Statistical Analysis Plan for Study M16-833

A Phase 2, Multicenter, Randomized, Placebo-Controlled, Double-Blind Study to Evaluate the Safety and Efficacy of Risankizumab in Adult Subjects with Moderate to Severe Hidradenitis Suppurativa

Date: 19 January 2021

Version 4.0

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 M16-833, A Phase 2, Multicenter, Randomized, Placebo-Controlled, Double-Blind Study to Evaluate the Safety and Efficacy of Risankizumab in Adult Subjects with Moderate to Severe Hidradenitis Suppurativa (HS).

Study M16-833 evaluates the efficacy and safety of risankizumab 180 mg and 360 mg versus placebo for the treatment of signs and symptoms of moderate to severe HS in adult subjects diagnosed for at least one year before the Baseline visit.

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 180 mg and 360 mg versus placebo for the treatment of signs and symptoms of moderate to severe HS in adult subjects diagnosed for at least one year before the Baseline visit.

The primary efficacy objective is based on the achievement of Hidradenitis Suppurativa Clinical Response (HiSCR; defined as at least a 50% reduction in the total AN count with no increase in abscess count and no increase in draining fistula count relative to Baseline) at Week 16 of treatment with risankizumab 360 mg or 180 mg when compared to placebo among the Intent-to-Treat (ITT) Population, which consists of all randomized subjects (Section 5.0).

The corresponding null hypothesis for the primary endpoint is:

- There is no difference between each risankizumab dose and placebo, with respect to the proportion of subjects who achieved HiSCR at Week 16.

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

- Achievement of HiSCR at Week 16 without the initiation of any antibiotics for HS-related infections and without premature discontinuation of study drug due to lack of efficacy prior to Week 16.

The secondary efficacy objectives are based on the ranked secondary endpoints as defined in Section 3.2.

The null hypothesis for the first ranked secondary endpoint is:

- There is no difference between each risankizumab dose and placebo, in the proportion of subjects achieving NRS30 in PGA Skin Pain at Week 8 among subjects with Baseline NRS ≥ 3.

The corresponding estimand is

- Achievement of NRS30 in PGA Skin Pain at Week 8 among subjects with Baseline NRS ≥ 3 without the initiation of any antibiotics for HS-related infections, without the impact of analgesics for HS-related skin pain at Week 8 and without premature discontinuation of study drug due to lack of efficacy prior to Week 8.

The null hypothesis for the second ranked secondary endpoint is:

- There is no difference between each risankizumab dose and placebo, in the proportion of subjects achieving NRS30 in PGA Skin Pain at Week 16 among subjects with Baseline NRS ≥ 3.

The corresponding estimand is

- Achievement of NRS30 in PGA Skin Pain at Week 16 among subjects with Baseline NRS ≥ 3 without the initiation of any antibiotics for HS-related infections, without the impact of analgesics for HS-related skin pain at Week 16 and without premature discontinuation of study drug due to lack of efficacy prior to Week 16.

The null hypothesis for the third ranked secondary endpoint is:

- There is no difference between each risankizumab dose and placebo in the proportion of subjects who experience flare, defined as an increase in AN count of at least a 25% with a minimum increase of 2 relative to Baseline (flare), during Period A.

The corresponding estimand is

- Experience of flare during Period A.

The null hypothesis for the fourth ranked secondary endpoint is:

- There is no difference between each risankizumab dose and placebo, in change from Baseline in Dermatology Life Quality Index (DLQI) at Week 16.

The corresponding estimand is

- Change from Baseline in DLQI at Week 16.

The null hypothesis for the fifth ranked secondary endpoint is:

- There is no difference between each risankizumab dose and placebo, in change from Baseline in HS-related worst swelling - assessed based on the Hidradenitis Suppurativa Symptom Assessment (HSSA) at Week 16.

The corresponding estimand is

- Change from Baseline in HS-related worst swelling - assessed based on the HSSA at Week 16.

The null hypothesis for the sixth ranked secondary endpoint is:

- There is no difference between each risankizumab dose and placebo, in change from Baseline in HS-related worst odor - assessed based on the HSSA at Week 16.

The corresponding estimand is

- Change from Baseline in HS-related worst odor - assessed based on the HSSA at Week 16.

The null hypothesis for the seventh ranked secondary endpoint is:

- There is no difference between each risankizumab dose and placebo, in change from Baseline in HS-related worst drainage- assessed based on the HSSA at Week 16.

The corresponding estimand is

- Change from Baseline in HS-related worst drainage - assessed based on the HSSA at Week 16.

The Primary Analysis will be performed after the last ongoing subject has completed the Week 16 visit, and the data up to Week 16 cutoff date have been cleaned. The cutoff date will be specified in the statistical programming plan (SPP) before the database lock for the Primary Analysis and before unblinding of subject randomization.

2.2 Study Design Overview

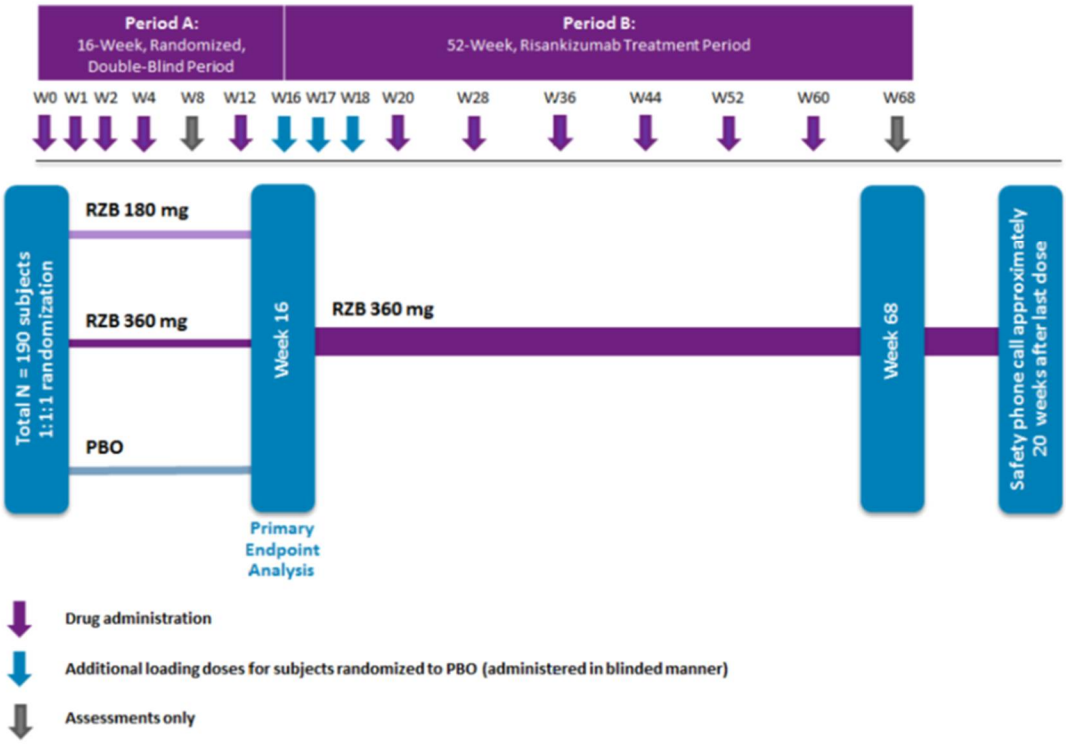
This is a Phase 2, multicenter, randomized, double-blind, parallel-group, placebo-controlled study to evaluate the safety and efficacy of 2 dose levels of risankizumab in adult subjects with moderate to severe HS diagnosed at least 1 year before the Baseline visit.

The duration of the study will be up to 85 weeks and will include an approximately 35-day screening period followed by 2 treatment periods.

- **Period A:** Subjects who meet the study's eligibility criteria (protocol Section 5.1) will be randomized at the Baseline visit, in a 1:1:1 ratio, to receive either risankizumab 180 mg or 360 mg via a subcutaneous (SC) injection, or matching placebo up to Week 16. Study drug administration for Period A will occur at Baseline, Weeks 1, 2, 4, and 12. The final efficacy evaluation of Period A will be at Week 16.
- **Period B:** Subjects who were initially randomized to placebo will then receive blinded risankizumab 360 mg at Weeks 16, 17, and 18, while subjects who were initially randomized to risankizumab will receive blinded matching placebo at the same time points to keep the treatment of Period A blinded. Starting at Week 20, all subjects will receive open-label risankizumab 360 mg every eight weeks (q8w), at Weeks 20, 28, 36, 44, 52, and 60. The final efficacy evaluation of Period B will take place at Week 68. A follow up call will be conducted approximately 20 weeks after last dose of study drug.

The schematic of the study is shown in Figure 1.

Figure 1. Study Schematic



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 the statistics department at AbbVie.

Subjects who meet the study's eligibility criteria will be initially randomized to receive either risankizumab 180 mg or 360 mg as an SC injection or matching placebo up to Week 16 in a 1:1:1 ratio.

Subjects who have been treated with anti-TNF therapy and had inadequate response to this therapy (TNF- IR) will be eligible. Subjects will be randomized by their anti-TNF use (yes or no) prior to Baseline. TNF-naïve (no) subjects will be further stratified by the worst Hurley stage across all affected anatomic regions. Therefore, there will be 4 strata in total:

- Prior exposure to anti-TNF (yes)
- Prior exposure to anti-TNF (no) with the worst Hurley Stage as I
- Prior exposure to anti-TNF (no) with the worst Hurley Stage as II
- Prior exposure to anti-TNF (no) with the worst Hurley Stage as III

All AbbVie personnel 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 initial treatment throughout the study. To maintain the blind, the risankizumab kits and placebo kits provided for the study will be identical in appearance. The IRT will provide access to unblinded subject treatment information in the case of a medical emergency.

2.4 Sample Size Determination

Under the study design with a 1:1:1 randomization, with the assumed Week 16 HiSCR response rate as summarized in Table 1, a total sample size of 222 subjects will provide 80% to 90% power to detect at least one risankizumab dose is different from placebo (with a 2-sided significance level of 0.025 for the test of each risankizumab dose versus placebo).

Table 1. Power under a 1:1:1 randomization

Assumed Dose Response Model	Assumed HiSCR Rate at Week 16 (%)			Power[1]
	Placebo (N = 74)	Risankizumab 180 mg (N = 74)	Risankizumab 360 mg (N = 74)	
Linear	26%	39.5%	53%	83.9%
EMax	26%	46.5%	53%	87.6%
EMax (with early plateau)	26%	50%	53%	90.6%

1. Power to detect at least one risankizumab dose is different from placebo, under a 2-sided significance level of 0.025 for each risankizumab dose.

3.0 Endpoints

3.1 Primary Endpoint

The primary endpoint is the proportion of subjects achieving HiSCR at Week 16. HiSCR is defined as at least a 50% reduction in the total AN count with no increase in abscess count and no increase in draining fistula count relative to Baseline.

3.2 Secondary Endpoint

The following ranked secondary endpoints will be evaluated at the time points listed:

1. Proportion of subjects achieving at least 30% reduction and at least 1 unit reduction from Baseline in Numerical Rating Scale (NRS30) in Patient's Global Assessment of Skin Pain (PGA Skin Pain) at Week 8 among subjects with Baseline Numerical Rating Scale (NRS) ≥ 3. NRS30 is based on worst skin pain in a 24-hour recall period (maximal daily pain).
2. Proportion of subjects achieving NRS30 in PGA Skin Pain at Week 16 among subjects with Baseline NRS ≥ 3.
3. Proportion of subjects who experience flare, defined as an increase in AN count of at least a 25% with a minimum increase of 2 relative to Baseline (flare), during Period A.
4. Change from Baseline in DLQI at Week 16.

5. Change from Baseline in HS-related swelling – assessed based on the HSSA at Week 16.
6. Change from Baseline in HS-related odor – assessed based on the HSSA at Week 16.
7. Change from Baseline in HS-related worst drainage – assessed based on the HSSA at Week 16.

3.3 Additional Endpoints

All variables listed as primary or ranked secondary endpoints will also be analyzed at all visits collected in addition to those listed above. The following endpoints will also be evaluated at all visits collected:

- Change and percent change from Baseline in PGA Skin Pain – at worst (maximal daily pain), among subjects who have Baseline PGA Skin Pain NRS ≥ 3
- Proportion of subjects achieving a total AN count of 0, 1, or 2.
- Proportion of subjects achieving complete elimination of inflammatory lesions by lesion type, among subjects who had the corresponding lesion type at Baseline.
- Proportion of subjects who experience at least 25% increase in inflammatory lesion counts with a minimum increase of 2 relative to Baseline, by lesion type.
- Change from Baseline in inflammatory lesion counts by lesion type.
- Percent change from Baseline in lesion counts by inflammatory lesion type, among subjects who had at least 3 of the corresponding lesion type at Baseline.
- Proportion of subjects achieving DLQI = 0 or 1.
- Proportion of subjects achieving a DLQI improvement (reduction) of ≥ 4 points among subjects with DLQI ≥ 4 at Baseline.
- Change from Baseline in symptoms assessed based on HSSA questionnaire.
- Change from Baseline in Hidradenitis Suppurativa Impact Assessment (HSIA) questionnaire.

- Change from Baseline in EuroQol 5 Dimensions 5 Levels Health State Instrument (EQ-5D-5L).
- Change from Baseline in Work Productivity and Activity Impairment (WPAI).
- Proportion of subjects achieving at least 1 grade improvement from Baseline in Patient Global Impression of Severity (PGIS) scale among subjects with Baseline PGIS of at least "Minimal."
- Proportion of subjects who report symptoms to be "Minimal" or "Absent" on the PGIS scale.
- Proportion of subjects achieving "much improved" or "very much improved" on the Patient Global Impression of Change (PGIC) scale.
- Change from Baseline in Hospital Anxiety and Depression Scale (HADS).
- Change from Baseline in high-sensitivity C-reactive protein (hsCRP).
- Proportion of subjects who experience worsening by at least 1 Hurley Stage in at least 1 affected anatomic region.

3.4 Safety Endpoints

The following evaluations will be included in the safety analyses:

- Treatment emergent adverse events (TEAEs);
- Serious adverse events (SAEs);
- Areas of safety interest (ASI);
- Adverse events (AEs) leading to discontinuation;
- Vital signs and laboratory tests.

3.5 Pharmacokinetic Endpoints

The pharmacokinetic endpoints will be analyzed separately.

4.0 Analysis Populations

The intent-to-treat (ITT) Population, which includes all randomized subjects, 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. Subjects will be analyzed according to treatment as randomized.

The Per-Protocol (PP) Population will include ITT subjects who meet all the following criteria:

- Receive at least 75% of the planned study drug prior to Week 16.
- Provide at least one post Baseline assessment on lesion count.
- Have Baseline AN count ≥ 5.
- Have Baseline draining fistula count ≤ 20.
- Do not take protocol prohibited medications that could potentially impact the efficacy assessment for the primary endpoint at Week 16.

The final criteria and the exclusion of subjects from the PP Population will be finalized before unblinding data for the Primary Analysis. Subjects in the PP population will be analyzed by treatment group as randomized.

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 one dose of study drug in Period A.
- The Safety Population in Period B (Safety_B) is defined as all subjects who received at least one dose of study drug in Period B.
- The All Risankizumab Treated (ALL_RZB) Population is defined as subjects who received at least one dose of risankizumab in the study. This population will be used to provide a comprehensive summary of safety.

For safety analysis, subjects will be analyzed as treated (determined by the first dose of study drug that the subject received).

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 for each treatment group:

- Subjects randomized in the study;
- Subjects who received randomized study drug in each period;
- Subjects who completed study;
- Subjects who discontinued study drug in each period (all reasons and primary reason);

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

6.0 Study Drug Duration and Compliance

For the ITT population, the duration of treatment will be summarized by treatment groups in each period. The duration of treatment for each period is defined as follows:

Period A: the minimum of (the last dose date in Period A + 56 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 + 56 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 All-Risankizumab Treated Period: the duration of treatment will be summarized by the following two groups:

- **Any risankizumab**: the minimum of (the last dose date of risankizumab + 56 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.

- **Risankizumab 360 mg**: the minimum of (the last dose date of risankizumab 360 mg + 56 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 360 mg.

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 ITT population. 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 ITT subjects in this treatment group who 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, waist circumference, and body mass index (BMI). Categorical demographic variables include sex, ethnicity, race, age (<40, 40 – 65, ≥ 65 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 prior exposure to TNF antagonists (yes, no), prior exposure to biologic agents that block IL-12/23/17 (yes, no), Baseline lesion counts (by lesion type), Baseline skin pain NRS (among all subjects and among subjects with baseline value ≥ 3), duration of HS (in years), Baseline Hurley stage (I, II, III), Baseline DLQI, Baseline HSSA, Baseline HSIA, Baseline WPAI, Baseline PGIS, Baseline EQ-5D-5L, Baseline HADS, Baseline hsCRP.

Demographics and baseline characteristics will be summarized among the ITT Population, overall and by treatment group. 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). Subjects with cardiovascular history and CV risk factors will also be summarized.

Medical history will be summarized among the ITT Population.

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 on or 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 date of the last dose of study drug plus 140 days. The number and percentage of subjects taking medications will be summarized by generic drug name based on the World Health Organization (WHO) Drug Dictionary for both prior and concomitant medications.

In addition, subjects' prior exposure to TNF antagonists and prior exposure to biologic agents that block IL-12/23/17 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

The primary efficacy endpoint, achievement of HiCSR at Week 16 (defined in Section 3.1), will be analyzed in the ITT Population and the following method will be used to address potential intercurrent events:

- Intercurrent events of either initiation of any antibiotics for HS-related infections or discontinuation of study drug due to lack of efficacy before Week 16 are included in the definition of primary endpoint; subjects who have either intercurrents above will be considered not achieving the response. No other intercurrent events are considered.

The first and second ranked secondary endpoints of HS-related skin pain (defined in Section 3.2), will be analyzed in the ITT Population and the following method will be used to address potential intercurrent events:

- Intercurrent events of either initiation of any antibiotics for HS-related infections or discontinuation of study drug due to lack of efficacy before Week 8 for the first ranked or before Week 16 for the second ranked secondary endpoint, as well as the intercurrent event of under the impact of analgesic use for HS-related skin pain (defined in Section 9.4.2) at Week 8 for the first ranked or at Week 16 for the second ranked secondary endpoint, are included in the definition of endpoint; subjects who have any of the intercurrents above will be considered not achieving the response.

The third ranked secondary endpoint, experiencing flare during Period A will be analyzed in the ITT Population based on observed flares. No intercurrent events are considered.

The fourth, fifth, sixth and seventh ranked secondary endpoints on changes from Baseline (defined in Section 3.2), will be analyzed in the ITT Population and the following method will be used to address potential intercurrent events:

- Intercurrent events of either initiation of any antibiotics for HS-related infections or discontinuation of study drug due to lack of efficacy before Week 16 are considered; subjects' observations after either intercurrents above will be considered as missing and handled by the MMRM model.

9.0 Efficacy Analyses

9.1 General Considerations

All efficacy analyses will be conducted in the ITT Population. In addition, Per-protocol analysis for primary endpoint will be performed. All statistical tests, between each risankizumab dose and placebo, will be performed at a two-sided alpha level of 0.025.

The Primary Analysis will be performed after the last ongoing subject has completed the Week 16 visit, and the data up to Week 16 have be cleaned. This will be the only and final analysis for the primary and ranked secondary endpoints. 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 each risankizumab dosing group and placebo using the Cochran-Mantel-Haenszel (CMH) test, adjusting for the actual values of stratification factors. In case of any stratum with zero subject in either treatment group, no stratification factor will be controlled. Non-responder imputation (NRI) will be the primary approach to handle missing values. To account for the impact of COVID-19 pandemic, the NRI will be adjusted to NRI-C, Non-Responder Imputation incorporating multiple imputation to handle missing data due to COVID-19.

For continuous endpoints, comparisons will be made between each risankizumab dosing group 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.

For variables other than PGA Skin Pain NRS and HSSA, "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 given. The "Baseline" of PGA Skin Pain NRS and HSSA is defined as the last non-missing weekly average score, calculated based on the daily scores from the past 7 days, on or before the date of the first administration of study drug or the date of randomization if no study drug is given. 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.

For any categorical endpoints of achievement, 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 receiving any prohibited antibiotics for HS-related infections or discontinued 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 non-responders at all post-baseline visits. The only exception is that when determining the HiSCR response status, a subject with missing Baseline value will be counted as a non-responder at any post-baseline visit unless a subject's abscess, inflammatory nodule, and draining fistula counts are all zero at a specific visit, then the subject will be a HiSCR responder at that visit. More details are provided in Appendix E.

For any categorical endpoints that connote worsening of disease, subjects with observed worsening outcomes, including those after receiving any prohibited antibiotics for HS-related infections or discontinued study drug due to lack of efficacy, will be counted as experienced such worsening.

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 (HiSCR 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. Taking abscess count as an example, the variables to be included in the imputation model are: treatment group, prior exposure to anti-TNF category, Baseline worst Hurley Stage, Baseline abscess count, and abscess counts at each visit up to the end of the double blind Period A. The imputed post-baseline abscess counts will be rounded to the same precision as the observed data to determine the responder status. The same imputation strategy will be applied to the inflammatory nodule count and the draining fistula count. HiSCR status will be determined accordingly. Using the Cochran-Mantel-Haenszel (CMH) model adjusted by actual stratification factors, 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 each risankizumab dose and placebo. Note that measurements will be considered as missing after receiving any antibiotics for HS-related infections or discontinued study drug due to lack of efficacy before MI. Regardless of MI imputed values, subjects receiving any antibiotics for HS-related infections or discontinued study drug due to lack of efficacy will be counted as non-responders at later visits.

Observed Case (OC) analysis will only be performed during the Primary Analysis for the summaries of categorical endpoints for visits after Week 16. 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 not be included for that visit. The OC analysis will not include values after receiving any prohibited antibiotics for HS-related infections or after discontinued study drug due to lack of efficacy.

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, 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 receiving any prohibited antibiotics for HS-related infections or 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.

Observed Case (OC) analysis will only be performed during the Primary Analysis for the summaries of continuous endpoints for visits after Week 16. 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 not be included for that visit. The OC analysis will not include values after receiving any prohibited antibiotics for HS-related infections or after discontinued study drug due to lack of efficacy.

9.3 Primary Efficacy Endpoint and Analyses

9.3.1 Primary Efficacy Endpoint

The primary endpoint is the proportion of subjects achieving HiSCR at Week 16 as defined in Section 3.1.

9.3.2 Main Analysis of Primary Efficacy Endpoint

The corresponding statistical null hypothesis to the primary endpoint is that there is no difference between each risankizumab dose and placebo, in the proportion of subjects achieving HiSCR at Week 16.

Comparison of the primary endpoint will be made between each risankizumab dose and placebo using the Cochran-Mantel-Haenszel (CMH) test, adjusting for the actual values stratification factors under a two-sided alpha level 0.025 for each dose.

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

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

	Attributes of the Estimand				
Estimand	**Treatment**	**Endpoint**	**Population**	**Intercurrent Events**	**Statistical Summary**
Primary efficacy endpoint of HiSCR at Week 16	Risankizumab 180 mg / risankizumab 360 mg / Placebo	Achievement of HiSCR at Week 16 without the initiation of any antibiotics for HS-related infections and without discontinuation of study drug due to lack of efficacy	Patients with moderate to severe HS	The intercurrents of either initiation of any antibiotics for HS-related infections or discontinuation of study drug due to lack of efficacy are included in the endpoint attribute using composite variable strategy. No other intercurrent event is considered.	Difference in proportion of subjects achieving HiSCR at Week 16 between each risankizumab dose and placebo

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.

The primary endpoint will also be analyzed among a PP population as additional sensitivity analyses, using the same approaches to handle missing data.

The attributes of the estimand corresponding to the sensitivity analysis of the primary efficacy endpoint among the PP Population is summarized in Table 3.

Table 3. Summary of the Estimand Attributes of the Sensitivity Analysis of the Primary Efficacy Endpoint among the PP Population

	Attributes of the Estimand				
Estimand	**Treatment**	**Endpoint**	**Population**	**Intercurrent Events**	**Statistical Summary**
Primary efficacy endpoint of HiSCR at Week 16	Risankizumab 180 mg / risankizumab 360 mg / Placebo	Achievement of HiSCR at Week 16 without the initiation of any antibiotics for HS-related infections and without discontinuation of study drug due to lack of efficacy	Patients with moderate to severe HS who have Baseline AN count ≥ 5, and have Baseline draining fistula count ≤ 20.	The intercurrents of either initiation of any antibiotics for HS-related infections or discontinuation of study drug due to lack of efficacy are included in the endpoint attribute using composite variable strategy. No other intercurrent event is considered.	Difference in proportion of subjects achieving HiSCR at Week 16 between each risankizumab dose and placebo

9.4 Secondary Efficacy 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

Ranked secondary efficacy endpoints will be tested between each risankizumab dose and placebo in a hierarchical order only if the null hypothesis of the primary endpoint for the corresponding risankizumab dose has been rejected.

The corresponding null hypotheses for ranked secondary endpoints with respect to each risankizumab dose are:

1. There is no difference between each risankizumab dose and placebo, in the proportion of subjects achieving NRS30 in PGA Skin Pain at Week 8 among subjects with Baseline NRS ≥ 3.
2. There is no difference between each risankizumab dose and placebo, in the proportion of subjects achieving NRS30 in PGA Skin Pain at Week 16 among subjects with Baseline NRS ≥ 3.
3. There is no difference between each risankizumab dose and placebo, in the proportion of subjects who experience flare during Period A.
4. There is no difference between each risankizumab dose and placebo, in change from Baseline in DLQI at Week 16.
5. There is no difference between each risankizumab dose and placebo, in change rom Baseline in HS-related worst swelling – assessed based on the HSSA at Week 16.
6. There is no difference between each risankizumab dose and placebo, in change from Baseline in HS-related worst odor – assessed based on the HSSA at Week 16.
7. There is no difference between each risankizumab dose and placebo, in change from Baseline in HS-related worst drainage– assessed based on the HSSA at Week 16.

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.

For the first and second ranked secondary endpoint of HS-related skin pain, any measurements under the impact of analgesic use for HS-related skin pain will be handled as follows:

- Prohibited analgesic (including opioids): A subject will be counted as a non-responder from the day that the subject initiates prohibited analgesic to 5 days after the end of such analgesic use.
- Protocol-allowed analgesic use:
  - A subject who enter the study without concomitant analgesic will be counted as a non-responder from the day that the subject initiates a protocol-allowed analgesic to 2 days after the end of such analgesic use.
  - A subject who entered the study on a stable dose of a protocol-allowed concomitant analgesic will be counted as a non-responder from the first day of an analgesic dose increase to 2 days after the end of the dose increase.

The attributes of the estimands corresponding to the ranked secondary efficacy endpoints are summarized in Table 4.

Table 4. Summary of the Estimand Attributes of the Ranked Secondary Efficacy Endpoints

	Attributes of the Estimand				
Estimand	**Treatment**	**Endpoint**	**Population**	**Intercurrent Events**	**Statistical Summary**
Ranked secondary endpoint of NRS30 at Week 8	Risankizumab 180 mg / risankizumab 360 mg / Placebo	Achievement of NRS30 in PGA Skin Pain at Week 8 without the initiation of any antibiotics for HS-related infections, without the impact of analgesics for HS-related skin pain, and without discontinuation of study drug due to lack of efficacy	Patients with moderate to severe HS among Baseline NRS ≥ 3	The intercurrents of either initiation of any antibiotics for HS-related infections, under the impact of analgesics for HS-related skin pain, or discontinuation of study drug due to lack of efficacy are included in the endpoint attribute using composite variable strategy.	Difference in proportion of subjects achieving NRS30 at Week 8 between each risankizumab dose and placebo
Ranked secondary endpoint of NRS30 at Week 16	Risankizumab 180 mg / risankizumab 360 mg / Placebo	Achievement of NRS30 in PGA Skin Pain at Week 16 without the initiation of any antibiotics for HS-related infections, without the impact of analgesics for HS-related skin pain, and without discontinuation of study drug due to lack of efficacy	Patients with moderate to severe HS among Baseline NRS ≥ 3	The intercurrents of either initiation of any antibiotics for HS-related infections, under the impact of analgesics for HS-related skin pain, or discontinuation of study drug due to lack of efficacy are included in the endpoint attribute using composite variable strategy.	Difference in proportion of subjects achieving NRS30 at Week 16 between each risankizumab dose and placebo

Table 4. Summary of the Estimand Attributes of the Ranked Secondary Efficacy Endpoints (Continued)

	Attributes of the Estimand				
Estimand	**Treatment**	**Endpoint**	**Population**	**Intercurrent Events**	**Statistical Summary**
Ranked secondary endpoint of flare during Period A	Risankizumab 180 mg / risankizumab 360 mg / Placebo	Experiencing flare during Period A	Patients with moderate to severe HS	No intercurrent event is considered.	Difference in proportion of subjects experiencing flare during Period A between each risankizumab dose and placebo
Ranked secondary endpoint of change from Baseline in DLQI at Week 16	Risankizumab 180 mg / risankizumab 360 mg / Placebo	Change from Baseline in DLQI at Week 16	Patients with moderate to severe HS	Subjects' observations after the intercurrents of either initiation of any antibiotics for HS-related infections or discontinuation of study drug due to lack of efficacy will be considered as missing and handled by MMRM.	Difference in change from Baseline in DLQI at Week 16 between each risankizumab dose and placebo
Ranked secondary endpoint of change from Baseline in HS-related worst swelling at Week 16	Risankizumab 180 mg / risankizumab 360 mg / Placebo	Change from Baseline in HS-related worst swelling at Week 16	Patients with moderate to severe HS	Subjects' observations after the intercurrents of either initiation of any antibiotics for HS-related infections or discontinuation of study drug due to lack of efficacy will be considered as missing and handled by MMRM.	Difference in change from Baseline in HS-related worst swelling at Week 16 between each risankizumab dose and placebo

Table 4. Summary of the Estimand Attributes of the Ranked Secondary Efficacy Endpoints (Continued)

	Attributes of the Estimand				
Estimand	**Treatment**	**Endpoint**	**Population**	**Intercurrent Events**	**Statistical Summary**
Ranked secondary endpoint of change from Baseline in HS-related worst odor at Week 16	Risankizumab 180 mg / risankizumab 360 mg / Placebo	Change from Baseline in HS-related worst odor at Week 16	Patients with moderate to severe HS	Subjects' observations after the intercurrents of either initiation of any antibiotics for HS-related infections or discontinuation of study drug due to lack of efficacy will be considered as missing and handled by MMRM.	Difference in change from Baseline in HS-related worst odor at Week 16 between each risankizumab dose and placebo
Ranked secondary endpoint of change from Baseline in HS-related worst drainage at Week 16	Risankizumab 180 mg / risankizumab 360 mg / Placebo	Change from Baseline in HS-related worst drainage at Week 16	Patients with moderate to severe HS	Subjects' observations after the intercurrents of either initiation of any antibiotics for HS-related infections or discontinuation of study drug due to lack of efficacy will be considered as missing and handled by MMRM.	Difference in change from Baseline in HS-related worst drainage at Week 16 between each risankizumab dose and placebo

9.4.3 Sensitivity and Supplementary Analyses for Key Secondary Efficacy Endpoints

There are no additional supportive secondary efficacy analyses planned other than those described above.

9.5 Additional Efficacy Analyses

Additional efficacy endpoints will be compared between each risankizumab dose and placebo 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. Continuous endpoints will be summarized by descriptive statistics including the mean, standard error, and the 95% confidence interval of the mean.

For efficacy endpoints of HS-related skin pain:

- Prohibited analgesic (including opioids): A subject will be counted as a non-responder from the day that the subject initiates prohibited analgesic to 5 days after the end of such analgesic use for categorical endpoints, and will have the last assessment prior to the initiation of analgesic carried forward until 5 days after the end of such analgesic use for continuous endpoints.
- Protocol-allowed analgesic use:
    - A subject who enter the study without concomitant analgesic will be counted as a non-responder from the day that the subject initiates a protocol-allowed analgesic to 2 days after the end of such analgesic use for categorical endpoints, and will have the last assessment prior to the initiation of analgesic carried forward until 2 days after the end of such analgesic use for continuous endpoints.

- A subject who entered the study on a stable dose of a protocol-allowed concomitant analgesic will be counted as a non-responder from the first day of an analgesic dose increase to 2 days after the end of the dose increase for categorical endpoints, and will have the last assessment prior to the dose increase carried forward until 2 days after the end of such analgesic dose increase for continuous endpoints.

9.6 Efficacy Subgroup Analyses

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

- Age group (< 40 years, ≥ 40 – < 65 years, ≥ 65 years)
- Sex (male, female)
- Race (white, non-white)
- Smoking (current, ex or never)
- BMI (normal: < 25, overweight: ≥ 25 – < 30, obese: ≥ 30)
- Duration of HS (by median)
- Baseline body weight (by median)
- Baseline CRP level (by median)
- Baseline AN count category (5 – 10, 11+)
- Prior exposure to TNF antagonists (yes, no)
- Prior exposure to biologic agents that block IL-12/23/17 (yes, no)
- Baseline Hurley stage (I, II, III)

In addition, the primary efficacy endpoint may also be summarized by anti-drug antibody (ADA) status and/or by neutralizing antibody (NAb) status, once the corresponding ADA and NAb data are completed.

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 20 weeks (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: TEAEs are defined in the following two groups:

- **Any risankizumab:** 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).
- **Risankizumab 360 mg:** Any event with an onset date on or after the first dose date of risankizumab 360 mg and the minimum of (140 days after the last dose date of risankizumab 360 mg, 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 AE 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 adjudicated MACE
- Any adjudicated anaphylactic reaction
- Any serious infection
- Any tuberculosis

- Any malignant tumor
- Any malignant tumor excluding NMSC
- Any serious hypersensitivity
- Any treatment-emergent AE leading to death

All deaths will also be summarized:

- 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 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:

- **Any risankizumab:** 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 cutoff date during the Primary Analysis + 1 day]) minus the first risankizumab dose date, normalized by 365.25 and rounded to one decimal place.
- **Risankizumab 360 mg:** Sum of study drug of risankizumab exposure, defined as the minimum of (the last risankizumab 360 mg dose date + 140 days, and the end of study date + 1 day [and cutoff date during the Primary Analysis + 1 day]) minus the first risankizumab 360 mg 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[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 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 ≥ 2.0 × ULN
- ALT and/or AST > 3× ULN and TBL ≥ 2× 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 × ULN, or
- AST ≥ 3 × ULN, or
- Alkaline phosphatase ≥ 1.5 × ULN, or
- Total bilirubin ≥ 2 × ULN.

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

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

The number and percentage of subjects with treatment emergent hypersensitivity events or injection site reactions may also be summarized by anti-drug antibody (ADA) status

and/or by neutralizing antibody (NAb) status, once the corresponding ADA and NAb data are completed.

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

An external data monitoring committee (DMC) composed of persons independent of AbbVie and with relevant expertise in their field will review unblinded safety data from the ongoing study. The primary responsibility of the DMC will be to protect the safety of the subjects participating in this study.

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

Since there are no efficacy analyses for early stopping, no alpha adjustment is needed.

13.0 Overall Type-I Error Control

For each risankizumab dose, the statistical comparisons for the primary efficacy variable and the ranked secondary variables will be carried out in the hierarchical order between the corresponding risankizumab dose and placebo, under a 2-sided significance level of 0.025. This means that for each risankizumab dose, statistically significant results for the comparison in the higher rank (primary, then ranked secondary variables) are necessary to initiate the testing of the next comparison in the lower rank. Since a step-down procedure is used, each comparison will be tested at a significance level of 0.025 and overall alpha level of 0.025 will be preserved for each risankizumab dose.

Therefore, the family-wised type-I error rate is controlled under a 2-sided significance level of 0.05.

14.0 Version History

Table 5. SAP Version History Summary

Version	Date	Summary
1.0	7 Mar 2019	Original version
2.0	26 Jul 2019	Added more statistical analysis details.
3.0	08 Oct 2020	Removed section about RAR & IERC and updated the control for overall type-I error rate throughout the SAP, to align with the same updates after protocol version 3.0. Removed the stratification cap of TNF-IR population, to align with the same updates after protocol version 3.0. Clarified the definition of HiSCR and NRS30 throughout the SAP, to align with the updated wording after protocol version 3.0. Clarified the timing and scope of the Primary Analysis. Updated the details about how to handle subjects who do not continue into Period B from the ITT Population in Section 4.0, to align with the same updates after protocol version 3.0. Updated the definition of concomitant medication in Section 7.3, to align with the most updated risankizumab PSSAP. Clarified in Section 9.1 that the non-responder imputation (NRI) approach will be adjusted to NRI-C, Non-Responder Imputation incorporating multiple imputation to handle missing data due to COVID-19. Added more details about handling missing data in Section 9.2 and Appendix E, including how to handle missing data caused by COVID-19 and subjects who received prohibited antibiotics or discontinued study drug to lack of efficacy. Clarified shift tables, PCI, and ASI definitions for safety analysis in Section 10.0, Appendix B, Appendix C, and Appendix D, to align with the most updated risankizumab PSSAP. Added estimands and their attributes for the primary endpoint and ranked secondary endpoints at Section 2.1, Section 8.0, Section 9.3, and Section 9.4.

Version	Date	Summary
4.0	19 Jan 2021	Clarified that study drug exposure will be calculated up to the study completion date for subjects who pre-maturely discontinued from study participation during Period A in Section 6.0 and Section 10.2.4. Corrected an error in Section 7.1, as PGIC does not apply at Baseline. Updated the definition of concomitant medication in Section 7.3, to align with the most updated risankizumab PSSAP. Included any adjudicated anaphylactic reaction to the AE overview table in Section 10.2.2 and Appendix B, to align with the most updated risankizumab PSSAP. Updated the lab test shift table categories and liver functioning test categories in Section 10.3 and added the reference to CTCAE version 4.03, to align with the most updated risankizumab PSSAP.

15.0 References

1. Common Terminology Criteria for Adverse Events (CTCAE) v4.03 (2010). 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 Area of Safety Interest

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

Area of Safety Interest	Search Criteria		Include in AE Overview (Y/N)
MACE	Adjudicated terms will be identified as described in PSSAP Table 3 using CECAT and CETERM from the CE SDTM dataset.		Y
Extended MACE	Adjudicated terms will be identified as described in PSSAP Table 3 (for MACE +) using CECAT and CETERM from the CE SDTM dataset.		N
Serious Infections	Serious AEs in the Infections and Infestations SOC		Y
Tuberculosis	Active Tuberculosis CMQ (code 80000188)		Y
Opportunistic Infections	Opportunistic infection excluding tuberculosis and herpes zoster CMQ (code 80000189)		N
Fungal Infections	Fungal infections CMQ (code 80000063)		N
Herpes Zoster	Herpes Zoster CMQ (code 80000175)		Y
Malignant Tumours	Narrow	Malignant tumours (SMQ 20000194)	Y
Non-melanoma Skin Cancer (NMSC)	Broad	Skin malignant tumours (SMQ 20000204) excluding terms identified by the Melanoma CMQ (code 80000119)	N
Malignant Tumours excluding NMSC	'Malignant Tumours excluding NMSC' is identified by the 'Malignant Tumours' search excluding terms identified by the 'Non-melanoma skin cancer (NMSC)' search.		Y
Hypersensitivity	Narrow	Hypersensitivity (SMQ 20000214)	Y – serious events only
Adjudicated Anaphylactic Reaction*	Adjudicated terms will be identified using SDTM data (e.g., CE and PR domains).		Y
Hepatic Events	Broad	Hepatic failure, fibrosis and cirrhosis and other liver damage-related conditions (SMQ 20000013)	N
	Broad	Hepatitis, non-infectious (SMQ 20000010)	N
	Broad	Cholestasis and jaundice of hepatic origin (SMQ 20000009)	N
	Broad	Liver related investigations, signs and symptoms (SMQ 20000008)	N
	Narrow	Liver-related coagulation and bleeding disturbances (SMQ 20000015)	N

* 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**
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 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 to handle missing data due to COVID-19 (NRI-C)

2.1 Overall Description of the Method

For a dichotomized outcome variable with missing data, the NRI-C 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.

Of note, later visits of subjects receiving any antibiotics for HS-related infections or discontinued study drug due to lack of efficacy will be set as missing before imputation. As a result, these assessments will not contribute to the imputation and the subjects will be counted as non-responders for the analysis at later visits. When determining the HiSCR response status, a subject with missing Baseline value will be counted as a non-responder at any post-baseline visit unless a subject's abscess, inflammatory nodule, and draining fistula counts are all zero at a specific visit, then the subject will be a HiSCR responder at that visit. NRI-C will be implemented as follows.

2.2 Multiple Imputation (MI) and MAR Assumption

When a dichotomized variable is derived from continuous scales, for example, HiSCR (as at least a 50% reduction in the total abscess and inflammatory nodule (AN) count with no increase in abscess count and no increase in draining fistula count relative to Baseline), the multiple imputation will be applied to the original scales of abscess, inflammatory nodule and draining fistula counts, assuming multivariate normal distributions. Then the

dichotomized variable will be derived from the imputed values. For demonstration purposes, we use HiSCR as an example in this Appendix E, and provide sample code for the imputation of abscess count in Appendix E Section 3.0. The same imputation strategy will be applied to the inflammatory nodule count and the draining fistula count. HiSCR status will be determined accordingly.

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., abscess count 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., abscess count 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, prior exposure to anti-TNF category, Baseline worst Hurley Stage, Baseline abscess count, all

post-baseline visits of abscess count 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, prior exposure to anti-TNF category, Baseline worst Hurley Stage, Baseline abscess count, all post-baseline visits of abscess count 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

Same as the abscess count, we also impute to obtain complete datasets for the inflammatory nodule count and the draining fistula count. For each 'complete' dataset, the imputed post-baseline values will be rounded to the same precision as the observed data. HiSCR status 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-C approach, all missing due to reasons other than COVID-19 will be categorized as non-responders, including visits after a subject receives any antibiotics for HS-related infections or discontinued study drug due to lack of efficacy. When determining the HiSCR response status, a subject with missing Baseline value will be counted as a non-responder at any post-baseline visit unless a subject's abscess, inflammatory nodule, and draining fistula counts are all zero at a specific visit, then the subject will be a HiSCR responder at that visit.
- 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.

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 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:

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

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 $v = (K-1)(1+\frac{W}{B})^2$. 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 COUNT_2; SET COUNT;
  /*THE MCMC STATMENT BELOW ASSUMES MULTI-VARIATE NORMAL*/
  /*  USE ALL DATA IN THE IMPUTATION MODEL*/
  IF TRT01P="180mg" THEN TRT1=1; ELSE TRT1=0; /* CREATE DUMMY TREATMENT
VARIABLE FOR THE ANALYSIS STAGE*/
  IF TRT01P="360mg" THEN TRT2=1; ELSE TRT2=0; /* CREATE DUMMY TREATMENT
VARIABLE FOR THE ANALYSIS STAGE*/
  IF HURLEY=1 THEN HURLEY1 = 1 ELSE HURLEY1 = 0; /* CREATE DUMMY
VARIABLE FOR HURLEY STAGE 1*/
  IF HURLEY=2 THEN HURLEY2 = 1 ELSE HURLEY2 = 0; /* CREATE DUMMY
VARIABLE FOR HURLEY STAGE 2*/
RUN;

/*AUGMENTATION STEP -- TO HAVE 30 MONOTONE MISSING DATASETS*/
PROC MI DATA= COUNT_2 OUT= COUNT_MONO NIMPUTE=30 SEED= 21423 /*RANDOM
SEED PRE-DEFINED*/
  ROUND=. . . . . 1 1 1 1 1 1 1 /*VALUE ROUND TO INTEGER*/
  MIN=. . . . . 0 0 0 0 0 0 0  /*MINIMUM VALUE OF COUNT IS 0*/
  MAX=. . . . . .   /*MAXIMUM VALUE*/
MCMC IMPUTE=MONOTONE ;
/*NOTE: CATEGORICAL VARIABLES SUCH AS TRT1 TRT2 HURLEY1 HURLEY2 ARE
DUMMY, CREATED ABOVE*/
/*NOTE: ALL OTHER NON-DUMMIED COVARIATES MUST BE CONTINUOUS*/
VAR TRT1 TRT2 TNFN HURLEY1 HURLEY2 BASE WK1 WK2 WK4 WK8 WK12 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= COUNT_MONO OUT= COUNT_FULL NIMPUTE=1 SEED= 21931  /*RANDOM
SEED PRE-DEFINED*/
  ROUND=. . . 1 1 1 1 1 1 1 /*VALUE ROUND TO INTEGER*/
  MIN=. . . 0 0 0 0 0 0 0  /*MINIMUM VALUE OF COUNT IS 0*/
  MAX=. . . . . .   /*MAXIMUM VALUE*/
  MINMAXITER=1000;
  /*CLASS CATEGORICAL VARIABLES TRT1 AND REG2*/
  CLASS TRT01P TNFN HURLEY;
  VAR TRT01P TNFN HURLEY BASE WK1 WK2 WK4 WK8 WK12 WK16;
  MONOTONE REG (WK1 WK2 WK4 WK8 WK12 WK16);  /* IMPUTED SEQUENTIALLY,
FROM WK 1 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;


/*We will get the imputed "complete" datasets for abscess, inflammatory
nodule, and draining fistula counts.  For sample code below, we denote
the following variables: ABS16, INF16 and DRA16 as the corresponding
lesion count at Week 16; and AN16 as the sum of ABS16 and INF16.
Determine the HiSCR status at Week 16.  */
DATA ALL; SET COUNT_FULL;
    IF 0<= AN16 <=0.5*AN_BASE AND ABS16<=ABS_BASE AND DRA16 <= DRA_BASE
THEN HISCR_16=1;
    ELSE HISCR_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 HISCR_NRI_16: TRADITIONAL NRI DATA AT WEEK 16, WHICH COVERS
THE SPECIAL HANDLING SUCH AS THE BEFORE-AND-AFTER EXCEPTION IN THE
PARTICULAR STUDY PERIOD*/
  IF COVID19_MISS NE 'Y' THEN HISCR_16 = HISCR_NRI_16;
RUN;
PROC SORT DATA=ALLF; BY _IMPUTATION_ SUBJID; RUN;

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

/*KEY CODE: ANALYZING EACH 'COMPLETE' DATASET*/
/*********************************/
/*COMPARE TREATMENT PLACEBO AND 360mg ONLY*/
DATA ALL; SET ALL;
   WHERE TRT01P NE "180mg";
RUN;

/*********************************/
/*INDIVIDUAL-LEVEL DATA --> # OF RESPONDERS & # OF SUBJECTS, TO BE READ-
IN TO PROC STDRATE */
PROC FREQ DATA=ALL;
  BY _IMPUTATION_;
  TABLES TRT2*STRATA*HISCR_16/LIST NOCUM NOPRINT OUT=COUNT_TABLE;
  /*TRT2 = 1 as 360mg and 0 as placebo*/
RUN;
DATA COUNT_TABLE; SET COUNT_TABLE;
  DROP PERCENT;
RUN;
PROC TRANSPOSE DATA=COUNT_TABLE OUT=FREQ_TABLE PREFIX=RESP;
ID HISCR_16;
BY _IMPUTATION_ TRT2 STRATA;
VAR COUNT;
RUN;
DATA FREQ_TABLE1; SET FREQ_TABLE;
  CASE=RESP1;
  SIZE=SUM(RESP0, RESP1);
  KEEP _IMPUTATION_ TRT2 STRATA CASE SIZE;
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=TRT2 EVENT=CASE TOTAL=SIZE;
  STRATA STRATA / 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.