor will remain blinded to each subject's treatment, clinical laboratory results, and all subject safety data during the course of the study.

2.4 Sample Size Determination

This study plans to enroll a total of N = 330 subjects, to be sufficiently powered to detect the treatment difference between risankizumab (RZB) and apremilast (APR) with respect to the co-primary endpoints in Period A, as well as between APR/RZB and APR/APR with respect to the primary endpoint in Period B:

- Assuming that 50% of APR-PASI75NR subjects in the APR/RZB group and 20% of APR-PASI75NR subjects in the APR/APR group will achieve PASI 90 at Week 52 in Period B, the sample size of 120 re-randomized APR-PASI75NR subjects (60 subjects per group) will have more than 90% power to detect the treatment difference between APR/RZB and APR/APR, using a Chi-square test with a 2-sided significance level of 0.05.
- Assuming 55% of subjects initially randomized to APR will be re-randomized at Week 16 under the APR-PASI75NR (as defined by the PASI 75 response at Week 16) stratum, the total sample size for the initial APR group is 120/55% = 220. Assuming the treatment difference is at least 40% with respect to the co-primary endpoints of PASI 90 and sPGA of 0 or 1 at Week 16 in Period A, the total sample size of N = 330 subjects (RZB: 110, APR: 220) will

also provide more than 90% power to detect the treatment difference between RZB and APR, using a Chi-square test with a 2-sided significance level of 0.05.

3.0 Endpoints

3.1 Primary Endpoints

The co-primary endpoints of this study in Period A are:

- Achievement of PASI 90 at Week 16, among subjects in the ITT_A Population.
- Achievement of sPGA 0 or 1 with at least 2-grade improvement from Baseline at Week 16, among subjects in the ITT_A Population.

The primary endpoint of this study in Period B is:

- Achievement of PASI 90 at Week 52, among subjects in the ITT_B_NR Population.

3.2 Secondary Endpoints

The ranked and only secondary endpoint of this study in Period A is:

1. Achievement of PASI 75 at Week 16, among subjects in the ITT_A Population.

The ranked secondary endpoints of this study in Period B are:

1. Achievement of PASI 75 at Week 52, among subjects in the ITT_B_NR Population.
2. Achievement of sPGA 0 or 1 with at least 2-grade improvement from Baseline at Week 52, among subjects in the ITT_B_NR Population.

3.3 Other Efficacy Endpoints

All variables listed in Section 3.1 and Section 3.2 as primary or ranked secondary endpoints will also be analyzed at all other visits collected. Additionally, the following endpoints will be evaluated at all visits collected:

- Achievement of 100% reduction from Baseline in PASI (PASI 100)
- Achievement of sPGA 0
- Change from Baseline in PASI
- Percent change from Baseline in PASI
- Change from Baseline in body surface area (BSA)
- Achievement of Palmoplantar Investigator's Global Assessment (ppIGA) 0 or 1, among subjects with Baseline ppIGA of at least 3
- Change from Baseline in Psoriasis Scalp Severity Index (PSSI), among subjects with Baseline scores > 0
- Change from Baseline in Nail Psoriasis Severity Index (NAPSI), among subjects with Baseline scores > 0
- 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
- Change from Baseline in Work Productivity and Activity Impairment (WPAI) scores
- Treatment Satisfaction Questionnaire for Medication version 9 (TSQM-9) scores (at each post-Baseline visit in Period A)
- Change from Entry of Period B in TSQM-9 scores, among subjects who are re-randomized at Week 16
- Change from Baseline in Psoriasis Symptoms Scale (PSS) scores
- Percentage change from Baseline in PSS, among subjects with baseline scores > 0
- Achievement of PSS 0

- Achievement of PSS 0 or 1

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 (hematology and chemistry)

3.5 Additional Endpoints

Not applicable.

4.0 Analysis Populations

The following analysis populations will be used for the analyses.

The Intent to Treat Population in Period A (ITT_A) includes all randomized subjects. The ITT_A Population will be used for all efficacy analyses in Period A.

The Intent to Treat Population for APR-PASI75NRs in Period B (ITT_B_NR) includes subjects randomized at Baseline (Day 1), fail to achieve PASI 75 at Week 16 and are re-randomized. The ITT_B_NR Population will be used for the analysis comparing switch-to-RZB (APR/RZB) vs. continue-with-APR (APR/APR) among APR-PASI75NRs in Period B.

The Intent to Treat Population for APR-PASI75 responders in Period B (ITT_B_R) includes subjects randomized to APR at baseline, achieve PASI 75 at Week 16 and are re-randomized (APR-PASI75Rs). The ITT_B_R Population will be used for the summary of efficacy for APR/RZB and APR/APR among APR-PASI75Rs in Period B.

The Intent to Treat Population for Period B combined Population (ITT_B_Combined) includes subjects randomized to APR at Baseline and are re-randomized at Week 16. The ITT_B_Combined Population will be used for combined analysis (APR-PASI75NRs and APR-PASI75Rs) between subjects re-randomized to APR/RZB and APR/APR in Period B.

The Intent to Treat Population for Long Term Efficacy (ITT_LT) includes subjects randomized to receive continuous RZB (i.e., all subjects randomized to RZB at Baseline) and subjects randomized and re-randomized to receive continuous APR, including half of the subjects randomized to APR at Baseline and prematurely discontinued from study before entering Period B. The ITT_LT Population will be used for the analysis comparing long term efficacy between continuous treatment of RZB/RZB and APR/APR in Period B. Of note, the algorithm to select half of the subjects randomized to APR at Baseline and prematurely discontinued from study before entering Period B is defined as follow:

1. Rank all such subjects by their subject ID in ascending order (as 1, 2, 3...).
2. Select subjects whose ranks are even numbers.

The Intent to Treat Population for Rescued Population (ITT_Res) includes subjects who are re-randomized to APR at Week 16, fail to achieve PASI 50 at Week 28 and/or Week 40, and receive RZB as the rescue medication. The ITT_Res Population will be used for the summary of efficacy after being rescued with RZB.

Each ITT Population will have a corresponding Safety Population, consists of subjects who received at least one dose of study drug during the specific Study Period. In addition, an all risankizumab treated (ALL_RZB) Population will be utilized for a comprehensive safety summary of RZB. The ALL_RZB Population is defined as all subjects who receive at least one dose of RZB as the study drug.

5.0 Subject Disposition

A summary of subject disposition in Period A will be provided among the ITT_A Population, where the number of subjects in each of the following categories will be summarized for each treatment group:

- Subjects randomized in the study;
- Subjects who took at least one dose of study drug in Period A;
- Subjects who completed protocol-specified treatment in Period A;
- Subjects who prematurely discontinued study drug in Period A (by reasons);
- Subjects who completed Period A

Summaries of subject disposition in Period B will be provided among the ITT_B_R and the ITT_B_NR Populations, as well as the RZB group in the ITT_LT Population, where the number of subjects in each of the following categories will be summarized for each treatment group:

- Subjects re-randomized in Period B (or Subjects who entered Period B among the ITT_LT Population);
- Subjects who took at least one dose of study drug in Period B;
- Subjects who received RZB as the rescue medication (not applicable to the ITT_LT Population);
- Subjects who completed protocol-specified treatment in Period B;
- Subjects who prematurely discontinued study drug in Period B (by reasons);
- Subjects who completed the study.

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

Duration of treatment will be summarized by treatment groups among the Safety_A, Safety_B_R, Safety_B_NR, and Safety_Res Populations as well as the RZB group in the Safety_LT Population. In addition, duration of any RZB will be summarized among the ALL_RZB Population.

The duration of RZB 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 for Period A + 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 for Period B + 1 day]) minus the first dose date in Period B.

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

ALL_RZB: 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 for Period B + 1 day]) minus the first dose date of risankizumab.

The duration of APR for each period is defined as follows:

Period A: the minimum of (the last dose date in Period A + 1 day, 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 for Period A + 1 day]) minus the first dose date in Period A.

Period B: the minimum of (the last dose date in Period B + 1 day, the start date of rescue RZB if applicable, and the end of study date + 1 day [and cutoff date during the Primary Analysis for Period B + 1 day]) minus the first dose date in Period B.

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_A, Safety_B_R, Safety_B_NR, and Safety_Res Populations as well as the RZB group in the Safety_LT Population.

The compliance for RZB will be summarized by the percentage of planned injections which are administered in each analysis period. The compliance for APR will be summarized by the number of APR tablets actually taken (i.e., the difference between the number of tablets dispensed and the number of tablets returned) by the subject divided by the number of tablets planned to be taken by the subject in each analysis period. When computing the compliance 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. For subjects who prematurely discontinued the study drug, the planned injections/tablets will only be counted prior to that scheduled visit of discontinuation. For subjects who are rescued, the planned APR tablets will only be counted prior to the first dose date of rescue RZB.

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 PASI, Baseline NAPSI, Baseline PSSI, Baseline BSA, Baseline WPAI, Baseline PSS and Baseline DLQI.

Demographics and baseline disease characteristics will be summarized among the ITT_A, ITT_B_R, ITT_B_NR, and ITT_Res Populations, 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 group. The system organ class (SOC) will be presented in alphabetical order, and the preferred terms will be presented in alphabetical order within each SOC. Subjects reporting more than one condition/diagnosis will be counted only once in each row (SOC or preferred term).

Medical history will be summarized among the ITT_A, ITT_B_R, and ITT_B_NR Populations.

7.3 Prior and Concomitant Medications

Prior and concomitant medications will be summarized by generic name. A prior medication is defined as any medication taken prior to the date of the first dose of study drug. A concomitant medication is defined as any medication that started prior to the date of the first dose of study drug and continued to be taken after the first dose of study drug or any medication that started on or after the date of the first dose of study drug, but not after the minimum of {the date of the last dose of RZB plus 140 days, the date of the last

dose of APR plus 28 days}. The number and percentage of subjects taking medications will be summarized by generic drug name based on the World Health Organization (WHO) Drug Dictionary for both prior and concomitant medications.

In addition, 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_A, ITT_B_R, and ITT_B_NR Population as well as the RZB group in the ITT_LT Population.

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

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

- Subjects who receive alternative PsO treatment regardless of premature discontinuation of study drug before Week 16 will be counted as non-responders.

The primary efficacy endpoint (defined in Section 3.1) and ranked secondary endpoints (defined in Section 3.2) in Period B will be analyzed in the ITT_B_NR Population and the following method will be used to address the potential intercurrent events:

- Subjects who receive alternative PsO treatment regardless of premature discontinuation of study drug before Week 52 will be counted as non-responders.
- Subjects who satisfy rescue criteria or receive the rescue medication at Week 28 or Week 40 will be counted as non-responders.

9.0 Efficacy Analyses

9.1 General Considerations

All efficacy analyses will be conducted among the ITT Population in each analysis period. All tests will be 2-sided at an alpha level of 0.05.

The Primary Analysis for Period A will be performed after all ongoing subjects have completed Week 16 and all data pertaining to Period A are cleaned. This will be the only and final analysis for efficacy in Period A.

The Primary Analysis for Period B will be performed after all ongoing subjects have completed Week 52 and all data pertaining to Period B are cleaned. This will be the only and final analysis for efficacy in Period B.

The final analysis will be performed upon study completion.

Categorical Endpoints:

For categorical endpoints, the number and proportion of subjects who achieved the endpoint, as well as the 95% confidence interval (CI) of that proportion, will be provided by each treatment group among all ITT Populations.

In addition, comparisons between treatment groups will be made in the following analyses, with the estimated treatment difference, as well as the corresponding 95% CI and the p-value provided among the following populations:

- ITT_A Population: Comparisons in Period A will be made between RZB and APR using the Cochran-Mantel-Haenszel (CMH) test, adjusting for the actual values of stratification factors at Baseline: Baseline body weight (≤ 100 kg, > 100 kg) and prior exposure to any systemic and/or biologic treatment for psoriasis (0, ≥ 1). In case of any stratum with zero subject in either treatment group, no stratification factor will be controlled.
- ITT_B_NR Population: Comparisons in Period B will be made between APR/RZB and APR/APR using the chi-square test.

- ITT_B_Combined Population: Comparisons in Period B will be made between APR/RZB and APR/APR using the CMH test, adjusting for the actual values of the stratification factor of re-randomization: PASI 75 response status at Week 16 (Yes/No). In case of any stratum with zero subject in either treatment group, no stratification factor will be controlled.
- ITT_LT Population: Comparisons in Period B will be made between RZB/RZB and APR/APR using the Cochran-Mantel-Haenszel (CMH) test, adjusting for the actual values of stratification factors at Baseline: Baseline body weight (≤ 100 kg, > 100 kg) and prior exposure to any systemic and/or biologic treatment for psoriasis (0, ≥ 1). In case of any stratum with zero subject in either treatment group, no stratification factor will be controlled.

Continuous Endpoints:

For continuous endpoints, the Baseline mean, visit mean, as well as the Least Square (LS) mean, 95% confidence interval (CI) and Standard Error (SE) of that mean, will be provided by each treatment group among all ITT Populations.

In addition, comparisons between treatment groups will be made in the following analyses, with the estimated treatment difference, as well as the corresponding 95% CI, SE and the p-value will be provided among the following populations:

- ITT_A Population: Comparisons in Period A will be made between RZB and APR 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 at Baseline, visit and treatment-by-visit interaction as covariates.
- ITT_B_NR Population: Comparisons in Period B comparisons will be made between APR/RZB and APR/APR 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 after Entry of Period B, adjusting for the fixed effects of treatment, visit and treatment-by-visit interaction as covariates.

- ITT_Combined Population: Comparisons in Period B comparisons will be made between APR/RZB and APR/APR 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 after Entry of Period B, adjusting for the fixed effects of treatment, actual values of stratification factor at re-randomization, visit and treatment-by-visit interaction as covariates.
- ITT_LT Population: Comparisons in Period B comparisons will be made between RZB/RZB and APR/APR 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 after Entry of Period B, adjusting for the fixed effects of treatment, actual values of stratification factor at Baseline, visit and treatment-by-visit interaction as covariates.

"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.

"Entry of Period B" refers to the last non-missing observation on or before:

- The first dose date on or after re-randomization, or the date of re-randomization if no study drug is administered in Period B, for subjects who are randomized to APR at Baseline.
- The first dose date on or after Week 16, or the first visit date on or after Week 16 if no study drug is administered in Period B, for subjects who are randomized to RZB at Baseline.

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-MI) will be the approach for handling missing data in the analysis of categorical endpoints among all ITT Populations except for the ITT_Res Population. The NRI-MI 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 receive alternative PsO treatment regardless of premature discontinuation of study drug or receive RZB as the rescue medication 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 D.

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

Observed Cases (OC) will be the approach for handling missing data in the analysis of categorical endpoints among the ITT_Res Population. The OC 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 OC analysis for that visit. OC will exclude values after a subject receive alternative PsO treatment regardless of premature discontinuation of study drug.

9.2.2 Continuous Endpoints

Mixed-Effect Model Repeat Measurement (MMRM) will be the primary and only approach to handle missing data for continuous endpoints among all ITT Populations except for the ITT_Res Population. 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 (but not all) post-baseline visits during the analysis period. Subjects' observations after the use of alternative PsO treatment regardless of premature discontinuation of study drug or after the use of RZB as the rescue medication 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.

- Comparison of additional endpoints among the ITT_A and ITT_LT Populations will be made adjusting for treatment, visit and treatment-by-visit interaction, actual values of stratification factors at Baseline, and baseline value (if not duplicating the stratification factors).
- Comparison of additional endpoints among the ITT_B_NR Population will be made adjusting for treatment, visit and treatment-by-visit interaction and baseline value.

- Comparison of additional endpoints among the ITT_B_Combined Population will be made adjusting for treatment, visit and treatment-by-visit interaction, actual values of stratification factors at re-randomization, and baseline value (if not duplicating the stratification factors).
- Summary statistics will be provided for additional endpoints among the ITT_B_R Population, adjusting for treatment, visit and treatment-by-visit interaction and baseline value.

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

Observed Cases (OC) will be the approach for handling missing data in the analysis of continuous endpoints among the ITT_Res Population: The OC 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 OC analysis for that visit. OC will exclude values after a subject receive alternative PsO treatment regardless of premature discontinuation of study drug.

9.3 Primary Efficacy Endpoints and Analyses

9.3.1 Primary Efficacy Endpoints

The co-primary endpoints of this study in Period A are:

- Achievement of PASI 90 at Week 16, among subjects in the ITT_A Population.
- Achievement of sPGA 0 or 1 with at least 2-grade improvement from Baseline at Week 16, among subjects in the ITT_A Population.

The primary endpoint of this study in Period B is:

- Achievement of PASI 90 at Week 52, among subjects in the ITT_B_NR Population.

9.3.2 Main Analysis of Primary Efficacy Endpoints

The primary efficacy endpoints will be analyzed as specified in Section 9.1.

The attributes of the estimands corresponding to the primary efficacy endpoints in Period A and Period B are summarized in Table 1.

Table 1. Summary of the Estimand Attributes of the Primary Efficacy Endpoints

Estimand Label	Attributes of the Estimands				
	Treatment	**Endpoint**	**Population**	**Handling of Intercurrent Events**	**Statistical Summary**
PASI 90 at Week 16	RZB and APR	PASI 90 at Week 16	ITT_A	• Subjects who receive alternative PsO treatment regardless of premature discontinuation of study drug before Week 16 will be considered as non-responders	Difference in the percentage of subjects achieving PASI 90 at Week 16 in RZB group in comparison with APR group
sPGA 0/1 at Week 16		sPGA 0/1 with at least 2-grade improvement from Baseline at Week 16			Difference in the percentage of subjects achieving sPGA 0/1 with at least 2-grade improvement from Baseline at Week 16 in RZB group in comparison with APR group
PASI 90 at Week 52	APR/RZB and APR/APR	PASI 90 at Week 52	ITT_B_NR	• Subjects who receive alternative PsO treatment regardless of premature discontinuation of study drug before Week 52 will be considered as non-responders • Subjects who satisfy rescue criteria or receive the rescue medication at Week 28 or Week 40 will be considered as non-responders	Difference in the percentage of subjects achieving PASI 90 at Week 52 in APR/RZB group in comparison with APR/APR group

9.3.3 Sensitivity and Supplementary Analyses of the Primary Efficacy Endpoints

There will be no sensitivity analysis for the primary efficacy endpoints.

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 secondary efficacy endpoints will be analyzed as specified in Section 9.1.

The attributes of the estimands corresponding to the key secondary efficacy endpoints are summarized in Table 2.

Table 2. Summary of the Estimand Attributes of the Key Secondary Efficacy Endpoints

Estimand Label	Attributes of the Estimands				
	Treatment	**Endpoint**	**Population**	**Handling of Intercurrent Events**	**Statistical Summary**
PASI 75 response at Week 16	RZB and APR	PASI 75 response at Week 16	ITT_A	• Subjects who receive alternative PsO treatment regardless of premature discontinuation of study drug before Week 16 will be considered as non-responders	Difference in the percentage of subjects achieving PASI 75 response at Week 16 in RZB group in comparison with APR group
PASI 75 at Week 52	APR/RZB and APR/APR	PASI 75 response at Week 16	ITT_B_NR	• Subjects who receive alternative PsO treatment regardless of premature discontinuation of study drug before Week 52 will be considered as non-responders • Subjects who satisfy rescue criteria or receive the rescue medication at Week 28 or Week 40 will be considered as non-responders	Difference in the percentage of subjects achieving PASI 75 at Week 52 in APR/RZB group in comparison with APR/APR group
sPGA 0/1 at Week 52		sPGA 0/1 with at least 2-grade improvement from Baseline at Week 52			Difference in the percentage of subjects achieving sPGA 0/1 with at least 2-grade improvement from Baseline at Week 52 in APR/RZB group in comparison with APR/APR group

9.4.3 Sensitivity and Supplementary Analyses for Key Secondary Efficacy Endpoints

There will be no sensitivity analysis for the ranked secondary efficacy endpoints.

9.5 Additional Efficacy Analyses

All additional efficacy endpoints will be analyzed as specified in Section 9.1.

Period A: Additional efficacy endpoints will be compared between the RZB and APR treatment groups among the ITT_A Population at each visit.

Period B: Additional efficacy endpoints will be compared between the APR/RZB and APR/APR treatment groups among the ITT_B_NR and ITT_B_Combined Populations, and between the RZB/RZB and APR/APR treatment groups among the ITT_LT Population, at each visit in Period B.

Additional efficacy endpoints will also be summarized by the APR/RZB and APR/APR treatment groups among the ITT_B_R Population at each visit in Period B.

During the Period of Rescue RZB: Additional efficacy endpoints will be summarized among the ITT_Res Population at each visit after the use of Rescue RZB.

9.6 Efficacy Subgroup Analyses

Subgroup analysis will be performed on co-primary endpoints in Period A based on the following subgroups, among the ITT_A Population:

- Age group: < 40, ≥ 40 to < 65, ≥ 65
- Sex (male, female)
- Race (white, non-white)
- Smoking (current, ex or never)
- Body mass index (normal: < 25, overweight: ≥ 25 to < 30, obese: ≥ 30)
- Baseline PASI score (by median)
- Psoriatic arthritis (yes, no)

- Prior systemic (including biologic) treatment (0, ≥ 1)

Age ≥ 65 years or BMI ≥ 30 subgroups will be combined with their adjacent subgroup when having fewer than 10% subjects.

10.0 Safety Analyses

10.1 General Considerations

Safety data will be summarized among the Safety_A, Safety_B_R, Safety_B_NR, and Safety_LT Populations. Safety summaries will be presented by treatment group. For the safety analysis, subjects are analyzed based on the treatment actually received, determined by the first dose of study drug that the subject received during the specific analysis period.

The overview of TEAEs, and potentially clinically important (PCI) findings in laboratory variables and vital sign variables will also be summarized among the Safety_B_Combined, Safety_Res, and ALL_RZB Populations.

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

A treatment-emergent adverse event (TEAE) in each analysis period is defined as follows:

A TEAE in Period A is defined as any AE with an onset date on or after the first dose of study drug in Period A, and

- Within the minimum of (i) 140 days after the last dose of RZB in Period A and (ii) the first dose date of study drug in Period B for subjects who receive RZB at Baseline.
- Within the minimum of (i) 28 days after the last dose of APR in Period A and (ii) the first dose date of study drug in Period B for subjects who receive APR at Baseline.

A TEAE in Period B is defined as any AE with an onset date on or after the first dose of study drug in Period B and

- Within 140 days after the last dose of RZB in Period B for subjects who receive RZB in Period B.
- Within the minimum of (i) 28 days after the last dose of APR in Period B and (ii) the first dose date of rescue risankizumab (if any) for subjects who receive APR in Period B.

A TEAE during the administration of rescue risankizumab is defined as any AE with an onset date on or after the first dose of rescue risankizumab and within 140 days after the last dose of rescue risankizumab.

A TEAE during the administration of risankizumab (i.e., among the ALL_RZB Population) is defined as any AE with an onset date on or after the first dose of risankizumab and within 140 days after the last dose of risankizumab.

10.2.2 Adverse Event Overview

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

- 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 leading to death
- Any treatment-emergent Areas of Safety Interest, as defined in Appendix B.

All deaths will also be summarized:

- Deaths occurring ≤ 140 days after last dose of RZB or ≤ 28 days after last dose of APR
- Deaths occurring > 140 days after last dose of RZB or > 28 days after last dose of APR

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

The number and percentage of subjects experiencing TEAEs will be tabulated using the Medical Dictionary for Regulatory Activities (MedDRA) system organ class and preferred term, by severity, and by relationship to the study drug as assessed by the investigator. 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 risankizumab group.

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

Exposure-adjusted AEs per 100 patient-years will be provided, where AEs per 100 patient-years of exposure are 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 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

RZB arm: 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 for Period A + 1 day]) minus the first dose date in Period A, normalized by 365.25 and rounded to one decimal place.

APR arm: the minimum of (the last dose date in Period A + 28 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 for Period A + 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

RZB/RZB or APR/RZB arm: 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 for Period B + 1 day]) minus the first dose date in Period B, normalized by 365.25 and rounded to one decimal place.

APR/APR arm: the minimum of (the last dose date in Period B + 28 days, the first dose date of rescue RZB if not missing, and the end of study date + 1 day [and the cutoff date during the Primary Analysis for Period B + 1 day]) minus the first dose date in Period B, normalized by 365.25 and rounded to one decimal place.

Rescue RZB: the minimum of (the last dose date of rescue RZB + 140 days, and the end of study date + 1 day [and the cutoff date during the Primary Analysis for Period B + 1 day]) minus the first rescue RZB dose date, 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 for Period A + 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 discontinuation of study drug 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 Safety Topics of Interest

Safety Topics of Interest (STI) will be summarized according to the search criteria 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.

Listings of selected Areas of safety interest will also 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 where SAE-related laboratory assessments on or before the first dose of study drug will be excluded. The clinical laboratory tests defined in the protocol operations manual (e.g., hematology and clinical chemistry) will be summarized.

Mean change from Baseline to each applicable post-baseline visit will be summarized for selected laboratory variables among the Safety_A, Safety_B_R, Safety_B_NR, and Safety_LT Populations. 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 among Safety_A, Safety_B_R and Safety_B_NR, and Safety_LT Populations categorized by the toxicity grade according to NCI CTCAE Version 4.03[1] 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 will be summarized among all safety populations.

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 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 will be categorized as follows:

- 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× ULN and Total bilirubin > 1.5× ULN
- ALT and/or AST > 3× ULN and Total bilirubin > 2× ULN
- ALT > 3× ULN and Total bilirubin > 1.5× ULN
- ALT > 3× ULN and Total bilirubin > 2× ULN

The listing will include all subjects who met any of the following four criteria:

- ALT > 3 × ULN, or
- AST > 3 × 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 × ULN or AST of > 3 × ULN
- Total bilirubin ≥ 2 × ULN

Urinalysis and pregnancy testing results will be provided in listings only.

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 among the Safety_A, Safety_B_R, Safety_B_NR, and Safety_LT Populations. 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 C) among all safety populations. 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

There will be no safety subgroup analyses for this study.

10.6 Other Safety Analyses

There will be no other safety analyses for this study.

11.0 Other Analyses

There will be no other safety analyses for this study.

12.0 Interim Analyses

12.1 Data Monitoring Committee

There is no data monitoring committee for this open label study.

13.0 Overall Type-I Error Control

Overall type-I error for Period A will be controlled by testing the co-primary endpoints, followed by the ranked and only secondary endpoint, in a hierarchical order as described in Section 3.1 and Section 3.2.

Overall type-I error for Period B will be controlled by testing the primary endpoint, followed by the ranked secondary endpoints, in a hierarchical order as described in Section 3.1 and Section 3.2.

14.0 Version History

Table 3. SAP Version History Summary

Version	Date	Summary
1.0	28 July 2021	Original version
2.0	05 May 2022	• Changed Area of safety interest (ASI) to Safety Topics of Interest (STI) to be aligned with the wording of PSSAP. • Changed liver function criteria in Section 10.3 to be aligned with PSSAP. • Changed 'Malignant Tumours' to 'Malignancies' in Appendix B to be consistent with the wording in PSSAP.
3.0	26 February 2023	• Changed 'Phase 3b' to 'Phase 4' in the title and Section 1.0 to be consistent with the protocol. • Added Primary Analysis for Period B in Section 9.1. • Added adjusting variables visit and treatment by visit interaction in Mixed-effect Model Repeat Measurement (MMRM) in Section 9.2.2 to be consistent with Section 9.1. • Changed 'cutoff for Period A' to 'cutoff date for Period B' in Section 6.0 and Section 10.2.4. • Added '[and cutoff date during the Primary Analysis for Period B + 1 day]' to Rescue RZB duration in Section 6.0. • CMQ codes were revised for MedDRA 25.1 in Appendix B.

15.0 References

1. 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. Definition of Safety Topics of Interest

Safety Topics of Interest (STI) will be identified using the following search criteria:

Safety Topics of Interest	Search Criteria	
MACE	Adjudicated terms will be identified as described in PSSAP Table 4a using CECAT and CETERM from the CE SDTM dataset.	
Extended MACE	Adjudicated terms will be identified as described in PSSAP Table 4a (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)
Malignancies	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, and the PCI criteria for vital sign findings are described in Table C-3.

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

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

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

Table C-2. 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
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 (> 3.0 × BL)
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 value must be more extreme than the baseline value to be considered a potentially clinically important finding.

Table C-3. 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 D. 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-MI)

2.1 Overall Description of the Method

For a dichotomized outcome variable with missing data, the NRI-MI 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 specific 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-MI) 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, PASI 90 (at least a 90% reduction in PASI relative to Baseline), the multiple imputation will be applied to the original scale, PASI (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., PASI, the original scale of PASI 90, 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., PASI 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.

- For analyses for the ITT_A Population: Covariates included in the model are treatment group, Baseline body weight as a continuous variable, prior exposure to any systemic and/or biologic treatment for PsO (0, ≥ 1), Baseline and all post-baseline visits of the outcome variable up to the end of Period A.
- For analyses for the ITT_B_Combined, ITT_B_NR and ITT_B_R Populations: Covariates included in the model are treatment group, PASI score at Baseline and at the Entry of Period B, and the outcome variable collected at the Entry of Period B and all later visits.

- For analyses for the ITT_LT Population: Covariates included in the model are treatment group, Baseline body weight as a continuous variable, prior exposure to any systemic and/or biologic treatment for PsO (0, ≥ 1), PASI score at Baseline and at the Entry of Period B, and the outcome variable collected at the Entry of Period B and all later visits. Subjects randomized to RZB at Baseline and did not enter Period B will be included in the analysis and will be counted as non-responders. To accommodate the 1:1 re-randomization, half (round down to integer) of the subjects randomized to APR at Baseline and did not enter Period B will be included in the analysis and will be counted as non-responders.

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 the same variables for each Population as listed in the augmentation step. 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., PASI 90 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

For the ITT_A and ITT_LT Populations: The statistical analysis will use the Cochran-Mantel-Haenszel (CMH) test adjusting for the actual values of stratification factors at Baseline: Baseline body weight (≤ 100 kg, > 100 kg), prior exposure to any systemic and/or biologic treatment for PsO (0, ≥ 1).

For the ITT_B_Combined Population: The statistical analysis will use the CMH test adjusting for the actual values of the stratification factor at re-randomization: PASI 75 response status at the Entry of Period B (Yes/No).

For ITT_B_NR Population: The statistical analysis will use the Chi-square test.

For ITT_B_R Population: Summary statistics will be calculated. No statistical comparison is conducted.

2.5.1 Analysis of Each Dataset

For each of the K 'complete' datasets, the statistical tests as specified above will be used to estimate the treatment difference between treatment groups and the corresponding standard error (except for ITT_B_R Population, in which only summary statistics of each group will be calculated).

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 k[th] 'complete' dataset, denoting the estimate of the treatment difference q by $\tilde{\theta}_k$ from the k[th] '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*/
/*PASI 90 at Week 16 is used as an example in the sample code*/

/*PRE-AUGMENTATION - CREATE DUMMY FOR CATEGORICAL VARIABLES*/
/*****************************************/
DATA PASI_2; SET PASI;
  /*THE MCMC STATMENT BELOW ASSUMES MULTI-VARIATE NORMAL*/
  IF TRT01PN=1 THEN TRT1=1; ELSE TRT1=0;
  /*BASELINE WEIGHT AND PRIOR BIO/SYS CATEGORY*/
  IF PSOBLGN  = 1 THEN REG1 = 1 ;ELSE REG1 = 0;
RUN;

/*AUGMENTATION STEP -- TO HAVE 30 MONOTONE MISSING DATASETS*/
PROC MI DATA= PASI 2 OUT= PASI MONO NIMPUTE=30 SEED= 12345 /*ACTUAL
RANDOM SEED PRE-DEFINED AS THE NUMERICAL VALUE OF THE DATE OF FIRST
SUBJECT FIRST DOSE*/
  ROUND=. . . 0.1 0.1 0.1 /*VALUE ROUND TO 1ST DECIMAL*/
  MIN=. . . 0 0 0  /*MINIMUM VALUE OF PASI IS 0*/
  MAX=. . . 72 72 72   /*MAXIMUM VALUE OF PASI 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 REG1 WTBL 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= PASI MONO OUT= PASI FULL NIMPUTE=1 SEED= 54321  /*ACTUAL
RANDOM SEED PRE-DEFINED AS THE NUMERICAL VALUE OF THE DATE OF LAST
SUBJECT FIRST DOSE */
  ROUND= =. . . 0.1 0.1 0.1 /*VALUE ROUND TO 1ST DECIMAL*/
  MIN=. . .  0 0 0  /*MINIMUM VALUE OF PASI IS 0*/
  MAX=. . .  72 72 72 /*MAXIMUM VALUE OF PASI IS 72*/
  MINMAXITER=1000;
  /*CLASS CATEGORICAL VARIABLES TRT1 AND REG2*/
  CLASS TRT1 REG1;
  VAR TRT1 REG1 WTBL 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, PASI 90 AT WEEK 16*/
DATA ALL; SET PASI FULL;
  IF 0<=WK16<=0.1*BASE THEN PASI90_16=1;
    ELSE PASI90_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 PASI90 NRI 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 PASI90_16= PASI90NRI_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* PASI90 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 PASI90_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.