STATISTICAL ANALYSIS PLAN

Study: PS0008

Product: BIMEKIZUMAB

A PHASE 3, MULTICENTER, RANDOMIZED, DOUBLE-BLIND STUDY WITH AN ACTIVE-CONTROLLED INITIAL TREATMENT PERIOD FOLLOWED BY A DOSE-BLIND MAINTENANCE TREATMENT PERIOD TO EVALUATE THE EFFICACY AND SAFETY OF BIMEKIZUMAB IN ADULT SUBJECTS WITH MODERATE TO SEVERE CHRONIC PLAQUE PSORIASIS

SAP/Amendment Number	Date
Final SAP	06 June 2018
Amendment 1	17 September 2019
Amendment 2	31 October 2019

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ADA	adalimumab
ADR	adverse drug reaction
AE	adverse event
AH	abnormal high
AL	abnormal low
ALP	alkaline phosphatase
ALT	alanine aminotransferase
AMS	Active Medication Set
ANCOVA	analysis of covariance
AST	aspartate aminotransferase
ATC	Anatomical Therapeutic
BKZ	bimekizumab
BKZS	bimekizumab set
BLQ	below level of quantification
BMI	body mass index
BSA	body surface area
CDF	cumulative distribution function
CI	confidence interval
CMH	Cochran-Mantel-Haenszel
CSR	Clinical Study Report
CTCAE	Common Terminology Criteria for Adverse Events
CV	cardiovascular
DILI	drug-induced liver injury
DLQI	Dermatology Life Quality Index
DMC	Data Monitoring Committee
EAER	exposure adjusted event rate
EAIR	exposure adjusted incidence rate
ECG	electrocardiogram
eCRF	electronic Case Report Form
eC-SSRS	electronic Columbia Suicide Severity Rating Scale

ePRO	electronic Patient-Reported Outcome
EQ-5D-3L	Euro-Quality of Life 5-Dimensions, 3 levels
ES	Enrolled Set
ESS	Escape Subject Set
EudraCT	European Union Drug Regulating Authorities Clinical Trials
FAS	Full Analysis Set
GGT	gamma-glutamyltransferase
HLT	High Level Term
ICF	Informed Consent Form
IGA	Investigator's Global Assessment
IGRA	interferon-gamma release assay
IL	interleukin
IMP	investigational medicinal product
LOCF	last observation carried forward
LFT	Liver Function Test
MACE	Major cardiovascular events
MAR	missing at random
MCID	minimal clinically important difference
MCMC	Markov-Chain Monte Carlo
MCS	Mental Component Summary
MedDRA	Medical Dictionary for Regulatory Activities
MI	multiple imputation
mNAPSI	Modified Nail Psoriasis Severity Index Score
MNAR	missing not at random
MS	Maintenance Set
NRI	Non-responder imputation
OC	Observed case
PASE	Psoriatic Arthritis Screening and Evaluation
PASI	Psoriasis Area Severity Index
PCS	Physical Component Summary
PEOT	Premature End of Treatment
PF	physical functioning

PGA	Patient's Global Assessment
PGADA	Patient's Global Assessment of Disease Activity
PHQ-9	Patient Health Questionnaire 9
PK	pharmacokinetics
PK-PPS	Pharmacokinetics Per-Protocol Set
pp-IGA	palmoplantar Investigator's Global Assessment
PPS	Per-Protocol Set
PRO	patient-reported outcome
PS	Patient Safety
PsA	psoriatic arthritis
PSD	patient symptom diary
PSO	psoriasis
Q4W	every 4 weeks
Q8W	every 8 weeks
QOL	quality of life
RS	Randomized Set
SAE	serious adverse event
SAP	Statistical Analysis Plan
sc	subcutaneously
scalp IGA	scalp-specific IGA
SD	standard deviation
SF-36	Short Form 36-item Health Survey
SFU	Safety Follow-Up
SS	Safety Set
TB	tuberculosis
TEAE	treatment-emergent adverse event
TNF	tumor necrosis factor
ULN	upper limit of normal
VAS	visual analog scale
WBC	white blood cell
WHO-DD	World Health Organization Drug Dictionary

1 INTRODUCTION

This statistical analysis plan (SAP) defines the scope of statistical analyses and provides a detailed description of statistical methodology for the statistical analyses to support the final clinical study report (CSR). The SAP is based on the following study document: Protocol Amendment 2, 6 April 2018.

2 PROTOCOL SUMMARY

2.1 Study objectives

2.1.1 Primary objective

The primary objective of the study is to compare the efficacy of bimekizumab (BKZ) administered sc for 16 weeks versus adalimumab (ADA) in the treatment of subjects with moderate to severe chronic plaque psoriasis (PSO).

2.1.2 Secondary objectives

The secondary objectives of the study are to:

- Evaluate the efficacy of bimekizumab compared to adalimumab after 4, 16, and 24 weeks of treatment
- Evaluate the efficacy of bimekizumab compared to adalimumab at achieving complete clearance according to Psoriasis Area and Severity Index (PASI100) after 16 weeks and 24 weeks of treatment
- Assess the maintenance of efficacy of bimekizumab dosing Q4W versus Q8W at Week 56
- Assess treatment emergent adverse events (TEAEs), serious adverse events (SAEs), and TEAEs leading to withdrawal adjusted by duration of subject exposure to study treatment

2.1.3 Other objectives

The other objectives of the study are to demonstrate the effects of bimekizumab on aspects of the disease:

- Assess the maintenance of efficacy of bimekizumab dosing Q4W versus Q8W during the Maintenance Treatment Period
- Assess the efficacy of bimekizumab over time
- Assess the change of skin-related quality of life (QOL)
- Assess the change of general health-related QOL
- Assess the change in nail PSO over time in subjects with nail PSO at Baseline
- Assess the change in scalp PSO over time in subjects with scalp PSO at Baseline
- Assess the change in palmoplantar PSO over time in subjects with palmoplantar PSO at Baseline
- Assess the symptoms of psoriatic arthritis (PsA) as measured by the Psoriatic Arthritis Screening and Evaluation (PASE questionnaire)

- Assess the change of patient-reported joint symptoms in subjects with PsA at Baseline
- Assess the change in symptoms of PSO as reported by subjects using the Patient Symptom Diary (all items) through Week 24
- Assess depression
- Assess the PK of bimekizumab
- Assess the immunogenicity of bimekizumab
- Assess work productivity
- Assess the effect of bimekizumab on gene and protein expression, and explore the relationship between genomic, genetic, and proteomic biomarkers and disease biology, drug treatment and inflammatory and immune responses (from consenting subjects who agree to participate in the biomarker substudy)
- Assess the safety and tolerability of bimekizumab

2.2 Study variables

2.2.1 Efficacy variables

2.2.1.1 Primary efficacy variables

The co-primary efficacy variables are the PASI90 response (defined as a subject that achieves 90% reduction from Baseline in the PASI score) at Week 16 and the Investigator's Global Assessment (IGA) response (defined as Clear or Almost Clear with at least a 2-category improvement relative to Baseline) at Week 16.

2.2.1.2 Secondary efficacy variable(s)

The secondary efficacy variables are:

- PASI90 response at Week 24
- IGA response (Clear or Almost Clear with at least a 2-category improvement relative to Baseline) at Week 24
- PASI75 response at Week 4
- PASI100 response at Week 16 and 24
- PASI90 response at Week 56
- IGA response at Week 56

2.2.1.3 Other efficacy variables

Change from Baseline variables evaluated are relative to the Baseline (first dose) Visit. In addition, for subjects who start with adalimumab and switch treatment at the Week 24 Visit, change from Baseline variables during the bimekizumab treatment period will be evaluated relative to both the Baseline (first dose) Visit and the Week 24 Visit.

For simplicity, "change from Baseline" is used below for all such variables. Greater detail on the definition of Baseline for different summaries is provided in the Section 3.2. Unless otherwise stated, PASI responder rates will be calculated relative to the Baseline (first dose) Visit.

The other efficacy variables are:

- PASI50, PASI75, PASI90, and PASI100 response
- Time to PASI50, PASI75, PASI90, and PASI100 response during the Initial Treatment Period
- Absolute and percent change from Baseline in PASI score
- Percent of subjects with absolute PASI score ≤1, ≤2, ≤3 and ≤5
- IGA response (Clear with at least a 2-category improvement relative to Baseline)
- IGA response (Clear or Almost Clear with at least a 2-category improvement relative to Baseline)
- Shift from Baseline in IGA score
- Absolute and percent change from Baseline in the BSA affected by PSO
- Percent of subjects with absolute BSA=0%, ≤1%, ≤3% and ≤5%
- Absolute and percent change from Baseline in the product of IGA and BSA (IGAxBSA)
- Change from Baseline in the Dermatology Life Quality Index (DLQI)
- Percent of subjects achieving a DLQI total score of 0 or 1
- Percentage of subjects achieving a minimal clinically important difference (MCID) (improvement from Baseline of 4 or more) in the DLQI
- Change from Baseline in Patient's Global Assessment of Disease Activity (PGADA) for the arthritis visual analog scale (VAS) in subjects with PsA as defined by PASE score at Baseline
- Change from Baseline in Patient Global Assessment of PSO score
- Change from Baseline in the Patient Symptom Diary scores
- Patient Symptom Diary response rates
- Scalp-specific Investigator's Global Assessment (scalp IGA) response (Clear or Almost Clear with at least a 2-category improvement relative to Baseline) for subjects with scalp PSO at Baseline
- Change from Baseline in modified Nail Psoriasis Severity Index (mNAPSI) score for subjects with nail PSO at Baseline
- mNAPSI75, mNAPSI90 and mNAPSI100 response (defined as a subject that achieves at least a 75%, 90% or 100% respectively reduction from Baseline in the mNAPSI score)
- Palmoplantar Investigator's Global Assessment (pp-IGA) response (Clear or Almost Clear with at least a 2-category improvement relative to Baseline) for subjects with palmoplantar PSO at Baseline
- Change from Baseline in the PASE questionnaire scores (function score, symptom score, and total score)

- Shift from Baseline in PASE score suggestive of PsA (<47 versus ≥47)
- Change from Baseline in Short Form 36-item Health Survey (SF-36) Physical Component Summary (PCS) score, and Mental Component Summary (MCS) score, and individual domains
- Responses to Euro-Quality of Life 5-Dimensions, 3 levels (EQ-5D-3L), absolute and changes from Baseline in EQ-5D-3L VAS scores
- Change from Baseline in Work Productivity and Activity Impairment Questionnaire-specific health problem (WPAI-SHP) V2.0 adapted to PSO scores

2.2.2 Pharmacokinetic variables

The PK variable is the plasma concentration of bimekizumab.

2.2.3 Safety variables

2.2.3.1 Secondary safety variables

The secondary safety variables include the following treatment emergent adverse events (TEAEs) and serious adverse events (SAEs).

- TEAEs adjusted by duration of subject exposure to study treatment
- SAEs adjusted by duration of subject exposure to study treatment
- TEAEs leading to withdrawal adjusted by duration of subject exposure to study treatment

2.2.3.2 Other safety variables

Safety variables to be assessed are:

- Severity and frequency of AEs (including serious AEs)
- Change from Baseline in vital signs
- ECG results
- Change from Baseline in clinical laboratory values (chemistry, hematology)
- Change from baseline in Patient Health Questionnaire (PHQ)-9 scores

2.2.4 Pharmacogenomic variables

For individuals consenting to the genomics, genetics, and proteomics substudy, blood samples will be drawn for exploratory genetic/epigenetic, genomic, proteomic, and metabolomics analysis and for candidate exploratory biomarker analyses.

2.2.5 Immunological variables

The immunological variable is the anti-bimekizumab antibody level prior to and following investigational medicinal product (IMP).

2.3 Study design and conduct

2.3.1 Study description

PS0008 is a Phase 3, multicenter, randomized, double-blind, active-comparator-controlled, parallel-group study to evaluate the efficacy, safety, and pharmacokinetics (PK) of bimekizumab

(also known as UCB4940) compared with adalimumab in adult subjects with moderate to severe chronic plaque psoriasis (PSO).

The study population consists of adult subjects (≥18 years of age) with a diagnosis of moderate to severe chronic plaque PSO (Baseline Psoriasis Area and Severity Index [PASI] ≥12 and body surface area [BSA] affected by PSO ≥10% and Investigator's Global Assessment [IGA] score ≥3 [on a 5-point scale]) who are candidates for adalimumab or for systemic PSO therapy and/or phototherapy.

2.3.2 Study periods

This study will include 4 periods, a Screening Period (2 to 5 weeks), an Initial Treatment Period (16 weeks), a Maintenance Treatment Period (40 weeks), and a SFU Period (20 weeks after the last dose of IMP). After completion of the Maintenance Treatment Period, eligible subjects will be allowed to enroll in an open-label study. Subjects enrolling into the open-label study will not have the PS0008 SFU Visit.

2.3.2.1 Screening period

The Screening Period will last 2 weeks, but can be extended up to a total of 5 weeks. During this time, eligible subjects will be informed about the study and sign the Informed Consent Form (ICF), laboratory data (hematology, urine, and biochemistry tests) will be obtained, and the doses of medications used to treat PsA, will be verified as stable. The Screening Period will also enable washout of any medications not permitted for use during the study. Subjects who require prophylaxis for latent tuberculosis infection must be on treatment for at least 8 weeks prior to their first dose of IMP. These subjects may be rescreened once they have completed the first 8 weeks of prophylaxis treatment.

One rescreening may be allowed after consultation with the Medical Monitor.

2.3.2.2 Double-blind Initial Treatment Period

During the active-controlled 16-week Initial Treatment Period, approximately 450 subjects will be randomized 1:1:1 to receive the following blinded IMP regimens:

- Bimekizumab 320mg administered Q4W throughout the study (150 subjects)
- Bimekizumab 320mg administered Q4W until Week 16 (150 subjects)
- Adalimumab 80mg administered as an initial dose, followed by 40mg Q2W starting 1 week after the initial dose (ie, adalimumab will be administered according to the labeling recommendations) until Week 24 (150 subjects)

Investigational medicinal products will be administered in the clinic by sc injection at the time points specified in the schedule of study assessments.

The active comparator-controlled Initial Treatment Period will be used to demonstrate the efficacy of bimekizumab versus adalimumab. Efficacy of bimekizumab versus adalimumab will be measured at Week 16 because it was the time point for the primary endpoint in prior pivotal studies of adalimumab in PSO where efficacy was demonstrated.

Subjects withdrawing early from the study will undergo the PEOT Visit assessments and will enter the SFU Period. Subjects withdrawing early from the study will not be eligible for inclusion in the open-label study.

Subjects will be followed in the double-blinded fashion. Subjects may receive placebo injections at certain visits in order to blind the IMP.
Subjects will be classified as completing the initial treatment period if they complete the Week 16 Visit without early withdrawal from the study or if they start treatment in the maintenance treatment period. The start of the maintenance treatment period marks the end of the initial treatment period.

2.3.2.3 Dose-blind Maintenance Treatment Period

After the 16-week Initial Treatment Period, subjects will enter the 40-week Maintenance Treatment Period. Treatment during the Maintenance Treatment Period will start with the first study drug administration on or after Week 16 and subjects will return to the clinic Q4W or Q8W through Week 56. The IMP will be administered in the clinic by sc injection at the time points specified in the schedule of study assessments.

Treatment during the Maintenance Treatment Period will be based on initial treatment per the following:

- Subjects in the bimekizumab 320mg Q4W treatment arm will continue to receive bimekizumab 320mg Q4W.
- Subjects in the bimekizumab 320mg Q4W/Q8W treatment arm will receive bimekizumab Q8W from Week 16 through Week 52.
- Subjects in the adalimumab treatment arm will receive bimekizumab 320mg Q4W from Week 24 to Week 52.

Subjects may receive placebo injections at certain visits in order to blind the IMP.

At the end of the Maintenance Treatment Period, all subjects enrolling in the open-label study will, after signing a new ICF, undergo the Week 56 study assessments and then receive their first dose of bimekizumab in the open-label study. All subjects not enrolling in the open-label study will have the Week 56 study assessments and will enter the SFU Period.

2.3.2.4 Safety Follow-Up Period

All subjects not continuing in the open-label study, including those withdrawn from IMP, will have a SFU Visit 20 weeks after their last dose of IMP.

2.3.2.5 Premature End of Treatment

Subjects withdrawing early from the study will undergo the PEOT Visit assessments and will enter the SFU Period.

2.3.3 Study duration per subject

For each subject, the study will last a maximum of up to 77 weeks, as follows:

- Screening Period: 2 to 5 weeks
- Double-blind, active-controlled Initial Treatment Period: 16 weeks
- Maintenance Treatment Period: 40 weeks
- Safety Follow-Up Period: a SFU Visit is planned 20 weeks after the last dose of IMP (for subjects not enrolling in the open-label study)

After the 40-week Maintenance Treatment Period, subjects will be allowed to enroll in an open-label study, in which case subjects will undergo the Week 56 study assessments before receiving their first open-label IMP dose. The SFU Visit will not be required for subjects who enroll in the open-label study.

The end of the study is defined as the date of the last visit of the last subject in the study.

2.3.4 Planned number of subjects and site(s)

Approximately 600 subjects will be screened in order to have 450 subjects randomized in the study. There will be approximately 150 subjects per treatment arm. The planned number of study sites is approximately 100. Every eligible subject who signs an ICF will be expected to be randomized.

2.3.5 Anticipated regions and countries

The regions planned for study conduct are Western Europe, Central/Eastern Europe, North America and Asia/Australia, with possible extension to other regions and countries.

2.4 Determination of sample size

A total of 450 subjects will be randomly assigned in a 1:1:1 ratio to the following treatment groups:

- Bimekizumab 320mg Q4W throughout the study (150 subjects)
- Bimekizumab 320mg Q4W/Q8W (ie, bimekizumab 320mg Q4W until Week 16, then bimekizumab 320mg Q8W from Week 16 through Week 52 (150 subjects)
- Adalimumab 80mg administered as an initial dose, followed by 40mg Q2W starting 1 week after the initial dose (ie, adalimumab will be administered according to the labeling recommendations) until Week 24, then bimekizumab 320mg Q4W from Week 24 to Week 52 (150 subjects)

When comparing bimekizumab to adalimumab, the assumed responder rates for PASI90 and IGA at Week 16 for the bimekizumab group are 75% and 85%, respectively. These estimates are based on the results of the Phase 2b PS0010 study. The PASI90 responder rates at Week 16 for adalimumab from the VOYAGE-1 and VOYAGE-2 studies were 50% and 47%, respectively (Blauvelt et al, 2017; Reich et al, 2017). Additionally, the IGA responder rates at Week 16 for adalimumab in VOYAGE-1 and VOYAGE-2 were 66% and 68%, respectively. For the purposes of these sample size calculations at Week 16, we assume a PASI90 responder rate and IGA responder rate for adalimumab of 50% and 68%, respectively.

The testing procedure described in protocol Section 14.3 indicates that the first test will be to demonstrate noninferiority to adalimumab for the co-primary variables of PASI90 and IGA at Week 16. If non-inferiority for both tests is demonstrated, then superiority to adalimumab will be tested.

Because the superiority evaluation is the more stringent test, sample size calculations are based on the testing of superiority to adalimumab for PASI90 and IGA at Week 16. Given these assumptions and a sample size of 150 subjects in the adalimumab arm and 300 subjects in the pooled bimekizumab arms, the power to detect a statistically significant difference between bimekizumab and adalimumab is >99% for PASI90 response and 97% for IGA response. This

assumes a 2-sided significance level of 0.05. Because both co-primary endpoints are highly powered independently, and because PASI and IGA response tend to be highly correlated, the overall power for achieving superiority on both co-primary endpoints is not calculated here.

Note that the non-inferiority testing procedure will be based on a 1-sided significance level of 0.025 and a non-inferiority margin of 10%. The power to demonstrate non-inferiority to adalimumab using the other assumptions described above is >99% for both PASI90 and IGA responses.

3 DATA ANALYSIS CONSIDERATIONS

3.1 General presentation of summaries and analyses

Statistical analysis and generation of tables, figures, subject data listings, and statistical output will be performed using SAS Version 9.3 or higher. All tables and listings will use Courier New font size 9.

Descriptive statistics will be displayed to provide an overview of the study results. For continuous variables, descriptive statistics will include number of subjects with available measurements (n), mean, standard deviation (SD), median, minimum, and maximum.

For categorical variables, the number and percentage of subjects in each category will be presented. Unless otherwise noted, the denominator for percentages should be based on the number of subjects included in the respective analysis set. Subjects with missing data can generally be accounted for using either of the following approaches:

- For summaries of demographics and Baseline characteristics: summarize percentages based on all subjects in the analysis set and include a "Missing" category (corresponding to subjects with missing data for the variable being summarized) as the last row in the list of categories being summarized.
- For summaries of efficacy and safety endpoints, unless otherwise specified: summarize percentages based only on those subjects with observed data for the variable being summarized. As the denominator may be different from the number of subjects in the analysis set being considered, the denominator should be displayed in the table. The general format for displaying this will be "n/Nsub (%)."

Percentages will be presented to 1 decimal place. If the percentage is 100%, do not present a decimal. If the percentage is 0, do not present the percentage. Typically, the % sign should be presented in the column header, but not with each individual value.

For bimekizumab PK concentrations, summary statistics will include geometric mean, geometric coefficient of variation (CV), 95% confidence intervals for geometric mean, arithmetic mean, SD, median, minimum, and maximum. All summaries of PK variables will be based on the observed values. No imputation will be used.

Decimal places for descriptive statistics will always apply the following rules:

- "n" will be an integer
- Mean, SD, and median will use one additional decimal place compared to the original data
- CV [%] will be presented with one decimal place

- Minimum and maximum will have the same number of decimal places as the original value.

Derived variables in general will display the mean, SD and median to 1 more decimal place than the variables used in the derivation. If the number of decimal places reported in the raw data is varied then use either the maximum raw number of reported decimal places or 3, whichever is the lowest, as a guide for the descriptive statistics.

Statistical tests of efficacy variables will be presented as two-sided p-values rounded to three decimal places. P-values less than 0.001 will be presented as "<0.001" and p-values greater than 0.999 will be presented as ">0.999." Statistical comparisons will be two-sided and will be performed at the 0.05 level of significance.

A complete set of data listings containing all documented data as well as calculated data (e.g. change from Baseline) will be generated.

Per protocol, visit windows of ±3 days from the first dose to Week 24 and ±7 days from Week 28 to Week 56 are permissible. For the SFU Visit, visit window is ±7 days from final dose. All by-visit summaries will contain nominal (i.e. scheduled) visits only. Unscheduled visits will not be mapped to scheduled visits except for some assessments that may occur within a 3-day time window of a scheduled visit. In that case, the assessment will be mapped to the corresponding scheduled visit and will be used for the analysis. This will only occur for some vendor data.

For missing PASI and IGA assessments at Week 16, available data from all visits (unscheduled or scheduled but without a planned assessment) within the 3 days window will be included.

3.2 Definition of Baseline values

A Baseline value for a subject is defined as the latest measurement for that subject up to and including the day of administration of first study medication, unless otherwise stated. If a Baseline assessment is taken on the same day as first administration of study medication, it is eligible to be used as the Baseline value, even in the case that the time of the assessment is recorded as taking place after the time of first study medication administration. This is considered acceptable as this measurement is still the best representation of the Baseline value of the given assessment since it is highly unlikely that the study medication could have an impact on any measurement in such a short period of time. However, such cases should be rare as study center personnel are instructed to do all assessments at the Baseline visit prior to administering study medication. One exception to this rule is plasma concentration. If Baseline plasma concentration is measured at a time after the first administration of study medication, then it should not be eligible to be considered as a Baseline plasma concentration. Such cases should be discussed with the quantitative clinical pharmacologist. If a Baseline measurement is missing or not collected, and a Screening value is available, the Screening value will be utilized as Baseline instead.

For subjects who start with adalimumab and switch treatment at the Week 24 Visit, an additional bimekizumab Baseline defined as the measurement at Week 24 will be used. If a Week 24 Visit measurement is missing, baseline will be the latest measurement prior to the first dose of bimekizumab. The variables listed below will also be summarized based on bimekizumab Baseline:

- PASI score

- DLQI
- Laboratory data (only in shift tables)

Note that for any laboratory value that occurs on the day of treatment switch, that lab value will be attributed and summarized for the treatment they were on previously. For subjects that switch from Adalimumab to Bimekizumab, the Baseline value is the laboratory value of the day of treatment switch.

Baseline values for component scores should be computed using components from the same visit where the relevant measurements were recorded prior to dosing. For example, if the Screening visit has all of the components, but the Baseline visit is missing one or more components, the Baseline value for the component score should be calculated using the Screening visit values.

3.3 Mapping of data from early withdrawal visits

If the early withdrawal visit occurs at a scheduled visit as outlined in the schedule of assessments, then no mapping is necessary, and any early withdrawal assessments should correspond to that scheduled visit. Premature study withdrawal visit assessments that occur on a date after a scheduled visit will be assigned to the next scheduled site visit per the protocol following the last visit where assessments were available. This approach means that there is a chance that data will be mapped to a visit where a given assessment was not actually collected per the protocol schedule of assessments. Such data would not be summarized in by-visit tables (though it would be available in the listings).

The only exception to the above rule is for anti-bimekizumab antibody assessments, in which all early withdrawal visit assessments will be assigned to the next scheduled visit at which anti-bimekizumab antibody are assessed. All by-visit summaries will contain nominal visits only. Unscheduled visits will not be mapped to scheduled visits. Note that based on the early withdrawal mapping conventions described above, a mapped early withdrawal visit is considered as observed at that visit and should be summarized as such in the tables.

3.4 Protocol deviations

Important protocol deviations are deviations from the protocol which potentially could have a meaningful impact on study conduct or on the primary efficacy, key safety, or PK outcomes for an individual subject. The criteria for identifying important protocol deviations will be defined within the appropriate protocol-specific document at study start. Important protocol deviations will be reviewed as part of the ongoing data cleaning process. Important protocol deviations including those that lead to exclusion from the analysis sets will be identified and documented prior to unblinding.

3.5 Analysis sets

For all analysis sets, each unique subject is only counted once, even if the same subject has more than one unique subject identification.

3.5.1 Enrolled Set

The Enrolled Set (ES) will consist of all subjects who have given informed consent.

3.5.2 Randomized Set

The Randomized Set (RS) will consist of all randomized subjects.

3.5.3 Safety Set

The Safety Set (SS) will consist of all subjects that received at least 1 dose of the IMP.

3.5.4 Full Analysis Set

The Full Analysis Set (FAS) will consist of all randomized subjects that receive at least 1 dose of the IMP and have a valid measurement of each of the co-primary efficacy variables at Baseline.

3.5.5 Bimekizumab Set

The Bimekizumab Set (BKZ Set) will consist of all subjects who have received at least 1 dose of bimekizumab in this study.

3.5.6 Bimekizumab Week 24 Set

The BKZ Week 24 Set will consist of all subjects who have received at least 1 dose of BKZ on or after Week 24.

3.5.7 Maintenance Set

The Maintenance Set (MS) will consist of all subjects who have received at least 1 dose of active IMP (bimekizumab or adalimumab) in the Maintenance Treatment Period (at Week 16 or later).

3.5.8 Per Protocol Set

The Per-Protocol Set (PPS) will consist of all subjects in the RS who had no important protocol deviations affecting the primary efficacy variables. Important protocol deviations will be pre-defined and subjects with important protocol deviations will be evaluated during ongoing data cleaning meetings prior to unblinding of the data.

3.5.9 Pharmacokinetics Per-Protocol Set

The Pharmacokinetics Per-Protocol Set (PK-PPS) will consist of all randomized subjects who received at least 1 dose of the IMP and provided at least 1 quantifiable plasma concentration post-dose without important protocol deviations that would affect the concentration.

3.6 Treatment assignment and treatment groups

It is expected that subjects receive treatment as randomized and hence safety analyses will be based on the SS, as randomized. However, if after unblinding it is determined that subjects randomized to a treatment arm had never received such treatment, then for safety analyses, these subjects will be reallocated to the appropriate received treatment groups, unless otherwise specified.

Efficacy analyses should be performed according to randomized treatment and not actual treatment received.

Summaries based on SS or RS will be provided by “BKZ 320mg Q4W”, “BKZ 320mg Q4W/Q8W”, “BKZ Total” and “ADA” treatment groups. The ADA group will only be presented from Baseline to Week 24 in the by-visit summaries.

Summaries based on BKZ Week 24 set will be provided by "BKZ 320mg Q4W", "BKZ 320mg Q4W/Q8W" and "ADA/BKZ 320mg Q4W" treatment groups for disposition, baseline assessments and efficacy. The ADA/BKZ 320mg Q4W group will be presented from Week 28 to Week 56 in the by-visit summaries for the PASI scores and the DLQI only.

For safety, summaries on the BKZ set will be provided by "BKZ 320mg Q4W", "BKZ 320mg Q8W" and "BKZ Total". BKZ Q4W will include data from subjects in the BKZ Q4W/Q8W arm (up to Week 16) and the ADA/BKZ Q4W arm (after Week 24).

Safety summaries on the BKZ Week 24 set will be provided for subjects switching from Adalimumab to Bimekizumab.

Summaries based on the Maintenance set will be provided by "BZK 320mg Q4W", "BKZ 320mg Q4W/Q8W" and "BKZ Total".

3.7 Center pooling strategy

Centers will be pooled into regions for analysis purposes. Geographic regions will be categorized as North America, Western Europe, Central/Eastern Europe, and Asia/Australia. Below is a table of geographic regions with corresponding countries.

Table 3–1: Geographic regions

Region	Countries
North America	Canada, United States
Western Europe	Germany
Central/Eastern Europe	Czech Republic, Hungary, Poland, Russian Federation
Asia/Australia	Australia, Republic of Korea, Taiwan

The following center pooling algorithm will be used for each geographic region:

- If a center has 15 or more subjects, then no pooling will be done for that center.
- Centers with fewer than 15 subjects will be ordered from largest to smallest with pooling proceeding in the following manner:
  - Two or more centers will be combined until the cumulative subject total is at least 15.
  - Once a pooled center has at least 15 subjects, the process will continue in an iterative fashion for the subsequent centers in the ordered list, where a new pooled center begins each time at least 15 subjects has been reached in the previous pool.
  - If this iterative process reaches the end of the ordered list of centers where the final pooled center has fewer than 15 subjects, then the subjects from the centers in that pool will be combined with the pooled center formed in the previous iteration.

This procedure is only to be performed within a geographic region –there will be no pooling of centers across regions.

In the event that the percentage of randomized subjects is less than 10% in either of the Asia/Australia or Western Europe regions, then the two regions will be combined as a

geographic region stratum for efficacy modeling, so that there are no modeling convergence issues across efficacy variables.

3.8 Coding dictionaries

All medications other than study drug will be classified by WHO Anatomical Therapeutic Chemical (ATC) Classification, presenting Anatomical Main Group (ATC Level 1), Pharmacological Subgroup (ATC level 3), and preferred term, using version SEP/2015 of the World Health Organization Drug Dictionary (WHO-DD), according to UCB standard operating procedures.

All AEs will be classified by primary system organ class (SOC), high level term (HLT) and preferred term (PT) using version 19.0 of MedDRA® according to UCB standard operating procedures.

Previous and ongoing medical history will be classified by version 19.0 of MedDRA® SOC and PT.

3.9 Relative day

The relative day will be included in different listings and will be calculated as follows:

- If the start (stop) date occurred on or after the first dose, but prior to the double-blind drug stop date, relative day is calculated as start (stop) date minus first dose date + 1
- If the start (stop) date occurred after the last dose of double-blind drug, the relative day to the most recent dose is calculated as start (stop) date minus most recent dose date. The relative day in this situation should be preceded by a '+'
- If the start (stop) date occurred before the first dose, the relative day is calculated as start (stop) date minus first dose date. The relative day in this situation should be preceded by a '-'.

For AEs, relative days for start and stop dates will be calculated as the number of days since the first injection of the medication. For non-treatment emergent AEs, relative day of onset will be negative if the event started and stopped before the first dose. Relative day will only be computed for fully completed dates and will be missing for partial dates.

Relative day will be calculated from first dose of IMP for all treatment groups, and additionally from first dose of bimekizumab for the ADA/BKZ 320mg Q4W arm.

3.10 Changes to protocol-defined analyses

For the following endpoints, the definition of response has been clarified:

- Scalp IGA response (clear or almost clear with at least a 2-category improvement from Baseline) for subjects with scalp PSO at Baseline
- Palmoplantar Investigator's Global Assessment (pp-IGA) response (clear or almost clear with at least a 2-category improvement from Baseline) for subjects with palmoplantar PSO at Baseline

The following endpoints are not listed in the protocol, but have been added to the SAP in order to achieve consistency with other studies from the program:

- mNAPSI75/90/100 response (defined as a subject that achieves at least a 75%, 90% and 100% reduction from Baseline in the mNAPSI score)
- Percent of subjects with absolute PASI score ≤1, ≤2, ≤3, and ≤5
- Percent of subjects with absolute BSA=0%, ≤1%, ≤3%, and ≤5%
- The sensitivity analysis for the primary endpoint using MI-MCMC/reference based methods detailed in the protocol has been removed from the analysis plan because it could generate bias due to the reference arm being the active comparator.

The following additional changes were made:

- The calculation of nominal p-values has been added for selected efficacy variables
- Modifications have been made to the text regarding the multiple imputation algorithm for clarity
- Change from Baseline in PHQ-9 score is listed an efficacy variable in the protocol. However, it is considered as a safety variable in this SAP.
- Selected TEAE tables were added to include calculation of risk differences.

Urinalysis summaries were removed from the plan.

4 STATISTICAL/ANALYTICAL ISSUES

4.1 Adjustments for covariates

The primary efficacy analyses and selected secondary analyses will be adjusted for the following covariates:

- Prior biologic exposure
- Region

4.2 Handling of dropouts or missing data

4.2.1 Handling of missing data for efficacy variables

Based on previous studies of biologics in subjects with moderate to severe chronic plaque psoriasis, it is expected that the number of subjects who discontinue prior to Week 16 will be low. For the small percentage of subjects for whom primary endpoint data are unavailable at Week 16, this lack of data is suggestive of an ineffective study treatment, thereby supporting the imputation of non-response. Achieving the clinical response and making it through 16 weeks of study treatment are both critical components of the primary outcome. Therefore, non-responder imputation (NRI) will be used as the method for handling missing data in the primary analysis for binary efficacy variables.

If a subject discontinued study treatment without terminating from the study and still continues with scheduled assessments, all efficacy data after last treatment date + 35 days will be treated as missing and subject to imputation as applicable.

4.2.1.1 Handling missing data for the primary efficacy variable

Missing data for co-primary efficacy variables will be imputed using NRI as primary method. That is, subjects with missing data at Week 16 or who discontinue treatment prior to Week 16 will be counted as non-responders for the analysis.

In addition, sensitivity analyses using multiple imputation (Markov-Chain Monte Carlo method) (MI-MCMC)/monotone regression, last observation carried forward (LOCF) and observed case (OC) method will be performed which will assess the impact of different methods of handling missing data. These methods are described in Section 8.1.3.1 to Section 8.1.3.2.

4.2.1.2 Handling missing data for the secondary efficacy variables

For secondary binary efficacy variables, missing data will be imputed using NRI as primary method. MI-MCMC/monotone regression and OC methods will be performed as sensitivity analysis.

For secondary continuous efficacy variables, MI-MCMC/monotone regression is the primary method for imputing missing data. If the imputation model cannot converge, last observation carried forward (LOCF) will be used. OC method will be performed as sensitivity analysis.

4.2.1.3 Handling missing data for the other efficacy variables

For other binary efficacy variables, missing data will be imputed using NRI as primary method.

For other continuous efficacy variables, MI-MCMC/monotone regression method will be used to impute missing data as primary method. If the imputation model cannot converge, last observation carried forward (LOCF) will be used.

For other ordinal variables (EQ-5D-3L and Patient Global Assessment of PSO), OC method will be applied as the primary analysis method. No imputation is applied.

For those other efficacy variables that were included in the sequence testing procedure, OC method will also be applied as sensitivity analysis.

4.2.1.4 Handling missing data for subgroup analyses

For subgroup analyses specified in Section 4.8, NRI will be used for responder variables. Only descriptive statistics will be provided.

4.2.1.5 Missing Data Overview and Summary

In summary, the approaches listed below will be used in this study for handling missing data for efficacy variables as appropriate:

- Non-responder imputation (NRI): Subjects who have missing data at the time point of interest are treated as though they did not respond to the treatment.
- Multiple Imputation (MI) – Markov Chain Monte Carlo (MCMC) / Monotone Regression: Using multiple imputation methodology, intermittent missing data are imputed based on the MCMC method, and monotone missing data are imputed using monotone regression. Note: for all analyses and summaries using MI, subjects with no baseline value will be excluded.
- Last observation carried forward (LOCF): Post-Baseline missing data are imputed by carrying forward the last available observation (including Baseline).

- Observed case (OC): Missing data are not imputed. Only subjects with available data who have not discontinued study treatment at the given time point are considered.

The following table depicts which missing data handling approaches should be used based on variable priority (primary, secondary, other) and variable type (responder, continuous, ordinal).

Table 4–1: Missing data handling approach by variable priority and type

Variable Priority	Variable Type	NRI	MI (MCMC/ Monotone Regression)	MI (MCMC/ Reference-based)	LOCF	OC
Primary	Responder	P	S[a]	n/a	S	S
Secondary	Responder	P	S[a]			S
	Continuous		P		B	S
Other	Responder	P				S[b, c]
	Continuous[e]		P		B	
	Ordinal					Pd

P=Primary method, S=Sensitivity method, B=Backup method, NRI=Non-responder imputation, MI=Multiple imputation, MCMC=Markov Chain Monte Carlo, LOCF=Last observation carried forward, OC=Observed Case
Note: Backup method is only applicable when the primary method is unable to converge due to challenges with the imputation model.
[a] Imputation method is applied on continuous data, and responder variable is derived from the continuous variable based on complete data set.
[b] Only applies to by-visit summaries of variables that are in the multiplicity-controlled testing procedure.
[c] Includes IGA response, PASI75, PASI90 and PASI100.
[d] Includes Patient Global Assessment of PSO, IGA score and EQ-5D-3L responses.
[e] For PASE, OC is the primary analysis method.

4.2.1.6 Missing data algorithms

MI – MCMC / Monotone Regression

Investigators will be given discretion to discontinue study treatment for subjects not responding in order to provide them with alternative therapy to treat the condition. Therefore, in many cases, missing efficacy data due to study treatment discontinuation should be dependent on the observed efficacy scores, but independent of unobserved data. This would be consistent with a missing at random (MAR) pattern of missingness. To investigate the efficacy results under the assumption of data being MAR, a multiple imputation method will be applied as follows:

1. Create a data set, one for each treatment group (note that a separate imputation procedure must be invoked in SAS for each treatment group as the seed cannot be set for by groups beyond the first when using a by statement), of subjects with observed values and those needing estimation by multiple imputation. The intermittent missing PASI/IGA values in each data set (ie, missing values for a given subject that has available data before and after the missing time point) will be filled in using the MCMC method, with a total of 100 sets of imputations being performed. The seed used for these imputations will be 852 (note that all other multiple imputation procedures described in this SAP related to

MCMC / Monotone Regression analyses will use this same seed as well). For monotone missing data (ie, where all subject data is missing after a given time point), monotone regression then will be used to impute missing data. This will be based on the 100 sets of imputations already created using the MCMC method such that there will be 100 imputations in total. In both cases, biologic exposure, geographic region and PASI/IGA values at Baseline and at each post-Baseline visit (in chronological order, see notes below about visits to include for different analysis sets) will be included in the imputation model. Note that PASI scores at earlier visits will also be used as predictors for the model of PASI at later visits. The resulting data sets for each treatment arm will be combined into one complete data set based on each of the 100 imputations.

Note: The imputation model will only allow continuous variables. Therefore, prior biologic exposure and region will be re-coded as indicator variables. For prior biologic exposure, this will simply be 0 for biologic-naïve subjects and 1 for biologic-exposed subjects. For region, which has 4 levels, one indicator variable will be defined as 0 for regions other than North America and 1 for North America. Two more indicator variables will be defined similarly replacing North America with Central/Eastern Europe, and Western Europe respectively. An indicator variable for Asia/Australia is not needed as the fourth region will be adequately represented by the other region indicator variables all being 0. In the event that the number of randomized subjects is less than 10% in either of the Asia/Australia or Western Europe regions, then the two regions will be combined as a geographic region stratum for efficacy modeling, so that there are no modeling convergence issues across efficacy variables. In order to achieve model convergence, prior biologic exposure may be dropped from the model. If convergence is still not obtained, then region may also be dropped from the model.

2. For each complete imputed data set, the dichotomous responder variable (PASI90/IGA 0 or 1) based on the PASI/IGA scores will be computed. Each complete imputed data set will then be analyzed based on the stratified CMH test.

   Note: For derivation of PASI90 response, the PASI value at Week 16 in the imputed data sets will be compared directly to the observed Baseline PASI value to determine whether or not a reduction of at least 90% was achieved. If values outside of the pre-defined range of values for PASI (0-72) are imputed, they will be cut off as appropriate after the multiple imputation procedure but before deriving the responder variable. For example, an imputed PASI value of -0.5 would be changed to 0 before deriving the PASI90 responder variable.

   Note: Standard rounding rules will be applied to the imputed IGA values in order to derive the binary IGA 0/1 responder variable. In addition, if values outside of the pre-defined range of values for IGA (0-4) are imputed, they will be cut off as appropriate after the multiple imputation procedure but before deriving the responder variable. For example, if a subject has an IGA score imputed as 1.4 (and assuming a Baseline IGA score of 3), this imputed value would be rounded down to 1, and the minimum change

from Baseline of 2 would have been met. Therefore, this subject would be considered an IGA 0/1 responder.

Additional ranges of values for other variables are defined in Table 4–2.

Table 4–2: Imputation allowable ranges by variable

Variable	Minimum value	Maximum value	Integer values only
PASI	0	72	No
IGA	0	4	Yes
PSD item	0	10	No
Scalp IGA	0	4	Yes
mNAPSI	0	130	No
BSA	0	100	Yes
IGAxBSA	0	400	Yes
DLQI	0	30	Yes
PGADA	0	100	Yes
SF-36	0	100	No
EQ-5D-3L VAS	0	100	Yes
WPAI dimension scores	0	100	No for variables “Percent work time missed due to problem” and “Percent overall work impairment due to problem”. Yes for variables “Percent impairment while working due to problem” and “Percent activity impairment due to problem” These two variables can only take values that are multiples of 10”
PHQ-9	0	27	Yes

3. The Week 16 results from the specified statistical analysis (e.g. stratified CMH, logistic regression) of each of the 100 imputed data sets will be combined for overall inference using Rubin’s rules, which account for the uncertainty associated with the imputed values (Rubin, 1987). This will be done using SAS PROC MIANALYZE.

Note that this procedure indicates that the imputation model will be applied for each treatment group separately in PROC MI, which will be the default method. However, in the event there are computational challenges with the imputation model (eg, due to a standard deviation of 0 for responses of a given imputation), it is acceptable to modify the imputation model to include treatment as a variable in the model rather than running a separate model for each treatment group. It should be noted that doing so assumes that treatment does not interact with any of the other variables in the imputation model.

Some further details about imputation for different analysis sets are provided below:

- Randomized Set: When programming multiple imputation based on the RS, PROC MI will be used with a separate data set for each of the 3 randomized treatment groups (Bimekizumab 320mg Q4W, Bimekizumab 320mg Q8W, Adalimumab). For Bimekizumab 320mg Q4W and Bimekizumab 320mg Q8W, include all scheduled assessment visits from Baseline to Week 56. For the Adalimumab group, include only scheduled assessments from Baseline to Week 24. NOTE: The combined Bimekizumab 320mg group will not be run through PROC MI. The combined group will be obtained by using the imputed values of the individual Q4W and Q8W groups
- Bimekizumab Week 24 Set: When programming multiple imputation based on the BKZ Week 24 Set, PROC MI will be used with a separate data set for each of the 3 randomized treatment groups (Bimekizumab 320mg Q4W, Bimekizumab 320mg Q8W, Adalimumab) including all scheduled assessment visits from Baseline to Week 56. NOTE: The combined Bimekizumab 320mg group will not be run through PROC MI. These tables will only present the 3 treatment sequences (no BKZ Total needed).

Some key points to consider relative to the calculation of the odds ratios and corresponding confidence intervals are noted below:

If stratified CMH or logistic regression are used, the estimates of the odds ratios from the logistic regression model in step 3 follow a log-normal distribution, and a log transformation is needed to normalize these estimates since the procedures for combining results from multiple imputed datasets are based on the assumption that the statistics estimated from each imputed dataset are normally distributed. Therefore, the log of the odds ratio estimates from the logistic regression model are used when combining into a single inference (the use of PROC MIANALYZE in step 3). Appropriate transformations to the standard errors and p-values should also be made in order to get the correct confidence intervals. For the CMH test using the p-value for the general association the Wilson-Hilferty transformation should be used (Ratitch, 2013).

Missing data for continuous primary or secondary efficacy variables will be imputed using MI as appropriate. The MI procedure for continuous variables will be similar to that described above with the following differences:

1. The absolute value of the given variable will be imputed. Once imputation has been performed across the 100 iterations specified, any values outside of the range of the given variable will be truncated accordingly.
2. The change from Baseline values will be computed based on the complete data sets.
3. The analysis model will be based on ANCOVA (see above) as opposed to the CMH test.

For other continuous efficacy variables, MI will be used to impute missing data when possible. If the imputation model cannot converge, LOCF will be used. The MI procedure will also be similar to that described above with the following differences: 1) No dichotomization will be necessary; 2) instead of using the stratified CMH test or logistic regression, the imputed data sets will be combined and simple means and standard errors will be calculated using Rubin's rules (via SAS PROC MIANALYZE). For calculation of other descriptive statistics such as the median, min and max, Rubin's rules do not apply. Multiple imputation estimates will be computed by simply averaging the estimates from the multiple repetitions of the imputation algorithm. This is the same approach that will be used when summarizing continuous secondary efficacy variables by subgroup.

4.2.2 Handling missing data for safety variables

For analyses of AEs and concomitant medication usage, a complete date must be established in order to correctly identify the AE or medication as occurring during treatment or not. For purposes of imputing missing components of partially-reported start and stop dates for AEs and for medication use, the algorithms listed below will be followed. Start and stop dates of AEs or concomitant medication will be displayed as reported in the subject data listings (i.e., no imputed values will be displayed in data listings).

Partial AE and concomitant medication start dates will be imputed as follows:

Imputation of Partial Start Dates

- If only the month and year are specified and the month and year of first dose is not the same as the month and year of the start date, then use the 1st of the month
- If only the month and year are specified and the month and year of first dose is the same as the month and year of the start date, then use the date of first dose
- If only the year is specified, and the year of first dose is not the same as the year of the start date, then use the 1st of January of the year of the start date
- If only the year is specified, and the year of first dose is the same as the year of the start date, then use the date of first dose
- If the start date is completely unknown and the stop date is unknown or not prior to the date of first dose, then use the date of first dose.

Imputation of Partial Stop Dates

- If only the month and year are specified, then use the last day of the month
- If only the year is specified, then use December 31st of that year
- If the stop date is completely unknown, do not impute the stop date.

In the event of ambiguity or incomplete data which makes it impossible to determine whether a medication was concomitant or an adverse event was treatment emergent, the medication will be considered as concomitant or the adverse event will be considered treatment emergent.

4.3 Interim analyses and data monitoring

After the final Week 56 visit, an interim analysis will be performed and a corresponding interim clinical study report (CSR) may be written. A final analysis and updated final CSR will be prepared once all data (through the safety follow-up (SFU) visit) have been collected.

In addition, an independent Data Monitoring Committee (DMC) will periodically review and monitor the safety data from this study and advise UCB.

Cardiovascular and Neuropsychiatric Adjudication Committees will also periodically review and monitor safety data from this study and advise UCB. Details will be provided in the DMC Charter and in the Adjudication Committee Charters.

Further details related to the DMC will be outlined in a separate analysis plan.

4.4 Multicenter studies

The center-by-treatment interaction will be tested by replacing region with center in the logistic regression model used for the sensitivity analysis (Section 8.1.3.5) and adding a center-by-treatment interaction term. In the model, center will be based on the original centers prior to pooling (see Section 3.7). However, if the model is unable to converge due to a low number of subjects at a given center, a pooling by center will be applied in order to allow the model to converge. If convergence still cannot be achieved, this analysis will not be performed. Detailed strategy in Section 3.7 will be applied in order to allow the model to converge.

4.5 Multiple comparisons/multiplicity

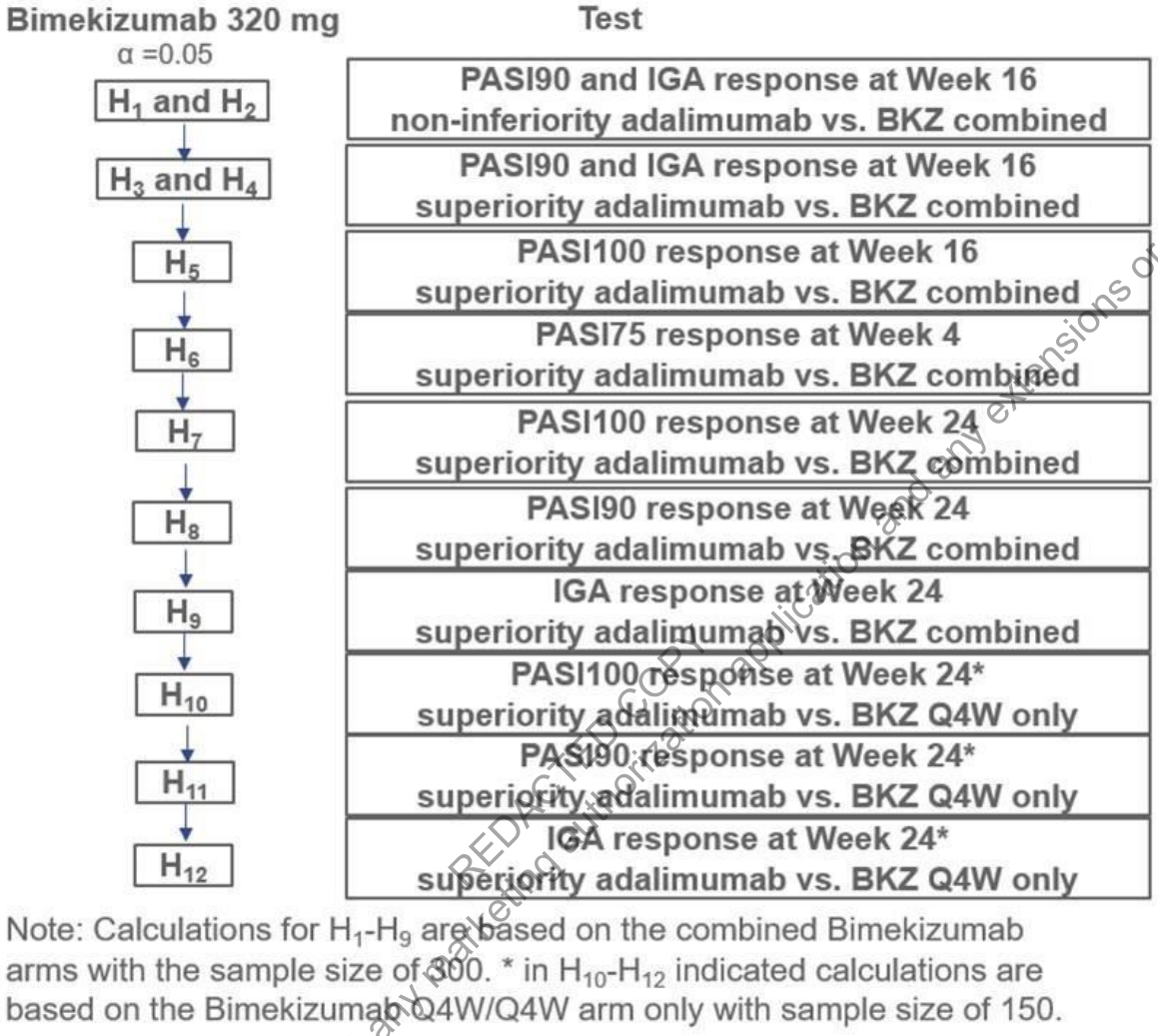
The statistical analysis of the co-primary efficacy variables and selected secondary efficacy variables will account for multiplicity and control the familywise Type I error rate at a 2-sided alpha level of 0.05 by using a fixed sequence testing procedure.

The hypotheses (H_1, H_2, H_3, H_4, H_5, H_6, H_7, H_8, H_9, H_{10}, H_{11} and H_{12}) comparing bimekizumab vs. adalimumab will be tested at a 2 -sided alpha level of 0.05.

The first 2 hypotheses (H_1 and H_2) will test whether bimekizumab is noninferior to adalimumab for the co-primary efficacy variables, PASI90 response at Week 16 and IGA response at Week 16. This evaluation of noninferiority will be tested at a 1-sided alpha level of 0.025 and will be based on a 1-sided 97.5% CI and a noninferiority margin of 10%. If noninferiority is achieved, the alpha will be passed to the next test in the sequence, allowing the testing procedure to proceed. The co-primary efficacy variables of PASI90 response at Week 16 and IGA response at Week 16 will then be evaluated for superiority relative to adalimumab at a 2-sided alpha level of 0.05, and testing will proceed only if superiority is achieved for both endpoints.

The hypotheses associated with the subsequent secondary efficacy endpoints are based on testing for superiority relative to adalimumab. See Figure 4-1for details on this procedure.

Figure 4-1: Sequence of testing



BKZ=bimekizumab; H=hypothesis; IGA=Investigator's Global Assessment; PASI=Psoriasis Area Severity Index; Q4W=every 4 weeks

4.6 Use of an efficacy subset of subjects

The primary efficacy analysis (described in Section 8.1.2) will be repeated based on FAS and PPS as a sensitivity analysis.

4.7 Active-control studies intended to show equivalence

A non-inferiority margin of 10% has been selected as this is considered to be a clinically relevant difference that could influence the choice of interventions used to treat chronic plaque PSO.

4.8 Examination of subgroups

Subgroup analyses will be performed on PASI75/90/100 response rates and IGA, using by visit summaries only. The following subgroups for analysis will be determined using baseline data:

- Age (<40 years, 40 to <65 years, ≥65 years)

- Gender (male, female)
- Disease duration (<median, ≥median)
- Region (North America [Canada, USA], Western Europe [Germany], Central/Eastern Europe [Czech Republic, Hungary, Poland, Russian Federation], Asia/Australia [Australia, Republic of Korea, Taiwan]
- Weight (≤100 kg, >100 kg)
- BMI (<25 kg/m^2, 25 to <30 kg/m^2, ≥30 kg/m^2)
- Prior systemic phototherapy or chemotherapy (yes, no)
- Prior biologic exposure (yes, no)
- Prior systemic therapy of any kind (yes, no)
- Baseline disease severity (PASI<20, PASI≥20)
- Antibody positivity (negative, positive)

Antibody positivity is the only subgroup that is not determined by Baseline data. It will be presented in a separate table.

The definition of prior systemic therapy of any kind is that if a subject received previous biologic therapy, previous systemic therapy (non-biologic), or previous systemic chemotherapy or phototherapy. Subjects who never received previous biologic therapy, previous systemic therapy (non-biologic), or previous systemic photochemotherapy or phototherapy will be classified as not receiving prior systemic treatment for psoriasis.

In addition, in order to assess whether an early response to treatment is predictive of response at later time points, subgroup analyses will be performed for the RS on PASI90/100 and IGA over time using the following early response subgroups:

- PASI75 responders (OC) at Week 4 (yes, no) to predict PASI90/100 and IGA (NRI) through initial and maintenance treatment period
- PASI90 responders (OC) at Week 16 (yes, no) to predict PASI100 (NRI) through maintenance treatment period

All summaries will be based on imputed data as appropriate and will include descriptive statistics only.

5 STUDY POPULATION CHARACTERISTICS

5.1 Subject disposition

Summaries of reasons for screen failures (for all subjects screened), disposition of subjects (for all subjects screened), disposition of analysis sets (for RS), disposition and discontinuation reasons through Week 24 (for RS) and through initial and maintenance treatment period (for BKZ Set), as well as the subjects who discontinued due to AEs up to Week 24 (for RS) and up to Week 56 (for BKZ Set) will be produced. The disposition of subjects for all subjects screened will include the number of subjects included in each analysis set (ES, RS, SS, FAS, BKZ Set, BKZ Week 24 Set, MS, PPS and PK-PPS) overall and by site.

A subject is considered to have completed the Week 16 initial treatment period if they have either completed the Week 16 visit or received study medication from Week 16 onwards. Similar for Week 24.

The following listings for subject disposition will be provided: subjects who did not meet study eligibility criteria (all subjects screened), subject disposition (all subjects screened), study discontinuation (RS), visit dates (RS), subjects excluded from efficacy analysis (RS).

5.2 Protocol deviations

A summary of number and percentage of subjects with an important protocol deviation (including a summary of subjects excluded from any analysis set due to important protocol deviations) by treatment group and through initial treatment period, through Week 24 and through initial and maintenance treatment period will be provided for the RS.

A by-subject listing of important protocol deviations will be provided.

6 DEMOGRAPHICS AND OTHER BASELINE CHARACTERISTICS

All summaries detailed in this section will be performed on the RS by treatment group. Summaries for demographics and other Baseline characteristics will also be repeated in SS and BKZ Set. If the SS and RS analysis sets are identical the summaries will not be repeated.

6.1 Demographics

Demographic variables will be summarized by treatment group and overall.

The following continuous variables will be summarized using descriptive statistics (number of subjects, mean, SD, minimum, median and maximum).

- Age at the time of study entry (years)
- Height (cm)
- Weight (kg)
- BMI (kg/m^2)

  BMI (kg/m^2) will be calculated as:

$$BMI = \frac{\text{Weight (kg)}}{(\text{Height (m)})^2}$$

The following categorical variables will be summarized using frequency counts and percentages.

- Age group (18-<65, 65-<85, ≥85 years)
- Age group (≤18, 19-<65, ≥65 years)
- Age group (<40, 40-<65, ≥65 years)
- Body Weight (≤100kg, >100kg)
- BMI (<25 kg/m2, 25 to <30 kg/m2, ≥30 kg/m2)

- Gender
- Race
- Ethnicity

By-subject listings of demographics will be provided.

6.2 Other Baseline characteristics

Baseline characteristics (including Baseline clinical measures) will be summarized by treatment group and overall.

Generally, the following continuous variables will be summarized using descriptive statistics (number of subjects, mean, SD, minimum, median and maximum).

- Psoriasis Body Surface Area (BSA; %)
- Psoriasis Area and Severity Index (PASI) score
- Modified Nail Psoriasis Severity Index (mNAPSI) total score
- Modified Nail Psoriasis Severity Index (mNAPSI) total score for subjects with nail involvement (i.e. mNAPSI>0)
- Patient Global Assessment of Disease Activity (PGADA) for arthritis visual analogue scale (VAS) score
- Dermatology Life Quality Index (DLQI) total score
- Duration of disease (years)

  Duration of disease (years) will be calculated as:

$$Disease\ Duration = \frac{(\text{Date of randomization} - Date\ of\ onset\ of\ Plaque\ Psoriasis^{1})}{365.25}$$

[1]If the date of onset of plaque psoriasis is partial, it should be imputed to the most recent feasible date (ie, last day of the month if only day is missing, or the last day of the year if day and month are missing). Note that if the date of randomization is missing then the duration of disease will be derived using the date of screening.

The following categorical variables will be summarized using frequency counts and percentages.

- Region (North America [Canada, USA], Western Europe [Germany], Central/Eastern Europe [Czech Republic, Hungary, Poland, Russian Federation], Asia/Australia [Australia, Republic of Korea, Taiwan]
- Country
- Duration of disease (<median, ≥median)
- Investigator's Global Assessment (IGA) score
- Baseline disease severity (PASI<20, PASI≥20)
- PGADA score (PGADA=0, PGADA>0)
- DLQI total score (DLQI=0, DLQI>0)

- Nail involvement (yes, no)
- Scalp involvement (yes, no)
- Palmoplantar involvement (yes, no)
- Prior biologic exposure (yes, no)
- Prior primary failure to biologic (yes/no)
- Prior anti-TNF therapy (yes, no)
- Prior phototherapy or chemotherapy (yes, no)
- Any prior systemic therapy (yes, no)
- PSD items: Pain, Itch, Scaling

Baseline nail, scalp, and palmoplantar involvement are based on the number of subjects achieving mNAPSI>0, Scalp IGA>0, and pp-IGA>0, respectively. The categorization of whether or not subjects had prior exposure to biologic therapy will be based on the Psoriasis Treatment History CRF module. Prior anti-TNFs include etanercept, adalimumab, infliximab, certolizumab pegol, and golimumab.

By-subject listings of Baseline characteristics will be provided.

6.3 Medical history and concomitant diseases

Previous and ongoing medical history will be summarized by treatment group(s), system organ class (SOC) and preferred term (PT) using MedDRA®. Medical procedures are not coded.

The following listings for medical history and concomitant diseases will be provided: medical history, psoriasis history, concomitant medical procedures, previous and ongoing medical history glossary, previous and ongoing medical history conditions, and procedure history.

6.4 Prior and concomitant medications

Medication start and stop dates will be compared to the date of first dose of treatment to allow medications to be classified as either Prior or Concomitant.

Details of imputation methods for missing or partial dates are described in Section 4.2.2.

Prior medications include any medications that started prior to the start date of study medication.

Concomitant medications are medications taken at least one day in common with the study medication dosing period.

The number and percentage of subjects taking prior medications (excluding past psoriasis medications) will be summarized by treatment group, overall and by Anatomical Therapeutic Chemical classification (ATC) class, presenting Anatomical Main Group (ATC Level 1), Pharmacological Subgroup (ATC level 3), and preferred term. The number and percentage taking concomitant medications will be summarized similarly.

Past psoriasis medications will be captured separately and will also be summarized by treatment group. These medications are not subject to dictionary coding. In addition, subjects who failed past psoriasis biologic treatment will be summarized by reason of failure as captured on the Psoriasis Treatment History CRF module.

By-subject listings of all Prior and Concomitant medications, prior and concomitant medications glossary, and psoriasis treatment history will be provided.

7 MEASUREMENTS OF TREATMENT COMPLIANCE

Due to the method of administration of the treatments, compliance will be examined in terms of completed injections.

Treatment compliance will be calculated as:

$$\frac{total\ number\ of\ completed\ injections}{total\ number\ of\ expected\ injections} \times 100$$

where the total number of expected injections is derived relative to when the subject finishes treatment. In this study, bimekizumab, adalimumab, or placebo are administered according to the schedule in Table 7–1.

It is expected that a subject should complete a total of 40 injections by the end of study. If a subject discontinues early, then the number of expected injections is based on the time of early discontinuation relative to the dosing visits. For example, if a subject discontinues after Week 8 visit and prior to Week 9 visit, the total number of expected injections will be 10.

Percent treatment compliance will be described with summary statistics and categorized as <75% and ≥75% , by treatment group and study periods (though Week 24 for SS, and through initial and maintenance treatment period for BKZ Set).

A by-subject listing of treatment compliance will be provided.

Table 7–1: Dosing Scheme

7Week Dose Assignment	0	1, 3	4	5, 7	8	9, 11	12	13, 15	16	17, 19	20	21, 23	24	28	32	36	40	44	48	52
bimekizumab 320mg Q4W/Q4W	●●	○	●●	○	●●	○	●●	○	●●	○	●●	○	●●	●●	●●	●●	●●	●●	●●	●●
bimekizumab 320mg Q4W/Q8W	●●	○	●●	○	●●	○	●●	○	●●	○	○○	○	●●	○○	●●	○○	●●	○○	●●	○○
Adalimumab 40mg / bimekizumab 320mg Q4W	▲▲	▲	○○	▲	○○	▲	○○	▲	○○	▲	○○	▲	●●	●●	●●	●●	●●	●●	●●	●●

Q4W=every 4 weeks; Q8W=every 8 weeks
Notes: A bimekizumab 160mg injection is depicted by a black circle (●). A placebo injection is depicted by a white circle (○). An adalimumab 40mg injection is depicted by a black triangle (▲)

8 EFFICACY ANALYSES

All efficacy analyses will be performed on the RS unless otherwise specified.

For efficacy endpoints tested through initial treatment period, comparisons will be made between BKZ Total (pooling the two identical treatment arms prior to week 16) and ADA treatment groups.

For by visit summaries, results will be presented by each individual treatment group as well as BKZ Total group (BKZ 320mg Q4W and BKZ 320mg Q4W/Q8W combined). Change from Baseline will be summarized through initial and maintenance treatment period for BKZ treatment groups; while for the ADA group, the change from baseline will only be summarized up to Week 24.

Additional summaries through the initial and maintenance treatment period will be produced for subjects who received bimekizumab from Week 24 onwards (BKZ week 24 set) in all three randomized treatment groups.

In addition, for subjects who started with adalimumab and switched treatment at the Week 24 Visit, change from Baseline (at Week 24) will be evaluated from Week 28 to Week 56 for PASI score and DLQI score only (see Section 3.2).

8.1 Statistical analysis of the co-primary efficacy variables

8.1.1 Derivations of co-primary efficacy variables

For missing PASI and IGA assessments at Week 16, available data from all visits (unscheduled or scheduled but without a planned assessment) within the 3 days window will be included.

8.1.1.1 Psoriasis Area and Severity Index (PASI)

PASI90 is defined to be equal to 1 if the percentage improvement from Baseline in the PASI scores is 90% or greater and 0 if the percentage improvement from Baseline is less than 90%. This definition is introduced for the purpose of identifying subjects who respond to the treatment (1 = responder, 0 = non-responder).

PASI scoring of psoriatic plaques is based on three criteria: redness (R), thickness (T), and scaliness (S). Severity is rated for each index (R, S, T) on a 0-4 scale (0 for no involvement up to 4 for very marked involvement). The body is divided into four areas comprising the head (h), upper extremities (u), trunk (t), and lower extremities (l). In each of these areas, the fraction of total surface area affected is graded on a 0-6 scale (0 for no involvement; up to 6 for 90% - 100% involvement).

The various body regions are weighted to reflect their respective proportion of BSA. The composite PASI score is then calculated by multiplying the sum of the individual-severity scores for each area by the weighted area-of-involvement score for that respective area, and then summing the four resulting quantities as follows (note for R, T, and S scores are as follows: 0 = none, 1 = slight, 2 = moderate, 3 = marked, and 4 = very marked):

$$\begin{aligned}\text{PASI} = {} & (0.1 \times (R_h + T_h + S_h) \times A_h) + (0.2 \times (R_u + T_u + S_u) \times A_u) \\ & + (0.3 \times (R_t + T_t + S_t) \times A_t) + (0.4 \times (R_l + T_l + S_l) \times A_l)\end{aligned}$$

where

R_h, R_u, R_t, R_l = redness score of plaques on the head, upper extremities, trunk, and lower extremities, scored 0-4 respectively;

T_h, T_u, T_t, T_l = thickness score of plaques on the head, upper extremities, trunk, and lower extremities, scored 0-4 respectively;

S_h, S_u, S_t, S_l = scaliness score of plaques on the head, upper extremities, trunk, and lower extremities, scored 0-4 respectively;

A_h, A_u, A_t, A_l = numerical value translation of % area of psoriatic involvement score for the head, upper extremities, trunk, and lower extremities, respectively (where 0 = 0% [clear], 1 = >0% to <10%, 2 = 10% to <30%, 3 = 30% to <50%, 4 = 50% to <70%, 5 = 70% to <90%, and 6 = 90% to 100%).

The highest potential PASI score is 72 for severe disease; the lowest is 0 for no psoriasis lesions. PASI scores are treated as continuous. The percent improvement in PASI scores from Baseline will be computed as:

$$\text{Percent improvement from Baseline } = 100 \times \frac{\text{Baseline PASI} - \text{Post Baseline PASI}}{\text{Baseline PASI}}$$

If a subject has experienced an improvement, this measure will be positive. If a subject has experienced a worsening in their condition, this measure will be negative.

If a subject is missing 1 or 2 severity measurements for a certain region, the average of the remaining severity measurement(s) within that region will be utilized to substitute for the missing severity measurement(s) in that region. If the area of affected skin and/or all severity measurements for up to 2 regions are missing, then the missing (R+T+S) x A for a region will be substituted by the average of the available (R+T+S) x A. Otherwise, the PASI will be set to missing.

8.1.1.2 PASI90 Response at Week 16

A categorical response variable, PASI90 at Week 16 is defined to be equal to 1 if the percentage improvement from Baseline to Week 16 in the PASI scores is 90% or greater and 0 if the percentage improvement from Baseline to Week 16 is less than 90%. This definition is introduced for the purpose of identifying subjects who respond to the treatment (1 = responder, 0 = non-responder).

8.1.1.3 Investigator's Global Assessment (IGA) Response at Week 16

A static IGA for psoriasis will be used to assess disease severity in all subjects during the study.

The Investigator will assess the overall severity of psoriasis using the following five-point scale:

Table 8–1: Investigator's Global Assessment

Score	Short Descriptor	Definition
0	Clear	No signs of psoriasis; post-inflammatory hyperpigmentation may be present
1	Almost clear	No thickening; normal to pink coloration; no to minimal focal scaling
2	Mild	Just detectable to mild thickening; pink to light red coloration; predominately fine scaling
3	Moderate	Clearly distinguishable to moderate thickening; dull to bright red, clearly distinguishable to moderate thickening; moderate scaling
4	Severe	Severe thickening with hard edges; bright to deep dark red coloration; severe/coarse scaling covering almost all or all lesions

IGA response is defined as clear [0] or almost clear [1] with at least a 2-category improvement from Baseline at Week 16.

8.1.2 Primary analysis of the co-primary efficacy variables

The co-primary efficacy variables for this study are PASI90 response and IGA response at Week 16, and the corresponding analyses are based on the RS. A subject will be classified as a PASI90 responder if the PASI score at Week 16 has improved at least 90% from Baseline, and IGA responder is any subject with a score of 0 or 1 (Clear or Almost Clear) with at least a 2-category improvement from Baseline to Week 16 in IGA score.

The primary analysis will be based on the stratified Cochran-Mantel-Haenszel (CMH) test where region and prior biologic exposure (yes/no) will be used as stratification variables.

For the assessment of noninferiority, a noninferiority margin of 10% will be used and evaluated based on the confidence interval for the stratified Cochran Mantel-Haenszel risk difference between bimekizumab and adalimumab. The Wald method will be used to calculate the confidence interval. A non-inferiority margin of 10% has been selected as this is considered to be a clinically relevant difference that could influence the choice of interventions used to treat chronic plaque PSO. Therefore, a difference within the 10% non-inferiority margin would suggest a similar impact on efficacy between the treatments.

The evaluation of superiority will use pairwise treatment comparisons based on the CMH test using the p-value for the general association. The odds ratio and associated confidence interval (CI) based on the Wald test will be presented. If one of the treatment groups has 0 or very low response where CMH can no longer be used, the logit method will be applied instead.

To calculate the stratified Mantel-Haenszel risk difference, the method of Greenland and Robins (1985) is used. For each combination of strata, a 2x2 table of treatment group and response is created. A theoretical 2x2 table for a given stratum is shown below, where n = a+b+c+d.

	Response	
Treatment Group	**Yes**	**No**
Bimekizumab	a	c
Adalimumab	b	d

Given that structure, the stratified Mantel-Haenszel risk difference, standard error, and two-sided (1-α)*100% confidence interval may be written as follows:

$$RD_{MH=}\frac{\sum_i((a_i*(b_i+d_i)/n_i)\ -(b_i*(a_i+c_i)/n_i)}{\sum_i((a_i+c_i)*(b_i+d_i))/n_i}$$

$$SE_{MH}=\sqrt{\frac{\sum_i\left\{\frac{[a_i*c_i*(b_i+d_i)^3]+[b_i*d_i*(a_i+c_i)^3]}{(a_i+c_i)*(b_i+d_i)*n_i^2}\right\}}{\left\{\sum_i\left[\frac{(a_i+c_i)*(b_i+d_i)}{n_i}\right]\right\}^2}}$$

$$CI_{MH}=RD_{MH}\pm\text{probit}\left(1-(\alpha/2)\right)*SE_{MH}$$

For the assessment of non-inferiority of bimekizumab to adalimumab, the lower 97.5% confidence limit for the stratified Mantel-Haenszel risk difference will be considered. If that value is greater than -10%, then non-inferiority will have been established.

Non-responder imputation will be used to account for missing data in the primary analysis. Specifically, any subject who withdraws from IMP prior to Week 16 or who has missing data for the co-primary efficacy variables at the Week 16 time point will be considered as a non-responder.

The number and percentage of subjects who are PASI90 responders at Week 16 will be summarized. IGA response will be summarized in the same manner as PASI90 responder variable.

A line plot of the percentage improvement from Baseline in PASI score over time by treatment group will be produced.

By-subject listings of PASI and IGA responder variables, PASI and IGA data will be provided.

8.1.3 Sensitivity analyses of the co-primary efficacy variables

The following sensitivity analyses for the co-primary efficacy variables will be performed to evaluate the assumptions related to the handling of missing data:

8.1.3.1 Sensitivity Analysis #1

Missing data will be addressed by using MI (Markov-Chain Monte Carlo [MCMC] method for intermittent missing data, followed by monotone regression for monotone missing data, see

Section 4.2.1.6 to evaluate the effect of the method for handling missing data on the analysis. The actual PASI/IGA scores will be imputed and then dichotomized to obtain the response status. The treatment differences for each imputed data set will subsequently be evaluated using the stratified CMH test as used in the primary analysis. The results from each of the imputed data sets will be combined for overall inference using Rubin's rules, which account for the uncertainty associated with the imputed values (Rubin, 1987). This will be done using SAS PROC MIANALYZE. This procedure assumes a missing at random (MAR) pattern of missingness and corresponds to an estimand of the difference in outcome improvement if all subjects tolerated or adhered to treatment (Mallinckrodt et al, 2012). This is an estimand of efficacy to evaluate the de jure hypothesis.

8.1.3.2 Sensitivity Analysis #2

This sensitivity analysis will be based on observed data at Week 16. Subjects with missing data or who have prematurely discontinued IMP will be treated as missing (see Section 4.2.1.6). The same stratified CMH test as in the primary efficacy analysis will be used.

8.1.3.3 Sensitivity Analysis #3

The primary efficacy analyses from Section 8.1.2 will be repeated using LOCF as the imputation method (see Section 4.2.1.5).

8.1.3.4 Sensitivity Analysis #4

The primary efficacy analyses from Section 8.1.2 will be repeated based on the FAS and the PPS.

8.1.3.5 Sensitivity Analysis #5

As a sensitivity analysis to the primary analysis method, logistic regression based on the RS will be used. The odds ratio of the responder rates at Week 16 will be estimated and tested between treatment groups using a logistic regression model with factors of treatment group, region, and prior biologic exposure (yes/no). The odds ratio, associated confidence interval (CI), and p-value based on the Wald test will be presented. If the logistic regression model is unable to converge, then prior biologic exposure may be dropped from the model to facilitate convergence. If the model is still unable to converge, then region may be removed from the model as well. In addition, if the logistic regression model cannot converge, then Fisher's exact test will be used for inferential comparisons. As with the primary analysis, missing data will be handled using NRI.

8.1.3.6 Sensitivity Analysis #6

The center-by-treatment interaction will be tested by replacing region with center in the logistic regression model described in Section 8.1.3.5 and adding a center-by-treatment interaction term. In the model, center will be based on the original centers prior to pooling. However, if the model is unable to converge due to a low number of subjects at a given center, a pooling (see Section 3.7) will be described in order to allow the model to converge. In order to obtain reasonable estimates of variability for a treatment arm at a given center, a minimum of 15 subjects will be considered acceptable for a center to be included in the model without pooling. Given the 1:1:1 randomization allocation scheme, this should provide a minimum of about 5 subjects in each treatment group. Centers with fewer than 15 subjects will be eligible for pooling. The pooling algorithm used is described in Section 3.7.

In order to achieve model convergence, other explanatory variables e.g. other prior biologic exposure may be dropped from the model as for the logistic regression with region.

If the center-by-treatment interaction is not found to be significant (α=0.10), then no further analyses will be performed. On the other hand, if the interaction is significant, further analyses will be conducted to determine which center or centers may be the source of interaction. This will be done by running the logistic regression model (including the interaction term) where each center will be systematically removed from the model. This impact of a given center will be based on the change in the interaction p-value when that center is removed. The center or centers that appear to be driving the significant interaction effect will then be removed from the model to verify that conclusions remain the same with or without the influential center(s). This sensitivity analysis will be based on RS with NRI for missing data.

8.2 Statistical analysis of the secondary efficacy variables

8.2.1 Derivations of secondary efficacy variables

8.2.1.1 PASI100 Response at Week 16 and 24

A categorical response variable, PASI100 at Week 16 is defined to be equal to 1 if the percentage improvement from Baseline to Week 16 in the PASI scores is 100% and 0 if the percentage improvement from Baseline to Week 16 is less than 100%. This definition is introduced for the purpose of identifying subjects who respond to the treatment (1 = responder, 0 = non-responder). The definition of percentage improvement from Baseline is given in Section 8.1.1.1.

PASI100 at Week 24 will be summarized in the same manner as PASI100 at Week 16.

8.2.1.2 PASI75 Response at Week 4

A categorical response variable, PASI75 at Week 4 is defined to be equal to 1 if the percentage improvement from Baseline to Week 4 in the PASI scores is 75% or greater and 0 if the percentage improvement from Baseline to Week 4 is less than 75%. This definition is introduced for the purpose of identifying subjects who respond to the treatment (1 = responder, 0 = non-responder). The definition of percentage improvement from Baseline is given in Section 8.1.1.1.

8.2.1.3 PASI90 Response at Week 24 and 56

PASI90 response definition and derivation are outlined in Section 8.1.1.1 and Section 8.1.1.2, use the data for weeks 24 and 56 respectively.

8.2.1.4 IGA Response at Week 24 and Week 56

IGA response and IGA 5-point scale are outlined in Section 8.1.1.3, define at Weeks 24 and 56 respectively.

8.2.2 Primary analysis of the secondary efficacy variables

Binary (responder) variables will be summarized using frequency tables by treatment group for each visit. Continuous variables will be summarized using descriptive statistics by treatment group for each visit. Primary analysis for secondary efficacy variables will be summarized based on imputed data (NRI and MCMC/Monotone Regression for binary and continuous variables, respectively), unless otherwise specified.

8.2.2.1 PASI and IGA responses

For PASI100 response at Week 16 and 24, PASI75 response at Week 4, PASI90 and IGA response at Week 24, the stratified CMH test similar to the primary analysis will be implemented to test for superiority of bimekizumab over adalimumab.

The treatment comparisons at Week 24 for PASI100, PASI90 and IGA will be performed in three different ways using the following randomized treatment arms:

1. Bimekizumab Q4W and bimekizumab Q4W/Q8W combined vs adalimumab (testing H_7, H_8 and H_9)
2. Bimekizumab Q4W only vs adalimumab (testing H_{10}, H_{11} and H_{12})
3. Bimekizumab Q4W/Q8W vs adalimumab (not part of the hierarchical testing sequence, these analyses will be exploratory and the p-values produced will be nominal)

Because all adalimumab subjects will receive bimekizumab treatment starting at Week 24, inferential comparisons against adalimumab after Week 24 will not be performed. PASI90 and IGA response at Week 56 will be summarized using descriptive statistics only.

8.2.3 Sensitivity analyses of the secondary efficacy variables

For binary response variables (PASI100 at Week 16 and 24, PASI75 at Week 4, PASI90 at Weeks 24 and 56, and IGA at Weeks 24 and 56), sensitivity analyses #1 and #2 (Section 8.1.3.1 and Section 8.1.3.2) will be performed. For PASI90 and IGA at Week 56, the sensitivity analyses are used to impute missing data for descriptive summaries only. Details regarding missing data algorithm can be found in Section 4.2.1.6.

It was noted during the conduct of the study that approximately 24 subjects received incorrect Placebo/Active treatment at Visit 20. All subsequent treatments were administered per protocol. The main analyses of the endpoints at Week 24 will include all subjects in the RS, however the potential impact of this on the Week 24 secondary endpoints (PASI90, PASI100 and IGA response) will be assessed through sensitivity analyses excluding these subjects from the analysis population.

8.3 Statistical analysis of other efficacy variables

The other efficacy variables are listed below and will be evaluated according to the planned assessments in the protocol, this excludes the timepoints for the primary and secondary variables specified above in Section 8.1.1 and Section 8.2.1.

Binary (responder) variables will be summarized using frequency tables by treatment group for each visit. Continuous variables will be summarized using descriptive statistics by treatment group for each visit. All variables will be summarized based on imputed data (NRI and MCMC/Monotone Regression for binary and continuous variables, respectively), unless otherwise specified.

For variables that are part of the sequence testing procedure, summaries based on observed case data will also be provided. There may be cases where the multiple imputation model fails to converge. In such situations, the last observation carried forward (LOCF) approach will instead be used to impute the missing data.

By visit summaries will include the bimekizumab and adalimumab treatment groups. There will be some key approaches for understanding the data after Week 24:

- Summary of data for the RS through initial and maintenance treatment period based on randomized treatment group including bimekizumab Q4W, bimekizumab Q4W/Q8W and bimekizumab total. The adalimumab group will only be summarized up to Week 24
- Summaries through the initial and maintenance treatment period for subjects who received bimekizumab from Week 24 onwards (BKZ week 24 set) in all three randomized treatment groups.
- For subjects who started with adalimumab and switched treatment at the Week 24 Visit, change from Baseline (at Week 24) will be evaluated from Week 28 to Week 56 for PASI score and DLQI score only (see Section 3.2).
- Summary of Maintenance treatment period data among responders only for PASI90, PASI100, IGA (0/1), and DLQI respectively;
  - For Week 16 responders in bimekizumab Q4W, bimekizumab Q8W and bimekizumab total (Maintenance Set)
  - For Week 24 responders in ADA/BKZ arm (Bimekizumab Set)

No adjustments to the analysis sets will be made for the subjects who were incorrectly dosed (see Section 8.2.3)

8.3.1 PASI

8.3.1.1 PASI50, PASI75, PASI90 and PASI100 response rate

Categorical response variables, PASI50, PASI75, PASI90 and PASI100 over time are defined to be equal to 1 if the percentage improvement from Baseline to visit timepoint in the PASI scores is 50%, 75%, 90% and 100% respectively or greater and 0 if the percentage improvement from Baseline to visit timepoint is less than 50%, 75%, 90% and 100% respectively. This definition is introduced for the purpose of identifying subjects who respond to the treatment (1 = responder, 0 = non-responder). The definition of PASI score percentage improvement from Baseline is given in Section 8.1.1.1.

A line plot of the PASI responder (PASI50, PASI75, PASI90, and PASI100) rate over time, by treatment group will be produced.

8.3.1.2 Time to PASI50, PASI75, PASI90 and PASI100 response through initial treatment period and through Week 24

Time to PASI50, PASI75, PASI90, and PASI100 response (in days) during the Initial Treatment Period will each be calculated as:

Min (Date of first PASIxx response, Date of Week 16 visit) – Date of Baseline visit + 1, here xx represents 50, 75, 90, 100 respectively.

All visits up to and including Week 16 (including unscheduled visits) will be considered.

For the PASI response through initial treatment period variables, subjects who discontinue study treatment without achieving a given PASI response prior to Week 16 visit will be censored at the date of the last observed PASI assessment on or prior to the date of study treatment

discontinuation. Subjects who reach the Week 16 Visit without achieving the given response will be censored at the date of the last observed PASI assessment on or prior to Week 16 Visit.

Subject will be censored at baseline (Day 0) if there is no Baseline PASI assessment or no post Baseline PASI assessment.

For the PASI response through Week 24 variables, data will be censored in the same manner as Week 16 initial treatment period variables.

Time to PASI50, PASI75, PASI90, and PASI100 response will each be estimated and presented using the Kaplan-Meier product-limit method for each treatment group.

Kaplan-Meier plots of time to PASI responses will also be presented by treatment group. In these Kaplan-Meier plots, the line will start at 0 (since there are no responders at Week 0) and will increase over time, representing time to achieving the response.

The median time to response, including the two-sided 95% confidence interval, will be calculated for each treatment. Comparisons of all bimekizumab vs adalimumab will be analyzed using a log-rank test stratified by region and prior biologic exposure.

8.3.1.3 PASI score

Absolute and percent change from Baseline in PASI score is defined in Section 8.1.1.1.

The percent of subjects with absolute PASI score ≤1, ≤2, ≤3 and ≤5 will be presented over time.

8.3.2 IGA response

- IGA response (Clear with at least a 2-category improvement relative to Baseline) is defined when IGA score is zero with at least a 2-category improvement from Baseline at visit timepoint. (Table 8–1).
- IGA response (Clear or Almost Clear with at least 2 category improvement relative to Baseline) is defined as IGA score of clear [0] or almost clear [1] with at least a 2-category improvement from Baseline at visit timepoint.
- Shift from Baseline in IGA score is defined at each Post-Baseline visit timepoint relative to Baseline.
- Scalp-specific IGA response

A static IGA for scalp PSO will be used to assess disease severity on the scalp

All subjects will complete the scalp IGA at Baseline. Only subjects with scalp involvement at Baseline will complete the scalp IGA at the other visits specified in the protocol. Subjects with scalp involvement at Baseline are defined as those with a scalp IGA score >0 at Baseline.

Scalp lesions will be assessed in terms of clinical signs of redness, thickness, and scaliness using a 5-point scale as outlined in Table 8–2 below.

Table 8–2: Scalp IGA

Score	Short Descriptor	Detailed Descriptor
0	Clear	Scalp has no signs of PSO; post-inflammatory hyperpigmentation may be present
1	Almost clear	Scalp has no thickening; normal to pink coloration; no to minimal focal scaling
2	Mild	Scalp has just detectable to mild thickening; pink to light red coloration; predominately fine scaling
3	Moderate	Scalp has clearly distinguishable to moderate thickening; dull to bright red, moderate scaling
4	Severe	Scalp has severe thickening with hard edges; bright to deep dark red coloration; severe/coarse scaling covering almost all or all lesions

Scalp IGA response is defined as clear [0] or almost clear [1] with at least a 2-category improvement from Baseline. For analysis purposes, the evaluation of scalp IGA will be limited to subjects with a Baseline scalp IGA of at least 2. Therefore, if a subject has a score of 2 at Baseline, they can only be considered a responder if their IGA is 0 (thereby meeting the criterion for a 2-category improvement from Baseline). Subjects with a Baseline scalp IGA of 1 will be assessed per the protocol but will not be part of the scalp IGA analysis.

- Palmoplantar Investigator's Global Assessment (pp-IGA) response

A static IGA for palmoplantar PSO will be used to assess palmoplantar disease severity.

Only subjects with palmoplantar involvement at Baseline will complete the pp-IGA at the other visits specified in the protocol. Subjects with palmoplantar involvement at Baseline are defined as those with a pp-IGA score >0 at Baseline.

Palmoplantar disease will be assessed in terms of clinical signs of redness, thickness, and scaliness using a 5-point scale as outlined in Table 8–3 below.

Table 8–3: Palmoplantar IGA

Score	Short Descriptor	Detailed Descriptor
0	Clear	Palmoplantar has no signs of PSO; post-inflammatory hyperpigmentation may be present
1	Almost clear	Palmoplantar has no thickening; normal to pink coloration; no to minimal focal scaling
2	Mild	Palmoplantar has just detectable to mild thickening; pink to light red coloration; predominately fine scaling
3	Moderate	Palmoplantar has clearly distinguishable to moderate thickening; dull to bright red and clearly distinguishable coloration; moderate scaling
4	Severe	Palmoplantar has severe thickening with hard edges; bright to deep dark red coloration; severe/coarse scaling covering almost all or all lesions; numerous fissures

PSO=psoriasis; pp-IGA=palmoplantar-specific Investigator's Global Assessment

Palmoplantar IGA response is defined as clear [0] or almost clear [1] with at least a 2-category improvement from Baseline. For analysis purposes, the evaluation of pp-IGA will be limited to subjects with a Baseline pp-IGA of at least 2. Therefore, if a subject has a score of 2 at Baseline, they can only be considered a responder if their IGA is 0 (thereby meeting the criterion for a 2-category improvement from Baseline). Subjects with a Baseline pp-IGA of 1 will be assessed per the protocol but will not be part of the palmoplantar IGA analysis.

A line plot of the IGA responder rate over time, by treatment group will be produced.

8.3.3 BSA

- Absolute change from Baseline in the BSA affected by PSO is defined as Post Baseline BSA minus Baseline BSA affected by PSO.
- Percent change from Baseline in BSA affected by PSO is defined as

$$\text{Percent change from Baseline } = 100 \times \frac{\text{Post Baseline BSA} - \text{Baseline BSA}}{\text{Baseline BSA}}$$

The percent of subjects with absolute BSA=0%, ≤1%, ≤3% and ≤5% will be presented over time.

8.3.4 Product of IGA and BSA (IGAxBSA)

- Absolute change from Baseline in the product IGAxBSA is defined as Post Baseline IGAxBSA minus product of Baseline IGAxBSA.
- Percent change from Baseline in the product of IGAxBSA is defined as

$$\text{Percent change from Baseline } = 100 \times \frac{\text{Post Baseline IGAxBSA} - \text{Baseline IGAxBSA}}{\text{Baseline IGAxBSA}}$$

8.3.5 Dermatology Life Quality Index (DLQI)

The DLQI questionnaire is used for subjects with psoriasis and consists of 10-questions. This is validated, quality-of-life questionnaire that covers 6 domains including symptoms and feelings, daily activities, leisure, work and school, personal relationships, and treatment, as assessed over the past week.

The scoring of each answer for the DLQI is as follows:

Table 8–4: Dermatology Life Quality Index

DLQI Scoring	
Response	Score
Very much	3
A lot	2
A little	1
Not at all	0
Not relevant	0
Question unanswered	0
Q7: 'prevented work or studying' = yes	3

The DLQI total score is calculated by adding the score of each question. The maximum score is 30, and the minimum score is 0. The higher the score, the more quality of life is impaired.

Meaning of DLQI Scores

0-1 = no effect at all on patient's life

2-5 = small effect on patient's life

6-10 = moderate effect on patient's life

11-20 = very large effect on patient's life

21-30 = extremely large effect on patient's life

This categorization will not be utilized in the analysis.

Because Q7 has a sub-question (referred to as Q7a here) after the leading yes/no question, some clarifying rules for scoring are provided:

- If Q7 is marked as "yes", a score of 3 is given regardless of the responses to Q7a.
- If Q7 is marked as "no", "not relevant", or is missing and Q7a is "A lot", a score of 2 is given.

- If Q7 is marked as "no", "not relevant", or is missing and Q7a is "A little", a score of 1 is given.
- If Q7 is marked as "no", "not relevant", or is missing and Q7a is "Not at all", a score of 0 is given.
- If Q7 is marked as "no" or "not relevant" and Q7a is missing, a score of 0 is given.
- If Q7 is missing and Q7a is missing, Q7 is considered unanswered (see below for details on how this impacts the overall DLQI score).

If 1 question is left unanswered, this is scored 0 and the scores are summed and expressed as usual out of a maximum of 30. If 2 or more questions are left unanswered, the questionnaire is not scored.

A subject is considered to have achieved the minimally clinical important difference (MCID) if their individual improvement from Baseline score is > 4. A 4-point improvement in the DLQI score (DLQI response) has been reported to be meaningful for the subject (within-subject MCID).

The DLQI related efficacy variables are defined as follows:

- Change from Baseline in DLQI is defined as Post-Baseline DLQI minus Baseline DLQI.
- Percent of subjects achieving a DLQI total score of 0 or 1 is defined as the number of subjects with DLQI absolute score of 0 or 1 divided by the number of subjects in RS.

Percent of subjects achieving a MICD in DLQI is defined as the number of subjects with improvement from Baseline score of 4 or more divided by the number of subjects in RS. Only subjects with a Baseline DLQI of 4 or greater will be considered for this analysis.

8.3.6 Patient's Global Assessment of Disease Activity (PGADA) for arthritis visual analog scale (VAS)

The PGADA for arthritis VAS will be used to provide an overall evaluation of arthritis disease symptoms. Subjects will respond to the question, "Considering all the ways your arthritis affects you, please mark a vertical line on the scale below to show how you are feeling today," using a VAS where 0 is "very good, no symptoms" and 100 is "very poor, severe symptoms.

All subjects will complete the PGADA at Baseline. Subjects with PsA at Baseline (defined as a past medical history of PsA or PASE ≥47) will complete the PGADA at the visits specified in the protocol. Change from Baseline in PGADA is defined as Post-Baseline PGADA minus Baseline PGADA.

8.3.7 Patient Global Assessment of PSO

The Patient Global Assessment of PSO is a PSO-specific item in which the subject responds to the multiple-choice question, "How severe are your psoriasis-related symptoms right now?" Possible responses to the question are "no symptoms," "mild symptoms," "moderate symptoms," "severe symptoms," or "very severe symptoms."

The Patient Global Assessment of psoriasis will be summarized based on OC as the primary analysis. No imputation is applied.

Shift from Baseline in PGA (Patient Global Assessment) of PSO score is defined to each Post-Baseline visit timepoint relative to Baseline.

8.3.8 Patient Symptom Diary (PSD)

A patient reported outcome (PRO) measure, the Patient Symptom Diary (PSD), will be used to assess key symptoms relevant to subjects with moderate to severe chronic plaque PSO. The electronic (ePRO) diary will be administered on a daily basis from Screening to the end of the Week 24 Visit.

The PSD consists of 14 different items, each measuring an aspect of the disease and its impact on the subject's quality of life. Each item will be scored separately on a 0-10 scale with 0 for no symptom and 10 for very severe or worst symptom. Weekly averages will be derived for each of the 14 items of the Psoriasis Diary up to Week 24. A weekly average is the sum of the scored item over the course of the study week (up to the actual visit date) divided by the number of days on which the item was completed and will be set to missing if four or more daily assessments (irrespective of whether these are consecutive or not) were missing of the corresponding question.

The Baseline value is also based on a weekly average (up to and including the first dose date) and will be computed in the same manner.

In cases where there is more than one diary record completed on a particular day, all available records within the 7-day window (including any double entries on one day) will be used in the calculation of weekly average scores. If this results in having more than 7 available scores to calculate the weekly average, the 7 records closest to the visit will be used. The diary entry completion date/time stamp will be used to determine this.

- Change from Baseline in PSD score is defined as post Baseline PSD score minus Baseline PSD score.
- Percent change from Baseline in PSD score is defined as

$$\text{Percent change from Baseline} = 100 \times \frac{\text{Post Baseline PSD score - Baseline PSD score}}{\text{Baseline PSD score}}$$

Absolute and percent changes from baseline PSD score for each item will be summarized by visit for each treatment group.

In addition, each of the PSD scores will be characterized in terms of the cumulative percent of subjects demonstrating a pre-specified point improvement at each visit. The threshold criteria in term of changes from baseline are set to 2.39, 1.98, 2.86, 3.05, 1.99, 2.01, 3.04, 2.82, 2.60, 2.69, 2.68, 1.51, 2.43, and 2.14 respectively for the itch, pain, scaling, redness, burning, cracking, dryness, irritation, sensitivity, lesions, thickening, fatigue, embarrassment, choice of clothing. For this responder analysis, it will be limited to the subjects with a Baseline PSD score at or above the applicable threshold score. Subjects with a missing score will be imputed using NRI.

Cumulative distribution plots will also be provided for absolute change from Baseline PSD at Week 16 for each item.

8.3.9 Modified Nail Psoriasis Severity Index (mNAPSI) score

Psoriatic nail disease will be evaluated at the Baseline visit using the mNAPSI. All affected nails will be scored (0 to 3) for onycholysis/oil drop dyschromia, nail plate crumbling, and pitting and will be scored (0 for "no" or 1 for "yes") for leukonychia, nail bed hyperkeratosis, splinter haemorrhages and red spots in the lunula. The score for an individual nail ranges from 0 to 13 with higher scores indicative of more severe nail disease. The total mNAPSI score is the sum of the scores for each individual nail. If a nail is unaffected, it will be recorded as such and will not contribute to the total mNAPSI score.

If any of the 7 response items that contribute to mNAPSI is present, while other items are missing (ie, partial mNAPSI data), then the missing items are assumed to be 0 for the mNAPSI calculation. In some cases, the data may be captured in such a way that only non-zero component scores are present in the database. Again, those components that are not present are simply assumed to be 0 for the mNAPSI calculation.

Change from Baseline in mNAPSI score for subjects with nail PSO at Baseline is defined as Post-Baseline mNAPSI score minus Baseline mNAPSI.

An mNAPSI75 responder is defined as a subject who achieved at least a 75% improvement from Baseline in the mNAPSI score. mNAPSI90 and mNAPSI100 are defined accordingly. The proportion of mNAPSI75/90/100 responders over time will be summarized for each treatment group.

8.3.10 Psoriatic Arthritis Screening and Evaluation (PASE)

The PASE questionnaire is a self-administered tool to screen for active PsA in subjects with PSO (Husni et al, 2014). The questionnaire consists of 15 items that are divided into a 7-item symptoms subscale and an 8-item functions subscale. Standardized responses are based on 5 categories relating to agreement (strongly agree [5], agree [4], no idea [3], disagree [2], and strongly disagree [1]). The total maximum score is 75 points (symptom score: 35 points and function score: 40 points). Psoriatic Arthritis Screening and Evaluation questionnaire scores ≥47 points are indicative of active PsA.

PASE will be collected at Baseline and Week 56/PEOT visit. PASE will be summarized based on OC only. No imputation is applied.

- Change from Baseline in the PASE questionnaire scores (function score, symptom score, and total score) is defined as Post-Baseline PASE questionnaire score (function, symptom, and total) minus Baseline PASE questionnaire score (function, symptom, and total).
- Shift from Baseline in PASE score suggestive of PsA (<47 versus ≥47) is defined to Week 56/PEOT visit timepoint relative to Baseline.

8.3.11 Short Form – 36 Items Health Survey (SF-36)

The SF-36v2, standard recall, measures the following 8 health domains as rated by the subjects over the past four weeks: Physical Functioning, Role Physical, Bodily Pain, General Health, Vitality, Social Functioning, Role Emotional, and Mental Health (Maruish, 2011).

The SF-36 Physical and Mental Component Summary scores (PCS and MCS, respectively) are used to measure the two broad components, or aspects, of health-physical and mental. PCS and

MCS are based on the aggregate of 8 health concepts described above and all of the 8 health domain scales are used to score both components summary measures.

One additional item asks respondents about health change over the past year.

The SF-36 will be used using QualityMetric's Health Outcomes™ Scoring Software version 5.1 or later. The software uses updates 2009 U.S. population norms and applies a Full Missing Score Estimation (Full MSE) method as follows:

  - A health domain score (except the physical functioning (PF) domain) will be estimated provided that at least one non-missing response is available within that domain
  - For the PF domain item response theory will be used to develop a model for estimates of the missing score (Thomas and Cyr, 2002)
  - Regression methods are then applied to estimate the PCS and the MCS on the basis of the available domains.

- Change from Baseline in Short Form 36-item Health Survey PCS score, and MCS score, and individual domains is defined as respective score at Post-Baseline timepoint minus the Baseline score.

8.3.12 Euro-Quality of Life 5-Dimensions, 3 levels (EQ-5D-3L)

The EQ-5D-3L health questionnaire provides a descriptive profile and a single index value for health status. The instrument is comprised of a 5-item health status measure and a visual analog scale (VAS).

Responses to EQ-5D-3L are scored as 1 for "no problem", 2 for "some or moderate problems", and 3 for "extreme problems".

Absolute EQ-5D-3L VAS score records the respondent's self-rated health status on a vertical 20cm scale, 0 to 100 graduated (0=worst imaginable health status, 100=best imaginable health status).

Change from Baseline in EQ-5D-3L VAS scores is defined as Post-Baseline EQ-5D-3L VAS score minus Baseline EQ-5D-3L VAS score.

Responses to EQ-5D-3L will be summarized based on OC only as primary analysis. No imputation is applied to responses to EQ-5D-3L but is applied to EQ-5D-3L VAS scores.

8.3.13 Work Productivity and Activity Impairment Questionnaire – Specific Health Problem (WPAI-SHP)

The WPAI-SHP is a subject-reported questionnaire that assesses subject's employment status, work absenteeism, work impairment while working, overall work, and daily activity impairment attributable to a specific health problem. Five out of 6 items of the WPAI-SHP are regrouped into the 4 dimensions, with scores expressed as percentage, where higher numbers indicate greater impairment and less productivity, i.e., worse outcomes, as described in the WPAI-SHP scoring rules.

The scoring rules for the WPAI-SHP are as follows:

Questions:

1 = currently employed

2 = hours missed due to specified problem

3 = hours missed other reasons

4 = hours actually worked

5 = degree problem affected productivity while working

6 = degree problem affected regular activities

Scores:

- Percent work time missed due to problem: [Q2 hours/(Q2 hours +Q4 hours)] *100
- Percent impairment while working due to problem: [Q5 score/10] *100
- Percent overall work impairment due to problem:
- [Q2 hours/(Q2 hours +Q4 hours)+[(1-(Q2 hours/(Q2 hours +Q4 hours))x(Q5 score/10)]] *100
- Percent activity impairment due to problem: [Q6 score/10] *100

Change from Baseline score is derived as post Baseline score minus Baseline score. A negative change score indicates a reduction in the score/improvement for the subject.

8.4 Additional statistical analysis of other efficacy variables

For selected efficacy variables, it is of interest to perform statistical tests and to calculate inferential statistics. As these tests are not part of the multiplicity-controlled procedure, the associated p-values are considered nominal and are not controlled for multiplicity.

For responder variables, the analysis will follow what was specified for the primary analysis. Specifically, a stratified Cochran-Mantel-Haenszel (CMH) test will be used, where region and prior biologic exposure (yes/no) will be stratification variables. The p-value will be based on the CMH test for a general association. Missing values will be imputed using NRI.

For continuous variables, the MI – MCMC / Monotone Regression approach used for other continuous variables will be applied for the imputation model. The analysis model will be based on analysis of covariance (ANCOVA) with fixed effects of treatment, region, and prior biologic exposure and Baseline value as a covariate.

Below is a list of variables for which these nominal p-values will be calculated (with the time points in parentheses). The results of these inferential tests will all be presented in a single table summarizing the testing performed outside of the multiplicity-controlled testing procedure.

- PASI90
  - Bimekizumab vs Adalimumab (Weeks 1, 3, 4, 8, and 12)
- IGA Clear or Almost Clear
  - Bimekizumab vs Adalimumab (Weeks 1, 3, 4, 8, and 12)

- PASI100
  - Bimekizumab vs Adalimumab (Weeks 4, 8, and 12)
- IGA Clear
  - Bimekizumab vs Adalimumab (Weeks 4, 8, 12, and 16)
- PASI75
  - Bimekizumab vs Adalimumab (Weeks 1, 3, and 16)
- PSD score - Pain
  - Bimekizumab vs Adalimumab (Week 16)
- PSD score - Itch
  - Bimekizumab vs Adalimumab (Week 16)
- PSD score - Scaling
  - Bimekizumab vs Adalimumab (Week 16)
- Scalp IGA Clear or Almost Clear (subjects with Baseline Scalp IGA ≥2)
  - Bimekizumab vs Adalimumab (Week 16)
- pp-IGA Clear or Almost Clear (subjects with Baseline pp-IGA ≥2)
  - Bimekizumab vs Adalimumab (Week 16)
- mNAPSI75 response (subjects with Baseline mNAPSI>0)
  - Bimekizumab vs Adalimumab (Week 16)
- mNAPSI90 response (subjects with Baseline mNAPSI>0)
  - Bimekizumab vs Adalimumab (Week 16)
- mNAPSI100 response (subjects with Baseline mNAPSI>0)
  - Bimekizumab vs Adalimumab (Week 16)
- DLQI 0/1 response
  - Bimekizumab vs Adalimumab (Week 16)
- PASI percentage change from Baseline
  - Bimekizumab vs Adalimumab (Weeks 1, 3, 4, 8, 12, and 16)

9 PHARMACOKINETICS AND PHARMACODYNAMICS

9.1 Pharmacokinetics

Pharmacokinetic variables will be analyzed for all subjects in the PK-PPS. Bimekizumab plasma concentrations will be summarized for each treatment at each scheduled visit.

PK summaries will be based on observed values. No imputation will be used. However, if plasma concentration measurements are below the level of quantification (BLQ), then for

calculation of the derived statistics the result will be set to ½ of the lower level of quantification (LLOQ). Descriptive statistics including geometric mean, geometric coefficient of variation, and geometric mean 95% CI if applicable will be calculated if at least 2/3 of the values of interest are above the LLOQ. If this is not the case, only median, minimum, and maximum will be presented.

In addition, geometric mean plasma concentration will be plotted on a linear and log-linear scale by treatment group, and by cumulative antibody status for subjects randomized to bimekizumab.

If the dosing for a visit is +/- 21 days out of window then the plasma concentration from that visit and all subsequent visits will be excluded from the PK summary.

If the PK sampling date is >1 day after the dosing date the data point will be excluded from the PK summary tables and figures.

However, all PK concentrations will be listed.

The PK will be subject to a population PK analysis, pooling data across phase 2 and 3 trials and will be described in a separate analysis plan and reported independently.

9.2 Pharmacodynamics

The relationship between PK and Efficacy and Safety will be explored. The analysis will be described in a separate analysis plan and reported independently.

9.3 Immunogenicity

9.3.1 Anti-nuclear antibodies

Not applicable.

9.3.2 Anti-bimekizumab antibodies

Anti-bimekizumab antibodies (ADAb) will be measured using a three tiered assay approach: screening assay, confirmatory assay and titration assay. Samples confirmed as positive within the confirmatory assay, will be further evaluated in a neutralizing assay to evaluate the potential of the ADAb to neutralize the activity of Bimekizumab (IL17A or IL17F, or both) in-vitro. Samples were taken at baseline, week 4, week 8, week 12, week 16, week 20, week 24, week 36, week 48, week 56 and in subjects who do not enter into the OLE (PS0014) a sample at SFU which is 20 weeks after the last dose.

In subjects randomized to Adalimumab for the first 24 weeks, no ADAb samples will be analysed until week 36, with the exception of the week 24 sample which will be used as a baseline sample.

- Screening, confirmatory and titer cut points of the respective assays will be determined by the bioanalytical laboratory based either on using commercially available drug-naïve samples or on the pre-dose samples of the study itself. In case cut point is being set on commercially available samples, the representative for the study population is being checked. The screening cut point will be used to determine the status of ADA in test sample as above the cut point (ACP) or below the cut point (BCP). Samples presenting ADA levels above the cut point, will be further evaluated in the confirmatory assay, the result of which will be reported as either “confirmed positive” (CP) or “not confirmed positive” (NCP).

- Values ≤ screening cut point (BCP) or above the screening cut point and confirmed negative (NCP) are defined as anti-bimekizumab antibody negative
- Values > predefined screening cut point (ACP) and confirmed positive (CP) are defined as anti-bimekizumab antibody positive
- Confirmed positive samples will be titrated, and the titer (reciprocal dilution factor including MRD) reported. The samples (or a subset) will also be subject to a neutralizing assay to evaluate whether the anti-bimekizumab antibody neutralizes the activity of bimekizumab (IL17 A or IL17F or both) in- vitro. This will be reported subsequent to the main DBL and CSR.

Subject Classification:

- For subjects who are negative at baseline, and antibody negative **at all** sampling points post treatment (including SFU) - **pre ADAb negative - treatment emergent ADAb negative**
- For subjects who are negative at baseline, and antibody positive **at any** sampling point post treatment (including SFU) - **pre ADAb negative - treatment emergent ADAb positive.** If a subject has a missing pre-treatment sample (either missing or insufficient volume) at baseline with one or more ADAb positive post-treatment samples will be also classified as **pre ADAb negative - treatment emergent ADAb positive**
- For subjects who are positive at baseline, and antibody negative **at all** sampling points post treatment (including SFU) - **pre ADAb positive - treatment emergent reduced ADAb**
- For subjects who are positive at baseline, and are positive **at any** sampling point post treatment (including SFU) with titer values of the same magnitude as baseline (i.e. ≤ than a 2.07 fold difference from the baseline value) - **pre ADAb positive - treatment emergent unaffected ADAb positive**
- For subjects who are positive at baseline, and are positive at any sampling point post treatment (including SFU) with increased titer values compared to baseline (above a 2.07 fold difference increase from baseline value which will be defined within the validation of the assay and will be included in the TFLs and/or SAP when available) - **pre ADAb positive - treatment emergent ADAb boosted positive.**
- For Subjects who have a positive pre-treatment sample and some post-treatment samples are missing, while other post-treatment samples are ADAb negative, **the subject will be classed as inconclusive.**

For the interim analysis at Week 56, SFU data will not be included.

Analysis

Immunogenicity will be assessed through summary tables and figures, and listing of individual results by subject. All analyses will be run on the safety population, unless specified otherwise.

- Number and percentage of subjects with anti-bimekizumab antibody level above the specified cut point at the time of each visit, separated by treatment group
- Number and percentage of subjects with anti-bimekizumab antibody level above the specified cut point at any visit during the treatment period, separated by treatment group.

- The time-point of the first occurrence of anti-bimekizumab anti-body positivity during the treatment period) will be summarized for each treatment group. A plot of the Titre versus elapsed time will be plotted
- All individual subject-level anti-bimekizumab antibody results will be listed.
- The number and percentage of subject in each of the 6 categories will be tabulated and separated by treatment group, with an additional category combining subjects who are **pre ADAb negative - treatment emergent ADAb positive** and - **pre ADAb positive - treatment emergent ADAb boosted positive, summarized as Total treatment emergent.** In addition, the number and percentage of subjects who are pre ADAb positive will be summarized (combination of **pre ADAb positive - treatment emergent reduced ADAb, pre ADAb positive - treatment emergent unaffected ADAb positive and pre ADAb positive - treatment emergent ADAb boosted positive**).
- The prevalence of immunogenicity, separated by treatment group, and defined sub-category, will be reported per time point, defined as (cumulative) proportion of subjects having confirmed positive ADAb samples at any point up to and including that time point. Missing samples will not be included in the denominator.
- The time to achieving treatment-emergent ADAb positivity, separated by treatment group and defined sub-category, will be analyzed based on Kaplan-Meier approach, subjects will be considered to have an event at time where treatment emergent ADAb positive is first achieved. Subjects classified as treatment-emergent ADAb negative will be censored at time of last available ADAb result.
- A summary of PASI 90 responder, separated by treatment group and defined sub-category, at weeks 16 and 24 as a function of ADAb titer will be presented graphically. This will be repeated for PASI 100 and 75 responders.
- Individual plots of Bimekizumab Concentrations/ADAb titer and % Change from baseline (all Y-axes) versus time (x-axis) for the full treatment period, including SFU where a patient has not progressed into the OLE. Plots should be labeled and grouped into the 6 sub-categories.
- Spaghetti plots of ADAb titer (Y-axis) against time (X-axis), separated by treatment group for all ADAb positive subjects, including baseline positive subjects.
- Box plots of all Bimekizumab concentrations where ADA titers exists versus ADA titer classification (group 1 > ; group 2 >; group 3 >) presented on a linear scale. Cut-points will be determined on availability of titre data

For purposes of efficacy subgroup analyses based on anti-bimekizumab antibody status, two categories will be used:

- Anti-bimekizumab antibody positive – This is defined as subjects who have anti-bimekizumab antibody levels above the specified cut point on at least 2 time points while on treatment (ie, excluding Baseline, excluding SFU).
- Anti-bimekizumab antibody negative – Subjects who are not defined as anti-bimekizumab positive (as described above) will be defined as anti-bimekizumab antibody negative.

The rationale for requiring at least 2 time points in which anti-bimekizumab antibody levels are above the specified cut point is to exclude those subjects who have only one occurrence of anti-bimekizumab antibody levels during the course of treatment. Including such subjects would increase the number of anti-bimekizumab antibody positive subjects with potentially no impact on efficacy.

10 SAFETY ANALYSES

All safety summaries and listings will be created using all subjects in the SS.

The SS, BKZ Set, BKZ Week 24 Set and MS will be used for summaries of safety data.

It was noted during the conduct of the study that approximately 24 subjects received incorrect Placebo/Active treatment at Visit 20. All subsequent treatments were administered per protocol. The impact of this on safety is expected to be minor and no adjustments will be made to the summaries of safety data. The actual treatments received will be shown in the data listings.

10.1 Extent of exposure

Summaries for exposure will be provided. This consists of a descriptive summary of study medication duration in days. In addition, total study medication duration and time at risk will be summarized in years by treatment group and treatment period (i.e, through initial treatment period, Week 24, and initial and maintenance treatment period). The summary of exposure through the initial and maintenance treatment period will be on BKZ set. The cumulative study medication duration will be summarized for subjects exposed for given durations of time, with four weekly categories through initial treatment period and Week 24. For the cumulative duration through initial and maintenance treatment period the following categories for duration will be used:

- >0 weeks
- >=16 weeks
- >=24 weeks
- >=40 weeks
- >=56 weeks

Definitions for study medication duration and time at risk in days are provided below for each period. Time at risk will be summarized in years. Time at risk in years is calculated by dividing the time at risk in days by 365.25.

Throughout this section, date of last clinical contact for each subject is defined as the maximum of (last visit date including SFU visit, last imputed AE start date, date of study termination or completion, last date of study drug administration).

10.1.1 Exposure during the initial treatment period for SS

Definitions for study medication duration (days) and time at risk (days) during the initial treatment period are provided as follows:

Study medication duration (days)

- Date of last dose of active medication in the Initial treatment Period – date of first dose in the Initial Treatment Period + 28 days for bimekizumab Q4W/ +14 days for adalimumab

Note: If date of last dose of active medication in the Initial Treatment Period + 28 days/ 14 days extends to a date beyond the date of the first dose in the Maintenance Treatment Period, then this calculation reverts to:

– Date of first dose in the Maintenance Treatment Period – date of first dose in the Initial Treatment Period + 1.

Note: For subjects who die during the Initial Treatment Period, if date of last dose of active medication in the Initial Treatment Period + 28 days/ 14 days extends to a date beyond the date of death, then this calculation reverts to:

– Date of death – date of first dose in the Initial Treatment Period + 1.

For subjects who permanently discontinue study treatment:

- Date of last dose of active medication – date of first dose + 28 days for bimekizumab Q4W/ +14 days for adalimumab.

  Note: If date of last dose of active medication + 28 days/ 14 days extends to a date beyond the final visit date (including PEOT, but not including SFU), then this calculation reverts to:

  – Final visit date (including PEOT, but not including SFU) – date of first dose + 1.

Time at risk (days)

- For subjects who complete the final visit of Initial Treatment Period and continue into the Maintenance Treatment Period: Date of first dose in the Maintenance Treatment Period – date of first dose in the Initial Treatment Period + 1.
- For subjects who discontinue on or prior to the final visit of the Initial Treatment Period, use the minimum of the following:
  – Date of last dose of Initial Treatment Period – date of first dose in the Initial Treatment Period + 140 days.
  – The total number of days in the Initial Treatment Period.
  – Date of last clinical contact – date of first dose in the Initial Treatment Period + 1.
- For subjects who die prior to the final visit of the Initial Treatment Period: Date of death – date of first dose in Initial Treatment Period +1.

10.1.2 Exposure through Week 24 (not including Week 24 Visit) for SS

Definitions for study medication duration (days) and time at risk (days) though Week 24 are provided as follows:

Study medication duration (days)

- Date of last dose of active medication prior to Week 24 Visit – date of first dose in the Initial Treatment Period + 28 days for bimekizumab Q4W/ +56 days for bimekizumab Q8W/ +14 days for adalimumab.

Note: If date of last dose of active medication prior to Week 24 + 28 days/ 56 days/ 14 days extends to a date beyond the date of the first dose on or after Week 24, then this calculation reverts to:

– Date of first dose on or after Week 24 – date of first dose in the Initial Treatment Period + 1.

Note: For subjects who die during the first 24 weeks, if date of last dose of active medication prior to Week 24 + 28 days/ 56 days/ 14 days extends to a date beyond the date of death, then this calculation reverts to:

– Date of death – date of first dose in the Initial Treatment Period + 1.

For subjects who permanently discontinue study treatment:

- Date of last dose of active medication – date of first dose + 28 days/ 56 days/ 14 days
- Note: If date of last dose of active medication + 28 days/ 56 days/ 14 days extends to a date beyond the final visit date (including PEOT, but not including SFU), then this calculation reverts to:

Final visit date (including PEOT, but not including SFU) – date of first dose + 1.

Time at risk (days)

- For subjects who complete through Week 24 and continue beyond Week 24: Date of first dose on or after Week 24 – date of first dose in Initial Treatment Period + 1.
- For subjects who discontinue prior to Week 24 visit, use the minimum of the following:
  – Date of last dose prior to Week 24 – date of first dose in the Initial Treatment Period + 140 days.
  – The total number of days up to Week 24.
  – Date of last clinical contact – date of first dose in the Initial Treatment Period + 1.
- For subjects who die prior to Week 24 visit: Date of death – date of first dose in Initial Treatment Period +1.

10.1.3 Exposure during the initial and maintenance period for BKZ Set

Definitions for study medication duration (days) and time at risk (days) through initial and maintenance treatment period are provided as follows:

For subjects randomized to BKZ Q4W/Q4W:

Study medication duration (days)

- Date of last dose of bimekizumab – date of first dose + 28 days

  Note: If date of last dose of bimekizumab+ 28 days extends to a date beyond the final visit date (including PEOT, but not including SFU), then this calculation reverts to:

  – Final visit date (including PEOT, but not including SFU) – date of first dose + 1.

  Note: For subjects who die, if date of last dose of bimekizumab+ 28 days extends to a date beyond the date of death, then this calculation reverts to:

– Date of death – date of first dose + 1.

Time at risk (days):

- For subjects who continue into an open-label study (and, therefore, do not have the SFU visit in the feeder study): Final visit date – date of first dose + 1.
- For subjects who die prior to the final visit: Date of death – date of first dose + 1.
- For all other subjects, use the minimum of the following:
  - Date of last dose– date of first dose + 140 days.
  - Date of last clinical contact – date of first dose + 1.

  Note: This group could include subjects who discontinue early, subjects who complete the Maintenance Period as scheduled but choose not to continue into an open-label study, or subjects who are ongoing in the SFU period at the time of the data snapshot.

For subjects randomized to BKZ Q4W/Q8W:

Study medication duration (days)

- Initial Treatment Period (attributed to BKZ Q4W)
  - Date of last dose of bimekizumab in the Initial treatment Period – date of first dose in the Initial Treatment Period + 28 days

Note: If date of last dose of bimekizumab in the Initial Treatment Period + 28 days extends to a date beyond the date of the first dose in the Maintenance Treatment Period, then this calculation reverts to:

– Date of first dose in the Maintenance Treatment Period – date of first dose in the Initial Treatment Period + 1.

Note: For subjects who die during the Initial Period, if date of last dose of bimekizumab in the Initial Treatment Period + 28 days extends to a date beyond the date of death, then this calculation reverts to:

– Date of death – date of first dose in the Initial Treatment Period + 1.

- Maintenance Treatment Period (attributed to BKZ Q8W)
  - Use the study medication duration algorithm described for Maintenance Treatment Period in 10.1.4.
- Initial and Maintenance Treatment Period combined (attributed to BKZ Total)
  - Date of last dose of bimekizumab – date of first dose + 28 days (if last dose was in Initial Period) or 56 days (if last dose was in Maintenance Period)

Note: If date of last dose of bimekizumab + 28 days (if last dose was in Initial Period) or 56 days (if last dose was in Maintenance Period) extends to a date beyond the final visit date (including PEOT, but not including SFU), then this calculation reverts to:

- Final visit date (including PEOT, but not including SFU) – date of first dose + 1.

Note: For subjects who die, if date of last dose of bimekizumab + 28 days (if last dose was in Initial Period) or 56 days (if last dose was in Maintenance Period) extends to a date beyond the date of death, then this calculation reverts to:

- Date of death – date of first dose + 1.

Time at risk (days):

- Initial Treatment Period (attributed to BKZ Q4W)
  - Date of first dose in Maintenance Treatment Period – Date of first dose in Initial Treatment Period + 1 (Note, this assumes that anyone in this category completes the Initial Treatment Period and doses in the Maintenance Treatment Period).
- Maintenance Treatment Period (attributed to BKZ Q8W)
  - Follow the time at risk algorithm described for the Maintenance Treatment Period in 10.1.4.
- Initial and Maintenance Treatment Period combined (attributed to BKZ Total)
  - For subjects who continue into an open-label study (and, therefore, do not have the SFU visit in the feeder study): Final visit date – date of first dose + 1.
  - For subjects who die prior to the final visit: Date of death – date of first dose + 1.
  - For all other subjects, use the minimum of the following:
    - Date of last dose– date of first dose + 140 days.
    - Date of last clinical contact – date of first dose + 1.

  Note: This group could include subjects who discontinue early, subjects who complete the Maintenance Period as scheduled but choose not to continue into an open-label study, or subjects who are ongoing in the SFU period at the time of the data snapshot.

For subjects randomized to ADA/BKZ Q4W:

Only the duration of Bimekizumab treatment will be derived over this period as the summaries produced are for the BKZ Set (not including any adalimumab treatment).

- Follow the study medication duration and time at risk algorithms described for the Maintenance Treatment Period in 10.1.4.

10.1.4 Exposure during the maintenance period for BKZ Set

Definitions for study medication duration (days) and time at risk (days) during the Maintenance Period (Week 16 to Week 56) are provided as follows:

For subjects randomized to BKZ Q4W/Q4W or BKZ Q4W/Q8W:

Study medication duration (days)

- Date of last dose of bimekizumab in the Maintenance Period – date of first dose in the Maintenance Period + 28 days (or 56 days in the case of Q8W dosing).

  Note: If date of last dose of bimekizumab in the Maintenance Period + 28 days (or 56 days in the case of Q8W dosing) extends to a date beyond the final visit date in the Maintenance Period (including PEOT, but not including SFU), then this calculation reverts to:

  – Final visit date in the Maintenance Period (including PEOT, but not including SFU) – date of first dose in the Maintenance Period + 1.

  Note: For subjects who die, if date of last dose of bimekizumab in the Maintenance Period + 28 days (or 56 days in the case of Q8W dosing) extends to a date beyond the date of death, then this calculation reverts to:

  – Date of death – date of first dose in the Maintenance Period + 1.

Time at risk (days):

- For subjects who continue into an open-label study (and, therefore, do not have the SFU visit in the feeder study): Final visit date in the Maintenance Period – date of first dose in the Maintenance Period + 1.
- For subjects who die prior to the final visit in the Maintenance Period: Date of death – date of first dose in the Maintenance Period + 1.
- For all other subjects, use the minimum of the following:

  – Date of last dose in the Maintenance Period – date of first dose in the Maintenance Period + 140 days.

  – Date of last clinical contact– date of first dose in the Maintenance Period + 1.

  Note: This group could include subjects who discontinue the maintenance period early, subjects who complete the Maintenance Period as scheduled but choose not to continue into an open-label study, or subjects who are ongoing in the SFU period at the time of the data snapshot.

For subjects randomized to ADA/BKZ Q4W:

Study medication duration (days)

Only the duration of Bimekizumab treatment will be derived over this period as the summaries produced are for the BKZ Set (not including any adalimumab treatment).

- Date of last dose of bimekizumab on or after Week 24 – date of first dose on or after Week 24 + 28 days

  Note: If date of last dose of bimekizumab + 28 days extends to a date beyond the final visit date in the Maintenance Period (including PEOT, but not including SFU), then this calculation reverts to:

  – Final visit date in the Maintenance Period (including PEOT, but not including SFU) – date of first dose on or after Week 24 + 1.

  Note: For subjects who die, if date of last dose of bimekizumab in the Maintenance Period + 28 days extends to a date beyond the date of death, then this calculation reverts to:

– Date of death – date of first dose on or after Week 24 + 1.

Time at risk (days):

- For subjects who continue into an open-label study (and, therefore, do not have the SFU visit in the feeder study): Final visit date in the Maintenance Period – date of first dose on or after Week 24 + 1.
- For subjects who die prior to the final visit in the Maintenance Period: Date of death – date of first dose on or after Week 24 + 1.
- For all other subjects, use the minimum of the following:
  - Date of last dose on or after Week 24 – date of first dose on or after Week 24 + 140 days.
  - Date of last clinical contact – date of first dose on or after Week 24 + 1.

  Note: This group could include subjects who discontinue the Maintenance Period early, subjects who complete the Maintenance Period as scheduled but choose not to continue into an open-label study, or subjects who are ongoing in the SFU period at the time of the data snapshot.

10.2 Adverse events

An AE is any untoward medical occurrence in a subject or clinical investigation subject administered a pharmaceutical product that does not necessarily have a causal relationship with this treatment. An AE can therefore be any unfavorable and unintended sign (including an abnormal laboratory finding), symptom, or disease temporally associated with the use of a medicinal (investigational) product, whether or not related to the medicinal (investigational) product.

Adverse events will be coded according to the Medical Dictionary for Regulatory Activities (MedDRA version 19.0). For subjects who switch treatment during the study i.e. those randomized to BKZ Q4W/Q8W or ADA/BKZ, any adverse events that occur after initiation of the new treatment are attributable to the new treatment. If an adverse event occurs on the date of a treatment switch, the event is attributed to the original treatment. The only exception to this is if the AE fulfills the criteria for an anaphylactic reaction as defined in the bimekizumab safety topics of interest document. Specifically, this includes the following:

- Events that fulfill the anaphylaxis criteria for acute events (refer to Appendix A in section 12)
- Events with a reported term containing the term “hypersensitivity (SMQ)”
- Events with an HLT of “Administration site reactions NEC”
- Events with an HLT of “Injection site reactions”

10.2.1 Data considerations

Treatment-emergent AEs (TEAEs) are defined as those AEs that have a start date on or following the first dose of study treatment through the final dose of study treatment + 140 days

(covering the 20-week SFU period). If it is not possible (due to partial dates) to determine whether or not an AE is treatment-emergent then it will be assumed to be a TEAE.

The rules for imputing partial start or stop dates are outlined in Section 4.2.1.2.

If the intensity of an adverse event is unknown, it is considered as severe. If the relationship to study drug is missing, it is considered as related.

AEs will be presented as "number of subjects (percentage of subjects) [number of events]". In this style of output, "[number of events]" will include all cases of an AE including repeat occurrences in individual subjects, while "number of subjects" will count each subject only once.

Subject time at risk represents the time a subject is at risk for having an AE. The definitions for subject time at risk (in days) are outlined in Section 10.1. These definitions will be used for exposure-adjusted AE summaries.

Selected AE summaries will include the exposure adjusted incidence rate (EAIR) with associated 95% confidence interval and the exposure adjusted event rate (EAER).

The EAIR is defined as the number of subjects (n) with a specific AE adjusted for the exposure and will be scaled to 100 patient-years:

$$EAIR = 100 \times n / \sum_{i=1}^{N} \left(T_{Exp(i)}\right)$$

Where $T_{Exp(i)}$ is the exposure time and N is the number of subjects at risk.

If a subject has multiple events, the time of exposure is calculated to the first occurrence of the AE being considered. If a subject has no events, the total time at risk is used.

Exact Poisson 95% confidence intervals for incidence rates are calculated using the relationship between the Poisson and the Chi-square distribution (Ulm, 1990; Fay and Feuer, 1997):

$$LCL = \chi^2_{2n,\alpha/2} / 2$$

$$UCL = \chi^2_{2(n+1),1-\alpha/2} / 2$$

where n is the number of subjects with a specific AE for the incidence rate of interest and is the basis for the number of the degrees of freedom for the chi-square quantile for the upper tail probability χ^2.

The EAER will be the number of AEs including repeat occurrences in individual subjects divided by the total time at risk scaled to 100 patient-years and calculated using:

$$EAER = 100 \times N_{AE} / \sum_{i=1}^{N} \left(T_{Risk(i)}\right)$$

where N_{AE} is the total number of AEs, T_{Risk} is the time at risk for each subject, and N is the total number of subjects at risk.

No confidence interval will be computed for EAER.

Selected summaries for the subjects who switch from adalimumab to bimekizumab will be presented by time of onset (≤70 days or > 70 days from the final dose of adalimumab). Only events starting after the first dose of bimekizumab will be included.

70 days was chosen as it is approximately equal to five half lives of adalimumab.

Selected summaries, as specified in Section 10.2.2, will include the risk difference between bimekizumab and adalimumab. The risk difference is calculated as:

$$RD = IP_{BKZ} - IP_{ADA}$$

where IP_{BKZ} is the incidence proportion for the bimekizumab-treated group and IP_{ADA} is the incidence proportion for the adalimumab group. Note that incidence proportion simply refers to the percentage of subjects within the specified treatment group that experienced a given adverse event.

The standard error for the risk difference is calculated as follows:

$$SE_{RD} = \sqrt{\left(IP_{BKZ} \times \left(\frac{1 - IP_{BKZ}}{n_{BKZ}}\right)\right) + \left(IP_{ADA} \times \left(\frac{1 - IP_{ADA}}{n_{ADA}}\right)\right)}$$

where n_{BKZ} is the number of subjects in the bimekizumab-treated group and n_{ADA} is the number of subjects in the adalimumab group.

The corresponding confidence interval for the risk difference is as follows:

$$CI_{RD} = RD \pm Z_{1-\alpha/2} \times (SE_{RD})$$

where $Z_{1-\alpha/2}$ is the Z statistic for the corresponding level of alpha. For the risk difference confidence intervals calculated in this SAP, 1.96 will be used (corresponding to a two-sided alpha of 0.05 and 95% confidence interval). Risk difference and the corresponding confidence interval will be presented as a percentage.

10.2.2 AE summaries

The following summaries will be provided in five different ways: through Week 24 for the SS, through initial and maintenance treatment period for the BKZ Set, through initial treatment period for the SS, through the maintenance period for the MS (excluding patients randomized to ADA) and from Week 24 to Week 56 for the BKZ Week 24 Set for patients randomized to ADA switching to BKZ (to examine the events occurring within 70 days and >70 days of the treatment switch).

- Incidence of TEAEs – Overview
- Incidence of TEAEs per 100 subject years by SOC, HLT, and PT
- Incidence of Serious TEAEs per 100 subject years by SOC, HLT, and PT

The following summary will be provided in four different ways: through Week 24 for the SS, through initial and maintenance treatment period for the BKZ Set, through initial treatment period for the SS and through the maintenance period for the MS (excluding patients randomized to ADA).

- Incidence of TEAEs Leading to Discontinuation per 100 subject years by SOC, HLT, and PT

The following summaries will be provided in two different ways: through Week 24 for the SS, and through initial and maintenance treatment period for the BKZ Set:

- Incidence of TEAEs Leading to Death by SOC, HLT, and PT
- Incidence of TEAEs by Maximum Relationship by SOC, HLT, and PT
- Incidence of Serious TEAEs by Relationship SOC, HLT, and PT – Note: For EudraCT reporting purposes
- Incidence of Related Serious TEAEs by SOC, HLT, and PT
- Incidence of TEAEs Leading to Death by Relationship by SOC, HLT, and PT – Note: For EudraCT reporting purposes
- Incidence of TEAEs by Maximum Severity, SOC, HLT, and PT
- Incidence of TEAEs by decreasing frequency of PT
- Incidence of TEAEs Above Reporting Threshold of 5% by SOC and PT
- Incidence of Non-Serious TEAEs by SOC, HLT, and PT
- Incidence of Non-Serious TEAEs by Relationship SOC, HLT, and PT
- Incidence of Non-Serious TEAEs Above Reporting Threshold of 5% by SOC and PT
- Incidence of Non-Serious TEAEs Above Reporting Threshold of 5% by Relationship SOC and PT
- Incidence of Related TEAEs by SOC, HLT, and PT
- Incidence of Related TEAEs Above Reporting Threshold of 5% by SOC and PT

The following summary will be provided through Week 24 for the SS:

- Incidence of TEAEs Above Reporting Threshold of 5% with Risk Differences by SOC and PT
- Incidence of Serious TEAEs and Risk Differences by SOC and PT

The following summary will be provided through initial and maintenance treatment period for the BKZ Set:

- Incidence of TEAEs per 100 subject years by SOC, HLT, and PT and by Time of Onset Relative to Anti-bimekizumab Antibody Status

In addition, all summaries of TEAEs based on “100 subject years” should include EAIR (with 95% confidence interval) and EAER.

The tables with risk differences will also be accompanied by figures (dot plots) which show the incidence of the adverse events and corresponding 95% risk difference confidence intervals. These will be ordered by descending order of risk difference (bimekizumab vs adalimumab).

10.2.3 Other Safety topics of interest

The following are AEs of other safety topics of interest that require special statistical analyses. Along with the tables described, there will be a table which displays the risk difference and 95% confidence intervals for each of the topics of interest. A corresponding figure (with dot plots) will be prepared.

10.2.3.1 Infections (serious, opportunistic, fungal and TB)

- **Incidence and Event Rate of Serious Infection TEAEs per 100 subject years by SOC, HLT and PT**

  Serious infections will be identified based on MedDRA classification (SOC "Infections and infestations") using the "Any SAE" table. A separate table does not need to be produced to summarize these events.

- **Incidence and Event Rate of Fungal Infection TEAEs per 100 subject years by SOC, HLT and PT**

  Fungal infections will be summarized in a stand-alone table. The table will include all TEAEs (serious and non-serious) which code into the High Level Group Term "Fungal infectious disorders"

- **Incidence and Event Rate of Opportunistic Infection TEAEs per 100 subject years by SOC, HLT and PT**

  Opportunistic infections (including tuberculosis) will be summarized in a stand-alone table. The table will include all TEAEs identified using UCB-defined search criteria (refer to Excel spreadsheet on "OI - MedDRA v19.0.xlsx" in "Bimekizumab-Safety-Topics-of-Interest.docx").

The following steps are followed for identifying and reviewing opportunistic infections:

Identification Process

The two steps below outline two ways in which opportunistic infections (or potential opportunistic infections) can be identified:

Step 1: Refer to column B of the spreadsheet which identifies the Preferred Terms (PTs) to be classified as opportunistic infections using either a single 'x' or a double 'xx'.

TEAEs which code to a PT flagged with a single 'x' need to also be serious to be considered an opportunistic infection.

All TEAEs which code to a PT flagged with a double 'xx' are considered to be an opportunistic infection, regardless of seriousness.

Step 2: Refer to column C of the spreadsheet which identifies the PTs that need to be evaluated on a case-by-case basis by the study physician to determine whether or not it is an opportunistic infection. If Column C has a single 'x', then the corresponding preferred term should be flagged for case-by-case review by the study physician.

Review Process

Opportunistic infections for a given study will be reviewed on the following occasions:

At quarterly Infectious Disease Committee (IDC) Meetings, listings will be produced for each study (see details below) and reviewed by the corresponding study physician ahead of the IDC Meeting. These listings will be posted as part of the broader SSD deliverable to a folder named for the given quarter (eg, 2018Q4) on the SharePoint. They should be based on the same data cut as the one used for SSD and should be delivered at the same time as the SSD outputs. The IDC will then agree on the final adjudication for each potential opportunistic infection.
For each study, a final listing for opportunistic infections (in the format described below) will be produced and agreed upon between the study physician and the IDC prior to finalizing the database.

In each of the circumstances described above, the study programming team will produce an Excel listing that will be provided to the project lead statistician, project lead programmer, and to the study physician (who will subsequently provide it to the IDC). The Excel listing will contain the following columns (using the descriptions below as the column headings in the Excel listing):

Study ID
Unique Subject ID
AE Term (Verbatim)
AE Preferred Term
AE System Organ Class
AE High Level Term
AE Low Level Term
Date of Onset
Outcome of Adverse Event
Date of Outcome
TEAE Flag
Serious Adverse Event?
Relationship to Study Medication
Intensity
Action Taken with IMP
Opportunistic Infection – Automatic
Opportunistic Infection – Manual Review
Flag
Data Cut Date
Opportunistic Infection – Final Adjudication
Note the following about the final 5 variables in this listing:

Opportunistic Infection – Automatic: This is flagged as “Y” if the criteria for automatic selection as described in “Step 1” of the identification process are met.
Opportunistic Infection – Manual Review: This is flagged as “Y” if the criteria for case-by-case selection as described in “Step 2” of the identification process are met.
Flag – This has a value of either “NEW” or “OLD”. It is marked as “NEW” if the event is appearing for the first time in that run of the listing. Otherwise, if it has appeared previously, it is marked as “OLD”. Unique records are determined by USUBJID AESPID for purposes of identifying whether an event has been modified from a previous run.
Date – Only for cases where Flag is “NEW”, this field will be populated with the data cut date for that particular run of the listing.

Opportunistic Infection – Final Adjudication – For new events, this is always left blank by the programmers. It should be completed by the study physician/IDC for every event that appears in the listing. For events adjudicated as opportunistic, the field should be populated with a "Y". Following each review by the study physician and IDC, the Opportunistic Infection – Final Adjudication column will be completed (as described above), and the spreadsheets for each study will be returned to the study programming team via e-mail (coordinated by the IDC secretary). Then, for subsequent runs of the listing, the study programming teams will incorporate adjudications from previous runs.

10.2.3.2 Malignancies

- **Incidence and Event Rate of Malignancy (including Unspecified Tumours) TEAEs per 100 subject years by SOC, HLT and PT**

This table will be based on the criteria Standardized MedDRA Query (SMQ) = "Malignant or unspecified tumours (SMQ)"

- **Incidence and Event Rate of Malignancy TEAEs per 100 subject years by SOC, HLT and PT**

This table will be based on the criteria SMQ = "Malignant tumours (SMQ)".

SMQ search should include all TEAEs which code to a PT included in the Scope=Narrow group within each SMQ.

Note that the events included in the "Malignant tumours" table will be a subset of the events included in the "Malignant or unspecified tumours" table. While the "Malignant tumours (SMQ)" is most relevant, "Malignant or unspecified tumours (SMQ)" must be reviewed for potential malignancies.

The output table based on the SMQ="Malignant or unspecified tumours (SMQ)" will include 2 different overall incidence rows:

- The first overall incidence row will summarize "Any malignancies (including unspecified)" and this row will summarize the incidence of all TEAEs flagged for inclusion in the table, regardless of the High Level Term (HLT) it codes to.
- The second overall incidence row will summarize "Any malignancy (including unspecified, excluding non-melanomic skin cancers)" and this row will summarize the incidence of AEs flagged for inclusion in the table, excluding those which code to an HLT of "skin neoplasms malignant and unspecified (excl melanoma)".

10.2.3.3 Major adverse cardiac events (MACE)

- **Incidence and Event Rate of Major Cardiac Event (MACE) TEAEs per 100 subject years by SOC, HLT and PT**

Major adverse cardiac events (MACE) will be presented in a table. The classification of an event as MACE is determined by an external cardiovascular event adjudication committee.

A separate table and listing will present the adjudicated cardiovascular events by type. For each cardiovascular event type (24 total), the individual PTs which fall within each event type will be summarized.

Additionally, a listing of all events identified for potential review by the cardiovascular event adjudication committee will be produced. This listing will indicate whether each event was escalated to the committee for formal review/adjudication.

10.2.3.4 Neutropenia

- **Incidence and Event Rate of Neutropenia TEAEs per 100 subject years by SOC, HLT and PT**

This table will be based on the following PTs (regardless of seriousness):

- Autoimmune neutropenia
- Band neutrophil count decreased
- Cyclic neutropenia
- Febrile neutropenia
- Idiopathic neutropenia
- Neutropenia
- Neutropenic infection
- Neutropenic sepsis
- Neutrophil count decreased

10.2.3.5 Suicidal Ideation and Behavior

- **Incidence and Event Rate of SIB-Adjudicated Neuropsychiatric TEAEs per 100 subject years by SOC, HLT and PT**

An external neuropsychiatric adjudication committee will evaluate potential neuropsychiatric events and determine whether any of those events were associated with suicidal ideation and behavior (SIB). If an event is adjudicated as SIB, further information will be provided. A table and listing for SIB events as determined by the adjudication committee will be included.

Additionally, a listing of all events identified for potential review by the neuropsychiatric adjudication committee will be produced. This listing will indicate whether each event was escalated to the committee for formal review/adjudication. A separate listing will also be produced to summarize the adjudicated results of all events escalated to the full committee.

10.2.3.6 Inflammatory bowel disease

- **Incidence and Event Rate of Inflammatory Bowel Disease TEAEs per 100 subject years by SOC, HLT and PT**

These events will be presented in a stand-alone table which will include all TEAEs which code into the HLT of "Colitis excl infective".

Inflammatory bowel disease events will be summarized stratified by subjects with or without a previous medical history of inflammatory bowel disease. Previous medical history of inflammatory bowel disease will be determined using the information recorded on the Extra-Articular Assessment at Screening CRF page ("Does subject have a history of IBD?").

10.2.3.7 Hypersensitivity (including anaphylaxis)

- **Incidence and Event Rate of Anaphylactic Reaction TEAEs per 100 subject years by SOC, HLT and PT**

A separate table will be prepared based on the MedDRA anaphylaxis Algorithm (see Appendix 1) for acute anaphylactic events (reported on the same day as when an injection was administered or one day after). An AE glossary table will also be produced to summarize the MedDRA coding for these events. The glossary table will include the following fields: reported term, PT, LLT, HLT, and SOC.

- **Incidence and Event Rate of Hypersensivity Reaction TEAEs per 100 subject years by SOC, HLT and PT**

A separate table will be prepared to summarize hypersensitivity events, identified using the SMQ "Hypersensitivity (SMQ)". All TEAEs which code to a PT included in the Scope=Narrow search will be included in this table. An AE glossary table will also be produced to summarize the MedDRA coding for these events. The glossary table will include the following fields: reported term, PT, LLT, HLT, and SOC.

Furthermore, injection site reactions will be evaluated based on the any TEAE table (no separate table needed) by looking under the following HLTs: "Administration site reactions NEC" and "Injection site reactions".

10.2.3.8 Hepatic events

- **Incidence and Event Rate of Hepatic Events TEAEs per 100 subject years by SOC, HLT and PT**

A table for hepatic events will be created based on the SMQ of "Drug related hepatic disorders - comprehensive search (SMQ)". However, these 2 sub-SMQs are to be excluded: "Liver neoplasms, benign (incl cysts and polyps) (SMQ)" and "Liver neoplasms, malignant and unspecified (SMQ)".

Note that all AEs meeting the above criteria are to be included. It should not be limited to events that the investigator determined to be related to study drug.

Cases of Hy's Law will be reported separately in a liver function test table (with adjudication for "potential DILI" confirmation).

A by-subject listing of all AE of safety topics of interest by type of safety topics of interest will be provided.

10.3 Clinical laboratory evaluations

Laboratory values, including markedly abnormal laboratory values will be presented descriptively by treatment group for SS up to Week 56. Markedly abnormal values and shifts from baseline (Week 0) will be presented for the SS through Week 24 and BKZ set through initial and maintenance treatment period.

The markedly abnormal tables and those based on CTCAE grade will be produced only for selected laboratory variables.

For tables where data are summarized by visit, unscheduled and repeat visits will not be summarized, but these data will be included in listings. For tables where multiple measurements over a period of time are considered (as in shift tables), unscheduled and repeat visits will be considered as long as they were collected in the period being summarized. All summaries will be presented in SI units and will be based on observed case values. In the case where laboratory values are below the lower limit of quantification, then these will be set to the midpoint between 0 and the lower limit of quantification for the purpose of summarizing the data. The following summaries are required:

- A summary of the absolute and change from Baseline values in each laboratory variable by treatment group and visit
- A summary of the number and percentage of subjects experiencing markedly abnormal values at any time while on treatment (assessment on or following the first dose of study treatment through the final dose of study treatment + 140 days) by laboratory variable and treatment group
- A summary of the number and percentage of subjects with a given CTCAE grade (0, 1, 2, 3, or 4) based on minimum/maximum post-baseline value by laboratory variable and treatment group
- A shift table of the number and percentage of subjects experiencing CTCAE grade 0, 1, 2, 3 or 4 values (as applicable) at Baseline to minimum/maximum post-Baseline CTCAE grade, by laboratory variable and treatment group

For laboratory summaries through week 24, the cut off will be the earliest of week 24 or last dose + 140 days.

For laboratory shift tables through week 24, the treatment groups BKZ 320mg Q4W, BKZ 320mg Q4W/Q8W, BKZ Total and ADA will be presented.
For laboratory shift tables through initial and maintenance treatment period, the treatment groups BKZ 320mg Q4W, BKZ 320mg Q4W/Q8W, ADA/ BKZ 320 mg Q4W and BKZ Total will be presented. For subjects randomized to ADA, only data after the switch to Bimekizumab will be considered. The last value prior to the switch will be considered as baseline.

A by-subject listing of all laboratory data will be provided. This listing will be presented by treatment group and will include: center, subject identifier, age, sex, race, weight, visit, laboratory variable, result (with abnormal values flagged as “L” or “H” accordingly) and unit.

Markedly abnormal values are defined as those with a severity of Grade 3 and above based on the common terminology criteria for adverse events (CTCAE) criteria (U.S. Department of Health and Human Services 2010). Definitions of markedly abnormal values using the Grade 3 cut points are given in the tables below for age ranges of ≥17 years (Table 10–1 for markedly abnormal biochemistry values and Table 10–2 for markedly abnormal hematology values). Tables summarizing markedly abnormal values should include a summary (counts and percentages) of markedly abnormal labs observed at any time while on treatment (ie, treatment-emergent markedly abnormal [TEMA]). For this summary, Baseline values and values observed more than 140 days after the last administration of study medication are not considered. The laboratory results classified as Grade 3 or Grade 4 will be summarized and listed separately.

Summaries in the ADA/BKZ arm will be provided to examine markedly abnormal labs occurring within 70 days and >70 days of the treatment switch. The time since the treatment switch is calculated as the number of days since the final dose of adalimumab. Only lab values recorded after the start of bimekizumab treatment will be included. 70 days was chosen as it is approximately equal to five half lives of adalimumab.

Table 10–1: Definitions of Markedly Abnormal Biochemistry Values

Parameter name	Conventional		Standard		Abnormal Designation
	Unit	Criteria	Unit	Criteria	
Creatinine	mg/dL	>3.0 x ULN	mmol/L	>3.0 x ULN	AH
Glucose	mg/dL	<40 >250	mmol/L	<1.7 >13.9	AL AH
Calcium	mg/dL	>12.5 <7.0	mmol/L	>3.1 <1.75	AH AL
Magnesium	mg/dL	>3.0 <0.9	mmol/L	>1.23 <0.4	AH AL
Potassium	mmol/L	>6.0 <3.0	mmol/L	>6.0 <3.0	AH AL
Sodium	mmol/L	>155 <130	mmol/L	>155 <130	AH AL

Table 10–2: Definitions of Markedly Abnormal Hematology Values

Parameter name	Conventional		Standard		Abnormal Designation
	Unit	Criteria	Unit	Criteria	
Hemoglobin	g/dL	<8.0 >4.0 above ULN	g/L	<80 >40 above ULN	AL AH
Lymphocytes Absolute	10^9/L	<0.5 >20.0	10^9/L	<0.5 >20.0	AL AH
Neutrophils Absolute	10^9/L	<1.0	10^9/L	<1.0	AL
Platelets	10^9/L	<50	10^9/L	<50	AL
WBC/Leukocytes	10^9/L	<2.0 >100	10^9/L	<2.0 >100	AL AH

Abbreviations: AH=abnormal high; AL=abnormal low; ALT= alanine aminotransferase; AST = aspartate aminotransferase; dL = deciliter; GGT: gamma-glutamyltransferase; L = liter; mg = milligram; mmol = millimoles; µg = microgram; ULN = upper limit of normal.

The following criteria will be considered for summaries of markedly abnormal liver function tests (LFTs):

- AST: >3xULN, >5xULN, >8xULN, >10xULN, >20xULN
- ALT: >3xULN, >5xULN, >8xULN, >10xULN, >20xULN

- AST or ALT: >3xULN, >5xULN, >8xULN, >10xULN, >20xULN
- Bilirubin: >1.5xULN, >2xULN

In addition, a table will be produced to summarize potential Hy's Law cases. The following definition will be used in that table:

- [AST ≥3xULN or ALT ≥3xULN] and Total Bilirubin ≥2xULN in the absence of ALP ≥2xULN

In order to meet the above criteria, a subject must experience the elevation in bilirubin and ALT or AST (and the absence of ALP elevation) at the same visit. For example, a subject who experiences a ≥2 x ULN elevation of bilirubin at one visit and a ≥3xULN elevation in ALT (or AST) at a subsequent visit has not fulfilled the Hy's law criteria.

10.4 Vital signs, physical findings, and other observations related to safety

10.4.1 Vital signs

The following vital signs variables should be summarized: systolic blood pressure (mmHg), diastolic blood pressure (mmHg), body temperature (°C) and heart rate (beats/min). The following summary will be provided:

- A summary of the absolute and change from Baseline value for each vital sign variable by treatment group and visit through initial and maintenance treatment period for the SS
- A summary of the number and percentage of subjects experiencing at least one markedly abnormal value for a vital sign variable as defined in Table 10–3, by treatment group, through Week 24 for the SS and through initial and maintenance treatment period for the BKZ set

Table 10–3: Definitions of Markedly Abnormal Blood Pressure Values

Parameter (unit)	Markedly Abnormal Low	Markedly Abnormal High
Systolic blood pressure (mmHg)	<90 and a decrease from Baseline of ≥20	>180 and an increase from Baseline of ≥20
Diastolic blood pressure (mmHg)	<50 and a decrease from Baseline of ≥15	>105 and an increase from Baseline of ≥15

A by-subject listing of all vital signs data should be provided. This listing should be presented by treatment group and will include: center, subject identifier, age, sex, race, weight, visit, vital sign variable and result (with abnormal values flagged as "L" or "H" accordingly).

10.4.2 Physical examination

Abnormal results of the physical examination together with details of abnormalities: abnormality clinically significant or not, will be listed by subject and visit.

10.4.3 Electrocardiograms

Electrocardiogram data will be analyzed by treatment group and visit for SS.

A summary of the number and percentage of subjects with normal and abnormal ECG results, as determined by the central reader, will be presented for all applicable visits.

The following ECG variables will be summarized (absolute values and change from baseline) by visit: QTcF, RR, PR, QRS and QT.

QTc outliers are defined as QTcF values following dosing that are greater than 450 ms or are increases from baseline greater than 30 ms. QTcF outliers will be highlighted in the data listings and summarized using the following categories:

- Values >450 ms, >480 ms, >500 ms
- Increase from baseline of >30 ms, increase from baseline of >60 ms, increase from Baseline of >90 ms
- Values >450 ms and increases of >30 ms. Values >500 ms and increases of >60 ms

The number and percentage of subjects who meet the ECG outlier criteria at any assessment post first dose will be summarized.

Two separate by-subject listings of all 12-lead ECG data will be provided based on interpretation from central reader and from site.

10.4.4 Other safety variables

10.4.4.1 Assessment and management of TB and TB risk factors

A summary of the number and percentage of subjects with negative, positive, and indeterminate IGRA (Interferon-Gamma Release Assay) results at all applicable visits will be presented for the entire study.

A by-subject listing of the "Evaluation of signs and symptoms of tuberculosis" questionnaire data and IGRA results will be provided.

A by-subject listing of the result of chest x-ray for tuberculosis will be provided by treatment.

10.4.4.2 Electronic Columbia Suicide Severity Rating Scale (eC-SSRS)

eC-SSRS questionnaire will be self-administered by the subject and assessed by trained study personnel. This scale will be used to assess suicidal ideation and behavior that may occur during the study. Results of the eC-SSRS will be summarized using the number of subject and percentage with (i) events in suicide behavior, (ii) suicidal ideation, (iii) suicidal behavior and ideation, and (iv) self-injurious behavior without suicidal intent.

Suicidal ideation is defined as an event in any of the following 5 categories:

- Wish to be dead
- Non-specific active suicidal thoughts
- Active suicidal ideation with any methods (not plan), without intent to act
- Active suicidal ideation with some intent to act, without specific plan
- Active suicidal ideation with specific plan and intent

Suicidal behavior is defined as an event in any of the following 4 categories:

- Actual attempt
- Interrupted attempt
- Aborted attempt
- Preparatory acts or behavior

Suicidal behavior or ideation is defined as an event in any of the above 9 categories.

Self-injurious behavior without suicidal intent is defined as an event in the category non-suicidal self-injurious injuries.

The incidence of subjects with suicidal ideation, suicidal behavior, suicidal ideation or behavior, and self-injurious behavior will be summarized by treatment group through week 24 and through initial and maintenance treatment period.

A by-subject listing of the eC-SSRS questionnaire data will be provided by treatment group.

10.4.4.3 Pregnancy testing

Pregnancy testing will consist of serum testing at the Screening. The pregnancy test will be urine at all other visits.

A by-subject listing of the pregnancy test data will be provided by treatment group.

10.4.4.4 Childbearing potential and Lifestyle

Childbearing potential and lifestyle will be collected at Screening. A by-subject listing will be provided for all the subjects screened.

10.4.5 Patient Health Questionnaire 9 (PHQ-9)

The PHQ-9 is a multipurpose instrument for screening, diagnosing, monitoring, and measuring the severity of depression. The PHQ-9 scores for depression range from 0 to 27 with higher scores indicating worse state. A score of 5-9 is considered to be minimal symptoms of depression. If any of the 9 questions are missing, then the score is treated as missing. A score of 10 to 14 is considered minor depression, dysthymia, or mild major depression. A score of 15 to 19 is considered to indicate moderately severe major depression, and a score ≥20 is considered to be severe major depression.

MI-MCMC/monotone regression method will be used to impute missing data. If the imputation model cannot converge, last observation carried forward (LOCF) will be used.

Change from Baseline in PHQ-9 is derived as post-Baseline score minus Baseline score. It will be summarized by treatment group through initial and maintenance treatment period for the SS and for the BKZ Week 24 Set for subjects switching from adalimumab to bimekizumab.

11 REFERENCES

Fay MP, Feuer EJ. Confidence intervals for directly standardized rates: A method based on the gamma distribution. Statistics in Medicine 1997; 16(7):791-801.

Felson DT, Anderson JJ, Boers M, Bombardier C, Chernoff M, Fried B, et al. The American College of Rheumatology preliminary core set of disease activity measures for rheumatoid arthritis clinical trials. The Committee on Outcome Measures in Rheumatoid Arthritis Clinical Trials. Arthritis Rheum. 1993;36(6):729-40.

Food and Drug Administration. Guidance for Industry. Drug-induced liver injury: premarketing clinical evaluation. US Dept of Health and Human Services, Center for Drug Evaluation and Research, Center for Biologics Evaluation and Research, 07/2009.

Greenland S, Robins JM. Estimation of common effect parameter from sparse follow up data. Biometrics 1985;41:55-68.

Husni ME, Qureshi AA, Keonig AS, Pedersen R, Robertson D. Utility of the PASE questionnaire, psoriatic arthritis (PsA) prevalenece and PsA improvement with anti-TNF therapy: results from the PRISTINE trial. J Dermatolog Treat. 2014;25(1):90-5.

Mallinckrodt CH, Watkin JG, Molenberghs G, Carroll RJ. Choice of the primary analysis in longitudinal clinical trials. Pharmaceutical Statistics 2004; 3:161-69.

Mallinckrodt CH, Lin Q, Lipkovich I, and Molenberghs G. A structured approach to choosing estimands and estimators in longitudinal clinical trials. Pharmaceutical Statistics 2012; 11:456-61.

Mallinckrodt CH. Preventing and Treating Missing Data in Longitudinal Clinical Trials: A Practical Guide. New York, NY: Cambridge University Press; 2013.

Maruish, M. E., User's manual for the SF-36v2 Health Survey (3rd ed.). Lincoln, RI: Quality Metric Incorporated, 2011.

Ratitch, B., Lipkovich, I., O'Kelly, M. Combining Analysis Results from Mutiply Imputed Categorical Data, PharmaSUG, 2013, SP03.

Rubin D.B. 1987. Multiple Imputation for Nonresponse in Surveys. New York: John Wiley and Sons.

Thomas, D. Roland and Cyr, Andre. Applying Item Response Theory Methods to Complex Survey Data. SSC Annual Meeting, May 2002.

Ulm, K., A simple method to calculate the confidence interval of a standardized mortality ratio. American Journal of Epidemiology, 1990; 131(12), 373-375

U.S. Department of Health and Human Services, Common Terminology Criteria for Adverse Events (CTCAE), Version 4.0, June 2010.

12 APPENDIX

12.1 Appendix A: MedDRA algorithmic approach to anaphylaxis

The SMQ *Anaphylactic reaction* consists of three parts:

- A **narrow search** containing PTs that represent core anaphylactic reaction terms

SMQ Anaphylactic reaction (SMQ)
- PT Anaphylactic reaction
- PT Anaphylactic shock
- PT Anaphylactic transfusion reaction
- PT Anaphylactoid reaction
- PT Anaphylactoid shock
- PT Circulatory collapse
- PT Dialysis membrane reaction
- PT Kounis syndrome
- PT Shock
- PT Shock symptom
- PT Type I hypersensitivity

- A **broad search** that contains additional terms that are added to those included in the narrow search. These additional terms are signs and symptoms possibly indicative of anaphylactic reaction and categorized in B, C or D
  - Cat B

PT B Acute respiratory failure
PT B Asthma
PT B Bronchial oedema
PT B Bronchospasm
PT B Cardio-respiratory distress
PT B Chest discomfort
PT B Choking
PT B Choking sensation
PT B Circumoral oedema
PT B Cough
PT B Cyanosis
PT B Dyspnoea
PT B Hyperventilation
PT B Irregular breathing
PT B Laryngeal dyspnoea
PT B Laryngeal oedema
PT B Laryngospasm
PT B Laryngotracheal oedema
PT B Mouth swelling
PT B Nasal obstruction
PT B Oedema mouth
PT B Oropharyngeal spasm
PT B Oropharyngeal swelling
PT B Respiratory arrest
PT B Respiratory distress
PT B Respiratory dyskinesia
PT B Respiratory failure
PT B Reversible airways obstruction
PT B Sensation of foreign body
PT B Sneezing
PT B Stridor
PT B Swollen tongue
PT B Tachypnoea
PT B Throat tightness
PT B Tongue oedema
PT B Tracheal obstruction
PT B Tracheal oedema
PT B Upper airway obstruction
PT B Wheezing

– Cat C

- Allergic oedema
- Angioedema
- Erythema
- Eye oedema
- Eye pruritus
- Eye swelling
- Eyelid oedema
- Face oedema
- Flushing
- Generalised erythema
- Injection site urticaria
- Lip oedema
- Lip swelling
- Nodular rash
- Ocular hyperaemia
- Oedema
- Periorbital oedema
- Pruritus
- Pruritus allergic
- Pruritus generalised
- Rash
- Rash erythematous
- Rash generalised
- Rash pruritic
- Skin swelling
- Swelling
- Swelling face
- Urticaria
- Urticaria papular

– Cat D

- Blood pressure decreased
- Blood pressure diastolic decreased
- Blood pressure systolic decreased
- Cardiac arrest
- Cardio-respiratory arrest
- Cardiovascular insufficiency
- Diastolic hypotension
- Hypotension

- An **algorithmic approach** which combines a number of anaphylactic reaction symptoms in order to increase specificity. A case must include one of the following where both occur on either the same day as when an injection was administered or one day after:
  – A narrow term or a term from Category A;
  – A term from Category B - (Upper Airway/Respiratory) AND a term from Category C - (Angioedema/Urticaria/Pruritus/Flush);
  – A term from Category D - (Cardiovascular/Hypotension) AND [a term from Category B - (Upper Airway/Respiratory) OR a term from Category C - (Angioedema/Urticaria/ Pruritus/Flush)]
- Hypersensitivity events will be identified using the “Hypersensitivity (SMQ)”. All TEAEs which code to a PT included in the Scope=Narrow search will be included.

12.2 Appendix A: Definition of CTCAE grades

Table 12–1: Definitions of CTCAE grades by biochemistry parameter

Parameter (unit)	Definition	Unit	Grade 1	Grade 2	Grade 3	Grade 4
Creatinine	High	mmol/L	>ULN-1.5 x ULN	(>1.5 – 3.0) x ULN	(>3.0 – 6.0) x ULN	6.0 x ULN
Sodium	Low	mmol/L	130-<LLN	N/A	120-<130	<120
Sodium	High	mmol/L	>ULN-150	>150-155	>155-160	>160
Potassium	Low	mmol/L	3.0-<LLN	3.0-<LLN	2.5-<3.0	<2.5
Potassium	High	mmol/L	>ULN-5.5	>5.5-6.0	>6.0-7.0	>7.0
Calcium	Low	mmol/L	2.0-<LLN	1.75-<2.0	1.5-<1.75	<1.5
Calcium	High	mmol/L	>ULN-2.9	>2.9-3.1	>3.1-3.4	>3.4
Magnesium	Low	mmol/L	0.5-<LLN	0.4-<0.5	0.3-<0.4	<0.3
Magnesium	High	mmol/L	>ULN-1.23	N/A	>1.23-3.30	>3.30

Subjects who meet the decreased potassium criteria of 3.0-<LLN, which is specified as the criteria for both CTCAE grade 1 and grade 2, will be counted as CTCAE grade 2.

Table 12–2: Definitions of CTCAE grades by hematology parameter

Parameter (unit)	Definition	Unit	Grade 1	Grade 2	Grade 3	Grade 4
Hemoglobin	Low	g/L	100-<LLN	80-<100	<80	N/A
Hemoglobin	High	g/L	>0-20 above ULN	>20-40 above ULN	>40 above ULN	N/A
Platelets	Low	10^9/L	75-<LLN	50-<75	25-<50	<25
WBC	Low	10^9/L	3-<LLN	2-<3	1-<2	<1
WBC	High	10^9/L	N/A	N/A	>100	N/A
Lymphocytes	Low	10^9/L	0.8-<LLN	0.5-<0.8	0.2-<0.5	<0.2
Lymphocytes	High	10^9/L	N/A	>4-20	>20	N/A
Neutrophils	Low	10^9/L	1.5-<LLN	1.0-<1.5	0.5-<1.0	<0.5

13 AMENDMENT(S) TO THE STATISTICAL ANALYSIS PLAN

13.1 Amendment 1

Rationale of the Amendment

The main purpose of this amendment is to achieve consistency with other SAPs of the program.

Modification and changes

Global Changes

The following changes were made throughout the SAP

- For the following endpoints, the definition of response has been clarified by adding the requirement of an improvement of at least two categories:
    - Scalp IGA response (clear or almost clear **with at least a 2-category improvement from Baseline)** for subjects with scalp PSO at Baseline
    - Palmoplantar Investigator's Global Assessment (pp-IGA) response (clear or almost clear **with at least a 2-category improvement from Baseline**) for subjects with palmoplantar PSO at Baseline
- "Through Week 16" was changed to "through initial treatment period".
- "Through Week 56" was changed to "through initial and maintenance treatment period".
- "Patient Symptom Diary response" was changed to "Patient Symptom Diary score".
- The following method of imputation of missing data for the primary efficacy variable will no longer be used.
    - MI – MCMC / Reference-based imputation: Using multiple imputation methodology, intermittent missing data are imputed based on the MCMC method, and monotone missing data are imputed using an imputation model based on adalimumab (reference) data.

Specific Changes

Change #1

List of abbreviation

The following abbreviations have been added:

LFT	liver function test
MS	Maintenance Set

The following abbreviations have been removed:

DAP	Data Analysis Plan
HLGT	Higher level group term
LTB	Latent tuberculosis
SOP	Standard operating procedure

Change #2

Section 2.2.1.3 Other efficacy variables

The following variables have been added:

- **mNAPSI90 response (defined as a subject that achieves at least a 90% reduction from Baseline in the mNAPSI score)**
- **mNAPSI100 response (defined as a subject that achieves a 100% reduction from Baseline in the mNAPSI score)**
- **Patient Symptom Diary response rates**

The following variable was removed in efficacy section:

- **Change from baseline in Patient Health Questionnaire (PHQ)-9 scores**

Change #3

Section 2.2.3.2 Other safety variables

The following variable was added in safety section:

- **Change from baseline in Patient Health Questionnaire (PHQ)-9 scores**

The following variable was removed from safety section:

- **Laboratory: urinalysis**

And

The following text was removed:

"Physical examination findings considered clinically significant changes since the physical examination at the Screening Visit will be recorded as AEs."

Change #4

Section 2.3.2.2 Double Blind Initial Treatment Period

Text added:

Subjects will be followed in the double-blinded fashion. Subjects may receive placebo injections at certain visits in order to blind the IMP.
Subjects will be classified as completing the initial treatment period if they complete the Week 16 Visit without early withdrawal from the study or if they start treatment in the maintenance treatment period. The start of the maintenance treatment period marks the end of the initial treatment period.

Change #5

Section 2.3.2.3 Dose-blind Maintenance Treatment Period

After the 16-week Initial Treatment Period, subjects will enter the 40-week Maintenance Treatment Period. Treatment during the Maintenance Treatment Period will start at Week 16 and subjects will return to the clinic Q4W or Q8W through initial and maintenance treatment period 56. The IMP will be administered in the clinic by sc injection at the time points specified in the schedule of study assessments.

Has been changed to

After the 16-week Initial Treatment Period, subjects will enter the 40-week Maintenance Treatment Period. Treatment during the Maintenance Treatment Period will start **with the first study drug administration on or after** Week 16 and subjects will return to the clinic Q4W or Q8W through **Week 56**. The IMP will be administered in the clinic by sc injection at the time points specified in the schedule of study assessments.

Change #6

Section 3.1 General presentation of summaries and analyses

- For summaries of efficacy and safety endpoints, unless otherwise specified: summarize percentages based only on those subjects with observed data for the variable being summarized. As the denominator may be different from the number of subjects in the analysis set being considered, the denominator should be displayed in the table. The general format for displaying this will be "n/Nobs (%)."

Has been changed to

- For summaries of efficacy and safety endpoints, unless otherwise specified: summarize percentages based only on those subjects with observed data for the variable being summarized. As the denominator may be different from the number of subjects in the analysis set being considered, the denominator should be displayed in the table. The general format for displaying this will be "n/N**sub** (%)."

The following text was added at end of section 3.1:

Per protocol, visit windows of ±3 days from the first dose to Week 24 and ±7 days from Week 28 to Week 56 are permissible. For the SFU Visit, visit window is ±7 days from final dose. All by-visit summaries will contain nominal (i.e. scheduled) visits only. Unscheduled visits will not be mapped to scheduled visits except for some assessments that may occur within a 3-day time window of a scheduled visit. In that case, the assessment will be mapped to the corresponding scheduled visit and will be used for the analysis. This will only occur for some vendor data.

For missing PASI and IGA assessments at Week 16, available data from all visits (unscheduled or scheduled but without a planned assessment) within the 3 days window will be included.

Change #7

Section 3.2 Definition of Baseline Value

Unless otherwise specified, the last valid measurement before study medication administration in the Initial period will be used as the Baseline value.

If a Baseline visit measurement is missing, and a Screening visit measurement is available, the Screening value will be utilized as Baseline. If a scheduled Baseline assessment is taken on the same day as first administration of study medication and no time is recorded, it will be assumed to have been taken prior to study medication.

For subjects who start with adalimumab and switch treatment at the Week 24 Visit, an additional bimekizumab Baseline defined as the measurement at Week 24 will be used. If a Week 24 Visit measurement is missing, the subject will be excluded from change from baseline summaries for

that measurement. The variables listed below will also be summarized based on bimekizumab Baseline:

Has been changed to

A Baseline value for a subject is defined as the latest measurement for that subject up to and including the day of administration of first study medication, unless otherwise stated. If a Baseline assessment is taken on the same day as first administration of study medication, it is eligible to be used as the Baseline value, even in the case that the time of the assessment is recorded as taking place after the time of first study medication administration. This is considered acceptable as this measurement is still the best representation of the Baseline value of the given assessment since it is highly unlikely that the study medication could have an impact on any measurement in such a short period of time. However, such cases should be rare as study center personnel are instructed to do all assessments at the Baseline visit prior to administering study medication. One exception to this rule is plasma concentration. If Baseline plasma concentration is measured at a time after the first administration of study medication, then it should not be eligible to be considered as a Baseline plasma concentration. Such cases should be discussed with the quantitative clinical pharmacologist. If a Baseline measurement is missing or not collected, and a Screening value is available, the Screening value will be utilized as Baseline instead.

For subjects who start with adalimumab and switch treatment at the Week 24 Visit, an additional bimekizumab Baseline defined as the measurement at Week 24 will be used. If a Week 24 Visit measurement is missing, baseline will be the latest measurement prior to the first dose of bimekizumab. The variables listed below will also be summarized based on bimekizumab Baseline:

And

The following variables have been removed from the list of variables that will have an additional Week 24 Baseline defined:

- PASI50
- PASI75
- PASI90
- PASI100
- Percentage of BSA
- Product of IGA and BSA (IGAxBSA)
- PGADA
- Patient Global Assessment of PSO
- mNAPSI
- SF-36
- EQ-5D-3L VAS
- (PHQ) 9

- WPAI-SHP V2.0

Only the following variables remained in the list with text added:

- PASI score
- DLQI
- Laboratory data (only in shift tables)

Note that for any laboratory value that occurs on the day of treatment switch, that lab value will be attributed and summarized for the treatment they were on previously. For subjects that switch from Adalimumab to Bimekizumab, the Baseline value is the laboratory value of the day of treatment switch.

Baseline values for component scores should be computed using components from the same visit where the relevant measurements were recorded prior to dosing. For example, if the Screening visit has all of the components, but the Baseline visit is missing one or more components, the Baseline value for the component score should be calculated using the Screening visit values.

Change #8

Section 3.5 Analysis sets

The following precision was added:

For all analysis sets, each unique subject is only counted once, even if the same subject has more than one unique subject identification.

and

The Bimekizumab Week 24 Set and the Maintenance Set were added:

3.5.6 Bimekizumab Week 24 Set

The BKZ Week 24 Set will consist of all subjects who have received at least 1 dose of BKZ on or after Week 24.

3.5.7 Maintenance Set

The Maintenance Set (MS) will consist of all subjects who have received at least 1 dose of active IMP (bimekizumab or adalimumab) in the Maintenance Treatment Period (at Week 16 or later).

Change #9

Section 3.6 Treatment assignment and treatment groups

Summaries based on BKZ set will be provided by "BKZ 320mg Q4W", "BKZ 320mg Q4W/Q8W" and "ADA/BKZ 320mg Q4W" treatment groups for disposition, baseline assessments and efficacy. The ADA/BKZ 320mg Q4W group will only be presented from Week 28 to Week 56 in the by-visit summaries.

Has been changed to

Summaries based on BKZ **Week 24** set will be provided by "BKZ 320mg Q4W", "BKZ 320mg Q4W/Q8W" and "ADA/BKZ 320mg Q4W" treatment groups for disposition, baseline

assessments and efficacy. The ADA/BKZ 320mg Q4W group will be presented from Week 28 to Week 56 in the by-visit summaries **for the PASI scores and the DLQI only**.

And

Text was added:

Safety summaries on the BKZ Week 24 set will be provided for subjects switching from Adalimumab to Bimekizumab.

Summaries based on the Maintenance set will be provided by "BZK 320mg Q4W", "BKZ 320mg Q4W/Q8W" and "BKZ Total".

Change #10

Section 3.7 Center pooling strategy and 4.2.1.6 Missing data algorithms

The following pooling strategy has been added to sections 3.7 and 4.2.1.6:

In the event that the percentage of randomized subjects is less than 10% in either of the Asia/Australia or Western Europe regions, the two regions will be combined as a geographic region stratum for efficacy modeling, so that there are no modeling convergence issues across efficacy variables.

Change #11

Section 3.10 Changes to protocol-defined analyses

The following text has been added to section 3.10:

For the following endpoints, the definition of response has been clarified:

- **Scalp IGA response (clear or almost clear with at least a 2-category improvement from Baseline) for subjects with scalp PSO at Baseline**
- **Palmoplantar Investigator's Global Assessment (pp-IGA) response (clear or almost clear with at least a 2-category improvement from Baseline) for subjects with palmoplantar PSO at Baseline**

The following endpoints are not listed in the protocol, but have been added to the SAP in order to achieve consistency with other studies from the program:

- **mNAPSI75/90/100 response (defined as a subject that achieves at least a 75%, 90% and 100% reduction from Baseline in the mNAPSI score)**
- **Percent of subjects with absolute PASI score ≤1, ≤2, ≤3, and ≤5**
- **Percent of subjects with absolute BSA=0%, ≤1%, ≤3%, and ≤5%**
- **The sensitivity analysis for the primary endpoint using MI-MCMC/reference based methods detailed in the protocol has been removed from the analysis plan because it could generate bias due to the reference arm being the active comparator.**

The following additional changes were made:

- **The calculation of nominal p-values has been added for selected other efficacy variables**
- **Modifications have been made to the text regarding the multiple imputation algorithm for clarity**
- **Change from Baseline in PHQ-9 score is listed an efficacy variable in the protocol. However, it is considered as a safety variable in this SAP.**
- **Selected TEAE tables were added to include calculation of risk differences.**

Urinalysis summaries were removed from the plan.

Change #12

Section 4.2.1: Handling of missing data for efficacy variables

The following sentence was added:

If a subject discontinued study treatment without terminating from the study and still continues with scheduled assessments, all efficacy data after discontinuation of study treatment will be treated as missing and subject to imputation as applicable.

Change #13

Section 4.2.1.6 Missing data algorithms

This section has been changed from:

MI – MCMC / Monotone Regression

Investigators will be given discretion to discontinue study treatment for subjects not responding in order to provide them with alternative therapy to treat the condition. Therefore, in many cases, missing efficacy data due to study treatment discontinuation should be dependent on the observed efficacy scores, but independent of unobserved data. This would be consistent with a missing at random (MAR) pattern of missingness. To investigate the efficacy results under the assumption of data being MAR, a multiple imputation method will be applied as follows:

1. Create a data set, one for each treatment group, of subjects with observed values and those needing estimation by multiple imputation. The intermittent missing PASI/IGA values in each data set (ie, missing values for a given subject that has available data before and after the missing time point) will be filled in using the MCMC method, with a total of 100 sets of imputations being performed. For monotone missing data (ie, where all subject data is missing after a given time point), monotone regression will be used to impute missing data. Again, this will be based on 100 sets of imputations. In both cases, biologic exposure, geographic region and PASI/IGA values at Baseline and at each post-Baseline visit (in chronological order) will be included in the imputation model. Note that PASI scores at earlier visits will also be used as predictors for the model of PASI at later visits. The resulting data sets for each treatment arm will be combined into one complete data set based on each of the 100 imputations.

Note: The imputation model will only allow continuous variables. Therefore, prior biologic exposure and region will be re-coded as indicator variables. For prior biologic exposure, this will simply be 0 for biologic-naïve subjects and 1 for biologic-exposed subjects. For region, which has 4 levels, one indicator variable will be defined as 0 for regions other than North America and 1 for North America. Three more indicator variables will be defined similarly replacing North America with Central/Eastern Europe, Western Europe, and Asia/Australia respectively. In order to achieve model convergence, prior biologic exposure may be dropped from the model. If convergence is still not obtained, then region may also be dropped from the model.

2. For each complete imputed data set, the dichotomous responder variable (PASI90/IGA 0 or 1) based on the PASI/IGA scores will be computed. Each complete imputed data set will then be analyzed based on the stratified CMH test.

   Note: For derivation of PASI90 response, the PASI value at Week 16 in the imputed data sets will be compared directly to the observed Baseline PASI value to determine whether or not a reduction of at least 90% was achieved. If values outside of the pre-defined range of values for PASI (0-72) are imputed, they will be cut off as appropriate after the multiple imputation procedure but before deriving the responder variable. For example, an imputed PASI value of -0.5 would be changed to 0 before deriving the PASI90 responder variable.

   Note: Standard rounding rules will be applied to the imputed IGA values in order to derive the binary IGA 0/1 responder variable. For example, if a subject has an IGA score imputed as 1.4 (and assuming a Baseline IGA score of 3), this imputed value would be rounded down to 1, and the minimum change from Baseline of 2 would have been met. Therefore, this subject would be considered an IGA 0/1 responder.

3. The Week 16 results from the specified statistical analysis (e.g. stratified CMH, logistic regression) of each of the 100 imputed data sets will be combined for overall inference using Rubin's rules, which account for the uncertainty associated with the imputed values (Rubin, 1987). This will be done using SAS PROC MIANALYZE.

Some key points to consider relative to the calculation of the odds ratios and corresponding confidence intervals are noted below:

If logistic regression is used, the estimates of the odds ratios from the logistic regression model in step 3 follow a log-normal distribution, and a log transformation is needed to normalize these estimates since the procedures for combining results from multiple imputed datasets are based on the assumption that the statistics estimated from each imputed dataset are normally distributed. Therefore, the log of the odds ratio estimates from the logistic regression model are used when combining into a single inference (the use of PROC MIANALYZE in step 3). Appropriate transformations to the standard errors should also be made in order to get the correct confidence intervals.

For other continuous efficacy variables, MI will be used to impute missing data when possible. If the imputation model cannot converge, LOCF will be used. The MI procedure will also be similar to that described above with the following differences: 1) the imputation

model will use the change from Baseline (instead of actual) values by visit and no dichotomization will be necessary; 2) instead of using the stratified CMH test or logistic regression, the imputed data sets will be combined and simple means and standard errors will be calculated using Rubin's rules (via SAS PROC MIANALYZE). For calculation of other descriptive statistics such as the median, min and max, Rubin's rules do not apply. Multiple imputation estimates will be computed by simply averaging the estimates from the multiple repetitions of the imputation algorithm. This is the same approach that will be used when summarizing continuous secondary efficacy variables by subgroup.

To the following (changes in bold):

MI – MCMC / Monotone Regression

Investigators will be given discretion to discontinue study treatment for subjects not responding in order to provide them with alternative therapy to treat the condition. Therefore, in many cases, missing efficacy data due to study treatment discontinuation should be dependent on the observed efficacy scores, but independent of unobserved data. This would be consistent with a missing at random (MAR) pattern of missingness. To investigate the efficacy results under the assumption of data being MAR, a multiple imputation method will be applied as follows:

1. Create a data set, one for each treatment group **(note that a separate imputation procedure must be invoked in SAS for each treatment group as the seed cannot be set for by groups beyond the first when using a by statement),** of subjects with observed values and those needing estimation by multiple imputation. The intermittent missing PASI/IGA values in each data set (ie, missing values for a given subject that has available data before and after the missing time point) will be filled in using the MCMC method, with a total of 100 sets of imputations being performed. **The seed used for these imputations will be 852 (note that all other multiple imputation procedures described in this SAP related to MCMC / Monotone Regression analyses will use this same seed as well).** For monotone missing data (ie, where all subject data is missing after a given time point), monotone regression **then** will be used to impute missing data. **This will be based on the 100 sets of imputations already created using the MCMC method such that there will be 100 imputations in total.** In both cases, biologic exposure, geographic region and PASI/IGA values at Baseline and at each post-Baseline visit (in chronological order, **see notes below about visits to include for different analysis sets)** will be included in the imputation model. Note that PASI scores at earlier visits will also be used as predictors for the model of PASI at later visits. The resulting data sets for each treatment arm will be combined into one complete data set based on each of the 100 imputations.

Note: The imputation model will only allow continuous variables. Therefore, prior biologic exposure and region will be re-coded as indicator variables. For prior biologic exposure, this will simply be 0 for biologic-naïve subjects and 1 for biologic-exposed subjects. For region, which has 4 levels, one indicator variable will be defined as 0 for regions other than North America and 1 for North America. **Two more indicator variables will be defined similarly replacing North America with Central/Eastern Europe, and Western Europe respectively. An indicator variable for Asia/Australia is not needed as the fourth region**

will be adequately represented by the other region indicator variables all being 0. In the event that the number of randomized subjects is less than 10% in either of the Asia/Australia or Western Europe regions, then the two regions will be combined as a geographic region stratum for efficacy modeling, so that there are no modeling convergence issues across efficacy variables. In order to achieve model convergence, prior biologic exposure may be dropped from the model. If convergence is still not obtained, then region may also be dropped from the model.

2. For each complete imputed data set, the dichotomous responder variable (PASI90/IGA 0 or 1) based on the PASI/IGA scores will be computed. Each complete imputed data set will then be analyzed based on the stratified CMH test.

Note: For derivation of PASI90 response, the PASI value at Week 16 in the imputed data sets will be compared directly to the observed Baseline PASI value to determine whether or not a reduction of at least 90% was achieved. If values outside of the pre-defined range of values for PASI (0-72) are imputed, they will be cut off as appropriate after the multiple imputation procedure but before deriving the responder variable. For example, an imputed PASI value of -0.5 would be changed to 0 before deriving the PASI90 responder variable.

Note: Standard rounding rules will be applied to the imputed IGA values in order to derive the binary IGA 0/1 responder variable. **In addition, if values outside of the pre-defined range of values for IGA (0-4) are imputed, they will be cut off as appropriate after the multiple imputation procedure but before deriving the responder variable.** For example, if a subject has an IGA score imputed as 1.4 (and assuming a Baseline IGA score of 3), this imputed value would be rounded down to 1, and the minimum change from Baseline of 2 would have been met. Therefore, this subject would be considered an IGA 0/1 responder.

Additional ranges of values for other variables are defined in Table 4-2.

Table 4–2: Imputation allowable ranges by variable

Variable	Minimum value	Maximum value	Integer values only
PASI	**0**	**72**	**No**
IGA	**0**	**4**	**Yes**
PSD item	**0**	**10**	**No**
Scalp IGA	**0**	**4**	**Yes**
mNAPSI	**0**	**130**	**No**
BSA	**0**	**100**	**Yes**
IGAxBSA	**0**	**400**	**Yes**
DLQI	**0**	**30**	**Yes**
PGADA	**0**	**100**	**Yes**
SF-36	**0**	**100**	**No**

Variable	Minimum value	Maximum value	Integer values only
EQ-5D-3L VAS	**0**	**100**	Yes
WPAI dimension scores	**0**	**100**	No for variables "Percent work time missed due to problem" and "Percent overall work impairment due to problem". Yes for variables "Percent impairment while working due to problem" and "Percent activity impairment due to problem" These two variables can only take values that are multiples of 10"
PHQ-9	**0**	**27**	**Yes**

3. The Week 16 results from the specified statistical analysis (e.g. stratified CMH, logistic regression) of each of the 100 imputed data sets will be combined for overall inference using Rubin's rules, which account for the uncertainty associated with the imputed values (Rubin, 1987). This will be done using SAS PROC MIANALYZE.

Note that this procedure indicates that the imputation model will be applied for each treatment group separately in PROC MI, which will be the default method. However, in the event there are computational challenges with the imputation model (eg, due to a standard deviation of 0 for responses of a given imputation), it is acceptable to modify the imputation model to include treatment as a variable in the model rather than running a separate model for each treatment group. It should be noted that doing so assumes that treatment does not interact with any of the other variables in the imputation model.

Some further detail on the imputation methods:

1. **<u>Randomized Set</u>: PROC MI with a separate data set for each of the 3 randomized groups (BKZ 320mg Q4W, BKZ 320mg Q8W, ADA). For BKZ 320mg Q4W and BKZ 320mg Q8W, include all assessment visits from Baseline to Week 56. For the ADA group, include only assessments from Baseline to Week 24. Based on the Randomized Set. NOTE: Do not run the combined BKZ**

320mg group through PROC MI. You will get the combined group using the imputed values of the individual Q4W and Q8W groups.

2. **<u>Bimekizumab Week 24 Set</u>: PROC MI with a separate data set for each of the 3 randomized groups (BKZ 320mg Q4W, BKZ 320mg Q8W, ADA) including all assessment visits from Baseline to Week 56. Based on the Bimekizumab Week 24 Set. NOTE: Do not run the combined BKZ 320mg group through PROC MI. These tables will only present the 3 treatment sequences (no BKZ Total needed).**

Some key points to consider relative to the calculation of the odds ratios and corresponding confidence intervals are noted below:

If stratified CMH or logistic regression are used, the estimates of the odds ratios from the logistic regression model in step 3 follow a log-normal distribution, and a log transformation is needed to normalize these estimates since the procedures for combining results from multiple imputed datasets are based on the assumption that the statistics estimated from each imputed dataset are normally distributed. Therefore, the log of the odds ratio estimates from the logistic regression model are used when combining into a single inference (the use of PROC MIANALYZE in step 3). Appropriate transformations to the standard errors **and p-values** should also be made in order to get the correct confidence intervals**. For the CMH test using the p-value for the general association the Wilson-Hilferty transformation should be used (Ratitch, 2013).**

Missing data for continuous primary or secondary efficacy variables will be imputed using MI as appropriate. The MI procedure for continuous variables will be similar to that described above with the following differences:

1. **The absolute value of the given variable will be imputed. Once imputation has been performed across the 100 iterations specified, any values outside of the range of the given variable will be truncated accordingly.**
2. **The change from Baseline values will be computed based on the complete data sets.**
3. **The analysis model will be based on ANCOVA (see above) as opposed to the CMH test.**

For other continuous efficacy variables, MI will be used to impute missing data when possible. If the imputation model cannot converge, LOCF will be used. The MI procedure will also be similar to that described above with the following differences: 1) the imputation model will use the change from Baseline (instead of actual) values by visit and no dichotomization will be necessary; 2) instead of using the stratified CMH test or logistic regression, the imputed data sets will be combined and simple means and standard errors will be calculated using Rubin's rules (via SAS PROC MIANALYZE). For calculation of other descriptive statistics such as the median, min and max, Rubin's rules do not apply. Multiple imputation estimates will be computed by simply averaging the estimates from the multiple repetitions of the imputation algorithm. This is the same approach that will be used when summarizing continuous secondary efficacy variables by subgroup.

Change #14

Section 4.4 Multicenter Studies

Sentence was added:

If convergence still cannot be achieved, this analysis will not be performed.

Change #15

Section 4.8 Examination of subgroups

The following text was added:

Antibody positivity is the only subgroup that is not determined by Baseline data. It will be presented in a separate table.

And

In addition, in order to assess whether an early response to treatment is predictive of response at later time points, subgroup analyses will be performed on PASI90/100 and IGA over time using the following early response subgroups:

- PASI75 responders at Week 4 (yes, no) to predict PASI90/100 and IGA through Week 16
- PASI90 responders at Week 4 (yes, no) to predict PASI90/100 and IGA through Week 16

Has been changed to

In addition, in order to assess whether an early response to treatment is predictive of response at later time points, subgroup analyses will be performed **for the RS** on PASI90/100 and IGA over time using the following early response subgroups:

- PASI75 responders (OC) at Week 4 (yes, no) to predict PASI90/100 and IGA (NRI) through initial and maintenance treatment period
- PASI90 responders (OC) at Week 4 (yes, no) to predict PASI90/100 and IGA (NRI) through maintenance treatment period

Change #16

Section 5.1 Subject Disposition

In the paragraph below, the population BKZ Week 24 Set and MS were added:

The disposition of subjects for all subjects screened will include the number of subjects included in each analysis set (ES, RS, SS, FAS, BKZ Set, **BKZ Week 24 Set, MS**, PPS and PK-PPS) overall and by site.

The following text was added:

A subject is considered to have completed the Week 16 initial treatment period if they have either completed the Week 16 visit or received study medication from Week 16 onwards. Similar for Week 24.

Change #17

Section 5.2 Protocol deviations

PDs will be also presented through initial treatment period.

Change #18

Section 6.2 Other Baseline Characteristics

The following variables were added:

- PGADA score (PGADA=0, PGADA>0)
- DLQI total score (DLQI=0, DLQI>0)
- Nail involvement (yes, no)
- Scalp involvement (yes, no)
- Palmoplantar involvement (yes, no)
- Prior primary failure to biologic (yes/no)
- PSD items: Pain, Itch, Scaling

The following sentence was added:

Baseline nail, scalp, and palmoplantar involvement are based on the number of subjects achieving mNAPSI>0, Scalp IGA>0, and pp-IGA>0, respectively.

Change #19

7 MEASUREMENTS OF TREATMENT COMPLIANCE

Threshold was changed from ≤80% and >80% to <75% and ≥75%.

Change #20

8 EFFICACY ANALYSES

The following sentence was added:

Additional summaries through the initial and maintenance treatment period will be produced for subjects who received bimekizumab from Week 24 onwards (BKZ week 24 set) in all three randomized treatment groups.

And

In addition, for subjects who started with adalimumab and switched treatment at the Week 24 Visit, change from Baseline (at Week 24) will be evaluated from Week 28 to Week 56 (see Section 3.2).

Was changed to:

In addition, for subjects who started with adalimumab and switched treatment at the Week 24 Visit, change from Baseline (at Week 24) will be evaluated from Week 28 to Week 56 **for PASI score and DLQI score only** (see Section 3.2).

Change #21

Section 8.1.1 Derivations of co-primary efficacy variables

The following text was added:
For missing PASI and IGA assessments at Week 16, available data from all visits (unscheduled or scheduled but without a planned assessment) within the 3 days window will be included.

Change #22

Section 8.1.2 Primary analysis of the co-primary efficacy variables

The following text in bold was added:

For the assessment of noninferiority, a noninferiority margin of 10% will be used and evaluated based on the confidence interval for the stratified Cochran Mantel-Haenszel risk difference between bimekizumab and adalimumab. The Wald method will be used to calculate the confidence interval. A non-inferiority margin of 10% has been selected as this is considered to be a clinically relevant difference that could influence the choice of interventions used to treat chronic plaque PSO. Therefore, a difference within the 10% non-inferiority margin would suggest a similar impact on efficacy between the treatments.

The evaluation of superiority will use pairwise treatment comparisons based on the CMH test using the p-value for the general association.

The odds ratio and associated confidence interval (CI) based on the Wald test will be presented. **If one of the treatment groups has 0 or very low response where CMH can no longer be used, the logit method will be applied instead.**

To calculate the stratified Mantel-Haenszel risk difference, the method of Greenland and Robins (1985) is used. For each combination of strata, a 2x2 table of treatment group and response is created. A theoretical 2x2 table for a given stratum is shown below, where n = a+b+c+d.

Treatment Group	Response	
	Yes	**No**
Bimekizumab	a	c
Adalimumab	b	d

Given that structure, the stratified Mantel-Haenszel risk difference, standard error, and two-sided (1-α)*100% confidence interval may be written as follows:

$$RD_{MH=}\frac{\sum_i((a_i*(b_i+d_i)/n_i)\ -(b_i*(a_i+c_i)/n_i)}{\sum_i((a_i+c_i)*(b_i+d_i))/n_i}$$

$$SE_{MH} = \sqrt{\frac{\sum_i \left\{ \frac{[a_i * c_i * (b_i + d_i)^3] + [b_i * d_i * (a_i + c_i)^3]}{(a_i + c_i) * (b_i + d_i) * n_i^2} \right\}}{\left\{ \sum_i \left[\frac{(a_i + c_i) * (b_i + d_i)}{n_i} \right] \right\}^2}}$$

$$CI_{MH} = RD_{MH} \pm \text{probit}\big(1 - (\alpha/2)\big) * SE_{MH}$$

For the assessment of non-inferiority of bimekizumab to adalimumab, the lower 97.5% confidence limit for the stratified Mantel-Haenszel risk difference will be considered. If that value is greater than -10%, then non-inferiority will have been established.

Change #23

Section 8.2.3 Sensitivity analyses of the secondary efficacy variables

The following text was added:

It was noted during the conduct of the study that approximately 25 subjects received incorrect Placebo/Active treatment at Visit 20. All subsequent treatments were administered per protocol. The main analyses of the endpoints at Week 24 will include all subjects in the RS, however the potential impact of this on the Week 24 secondary endpoints (PASI90, PASI100 and IGA response) will be assessed through sensitivity analyses excluding these subjects from the analysis population.

Change #24

Section 8.3 Statistical analysis of other efficacy variables

By visit summaries will include the bimekizumab and adalimumab treatment groups. There will be 2 key approaches for understanding the data after Week 24:

- Summary of data for the RS through Week 56 based on randomized treatment group including bimekizumab Q4W, bimekizumab Q4W/Q8W and bimekizumab total. The adalimumab group will only be summarized up to Week 24
- Summary of Maintenance treatment period data among responders only for PASI90, PASI100, IGA (0/1), and DLQI respectively;
  - For Week 16 responders in bimekizumab Q4W, bimekizumab Q8W and bimekizumab total
  - For Week 24 responders in ADA/BKZ arm

Has been changed to

By visit summaries will include the bimekizumab and adalimumab treatment groups. There will be some key approaches for understanding the data after Week 24:

- Summary of data for the RS through initial and maintenance treatment period based on randomized treatment group including bimekizumab Q4W, bimekizumab Q4W/Q8W and bimekizumab total. The adalimumab group will only be summarized up to Week 24

- Summaries through the initial and maintenance treatment period for subjects who received bimekizumab from Week 24 onwards (BKZ week 24 set) in all three randomized treatment groups.
- For subjects who started with adalimumab and switched treatment at the Week 24 Visit, change from Baseline (at Week 24) will be evaluated from Week 28 to Week 56 for PASI score and DLQI score only (see Section 3.2).
- Summary of Maintenance treatment period data among responders only for PASI90, PASI100, IGA (0/1), and DLQI respectively;
  - For Week 16 responders in bimekizumab Q4W, bimekizumab Q8W and bimekizumab total (Maintenance Set)
  - For Week 24 responders in ADA/BKZ arm (Bimekizumab Set)

No adjustments to the analysis sets will be made for the subjects who were incorrectly dosed (see Section 8.2.3)

Change #25

Section 8.3.1.2 Time to PASI50, PASI75, PASI90 and PASI100 response through initial treatment period and through Week 24

The following sentences were added (changes in bold):

All visits up to and including Week 16 (including unscheduled visits) will be considered.

For the PASI response through Week 16 initial treatment period variables, subjects who discontinue study treatment without achieving a given PASI response prior to Week 16 visit will be censored **at the date of the last observed PASI assessment on or prior to the date of study treatment discontinuation.** Subjects who reach the Week 16 Visit without achieving the given response will be censored at the date of the **last observed PASI assessment on or prior to** Week 16 Visit.

Subject will be censored at baseline (Day 0) if there is no Baseline PASI assessment or no post Baseline PASI assessment.

Change #26

Section 8.3.5 Dermatology Life Quality Index (DLQI)

A clarification was added about DLQI Total score.

The DLQI **total score** is calculated by adding the score of each question. The maximum score is 30, and the minimum score is 0. The higher the score, the more quality of life is impaired.

And

A clarification was added at the end of final bullet.

Percent of subjects achieving a MICD in DLQI is defined as the number of subjects with improvement from Baseline score of 4 or more divided by the number of subjects in RS. **Only subjects with a Baseline DLQI of 4 or greater will be considered for this analysis.**

Change #27

Section 8.3.8 Patient Symptom Diary (PSD)

Baseline in PSD score is defined as the average score of the last 7 days before study medication administration at Baseline visit.

Has been changed to:

The Baseline value is also based on a weekly average (up to and including the first dose date) and will be computed in the same manner.

In cases where there is more than one diary record completed on a particular day, all available records within the 7-day window (including any double entries on one day) will be used in the calculation of weekly average scores. If this results in having more than 7 available scores to calculate the weekly average, the 7 records closest to the visit will be used. The diary entry completion date/time stamp will be used to determine this.

And

In addition, each of the PSD scores will be characterized in terms of the cumulative percent of subjects demonstrating a pre-specified point improvement at Week 16.

Has been changed to:

In addition, each of the PSD scores will be characterized in terms of the cumulative percent of subjects demonstrating a pre-specified point improvement at each visit.

Change #28

Section 8.3.9 Modified Nail Psoriasis Severity Index (mNAPSI) score

If 1 or 2 response items scored on the 0 to 1 scale are missing, the missing response(s) will be imputed by the average of the available responses. Otherwise, the total mNAPSI score will be set to missing.

Has been changed to:

If any of the 7 response items that contribute to mNAPSI is present, while other items are missing (ie, partial mNAPSI data), then the missing items are assumed to be 0 for the mNAPSI calculation. In some cases, the data may be captured in such a way that only non-zero component scores are present in the database. Again, those components that are not present are simply assumed to be 0 for the mNAPSI calculation.

And

Variables mNAPSI90 and mNAPSI100 and their definition were added in text below:

An mNAPSI75 responder is defined as a subject who achieved at least a 75% improvement from Baseline in the mNAPSI score. **mNAPSI90 and mNAPSI100 are defined accordingly.** The proportion of mNAPSI75/**90/100** responders over time will be summarized for each treatment group.

Change #29

Section 8.3.11: Short Form – 36 Items Health Survey (SF-36)

The version of the SF-36 was provided.

The SF-36 will be used using QualityMetric's Health Outcomes™ Scoring Software **version 5.1 or later.**

Change #30

Section 8.3.12 Euro-Quality of Life 5-Dimensions, 3 levels (EQ-5D-3L)

The following sentence was clarified:

Responses to EQ-5D-3L will be summarized based on OC only as primary analysis. No imputation is applied **to responses to EQ-5D-3L but is applied to EQ-5D-3L VAS scores**.

Change #31

New section added:

8.4 Additional statistical analysis of other efficacy variables

For selected efficacy variables, it is of interest to perform statistical tests and to calculate inferential statistics. As these tests are not part of the multiplicity-controlled procedure, the associated p-values are considered nominal and are not controlled for multiplicity.

For responder variables, the analysis will follow what was specified for the primary analysis. Specifically, a stratified Cochran-Mantel-Haenszel (CMH) test will be used, where region and prior biologic exposure (yes/no) will be stratification variables. The p-value will be based on the CMH test for a general association. Missing values will be imputed using NRI.

For continuous variables, the MI – MCMC / Monotone Regression approach used for other continuous variables will be applied for the imputation model. The analysis model will be based on analysis of covariance (ANCOVA) with fixed effects of treatment, region, and prior biologic exposure and Baseline value as a covariate.

Below is a list of variables for which these nominal p-values will be calculated (with the time points in parentheses). The results of these inferential tests will all be presented in a single table summarizing the testing performed outside of the multiplicity-controlled testing procedure.

- PASI90
  - Bimekizumab vs Adalimumab (Weeks 1, 3, 4, 8, and 12)
- IGA Clear or Almost Clear
  - Bimekizumab vs Adalimumab (Weeks 1, 3, 4, 8, and 12)
- PASI100
  - Bimekizumab vs Adalimumab (Weeks 4, 8, and 12)

- IGA Clear
  - Bimekizumab vs Adalimumab (Weeks 4, 8, 12 and 16)
- PASI75
  - Bimekizumab vs Adalimumab (Weeks 1, 3, and 16)
- PSD score - Pain
  - Bimekizumab vs Adalimumab (Week 16)
- PSD score - Itch
  - Bimekizumab vs Adalimumab (Week 16)
- PSD score - Scaling
  - Bimekizumab vs Adalimumab (Week 16)
- Scalp IGA Clear or Almost Clear (subjects with Baseline Scalp IGA ≥2)
  - Bimekizumab vs Adalimumab (Week 16)
- pp-IGA Clear or Almost Clear (subjects with Baseline pp-IGA ≥2)
  - Bimekizumab vs Adalimumab (Week 16)
- mNAPSI75 response (subjects with Baseline mNAPSI>0)
  - Bimekizumab vs Adalimumab (Week 16)
- mNAPSI90 response (subjects with Baseline mNAPSI>0)
  - Bimekizumab vs Adalimumab (Week 16)
- mNAPSI100 response (subjects with Baseline mNAPSI>0)
  - Bimekizumab vs Adalimumab (Week 16)
- DLQI 0/1 response
  - Bimekizumab vs Adalimumab (Week 16)
- PASI percentage change from Baseline
  - Bimekizumab vs Adalimumab (Weeks 1, 3, 4, 8, 12, and 16)

Change #32

Section 9.1 Pharmacokinetics

In addition, geometric mean plasma concentration will be plotted by treatment group, and by cumulative antibody status for subjects randomized to bimekizumab.

PK concentrations will also be listed.

Has been changed to

In addition, geometric mean plasma concentration will be plotted by treatment group, and by cumulative antibody status for subjects randomized to bimekizumab **on linear and log linear scale.**

If the dosing for a visit is +/- 21 days out of window, then the plasma concentration from that visit and all subsequent visits will be excluded from the PK **summary.**

If the PK sampling date is >1 day after the dosing date the data point will be excluded from the PK summary tables and figures.

However, all PK concentrations will be listed.

Change #33

Section 9.3.2 Anti-bimekizumab antibodies

The following text was deleted:

In addition to the ADAb classifications, subjects will also receive an overall neutralizing (nADA) classification for each nADA assay separately, inclusive of baseline and post-baseline results, on the nADA assay results:

- nADA negative: no nADA positive samples at baseline or post-baseline
- nADA positive: one or more positive samples at baseline or post-baseline
- Missing: relevant nADA samples are missing, e.g. if subject had samples selected for nADA testing based on their ADA levels, but there was insufficient sample left for nADA testing.

And

The following text was added:

For the interim analysis at Week 56, SFU data will not be included.

And

- For subjects who are positive at baseline, and are positive **at any** sampling point post treatment (including SFU) with titer values of the same magnitude as baseline (i.e. ≤ then a predefined fold difference from the baseline value) - **pre ADAb positive - treatment emergent unaffected ADAb positive**
- For subjects who are positive at baseline, and are positive at any sampling point post treatment (including SFU) with increased titer values compared to baseline (above a predefined fold difference increase from baseline value which will be defined within the validation of the assay and will be included in the TFLs and/or SAP when available) - **pre ADAb positive - treatment emergent ADAb boosted positive.**

Have been changed to:

- For subjects who are positive at baseline, and are positive **at any** sampling point post treatment (including SFU) with titer values of the same magnitude as baseline (i.e. ≤ than a **2.07** fold difference from the baseline value) - **pre ADAb positive - treatment emergent unaffected ADAb positive**

- For subjects who are positive at baseline, and are positive at any sampling point post treatment (including SFU) with increased titer values compared to baseline (above a **2.07** fold difference increase from baseline value which will be defined within the validation of the assay and will be included in the TFLs and/or SAP when available) - **pre ADAb positive - treatment emergent ADAb boosted positive.**

And

- The number and percentage of subject in each of the 6 categories will be tabulated and separated by treatment group, with an additional category combining subjects who are **pre ADAb negative - treatment emergent ADAb positive** and - **pre ADAb positive - treatment emergent ADAb boosted positive, summarized as Total treatment emergent.**

Has been changed to

- The number and percentage of subject in each of the 6 categories will be tabulated and separated by treatment group, with an additional category combining subjects who are **pre ADAb negative - treatment emergent ADAb positive** and - **pre ADAb positive - treatment emergent ADAb boosted positive, summarized as Total treatment emergent.** In addition, the number and percentage of subjects who are pre ADAb positive will be summarized (combination of **pre ADAb positive - treatment emergent reduced ADAb, pre ADAb positive - treatment emergent unaffected ADAb positive and pre ADAb positive - treatment emergent ADAb boosted positive**).

And

The following bullet point was removed

- A summary table to present the incidence (%) of subjects classified as treatment emergent ADAb positive with incidence of IL17-A and IL17F neturalizing ADAb, respectively over the 56 weeks.

And

The following text was added

For purposes of efficacy subgroup analyses based on anti-bimekizumab antibody status, two categories will be used:

- Anti-bimekizumab antibody positive – This is defined as subjects who have anti-bimekizumab antibody levels above the specified cut point on at least 2 time points while on treatment (ie, excluding Baseline, excluding SFU).
- Anti-bimekizumab antibody negative – Subjects who are not defined as anti-bimekizumab positive (as described above) will be defined as anti-bimekizumab antibody negative.

The rationale for requiring at least 2 time points in which anti-bimekizumab antibody levels are above the specified cut point is to exclude those subjects who have only one occurrence of anti-bimekizumab antibody levels during the course of treatment. Including such subjects would

increase the number of anti-bimekizumab antibody positive subjects with potentially no impact on efficacy.

Change #34

Section 10 SAFETY ANALYSES

All safety summaries and listings will be performed using all subjects in the SS.

The SS and BKZ Set will be used for summaries of safety data.

Has been changed to

All safety summaries and listings will be **created** using all subjects in the SS.

The SS, BKZ Set, **BKZ Week 24 Set and MS** will be used for summaries of safety data.

It was noted during the conduct of the study that approximately 24 subjects received incorrect Placebo/Active treatment at Visit 20. All subsequent treatments were administered per protocol. The impact of this on safety is expected to be minor and no adjustments will be made to the summaries of safety data. The actual treatments received will be shown in the data listings.

Change #35

Section 10.1.2 Exposure through Week 24 (not including Week 24 Visit) for SS

Calculations adding +28 days were changed to 28 days **(or 56 days in the case of Q8W dosing)**

Change #36

Section 10.1.3 Exposure during the initial and maintenance period for BKZ set

Calculations adding +28 days were changed to 28 days **(or 56 days in the case of Q8W dosing)**

And

Calculations using final visit date (not including SFU) were changed to final visit date (**including PEOT, but** not including SFU)

Change #37

The following section and text was added:

Section 10.1.4 Exposure during the maintenance period for BKZ Set

Definitions for study medication duration (days) and time at risk (days) during the maintenance period (Week 16 to Week 56) are provided as follows:

Study medication duration (days)

- Date of last dose of bimekizumab in the maintenance period – date of first dose of bimekizumab in the maintenance period + 28 days (or 56 days in the case of Q8W dosing).

Note: If date of last dose of bimekizumab in the maintenance period + 28 days (or 56 days in the case of Q8W dosing) extends to a date beyond the final visit date in the maintenance period (including PEOT, but not including SFU), then this calculation reverts to:

– Final visit date in the maintenance period (including PEOT, but not including SFU) – date of first dose of bimekizumab in the maintenance period + 1.

Note: For subjects who die, if date of last dose of bimekizumab in the maintenance period + 28 days (or 56 days in the case of Q8W dosing) extends to a date beyond the date of death, then this calculation reverts to:

– Date of death – date of first dose of bimekizumab in the maintenance period + 1.

Time at risk (days):

- For subjects who continue into an open-label study (and, therefore, do not have the SFU visit in the feeder study): Final visit date in the maintenance period – date of first dose of bimekizumab in the maintenance period + 1.
- For subjects who die prior to the final visit in the maintenance period: Date of death – date of first dose of bimekizumab in the maintenance period + 1.
- For all other subjects, use the minimum of the following:
  - Date of last dose of bimekizumab in the maintenance period – date of first dose of first bimekizumab in the maintenance period + 140 days.
  - Date of last clinical contact in the maintenance period – date of first dose of bimekizumab in the maintenance period + 1.

  Note: This group could include subjects who discontinue the maintenance period early, subjects who complete the Maintenance Period as scheduled but choose not to continue into an open-label study, or subjects who are ongoing in the SFU period at the time of the data snapshot.

Change #38

Section 10.2.1 Data considerations

Adverse drug reactions (ADRs) are defined as a response to a drug which is noxious and unintended and which occurs at doses normally used in man for prophylaxis, diagnosis, or therapy of disease or for modification of physiological function. Any AE that is considered "Related" to study treatment will be classed as an ADR. Risk difference and the corresponding confidence interval will be presented as a percentage.

Has been changed to:

Selected summaries for the subjects who switch from adalimumab to bimekizumab will be presented by time of onset (≤ 70 days or > 70 days from the final dose of adalimumab). Only events starting after the first dose of bimekizumab will be included. 70 days was chosen as it is approximately equal to five half lives of adalimumab.

Selected summaries, as specified in Section 10.2.2, will include the risk difference between bimekizumab and adalimumab. The risk difference is calculated as:

$$RD = IP_{BKZ} - IP_{ADA}$$

where IP_{BKZ} is the incidence proportion for the bimekizumab-treated group and IP_{ADA} is the incidence proportion for the adalimumab group. Note that incidence proportion simply refers to the percentage of subjects within the specified treatment group that experienced a given adverse event.

The standard error for the risk difference is calculated as follows:

$$SE_{RD} = \sqrt{\left(IP_{BKZ} \times \left(\frac{1 - IP_{BKZ}}{n_{BKZ}}\right)\right) + \left(IP_{ADA} \times \left(\frac{1 - IP_{ADA}}{n_{ADA}}\right)\right)}$$

where n_{BKZ} is the number of subjects in the bimekizumab-treated group and n_{ADA} is the number of subjects in the adalimumab group.

The corresponding confidence interval for the risk difference is as follows:

$$CI_{RD} = RD \pm Z_{1-\alpha/2} \times (SE_{RD})$$

where $Z_{1-\alpha/2}$ is the Z statistic for the corresponding level of alpha. For the risk difference confidence intervals calculated in this SAP, 1.96 will be used (corresponding to a two-sided alpha of 0.05 and 95% confidence interval). Risk difference and the corresponding confidence interval will be presented as a percentage.

Change #39

Section 10.2.2 AE summaries

All AEs will be summarized through Week 24 for the SS, and through Week 56 for the BKZ Set. In addition, all summaries of TEAEs based on "100 subject years" should include EAIR (with 95% confidence interval) and EAER.

- Incidence of TEAEs – Overview
- Incidence of TEAEs per 100 subject years by SOC, HLT, and PT
- Incidence of Serious TEAEs per 100 subject years by SOC, HLT, and PT
- Incidence of TEAEs Leading to Discontinuation per 100 subject years by SOC, HLT, and PT
- Incidence of TEAEs Leading to Death by SOC, HLT, and PT
- Incidence of TEAEs by Maximum Relationship by SOC, HLT, and PT
- Incidence of Serious TEAEs by Relationship SOC, HLT, and PT – Note: For EudraCT reporting purposes
- Incidence of Serious Adverse Drug Reactions by SOC, HLT, and PT
- Incidence of TEAEs Leading to Death by Relationship by SOC, HLT, and PT – Note: For EudraCT reporting purposes
- Incidence of TEAEs by Maximum Severity, SOC, HLT, and PT

- Incidence of TEAEs by decreasing frequency of PT
- Incidence of TEAEs Above Reporting Threshold of 5% by SOC and PT
- Incidence of Non-Serious TEAEs by SOC, HLT, and PT
- Incidence of Non-Serious TEAEs by Relationship SOC, HLT, and PT
- Incidence of Non-Serious TEAEs Above Reporting Threshold of 5% by SOC and PT
- Incidence of Non-Serious TEAEs Above Reporting Threshold of 5% by Relationship SOC and PT
- Incidence of Adverse Drug Reactions by SOC, HLT, and PT
- Incidence of Adverse Drug Reactions Above Reporting Threshold of 5% by SOC and PT

Incidence of TEAEs per 100 subject years by SOC, HLT, and PT and by Time of Onset Relative to Anti-bimekizumab Antibody Status).

Has been changed to:

The following summaries will be provided in five different ways: through Week 24 for the SS, through initial and maintenance treatment period for the BKZ Set, through initial treatment period for the SS, through the maintenance period for the MS (excluding patients randomized to ADA) and from Week 24 to Week 56 for the BKZ Week 24 Set for patients randomized to ADA switching to BKZ (to examine the events occurring within 70 days and >70 days of the treatment switch).

- Incidence of TEAEs – Overview
- Incidence of TEAEs per 100 subject years by SOC, HLT, and PT
- Incidence of Serious TEAEs per 100 subject years by SOC, HLT, and PT

The following summary will be provided in four different ways: through Week 24 for the SS, through initial and maintenance treatment period for the BKZ Set, through initial treatment period for the SS and through the maintenance period for the MS (excluding patients randomized to ADA).

- Incidence of TEAEs Leading to Discontinuation per 100 subject years by SOC, HLT, and PT

The following summaries will be provided in two different ways: through Week 24 for the SS, and through initial and maintenance treatment period for the BKZ Set:

- Incidence of TEAEs Leading to Death by SOC, HLT, and PT
- Incidence of TEAEs by Maximum Relationship by SOC, HLT, and PT
- Incidence of Serious TEAEs by Relationship SOC, HLT, and PT – Note: For EudraCT reporting purposes
- Incidence of Related Serious TEAEs by SOC, HLT, and PT
- Incidence of TEAEs Leading to Death by Relationship by SOC, HLT, and PT – Note: For EudraCT reporting purposes
- Incidence of TEAEs by Maximum Severity, SOC, HLT, and PT

- Incidence of TEAEs by decreasing frequency of PT
- Incidence of TEAEs Above Reporting Threshold of 5% by SOC and PT
- Incidence of Non-Serious TEAEs by SOC, HLT, and PT
- Incidence of Non-Serious TEAEs by Relationship SOC, HLT, and PT
- Incidence of Non-Serious TEAEs Above Reporting Threshold of 5% by SOC and PT
- Incidence of Non-Serious TEAEs Above Reporting Threshold of 5% by Relationship SOC and PT
- Incidence of Related TEAEs by SOC, HLT, and PT
- Incidence of Related TEAEs Above Reporting Threshold of 5% by SOC and PT

The following summary will be provided through Week 24 for the SS:

- Incidence of TEAEs Above Reporting Threshold of 5% with Risk Differences by SOC and PT
- Incidence of Serious TEAEs and Risk Differences by SOC and PT

The following summary will be provided through initial and maintenance treatment period for the BKZ Set:

- Incidence of TEAEs per 100 subject years by SOC, HLT, and PT and by Time of Onset Relative to Anti-bimekizumab Antibody Status

In addition, all summaries of TEAEs based on "100 subject years" should include EAIR (with 95% confidence interval) and EAER.

The tables with risk differences will also be accompanied by figures (dot plots) which show the incidence of the adverse events and corresponding 95% risk difference confidence intervals. These will be ordered by descending order of risk difference (bimekizumab vs adalimumab).

Change #40

Section 10.2.3 Other safety topics of interest

The following are AEs of other safety topics of interest that require special statistical analyses.
Has been changed to:

The following are AEs of other safety topics of interest that require special statistical analyses. Along with the tables described, there will be a table which displays the risk difference and 95% confidence intervals for each of the topics of interest. A corresponding figure (with dot plots) will be prepared.

Change #41

Section 10.2.3.1 Infections (serious, opportunistic, fungal and TB)

The embedded OI - MedDRA v19.0.xlsx file has been removed from the document.

And

"Incidence**" has been changed with "**Incidence **and Event Rate"**

And

1. Study programming team creates a spreadsheet which lists all of the subjects with a TEAE present in the database which codes to a PT identified as case-by-case. Information from the AE dataset to be included in the spreadsheet: Subject ID, AE verbatim term, SOC, High Level Term (HLT), Lower Level Term (LLT), PT, AE start date, AE end date, seriousness, severity, relationship to study medication, action taken. Additionally, a column will be included where the study physician can document their decision on the case.
2. Study physician reviews the cases in the spreadsheet and indicates in the additional column which AEs are confirmed to be opportunistic infections via a single 'x'.
3. Study programming team incorporates these decisions into the AE dataset by merging the study physician decisions for individual subjects / PTs and flagging the confirmed opportunistic infections as such in the dataset.

Has been changed to

1. Study programming team creates a spreadsheet which lists all of the subjects with a TEAE present in the database which codes to a PT identified as case-by-case. Information from the AE dataset to be included in the spreadsheet: Subject ID, AE verbatim term, SOC, High Level Term (HLT), Lower Level Term (LLT), PT, AE start date, AE end date, seriousness, severity, relationship to study medication, action taken. Additionally, a column will be included where the study physician can document their decision on the case.
2. Study physician reviews the cases in the spreadsheet and indicates in the additional column which AEs are confirmed to be opportunistic infections via a single 'x'.
3. **At quarterly Infectious Disease Committee (IDC) Meetings, outputs will be produced and reviewed by the study physician ahead of the IDC Meeting. The IDC will then agree on the final adjudication for each potential opportunistic infection. A final output for opportunistic infections will be produced and agreed upon between the study physician and the IDC prior to finalizing the database.**
4. Study programming team incorporates these decisions into the AE dataset by merging the **final** decisions for individual subjects / PTs and flagging the confirmed opportunistic infections as such in the dataset.

And

Following the initial physician review of case-by-case events, subsequent reviews will be based on the cumulative set of case-by-case events present in the database at each time point of spreadsheet creation. Physician decisions from previous runs should be retained in each subsequent run. The final run of the spreadsheet, with all study physician decisions on the full set of case-by-case events, will be archived at the conclusion of the study.

Has been changed to

Following the initial physician review of case-by-case events, subsequent reviews will be based on the cumulative set of case-by-case events present in the database at each time point of spreadsheet creation. Physician decisions from previous runs should be retained in each subsequent run. The final run of the spreadsheet, with all **final** study physician decisions **(with IDC agreement)** on the full set of case-by-case events, will be archived at the conclusion of the study.

And

Step 1: Refer to column B of the spreadsheet which identifies the Preferred Terms (PTs) to be classified as opportunistic infections using either a single 'x' or a double 'xx'.

- TEAEs which code to a PT flagged with a single 'x' need to also be serious to be considered an opportunistic infection.
- All TEAEs which code to a PT flagged with a double 'xx' are considered to be an opportunistic infection, regardless of seriousness.

All serious TEAEs in the study database which code to a PT flagged with a single 'x' and all TEAEs in the study database which code to a PT flagged with a double 'xx' will be summarized as an opportunistic infection in the stand-alone table.

Step 2: Refer to column C of the spreadsheet which identifies the PTs that need to be evaluated on a case-by-case basis by the study physician to determine whether it is a true opportunistic infection or not. The process for physician review is as follows:

1. Study programming team creates a spreadsheet which lists all of the subjects with a TEAE present in the database which codes to a PT identified as case-by-case. Information from the AE dataset to be included in the spreadsheet: Subject ID, AE verbatim term, SOC, High Level Term (HLT), Lower Level Term (LLT), PT, AE start date, AE end date, seriousness, severity, relationship to study medication, action taken. Additionally, a column will be included where the study physician can document their decision on the case.
2. Study physician reviews the cases in the spreadsheet and indicates in the additional column which AEs are confirmed to be opportunistic infections via a single 'x'.
3. At quarterly Infectious Disease Committee (IDC) Meetings, outputs will be produced and reviewed by the study physician ahead of the IDC Meeting. The IDC will then agree on the final adjudication for each potential opportunistic infection. A final output for opportunistic infections will be produced and agreed upon between the study physician and the IDC prior to finalizing the database. Study programming team incorporates these decisions into the AE dataset by merging the study physician decisions for individual subjects / PTs and flagging the confirmed opportunistic infections as such in the dataset.

All subjects with a case-by-case PT reported that has been confirmed by the study physician to be an opportunistic infection in Step 2 will be summarized as such in the stand-alone table, along with all events identified in Step 1 of this process.

The timing and frequency of Step 2 will be outlined and agreed to by the study team at the beginning of the study. It is suggested that this process be executed multiple times throughout the course of the study, more frequently in the weeks prior to the DMC meeting.

Following the initial physician review of case-by-case events, subsequent reviews will be based on the cumulative set of case-by-case events present in the database at each time point of spreadsheet creation. Physician decisions from previous runs should be retained in each subsequent run. The final run of the spreadsheet, with all final study physician decisions (with IDC agreement) on the full set of case-by-case events, will be archived at the conclusion of the study.

Has been changed to

Identification Process

The two steps below outline two ways in which opportunistic infections (or potential opportunistic infections) can be identified:

Step 1: Refer to column B of the spreadsheet which identifies the Preferred Terms (PTs) to be classified as opportunistic infections using either a single 'x' or a double 'xx'.

TEAEs which code to a PT flagged with a single 'x' need to also be serious to be considered an opportunistic infection.
All TEAEs which code to a PT flagged with a double 'xx' are considered to be an opportunistic infection, regardless of seriousness.
Step 2: Refer to column C of the spreadsheet which identifies the PTs that need to be evaluated on a case-by-case basis by the study physician to determine whether or not it is an opportunistic infection. If Column C has a single 'x', then the corresponding preferred term should be flagged for case-by-case review by the study physician.

Review Process

Opportunistic infections for a given study will be reviewed on the following occasions:

At quarterly Infectious Disease Committee (IDC) Meetings, listings will be produced for each study (see details below) and reviewed by the corresponding study physician ahead of the IDC Meeting. These listings will be posted as part of the broader SSD deliverable to a folder named for the given quarter (eg, 2018Q4) on the SharePoint. They should be based on the same data cut as the one used for SSD and should be delivered at the same time as the SSD outputs. The IDC will then agree on the final adjudication for each potential opportunistic infection.
For each study, a final listing for opportunistic infections (in the format described below) will be produced and agreed upon between the study physician and the IDC prior to finalizing the database.
In each of the circumstances described above, the study programming team will produce an Excel listing that will be provided to the project lead statistician, project lead programmer, and to the study physician (who will subsequently provide it to the IDC). The Excel listing will contain the following columns (using the descriptions below as the column headings in the Excel listing):

Study ID
Unique Subject ID
AE Term (Verbatim)
AE Preferred Term

AE System Organ Class
AE High Level Term
AE Low Level Term
Date of Onset
Outcome of Adverse Event
Date of Outcome
TEAE Flag
Serious Adverse Event?
Relationship to Study Medication
Intensity
Action Taken with IMP
Opportunistic Infection – Automatic
Opportunistic Infection – Manual Review
Flag
Data Cut Date
Opportunistic Infection – Final Adjudication
Note the following about the final 5 variables in this listing:

Opportunistic Infection – Automatic: This is flagged as "Y" if the criteria for automatic selection as described in "Step 1" of the identification process are met.
Opportunistic Infection – Manual Review: This is flagged as "Y" if the criteria for case-by-case selection as described in "Step 2" of the identification process are met.
Flag – This has a value of either "NEW" or "OLD". It is marked as "NEW" if the event is appearing for the first time in that run of the listing. Otherwise, if it has appeared previously, it is marked as "OLD". Unique records are determined by USUBJID AESPID for purposes of identifying whether an event has been modified from a previous run.
Date – Only for cases where Flag is "NEW", this field will be populated with the data cut date for that particular run of the listing.
Opportunistic Infection – Final Adjudication – For new events, this is always left blank by the programmers. It should be completed by the study physician/IDC for every event that appears in the listing. For events adjudicated as opportunistic, the field should be populated with a "Y". Following each review by the study physician and IDC, the Opportunistic Infection – Final Adjudication column will be completed (as described above), and the spreadsheets for each study will be returned to the study programming team via e-mail (coordinated by the IDC secretary). Then, for subsequent runs of the listing, the study programming teams will incorporate adjudications from previous runs.

Change #42

Section 10.2.3.2 Malignancies, including Lymphoma

Incidence of Malignant or Unspecified Tumours TEAEs per 100 subject years by SOC, HLT and PT

These events will be presented in the following tables:
One table will be based on the criteria Standardized MedDRA Query (SMQ) = "Malignant or unspecified tumours (SMQ)"
One table will be based on the criteria SMQ = "Malignant tumours (SMQ)".

Has been changed to:

Section 10.2.3.2 Malignancies

- **Incidence and Event Rate of Malignancy (including Unspecified Tumours) TEAEs per 100 subject years by SOC, HLT and PT**

This table will be based on the criteria Standardized MedDRA Query (SMQ) = "Malignant or unspecified tumours (SMQ)"

- **Incidence and Event Rate of Malignancy TEAEs per 100 subject years by SOC, HLT and PT**

This table will be based on the criteria SMQ = "Malignant tumours (SMQ)".

Change #43

Section 10.2.3.3 Major adverse cardiac event

- **Incidence of Major Cardiac Event (MACE) TEAEs per 100 subject years by SOC, HLT and PT**

Major adverse cardiac events (MACE) will be presented in a table and a listing. The classification of an event as MACE is determined by an external cardiovascular event adjudication committee.

Has been changed to:

Section 10.2.3.3 Major adverse cardiac events (MACE)

- **Incidence and Event Rate of Major Cardiac Event (MACE) TEAEs per 100 subject years by SOC, HLT and PT**

Major adverse cardiac events (MACE) will be presented in a table. The classification of an event as MACE is determined by an external cardiovascular event adjudication committee.

A separate table and listing will present the adjudicated cardiovascular events by type. For each cardiovascular event type (24 total), the individual PTs which fall within each event type will be summarized.

Additionally, a listing of all events identified for potential review by the cardiovascular event adjudication committee will be produced. This listing will indicate whether each event was escalated to the committee for formal review/adjudication.

Change #44

Section 10.2.3.5 Neuropsychiatric events (in particular depression and suicide)

- **Incidence of Neuropsychiatric TEAEs per 100 subject years by SOC, HLT and PT**

This table is based on the SMQ of "Depression and suicide/self-injury" (all TEAEs which code to a PT included in the Scope=Broad and/or Scope=Narrow).

An external neuropsychiatric adjudication committee will evaluate potential neuropsychiatric events and determine whether any of those events were associated with suicidal ideation and

behavior (SIB). If an event is adjudicated as SIB, further information will be provided. A separate listing for SIB events as determined by the adjudication committee will be included.

Has been changed to:

Section 10.2.3.5 Suicidal Ideation and Behavior

- **Incidence and Event Rate of SIB-Adjudicated Neuropsychiatric TEAEs per 100 subject years by SOC, HLT and PT**

An external neuropsychiatric adjudication committee will evaluate potential neuropsychiatric events and determine whether any of those events were associated with suicidal ideation and behavior (SIB). If an event is adjudicated as SIB, further information will be provided. A table and listing for SIB events as determined by the adjudication committee will be included.

Additionally, a listing of all events identified for potential review by the neuropsychiatric adjudication committee will be produced. This listing will indicate whether each event was escalated to the committee for formal review/adjudication.

Change #45

Section 10.2.3.6 Inflammatory bowel disease

The following sentence was added:
Inflammatory bowel disease events will be summarized stratified by subjects with or without a previous medical history of inflammatory bowel disease. Previous medical history of inflammatory bowel disease will be determined using the information recorded on the Extra-Articular Assessment at Screening CRF page ("Does subject have a history of IBD?").

Change #46

Section 10.2.3.7 Anaphylactic reactions

- **Incidence of Anaphylactic Reaction TEAEs per 100 subject years by SOC, HLT and PT**

This table will be prepared based on the MedDRA anaphylaxis algorithm (refer to Appendix 1 in "Bimekizumab-Safety-Topics-of-Interest.docx") for acute events (reported on the same day) and for treatment-emergent PTs including the term "hypersensitivity" reported at any time.

Furthermore, injection site reactions will be evaluated based on the any TEAE table (no separate table needed) by looking under the following HLTs: "Administration site reactions NEC" and "Injection site reactions".

Has been changed to:

- **Incidence and Event Rate of Anaphylactic Reaction TEAEs per 100 subject years by SOC, HLT and PT**

A separate table will be prepared based on the MedDRA anaphylaxis Algorithm (see Appendix 1) for acute anaphylactic events (reported on the same day as when an injection was administered or one day after).

- **Incidence and Event Rate of Hypersensivity Reaction TEAEs per 100 subject years by SOC, HLT and PT**

A separate table will be prepared to summarize hypersensitivity events, identified using the SMQ "Hypersensitivity (SMQ)". All TEAEs which code to a PT included in the Scope=Narrow search will be included in this table.

Furthermore, injection site reactions will be evaluated based on the any TEAE table (no separate table needed) by looking under the following HLTs: "Administration site reactions NEC" and "Injection site reactions".

Change #47

Section 10.2.3.8 Hepatic events and DILI

- Incidence of hepatic events TEAEs per 100 subject years by SOC, HLT and PT

Has been changed to:

- Incidence and Event Rate of Hepatic Events and Drug Induced Liver Injury TEAEs per 100 subject years by SOC, HLT and PT

Change #48

Section 10.3 Clinical laboratory evaluations

Laboratory values and urinary values will be presented descriptively by treatment group for SS up to Week 56. Markedly abnormal values and shifts from baseline (Week 0) will be presented for the SS through Week 24 and BKZ set through Week 56.

Has been changed to

Laboratory values, **including markedly abnormal laboratory values** will be presented descriptively by treatment group for SS up to Week 56. Markedly abnormal values and shifts from baseline (Week 0) will be presented for the SS through Week 24 and BKZ set through initial and maintenance treatment period.

The markedly abnormal tables and those based on CTCAE grade will be produced only for selected laboratory variables.

And

- A summary of the number and percentage of subjects experiencing markedly abnormal values by laboratory variable, treatment group and visit
- A shift table of the number and percentage of subjects experiencing low, normal or high values at Baseline to maximum post-Baseline value (ie, low, normal, high), by laboratory variable and treatment group
- A Shift table of the number and percentage of subjects experiencing low, normal or high values at Baseline to minimum post-Baseline value (ie, low, normal, high), by laboratory variable and treatment group

- A Shift table of the number and percentage of subjects experiencing low, normal or high values at Baseline to the End of Treatment[1] value (ie, low, normal, high), by laboratory variable and treatment group.

[1]The End of Treatment value refers to the value from the last observed non-missing post baseline visit prior to the end of treatment.

Has been changed to

- A summary of the number and percentage of subjects experiencing markedly abnormal values at any time while on treatment (assessment on or following the first dose of study treatment through the final dose of study treatment + 140 days) by laboratory variable and treatment group
- A summary of the number and percentage of subjects with a given CTCAE grade (0, 1, 2, 3, or 4) based on minimum/maximum post-baseline value by laboratory variable and treatment group
- A shift table of the number and percentage of subjects experiencing CTCAE grade 0, 1, 2, 3 or 4 values (as applicable) at Baseline to maximum post-Baseline CTCAE grade, by laboratory variable and treatment group

For laboratory summaries through week 24, the cut off will be the earliest of week 24 or last dose + 140 days.

For laboratory shift tables through week 24, the treatment groups BKZ 320mg Q4W, BKZ 320mg Q4W/Q8W, BKZ Total and ADA will be presented. For laboratory shift tables through initial and maintenance treatment period, the treatment groups BKZ 320mg Q4W, BKZ 320mg Q4W/Q8W, ADA/ BKZ 320 mg Q4W and BKZ Total will be presented. For subjects randomized to ADA, only data after the switch to Bimekizumab will be considered. The last value prior to the switch will be considered as baseline.

The following paragraph has been updated (bold):

Health and Human Services 2010). Definitions of markedly abnormal values using the Grade 3 cut points are given in the tables below for age ranges of ≥17 years (Table 10–1 for markedly abnormal biochemistry values and Table 10–2 for markedly abnormal hematology values). **Tables summarizing markedly abnormal values should include a summary (counts and percentages) of markedly abnormal labs observed at any time while on treatment (ie, treatment-emergent markedly abnormal [TEMA]). For this summary, Baseline values and values observed more than 140 days after the last administration of study medication are not considered.** The laboratory results classified as Grade 3 or Grade 4 will be summarized and listed separately.

Summaries in the ADA/BKZ arm will be provided to examine markedly abnormal labs occurring within 70 days and >70 days of the treatment switch. The time since the treatment switch is calculated as the number of days since the final dose of adalimumab. Only lab values recorded after the start of bimekizumab treatment will be included. 70 days was chosen as it is approximately equal to five half lives of adalimumab.

The following text

A summary table highlighting the potential cases of Hy's Law, within each treatment group will be presented. Hy's Law is defined as:

- AST >3xULN or ALT >3xULN and
- Total Bilirubin >2xULN

In order to meet the above criteria, a subject must experience the elevation in bilirubin and ALT or AST at the same visit. For example, a subject who experiences a >=2 x ULN elevation of bilirubin at one visit and a 3 x ULN elevation in ALT (or AST) at a subsequent visit has not fulfilled the Hy's law criteria.

Has been changed to

The following criteria will be considered for summaries of markedly abnormal liver function tests (LFTs):

- **AST: >3xULN, >5xULN, >8xULN, >10xULN, >20xULN**
- **ALT: >3xULN, >5xULN, >8xULN, >10xULN, >20xULN**
- **AST or ALT: >3xULN, >5xULN, >8xULN, >10xULN, >20xULN**
- **Bilirubin: >1.5xULN, >2xULN**

In addition, a table will be produced to summarize potential Hy's Law cases. The following definition will be used in that table:

- **[AST ≥3xULN or ALT ≥3xULN] and Total Bilirubin ≥2xULN in the absence of ALP ≥2xULN**

In order to meet the above criteria, a subject must experience the elevation in bilirubin and ALT or AST (and the absence of ALP elevation) at the same visit. For example, a subject who experiences a ≥2 x ULN elevation of bilirubin at one visit and a ≥3xULN elevation in ALT (or AST) at a subsequent visit has not fulfilled the Hy's law criteria.

And the following sentence was deleted:

Other laboratory values of interest may be provided depending on the study protocol.

And the following table was deleted:

Table 10.1 Definitions of Markedly Abnormal Liver Functions Values

Change #49

Section 10.4.1 Vital signs

- A summary of the number and percentage of subjects experiencing at least one markedly abnormal value for a vital sign variable as defined in Table 10–3, by treatment group **and visit,** through Week 24 for the SS and through Week 56 for the BKZ set

Has been changed to

- A summary of the number and percentage of subjects experiencing at least one markedly abnormal value for a vital sign variable as defined in Table 10–3, by treatment group,

through Week 24 for the SS and through initial and maintenance treatment period for the BKZ set

Change #50

Section 10.4.3 Electrocardiograms

A summary of the number and percentage of subjects with normal, abnormal not clinically significant and abnormal clinically significant ECG results at all applicable visits will be provided.

Has been changed to:

A summary of the number and percentage of subjects with normal and abnormal ECG results, as determined by the central reader, will be presented for all applicable visits.

And

A by-subject listing of all 12-lead ECG data will be provided.

Has been changed to:

Two separate by-subject listings of all 12-lead ECG data will be provided based on interpretation from central reader and from site.

Change #51

Section 10.4.5 Patient Health Questionnaire 9 (PHQ-9)

This was added:

MI-MCMC/monotone regression method will be used to impute missing data. If the imputation model cannot converge, last observation carried forward (LOCF) will be used.

Change from Baseline in PHQ-9 is derived as post-Baseline score minus Baseline score. It will be summarized by treatment group through initial and maintenance treatment period for the SS and for the BKZ Week 24 Set for subjects switching from adalimumab to bimekizumab.

Change #52

Section 11 References

The following references were added:

Greenland S, Robins JM. Estimation of common effect parameter from sparse follow up data. Biometrics 1985;41:55-68.

Ratitch, B., Lipkovich, I., O'Kelly, M. Combining Analysis Results from Mutiply Imputed Categorical Data, PharmaSUG, 2013, SP03.

Change #53

Section 12 Appendix

Appendix A: MedDRA algorithmic approach to anaphylaxis has been added

13.2 Amendment 2

Change #1

Section 4.2.1: Handling of Missing Data for Efficacy Variables
The paragraph below:

If a subject discontinued study treatment without terminating from the study and still continues with scheduled assessments, all efficacy data after discontinuation of study treatment will be treated as missing and subject to imputation as applicable.

Was changed to:

If a subject discontinued study treatment without terminating from the study and still continues with scheduled assessments, all efficacy data after **last treatment date + 35 days** will be treated as missing and subject to imputation as applicable.

Change #2

Section 4.2.1.5: Missing Data Overview and Summary

The sentence highlighted in bold was added.

- Multiple Imputation (MI) – Markov Chain Monte Carlo (MCMC) / Monotone Regression: Using multiple imputation methodology, intermittent missing data are imputed based on the MCMC method, and monotone missing data are imputed using monotone regression. **Note: for all analyses and summaries using MI, subjects with no baseline value will be excluded.**

Change #3

Section 10.1: Extent of exposure

An extra category **">= 40 weeks"** was added to the cumulative study medication duration summary for the initial and maintenance treatment period.

Change #4

Section 10.1.1: Exposure during the initial treatment period for SS

Study medication duration (days)

- Date of last dose prior to Week 16 Visit – date of first dose in the Initial Treatment Period + 28 days

Note: If date of last dose prior to Week 16 + 28 days extends to a date beyond the date of Week 16 visit, then this calculation reverts to:

  – Date of Week 16 visit – date of first dose in the Initial Treatment Period + 1.

  Note: For subjects who die during the first 16 weeks, if date of last dose + 28 days extends to a date beyond the date of death, then this calculation reverts to:

  – Date of death – date of first dose in the Initial Treatment Period + 1.

Time at risk (days)

- For subjects who complete the initial treatment period: Date of Week 16 visit – date of first dose + 1.
- For subjects who discontinue prior to Week 16 visit, use the minimum of the following:
  – Date of last dose – date of first dose in the Initial Treatment Period + 140 days.
  – The total number of days from first dose up to Week 16.
  – Date of last clinical contact – date of first dose + 1.
- For subjects who die prior to Week 16 visit: Date of death – date of first dose +1.

Was changed to:

Study medication duration (days)

- Date of last dose of active medication in the Initial treatment Period – date of first dose in the Initial Treatment Period + 28 days for bimekizumab Q4W/ + 14 days for adalimumab.

Note: If date of last dose of active medication in the Initial Treatment Period + 28 days/ 14 days extends to a date beyond the date of the first dose in the Maintenance Treatment Period, then this calculation reverts to:

  – Date of first dose in the Maintenance Treatment Period – date of first dose in the Initial Treatment Period + 1.

  Note: For subjects who die during the Initial Treatment Period, if date of last dose of active medication in the Initial Treatment Period + 28 days/ 14 days extends to a date beyond the date of death, then this calculation reverts to:

  – Date of death – date of first dose in the Initial Treatment Period + 1.

For subjects who permanently discontinue study treatment:

- Date of last dose of active medication – date of first dose + 28 days for bimekizumab Q4W/ + 14 days for adalimumab.

  Note: If date of last dose + 28 days/ 14 days extends to a date beyond the final visit date (including PEOT, but not including SFU), then this calculation reverts to:

  – Final visit date (including PEOT, but not including SFU) – date of first dose + 1.

Time at risk (days)

- For subjects who complete the final visit of Initial Treatment Period and continue into the Maintenance Treatment Period: Date of first dose in the Maintenance Treatment Period – date of first dose in the Initial Treatment Period + 1.
- For subjects who discontinue on or prior to the final visit of the Initial Treatment Period, use the minimum of the following:
  - Date of last dose of Initial Treatment Period – date of first dose in the Initial Treatment Period + 140 days.
  - The total number of days in the Initial Treatment Period.
  - Date of last clinical contact – date of first dose in the Initial Treatment Period + 1.
- For subjects who die prior to the final visit of the Initial Treatment Period: Date of death – date of first dose in Initial Treatment Period +1.

Change #5

Section 10.1.2: Exposure through Week 24 (not including Week 24 Visit) for SS

Study medication duration (days)

- Date of last dose prior to Week 24 Visit – date of first dose in the Initial Treatment Period + 28 days (or 56 days in the case of Q8W dosing).

  Note: If date of last dose prior to Week 24 + 28 days (or 56 days in the case of Q8W dosing) extends to a date beyond the date of Week 24 visit, then this calculation reverts to:

  - Date of Week 24 visit – date of first dose in the Initial Treatment Period + 1.

  Note: For subjects who die during the first 24 weeks, if date of last dose + 28 days (or 56 days in the case of Q8W dosing) extends to a date beyond the date of death, then this calculation reverts to:

  - Date of death – date of first dose in the Initial Treatment Period + 1.

Time at risk (days)

- For subjects who complete through Week 24: Date of Week 24 visit – date of first dose + 1.
- For subjects who discontinue prior to Week 24 visit, use the minimum of the following:
  - Date of last dose – date of first dose in the Initial Treatment Period + 140 days.
  - The total number of days up to Week 24.
  - Date of last clinical contact – date of first dose + 1.
- For subjects who die prior to Week 24 visit: Date of death – date of first dose +1.

Was changed to:

Study medication duration (days)

- Date of last dose of active medication prior to Week 24 Visit – date of first dose in the Initial Treatment Period + 28 days for bimekizumab Q4W/ 56 days for bimekizumab Q8W/ 14 days for adalimumab

  Note: If date of last dose of active medication prior to Week 24 + 28 days/ 56 days/ 14 days extends to a date beyond the date of the first dose on or after Week 24, then this calculation reverts to:

  – Date of first dose on or after Week 24 – date of first dose in the Initial Treatment Period + 1.

  Note: For subjects who die during the first 24 weeks, if date of last dose of active medication prior to Week 24 + 28 days/ 56 days/ 14 days extends to a date beyond the date of death, then this calculation reverts to:

  – Date of death – date of first dose in the Initial Treatment Period + 1.

For subjects who permanently discontinue study treatment:

- Date of last dose of active medication – date of first dose + 28 days/ 56 days/ 14 days
  Note: If date of last dose + 28 days/ 56 days/ 14 days extends to a date beyond the final visit date (including PEOT, but not including SFU), then this calculation reverts to:
  – Final visit date (including PEOT, but not including SFU) – date of first dose + 1.

Time at risk (days)

- For subjects who complete through Week 24 and continue beyond Week 24: Date of first dose on or after Week 24 – date of first dose in Initial Treatment Period + 1.
- For subjects who discontinue prior to Week 24 visit, use the minimum of the following:
  – Date of last dose prior to Week 24 – date of first dose in the Initial Treatment Period + 140 days.
  – The total number of days up to Week 24.
  – Date of last clinical contact – date of first dose in the Initial Treatment Period + 1.
- For subjects who die prior to Week 24 visit: Date of death – date of first dose in Initial Treatment Period +1.

Change #6

Section 10.1.3: Exposure during the initial and maintenance period for BKZ Set

Study medication duration (days)

- Date of last dose of bimekizumab – date of first dose of bimekizumab + 28 days (or 56 days in the case of Q8W dosing).

Note: If date of last dose of bimekizumab + 28 days (or 56 days in the case of Q8W dosing) extends to a date beyond the final visit date (including PEOT, but not including SFU), then this calculation reverts to:

– Final visit date (including PEOT, but not including SFU) – date of first dose of bimekizumab + 1.

Note: For subjects who die, if date of last dose of bimekizumab + 28 days (or 56 days in the case of Q8W dosing) extends to a date beyond the date of death, then this calculation reverts to:

– Date of death – date of first dose of bimekizumab + 1.

Time at risk (days):

- For subjects who continue into an open-label study (and, therefore, do not have the SFU visit in the feeder study): Final visit date – date of first dose of bimekizumab + 1.
- For subjects who die prior to the final visit: Date of death – date of first dose of bimekizumab + 1.
- For all other subjects, use the minimum of the following:
  – Date of last dose of bimekizumab – date of first dose of first bimekizumab + 140 days.
  – Date of last clinical contact – date of first dose of bimekizumab + 1.

  Note: This group could include subjects who discontinue early, subjects who complete the Maintenance Period as scheduled but choose not to continue into an open-label study, or subjects who are ongoing in the SFU period at the time of the data snapshot.

Was changed to:

For subjects randomized to BKZ Q4W/Q4W:

Study medication duration (days)

- Date of last dose of bimekizumab – date of first dose + 28 days

  Note: If date of last dose of bimekizumab+ 28 days extends to a date beyond the final visit date (including PEOT, but not including SFU), then this calculation reverts to:

  – Final visit date (including PEOT, but not including SFU) – date of first dose + 1.

  Note: For subjects who die, if date of last dose of bimekizumab + 28 days extends to a date beyond the date of death, then this calculation reverts to:

  – Date of death – date of first dose + 1.

Time at risk (days):

- For subjects who continue into an open-label study (and, therefore, do not have the SFU visit in the feeder study): Final visit date – date of first dose + 1.
- For subjects who die prior to the final visit: Date of death – date of first dose + 1.

- For all other subjects, use the minimum of the following:
  - Date of last dose– date of first dose + 140 days.
  - Date of last clinical contact – date of first dose + 1.

  Note: This group could include subjects who discontinue early, subjects who complete the Maintenance Period as scheduled but choose not to continue into an open-label study, or subjects who are ongoing in the SFU period at the time of the data snapshot.

For subjects randomized to BKZ Q4W/Q8W:

Study medication duration (days)

- Initial Treatment Period (attributed to BKZ Q4W)
  - Date of last dose of bimekizumab in the Initial treatment Period – date of first dose in the Initial Treatment Period + 28 days

Note: If date of last dose of bimekizumab in the Initial Treatment Period + 28 days extends to a date beyond the date of the first dose in the Maintenance Treatment Period, then this calculation reverts to:

- Date of first dose in the Maintenance Treatment Period – date of first dose in the Initial Treatment Period + 1.

Note: For subjects who die during the Initial Period, if date of last dose of bimekizumab in the Initial Treatment Period + 28 days extends to a date beyond the date of death, then this calculation reverts to:

- Date of death – date of first dose in the Initial Treatment Period + 1.

- Maintenance Treatment Period (attributed to BKZ Q8W)
  - Use the study medication duration algorithm described for Maintenance Treatment Period in 10.1.4.

- Initial and Maintenance Treatment Period combined (attributed to BKZ Total)
  - Date of last dose of bimekizumab – date of first dose + 28 days (if last dose was in Initial Period) or 56 days (if last dose was in Maintenance Period)

Note: If date of last dose of bimekizumab + 28 days (if last dose was in Initial Period) or 56 days (if last dose was in Maintenance Period) extends to a date beyond the final visit date (including PEOT, but not including SFU), then this calculation reverts to:

- Final visit date (including PEOT, but not including SFU) – date of first dose + 1.

Note: For subjects who die, if date of last dose of bimekizumab + 28 days (if last dose was in Initial Period) or 56 days (if last dose was in Maintenance Period) extends to a date beyond the date of death, then this calculation reverts to:

- Date of death – date of first dose + 1.

Time at risk (days):

- Initial Treatment Period (attributed to BKZ Q4W)
  - Date of first dose in Maintenance Treatment Period – Date of first dose in Initial Treatment Period + 1 (Note, this assumes that anyone in this category completes the Initial Treatment Period and doses in the Maintenance Treatment Period).
- Maintenance Treatment Period (attributed to BKZ Q8W)
  - Follow the time at risk algorithm described for the Maintenance Treatment Period in 10.1.4.
- Initial and Maintenance Treatment Period combined (attributed to BKZ Total)
  - For subjects who continue into an open-label study (and, therefore, do not have the SFU visit in the feeder study): Final visit date – date of first dose + 1.
  - For subjects who die prior to the final visit: Date of death – date of first dose + 1.
  - For all other subjects, use the minimum of the following:
    - Date of last dose– date of first dose + 140 days.
    - Date of last clinical contact – date of first dose + 1.

  Note: This group could include subjects who discontinue early, subjects who complete the Maintenance Period as scheduled but choose not to continue into an open-label study, or subjects who are ongoing in the SFU period at the time of the data snapshot.

For subjects randomized to ADA/BKZ Q4W:

Only the duration of Bimekizumab treatment will be derived over this period as the summaries produced are for the BKZ Set (not including any adalimumab treatment).

- Follow the study medication duration and time at risk algorithms described for the Maintenance Treatment Period in 10.1.4.

Change #7

Section 10.1.4: Exposure during the maintenance period for BKZ Set

<u>Study medication duration (days)</u>

- Date of last dose of bimekizumab in the maintenance period – date of first dose of bimekizumab in the maintenance period + 28 days (or 56 days in the case of Q8W dosing).

  Note: If date of last dose of bimekizumab in the maintenance period + 28 days (or 56 days in the case of Q8W dosing) extends to a date beyond the final visit date in the maintenance period (including PEOT, but not including SFU), then this calculation reverts to:

  – Final visit date in the maintenance period (including PEOT, but not including SFU) – date of first dose of bimekizumab in the maintenance period + 1.

Note: For subjects who die, if date of last dose of bimekizumab in the maintenance period + 28 days (or 56 days in the case of Q8W dosing) extends to a date beyond the date of death, then this calculation reverts to:

– Date of death – date of first dose of bimekizumab in the maintenance period + 1.

Time at risk (days):

- For subjects who continue into an open-label study (and, therefore, do not have the SFU visit in the feeder study): Final visit date in the maintenance period – date of first dose of bimekizumab in the maintenance period + 1.
- For subjects who die prior to the final visit in the maintenance period: Date of death – date of first dose of bimekizumab in the maintenance period + 1.
- For all other subjects, use the minimum of the following:
  – Date of last dose of bimekizumab in the maintenance period – date of first dose of first bimekizumab in the maintenance period + 140 days.
  – Date of last clinical contact in the maintenance period – date of first dose of bimekizumab in the maintenance period + 1.

  Note: This group could include subjects who discontinue the maintenance period early, subjects who complete the Maintenance Period as scheduled but choose not to continue into an open-label study, or subjects who are ongoing in the SFU period at the time of the data snapshot.

Was changed to:

For subjects randomized to BKZ Q4W/Q4W or BKZ Q4W/Q8W:

Study medication duration (days)

- Date of last dose of bimekizumab in the Maintenance Period – date of first dose in the Maintenance Period + 28 days (or 56 days in the case of Q8W dosing).

  Note: If date of last dose of bimekizumab in the Maintenance Period + 28 days (or 56 days in the case of Q8W dosing) extends to a date beyond the final visit date in the Maintenance Period (including PEOT, but not including SFU), then this calculation reverts to:

  – Final visit date in the Maintenance Period (including PEOT, but not including SFU) – date of first dose in the Maintenance Period + 1.

  Note: For subjects who die, if date of last dose of bimekizumab in the Maintenance Period + 28 days (or 56 days in the case of Q8W dosing) extends to a date beyond the date of death, then this calculation reverts to:

  – Date of death – date of first dose in the Maintenance Period + 1.

Time at risk (days):

- For subjects who continue into an open-label study (and, therefore, do not have the SFU visit in the feeder study): Final visit date in the Maintenance Period – date of first dose in the Maintenance Period + 1.
- For subjects who die prior to the final visit in the Maintenance Period: Date of death – date of first dose in the Maintenance Period + 1.
- For all other subjects, use the minimum of the following:
  - Date of last dose in the Maintenance Period – date of first dose in the Maintenance Period + 140 days.
  - Date of last clinical contact– date of first dose in the Maintenance Period + 1.

  Note: This group could include subjects who discontinue the maintenance period early, subjects who complete the Maintenance Period as scheduled but choose not to continue into an open-label study, or subjects who are ongoing in the SFU period at the time of the data snapshot.

For subjects randomized to ADA/BKZ Q4W:

Study medication duration (days)

Only the duration of Bimekizumab treatment will be derived over this period as the summaries produced are for the BKZ Set (not including any adalimumab treatment).

- Date of last dose of bimekizumab on or after Week 24 – date of first dose on or after Week 24 + 28 days

  Note: If date of last dose of bimekizumab + 28 days extends to a date beyond the final visit date in the Maintenance Period (including PEOT, but not including SFU), then this calculation reverts to:
  - Final visit date in the Maintenance Period (including PEOT, but not including SFU) – date of first dose on or after Week 24 + 1.

  Note: For subjects who die, if date of last dose of bimekizumab in the Maintenance Period + 28 days extends to a date beyond the date of death, then this calculation reverts to:
  - Date of death – date of first dose on or after Week 24 + 1.

Time at risk (days):

- For subjects who continue into an open-label study (and, therefore, do not have the SFU visit in the feeder study): Final visit date in the Maintenance Period – date of first dose on or after Week 24 + 1.
- For subjects who die prior to the final visit in the Maintenance Period: Date of death – date of first dose on or after Week 24 + 1.
- For all other subjects, use the minimum of the following:
  - Date of last dose on or after Week 24 – date of first dose on or after Week 24 + 140 days.
  - Date of last clinical contact – date of first dose on or after Week 24 + 1.

Note: This group could include subjects who discontinue the Maintenance Period early, subjects who complete the Maintenance Period as scheduled but choose not to continue into an open-label study, or subjects who are ongoing in the SFU period at the time of the data snapshot.

Change #8

Section 10.2.3.5: Suicidal Ideation and Behavior

The following text was added:

A separate listing will also be produced to summarize the adjudicated results of all events escalated to the full committee.

Change #9

Section 10.2.3.7: Anaphylactic Reaction

Was changed to:

Section 10.2.3.7: Hypersensitivity (including anaphylaxis)

The following text was added for both Anaphylactic Reaction TEAEs and Hypersensitivity TEAEs:

An AE glossary table will also be produced to summarize the MedDRA coding for these events. The glossary table will include the following fields: reported term, PT, LLT, HLT, and SOC.

Change #10

Section 10.2.3.8: Hepatic Events and DILI

- Incidence and Event Rate of Hepatic Events and Drug Induced Liver Injury TEAEs per 100 subject years by SOC, HLT and PT

Was changed to

Section 10.2.3.8: Hepatic Events

- Incidence and Event Rate of Hepatic Events TEAEs per 100 subject years by SOC, HLT and PT

STATISTICAL ANALYSIS PLAN SIGNATURE PAGE

This document has been reviewed and approved per the Review and Approval of Clinical Documents Standard Operating Procedures. Signatures indicate that the final version of the Statistical Analysis Plan (SAP) or amended SAP is released for execution.

Approval Signatures

Name: ps0008-sap-amendment-2

Version: 1.0

Document Number: CLIN-000143707

Title: ps0008-sap-amendment-2

Approved Date: 31 Oct 2019

Document Approvals	
Approval Verdict: Approved	Name: Jackie Thirlwell Capacity: Clinical Date of Signature: 31-Oct-2019 18:22:07 GMT+0000
Approval Verdict: Approved	Name: Veerle Vanvoorden Capacity: Clinical Date of Signature: 31-Oct-2019 20:46:10 GMT+0000